THE JOURNAL OF FINANCE • VOL. LIV, NO. 5 • OCTOBER 1999

Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach

MICHAEL W. BRANDT*

ABSTRACT

This paper develops a nonparametric approach to examine how portfolio and consumption choice depends on variables that forecast time-varying investment opportunities. I estimate single-period and multiperiod portfolio and consumption rules of an investor with constant relative risk aversion and a one-month to 20-year horizon. The investor allocates wealth to the NYSE index and a 30-day Treasury bill. I find that the portfolio choice varies significantly with the dividend yield, default premium, term premium, and lagged excess return. Furthermore, the optimal decisions depend on the investor's horizon and rebalancing frequency.

How does portfolio and consumption choice depend on variables that forecast time-varying investment opportunities? Prior studies that address this question assume a statistical model relating returns to forecasting variables and solve for an investor's portfolio and consumption choice using estimates of the implied conditional distribution of returns. As a result, their answers are shaped as much by modeling assumptions as by the data. An incorrect model of how returns relate to forecasting variables can yield inconsistent portfolio and consumption choice estimates and invalid inferences. This paper develops and implements an econometric approach that is robust to such model misspecification. Sample analogues of the conditional Euler equations, the first-order conditions of the investor's expected utility maximization, yield consistent estimates shaped by the data.

I modify the method of moments approach of Hansen and Singleton (1982). I fix the parameters of an individual investor's utility function and estimate the optimal wealth and consumption process, and thereby the investor's portfolio and consumption rules, from sample analogues of the conditional Euler equations. In contrast, Hansen and Singleton use observations of the aggregate wealth and consumption process to estimate the parameters of the representative investor's utility function from otherwise identical moment conditions.

* The Wharton School, University of Pennsylvania. I thank Yacine Aït-Sahalia, George Constantinides, Lars Peter Hansen, Campell Harvey, and especially John Cochrane for their guidance and encouragement. I also thank Robert Hodrick, René Stulz (the editor), two anonymous referees, and seminar participants at the 1998 Western Finance Association conference, Columbia University, Duke University, MIT, Northwestern University, Ohio State University, Stanford University, the University of California at Berkeley and Los Angeles, the University of Chicago, the University of Pennsylvania, the University of Southern California, and Yale University for many helpful comments. I am responsible for any remaining errors.

When investment opportunities are time-varying, an investor follows decision rules that map the state of nature, characterized by observable forecasting variables, into the portfolio and consumption choice in that state. To infer the decision rules without parameterizing how the decisions depend on the forecasting variables, I separately estimate the portfolio and consumption choice in every state. I use Hansen and Singleton's approach conditional on a given state, and hence for some realization of the forecasting variables, to infer the portfolio and consumption choice in that state. Repeating this for every state, one state at a time, yields consistent pointwise estimates of the decision rules. In other words, I nonparametrically estimate the decision rules in a method of moments framework.

To understand the main ideas of this paper, consider the following stylized single-period portfolio choice. An investor with preferences $u(W)$ decides on the fraction α of wealth to allocate to a portfolio of equities. Any remaining wealth is invested in a riskless security. Equities yield an uncertain return $\tilde{R}^e_{t+1}$ in excess of the risk-free rate R^f. Investment opportunities, meaning the distribution of returns, may vary through time. If they do, they are correlated with an observable forecasting variable z_t. The following conditional Euler equation characterizes the portfolio choice as a function of this forecasting variable:

$$\alpha(z) = \{\alpha : E[u'(R^f + \alpha\tilde{R}^e_{t+1})\tilde{R}^e_{t+1} | z_t = z] = 0\}. \tag{1}$$

This equation says: "the optimal portfolio choice $\alpha(z)$ in the state z is the decision α that equates to zero the expectation of the investor's marginal utility, given that the observable forecasting variable z_t is equal to z."

With constant investment opportunities, the portfolio choice is independent of z. This implies the decision rule is constant: $\alpha(z) = \alpha$. Then, by the law of iterated expectations, the portfolio choice is also identified by the unconditional expectation of equation (1). Replacing this unconditional expectation with a sample average yields a consistent estimator of α:

$$\hat{\alpha}_T = \left\{\alpha : \frac{1}{T}\sum_{t=1}^{T} u'(R^f + \alpha R^e_{t+1})R^e_{t+1} = 0\right\}. \tag{2}$$

The simplest form of time-varying investment opportunities is having two states of nature. Suppose one state is identified by $z_t \leq \bar{z}$ and the other state by $z_t > \bar{z}$. The decision rule consists of the portfolio choice in each state. The portfolio choice when z is less than $\bar{z}$, or when z is greater than $\bar{z}$, equates to zero the expected marginal utility conditional on being in the respective state. Replacing each conditional expectation with a sample average computed only with returns realized in the given state yields a consistent estimator of the decision rule:

$$\hat{\alpha}_T(z) = \left\{\begin{array}{ll} \alpha : \dfrac{1}{T}\displaystyle\sum_{\{t: z_t \leq \bar{z}\}} u'(R^f + \alpha R^e_{t+1})R^e_{t+1} = 0 & \text{if } z \leq \bar{z} \\[2ex] \alpha : \dfrac{1}{T}\displaystyle\sum_{\{t: z_t > \bar{z}\}} u'(R^f + \alpha R^e_{t+1})R^e_{t+1} = 0 & \text{if } z > \bar{z} \end{array}\right\}. \tag{3}$$

This example illustrates the conditional method of moments. I estimate pointwise the decision rules by replacing conditional population moments with conditional sample moments.[1] This idea extends to continuously varying investment opportunities. In that case, I replace conditional expectations with sample averages computed not only with returns realized in a given state, but also with returns realized in states "similar" to it. This modification is necessary because in a finite sample there are not enough repeat realizations of every possible state.

The economic motivation of this paper arises from overwhelming evidence of time-varying risk premia and volatility.[2] Kandel and Stambaugh (1996) and Solnik (1993) show that weak predictive regressions still yield economically significant variations in the portfolio choice of a single-period investor. Barberis (1999) extends Kandel and Stambaugh's analysis to long-horizon predictability and documents even stronger effects on portfolio choice.

Merton (1969) theorizes that if investment opportunities are time-varying, the portfolio choice of a multiperiod investor can differ from that of a single-period investor because of hedging demands. The investor tries to hedge against predictable changes in future investment opportunities. Balduzzi and Lynch (1996) and Barberis (1999) estimate the multiperiod portfolio choice corresponding to standard predictive regressions. They find that multiperiod decisions differ substantially from single-period decisions. Campbell and Viceira (1996) confirm this result by calibrating an approximation of the portfolio and consumption choice of an infinitely lived investor. Finally, Brennan, Schwartz, and Lagnado (1997) come to a similar conclusion for the long-horizon asset allocation of a continuously rebalancing investor.

The above-mentioned papers model the conditional moments of returns as functions of the forecasting variables and specify the dynamics of investment opportunities. Then, they solve for an investor's decisions using estimates of the implied conditional distribution of returns. In contrast, I directly estimate the portfolio and consumption choice from the data. I focus on how the mean-variance ratio, which ultimately determines the portfolio and consumption choice, depends on the forecasting variables. To avoid structural assumptions, I develop a nonparametric approach of recovering single-period and multiperiod decision rules. My estimates and inferences are robust to model misspecification and provide reliable guidance for future research on this topic.

My empirical analysis focuses on single-period and multiperiod decision rules of an investor with constant relative risk aversion (CRRA) and a one-month to 20-year horizon. The investor allocates wealth to the New York

[1] Returns and forecasting variables must have a time-invariant Markov structure. If the relation between returns and forecasting variables is time-varying or if investment opportunities depend on the whole history of forecasting variables, conditional expectations cannot be estimated with conditional sample averages.

[2] Campbell (1987), Campbell and Shiller (1988), Fama and Schwert (1977), Fama and French (1988a, 1989), Ferson and Harvey (1991), and Keim and Stambaugh (1986), among others, document time-varying risk premia. Black (1976), Bollerslev, Chou, and Kroner (1992), French, Schwert, and Stambaugh (1987), Schwert (1989), and Pagan and Schwert (1990), among others, present evidence of time-varying volatility.

Stock Exchange (NYSE) index and a 30-day Treasury bill. Decisions are made conditional on the dividend yield, default premium, term premium, and lagged excess return. I estimate how the single-period portfolio and consumption choice depends on these forecasting variables and on the investor's horizon. I also illustrate how the hedging demand, the difference between the multiperiod and myopic portfolio choices, depends on the forecasting variables, on the horizon, and on the rebalancing frequency. Finally, I analyze the sensitivity of consumption to the forecasting variables.

Section I sets up the portfolio and consumption choice of a finite-horizon investor and derives the conditional Euler equations that characterize these decisions as a function of the forecasting variables. Section II then develops the conditional method of moments. The empirical analysis follows in Section III. Section IV concludes the paper.

I. Portfolio and Consumption Choice

A. Expected Utility Maximization

Consider the portfolio and consumption choice of a finite-horizon investor who maximizes the expected utility of lifetime consumption. Assume the utility of lifetime consumption is additively time separable, exhibits constant relative risk aversion, and, without loss of generality, precludes bequests to future generations. Then, an investor maximizes the conditional expectation of:

$$u(C_0, C_1, \ldots, C_T) = \sum_{t=0}^{T} \beta^t u(C_t), \tag{4}$$

where

$$u(C_t) = \begin{cases} \dfrac{C_t^{1-\gamma}}{1-\gamma} & \text{for } \gamma > 0 \text{ and } \gamma \neq 1; \\ \ln(C_t) & \text{for } \gamma = 1. \end{cases} \tag{5}$$

Two parameters describe CRRA preferences. The discount factor β measures patience, the willingness to give up consumption today for consumption tomorrow. The coefficient γ captures risk aversion, the reluctance to trade consumption for a fair gamble over consumption today.

Each period t, the investor consumes a fraction c_t of wealth W_t and allocates savings to a risk-free security with return $R^f \equiv \exp(r^f)$ and to N risky securities with excess returns $\tilde{R}^e_{t+1} \equiv \exp(\tilde{r}_{t+1}) - \exp(r^f)$, where $\tilde{r}_{t+1}$ and r^f are continuously compounded rates of return. In allocating fractions α_t of savings to the risky securities, the investor faces the intertemporal budget constraint:

$$\widetilde{W}_{t+1} = (1 - c_t) W_t \tilde{R}^p_{t+1}, \tag{6}$$

where $\tilde{R}^p_{t+1} \equiv R^f + \alpha_t' \tilde{R}^e_{t+1}$ is the gross return on wealth generated by the portfolio choice.

For CRRA preferences the portfolio and consumption rules are only functions of the current state of nature and the horizon. As a result, given a K-dimensional vector of forecasting variables z_t that captures the current state, the decision rules are $\alpha_t = \alpha(z_t, T-t)$ and $c_t = c(z_t, T-t)$. Since there are no bequests, the investor consumes everything at the end of the horizon. This implies $\alpha(z,0) = 0$ and $c(z,0) = 1$.

Unless the investor is myopic (the log utility case) the portfolio and consumption choice anticipates future changes in investment opportunities. To formalize this idea, define the investor's indirect utility as:

$$\begin{aligned} V(z_t, W_t, T-t) &= \max_{\{\alpha_\tau, c_\tau\}_{\tau=t}^{T}} E\left[\sum_{\tau=t}^{T} \beta^{\tau-t} u(C_\tau) \middle| z_t\right] \\ &= \max_{\alpha_t, c_t} u(c_t W_t) + \beta E[V(\tilde{z}_{t+1}, \widetilde{W}_{t+1}, T-t-1) | z_t], \end{aligned} \tag{7}$$

subject to the budget constraint. This indirect utility V is the expected utility of current and future portfolio and consumption choices given the current state, wealth, and horizon. Its definition demonstrates that the investor's intertemporal optimization is equivalent to a single-period portfolio and consumption choice with state-, wealth-, and horizon-dependent utility $u(C_t) + V(\tilde{z}_{t+1}, \widetilde{W}_{t+1}, T-t-1)$.

A convenient feature of CRRA preferences is that its indirect utility separates into two terms, functions of next period's consumption choice and of next period's utility of wealth:

$$V(z_t, W_t, T-t) = \psi(z_t, T-t) u(W_t), \tag{8}$$

where

$$\psi(z_t, T-t) = \begin{cases} c(z_t, T-t)^{-\gamma} & \text{for } \gamma > 0 \text{ and } \gamma \neq 1; \\ 1 & \text{for } \gamma = 1. \end{cases} \tag{9}$$

Equation (8) follows from the homotheticity of CRRA preferences and equation (9) from solving the envelope condition $\partial V/\partial W = \partial u/\partial C$ for ψ.

B. Conditional Euler Equations

CRRA utility and indirect utility are globally concave. As a result, there exists an interior solution to the portfolio and consumption choice. This implies that the decision rules are characterized by the following $N + 1$ conditional Euler equations, the first-order conditions of the investor's expected utility maximization:

$$\begin{bmatrix} \alpha(z_t, T-t) \\ c(z_t, T-t) \end{bmatrix} = \left\{ \begin{bmatrix} \alpha \\ c \end{bmatrix} : E[m_{t+1}(\alpha, c, T-t) | z_t] = 0 \right\}, \tag{10}$$

where

$$m_{t+1}(\alpha,c,T-t) = \begin{bmatrix} \psi(\tilde{z}_{t+1},T-t-1)u'(\widetilde{W}_{t+1})\tilde{R}^{e}_{t+1} \\ \beta\psi(\tilde{z}_{t+1},T-t-1)\,\dfrac{u'(\widetilde{W}_{t+1})}{u'(cW_t)}\,\tilde{R}^{p}_{t+1} - 1 \end{bmatrix}, \tag{11}$$

and W_{t+1} and ψ are given by equations (6) and (9), respectively.

II. Econometric Approach

A. Method of Moments

To illustrate how sample analogues of the Euler equations generate consistent estimates of the portfolio and consumption decisions, assume investment opportunities are constant. Constant investment opportunities imply that the CRRA portfolio and consumption choice is independent of both the forecasting variables and the horizon. This means in equation (10) the $\{\alpha,c\}$ that set the conditional expectation of m_{t+1} to zero also set their unconditional expectation to zero. As long as returns are stationary, replacing this unconditional expectation with an historic average yields a consistent estimator:

$$\begin{bmatrix} \hat{\alpha}_T \\ \hat{c}_T \end{bmatrix} = \left\{ \begin{bmatrix} \alpha \\ c \end{bmatrix} : \frac{1}{T}\sum_{t=1}^{T} m_{t+1}(\alpha,c,\tau) = 0 \right\}. \tag{12}$$

Of course, the premise of this paper is that investment opportunities are time-varying and that the CRRA portfolio and consumption choice is a function of the observed state. To extend this traditional method of moments approach, we need to parameterize the decision rules. Unfortunately, unless we know exactly how returns relate to forecasting variables, there is no theoretically correct parameterization of the portfolio and consumption choice.[3]

B. Conditional Method of Moments

The basic idea of the conditional method of moments is that sample analogues to the conditional, not to the unconditional, expectation of m_{t+1} generate consistent estimates of the decisions under time-varying investment opportunities. To understand this point, consider first the case of discretely varying investment opportunities.

Suppose investment opportunities assume only one of S states: $z = \{z_1,z_2,\ldots,z_S\}$. The portfolio and consumption rules are the S sets of decisions $\{\alpha(z_s,\tau),c(z_s,\tau)\}_{s=1}^{S}$ that equate to zero the conditional expectations

[3] For some return generating processes the implied parameterization of the decision rules can be derived. For example, Campbell and Viceira (1996) show that if returns are homoskedastic log-normal with a risk premium that is linear in an AR(1) state variable, the portfolio (consumption) choice of an infinitely lived investor with CRRA utility is approximately linear (quadratic) in the state variable.

$E[m_{t+1}|z = z_s]$. Replacing these conditional expectations with historic averages, each constructed only with returns observed in state s, gives the estimator:

$$\begin{bmatrix} \hat{\alpha}_T(z,\tau) \\ \hat{c}_T(z,\tau) \end{bmatrix} = \begin{cases} \begin{bmatrix} \alpha \\ c \end{bmatrix} : \dfrac{1}{T_1} \sum\limits_{\{t: z_t = z_1\}} m_{t+1}(\alpha, c, \tau) = 0 & \text{if } z = z_1 \\ \begin{bmatrix} \alpha \\ c \end{bmatrix} : \dfrac{1}{T_2} \sum\limits_{\{t: z_t = z_2\}} m_{t+1}(\alpha, c, \tau) = 0 & \text{if } z = z_2 \\ \qquad \cdots \\ \begin{bmatrix} \alpha \\ c \end{bmatrix} : \dfrac{1}{T_S} \sum\limits_{\{t: z_t = z_S\}} m_{t+1}(\alpha, c, \tau) = 0 & \text{if } z = z_S. \end{cases} \tag{13}$$

where T_s is the number of times state s is observed in the sample of size $T = \sum_{s=1}^{S} T_s$.

This example illustrates how the conditional method of moments works. I consider the portfolio and consumption choice in each state s as independent of the others and apply the traditional method of moments, equation (12), only to returns observed in state s. This yields consistent estimates of the portfolio and consumption choice in state s. Collectively, the S estimates reveal how the investor's decisions depend on the forecasting variables. In other words, the conditional method of moments yields a pointwise, or nonparametric, estimate of the portfolio and consumption rules.

In principle, the same approach applies when investment opportunities vary continuously. The only problem is that for any state z, the finite sample at hand may not contain enough observations from state z to reliably compute a historic average of m_{t+1} with only returns observed in that state. An intuitive solution to this problem is to estimate the portfolio and consumption choice in state z not only with returns observed in state z, but also with returns observed in states "similar" to z. The remainder of this section formalizes this solution.

Define a weighting function $\omega((z - z_t)/h_T)$ to measure how similar an observed state z_t is to some reference state z. Then, to estimate the portfolio and consumption choice in state z, simply replace the conditional expectation of m_{t+1} with an ω-weighted average of m_{t+1}:[4]

$$\begin{bmatrix} \hat{\alpha}_T(z,\tau) \\ \hat{c}_T(z,\tau) \end{bmatrix} = \left\{ \begin{bmatrix} \alpha \\ c \end{bmatrix} : \frac{1}{T h_T^K} \sum_{t=1}^{T} \omega\left(\frac{z - z_t}{h_T} \right) m_{t+1}(\alpha, c, \tau) = 0 \right\}. \tag{14}$$

[4] Luttmer (1997) uses a similar econometric approach to estimate transaction cost bounds that reconcile consumption asset pricing models with historic data on aggregate consumption and returns.

The parameters h_T, so-called bandwidths, scale the difference between the observed state z_t and the reference state z. The scaled differences determine how similar these two states are. The factor Th_T^K ensures that the weighted sum is a nondegenerate average as $h_T \to 0$.

One way to understand the estimator is to think of the weighting function as a data window centered on z and of the bandwidth as the window's width. Such an estimator uses an observation t only if z_t falls within the data window. If z_t falls outside the data window, the state is judged too "dissimilar" from z and the estimator ignores this observation.

A more sophisticated weighting scheme uses all observations, but places greater emphasis on data from states that are very similar to the reference state. One weighting function that accomplishes this is the product of K standard normal densities:

$$\omega(u) = \prod_{i=1}^{K} k(u_i) = \sum_{i=1}^{K} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} u_i^2\right). \tag{15}$$

The bandwidths can still be interpreted as window widths. On one hand, a greater h_T means an observation from state $z_t \neq z$ is weighted downward less for being different from the z. On the other hand, a smaller h_T means that this observation is weighted downward more. At the extremes, $h_T = \infty$ corresponds to the unconditional estimator, equation (12), and $h_T = 0$ corresponds to the discrete state estimator, equation (13). Interpreting bandwidths as window widths also helps clarify how the optimal h_T varies with the sample size and across forecasting variables. Suppose a window must contain at least 10 observations. A small sample or widely dispersed forecasting variable requires a wider window than does a large sample or less dispersed forecasting variable.[5]

Like any statistical approach, the conditional method of moments has its advantages and its disadvantages. The advantage is being nonparametric. The estimator is less biased than incorrectly specified parametric estimators, and at least it is consistent. However, this comes at the cost of losing observations. As a result, the variance of the estimator exceeds the variance of a correctly specified parametric estimator. h_T is set to achieve a mean squared error optimal balance between the bias and variance for a sample of size T.

Up to this point, the discussion implicitly assumes that we can construct m_{t+1} from the data. However, because the indirect utility function is separable and recursive, $m_{t+1}(\alpha,c,\tau)$ in equation (11) depends on the unknown consumption choice $c(z_{t+1},\tau - 1)$, for all $\tau > 1$. To understand the extent of

[5] Consistency of the estimator only requires that $h_T \to 0$ fast enough as $T \to \infty$. This means that, without loss of generality, the cross-sectional subscripts on the bandwidths can be ignored for ease of notation. However, in practice the optimal bandwidths in a sample of size T are different across forecasting variables.

this problem, consider the two-period portfolio and consumption choice. If we want to estimate $\{\alpha(z,2), c(z,2)\}$ we need $\{m_{t+1}(\alpha,c,2)\}_{t=1}^{T}$, each of which depends on $\{c(z_{t+1},1)\}_{t=1}^{T}$. In turn, to estimate any $c(z_{t+1},1)$ we need $\{m_{t+1}(\alpha,c,1)\}_{t=1}^{T}$, which we can construct because $c(z,0) = 1$.

Fortunately, this problem appears more serious than it actually is. Because the portfolio and consumption choice is only forward looking, we can estimate the τ-period decision rules two ways: either jointly estimate $\{\alpha(z_t,j), c(z_t,j)\}$, for $t = 1,2,\ldots,T$ and $j = 1,2,\ldots,\tau$, by stacking the τT sets of moment conditions $E[m_{t+1}(\alpha,c,j)|z = z_t]$, or recursively estimate the decision rules starting with $\alpha(z,0) = 0$ and $c(z,0) = 1$. The resulting estimates and their properties are identical.

C. *Asymptotics*

Intuitively, the conditional method of moments yields consistent estimates of the decision rules because the weighted average of m_{t+1} converges to the conditional expectation of m_{t+1} uniformly for all $\{\alpha,c\}$, τ, and z. This fact, and the features of CRRA utility, are the key ingredients to establishing the following asymptotic result.

Assume $\{R^e_{t+1}, z_t\}_{t=1}^{T}$ are realizations of a strictly stationary process, the weighting function $\omega(u)$ is the product of K univariate continuous and bounded functions $k(u_i)$ with:

$$\int_{\mathbb{R}} k(u)\,du = 1, \qquad \int_{\mathbb{R}} uk(u)\,du = 0, \quad \text{and} \qquad \int_{\mathbb{R}} u^2 k(u)\,du < \infty, \tag{16}$$

the bandwidths h_T satisfy:

$$h_T \to 0, \qquad Th_T^{K+4} \to 0, \quad \text{and} \quad Th_T^{K} \to \infty \quad \text{as } T \to \infty, \tag{17}$$

the function $\psi(z,\tau-1)$ is known, and the regularity conditions in the Appendix apply, then:

$$\sqrt{Th_T^K}\left(\begin{bmatrix}\hat{\alpha}_T(z,\tau)\\ \hat{c}_T(z,\tau)\end{bmatrix} - \begin{bmatrix}\alpha(z,\tau)\\ c(z,\tau)\end{bmatrix}\right) \tag{18}$$

converges in distribution to a multivariate normal random vector with zero mean and variance-covariance matrix:

$$\Sigma(z,\tau) = \frac{D(z,\tau)^{-1}V(z,\tau)D(z,\tau)^{-1}}{f_z(z)} \int_{\mathbb{R}^K} \omega^2(u)\,du, \tag{19}$$

where $f_z(z)$ is the unconditional density of z,

$$D(z,\tau) = E\left[\frac{\partial m_{t+1}(\alpha,c,\tau)}{\partial[\alpha',c]}\middle| z\right], \tag{20}$$

$$\text{and} \quad V(z,\tau) = E[m_{t+1}(\alpha,c,\tau)m_{t+1}(\alpha,c,\tau)'|z]. \tag{21}$$

The Appendix sketches a proof of this result.

The asymptotics resemble those of the traditional method of moments, except for a slower rate of convergence in (18) and the fact that the expectations in equation (20) are conditional on the forecasting variables.[6] The slower convergence rate of $\sqrt{Th_T^K}$, instead of the parametric convergence rate of $\sqrt{T}$, is the asymptotic cost of being nonparametric. This cost increases exponentially with the number of forecasting variables K. As a result, this is often referred to as the "curse of dimensionality" of nonparametric estimators.

A subtle difference between the limiting distribution above and that of the traditional method of moments is that in the latter $V = \sum_{j=-\infty}^{\infty} E[m_{t+1}m'_{t+j}]$. This expression simplifies to $V = E[m_{t+1}m'_{t+1}]$ only if marginal utility is serially uncorrelated. By taking weighted averages of m_{t+1} that are increasingly more focused on a single state z as $T \to \infty$, the conditional method of moments asymptotically eliminates the effect of serial dependence in the data. Although this feature is common to other nonparametric methods and conveniently simplifies the asymptotics here, Robinson (1983) warns about exploiting such extreme results to draw inferences in finite samples.

The above result assumes that the function ψ is known. Of course, for $\tau > 1$ this is not the case. However, the asymptotics of the estimator extend to the case of an unknown ψ. Recall that one way to estimate the τ-period decision rules is to jointly estimate $\{\alpha(z_t,j),c(z_t,j)\}$, for $t = 1,2,\ldots,T$ and $j = 1,2,\ldots,\tau$, by stacking the τT sets of moment conditions $E[m_{t+1}(\alpha,c,j)|z = z_t]$. As a result, it is straightforward to alter the above result to state that jointly $\{\alpha(z_t,j),c(z_t,j)\}$, for $t = 1,2,\ldots,T$ and $j = 1,2,\ldots,\tau$, are consistent and asymptotically normally distributed.

To understand how the conditional method of moments relates to other nonparametric methods, the asymptotics suggest an alternative and insightful interpretation of the estimator. Suppose we observe at time $t + 1$ a noisy measure of the CRRA portfolio and consumption choice at time t:

$$\begin{bmatrix} \tilde{\alpha}_{t+1} \\ \tilde{c}_{t+1} \end{bmatrix} = \begin{bmatrix} \alpha(z_t,\tau) \\ c(z_t,\tau) \end{bmatrix} + \bar{\nu}_{t+1}, \tag{22}$$

[6] The estimator inherits the standard asymptotic bias of a nonparametric regression. However, conditions (16) and (17) guarantee that this bias is negligible as $T \to \infty$. Alternatively, one could explicitly account for the bias by modifying the moment conditions. Such bias corrections are suggested by Fan, Heckman, and Wand (1995) and by Gozalo and Linton (1994) in different contexts.

where $\nu = D(z_t,\tau)^{-1}m_{t+1}(\alpha(z_t,\tau),c(z_t,\tau),\tau)$ is a zero mean measurement error. Then, the asymptotics of the conditional method of moments are the same as those of a nonparametric regression of the observed portfolio and consumption choices $\{\alpha_{t+1},c_{t+1}\}$ on the forecasting variables z_t.[7] This implies that standard results for nonparametric regression carry over to this interpretation of the conditional method of moments.

D. Weighting Function and Bandwidth

The product of normal densities, equation (15), is not the only weighting function that yields consistent estimates. Admissible weighting functions come in a variety of shapes. Their common feature is that the weighted average of m_{t+1} converges uniformly to the conditional expectation of m_{t+1}. Although the estimator's efficiency depends on the shape of the weighting function, simulation studies for other nonparametric estimators show that the shape of the weighting function is far less important to the asymptotics than choosing the bandwidths as a function of the sample size.[8] Using a product of standard normal densities as weighting function is convenient because it is symmetric, uniformly positive, and computationally efficient.

Choosing the bandwidths is the real obstacle to applying the conditional method of moments. To appreciate the importance of bandwidth choice, consider how the properties of the estimator change as I vary the bandwidths. On one hand, if I increase h_T and thereby widen the spread of the weighting function, the estimator uses more data on the whole but differentiates less between data from different states. This decreases the variance of the estimator, but increases the potential bias by averaging across observations that are less "similar." On the other hand, if I decrease h_T, and thereby narrow the spread of the weighting function, the estimator differentiates more between data from different states but uses fewer data on the whole. This decreases the potential bias of the estimator, but it increases the variance. Optimal bandwidths achieve a desirable balance between bias and variance.

The conventional metric for trading off bias against variance is to minimize the mean squared error (MSE) of the estimates. Applying standard results from the nonparametric regression literature to the transformed regression interpretation, equation (22), the bandwidths that minimize the average MSE of the estimates are of the form

$$h_T = \lambda\sigma(z)T^{-(1/(K+4))}, \tag{23}$$

[7] Tsybakov (1982) and Härdle (1984) establish that robust nonparametric regressions have this dual interpretation. Both estimators have a limiting normal distribution with the same asymptotic bias and variance.

[8] Silverman (1986) and Härdle (1990) summarize the literature on optimal weighting functions for nonparametric regression estimators. They conclude that over a large class of admissible weighting functions, differences in the mean squared error of the estimators are small. The asymptotic consequences of choosing a substantially suboptimal weighting function are less serious than misspecifying the optimal bandwidths by only 10 percent.

where $\sigma(z)$ is the vector of unconditional standard deviations of the forecasting variables. The constant λ depends on $\Sigma(z,\tau)$, the first and second derivatives of $\{\alpha(z,\tau),c(z,\tau)\}$ with respect to z, and $f_z(z)$. Unfortunately, none of these quantities are known.

There are many ways to choose λ in practice. I use "leave-one-out" cross-validation, which is a data-driven approach. It minimizes the average squared error in out-of-sample prediction. In this case, leave-one-out cross-validation works as follows: For every observation $\{R_{t+1},z_t\}$, use the remaining observations $\{R_{j+1},z_j\}_{j=1,j\neq t}^{T}$ to estimate $\{\alpha(z_t,\tau),c(z_t,\tau)\}$ and compute an out-of-sample predictive error $\nu_{t+1} = \{\alpha_{t+1} - \hat{\alpha}(z_t,\tau),\ c_{t+1} - \hat{c}(z_t,\tau)\}$. Although the observation $\{\alpha_{t+1},c_{t+1}\}$ does not exist, the predictive error can alternatively be constructed from $\nu = D^{-1}m_{t+1}$ with an estimate of $D(z,\tau)$. Then, choose λ to minimize the sum of squared predictive errors $\sum_{t=1}^{T}\nu_{t+1}'\nu_{t+1}$.

Following Stone's (1976) introduction of cross-validation for model selection, Clark (1975) first uses this technique to choose bandwidths for nonparametric estimators. Härdle and Marron (1985) prove that leave-one-out cross-validation yields asymptotically optimal bandwidths in a standard nonparametric regression. Härdle (1984) extends this result to a nonparametric regression on pseudo-data, such as that defined by equation (22). This last paper formally justifies the above-described leave-one-out cross-validation procedure and guarantees its asymptotic optimality for the conditional method of moments.

E. Finite-Sample Properties

Although the conditional method of moments is consistent, its finite-sample properties are unknown.[9] The accuracy of the estimator can be judged in two ways: compare my results to other consistent estimates or conduct controlled experiments. I use the second approach because other estimates of portfolio and consumption rules may themselves be inconsistent.

Table I compares the finite-sample properties of the conditional method of moments to that of the traditional regression approach. The results are based on 5000 independent samples of $T = 300$, 600, or 1200 observations. Returns are generated by the regression:

$$\ln(1 + R_{t+1}^{e}) = 0.059 + 0.016\ln(z_t) + \epsilon_{t+1}, \quad \text{where } \epsilon_{t+1} \sim \mathrm{N}(0,0.001). \tag{24}$$

In Panel A the forecasting variable is distributed as $\ln(z_t) \sim \mathrm{N}(-3.455,0.125)$. In Panel B it is generated by the autoregression

$$\ln(z_{t+1}) = -0.023 + 0.993\ln(z_t) + \xi_{t+1}, \quad \text{where } \xi_{t+1} \sim \mathrm{N}(0,0.002). \tag{25}$$

[9] Pritsker (1998) illustrates that persistence in financial data can substantially deteriorate the finite-sample properties of nonparametric estimators. Then, asymptotic results are unreliable even in samples of several thousand observations.

Table I

Finite-Sample Properties of the Conditional Method of Moments

This table shows average conditional method of moments and regression approach estimates of the single-period portfolio choice of an investor with constant relative risk aversion $\gamma = 5$. The results are based on 5000 independent samples of $T = 300$, 600, or 1200 excess returns R^e_{t+1} and realizations of a forecasting variable z_t. Returns are generated by the regression:

$$\ln(1 + R^e_{t+1}) = 0.059 + 0.016\ln(z_t) + \epsilon_{t+1}, \quad \text{where } \epsilon_{t+1} \sim \mathrm{N}(0,0.001).$$

In Panel A the forecasting variable is distributed as $\ln(z_t) \sim \mathrm{N}(-3.455,0.125)$. In Panel B it is generated by the autoregression:

$$\ln(z_{t+1}) = -0.023 + 0.993\ln(z_t) + \xi_{t+1}, \quad \text{where } \xi_{t+1} \sim \mathrm{N}(0,0.002).$$

The estimates represent the fraction of savings allocated to equities when the forecasting variable is equal to its 10th, 25th, 50th, 75th, and 90th percentiles. Standard deviations of the estimates are in parentheses. p-values of the difference between two portfolio choice estimates are below the underbraces. The results for $T = \infty$ correspond to the true portfolio choice.

Sample Size	Conditional Method of Moments					Regression Approach				
	Percentile of z					Percentile of z				
	10%	25%	50%	75%	90%	10%	25%	50%	75%	90%
Panel A: Average Estimates for an Independent Forecasting Variable										
$T = 300$	0.10	0.25	0.40	0.56	0.72	0.11	0.27	0.44	0.61	0.76
	(0.45)	(0.45)	(0.46)	(0.49)	(0.53)	(0.33)	(0.31)	(0.29)	(0.30)	(0.32)
$T = 600$	0.11	0.25	0.41	0.58	0.73	0.11	0.27	0.44	0.61	0.76
	(0.34)	(0.34)	(0.34)	(0.36)	(0.39)	(0.22)	(0.21)	(0.20)	(0.21)	(0.22)
		0.09		0.12			0.02		0.02	
			0.08					0.01		
$T = 1200$	0.11	0.26	0.42	0.59	0.74	0.12	0.27	0.44	0.61	0.76
	(0.25)	(0.25)	(0.26)	(0.27)	(0.29)	(0.16)	(0.15)	(0.14)	(0.15)	(0.16)
$T = \infty$	0.12	0.27	0.44	0.61	0.76	0.12	0.27	0.44	0.61	0.76
Panel B: Average Estimates for a Persistent Forecasting Variable										
$T = 300$	0.29	0.40	0.53	0.63	0.71	0.20	0.32	0.49	0.64	0.78
	(0.35)	(0.37)	(0.38)	(0.36)	(0.35)	(0.41)	(0.44)	(0.44)	(0.45)	(0.45)
$T = 600$	0.18	0.30	0.45	0.59	0.72	0.14	0.28	0.45	0.63	0.78
	(0.26)	(0.28)	(0.29)	(0.25)	(0.26)	(0.27)	(0.28)	(0.30)	(0.30)	(0.29)
		0.07		0.09			0.01		0.01	
			0.06					0.01		
$T = 1200$	0.13	0.27	0.43	0.60	0.74	0.11	0.26	0.44	0.61	0.77
	(0.21)	(0.22)	(0.23)	(0.25)	(0.26)	(0.18)	(0.19)	(0.20)	(0.20)	(0.19)
$T = \infty$	0.12	0.27	0.44	0.61	0.76	0.12	0.27	0.44	0.61	0.76

Together, equations (24) and (25) are a restricted form of the vector autoregression assumed by Balduzzi and Lynch (1996), Barberis (1999), Brennan et al. (1997), Campbell and Viceira (1996), Kandel and Stambaugh (1996), and Solnik (1993).[10] The first equation implies that the single-period portfolio choice is approximately linear in $\ln(z_t)$. The second equation makes the forecasting variable highly persistent.

The left half of Table I shows average conditional method of moments estimates of the single-period portfolio choice for $\gamma = 5$ and z equal to the 10th, 25th, 50th, 75th, and 90th percentiles of its stationary distribution.[11] To abstract from bandwidth selection, I set $\lambda = 1$ and hence $h = \sigma(z)T^{-1/5}$. This yields a lower bounds on the finite-sample properties. The right half of the table shows average regression approach estimates. Like the above authors, I regress $\ln(1 + R^e_{t+1})$ on $\ln(z_t)$ and solve for the optimal decision given the implied conditional distribution of returns. Finally, $T = \infty$ gives the true portfolio choice implied by the model.

The standard deviations of the estimates are in parentheses. For $T = 600$ the fraction of estimated decision rules for which the portfolio choice for the lower percentile exceeds that for the higher percentile is reported below the underbraces. This statistic is a p-value of the slope of the decision rule between two percentiles.

Panel A of Table I shows that the estimator captures the shape of the decision rule on average. The average estimates nearly match the true portfolio choices, regardless of the sample size. The panel also shows that for $T = 600$ the estimator is reasonably accurate. The p-values are high, although, not surprisingly, they are lower than those of the regression approach. Comparing the results in Panel A to those in Panel B illustrates that for $T = 600$ or 1200 the finite-sample properties of the estimator do not deteriorate if the forecasting variable is persistent.[12]

III. Empirical Results

A. Preferences, Assets, and Forecasting Variables

The empirical results are for an investor with relative risk aversion γ ranging from one (the log utility case) to 10. The horizon ranges from one month to 20 years, and the rebalancing period, the time that elapses

[10] The parameter values are least-squares estimates of the model from monthly data on NYSE index returns, 30-day Treasury bill rates, and aggregate dividend yields. Section III.A describes this data set.

[11] The conditional method of moments cannot estimate the portfolio choice for an unobserved level of the forecasting variable. Therefore, I compute the results for percentile z only from samples for which z lies between the 5th and 95th percentiles of the observations. I apply the same rule to the regression approach.

[12] The precision of the estimator is better in Panel B than in Panel A. This is because of the restriction in footnote 11. Given that z lies well within the range of the sample, there are more similar observations in the sample if the forecasting variable is persistent than if it is independent and identically distributed.

between consecutive portfolio and consumption choices, ranges from one month to four years. The coefficient of time preference is arbitrarily set to $\beta = 0.99$.

The investor has a choice between two securities, the value-weighted NYSE index and a 30-day Treasury bill. Monthly returns for January 1947 to December 1996, a total of 600 observations, are from the Center for Research in Security Prices (CRSP). Nominal returns are deflated by the rate of change in the Consumer Price Index from CITIBASE.

Investment opportunities are described by four variables known to forecast time-varying risk premia and volatility: the dividend yield, default premium, term premium, and lagged excess return. The dividend yield is the sum of reinvested dividends paid on the NYSE index over the past 12 months, divided by the current value of the index. It is computed with returns from CRSP. The default premium is the annualized yield spread between Moody's Baa and Aaa rated bonds. The term premium is the difference in annualized yields of a portfolio of long-term government bonds and a 90-day Treasury bill. All bond yields are from CITIBASE. The lagged excess return is the excess NYSE return measured over the previous rebalancing period. Table II describes the data.

B. Unconditional Estimates

Panel A of Table III shows method of moments estimates of the single-period portfolio choice, the fraction of savings allocated to equities. Panel B shows estimates of the single-period consumption choice, the fraction of wealth consumed. Across columns of Table III both the horizon and rebalancing periods range from one month to four years.

To estimate n-month decisions I apply the unconditional estimator, equation (12), to n-month returns $R_{t+n} \equiv \exp(\sum_{j=1}^{n} r_{t+j})$ and $R^f_{t+n} \equiv \exp(\sum_{j=1}^{n} r^f_{t+j})$. When $n > 1$, these returns overlap, and the estimator uses serially correlated marginal utilities. Therefore, I report in parentheses the Newey and West (1987) estimates of the asymptotic standard errors.

The point estimates in Panel A of Table III are reasonable. Merton (1969) shows that with independent log-normally distributed returns and continuous rebalancing, the CRRA portfolio choice is $\alpha = \gamma^{-1}\ E(\tilde{R}^e_{t+1})/Var(\tilde{R}^e_{t+1})$. Plugging in monthly sample moments from Table II yields portfolio choices very similar to those in the first column of Table III.

This analytical solution also helps to understand the large standard errors in Table III. Suppose we know $Var(\tilde{R}^e_{t+1})$ and only need to estimate $E(\tilde{R}^e_{t+1})$. The standard errors of the resulting portfolio choice estimates are $1/[\gamma\sqrt{T}\,\mathrm{Std}(\tilde{R}^e_{t+1})]$. Plugging in sample moments from Table II yields standard errors as large as those in Table III. This approximation shows that the large standard errors are a feature of the data and are not a shortcoming of the method of moments. Since we cannot accurately measure the unconditional equity risk premium, it is also difficult to estimate the fraction of savings to be allocated to equities.

Table II

Descriptive Statistics of Forecasting Variables and Returns

This table shows descriptive statistics of the dividend yield (Div), default premium (Def), term premium (Term), and excess return (Ret) computed with monthly observations from January 1947 to December 1996. The dividend yield is the sum of reinvested monthly dividends paid on the NYSE index over the past 12 months divided by the current level of the index. The default premium is the difference in annualized yields of Moody's Baa and Aaa rated bonds. The term premium is the difference in annualized yields of a portfolio of long-term government bonds and the 90-day Treasury bill. The excess return is the monthly return on the NYSE index in excess of the 30-day Treasury bill rate, deflated by the rate of change in the Consumer Price Index. The lagged correlations are for a one-month lag.

	Variables			
Statistic	Div	Def	Term	Ret
Mean	3.97	0.92	1.17	0.68
Std. deviation	1.08	0.43	1.30	4.02
Skewness	0.64	1.49	−0.31	−0.37
Kurtosis	2.57	5.18	3.90	5.10
Percentile:				
5%	2.66	0.45	−1.04	−5.89
25%	3.05	0.64	0.47	−1.78
50%	3.81	0.77	1.17	0.92
75%	4.75	1.12	1.88	3.21
95%	6.09	1.81	3.40	6.41
Correlation with:				
Div	1.00			
Def	0.24	1.00		
Term	−0.10	0.08	1.00	
Ret	−0.03	0.04	0.12	1.00
Lagged Div	0.99	0.25	−0.08	0.11
Lagged Def	0.22	0.98	0.11	0.07
Lagged Term	−0.12	0.04	0.96	0.14
Lagged Ret	−0.04	−0.01	0.09	0.05

Differences in estimates along any row of Panel A in Table III suggest that the portfolio choice depends on the horizon. For example, an investor with $\gamma = 5$ chooses a portfolio of 80 percent equities if the horizon is one month. The same investor chooses a portfolio of 70 percent equities if the horizon is one year and a portfolio of 100 percent equities if the horizon is four years.

Figure 1 supports this conclusion. It shows the portfolio choice as a function of the horizon for $\gamma = 1$ (solid line), $\gamma = 2$ (dashed line), $\gamma = 5$ (dashed-dotted line), and $\gamma = 10$ (dotted line). After an initial decline, the portfolio choice increases with the horizon. This is consistent with the regressions of Fama and French (1988b) and the variance ratios of Poterba and Summers (1988). Both studies find that long-horizon returns are negatively serially correlated, which makes them less risky than monthly returns. Short- to medium-horizon returns are positively serially correlated, which makes them more risky than monthly returns.

Table III

Single-Period Portfolio and Consumption Choice as a Function of the Horizon

This table shows method of moments estimates of the unconditional single-period portfolio choice α and consumption choice c of an investor with constant relative risk aversion γ. Both the horizon and rebalancing period range from one month to four years. The intertemporal discount factor is $\beta = 0.99$. The estimates in Panel A represent the fraction of savings allocated to equities. The estimates in Panel B represent the fraction of wealth consumed. The method of moments estimator $\{\hat{\alpha}_T, \hat{c}_T\}$ of $\{\alpha, c\}$ is:

$$\begin{bmatrix} \hat{\alpha}_T \\ \hat{c}_T \end{bmatrix} = \left\{ \begin{bmatrix} \alpha \\ c \end{bmatrix} : \frac{1}{T} \sum_{t=1}^{T} m_{t+1}(\alpha, c, 1) = 0 \right\},$$

where m_{t+1} is the investor's marginal utility, given a portfolio choice α, a consumption choice c, and a horizon of one rebalancing period. Serial-correlation corrected asymptotic standard errors are in parentheses. Note that the consumption choice of an investor with logarithmic utility is deterministic: $c = 1/(1 + \beta)$.

Risk Aversion	Horizon					
	1-Month	3-Month	6-Month	1-Year	2-Year	4-Year
	Panel A: Portfolio Choice					
$\gamma = 1$	3.45	2.88	2.73	2.21	2.08	3.09
	(0.83)	(0.72)	(0.64)	(0.49)	(0.45)	(0.63)
$\gamma = 2$	1.93	1.70	1.69	1.57	1.56	2.41
	(0.45)	(0.39)	(0.36)	(0.29)	(0.25)	(0.36)
$\gamma = 5$	0.79	0.72	0.71	0.69	0.73	1.02
	(0.20)	(0.17)	(0.16)	(0.15)	(0.17)	(0.24)
$\gamma = 10$	0.40	0.37	0.36	0.35	0.37	0.49
	(0.10)	(0.09)	(0.08)	(0.07)	(0.09)	(0.12)
	Panel B: Consumption Choice					
$\gamma = 1$	0.50	0.51	0.52	0.53	0.56	0.62
	na	na	na	na	na	na
$\gamma = 2$	0.50	0.51	0.51	0.53	0.55	0.61
	(0.09)	(0.08)	(0.08)	(0.07)	(0.06)	(0.07)
$\gamma = 5$	0.50	0.50	0.51	0.51	0.53	0.56
	(0.11)	(0.10)	(0.10)	(0.09)	(0.08)	(0.09)
$\gamma = 10$	0.50	0.50	0.50	0.51	0.52	0.54
	(0.11)	(0.10)	(0.10)	(0.09)	(0.08)	(0.09)

Given the large standard errors in Table III, it is unclear whether the portfolio choice differs statistically across horizons. To explore the accuracy of the apparent horizon pattern, I repeat the analysis for an investor with $\gamma = 5$ for 5000 resampled data sets of 600 monthly returns, sampled randomly with replacement from the original data. As expected, the one-month decision is statistically indistinguishable from the two- through four-year decisions. However, in only 11 and nine percent of the resampled data sets

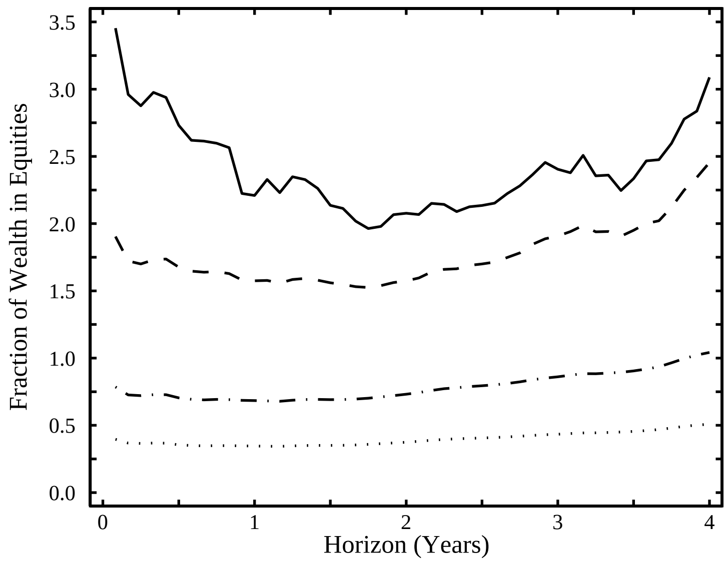

Figure 1. Single-period portfolio choice as a function of the horizon. This figure shows method of moments estimates of the unconditional single-period portfolio choice of an investor with constant relative risk aversion γ = 1 (solid line), 2 (dashed line), 5 (dashed-dotted line), and 10 (dotted line). The horizon and rebalancing period range from one month to four years. Each portfolio choice represents the fraction of savings allocated to equities.

does the one-month portfolio choice exceed the six-month or one-year decisions by as much as in the real data. In other words, the initial decline in the portfolio choice is marginally significant.

Unconditional decisions are optimal if returns are independent and identically distributed or if the trade-off between risk and expected reward is constant. Numerous studies document time-varying risk premia and return volatility. They all but rule out the first scenario. Moreover, the results of Campbell (1987) and Harvey (1989) statistically reject the second scenario. Although there already exists indirect evidence that the portfolio and consumption choice depends on the forecasting variables, I can directly test this hypothesis.

If the unconditional decisions are optimal, they are characterized by both the conditional and the unconditional Euler equations. Additionally, they are overidentified by the moment conditions $E[m_{t+1} \otimes g(z_t)] = 0$, which result from multiplying the conditional Euler equations by functions $g(z)$ of the forecasting variables and then taking unconditional expectations. Hansen (1982) shows how to use overidentifying moment conditions to construct a Wald test of the assumptions used to generate them. The test statistic is:

$$J = \min_{\alpha,c} T\left[\frac{1}{T}\sum_{t=1}^{T} m_{t+1}(\alpha,c,1) \otimes g(z_t)\right]' S^{-1}\left[\frac{1}{T}\sum_{t=1}^{T} m_{t+1}(\alpha,c,1) \otimes g(z_t)\right], \tag{26}$$

where S is the asymptotic covariance matrix of $m_{t+1} \otimes g(z_t)$. Under the null hypothesis that the decisions are independent of z, this statistic is distributed Chi-square with degrees of freedom equal to the number of functions g less the number of decisions.

Table IV presents univariate Wald tests computed with Newey and West (1987) estimates of S. Asymptotic p-values are in parentheses. The test functions in Panel A are $g(z) = [1,z]$. To increase the power of the test against nonlinear alternatives, the test functions in Panels B and C are $g(z) = [1,z,z^2]$ and $g(z) = [1,z,z^2,z^3]$, respectively. The tests decisively reject that the monthly or annual decisions are independent of the forecasting variables. Panel A shows that the portfolio and consumption choice is correlated with the dividend yield, term premium, and lagged excess return. Including squared and cubed forecasting variables in Panels B and C further increases the evidence against unconditional decisions.

C. Conditional Estimates

C.1. Single-Period Decision Rules

Figure 2 shows univariate decision rules of an investor with $\gamma = 1$ (solid line), 2 (dashed line), 5 (dashed-dotted line), and 10 (dotted line).[13] Both the horizon and rebalancing period are one month. Every line represents portfolio choices for the forecasting variable ranging from its 10th to 90th percentiles. In turn, each portfolio choice is computed by applying to monthly returns the conditional estimator, equation (14), with the univariate weighting function, equation (15), and cross-validated bandwidths.

As expected, the decision rules are not constant. All variables forecast changes in investment opportunities. The portfolio choice is approximately linear in the dividend yield and excess return. It is nonlinear in both the default premium and term premium. However, the degree of nonlinearity in these variables diminishes as γ increases.

Table V summarizes the results from Figure 2. It reports the portfolio choice at the 25th, 50th, and 75th percentiles of each forecasting variable. It also gives the average portfolio choice over the sample period. Asymptotic standard errors are in parentheses.[14] The numbers confirm the visual conclusion from Figure 2. The investor chooses substantially different portfolios depending on the realization of each forecasting variable.

Consider again an investor with $\gamma = 5$. At a 4.8 percent dividend yield (75th percentile) the investor holds a portfolio of 83 percent equities, compared to only 44 percent at a 3.1 percent dividend yield (25th percentile). The portfolio choice is equally sensitive to the other variables. The differ-

[13] The pairwise correlations between the forecasting variables are low. As a result, the marginal effect of the variables in multivariate decisions is similar to their effect in univariate decisions. To save space, I only present univariate results here.

[14] Although the asymptotics treat the data as serially uncorrelated, I follow Conley et al. (1997) and report Newey and West (1987) estimates of the asymptotic standard errors.

Table IV

Wald Tests of whether the Portfolio and Consumption Choice Is Independent of the Forecasting Variables

This table shows Wald tests of whether the single-period portfolio and consumption choice of an investor with constant relative risk aversion γ is uncorrelated with a function $g(z)$ of the forecasting variable z. Both the horizon and rebalancing period are one month or one year. The intertemporal discount factor is $\beta = 0.99$. The forecasting variables are the dividend yield (Div), default premium (Def), term premium (Term), and lagged excess return (Ret). The test statistic is:

$$J_T = \min_{\alpha,c} T\left[\frac{1}{T}\sum_{t=1}^{T} m_{t+1}(\alpha,c,1)\otimes g(z_t)\right]'\hat{S}^{-1}\left[\frac{1}{T}\sum_{t=1}^{T} m_{t+1}(\alpha,c,1)\otimes g(z_t)\right],$$

where m_{t+1} is the investor's marginal utility, given a portfolio choice α, a consumption choice c, and a horizon of one rebalancing period. $\hat{S}$ is a serial-correlation corrected estimate of the asymptotic covariance matrix of $m_{t+1} \otimes g(z_t)$. Asymptotic p-values are in parentheses. Note that the consumption choice of an investor with $\gamma = 1$ is deterministic: $c = 1/(1 + \beta)$.

Risk Aversion	1-Month Horizon				1-Year Horizon			
	Div	Def	Term	Ret	Div	Def	Term	Ret
				Panel A: $g(z) = [1,z]$				
$\gamma = 1$	5.98	0.36	2.66	4.67	14.75	0.59	24.53	0.84
	(0.02)	(0.55)	(0.10)	(0.04)	(0.00)	(0.44)	(0.00)	(0.36)
$\gamma = 2$	7.89	3.67	8.83	5.28	11.77	3.30	12.82	11.96
	(0.02)	(0.16)	(0.01)	(0.07)	(0.00)	(0.19)	(0.00)	(0.00)
$\gamma = 5$	7.76	3.51	8.06	5.82	11.43	2.99	9.32	11.10
	(0.02)	(0.17)	(0.02)	(0.05)	(0.00)	(0.22)	(0.01)	(0.00)
$\gamma = 10$	7.74	3.48	7.90	6.09	11.38	2.93	9.06	10.88
	(0.02)	(0.18)	(0.02)	(0.05)	(0.00)	(0.23)	(0.01)	(0.00)
				Panel B: $g(z) = [1,z,z^2]$				
$\gamma = 1$	7.17	2.81	12.42	4.76	17.41	2.35	25.09	2.11
	(0.03)	(0.25)	(0.00)	(0.09)	(0.00)	(0.31)	(0.00)	(0.35)
$\gamma = 2$	19.41	15.01	20.45	5.47	12.63	4.65	11.17	18.11
	(0.00)	(0.00)	(0.00)	(0.24)	(0.01)	(0.32)	(0.02)	(0.00)
$\gamma = 5$	21.25	15.20	18.52	6.37	13.54	3.72	8.26	17.16
	(0.00)	(0.00)	(0.00)	(0.17)	(0.01)	(0.44)	(0.08)	(0.00)
$\gamma = 10$	21.77	15.29	17.91	6.97	13.78	3.55	8.14	16.87
	(0.00)	(0.00)	(0.00)	(0.13)	(0.01)	(0.47)	(0.09)	(0.00)
				Panel C: $g(z) = [1,z,z^2,z^3]$				
$\gamma = 1$	7.28	4.56	18.09	4.89	19.47	3.94	25.80	4.25
	(0.06)	(0.21)	(0.00)	(0.18)	(0.00)	(0.27)	(0.00)	(0.24)
$\gamma = 2$	27.56	18.70	21.88	5.15	13.16	13.61	9.94	23.78
	(0.00)	(0.00)	(0.00)	(0.52)	(0.04)	(0.03)	(0.12)	(0.00)
$\gamma = 5$	28.45	18.94	20.05	6.26	15.06	11.30	9.37	22.90
	(0.00)	(0.00)	(0.00)	(0.39)	(0.02)	(0.08)	(0.15)	(0.00)
$\gamma = 10$	28.72	19.06	19.51	6.90	15.52	10.78	9.38	22.64
	(0.00)	(0.00)	(0.00)	(0.33)	(0.02)	(0.10)	(0.15)	(0.00)

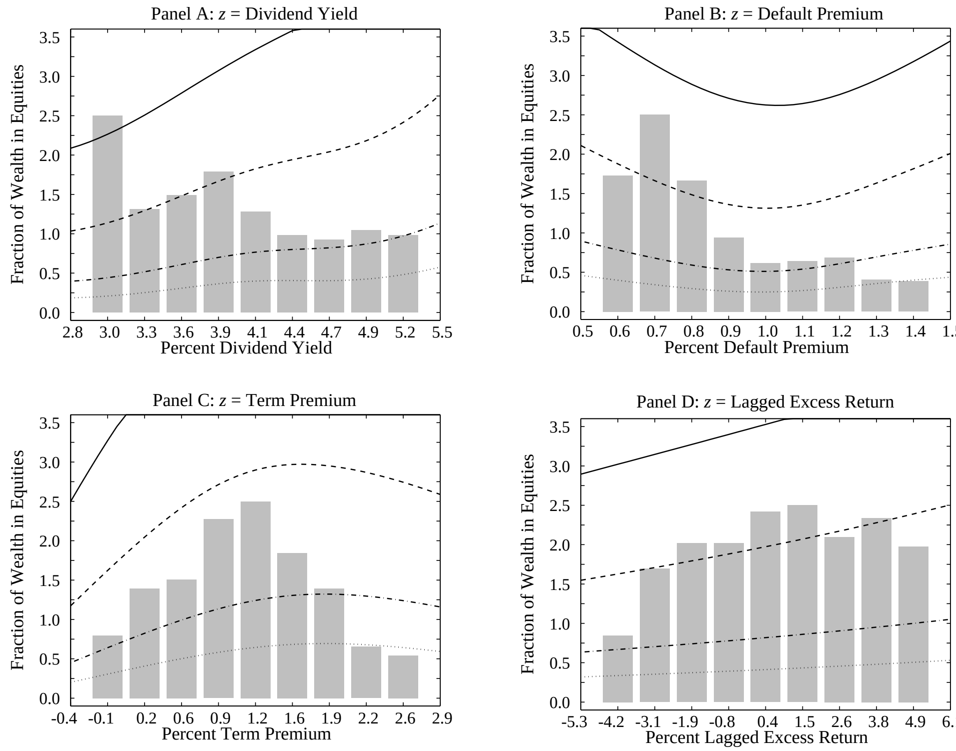

Figure 2. Single-period portfolio choice as a function of the forecasting variables. This figure shows conditional method of moments estimates of the conditional single-period portfolio choice of an investor with constant relative risk aversion $\gamma = 1$ (solid line), 2 (dashed line), 5 (dashed-dotted line), and 10 (dotted line). The horizon and rebalancing period are one month. Each portfolio choice represents the fraction of savings allocated to equities as a function of the dividend yield (Panel A), default premium (Panel B), term premium (Panel C), or lagged excess return (Panel D). Histograms of the forecasting variables are in the background.

ences between the equity holdings at the 75th and 25th percentiles of the default premium, term premium, and excess return are 20, 31, and 22 percent of savings, respectively.

Given there is only weak statistical evidence of predictability at a monthly frequency, the portfolio choice seems overly sensitive to the forecasting variables. This finding is explained by Kandel and Stambaugh (1996), who show that even after considering estimation risk, a CRRA investor acts aggressively to weak predictive evidence. Since my estimator abstracts from estimation risk, the results in Figure 2 and Table V are not surprising.

In fact, the estimates are reasonable. The figures of Barberis (1999) suggest that, ignoring parameter uncertainty, a one-year investor with $\gamma = 5$ holds between 10 and 100 percent equities as the dividend yield ranges from 3.0 to 4.6 percent. The results of Campbell and Viceira (1996) imply that an

Table V

Single-Period Portfolio Choice as a Function of the Forecasting Variables

This table shows conditional method of moments estimates of the conditional single-period portfolio choice α of an investor with constant relative risk aversion γ. Both the horizon and rebalancing period are one month. The estimates represent the fraction of savings allocated to equities when the forecasting variable z is equal to its 25th, 50th, and 75th percentiles. The forecasting variables are the dividend yield (Panel A), default premium (Panel B), term premium (Panel C), and lagged excess return (Panel D). Each panel also shows the average conditional portfolio choice over the sample period. The conditional method of moments estimator $\hat{\alpha}_T$ of α is:

$$\hat{\alpha}_T(z,1) = \left\{\alpha : \frac{1}{Th_T}\sum_{t=1}^{T}\omega\left(\frac{z_t - z}{h_T}\right)m_{t+1}(\alpha, 0.5, 1) = 0\right\},$$

where m_{t+1} is the investor's marginal utility, given a portfolio choice α, a consumption choice $c = 0.5$, and a horizon of one rebalancing period. ω is a standard normal density and the bandwidth h_T is chosen by leave-one-out cross-validation. Serial-correlation corrected asymptotic standard errors are in parentheses. Bootstrapped p-values of the difference between two portfolio choice estimates are below the underbraces.

Risk Aversion	Percentile of z: 25%	50%	75%	Average	Percentile of z: 25%	50%	75%	Average
	Panel A: z = Dividend Yield				Panel B: z = Default Premium			
$\gamma = 1$	2.23	2.97	3.85	3.08	3.33	2.97	2.66	3.26
	(1.16)	(0.93)	(0.86)		(0.93)	(0.96)	(1.04)	
$\gamma = 2$	1.12	1.61	2.05	1.77	1.81	1.55	1.36	1.83
	(0.66)	(0.57)	(0.65)		(0.59)	(0.57)	(0.62)	
$\gamma = 5$	0.44	0.67	0.83	0.74	0.75	0.62	0.55	0.76
	(0.28)	(0.26)	(0.28)		(0.25)	(0.24)	(0.27)	
	0.04	0.19			0.05	0.35		
		0.09				0.25		
$\gamma = 10$	0.21	0.35	0.41	0.37	0.38	0.31	0.28	0.38
	(0.15)	(0.13)	(0.15)		(0.13)	(0.12)	(0.14)	
	Panel C: z = Term Premium				Panel D: z = Lagged Excess Return			
$\gamma = 1$	4.13	4.31	4.24	3.71	3.24	3.60	3.90	3.54
	(1.04)	(0.96)	(0.98)		(0.92)	(0.89)	(0.85)	
$\gamma = 2$	2.25	2.83	2.96	2.32	1.77	2.03	2.28	2.04
	(0.55)	(0.48)	(0.49)		(0.53)	(0.52)	(0.51)	
$\gamma = 5$	0.91	1.20	1.32	0.99	0.73	0.84	0.95	0.85
	(0.24)	(0.23)	(0.24)		(0.23)	(0.22)	(0.23)	
	0.01	0.22			0.08	0.09		
		0.02				0.08		
$\gamma = 10$	0.46	0.62	0.69	0.51	0.37	0.43	0.48	0.43
	(0.12)	(0.12)	(0.13)		(0.11)	(0.11)	(0.12)	

infinitely lived investor with $\gamma = 4$ holds a portfolio of more than 100 percent equities at a 4.8 percent dividend yield and of almost no equities at a 3.1 percent dividend yield.

The standard errors in Table V are larger than those in Table III. This is not surprising because the conditional method of moments uses only data from state z and states similar to z to estimate the decisions in state z. As a result, each estimate in Table V is based on fewer data than the estimates in Table III. This explains the increase in standard errors.

To further explore the accuracy of the estimates, I repeat the analysis for an investor with $\gamma = 5$ for 5000 resampled data sets in which by construction returns are independent of the forecasting variables. Each data set consists of 600 monthly returns and realizations of the forecasting variables, sampled randomly and independently with replacement from the original data. The results are summarized beneath the underbraces in Table V. These statistics are the fractions of "resampled data" estimates for which the slope of the decision rule between two percentiles exceeds the slope of the "real data" decision rule. They are bootstrapped p-values of the slope between two points of the decision rule.

With 90 percent confidence, the slope of all decision rules is nonzero between the 25th and 50th percentiles. Between the 25th and 75th percentiles the slope is positive for the dividend yield, term premium, and excess return. Between the 50th and 75th percentiles it is positive only for the excess return. With 95 percent confidence, the slope between the 25th and 50th percentiles is nonzero for the dividend yield, default premium, and term premium. Between the 25th and 75th percentiles it is positive only for the term premium.

Campbell and Viceira's (1996) standard errors are of similar magnitude. Furthermore, the intercepts of their linear decision rules are less than one standard deviation from zero but the slope coefficients are statistically significant. As with the unconditional portfolio choice, we cannot estimate the fraction of savings to be allocated to equities on average. However, we can estimate how the portfolio choice depends on the forecasting variables.

To examine how sensitive the estimates are to the choice of bandwidths, Panels A through D of Figure 3 show three sets of decision rules for an investor with $\gamma = 5$. The first set (dotted lines) is computed with the cross-validated bandwidths. They are the same as the dashed-dotted lines in Figure 2. The other two sets are computed with bandwidths 25 percent smaller (minus symbols) or 25 percent larger (plus symbols) than the cross-validated bandwidths.

The panels illustrate that the above conclusions are insensitive to reasonable variations in bandwidths. Increasing or decreasing the bandwidths by 25 percent does not change the fact that the portfolio choice depends on the forecasting variable.[15] The decisions are always less nonlinear in the dividend yield and excess return than they are in the default premium and term premium. Of course, the degree of nonlinearity depends on the bandwidths.

[15] Also, the bootstrapped p-values in Table III are insensitive to a 25 percent increase or decrease in bandwidths.

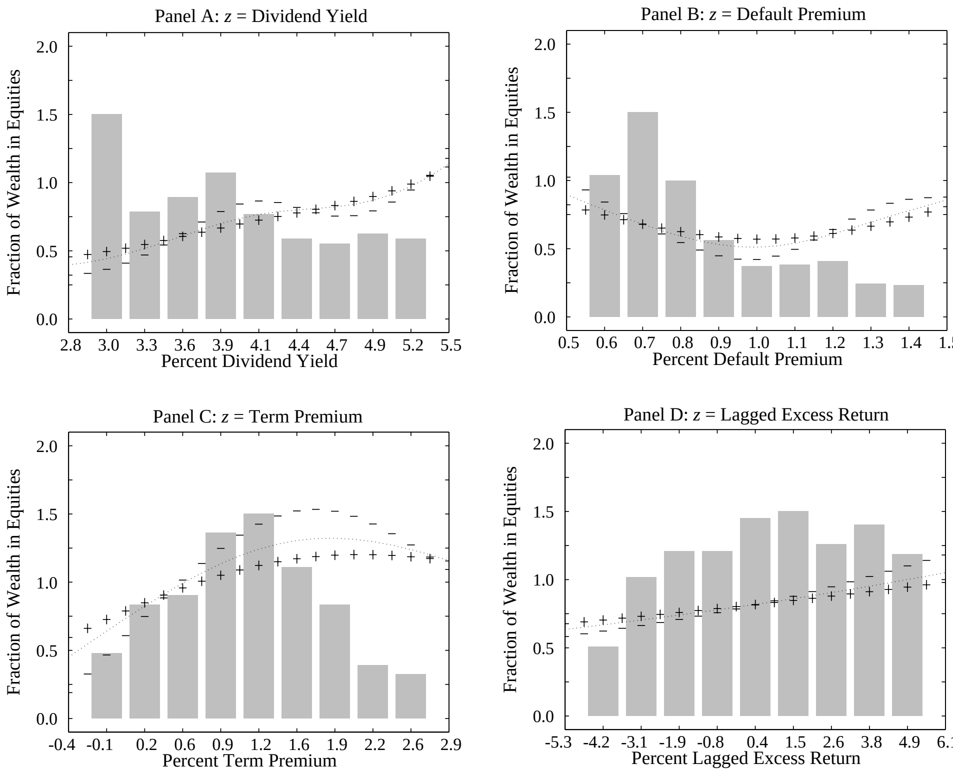

Figure 3. Sensitivity of the portfolio choice estimates to the bandwidths. This figure shows three sets of conditional method of moments estimates of the conditional single-period portfolio choice of an investor with constant relative risk aversion $\gamma = 5$. The horizon and rebalancing period are one month. The first set of estimates (dotted line) is computed with cross-validated bandwidths and hence corresponds to the dashed lines in Figure 2. The other two sets of estimates are computed with bandwidths 25 percent smaller (minus symbols) or 25 percent larger (plus symbols) than the cross-validated bandwidths. Each portfolio choice represents the fraction of savings allocated to equities as a function of the dividend yield (Panel A), default premium (Panel B), term premium (Panel C), or lagged excess return (Panel D). Histograms of the forecasting variables are in the background.

Figure 4 shows the decision rules as a function of the horizon for $\gamma = 5$. The horizon and rebalancing period range from three months to four years. Each n-month portfolio choice is computed by applying to overlapping n-month returns the conditional estimator, equation (14), with univariate weighting function, equation (15), and the same bandwidths as in Figure 2 and Table V.[16]

[16] The MSE optimal bandwidths change with the horizon. Unfortunately, with dependent data leave-one-out cross-validation leads to bandwidths that are too small relative to the optimal. Although Hart and Vieu (1990) claim that "leave-some-out" cross-validation eliminates this downward bias, bandwidth selection with overlapping data is controversial. As a result, I use the same bandwidths throughout my analysis. Of course, I make sure that my conclusions are insensitive to reasonable variations in these bandwidths.

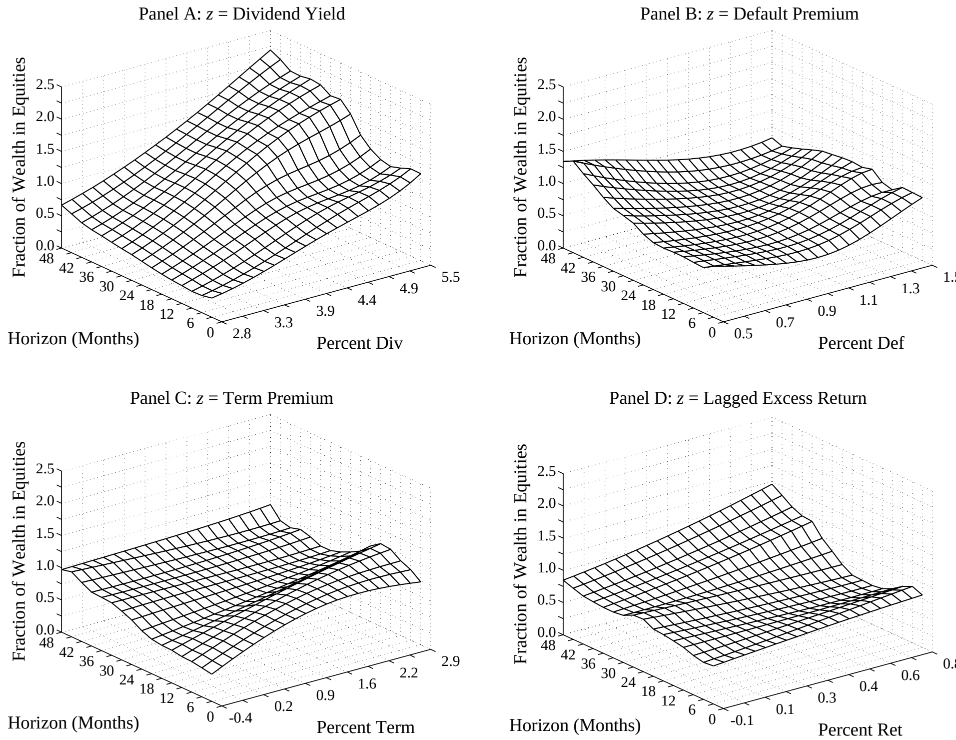

Figure 4. Single-period portfolio choice as a function of the horizon and the forecasting variables. This figure shows conditional method of moments estimates of the conditional single-period portfolio choice of an investor with constant relative risk aversion $\gamma = 5$. The horizon and rebalancing period range from two months to four years. Each portfolio choice represents the fraction of savings allocated to equities as a function of the horizon and the dividend yield (Panel A), default premium (Panel B), term premium (Panel C), or lagged excess return (Panel D).

In Panel A the portfolio choice increases with the dividend yield for all horizons. More interestingly, both the level and the slope of the decision rule increase with the horizon. At a 3.1 percent dividend yield, an investor with $\gamma = 5$ and a three-month, two-year, or four-year horizon holds 36, 50, or 76 percent of savings in equities. At a 4.8 percent dividend yield, the same investor holds 100, 153, or 172 percent of savings in equities.

This horizon pattern matches that of Barberis (1999). It is also consistent with the results of Fama and French (1988a). They document that the slope coefficient and the R^2 of a regression of returns on dividend yields increase significantly with the return horizon. Interestingly, the increase in portfolio choice for a given dividend yield is nonlinear, even nonmonotonic for some dividend yields, in the horizon. This finding differs from the linear relation reported by Barberis and could be evidence of model misspecification in that study.

Overall, the average portfolio choice in Panels B through D increases with the horizon. However, the results are less systematic than in Panel A. In particular, the slopes of the decision rules between the 25th and 75th percentiles are of opposite signs at different horizons. This illustrates the advantage of the conditional method of moments. It would be hard to formulate a model of investment opportunities that generates such irregular horizon patterns.

The focus has been on portfolio choice. This is because the single-period consumption choice is surprisingly insensitive to the forecasting variables. The conditional estimates are within a few percent of the unconditional estimates in Table III. For instance, with a one-month horizon the consumption choice for $\gamma = 5$ increases from 50.1 to 50.4 percent of wealth as the dividend yields ranges between 3.1 and 4.8 percent. With a one-year horizon the consumption choice increases from 51.0 to 51.9 percent of wealth. Furthermore, the estimated shapes of the consumption rules mirror those of the portfolio rules.

C.2. Multiperiod Decision Rules

Merton (1969) illustrates that if returns are independent and identically distributed, investors with the same CRRA preferences and rebalancing period hold identical portfolios, whether they live for three months or 20 years. However, if investment opportunities are time-varying, the multiperiod portfolio choice can diverge from the single-period or myopic portfolio choice due to hedging demands. Hedging demands arise when the investor tries to smooth the effects of predictable changes in future investment opportunities.

A natural place to start analyzing hedging demands is the two-period problem. Table VI reports two-period hedging demands of an investor with $\gamma = 5$, a six-month to four-year horizon, and a corresponding three-month to two-year rebalancing period. It shows the difference between the estimated two-period portfolio choice $\hat{\alpha}_T(z,2)$ and the estimated myopic portfolio choice $\hat{\alpha}_T(z,1)$ at the 25th, 50th, and 75th percentiles of each forecasting variable. The table also gives the average hedging demand over the sample period.

To assess the statistical significance of the estimates, I repeat the analysis using 5000 resampled data sets in which investment opportunities are by construction time-varying but serially uncorrelated. Each data set consists of 600 n-period returns and their corresponding realizations of the forecasting variable, sampled as pairs randomly with replacement from the original data. The fraction of "resampled data" hedging demands that exceed the "real data" hedging demand is reported in parentheses beneath each estimate.

Panel A shows a two-period investor holds a greater fraction of savings in equities than does an otherwise identical myopic investor. Depending on the horizon and the dividend yield, the hedging demand in most cases exceeds one percent and in some cases five percent of savings. Furthermore, the bootstrapped p-values indicate that the estimates are statistically significant.

Table VI

Two-Period Hedging Demand as a Function of the Horizon and the Forecasting Variables

This table shows conditional method of moments estimates of the conditional two-period hedging demand of an investor with constant relative risk aversion $\gamma = 5$. The horizon ranges from one to ten years and the rebalancing period is half the horizon. The intertemporal discount factor is $\beta = 0.99$. The estimates represent the difference between the two-period portfolio choice $\alpha(z,2)$ and the myopic portfolio choice $\alpha(z,1)$, expressed as a percentage of savings, when the forecasting variable z is equal to its 25th, 50th, and 75th percentiles. The forecasting variables are the dividend yield (Panel A), default premium (Panel B), term premium (Panel C), and lagged excess return (Panel D). Each panel also shows the average hedging demand over the sample period. The conditional method of moments estimator $\{\hat{\alpha}_T(z,\tau), \hat{c}_T(z,\tau)\}$ of the τ-period portfolio choice $\alpha(z,\tau)$ and consumption choice $c(z,\tau)$ is:

$$\begin{bmatrix} \hat{\alpha}_T(z,\tau) \\ \hat{c}_T(z,\tau) \end{bmatrix} = \left\{ \begin{bmatrix} \alpha \\ c \end{bmatrix} : \frac{1}{Th_T} \sum_{t=1}^{T} \omega\left(\frac{z_t - z}{h_T}\right) m_{t+1}(\alpha, c, \tau) = 0 \right\},$$

where m_{t+1} is the investor's marginal utility, given a portfolio choice α, a consumption choice c, and a horizon of $\tau = 1$ or $\tau = 2$ rebalancing periods. ω is a standard normal density and the bandwidth h_T is chosen by leave-one-out cross-validation. Bootstrapped p-values are in parentheses.

	Percentile of z				Percentile of z			
Horizon	25%	50%	75%	Average	25%	50%	75%	Average
	Panel A: z = Dividend Yield				Panel B: z = Default Premium			
6-month	0.69	0.83	1.62	1.09	−0.07	−0.02	0.06	−0.01
	(0.05)	(0.04)	(0.01)		(0.59)	(0.53)	(0.38)	
1-year	1.54	1.70	2.89	2.12	0.08	0.12	0.21	0.31
	(0.04)	(0.03)	(0.01)		(0.30)	(0.25)	(0.17)	
2-year	2.76	2.75	5.58	3.55	−0.08	−0.04	0.30	0.39
	(0.02)	(0.02)	(0.01)		(0.64)	(0.54)	(0.09)	
4-year	5.68	4.37	4.48	4.74	0.03	0.06	0.39	0.99
	(0.02)	(0.01)	(0.01)		(0.41)	(0.35)	(0.07)	
	Panel C: z = Term Premium				Panel D: z = Lagged Excess Return			
6-month	−0.62	−0.44	−0.25	−0.43	−0.64	−0.56	−0.43	−0.52
	(0.74)	(0.61)	(0.21)		(0.86)	(0.82)	(0.92)	
1-year	−0.86	−0.61	−0.31	−0.56	−0.78	−0.79	−0.79	−0.78
	(0.87)	(0.76)	(0.18)		(0.97)	(0.98)	(0.98)	
2-year	−1.24	−1.16	−0.95	−0.97	1.20	1.19	1.16	1.18
	(0.92)	(0.89)	(0.86)		(0.03)	(0.02)	(0.02)	
4-year	−0.75	−0.98	−1.23	−0.96	2.42	2.44	2.48	2.44
	(0.79)	(0.85)	(0.89)		(0.02)	(0.01)	(0.01)	

Of the other forecasting variables, only the excess return generates significant hedging demands. Six-month and one-year hedging demands are negative. Two- and four-year hedging demands are positive. This pattern corresponds to upward sloping three- and six-month and downward sloping one- and two-year decision rules in Figure 4.

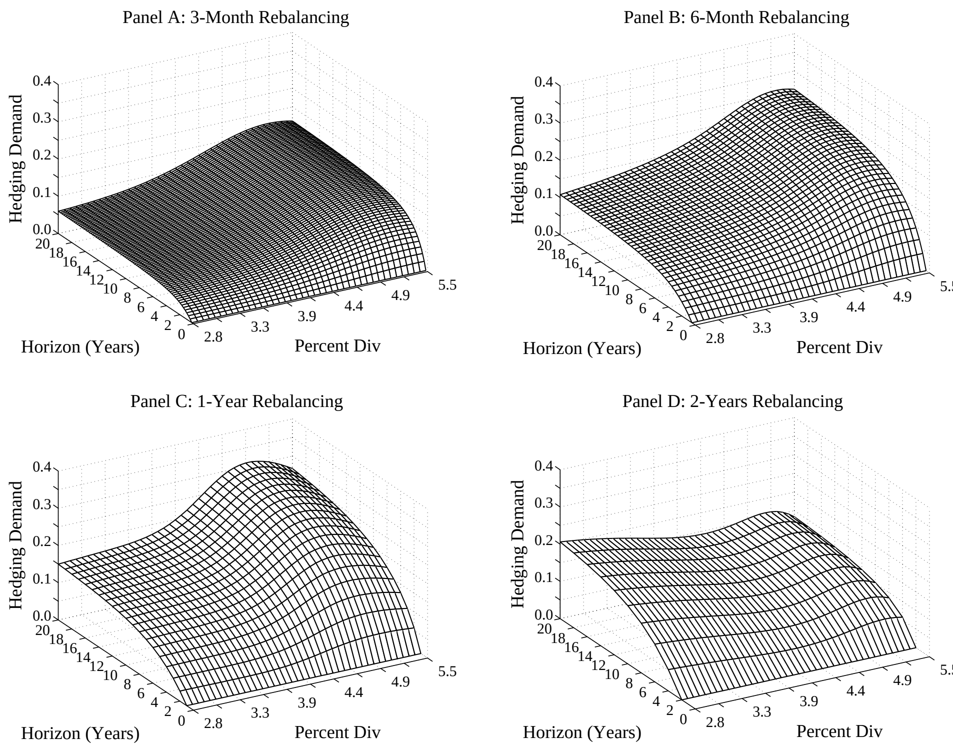

Figure 5. Hedging demand as a function of the horizon, rebalancing period, and the dividend yield. This figure shows conditional method of moments estimates of the conditional multiperiod hedging demand of an investor with constant relative risk aversion $\gamma = 5$. The horizon ranges from three months to 20 years and the rebalancing period is three months (Panel A), six months (Panel B), one year (Panel C), or two years (Panel D). Each hedging demand represents the difference between the multiperiod and the myopic portfolio choice as a function of the horizon and the dividend yield (Div).

Figure 5 takes a closer look at the multiperiod portfolio choice for the dividend yield. It shows the hedging demand as a function of the horizon for $\gamma = 5$. The horizon ranges from three months to 20 years, and the rebalancing period is three months (Panel A), six months (Panel B), one year (Panel C), and two years (Panel D). Each hedging demand is the difference between the τ-period portfolio choice $\hat{\alpha}_T(z,\tau)$ and the myopic portfolio choice $\hat{\alpha}_T(z,1)$, where τ is the number of decisions the investor makes until the end of the horizon. Table VII reports the hedging demand for the 25th, 50th, and 75th percentiles of the dividend yield. It also shows the average hedging demand over the sample period.

Figure 5 and Table VII illustrate that for a given dividend yield the hedging demand increases monotonically with the horizon. This result does not depend on the rebalancing frequency. Consider the case of annual rebalancing in Panel C in Figure 5 or Table VII. For a 3.1 percent dividend yield, a

Table VII

Hedging Demand as a Function of the Horizon, Rebalancing Period, and the Dividend Yield

This table shows conditional method of moments estimates of the conditional multiperiod hedging demand of an investor constant relative risk aversion $\gamma = 5$. The horizon ranges from two to ten years and the rebalancing period is three months (Panel A), six months (Panel B), one year (Panel C), and two years (Panel D). The intertemporal discount factor is $\beta = 0.99$. The estimates represent the difference between the multiperiod portfolio choice $\alpha(z,\tau)$ and the myopic portfolio choice $\alpha(z,1)$ when the dividend yield z is equal to its 25th, 50th, and 75th percentiles. Each panel also shows the average hedging demand over the sample period. The conditional method of moments estimator $\{\hat{\alpha}_T(z,\tau), \hat{c}_T(z,\tau)\}$ of the τ-period portfolio choice $\alpha(z,\tau)$ and the τ-period consumption choice $c(z,\tau)$ is

$$\begin{bmatrix} \hat{\alpha}_T(z,\tau) \\ \hat{c}_T(z,\tau) \end{bmatrix} = \left\{ \begin{bmatrix} \alpha \\ c \end{bmatrix} : \frac{1}{Th_T} \sum_{t=1}^{T} \omega\left(\frac{z_t - z}{h_T}\right) m_{t+1}(\alpha, c, \tau) = 0 \right\},$$

where m_{t+1} is the investor's marginal utility, given a portfolio choice α, a consumption choice c, and a horizon of $\tau = 1$ (myopic) or $\tau =$ (horizon)/(rebalancing period) rebalancing periods. ω is a standard normal density and the bandwidth h_T is chosen by leave-one-out cross-validation. Note that a two-year portfolio choice with a two-year rebalancing period is by definition a myopic decision.

	Dividend Yield Percentile				Dividend Yield Percentile			
Horizon	25%	50%	75%	Average	25%	50%	75%	Average
	Panel A: 3-Month Rebalancing				Panel B: 6-Month Rebalancing			
2-Year	3.63	4.84	9.19	5.78	4.26	5.05	8.61	6.00
4-Year	5.35	7.22	13.38	8.37	7.77	9.54	15.55	10.75
10-Year	6.67	9.05	16.62	10.37	11.52	14.37	23.01	15.85
20-Year	7.10	9.65	17.67	11.03	12.85	16.11	25.75	17.72
	Panel C: 1-Year Rebalancing				Panel D: 2-Year Rebalancing			
2-Year	2.76	2.75	5.58	3.55	0.00	0.00	0.00	0.00
4-Year	7.51	7.93	15.53	9.61	5.68	4.37	4.48	4.74
10-Year	14.37	15.65	28.51	17.96	15.64	12.78	11.72	12.76
20-Year	17.47	19.24	34.75	21.93	20.95	17.39	15.34	16.92

four-year, ten-year, or 20-year investor holds an additional 8, 14, or 17 percent of saving in equities, relative to the one-year investor who already holds 37 percent of savings in equities. For a 4.8 percent dividend yield, the hedging demands are 16, 29, and 35 percent of savings in addition to the one-year investor's 100 percent of savings in equities.

The magnitude of these hedging demands matches Campbell and Viceira's (1996) estimates. They report that on average the percentage hedging demand, the fraction of total demand for equity due to hedging motives, for an infinitely lived investor with $\gamma = 4$ and annual rebalancing is 35 percent. The above numbers for a 20-year horizon imply percentage hedging demands of 31 and 26 percent for 3.1 and 4.8 percent dividend yields, respectively.

As τ increases to ∞, my estimates converge to estimates of the infinite-horizon portfolio choice. It is surprising how quickly this convergence occurs. The graphs suggest that a 15-year, and certainly a 20-year investor acts like an infinitely lived investor.

Another interesting finding is that for a fixed horizon, the shape of the hedging demand as a function of the dividend yield depends on the rebalancing frequency. Overall, in Panels A through C of Figure 5 the hedging demand is greater for higher dividend yields. This implies that a 10- or 20-year investor times the market more aggressively than does a single-period investor. In Panel D the hedging demand is higher for lower dividend yields. This has the opposite implication and is more consistent with Siegel's (1994) investment advice to hold stocks for the long run.

In absolute terms, the multiperiod consumption choice is as insensitive to the dividend yield as is the single-period consumption choice. Recall that single-period consumption for $\gamma = 5$ increases from 51.0 to 51.9 percent of wealth as the dividend yield ranges between its 25th and 75th percentiles. With annual rebalancing, the 4-, 10-, and 20-year consumption choice increases almost linearly from 21.9 to 23.3, from 11.6 to 12.8, and from 7.8 to 9.1 percent of wealth, respectively. However, since the level of consumption decreases with the horizon due to consumption smoothing, the relative sensitivity of the consumption choice to the dividend yield increases dramatically with the horizon.

All empirical results are for $\beta = 0.99$. Decreasing the discount factor, thereby making the investor less patient, leads to even less responsive consumption rules. As a result, the hedging demands in Table VII and Figure 5 are lowered. For instance, with $\beta = 0.99$ and annual rebalancing, the average hedging demands for a 4-, 10- and 20-year horizon are 9.6, 18.0, and 21.9 percent of savings, respectively. If $\beta = 0.95$, the average hedging demands are 8.2, 14.2, and 16.4 percent, respectively. However, the shapes of the decision rules are unaffected by the discount factor.

IV. Conclusion

This paper addresses an important question economic theory leaves unanswered: How does portfolio and consumption choice depend on variables that forecast time-varying investment opportunities? My research is distinct from the existing empirical literature on portfolio and consumption choice because it does not assume a model of how returns relate to forecasting variables or of how investment opportunities change through time. Instead, I develop a nonparametric approach that generates consistent estimates shaped by the data.

I estimate single-period and multiperiod portfolio and consumption rules of an investor with CRRA utility who allocates wealth to the NYSE index or a 30-day Treasury bill. I find that the portfolio choice varies significantly with the dividend yield, default premium, term premium, and lagged excess return. Furthermore, the optimal decisions depend on the investor's horizon and rebalancing frequency. The fact that the optimal multiperiod portfolio choice differs from the myopic portfolio choice suggests that future research on empirical anomalies in financial data must consider the multiperiod aspect of the investors portfolio and consumption choice.

Several aspects of portfolio and consumption choice remain unexplored. For example, Canner, Mankiw, and Weil (1997) document and discuss anomalous investment advice of professionals. The following two extensions of my empirical analysis will help clarify less stylized and more realistic portfolio and consumption choice problems. First, I will allow the investor to allocate wealth to more assets, such as equity portfolios that reflect mutual fund styles, short-term bonds, long-term bonds, and foreign equities. Second, I will study decision rules of alternative preference specifications.

Transaction costs further complicate the multiperiod portfolio choice. Rebalancing a portfolio every period is typically not optimal. Instead, the investor holds a portfolio until the expected gains from rebalancing outweigh the transaction costs of trading. It is possible to modify the conditional Euler equations to reflect both fixed and proportional transaction costs. The conditional method of moments can then be used to estimate not only portfolio and consumption rules, but also optimal rebalancing rules.

My estimator is a conditional version of Morvai's (1991, 1992) investment strategies. This author proves that for log utility and in the case of constant investment opportunities, the strategy of equation (12) yields an asymptotically optimal portfolio choice. It achieves the optimal growth rate of wealth in an infinite sequence of portfolio choices. Studying similar decision theoretic properties of the conditional method of moments is an ambitious but interesting topic of future research.

Appendix

This appendix shows how to derive the asymptotics of the conditional method of moments. It combines the properties of Huber's (1964) M-estimators with those of nonparametric time series regressions. Robinson (1983) first proves the consistency and derives the asymptotic distribution of such nonparametric M-estimators for time series. Gourieroux, Monfort, and Tenreiro (1995) establish the same results under less restrictive assumptions. It is straightforward to adapt their arguments to the conditional method of moments.

Below are the assumptions and main steps sufficient to prove consistency and derive the asymptotic distribution of the conditional method of moments. For further details, see Gourieroux et al. (1995) and the references cited therein.

A. Assumptions on the Data

Assumption 1: *Returns and forecasting variables $\{R^e_{t+1}, z_t\}_{t=1}^T$ are realizations from a strictly stationary $\mathbb{R}^N \times \mathbb{R}^K$-valued process $\{X_t\}$.*

ASSUMPTION 2: *$\{X_t\}$ is geometric-mixing. This means that for every A in $\mathcal{F}_1^k$ and every B in $\mathcal{F}_{k+n}^{\infty}$ the following inequality is satisfied:*

$$|P(A \cap B) - P(A)P(B)| \leq \phi\rho^n \quad \textit{for all positive integers } n, \tag{A1}$$

where $\phi \geq 0$, $\rho \in (0,1)$, and $\mathcal{F}_m^n$ is the σ-field generated by $\{X_t\}_{t=n}^m$.

ASSUMPTION 3: *$\{z_t\}$ has a continuous distribution that is described by a continuous and strictly positive density function f_z on a compact subset Z of $\mathbb{R}^K$.*

ASSUMPTION 4: *For all $t \geq 0$, $\{z_0, z_t\}$ has a continuous transition density function f_{z_t,z_0} that is uniformly bounded in a neighborhood of $\{z,z\}$.*

B. Assumptions on Marginal Utility

ASSUMPTION 5: *For every return and horizon, m_{t+1} is a continuous and twice differentiable function of the portfolio and consumption choice $\{\alpha,c\}$ on a compact choice set $A \times C$ of $\mathbb{R}^{N+1}$.*

ASSUMPTION 6: *For every feasible portfolio and consumption choice and horizon, m_{t+1} is a measurable function of the return R_{t+1}^e.*

ASSUMPTION 7: *m_{t+1} is Lipschitz continuous. This means that for every return, horizon, any two portfolio and consumption choices $\{\alpha,c\}$ and $\{\alpha',c'\}$ in $A \times C$, and for some positive constant a, independent of these choices, the following inequality is satisfied:*

$$|m_{t+1}(\alpha,c,\tau) - m_{t+1}(\alpha',c',\tau)| \leq a\,\|\{\alpha,c\} - \{\alpha',c'\}\|. \tag{A2}$$

ASSUMPTION 8: *There exist constants $b \in \,]0,1[$ and $d > 0$, such that the following inequality is satisfied:*

$$E\left[\sup_{\alpha \in A}\sup_{c \in C}|m_{t+1}(\alpha,c,\tau)|^{2/b+d}\right] < \infty. \tag{A3}$$

C. Assumptions on Expected Marginal Utility

ASSUMPTION 9: *$M_\infty(z,\alpha,c,\tau) = E\,[m_{t+1}(\alpha,c,\tau)|z]$ is uniformly equicontinuous. This means that for all positive ϵ there exists a constant d such that the following inequality is satisfied:*

$$\sup_{z \in Z}\ \sup_{z' : \|z'-z\| < d}\ \sup_{\alpha \in A}\ \sup_{c \in C}|M_\infty(z,\alpha,c,\tau) - M_\infty(z',\alpha,c,\tau)| < \epsilon. \tag{A4}$$

ASSUMPTION 10: *For every realization of the forecasting variables and horizon, there exists a portfolio and consumption choice $\{\alpha(z,\tau), c(z,\tau)\} \in A \times C$ that is a unique zero of $M_\infty(z,\alpha,c,\tau)$.*

ASSUMPTION 11: *For* $\{\alpha, c\} = \{\alpha(z,\tau), c(z,\tau)\}$ *and some constant* $d > 0$:

$$E[\|m_{t+1}\|^{2+d}] < \infty; \tag{A5}$$

$$E[m_{t+1}\, m'_{t+1}|z] f_z(z) \text{ is continuous in } z; \tag{A6}$$

$$E[\|m_{t+1}\|^{2+d}|z] f_z(z) \text{ is bounded in a neighborhood of } z. \tag{A7}$$

ASSUMPTION 12: *For* $\{\alpha, c\} = \{\alpha(z,\tau), c(z,\tau)\}$ *and some constant* $d > 0$:

$$E\left[\left\| \frac{\partial m_{t+1}}{\partial[\alpha', c]} \right\|^{2+d}\right] < \infty; \tag{A8}$$

$$E\left[\left. \frac{\partial m_{t+1}}{\partial[\alpha', c]} \right| z\right] f_z(z) \text{ is continuous in } z; \tag{A9}$$

$$E\left[\left. \left\| \frac{\partial m_{t+1}}{\partial[\alpha', c]} \right\|^{2+d} \right| z\right] f_z(z) \text{ is bounded in a neighborhood of } z. \tag{A10}$$

D. Assumptions on the Weighting Function

ASSUMPTION 13: *The weighting function* $\omega(u)$ *is the product of* K *univariate, continuous, and bounded functions* $k(u_i)$ *with*

$$\int_{\mathbb{R}} k(u)\, du = 1, \qquad \int_{\mathbb{R}} uk(u)\, du = 0, \quad \text{and} \quad \int_{\mathbb{R}} u^2 k(u)\, du < \infty. \tag{A11}$$

ASSUMPTION 14: *The weighting function is Lipschitz continuous. This means that for any two points* u *and* u' *in* $\mathbb{R}^K$ *and for some positive constant* a, *independent of these points, the following inequality is satisfied*:

$$|\omega(u) - \omega(u')| \le a\|u - u'\|. \tag{A12}$$

E. Assumptions on the Bandwidths

ASSUMPTION 15: *The bandwidths* h_T *satisfy*:

$$h_T \to 0, \qquad Th_T^{K+4} \to 0, \quad \text{and} \quad Th_T^K \to \infty \quad \text{as } T \to \infty. \tag{A13}$$

ASSUMPTION 16: *For the constant* b *from Assumption 12,*

$$\frac{T^{(1-b)/2} h_T^K}{\log(T)} \to \infty \quad \text{as } T \to \infty. \tag{A14}$$

The bandwidths Assumption 13 and the weighting function Assumption 15 are more restrictive than necessary to derive an asymptotic distribution of the conditional method of moments. They assure that the asymptotic bias of the estimator vanishes as $T \to \infty$.

The marginal utility assumptions are general enough to be satisfied by CRRA utility and other popular preferences. Assumptions 8 and 16 are particularly important. They distinguish Gourieroux et al.'s (1995) arguments from similar proofs in the literature. With these assumptions we can establish uniform consistency of the conditional method of moments for non-bounded marginal utility.

F. Consistency

To establish consistency, it is sufficient to prove uniform almost sure convergence of:

$$M_T(z,\alpha,c,\tau) = \frac{1}{Th_T^K}\sum_{t=1}^{T}\omega\left(\frac{z-z_t}{h_T}\right)m_{t+1}(\alpha,c,\tau) \tag{A15}$$

to $M_\infty(z,\alpha,c,\tau)$. Assumption 9, that M_∞ is continuous, Assumption 10, that $\{\alpha(z),c(z)\}$ exists, and the compactness of $\{Z,A,C\}$ guarantee the result.

Following the arguments in Gourieroux et al. (1995), consider the following decomposition of M_T:

$$M_T(z,\alpha,c,\tau) = \underbrace{(M_T(z,\alpha,c,\tau) - E[M_T(z,\alpha,c,\tau)])}_{\xrightarrow{T\to\infty} 0 \quad \text{a.s.}} + \underbrace{E[M_T(z,\alpha,c,\tau)]}_{\xrightarrow{T\to\infty} M_\infty \quad \text{a.s.}}. \tag{A16}$$

The uniform consistency of M_T and uniform almost sure convergence of $E[M_T]$ to M_∞ imply the required uniform almost sure convergence of M_T to M_∞.

So then, the first step is to show that:

$$\sup_{z\in Z}\sup_{\alpha\in A}\sup_{c\in C}|M_T - E[M_T]| = 0 \quad \text{a.s.} \tag{A17}$$

The proof of Proposition 1 in Mack and Silverman (1982) illustrates one way to establish equation (A17) without assuming that marginal utility is bounded. M_T can be expressed as a sum of the following two terms:

$$M_T'(z,\alpha,c,\tau) = \frac{1}{Th_T^K}\sum_{\{t:|m_{t+1}|\le T^\lambda\}}\omega\left(\frac{z-z_t}{h_T}\right)m_{t+1}(\alpha,c,\tau), \tag{A18}$$

$$M_T''(z,\alpha,c,\tau) = \frac{1}{Th_T^K}\sum_{\{t:|m_{t+1}|> T^\lambda\}}\omega\left(\frac{z-z_t}{h_T}\right)m_{t+1}(\alpha,c,\tau), \tag{A19}$$

with $\lambda \in \,](2/b + d)^{-1}, (2/b)^{-1}[$, where b and d are the constants in Assumption 8. Given this choice of λ and the inequality in Assumption 8, it is straightforward to adapt the above-mentioned proof to show that:

$$\sup_{z\in Z}\sup_{\alpha\in A}\sup_{c\in C}|M''_T(z,\alpha,c,\tau) - E[M''_T(z,\alpha,c,\tau)]| = 0 \quad \text{a.s.} \tag{A20}$$

Standard asymptotic arguments and an inequality from Bosq (1988), which uses Assumption 16 to bound $P(|M'_T - E[M'_T]| > \epsilon)$, for $\epsilon > 0$, ensure the consistency of M'_T. Adding the continuity Assumptions 7, 8, and 14, together with the bandwidths Assumptions 15 and 16, yields uniform consistency, and hence completes the proof of equation (A17).

Finally, the same arguments as in the proof of Lemma 4 in Collomb and Härdle (1986), the equicontinuity Assumption 9, weighting function Assumption 13, the bandwidths Assumption 15, and the results of Parzen (1962), imply that:

$$\sup_{z\in Z}\sup_{\alpha\in A}\sup_{c\in C}|E[M_T] - M_\infty| = 0 \quad \text{a.s.} \tag{A21}$$

The main result follows.

G. Limiting Distribution

Expanding M_T around the portfolio and consumption choice $\{\alpha(z),c(z)\}$ yields

$$\begin{aligned}
&\underbrace{\frac{1}{Th_T^K}\sum_{t=1}^{T}\left(\omega\left(\frac{z-z_t}{h_T}\right)m_{t+1} - E\left[\omega\left(\frac{z-z_t}{h_T}\right)m_{t+1}\right]\right)}_{\xrightarrow{T\to\infty} N\left(0,\frac{1}{Th_T^K}\frac{V(z,\tau)}{f_z(z)}\int_{I\!R^K}\omega^2(u)\,du\right)} \\
&\quad + \underbrace{\frac{1}{Th_T^K}\sum_{t=1}^{T}\omega\left(\frac{z-z_t}{h_T}\right)\frac{\partial m_{t+1}}{\partial[\alpha',c]}}_{\xrightarrow{T\to\infty} D(z,\tau)}\sqrt{Th_T^K}\left(\begin{bmatrix}\hat{\alpha}_T(z,\tau)\\ \hat{c}_T(z,\tau)\end{bmatrix} - \begin{bmatrix}\alpha(z,\tau)\\ c(z,\tau)\end{bmatrix}\right) \\
&\qquad + \underbrace{\frac{1}{Th_T^K}\sum_{t=1}^{T}E\left[\omega\left(\frac{z-z_t}{h_T}\right)m_{t+1}\right]}_{\xrightarrow{T\to\infty} 0} \simeq 0,
\end{aligned} \tag{A22}$$

where m_{t+1} is evaluated at $\{\alpha(z),c(z)\}$.

The limiting distribution of the first term follows from the central limit theorem derived in Robinson (1983). The convergence of the second term to $D(z,\tau)$ can be established like the above consistency of M_T. Finally, the weighting function Assumption 13 and the bandwidths Assumption 15 ensure that the the third term vanishes as $T \to \infty$.

REFERENCES

Balduzzi, Pierluigi, and Anthony Lynch, 1996, Transaction costs and predictability: Some utility cost calculations, Working paper, New York University.

Barberis, Nicholas, 1999, Investing for the long run when returns are predictable, *Journal of Finance*, forthcoming.

Black, Fisher, 1976, Studies of stock price volatility changes, *Proceedings of the Business and Economic Statistics Section, American Statistical Association* 1, 177–181.

Bollerslev, Tim, Ray Y. Chou, and Kenneth F. Kroner, 1992, ARCH modeling in finance: A review of the theory and empirical evidence, *Journal of Econometrics* 52, 5–59.

Bosq, Denis, 1988, Prévision non paramétrique d'un processus stationaire non boré, Working paper, University of Paris.

Brennan, Michael J., Eduardo S. Schwartz, and Ronald Lagnado, 1997, Strategic asset allocation, *Journal of Economic Dynamics and Control* 21, 1377–1403.

Campbell, John Y., 1987, Stock returns and the term structure, *Journal of Financial Econometrics* 18, 373–399.

Campbell, John Y., and Robert J. Shiller, 1988, The dividend price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.

Campbell, John Y., and Luis M. Viceira, 1996, Consumption and portfolio decisions when expected returns are time varying, NBER working paper #5857.

Canner, Niko, N. Gregory Mankiw, and David N. Weil, 1997, An asset allocation puzzle, *American Economic Review* 87, 181–191.

Clark, R. Malcom, 1975, A calibration curve for radiocarbon dates, *Antiquity* 49, 251–266.

Collomb, Gérard, and Wolfgang Härdle, 1986, Strong uniform convergence rates in robust nonparametric time series analysis and prediction: Kernel regression estimation from dependent observations, *Stochastic Processes and Their Applications* 23, 77–89.

Conley, Timothy G., Lars Peter Hansen, Erzo G. J. Luttmer, and Jose A. Scheinkman, 1997, Short-term interest rates as subordinated diffusions, *Review of Financial Studies* 10, 525–577.

Fama, Eugene F., and Kenneth R. French, 1988a, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3–25.

Fama, Eugene F., and Kenneth R. French, 1988b, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246–273.

Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.

Fama, Eugene F., and G. William Schwert, 1977, Asset returns and inflation, *Journal of Financial Economics* 5, 115–146.

Fan, Jianqing, Nancy E. Heckman, and Matt P. Wand, 1995, Local polynomial smoothing in generalized linear models and quasi-likelihood functions, *Journal of the American Statistical Association* 90, 141–150.

Ferson, Wayne E., and Campbell R. Harvey, 1991, The variation of economic risk premiums, *Journal of Political Economy* 99, 385–415.

French, Kenneth R., G. William Schwert, and Robert F. Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 19, 3–29.

Gourieroux, Christian, Alain Monfort, and Carlos Tenreiro, 1995, Kernel M-estimators and functional residuals plots, Working paper, CREST.

Gozalo, Pedro L., and Oliver B. Linton, 1994, Local nonlinear least squares estimation: Using parametric information nonparametrically, Working paper, Cowles Foundation.

Hansen, Lars Peter, 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–1054.

Hansen, Lars Peter, and Kenneth J. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica* 50, 1269–1286.

Härdle, Wolfgang, 1984, How to determine the bandwidths of nonlinear smoothers in practice; in Jürgen Franke, Wolfgang Härdle, and Douglas Martin, eds.: *Robust and Nonlinear Time Series Analysis* (Springer Verlag, Heidelberg).

Härdle, Wolfgang, 1984, Robust regression function estimation, *Journal of Multivariate Analysis* 14, 169–180.

Härdle, Wolfgang, 1990, *Applied Nonparametric Regression* (Cambridge University Press, New York).

Härdle, Wolfgang, and James S. Marron, 1985, Optimal bandwith selection in nonparametric regression function estimation, *Annals of Statistics* 13, 1465–1481.

Hart, Jeffrey D., and Philippe Vieu, 1990, Data-driven bandwidth choice for density estimation based on dependent data, *Annals of Statistics* 18, 873–890.

Harvey, Campbell R., 1989, Time-varying conditional covariance in tests of asset pricing models, *Journal of Financial Economics* 24, 289–317.

Huber, Peter J., 1964, Robust estimation of a location parameter, *Annals of Mathematical Statistics* 35, 73–101.

Kandel, Shmuel, and Robert F. Stambaugh, 1996, On the predictability of stock returns: An asset-allocation perspective, *Journal of Finance* 51, 385–424.

Keim, Donald B., and Robert F. Stambaugh, 1986, Predicting returns in the stock and bond markets, *Journal of Financial Economics* 17, 357–390.

Luttmer, Erzo G. J., 1997, What level of fixed costs can reconcile asset returns and consumption choices?, Working paper, London School of Economics.

Mack, Yue-pok, and Bernard W. Silverman, 1982, Weak and strong uniform consistency of kernel regression estimates, *Zeitschrift für Wahrscheinlichkeitstheorie and Verwandte Gebiete* 61, 405–415.

Merton, Robert C., 1969, Lifetime portfolio selection under uncertainty: The continuous-time case, *Review of Economics and Statistics* 51, 247–257.

Morvai, Gusztáv, 1991, Empirical log-optimal portfolio selection, *Problems in Control Information Theory* 20, 453–463.

Morvai, Gusztáv, 1992, Portfolio choice based on the empirical distribution, *Kybernetika* 28, 484–493.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Pagan, Adrian R., and G. William Schwert, 1990, Alternative models for conditional stock volatility, *Journal of Econometrics* 45, 267–290.

Parzen, Emanuel, 1962, On estimation of a probability density and mode, *Annals of Mathematical Statistics* 33, 1065–1076.

Poterba, James, and Lawrence H. Summers, 1988, Mean reversion in stock prices, *Journal of Financial Economics* 22, 27–59.

Pritsker, Matt, 1998, Nonparametric density estimation and tests of continuous time interest rate models, Working paper, Federal Reserve Board.

Robinson, Peter M., 1983, Nonparametric estimators for time series, *Journal of Time Series Analysis* 4, 185–207.

Schwert, G. William, 1989, Why does stock market volatility change over time?, *Journal of Finance* 44, 1115–1153.

Siegel, Jeremy J., 1994, *Stocks for the Long Run* (Richard D. Irwin, Burr Ridge, Ill.).

Silverman, Bernard W., 1986, *Density Estimation for Statistics and Data Analysis* (Chapman and Hall, New York).

Solnik, Bruno, 1993, The performance of international asset pricing strategies using conditioning information, *Journal of Empirical Finance* 1, 33–35.

Stone, Mervyn, 1976, Cross-validatory choice and assessment of statistical prediction, *Journal of the Royal Statistical Society, Series B* 38, 111–147.

Tsybakov, Alexander B., 1982, Robust estimates of a function, *Problems of Information Transmission* 18, 190–201.