Review of Economic Dynamics **5,** 618–645 (2002)
doi:10.1006/redy.2001.0155

Does Intrinsic Habit Formation Actually Resolve the Equity Premium Puzzle?[1]

David A. Chapman

Finance Department, McCombs School of Business, University of Texas, Austin, Texas 78712-1179
E-mail: chapman@eco.utexas.edu

Received December 28, 1999; published online April 3, 2002

G. M. Constantinides (1990, *Journal of Political Economy* **98**, 519–543) describes a simple model of intrinsic habit formation that appears to resolve the "equity premium puzzle" of R. Mehra and E. C. Prescott (1985, *Journal of Monetary Economics* **15**, 145–161). This finding is particularly important, since it has motivated a broader consideration of the implications of habit formation preferences in dynamic equilibrium models. However, consumption growth actually behaves very differently pre- and post-1948, and the explanatory power of the habit formation model is driven by the pre-1948 data. Using data from 1949 to 2000, constructed in a manner comparable to R. Mehra and E. C. Prescott, I demonstrate that intrinsic habit *cannot* rationalize the unconditional moments of discrete consumption and real asset returns with values of the risk aversion coefficient that are less than four times larger than the values found by G. M. Constantinides for *any* feasible calibration of the model. *Journal of Economic Literature* Classification Numbers: E21, G12.

1. INTRODUCTION

The equity premium puzzle, as posed in Mehra and Prescott (1985) (hereafter MP), is based on the observation that the average realized return premium on a broad portfolio of stocks above the short-term, risk-free asset is 6%, using U.S. data from 1889 to 1978. The "puzzle" is that this simple fact has proven to be very difficult to rationalize within the framework of a time-separable, complete-markets models of the determination of asset returns; see Kocherlakota (1996) for a survey of this literature. It is not

[1] This paper is available online at http://www.bus.utexas.edu/Faculty/David.Chapman/habit2.html. I thank George Constantinides, John Graham, Sheridan Titman, two anonymous referees, and the editor (Tom Cooley) for helpful comments.

1094-2025/02 $35.00

surprising, therefore, that a number of alternative theoretical models have been suggested as possible explanations.

Constantinides (1990) proposes a solution to the puzzle that is based on changing the form of the utility function of the representative agent. In this simple "habit formation" model, momentary utility depends on the difference between current consumption and a weighted-average of the consumer's own past consumption choices. This is the definition of "intrinsic" habit formation, as opposed to "external" habit models, where momentary utility depends on aggregate consumption. Constantinides (1990) demonstrates that a suitably parameterized version of a simple intrinsic habit formation model can reconcile the unconditional moments of asset returns and consumption growth with moderate levels of risk aversion. In fact, this finding has motivated a broader consideration of habit formation specifications in dynamic equilibrium models. See, for example, Campbell and Cochrane (1999), Jermann (1998), and Lettau and Uhlig (2000).

The economic intuition behind the solution to the puzzle offered by habit formation utility is that it offers additional flexibility in modeling risk aversion and the agent's attitude toward intertemporal substitution in consumption, since these constructs are no longer simply the inverse of one another. As long as the representative agent is willing to substitute consumption across time, smooth consumption can be consistent with less curvature in the agent's value function. In this paper, I reconsider the ability of this simple version of habit formation to explain the equity premium puzzle.

The issue here is not the model's economic intuition, but rather its calibration and quantitative significance. There are two specific issues considered below. First, the model is formulated in continuous time. As Constantinides (1990) notes, this difference is not an essential feature of the resolution of the puzzle, but it does complicate the process of calibrating the model to match the unconditional moments of the data. This is because the stylized facts in MP are developed using discretely sampled (annual) observations of consumption and asset returns, and sample moments based on these observations need to be calibrated to the discrete moments implied by the continuous-time model.

In the case of asset returns, this adjustment is straightforward. However, estimating the unconditional moments of a continuous-time consumption growth process from the discrete consumption data is complicated by the well-known "time-aggregation" problem; see Christiano *et al.* (1991) or Heaton (1993). Observed consumption is *not* a periodic sampling from the (unobserved) continuous-time consumption process, but rather it is the integral of continuous-time consumption between the discrete measurement dates.

The second important issue in calibrating the model is the covariance stationarity of the consumption growth process. In the five decades from

1889 to 1938, the average standard deviation of consumption growth was 4.51% per year. In the three decades from 1949 to 1978, the average standard deviation of consumption growth was 1.13% per year. This dramatic decline in the ex post standard deviation of consumption growth may reflect the resolution of ex ante uncertainty about the survival of the political systems of the United States and Western Europe [see Brown *et al.* (1995) or Jorion and Goetzmann (1999)].

Alternatively, it may reflect changes in the construction of the aggregate consumption series over time. In either case, these facts suggest that using data from the first half of the 20th century is not innocuous when examining a consumption-based explanation of the properties of asset returns. Of course, using only post-1949 data can be expected to worsen the equity premium puzzle because it *reduces* consumption volatility.

The first result reported below is that adjusting the calibration of the model to match discrete moments has a noticeable, but economically insignificant impact on the conclusions in Constantinides (1990). Using the MP data set and the point estimates of consumption growth and asset return moments, the most important change introduced by using instantaneous moments is that the expected return on the risky asset increases from 7.00% per year to 8.36% per year. This raises the Sharpe ratio of the risky asset, and it implies that the smallest level of the mean value function curvature coefficient [what Constantinides (1990) refers to as the coefficient of risk aversion] associated with Constantinides's parameterization increases from 2.811 to 4.705. This increase is not economically significant in the sense that it is hard to imagine a reader with a strong opinion that 2.811 is a reasonable value for the curvature coefficient but that 4.705 is excessive. Furthermore, there are alternative permissible parameterizations of the model that drive the curvature coefficient lower.

The nonstationarity of the consumption growth process turns out to be more problematic for the habit formation model. I construct annual data from 1949 to 2000 that is similar to the MP data set. Using these data (and converting from instantaneous to discrete moments), the smallest mean curvature coefficient (including leverage effects) is approximately 10.75. This—I would argue—is a substantial increase above the value of 2.811 reported in Constantinides (1990), *and it cannot be reduced by considering alternate choices of the habit formation parameters*.

This is not the first paper to suggest that there are deficiencies in simple habit models. Indeed, the model developed in Campbell and Cochrane (1999) to explain the *dynamics* of conditional asset return moments represents an evolution of the habit specification that is designed to overcome some of the problems of the Constantinides (1990) framework. Jermann (1998) and Lettau and Uhlig (2000) are examples of the problems that can arise when representative agent habit models are extended to production economies.

Indeed, Otrok *et al.* (2001) also challenge the conclusion that the basic habit model resolves the equity premium puzzle, and an important component of their analysis is to argue that the consumption process has changed dramatically over the course of the 20th century. They use the concept of spectral utility to decompose the variability of an exogenous consumption process into the contributions coming from different frequencies of the spectrum. Otrok *et al.* (2001) demonstrate that habit utility implies much more sensitivity to high-frequency, as opposed to low-frequency, volatility in consumption. This implies that it is possible to generate dramatic changes in the equity premium for small changes in overall consumption volatility and vice versa. They also point out that habit utility predicts changes in the equity premium due to the changes observed in aggregate consumption growth that are completely at odds with the data on asset returns.

The advantage of this paper, relative to that of Otrok *et al.* (2001), is simplicity. Their analysis is conducted in discrete time, using an exogenous consumption growth process, a habit index based on discrete lags, and asset returns that are determined endogenously. This is—by far—the most common mechanism for describing an endowment economy, and it provides their results with a wide range of applications.

This advantage comes at a cost of differing in a number of ways from the structure used in Constantinides (1990). By contrast, I adopt all of the assumptions of the Constantinides analysis. The model is examined in continuous time. The return to the risky asset is determined exogenously as the return to the risky production process, which is iid normal, the risk-free return is also determined exogenously, and the habit index is described by two (strictly positive) parameters, one of which defines the "persistence" of habit and the other the "intensity" of habit. As in Constantinides (1990), it is the consumption process that is determined endogenously.

The only differences between the analysis in Constantinides (1990) and the quantitative experiments conducted below are related to correcting the time-aggregation problem and examining the appropriate time period to use in conducting the calibration. It is not clear that all of the differences between a standard endowment economy and the Constantinides setup are important in examining the equity premium.[2] However, by adopting all

[2]Indeed, Constantinides (1990) argues explicitly that some of these differences are unimportant. The differences that seem to matter, more for some parameterizations of the model than for others, are the time-aggregation measurement issues and the fact that the time-discount rate in the Constantinides structure has no connection whatsoever to the return to the risk-free asset.

of these assumptions, I can provide a much more direct analysis of the robustness of the Constantinides (1990) results.

The remainder of the paper is organized as follows: Section 2 re-examines the stylized facts about the moments of consumption growth from 1889 to 1978. Section 3 contains a brief description of the intrinsic habit formation model, and a detailed discussion of the model calibration, explicitly adjusting for time aggregation in measured consumption growth moments, is presented in Section 4. The results of various simulations of the model are examined in Section 5, and the conclusions are in Section 6. Appendix A describes the data used in defining the stylized facts.

2. RE-EXAMINING CONSUMPTION GROWTH

As noted earlier, the equity premium puzzle, in the narrowest sense, is defined using the annual consumption growth rate in real per capita consumption of nondurable goods and services from 1889 to 1978 contained in MP. The original source for the data is Grossman and Shiller (1981); the data are plotted in Fig. 1.[3]

This consumption growth process does *not* appear, to the naked eye, to be covariance stationary over the 90 years of the MP data set. The *volatility* of consumption growth is much higher in the pre-1940 data than in the period following World War II. The simple average of the average volatility of consumption growth per decade for the five decades from 1889 to 1939 was 4.51% per year in the MP data, while the simple average of the average volatility in the three decades from 1949 to 1978 was only 1.13% per year.[4]

Table I examines the Shiller (1982) data over the entire sample and over the subperiods of 1890–1948 and 1949–1978. The moments of the total period are dominated by the first (longer) subperiod. In addition to the increased volatility of the first subperiod, the first four autocorrelations of the two subperiods are exactly opposite in sign, and the magnitudes of the autocorrelations in the latter period are generally larger. Suppose that the null hypothesis is that annual, real, per capita consumption growth, g_t, evolves as a first-order autoregressive process with constant coefficients

$$g_{t+1} = \alpha + \phi g_t + \varepsilon_{t+1}, \tag{1}$$

where $|\phi| < 1$, $\alpha/(1-\phi)$ is the unconditional mean of annual consumption growth, and $\varepsilon_{t+1} \sim \text{iid}(0, \sigma^2)$. Simple estimates of the model parameters in

[3]The data in Fig. 1 actually come from Shiller (1982). There are slight differences between the MP data and the data reported in Shiller (1982), but these differences are not economically significant.

[4]The corresponding values in the Shiller (1982) data sets are 4.24% and 1.06%, respectively.

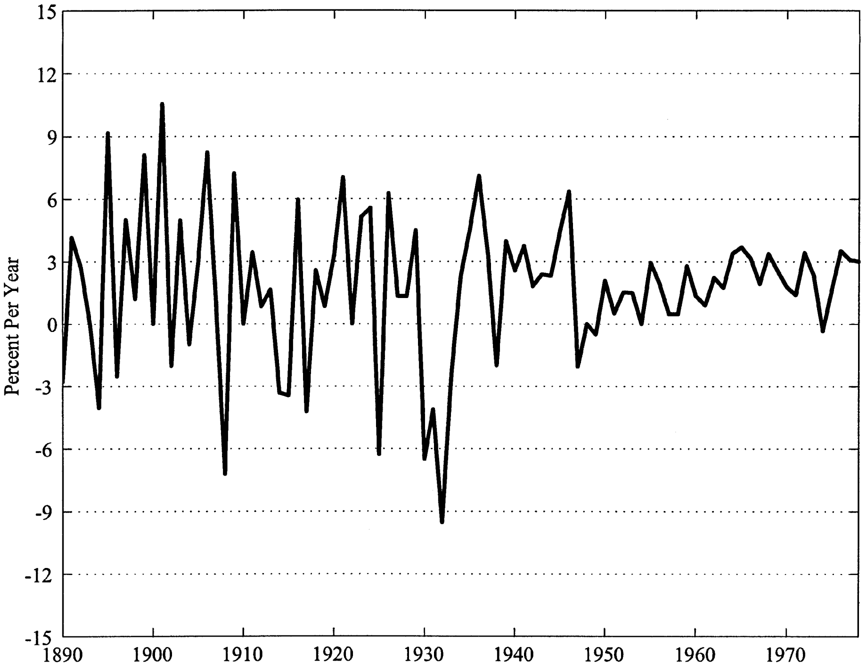

FIG. 1. Growth rate in real per capita consumption of nondurable goods and services from 1890 to 1978.

the 1890–1948 and 1949–1978 subperiods are shown in panel A of Table II. These estimates are consistent with the results in Table I, and they verify that consumption growth variances appear to be different in the two subperiods, even after controlling for differences in the conditional mean.

A comparison of the coefficients across the two subperiods clearly does not constitute a formal test of the hypothesis that ϕ and σ have changed. Furthermore, the results in Tables I and II make the arbitrary assumption

TABLE I
Examining Consumption Growth from 1890 to 1978

			Autocorrelations			
Period	Mean	SD	ρ_1	ρ_2	ρ_3	ρ_4
1890–1978	1.729	3.565	−0.14	0.17	−0.05	−0.04
1890–1948	1.630	4.306	−0.16	0.17	−0.05	−0.06
1849–1978	1.922	1.195	0.19	−0.03	0.24	0.26

Note. The data are from Shiller (1982), and means and standard deviations are expressed in percent at an annual rate.

TABLE II
Consumption Growth Rate Autoregressions: Shiller Data

Panel A: Regression estimates for aggregate consumption growth

Period	Parameters α	ϕ	σ	R-squared	Nobs.
1890–1948	1.967	−0.157	4.287	2.5%	58
	(0.603)	(0.131)			
1849–1978	1.433	0.269	1.169	7.6%	30
	(0.387)	(0.177)			

Panel B: Cumsum-squares test for structural stability in consumption growth

	Statistic	p-value
Cumsum squares	1.6205	0.0105

Note. The data are from Shiller (1982), and the model being estimated in each subperiod is given in Eq. (1). Nobs. is the number of observations in the estimation period. Standard errors are reported in parentheses.

that the break in the data occurs in 1948. This is in the tradition of the Chow test, which is discussed, for example, in Johnston (1984). There has been a large recent literature on testing for structural breaks in dynamic models *without* prespecifying the break point. See, for example, Andrews (1993) and Andrews and Ploberger (1994). These tests can be applied to dynamic models of the form of Eq. (1). They are designed to have power against (local or distant) alternative hypothesis of breaks in the coefficients of the conditional mean, as opposed to heteroscedasticity (conditional or unconditional).

The results of applying the cusum-squares test to the annual consumption data are shown in panel B of Table II.[5] The value of the test statistic is 1.6205, which has a p-value of 0.0105 under its asymptotic distribution. This is additional evidence—beyond the heuristic analyses of Fig. 1 and Tables I and II—that there is a structural break in the aggregate consumption growth data. It is not clear whether this break reflects possible changes in the construction of the data or whether there has been a fundamental change in the volatility of the data-generating process itself. In either case, it appears that using data from the first half of the 20th century in examining

[5]The details of the construction of the cusum-squares test are described in Ploberger and Krämer (1990).

asset pricing models is problematic in the sense that it may not be useful for constructing reasonable forecasts of future consumption volatility.[6, 7]

There are similar concerns in the literature about the use of *aggregate* consumption data in testing equilibrium pricing models. If financial assets are held by only a small subset of the population, then asset returns may only be related to a component of aggregate consumption. See, for example, Heaton and Lucas (1998), Mankiw and Zeldes (1991), and Polkovnichenko (2000). This problem is likely to be substantially worse in matching the consumption and return data from the period of the 1890s through World War I.[8] These arguments about the noncomparability of early return and consumption data with post-war data are distinct from the "survivorship" arguments of Brown *et al.* (1995).

Finally, what is the economic significance of using only the post-war data? The major impact on pricing tests is through the dramatic reduction in the volatility of consumption growth. This makes it more difficult for simple,

[6] Romer (1989) demonstrates that the volatility of the pre–World War I business cycle is substantially overstated, when based on standard estimates of aggregate output. While these arguments do not carry over directly to measures of annual consumption data, since the revised GNP series that Romer constructs is based on differences in the cyclical properties of commodity output, producer prices, and the value added in transportation and distribution, it suggests that measurement errors may account for a portion of the changing characteristics of the consumption data.

[7] In addition to a statistical analysis, there is anecdotal evidence suggesting that the use of data from the early period of the Shiller (1982) data set (and MP data set) may be problematic. For example, consider the following excerpt from Chernow (1998) describing the "panic of 1907":

> As panic overtook Wall Street in late October 1907, throngs of petrified depositors lined up in front of banks to empty their accounts. . . . That night, in an extraordinary pledge of faith in a private citizen, Treasury Secretary George Cortelyou met with [J. P.] Morgan in a Manhattan hotel room and placed at his disposal twenty-five million dollars [approximately $310 million in 1999 dollars] in government funds to stem the panic. While Morgan was the impresario of the salvage operation, [John D.] Rockefeller [Sr.] provided more private money than anybody else. . . . Rockefeller . . . called Melvin E. Stone, general manager of the Associated Press. He told Stone, for quotation, that the country's credit was sound and that, if necessary, he would give half of all he possessed to maintain America's credit. . . . Rockefeller offered his services to J. P. Morgan, and his millions formed part of the twenty-five million-dollar fund that Morgan marshaled that day to keep the stock market open, averting the bankruptcy of at least fifty brokerage houses. [Chernow (1998, pp. 534–544).]

This incident demonstrates the substantial differences between U.S. capital markets of the early 20th century and modern U.S. capital markets. It is unimaginable that, in modern times, any two individuals could "bail out" Wall Street.

[8] Again, Chernow (1998) contains substantial anecdotal evidence of both the concentrated ownership of the Standard Oil Company and its relative importance in the composition of a value-weighted portfolio of NYSE stocks.

dynamic models to satisfy the joint restrictions on asset returns and consumption growth. Clearly, since MP already reject a time-separable model using the more volatile consumption data, this restriction in the data can only make this rejection worse. The interesting issue is whether or not the habit formation model of Constantinides (1990) can survive the more stringent test of the post-war data.

3. A BRIEF REVIEW OF INTRINSIC HABIT FORMATION WITH EXOGENOUS RETURNS

Since the model is described in detail in Constantinides (1990), only the highlights and main conclusions are presented below, primarily to introduce notation and basic definitions.[9] There is a single representative agent whose consumption and investment decisions rationalize the observed aggregate consumption process and the observed equilibrium asset prices. There is a single good available for consumption or investment.

The preferences of the representative agent over alternate consumption paths $\{c_t\}_{t=0}^{\infty}$ are defined by the functional

$$U\big(\{c_t\}_{t=0}^{\infty}\big) = E_0\left[\int_0^{\infty} e^{-\rho t}\frac{(c_t - x_t)^{\gamma}}{\gamma}\,dt\right], \tag{2}$$

where $E_0[\cdot]$ denotes expectation with respect to the information available at $t = 0$ and $\{x_t\}_{t=0}^{\infty}$ is an index of past consumption. This “habit index” is defined as

$$x_t = e^{-at}x_0 + b\int_0^t e^{a(s-t)}c_s\,ds, \tag{3}$$

where the parameter b defines the intensity of habit persistence and a defines the persistence of past consumption in the construction of the habit index. ρ, γ, a, and b are all constants. Expected utility is not defined for levels of consumption below the level of the habit index.

There are only two investment opportunities available to the representative agent. There is a risk-free technology, which has an expected return over any instant dt of $r\,dt$, and a risky technology whose instantaneous return is $\mu\,dt + \sigma\,d\omega_t$, where ω_t is a scalar Brownian motion. This Brownian motion is the *only* source of uncertainty in the model. r, μ, and σ are constants.

[9]The notation in this section (and the following sections) is chosen to be consistent with Constantinides (1990).

The unique optimal consumption and investment processes are given by

$$c_t = x_t + h\left[W_t - \frac{x_t}{r+a-b}\right] \tag{4}$$

and

$$\alpha_t = m\left[1 - \frac{x_t/W_t}{r+a-b}\right], \tag{5}$$

where

$$h \equiv \left[\frac{r+a-b}{(r+a)(1-\gamma)}\right]\left[\rho - \gamma r - \frac{\gamma(\mu-r)^2}{2(1-\gamma)\sigma^2}\right].$$

W_t is the optimal wealth process, defined as

$$W_t = \frac{x_t}{r+a-b} + \left(W_0 - \frac{x_0}{r+a-b}\right)\cdot \exp\left[\left(n - \frac{m^2\sigma^2}{2}\right)t + m\sigma\omega_t\right], \tag{6}$$

where $m \in [0, 1]$ is

$$m \equiv \frac{\mu - r}{(1-\gamma)\sigma^2}$$

and

$$n \equiv \frac{r-\rho}{1-\gamma} + \frac{(\mu-r)^2(2-\gamma)}{2(1-\gamma)^2\sigma^2}.$$

The parameter restrictions invoked in Constantinides (1990) in the proof of the existence and uniqueness of the optimal consumption policy are:

(1) $1-\gamma > 0,\ \gamma \neq 0$;
(2) $W_0 > 0,\ W_0 - [x_0/(r+a-b)] > 0$;
(3) $0 < b < r+a$;
(4) $\rho - \gamma r - \frac{1}{2}\gamma(\mu-r)^2(1-\gamma)^{-1}\sigma^{-2} > 0$;
(5) $x_0 \geq 0$;
(6) $0 \leq m \equiv (\mu-r)/[(1-\gamma)\sigma^2] \leq 1$;
(7) $n + a - b - m^2\sigma^2 > 0$.

The consumption growth rate consistent with Eq. (4) is

$$\frac{dc_t}{c_t} = [n + b - (n+a)z_t]\,dt + [1 - z_t]m\sigma\,d\omega_t, \tag{7}$$

where $z_t \equiv x_t/c_t$.

The curvature coefficient of the value function for this planning problem is equal to

$$\Psi = (1-\gamma)\left[1 + \frac{hz_t}{(1-z_t)(r+a-b)}\right], \tag{8}$$

which is time varying with z_t.[10] The unconditional mean of this curvature parameter is

$$\overline{\Psi} = (1-\gamma)\left[1 + \frac{hb}{(r+a-b)(n+a-b-m^2\sigma^2)}\right]. \tag{9}$$

The fundamental point of Constantinides (1990) is that habit formation permits (simultaneously) a more flexible modeling of attitudes toward atemporal wealth gambles and the intertemporal elasticity of substitution in consumption.[11] This flexibility may allow for a reconciliation of the unconditional moments of excess returns, the risk-free rate, and consumption growth.

The final step in connecting the model economy with aggregate data on consumption and asset returns is to assume the existence of a single firm whose value is equal to the capital stock in the economy. This firm invests a constant percentage, δ_1, of its capital in the risky technology, which implies that $(1-\delta_1)$ of its capital is invested in the risk-free technology. The firm is financed with risk-free debt and equity, and the ratio of equity to total firm value is a constant δ_2. The rate of return on the stock of this firm, where the stock price at time t is denoted S_t, is

$$\frac{dS_t}{S_t} = \left(\frac{\delta_1}{\delta_2}\right)[(\mu - r)\,dt + \sigma\,d\omega_t] + r\,dt. \tag{10}$$

4. CALIBRATING THE MODEL

The process of model calibration as described, for example, in Cooley and Prescott (1995) involves three steps: (1) Restrict the model primitives to a parametric class. (2) "Construct measurements that are consistent with the

[10]Constantinides (1990) describes this as the coefficient of relative risk aversion, defined "in terms of an atemporal gamble that changes the current level of (wealth) by the outcome of the gamble ..." (p. 527).

[11]The intertemporal elasticity of substitution in consumption in the habit formation model is defined by Constantinides (1990) as

$$s \equiv \left.\frac{\partial[E(dc/c)/dt]}{\partial r}\right|_{z_t\,\mu-r,\,\sigma^2} = \frac{1-z_t}{1-\gamma}.$$

parametric class of models.... (3) ... [A]ssign values to the model parameters" [Cooley and Prescott (1995, p. 15)]. The model structure defined in the previous section constitutes the first step in this process. The parameters of the model, $[\rho, \gamma, a, b, r, \mu, \sigma]$, must be matched to the moments measured in the data.

4.1. Real Asset Return: 1949–2000

The asset return data consist of real annual returns (with dividends) to the CRSP value-weighted portfolio and the (average) real annual return to three-month Treasury bills. These series are shown in Fig. 2, and their construction is described in Appendix A. Table III shows that the average excess return to stocks from 1949 to 2000 is 8.73% per year, which is *larger* than the 6.18% annual equity premium in MP. The average real return to stocks was 10.46% per year, and the real return to Treasury bills was 1.74% per year.

Real stock return volatility is *slightly* higher over the 1949–2000 period than over the 1889–1978 period. The average annual standard deviation of real stock returns has increased from 16.54 to 17.37. At first glance, this seems surprising, given that the longer time period includes the extremely volatile period of the Great Depression. However, the longer period also includes the relatively low-volatility decades of 1889–1898 and 1909–1918, and the latter period includes the relatively volatile decade of 1969–1978.

The risk-free rate, r, is set equal to its sample mean of 17.4% per year over the 1949–2000 period. Equation (10) implies that the instantaneous moments of stock returns are

$$\frac{E(dS_t/S_t)}{dt} = \frac{\delta_1}{\delta_2}(\mu - r) + r \tag{11}$$

and

$$\frac{\operatorname{Var}(dS_t/S_t)}{dt} = \left(\frac{\delta_1}{\delta_2}\right)^2 \sigma^2. \tag{12}$$

Since 1970, the average leverage ratio of equity to total firm value has been 0.6, which provides a value for δ_2.[12] Given this parameter value, the leverage ratio is set equal to $\delta_1/\delta_2 = 1.67$.[13] Given the process for the risky asset, Eqs. (11) and (12), the level of the risk-free rate, and aggregate

[12] See Grinblatt and Titman (1997, pp. 5–6).

[13] As Constantinides (1990) notes, this ratio is free to vary in the half-open interval $(0, \delta_2^{-1}]$. Leverage affects the model through Eqs. (11) and (12) and condition (6) in Section 3. In particular, when δ_1/δ_2 is allowed to vary, condition (6) becomes

$$\frac{\mu - r}{\sigma^2} \leq \frac{\delta_1}{\delta_2}(1 - \gamma),$$

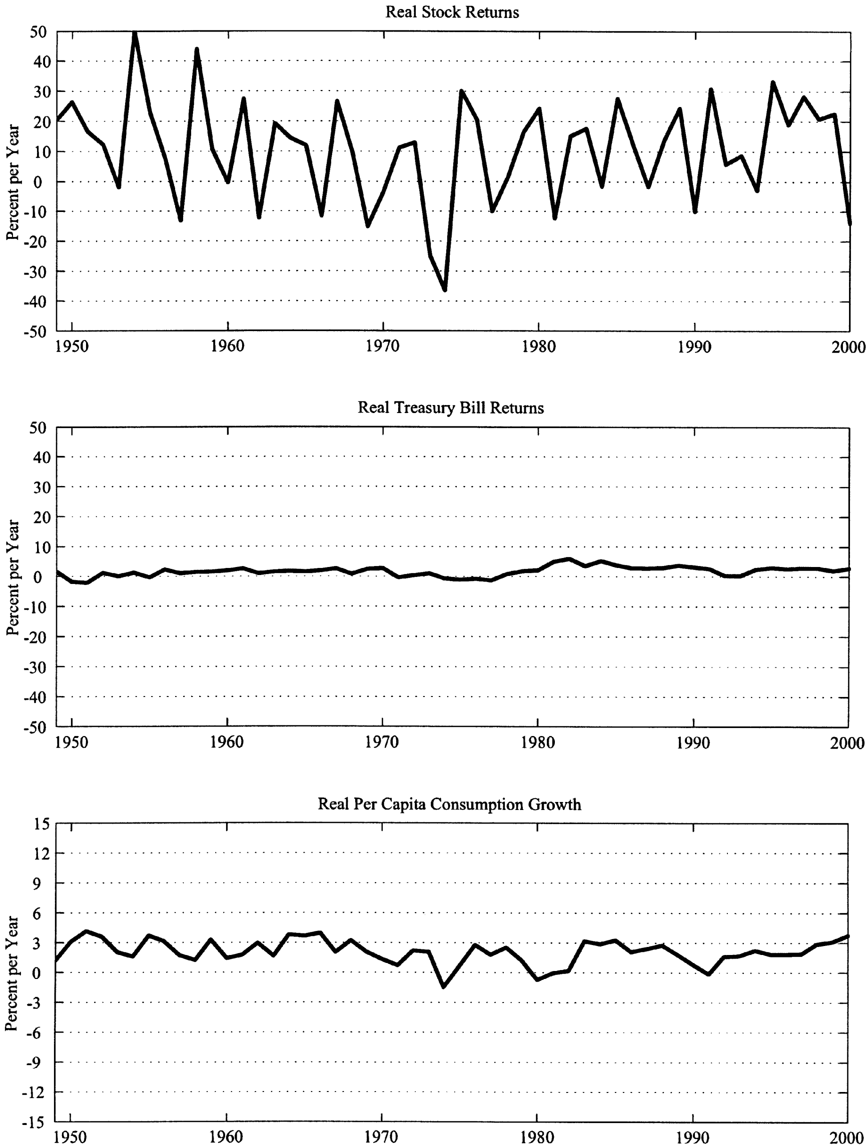

FIG. 2. Ex post real return to the CRSP value-weighted index, ex post real yield on three-month Treasury bills, and growth rate in real per capita consumption of nondurable goods and services from 1949 to 2000.

TABLE III
Consumption Growth and Asset Returns from 1949 to 2000

	Real consumption growth		Risk-free return		Risk premium		Stock return	
Time period	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1949–2000	2.07	1.23	1.74	1.73	8.73	17.41	10.46	17.37
SE	0.17		0.24		2.41		2.41	
1949–1958	2.48	1.13	0.43	1.53	18.08	19.22	18.51	19.15
1959–1968	2.73	1.01	1.77	0.61	7.79	13.69	9.56	13.98
1969–1978	1.37	1.19	0.29	1.48	−1.63	21.21	−1.34	20.75
1979–1988	1.51	1.36	3.63	1.38	7.48	13.14	11.11	12.52
1989–2000	1.93	1.02	2.43	1.07	11.38	15.93	13.81	16.10

Note. All data series are described in Appendix A. All series are converted to real returns using the inflation rate on personal consumption expenditures on nondurables plus services. The series are expressed in percent at an annual rate.

leverage, μ is equal to 0.0730 and σ is equal to 0.1066, where these are the values that solve the equations

$$1.67(\mu - r) + r = 0.1046 + \frac{1}{2}\sigma^2$$

(the $\frac{1}{2}\sigma^2$ term follows from the lognormality of the risky asset return) and

$$(1.67)^2\sigma^2 = (0.1737)^2.$$

4.2. The Rate of Time Preference

Ordinarily, in an economy in which the consumption process is specified exogenously (an "endowment economy") or even in the production economies examined in the real business cycle literature, the time preference parameter, ρ, is chosen so that the real risk-free rate implied by the standard Euler equation from the representative agent's optimal consumption-portfolio problem implies an average risk-free rate that is equal to observed data.[14]

which means that γ can be smaller (in absolute value) for $\delta_1/\delta_2 > 1$. δ_2 equal to 0.6 implies a *maximum* value for δ_1/δ_2 of 1.67, which occurs for the extreme case in which "the firm" invests solely in the risky technology. Choosing δ_1/δ_2 at the extreme upper end of its permissible range allows $(1-\gamma)$ to be as small as possible.

[14] As an example, if aggregate consumption evolves in discrete time as a lognormal process, then Campbell (1986) shows that the short-term risk-free rate will be determined as

$$r_t = \rho + \gamma E_t(g_{t+1}) - \frac{1}{2}\gamma^2 \mathrm{Var}_t(g_{t+1}),$$

This is explicitly *not* the case in the model in Section 3. There, the risk-free (and risky) return is determined exogenously by the return to the risk-free (and risky) technology, which is assumed to have an infinitely elastic supply. Absence of arbitrage between real production possibilities and financial assets fixes the risk-free rate at the exogenously specified level of r. In essence, ρ is a free parameter in the model, and it can be used to alter the properties of the consumption process without affecting asset returns.

The only remaining mechanism for selecting the discount parameter is economic intuition about what would seem to constitute a "reasonable" range for the parameter. At one extreme, a value of ρ equal to 0 implies that future periods are valued as highly (in utility terms) as current consumption. An upper bound for ρ, say $\bar{\rho}$, is more complicated, since there is no obvious candidate for an effective lower bound on the investment horizon. For example, a value of ρ equal to 0.5 implies that consumption 10 years in the future is virtually worthless. Is ρ equal to 0.5 too large or too small?

Any value of ρ in the interval $[0.0, \bar{\rho}]$ is equally reasonable (or unreasonable). In the baseline parameterization of the model, I will fix ρ at 0.0 for two reasons. First, this value is actually most favorable to the model's ability to resolve the equity premium puzzle. Second, resolving the equity premium puzzle by varying the time-discount parameter is (in my opinion) inconsistent with the spirit of the analysis in Constantinides (1990). The *point* is to use the *habit parameters* to resolve the puzzle. In Section 5.2, I will examine the sensitivity of the results to varying the rate of time preference.

4.3. Consumption Growth: 1949–2000

Calibrating the consumption process in the Constantinides (1990) model is complicated by the well-known time-aggregation problem: Measured consumption is an integral of the model's instantaneous consumption rates.[15] Observable consumption, over a measurement interval Δ, is

$$c^a_{t+\Delta} \equiv \frac{1}{\Delta} \int_t^{t+\Delta} c_s \, ds, \tag{13}$$

where, in the following analysis, Δ will correspond to one year.

I choose the parameters of the consumption process in Eq. (4) so that the mean and standard deviation of an annual time average of the process

where g is the consumption growth rate and γ is the coefficient of relative risk aversion in a time-separable power utility function. In this case, variations in the value of ρ assigned to a representative agent will affect the real rate.

[15] See, for example, Christiano *et al.* (1991).

match the moments in the post-war data. The calibration is done in the following steps: (i) Simulate daily observation for $T = 6000$ years of the true, continuous-time process.[16] (ii) For each "year" of simulated data, evaluate the integral in Eq. (13) using a discrete approximation:

$$\hat{c}^a_{t+\Delta} \equiv \frac{1}{\Delta}\sum_{j=1}^{M} \hat{c}_{t+j} \cdot \tau \approx \frac{1}{\Delta}\int_t^{t+\Delta} c_s\, ds,$$

where $\hat{c}$ is the simulated daily series, $M = \Delta/\tau = 250$ is the number of daily observations in a year, and τ is the one-day time increment. (iii) Construct sample estimates of the mean and standard deviation of time-aggregated consumption growth from the time-aggregated simulated data.

In Constantinides (1990), the *instantaneous* consumption growth moments are calibrated using Eq. (7); i.e.,

$$\frac{E(dc/c)}{dt} = (n+b) - (n+a)E(z) \tag{14}$$

and

$$\frac{\mathrm{Var}(dc/c)}{dt} = m^2\sigma^2 E[(1-z)^2], \tag{15}$$

where the expectations in Eqs. (14) and (15) are calculated numerically from the stationary density for z. Constantinides (1990) proves that this density is

$$p_z(z) = k\ \exp\left[\frac{2b}{m^2\sigma^2}\left(1-\frac{1}{z}\right)\right](1-z)^{2(n+a-b-m^2\sigma^2)/m^2\sigma^2} z^{2(b-a-n)/m^2\sigma^2},$$

where

$$k \equiv \left(\frac{2b}{m^2\sigma^2}\right)^{1-2(n+a-b)/m^2\sigma^2} \Gamma\left[\frac{2(n+a-b)}{m^2\sigma^2} - 1\right],$$

and (by construction) z is contained in the half-closed interval $[0, 1)$. The density has a unique mode at

$$\mathrm{mode}(z) = \frac{n+a-\sqrt{(n+a)^2 - 4m^2\sigma^2 b}}{2m^2\sigma^2}.$$

Table IV compares the instantaneous and the time-averaged discrete moments of habit formation consumption growth for different choices of

[16]The sample path of the continuous-time process is constructed using an order 1.0 strong Itô–Taylor expansion. This approach is described in detail in Kloeden and Platen (1992). An appendix, from an earlier version of this paper, describes the application of this procedure to the problem at hand, and it is available online at http://www.bus.utexas.edu/~chapmand/habit2.html.

TABLE IV
Continuous vs. Discrete Time-Averaged Consumption Growth Moments

		Panel A: $\gamma = -4.55$							
		b							
		0.10		0.20		0.30		0.40	
a		Mean	SD	Mean	SD	Mean	SD	Mean	SD
0.10	Discrete	2.77	1.99	n.a.		n.a.		n.a.	
	Instant.	2.79	2.08						
0.20	Discrete	2.77	4.68	2.77	1.38	n.a.		n.a.	
	Instant.	2.90	5.24	2.77	1.19				
0.30	Discrete	2.76	5.72	2.76	3.70	2.76	1.29	n.a.	
	Instant.	2.98	6.51	2.83	3.66	2.77	0.83		
0.40	Discrete	2.26	6.27	2.76	4.71	2.77	3.24	2.75	1.27
	Instant.	3.02	7.19	2.89	4.99	2.80	2.81	2.77	0.64
		Panel B: $\gamma = -9.10$							
0.10	Discrete	1.52	0.68	n.a.		n.a.		n.a.	
	Instant.	1.52	0.70						
0.20	Discrete	1.52	2.42	1.53	0.54	n.a.		n.a.	
	Instant.	1.60*	2.77*	1.52	0.38				
0.30	Discrete	1.52	3.14	1.52	1.90	1.51	0.49	n.a.	
	Instant.	1.63*	3.53*	1.55*	1.88*	1.52	0.26		
0.40	Discrete	1.52	3.41	1.52	2.58	1.52	1.63	1.54	0.48
	Instant.	1.64*	3.92*	1.59*	2.68*	1.51*	1.43*	1.52	0.20

Note. "Discrete" refers to the mean and standard deviation of the time-averaged consumption process computed by approximating the integral in Eq. (13) along a long simulated path of the continuous-time process in Eq. (7). "Instant." refers to the instantaneous mean and standard deviation calculated according to Eqs. (14) and (15). "n.a." means that the parameter combination associated with that table entry is inadmissible under the restrictions in Section 3. An * means that the numerical approximations to the integrals used in evaluating the instantaneous moments failed to converge, and the values reported in the table are the values of the moments evaluated at the mode of the state variable x/c. The additional parameters used in constructing the table are $\mu = 0.0730$, $\sigma = 0.1066$, $r = 0.0174$, and $\rho = 0.0$.

γ, a, and b.[17] Expected consumption growth—measured either discretely or instantaneously—is insensitive to variation in the habit persistence parameter, a, or habit intensity parameter, b. Variation in γ, however, has a dramatic effect on expected consumption growth. A decline from -4.55 to -9.10 results in an expected growth rate decline of roughly 45%. The *volatility* of consumption growth is sensitive to the habit parameters a and b. Volatility increases as the persistence of habit declines, and it declines as

[17] The Values of μ, σ, r, and ρ are fixed at 0.0730, 0.1066, 0.0174, and 0.0, respectively.

the intensity of habit increases. Both of these results are intuitively reasonable: Reductions in the importance of prior habit results in the consumer being willing to accept a more volatile consumption stream.

A comparison of the discrete and instantaneous moments of the consumption process implies that the differences between these moments depend on the habit parameters in complex ways. In particular, the results in Table IV demonstrate: (i) Annual expected consumption growth measured from the (discrete) time-averaged process is typically smaller than the annual expected consumption growth implied by the instantaneous moments, although the differences are generally not large. (ii) Looking down the main diagonals of panels A and B, when $a = b$, the volatilities of the discrete moments are higher than their instantaneous counterparts, *and the magnitude of this difference is increasing in the level of the parameters*, but (iii) when $a > b$, the volatility of instantaneous moments exceeds the discrete moments, except where the instantaneous volatility cannot be computed at the mean of the state variable z. The second point noted above is particularly important, because a and b are close to equal in parameterizations that match consumption moments from the data.

Table III contains (among other things) summary statistics for real per capita consumption growth using annual data from 1949 to 2000. This series is constructed in a manner similar to the series in MP. There are no authoritative estimates of the habit index parameters a and b. The object of the calibration exercise in Constantinides (1990) was to show that a variety of combinations of a and b could reconcile the unconditional moments of real per capita consumption growth with a low value function curvature coefficient. This was accomplished by allowing a and b to vary while simultaneously satisfying consumption growth moments, all of the parameter restrictions listed in Section 3, and a “low” level of value function curvature. The calibrations reported below are in the same spirit, except that the consumption parameters are calibrated to match the moments of the discrete, time-averaged consumption growth process.

5. WHAT *DOES* THE MODEL IMPLY?

5.1. Basic Results

The calibration exercise in Constantinides (1990) demonstrates that, for various choices of a and b, the mean of the curvature coefficient can assume a variety of values, ranging from a high of 8.692 to a low of 2.811. As Constantinides (1990) notes, the curvature of the value function can be

TABLE V

Mean and Variance of Consumption Growth and Risk Aversion Measures for MP Data Calibrated to Discrete-Time Moments: $\mu = 0.084$, $\sigma = 0.165$, $r = 0.01$, $\gamma = -2.97$, $\delta_1/\delta_2 = 1$, and $\rho = 0.037$

Parameter a, per month	0.100	0.200	0.300	0.400	0.500	0.600
Parameter b	0.076	0.147	0.222	0.301	0.382	0.465
Mode of z ($\hat{z}$)	0.653	0.680	0.703	0.724	0.741	0.756
Discrete consumption growth	0.018	0.018	0.018	0.018	0.018	0.018
Standard deviation of discrete consumption growth	0.036	0.036	0.036	0.036	0.036	0.036
Value function curvature (Ψ)						
Mean	6.683	5.495	5.098	4.907	4.787	4.705
Value at $\hat{z}$	6.386	5.405	5.052	4.879	4.767	4.690
Elasticity of substitution						
s (at $\hat{z}$)	0.09	0.08	0.07	0.07	0.07	0.06
$s \cdot \Psi$ (at $\hat{z}$)	0.56	0.44	0.38	0.34	0.31	0.29

Note. parameters μ, σ, r, γ, δ_1, δ_2, and ρ are chosen to match the MP data set calibrated to the discrete moments of time-aggregated consumption growth and the instantaneous moments of asset returns. a and b are the parameters that determine the form of the habit persistence index. z is the state variable in the model, and it is defined as the ratio of the habit index to instantaneous consumption. The curvature of the value function is defined in Eq. (8), and the intertemporal elasticity of substitution is defined as

$$s \equiv \left. \frac{\partial[E(dc/c)/dt]}{\partial r} \right|_{z_t,\, \mu - r,\, \sigma^2} = \frac{1 - z_t}{1 - \gamma}.$$

reduced by considering the level of aggregate leverage.[18] The issue examined in this subsection is whether or not this finding is robust to time aggregation and to the changes in the consumption growth process documented in Section 2.

The effect of calibrating the model to discrete, rather than instantaneous, moments is shown in Table V.[19] The first difference from Constantinides (1990) is that μ has to satisfy Eq. (11), which means that it must increase from 0.07 to 0.084. This parameter change increases the Sharpe ratio of the risky asset, and partly to satisfy this higher Sharpe ratio and partly to match discrete, time-aggregated consumption growth, γ declines from -1.2 to -2.97.

In this parameterization of the model, the smallest value of the curvature parameter that reconciles the unconditional mean and standard deviation of consumption growth with asset return means and standard deviations is

[18] There is also some scope for reducing the mean coefficient of relative risk aversion by continuing to increase a and b.

[19] The results in Tables V–VII were constructed using MathCad 2001 Professional. Their accuracy was verified by first reproducing the results in Constantinides (1990). In the interest of space, these replications are not reported here, but they are available upon request.

TABLE VI
Mean and Variance of Consumption Growth and Risk Aversion Measures for the 1949–2000 Data: $\mu = 0.0730$, $\sigma = 0.1066$, $r = 0.0174$, $\gamma = -6.4$, $\delta_1/\delta_2 = 1.67$, and $\rho = 0.0$

Parameter a, per month	0.100	0.200	0.300	0.400	0.500	0.600
Parameter b	0.099	0.188	0.279	0.371	0.466	0.560
Mode of $z(\hat{z})$	0.831	0.859	0.875	0.886	0.898	0.905
Discrete consumption growth	0.0207	0.0207	0.0208	0.0208	0.0207	0.0207
Standard deviation of discrete consumption growth	0.0123	0.0125	0.0125	0.0127	0.0125	0.0123
Autocorrelation of consumption growth						
First order	0.45	0.62	0.71	0.78	0.79	0.80
Second order	0.25	0.48	0.59	0.68	0.69	0.70
Value function curvature (Ψ)						
Mean	17.43	13.95	12.53	11.71	11.35	10.97
Value at $\hat{z}$	17.01	13.80	12.45	11.66	11.31	10.94
Elasticity of substitution						
s (at $\hat{z}$)	0.023	0.019	0.017	0.015	0.014	0.013
$s \cdot \Psi$ (at $\hat{z}$)	0.388	0.264	0.210	0.180	0.155	0.140

Note. The parameters μ, σ, r, γ, δ_1, δ_2, and ρ are chosen to match the 1949–2000 data set calibrated to the discrete moments of time-aggregated consumption growth and the instantaneous moments of asset returns, allowing for aggregate leverage. a and b are the parameters that determine the form of the habit persistence index. z is the state variable in the model, and it is defined as the ratio of the habit index to instantaneous consumption. The curvature of the value function is defined in Eq. (8), and the intertemporal elasticity of substitution is defined as

$$s \equiv \frac{\partial[E(dc/c)/dt]}{\partial r}\bigg|_{z_t,\,\mu-r,\,\sigma^2} = \frac{1-z_t}{1-\gamma}.$$

4.705. Matching the discrete, as opposed to the instantaneous, moments of the data has a noticeable impact on the model's ability to resolve the equity premium puzzle, but it is not economically significant. First, 4.705 remains a comparatively low value of the curvature parameter (in the equity premium literature). More important, this number *can* be reduced further by considering aggregate leverage.

In Table VI, the situation is completely different. The habit formation model is calibrated to the discrete moments of asset return and consumption growth, using post-1948 data and the results in Section 4.3. This parameterization matches the asset returns and allows for aggregate leverage consistent with post-war U.S. data.[20] When ρ is set to zero, γ must be set equal to -6.4 in order to match the unconditional means of consumption growth. *Even allowing for aggregate leverage calibrated to the recent data*, it is now impossible to reconcile asset return moments and consumption growth

[20]If anything, given the trend in U.S. leverage ratios over the past 50 years, the estimate of 0.6 errs on the side of too much leverage, which works in favor of the model.

moments for a in the range of [0.1, 0.6], b consistent with the parameter restrictions in Section 3, and a level of risk aversion below (roughly) 11.

Finally, Table VI also reports the first- and second-order autocorrelation coefficients for discrete consumption growth implied by the model. The values range from 0.45 to 0.80 for the first-order coefficient, and from 0.25 to 0.70 for the second-order coefficient. The corresponding values in the 1949–2000 data are 0.41 and 0.07. These results imply that the habit parameterizations that match the mean and standard deviation of consumption growth are incapable of also matching the measured correlation structure of aggregate consumption data (in any period).

5.2. Is There Any Calibration That Works?

The results in Table VI suggest that values for the habit parameters in the range considered by Constantinides (1990) no longer serve to reconcile real asset return and consumption data with "reasonable" levels of value function curvature. Does this necessarily imply, however, that there are *no* permissible choices of γ, a, and b that would generate values for Ψ that are less than, say, 10?

To gain some intuition about possible alternative parameterizations, consider Figs. 3 and 4, which show the mean and variance (respectively) of discrete consumption growth as functions of a, b, and γ.[21] Given the values of μ, σ, r, and ρ, one representation for the relevant parameter restrictions required for the model solution to be well defined (derived from restrictions (3) and (7) in Section 3) is

$$a > (a - b) > \max\left[-r,\ \frac{(\mu - r)^2\gamma}{2(1-\gamma)^2\sigma^2} - \frac{r - p}{1 - \gamma}\right]$$

and

$$a > 0 \quad \text{and} \quad b > 0.$$

There are no upper bounds on a (or b, by extension). In Figs. 3 and 4, a and b are allowed to be as large as 5.0.

Consistent with the results in Table IV, Fig. 3 shows that (to a first-order approximation) expected consumption growth is completely insensitive to variations in habit persistence, a, and intensity, b. Expected consumption growth only responds to movements in γ. Therefore, to match the observed average growth rate of 0.0207 over the 1949–2000 period, γ cannot be (substantially) lower than the value of −6.4 chosen for use in Table VII.

[21] As the results in Table V demonstrate, the differences between instantaneous and discrete consumption growth moments can be substantial, particularly for larger values of a.

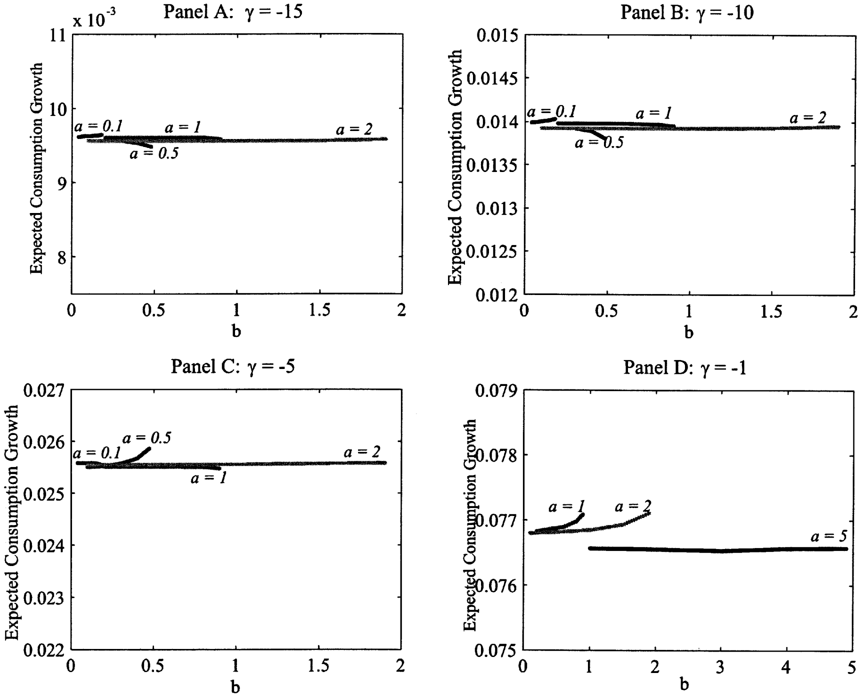

FIG. 3. Mean of discrete, time-aggregated consumption growth, as a function of the habit parameters a and b and the utility parameter γ, where all other parameters are set equal to the values implied by the 1949–2000 data.

Figure 4 shows that, given γ and a, the variance of consumption growth is very sensitive to the choice of b, the habit intensity parameter.

These results suggest the following calibration strategy: (i) Choose the smallest level of γ consistent with discrete (unconditional) expected consumption growth. (ii) Given γ, solve for an increasing sequence of $\{(a_i, b_i)\}_{i=1}^{\infty}$ pairs that match the standard deviation of discrete consumption growth (and satisfy the parameter restrictions in Section 3). Choose the parameter combination with the lowest implied level of value function curvature. This is a computationally intensive procedure, since discrete consumption growth moments are calculated numerically via Monte Carlo simulation.

This strategy, however, is precisely the one that (implicitly) underlies the construction of Table VI. The issue is: How far can the sequence of (a_i, b_i) values be extended and what is the resulting value of the curvature parameter? Table VII provides the answer to this question, and it suggests that it is not feasible to reduce the curvature parameter below (roughly) 10.75.

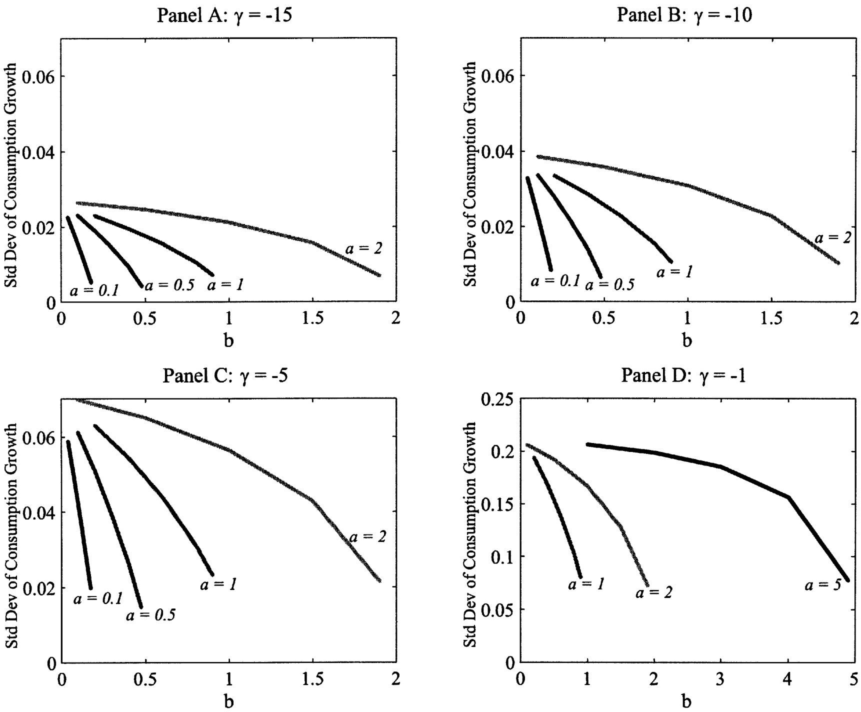

FIG. 4. Variance of discrete, time-aggregated consumption growth, as a function of the habit parameters a and b and the utility parameter γ, where all other parameters are set equal to the values implied by the 1949–2000 data.

There are two problems with extending the analysis beyond the values considered in Table VII. First, while there is no theoretical upper bound on the persistence and intensity parameters, for values of b in excess of 2, it is impossible to compute the moments of instantaneous consumption growth defined in Eqs. (14) and (15).[22] This is not a critical issue, since it is discrete consumption growth that is the appropriate construct to compare to the actual data.

The second—and more important—problem with increasing sequences of (a_i, b_i) pairs is that habit becomes more difficult to interpret. For example, when a equals 25, the half-life of past consumption is less than one third of a month, and the weight, in the habit index, on consumption one month in the past is 0.12. This may or may not be unreasonable, but in the limit, habit becomes infinitely short and infinitely intense, which seems unintuitive. In

[22]Attempts at calculating the integral by numerical quadrature fail with the message that the algorithm encountered a value greater than 10^{307}.

TABLE VII
Expected Value Function Curvature at (a, b) Combinations That Match Consumption Growth Moments, 1949–2000

a	b	$\overline{\Psi}$
0.100	0.099	17.43
0.200	0.188	13.95
0.300	0.279	12.53
0.400	0.371	11.71
0.500	0.466	11.35
1.000	0.957	10.92
2.000*	1.955	10.91
5.000*	4.952	10.81
7.000*	6.951	10.77
10.00*	9.951	10.78
15.00*	14.950	10.74
20.00*	19.950	10.74
25.00*	24.950	10.75
30.00*	29.950	10.75
35.00*	34.950	10.75

Note. $\overline{\Psi} = E[\Psi]$, and the values of a and b are chosen so the discrete, time-averaged consumption data generated by the habit formation model economy match the mean and standard deviation of real consumption growth from 1949 to 2000. The additional parameter values used in constructing the table are $\mu = 0.0730$, $\sigma = 0.1066$, $r = 0.0174$, $\gamma = -6.4$, $\delta_1/\delta_2 = 1.67$, and $\rho = 0.0$. * implies that the corresponding moments of the instantaneous consumption growth process could not be computed numerically for this parameter value combination. The results in the table extend up to the point where it is no longer possible to calculate the discrete consumption growth moments.

any case, the evidence in Table VII suggests that it is impossible to reduce the curvature parameter below (roughly) 10.75.

In summary, Figs. 3 and 4 Tables V–VII demonstrate that habit preferences offer increased flexibility above a purely time separable utility specification [Constantinides (1990) and Table V]. However, the return and consumption growth data from 1949 to 2000 are too great a challenge for a habit model that is required to match the moments of discrete, time-averaged consumption growth (Table VI). In particular, the increase in expected consumption growth and the dramatic decline in consumption volatility are impossible to reconcile with real asset return moments that are (essentially) unchanged from the earlier time periods (Figs. 3 and 4 and Table VII).

The final issue in considering the performance of the habit model is varying ρ, the rate of time discounting. In Section 4.2, I claimed that setting

TABLE VIII
Minimized Expected Value Function Curvature as a Function of the Time Preference Parameter

	Utility parameters			Consumption moments		
ρ	γ	a	b	Mean	SD	min $\overline{\Psi}$
0.00	−6.40	0.600	0.560	2.07	1.23	10.970
0.02	−5.45	0.600	0.575	2.07	1.22	11.495
0.04	−4.48	0.600	0.584	2.06	1.25	12.218
0.06	−3.50	0.600	0.594	2.07	1.24	15.382
0.08	−2.51	0.600	0.603	2.07	1.25	34.888
0.09	−1.98	0.600	0.604	2.08	1.44	109.93

Note. $\overline{\Psi}$ is the $E[\Psi]$. The values of the utility parameters, a, b, and γ, are chosen so that $\overline{\Psi}$ is the minimized subject to the constraint that discrete, time-averaged consumption data generated by the habit formation model economy match the mean and standard deviation of real consumption growth from 1949 to 2000. The additional parameter values used in constructing the table are $\mu = 0.0730$, $\sigma = 0.1066$, $r = 0.0174$, and $\delta_1/\delta_2 = 1.67$.

ρ equal to 0.0 maximized the chance for the habit model to match the mean and variance of post-war consumption growth with a low level of value function curvature. This issue is examined in Table VIII, which shows the minimized value of $\overline{\Psi}$ for a range of values of ρ and for choices of μ, σ, r, and δ_1/δ_2 consistent with the moments of asset returns.

The first row of Table VIII reproduces information from Table VI, and it shows that the minimized value of $\overline{\Psi}$ is 10.970. Each row in the table is associated with a higher level of ρ. The habit parameter a is fixed at 0.600. γ is chosen to ensure that the mean of consumption growth is approximately 2.07% per year, and then b is varied to match the volatility of consumption growth (at 1.23% per year).[23] The clear message from Table VIII is that parameter combinations that match consumption moments with increased values for ρ result in *higher* levels of $\overline{\Psi}$. For choices of ρ equal to 0.02 or 0.04, this increase is not dramatic, but by ρ equal to 0.08 the value of $\overline{\Psi}$ is more than three times the value for ρ equal to 0.0. It is impossible to match the unconditional consumption growth moments for values of ρ in excess of 0.08. Even ρ equal to 0.09 implies a minimum consumption growth volatility of 1.44% per year.

[23]This calibration approach is consistent with the results in Figs. 3 and 4. The choice of a equal to 0.600 is not restrictive, as demonstrated in Table VII.

6. CONCLUSIONS

The equity premium puzzle has proved to be a formidable challenge for standard dynamic asset pricing models. The simple habit model of Constantinides (1990) appeared to be an exception to this rule, and it motivated an explosion in dynamic modeling based on one form or another of a representative agent habit specification. In this paper, I demonstrate that, when the model is calibrated to discrete consumption growth and when the moments of consumption growth are based on the 52 years of data from 1949 to 2000, it is not possible to reconcile asset returns with consumption growth mean and standard deviation for value function curvature parameters below 10.75.

This is not the first paper to point out problems in simple representative agent habit models; see, for example, Jermann (1998) and Lettau and Uhlig (2000). The advantage of this paper is the same as the original Constantinides (1990) article: simplicity. The point of this paper is to show that the simplest habit model is not robust on its "home court" of unconditional asset return and consumption moments. I accept virtually all of the ground rules of the Constantinides analysis: The model evolves continuously in time, the risky real asset return is specified exogenously as iid normal, with given mean and variance, and the risk-free asset is also specified exogenously and independently of the representative agent's utility parameters (in particular, the rate of time preference).

The two variations from Constantinides (1990) that I impose in the calibration exercise seem reasonable. First, the model in Constantinides (1990) was calibrated to consumption data measured at annual intervals. As such, it is impossible to avoid the time-aggregation issues that have been explicitly addressed here. Second, the formal and informal evidence presented here—see also Otrok *et al*. (2001)—documents that it is neither appropriate nor innocuous to regard the consumption data from the entire period from 1890 to 2000 as coming from a stationary process. Since most quantitative business cycle studies are calibrated to post-World War II data, the tougher hurdle of the 1949–2000 data, again, seems reasonable. The fundamental conclusion of the paper is that the simple habit model's only success is not a robust result.

APPENDIX A

Data Description

Consumption and Population Data. The consumption data are based on the most recent comprehensive revision of the National Income and Product Accounts, and they are available from 1929 to 1998. They are available online at http://www.bea.doc.gov/.

• Nominal Personal Consumption Expenditures (PCE) on Nondurable Goods and Nominal PCE on Services, expressed in billions of dollars, are deflated using the Implicit Price Deflator for each series (deflated to 1992 dollars).

• Real consumption expenditures on nondurables and services is defined as the sum of the two deflated series.

• The consumption price deflator is defined as the ratio of nominal nondurables and services to real nondurables and services. The inflation rate is defined as the growth rate in the composite price deflator.

• Total Civilian Non-Institutional Population 16 Years of Age or Older, in thousands, not seasonally adjusted, was taken from the FRED database maintained by the Federal Reserve Bank of St. Louis. Online access to this database is at http://www.stls.frb.org/fred/.

Real per capita consumption, in month t, expressed in thousands of 1992 dollars per person, is

$$\frac{(\text{Real nondurables})_t + (\text{Real services})_t}{(\text{Population}/1000)}.$$

Asset Price and Price Level Data. All asset return series are available from January 1948 to December 1998, for a total of 50 annual observations.

• The short-term interest rate is the yield to maturity (based on the average of the bid and ask quotes) on a three-month Treasury bill. This series is from the FRED database.

• The real risk-free rate series is constructed by subtracting the annual inflation rate (defined above) from the nominal series.

• Annual stock returns are defined from the monthly returns to the CRSP value-weighted portfolio (with dividends) by multiplying the monthly gross returns.

• The real stock return series is constructed by subtracting the annual inflation rate (defined above) from the nominal series.

• The excess return to stocks is constructed by subtracting the real risk-free rate from the real stock return.

REFERENCES

Andrews, D. W. K. (1993). "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica* **61**, 821–856.

Andrews, D. W. K., and Ploberger, W. (1994). "Optimal Tests when a Nuisance Parameter Is Present Only under the Alternative," *Econometrica* **62**, 1383–1414.

Brown, S. J., Goetzmann, W. N., and Ross, S. A. (1995). "Survival," *Journal of Finance* **50**, 853–873.

Campbell, J. Y. (1986). "Bond and Stock Returns in a Simple Exchange Model," *Quarterly Journal of Economics* **101**, 785–803.

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

Chernow, R. (1998). *Titan: The Life of John D. Rockefeller, Sr.*, New York: Random House.

Christiano, L. J., Eichenbaum, M., and Marshall, D. A. (1991). "The Permanent Income Hypothesis Revisited," *Econometrica* **59**, 397–423.

Constantinides, G. M. (1990). "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy* **98**, 519–543.

Cooley, T. F., and Prescott, E. C. (1995). "Economic Growth and Business Cycles," in *Frontiers of Business Cycle Research* (T. F. Cooley, Ed.), Princeton, NJ: Princeton Univ. Press.

Grinblatt, M., and Titman, S. (1997). *Financial Markets and Corporate Strategy*, Boston: Irwin/McGraw–Hill.

Grossman, S. J., and Shiller, R. J. (1981). "The Determinants of the Variability of the Stock Market Prices," *American Economic Review* **71**, 222–227.

Heaton, J. (1993). "The Interaction between Time-Nonseparable Preferences and Time Aggregation," *Econometrica* **61**, 353–385.

Heaton, J., and Lucas D. J. (1998). "Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk," *Journal of Finance* **55**, 1163–1198.

Jermann, U. J. (1998). "Asset Pricing in Production Economics," *Journal of Monetary Economics* **41**, 257–275.

Johnston, J. (1984). *Econometric Methods*, 3rd ed., New York: McGraw–Hill.

Jorion, P., and Goetzmann, W. N. (1999). "Global Stock Markets in the Twentieth Century," *Journal of Finance* **54**, 953–980.

Kloeden, P. E., and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations*, New York: Springer-Verlag.

Kocherlakota, N. R. (1996). "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature* **34**, 42–71.

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

Mankiw, G. N., and Zeldes, S. P. (1991). "The Consumption of Stockholders and Nonstockholders," *Journal of Financial Economics* **29**, 97–112.

Mehra, R., and Prescott, E. C. (1985). "The Equity Premium: A Puzzle," *Journal of Monetary Economics* **15**, 145–161.

Otrok, C., Ravikumar, B., and Whiteman, C. H. (2001). "Habit Formation: A Resolution of the Equity Premium Puzzle?" manuscript, Department of Economics, University of Virginia.

Polberger, W., and Krämer, W. (1990). "The Local Power of the Cusum and Cusum Squares Tests," *Econometric Theory* **6**, 335–347.

Polkovnichenko, V. (2000). "Heterogeneous Labor Income and Preferences: Implications for Stock Market Participation," manuscript, Carlson School of Management, University of Minnesota.

Romer, C. D. (1989). "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1869–1908," *Journal of Political Economy* **97**, 1–37.

Shiller, R. J. (1982). "Consumption, Asset Markets, and Macroeconomic Fluctuations," *Carnegie–Rochester Series on Public Policy* **17**, 203–238.