Consumption, Aggregate Wealth, and Expected Stock Returns

MARTIN LETTAU and SYDNEY LUDVIGSON*

ABSTRACT

This paper studies the role of fluctuations in the aggregate consumption–wealth ratio for predicting stock returns. Using U.S. quarterly stock market data, we find that these fluctuations in the consumption–wealth ratio are strong predictors of both real stock returns and excess returns over a Treasury bill rate. We also find that this variable is a better forecaster of future returns at short and intermediate horizons than is the dividend yield, the dividend payout ratio, and several other popular forecasting variables. Why should the consumption–wealth ratio forecast asset returns? We show that a wide class of optimal models of consumer behavior imply that the log consumption–aggregate wealth (human capital plus asset holdings) ratio summarizes expected returns on aggregate wealth, or the market portfolio. Although this ratio is not observable, we provide assumptions under which its important predictive components for future asset returns may be expressed in terms of observable variables, namely in terms of consumption, asset holdings and labor income. The framework implies that these variables are cointegrated, and that deviations from this shared trend summarize agents' expectations of future returns on the market portfolio.

UNDERSTANDING THE EMPIRICAL LINKAGES between macroeconomic variables and financial markets has long been a goal of financial economics. One reason for the interest in these linkages is that expected excess returns on common stocks appear to vary with the business cycle. This evidence suggests that stock returns should be forecastable by business cycle variables at cyclical frequencies. Indeed, the forecastability of stock returns is well documented. Financial indicators such as the ratios of price to dividends, price to earnings, or dividends to earnings have predictive power for excess returns over a Treasury-bill rate. These financial variables, however, have been most successful at predicting returns over long horizons. Over horizons spanning the

length of a typical business cycle, stock returns have typically been found to be only weakly forecastable.[1] Moreover, traditional macroeconomic variables have proven especially dismal as predictive variables.

The question of whether expected returns vary at cyclical frequencies and with macroeconomic variables is also pertinent to the debate over *why* excess returns are predictable. One possibility is that financial markets are inefficient. Alternatively, predictable variation in returns could simply reflect the rational response of agents to time-varying investment opportunities, possibly driven by cyclical variation in risk aversion (e.g., Sundaresan (1989), Constantinides (1990), Campbell and Cochrane (1999)) or in the joint distribution of consumption and asset returns. If these rational explanations are correct, it is reasonable to expect that key macroeconomic variables should perform an important function in forecasting excess stock returns. As yet, however, there is little empirical evidence that real macroeconomic variables perform such a function.

This paper adopts a new approach to investigating the linkages between macroeconomics and financial markets. We begin by noting that aggregate consumption, asset holdings, and labor income share a common long-term trend, but may deviate substantially from one another in the short run. We study the role of these transitory deviations from the common trend in consumption, asset holdings, and labor income for predicting stock market fluctuations. Our results show that these "trend deviations" are a strong univariate predictor of both raw stock returns and excess stock returns over a Treasury bill rate, and can account for a substantial fraction of the variation in future returns. This variable provides information about future stock returns that is not captured by lagged values of other popular forecasting variables, and displays its greatest predictive power for returns over business cycle frequencies, those ranging from one to five quarters. In addition, we find that observations on this variable would have improved out-of-sample forecasts of excess stock returns in postwar data relative to a variety of alternative forecasting models. These results occur despite the fact that the individual growth rates of consumption, labor income, and wealth, like other macroeconomic variables, bear little relationship to future stock returns.

Why should deviations from the common long-term trend in consumption, asset wealth, and labor income forecast asset returns? We show that this feature of the data may arise as an implication of a wide range of forward-looking models of investor behavior where consumption is a function of aggregate wealth (the "market" portfolio), defined as the sum of human and asset wealth. To make the framework tractable, we employ a log-linear approximation of the intertemporal budget constraint. For a wide class of pref-

---

[1] One exception to this is a study by Campbell (1987) which finds that Treasury bill rates and several measures of the term spread can explain a substantial fraction of the variation in next month's excess stock return. We also confirm in our extended sample that a stochastically detrended short rate has modest forecasting power for returns at business cycle frequencies.

erences, the log consumption–aggregate wealth ratio predicts asset returns because it is a function of expected future returns on the market portfolio. This result has been noted previously by Campbell and Mankiw (1989) and is the starting point of our theoretical framework.

There are two important obstacles that must be overcome before the log consumption–aggregate wealth ratio can be empirically linked with future asset returns. The most immediate is that aggregate wealth—specifically the human capital component of it—is unobservable. This paper argues that the important predictive components of the consumption–aggregate wealth ratio for future market returns may be expressed in terms of observable variables, namely in terms of consumption, asset holdings, and current labor income.

The model we investigate implies that the log of consumption, labor income, and asset holdings share a common stochastic trend. They are cointegrated. The parameters of this shared trend are the average shares of human capital and asset wealth in aggregate wealth. If expected consumption growth is not too volatile, stationary deviations from the shared trend among these three variables produce movements in the consumption–aggregate wealth ratio and predict future asset returns. This follows from the fact that the consumption–aggregate wealth ratio summarizes agents' expectations of future returns on the market portfolio. Accordingly, deviations from the shared trend will forecast returns to asset holdings, as long as the expected return to human capital is not too volatile.

A remaining obstacle to using deviations in the common trend among consumption, labor income, and asset wealth as a forecasting variable is that the parameters of this shared trend are unobservable and must be estimated. In ordinary empirical applications this estimation is problematic due to the presence of endogenous regressors. In our application, however, consumption, labor income, and asset wealth are cointegrated. We can obtain a "superconsistent" estimate of the cointegrating parameters that will be robust to the presence of regressor endogeneity.

What is the economic intuition for our results? Investors who want to maintain a flat consumption path over time will attempt to "smooth out" transitory movements in their asset wealth arising from time variation in expected returns. When excess returns are expected to be higher in the future, forward-looking investors will react by increasing consumption out of current asset wealth and labor income, allowing consumption to rise above its common trend with those variables. When excess returns are expected to be lower in the future, these investors will react by decreasing consumption out of current asset wealth and labor income, and consumption will fall below its shared trend with these variables. In this way, investors may insulate future consumption from fluctuations in expected returns, and stationary deviations from the shared trend among consumption, asset holdings, and labor income are likely to be a predictor of excess stock returns, consistent with what we find.

The rest of the paper is organized as follows. The next section presents the theoretical framework linking consumption, aggregate wealth, and expected returns, and shows how we express the important predictive components of the consumption–aggregate wealth ratio in terms of observable variables. In Section II, we present the results of estimating the trend relationship among consumption, labor income, and asset holdings. We then move on to test the important implication of the framework presented in Section II, that deviations from trend asset wealth are likely to lead stock returns. Section III discusses the data used in our forecasting regressions for asset returns and presents some summary statistics. Sections IV, V, and VI document our main findings on the predictability of stock returns. Section VII concludes.

I. The Consumption–Wealth Ratio

This section presents a general framework linking consumption, asset holdings, and labor income with expected returns.

Consider a representative agent economy in which all wealth, including human capital, is tradable. Let W_t be aggregate wealth (human capital plus asset holdings) in period t. C_t is consumption and $R_{w,t+1}$ is the net return on aggregate wealth. The accumulation equation for aggregate wealth may be written[2]

$$W_{t+1} = (1 + R_{w,t+1})(W_t - C_t). \tag{1}$$

We define $r \equiv \log(1 + R)$, and use lowercase letters to denote log variables throughout. Campbell and Mankiw (1989) show that, if the consumption–aggregate wealth ratio is stationary, the budget constraint may be approximated by taking a first-order Taylor expansion of the equation. The resulting approximation gives an expression for the log difference in aggregate wealth

$$\Delta w_{t+1} \approx k + r_{w,t+1} + (1 - 1/\rho_w)(c_t - w_t) \tag{2}$$

where ρ_w is the steady-state ratio of new investment to total wealth, $(W - C)/W$, and k is a constant that plays no role in our analysis.[3] Solving this difference equation forward and imposing that $\lim_{i \to \infty} \rho_w^i (c_{t+i} - w_{t+i}) = 0$, the log consumption–wealth ratio may be written

$$c_t - w_t = \sum_{i=1}^{\infty} \rho_w^i (r_{w,t+i} - \Delta c_{t+i}). \tag{3}$$

---

[2] Labor income does not appear explicitly in this equation because of the assumption that the market value of tradable human capital is included in aggregate wealth.

[3] We omit unimportant linearization constants in the equations from now on.

Equation (3) holds simply as a consequence of the agent's intertemporal budget constraint and therefore holds ex post, but it also holds ex ante. Accordingly, we can take conditional expectations of both sides of (3) to obtain

$$c_t - w_t = E_t \sum_{i=1}^{\infty} \rho_w^i (r_{w,t+i} - \Delta c_{t+i}), \tag{4}$$

where E_t is the expectation operator conditional on information available at time t. Equation (4) shows that, if the aggregate consumption–wealth ratio is not constant, it must forecast changing returns to the market portfolio or changing consumption growth. Put another way, the consumption–wealth ratio can only vary if consumption growth or returns or both are predictable.

As suggested by equation (4), the consumption–wealth ratio is a function of expected future returns to the market portfolio in a broad range of optimal consumption models. The information set upon which expectations are conditioned will depend on the state variables in the model. These models may differ in their specification of preferences, or in the assumptions about the stochastic properties of consumption and asset returns. All of them, however, will imply that the consumption–aggregate wealth ratio is a function of expected future returns, and that agents' expectations about future returns and consumption growth may be inferred from observable consumption behavior. Moreover, we do not need to explicitly model how returns to wealth and consumption growth are determined by some specific set of preferences.

Because aggregate wealth, in particular, human capital, is not observable, the framework presented above is not directly suited for predicting asset returns. To overcome this obstacle, we assume that the nonstationary component of human capital, denoted H_t, can be well-described by aggregate labor income, Y_t, implying that $h_t = \kappa + y_t + z_t$, where κ is a constant and z_t is a mean zero stationary random variable. This assumption may be rationalized by a number of different specifications linking labor income to the stock of human capital. First, labor income may be described as the annuity value of human wealth, $Y_t = R_{h,t+1} H_t$, where $R_{h,t+1}$ is the net return to human capital. In this case (ignoring a linearization constant), $r_{h,t} \equiv \log(1 + R_{h,t+1}) \approx 1/\rho_y (y_t - h_t)$, where $\rho_y \equiv (1 + Y/H)/(Y/H)$, implying $z_t = -\rho_y r_{h,t}$. This specification places no restrictions on the functional form of expected or realized returns, and it makes no assumptions about the relationship between returns to human capital and returns to asset wealth.[4] Alternatively, one could specify a "Gordon growth model" for human capital by assuming that expected returns to human capital are constant and labor income follows a random walk, in which case z_t is a constant equal to $\log(R_h)$. Finally, aggregate labor income can be thought of as the dividend on human capital, as in Campbell (1996) and Jagannathan and

---

[4] This specification ignores effects that arise from an endogenous choice of labor supply in models where utility over leisure and consumption is nonseparable.

Wang (1996). In this case, the return to human capital may be defined as $R_{h,t+1} = (H_{t+1} + Y_{t+1})/H_t$, and a log-linear approximation of $R_{h,t+1}$ implies that $z_t = E_t \sum_{j=0}^{\infty} \rho_h^j (\Delta y_{t+1+j} - r_{h,t+1+j})$. In each of these specifications, the log of aggregate labor income captures the nonstationary component of human capital.

We are now in a position to express the important predictive components of the consumption–aggregate wealth ratio in terms of observable variables. Let A_t be asset holdings, and let $1 + R_{a,t}$ be its gross return. Aggregate wealth is therefore $W_t = A_t + H_t$ and log aggregate wealth may be approximated as

$$w_t \approx \omega a_t + (1 - \omega) h_t, \tag{5}$$

where ω equals the average share of asset holdings in total wealth, A/W. This ratio may also be expressed in terms of steady-state labor income and returns as $R_h A/(Y + R_h A)$.

The return to aggregate wealth can be decomposed into the returns of its two components

$$1 + R_{w,t} = \omega_t (1 + R_{a,t}) + (1 - \omega_t)(1 + R_{h,t}). \tag{6}$$

Campbell (1996) shows that (6) may be transformed into an approximate equation for log returns taking the form

$$r_{w,t} \approx \omega r_{a,t} + (1 - \omega) r_{h,t}. \tag{7}$$

Substituting (7) into the ex ante budget constraint (4), (again ignoring constants) gives

$$c_t - \omega a_t - (1 - \omega) h_t = E_t \sum_{i=1}^{\infty} \rho_w^i \{ [\omega r_{a,t+i} + (1 - \omega) r_{h,t+i}] - \Delta c_{t+i} \}. \tag{8}$$

This equation still contains the unobservable variable h_t on the left-hand side. To remove it, we substitute our formulation linking the log of labor income to human capital, $h_t = \kappa + y_t + z_t$, into (8), which yields an approximate equation describing the log consumption–aggregate wealth ratio using only observable variables on the left-hand side:

$$c_t - \omega a_t - (1 - \omega) y_t = E_t \sum_{i=1}^{\infty} \rho_w^i \{ [\omega r_{a,t+i} + (1 - \omega) r_{h,t+i}] - \Delta c_{t+i} \} + (1 - \omega) z_t. \tag{9}$$

Since all the terms on the right-hand side of (9) are presumed stationary, c, a, and y must be cointegrated, and the left-hand side of (9) gives the deviation in the common trend of c_t, a_t, and y_t. In what follows, we denote the trend deviation term $c_t - \omega a_t - (1 - \omega) y_t$ as cay_t. Moreover, equation (9) shows that cay_t will be a good proxy for market expectations of future asset

returns, $r_{a,t+i}$, as long as expected future returns on human capital, $r_{h,t+i}$, and consumption growth, Δc_{t+i}, are not too variable, or as long as these variables are highly correlated with expected returns on assets.

It is instructive to compare (9) to an expression for another variable that has been widely used to forecast asset returns, the log dividend–price ratio. Let d_t and p_t be log dividend and log price, respectively, of the stock of asset wealth. Campbell and Shiller (1988) show that the log dividend–price ratio may be written

$$d_t - p_t = E_t \sum_{j=1}^{\infty} \rho_a^j (r_{a,t+j} - \Delta d_{t+j}), \tag{10}$$

where $\rho_a = P/(P + D)$. This equation is often referred to as the "dynamic dividend growth model" and is derived by taking a first-order Taylor approximation of the equation defining the log stock return, $r_t = \log(P_t + D_t) - \log(P_t)$. This equation says that if the dividend–price ratio is high, agents must be expecting either high returns on assets in the future or low dividend growth rates.

Note the similarity between (10) and our expression for the consumption–wealth ratio, (4). Both hold ex post as well as ex ante. The role of consumption in (4) is directly analogous to that of d_t in (10): When the consumption–aggregate wealth ratio is high, agents must be expecting either high returns on the market portfolio in the future or low consumption growth rates. Thus, consumption may be thought as the dividend paid from aggregate wealth. Unlike dividends, however, the determinants of consumption are more readily defined by theory and one can combine the budget constraint formulation in (4) with various models of consumer behavior. Note that in an exchange economy without labor income, aggregate consumption is equal to aggregate dividends and the consumption–wealth ratio is a scale transformation of the dividend–price ratio. This tight link between $c_t - w_t$ and $d_t - p_t$ is broken in economies with labor income, or if there is a saving technology so that agents are not forced to consume their endowment.

II. Estimating the Trend Relationship Among Consumption, Labor Income, and Asset Holdings

An important task in using cay_t to forecast asset returns is the estimation of the parameters of the shared trend in consumption, labor income, and wealth in (9).[5] At first glance, it may appear that obtaining a consistent estimate of these parameters would be difficult because c_t, a_t, and y_t are endogenously determined. This section discusses how we apply the asymptotic properties of cointegrated variables to circumvent this difficulty.

[5] This section draws from Ludvigson and Steindel (1999). That paper studies the properties of a vector-error-correction representation for consumption, asset wealth, and labor income and focuses on its implications for consumption, but does not address the issue of stock market predictability.

Before estimating the parameters of the shared trend, we deal with a measurement issue that arises from the nature of the data on consumption. Previous empirical work that has investigated consumption-based models has used expenditures on nondurables and services as a measure of consumption. The use of these expenditure categories is justified on the grounds that the theory applies to the *flow* of consumption; expenditures on durable goods are not part of this flow because they represent replacements and additions to a stock, rather than a service flow from the existing stock. But because nondurables and services expenditure is only a component of unobservable total consumption, the standard solution to this problem requires the researcher to assume that total consumption is a constant multiple of nondurable and services consumption (Blinder and Deaton (1985), Galí (1990)).

We follow in this tradition and use nondurables and services as our consumption measure, and assume a constant scale factor governing the relationship between the log of total consumption and the log of nondurables consumption, denoted $c_{n,t}$. Thus we write log total consumption, $c_t = \lambda c_{n,t}$, where $\lambda > 1$, implying that the estimated cointegrating vector for $c_{n,t}$, a_t, and y_t will be given by $[1, -(1/\lambda)\omega, -(1/\lambda)(1-\omega)]$.[6] We define $\beta_a = (1/\lambda)\omega$, and $\beta_y = (1/\lambda)(1-\omega)$, the parameters of the cointegrating relation to be estimated. Note that $\beta_a + \beta_y$ identifies $1/\lambda$.

The data used for this estimation are quarterly, seasonally adjusted, per capita variables, measured in 1992 dollars. The Appendix provides details on data construction and the source for all the data we use. As a preliminary step, we test whether each variable passes a unit root test. Given that consumption, labor income, and household net worth all appear to contain a unit root, we then move on to test for the presence of cointegration in our sample, which spans the period from the fourth quarter of 1952 to the third quarter of 1998. We provide the results of these tests in the Appendix, and simply note here that there is strong evidence supporting the hypothesis of a single cointegrating vector for consumption, labor income, and asset holdings.

To estimate β_a and β_y, we employ a method that generates optimal estimates of the cointegrating parameters in a multivariate setting. We follow Stock and Watson (1993) and use a dynamic least squares (*DLS*) technique that specifies a single equation taking the form

$$c_{n,t} = \alpha + \beta_a a_t + \beta_y y_t + \sum_{i=-k}^{k} b_{a,i} \Delta a_{t-i} + \sum_{i=-k}^{k} b_{y,i} \Delta y_{t-i} + \epsilon_t, \qquad (11)$$

where Δ denotes the first difference operator.

---

[6] Previous research has worked with formulations in levels, rather than in logs as we do here. Because Blinder and Deaton (1985) report that the share of nondurables and services in measured expenditures has displayed a secular decline over the sample period, the assumption that total consumption is a constant multiple of nondurable consumption may be questionable. By contrast, we postulate that the *log* of total consumption is a constant multiple of the log of nondurable and services consumption. Unlike the ratio of levels, the ratio of logs appears to have exhibited little secular movement during our sample period.

Equation (11) is estimated by *OLS*, and this methodology provides a consistent estimate of the cointegrating parameters through its estimates of β_a and β_y. The *DLS* specification adds leads and lags of the first difference of the right-hand side variables to a standard *OLS* regression of consumption on labor income and asset holdings to eliminate the effects of regressor endogeneity on the distribution of the least squares estimator. We denote the estimated trend deviation by $\widehat{cay}_t \equiv c_{n,t} - \hat{\beta}_a a_t - \hat{\beta}_y y_t$, where "hats" denote estimated parameters.[7]

It is important to recognize that estimates of β_a and β_y will be consistent despite the fact that ϵ_t will typically be correlated with the regressors a_t and y_t. This follows from the fact that *OLS* estimates of cointegrating parameters are "superconsistent," converging to the true parameter values at a rate proportional to the sample size T rather than proportional to $\sqrt{T}$ as in ordinary applications (Stock (1987)). This means that the data should provide a consistent estimate of ω, in effect making observable the average ratios of each component of wealth.

Implementing the regression in (11) using data from the fourth quarter of 1952 to the third quarter of 1998 generates the following point estimates (ignoring coefficient estimates on the first differences) for the parameters of the shared trend consumption, labor income, and wealth:

$$c_{n,t} = 0.61 + 0.31a_t + 0.59y_t, \qquad (12)$$
$$\quad\;\,(7.96)\quad(11.70)\quad\;(23.92)$$

where the corrected *t*-statistics appear in parentheses below the coefficient estimates.[8] The coefficient estimates suggest that λ is about 1.10, implying that the share of asset holdings in aggregate wealth is close to one-third, whereas the share of human capital is close to two-thirds. These values are consistent with what one would expect if aggregate production is governed by a Cobb–Douglas technology. In this case, total output, $Y_t = K_t^\alpha L_t^{1-\alpha}$, the payments to labor are given by $(1 - \alpha)Y_t$ and the payments to capital are given by αY_t. If human and asset wealth are discounted at the same rate, R_{t+i}, the stock of asset wealth is $A_t = \alpha E_t \sum_{i=1}^{\infty}[Y_{t+i}/(\prod_{j=1}^{i} R_{t+i})]$, and the stock of human capital is $H_t = (1 - \alpha)E_t \sum_{i=1}^{\infty}[Y_{t+i}/(\prod_{j=1}^{i} R_{t+i})]$. Thus the estimates in (12) imply that $\alpha = 0.34$, and are very close to values used in the real business cycle literature (see, e.g., Kydland and Prescott (1982) and Hansen (1985)).

How can we interpret deviations from the shared trend in consumption, labor income, and assets? Are they better described as transitory movements in asset wealth or as transitory movements in consumption and labor in-

---

[7] The data on $\widehat{cay}_t$ can be downloaded from our websites: http://www.ny.frb.org/rmaghome/ economist/lettau/lettau.html and http://www.ny.frb.org/rmaghome/economist/ludvigson/ ludvigson.html.

[8] We experimented with various lead/lag lengths in estimating the *DLS* specification. For the results reported in (12), the value of $k = 8$ was used. Neither the cointegrating parameter estimates nor the forecasting results we present below are sensitive to the particular value of k for k ranging from one to eight. Similar estimates of the cointegrating parameters were also obtained using Johansen's (1988) full information maximum-likelihood technique.

Table I

Estimates from a Cointegrated VAR

This table reports the sum of estimated coefficients from cointegrated vector autoregressions (VAR) of the column variable on the row variable. t-statistics for the sum appear in parentheses. Significant coefficients at the five percent level are highlighted in bold face. $\widehat{cay}_t$ is $c_t - \hat{\beta}_a a_t - \hat{\beta}_y y_t$, where c_t is consumption, a_t is asset wealth, and y_t is labor income. The sample period is fourth quarter of 1952 to third quarter 1998.

Dependent variable	Equation		
	Δc_t	Δa_t	Δy_t
$\Delta c_{t-i, i=1,2}$	**0.296**	0.208	**0.575**
(t-stat)	(2.236)	(0.463)	(2.353)
$\Delta a_{t-i, i=1,2}$	0.019	0.177	0.081
(t-stat)	(0.623)	(1.635)	(1.400)
$\Delta y_{t-i, i=1,2}$	0.074	0.239	−0.060
(t-stat)	(1.020)	(0.926)	(−0.425)
$\widehat{cay}_{t-1}$	−0.015	**0.445**	0.100
(t-stat)	(−0.411)	(3.554)	(1.440)
$\bar{R}^2$	0.12	0.14	0.05

come? To answer this question, it is instructive to examine a three-variable, cointegrated vector autoregression (VAR) where the log difference in consumption, asset wealth, and labor income are each regressed on their own lags and an "error-correction term," equal to the lagged value of the estimated trend deviation, $\widehat{cay}_{t-1}$. We focus on the relationship between the estimated trend deviation and future growth rates of each variable.

Table I presents these results using a two lag VAR.[9] Note that, in this cointegrated vector autoregression, as well as in the forecasting regressions for asset returns presented in the next section, standard errors do not need to be adjusted to account for the use of the generated regressor, $\widehat{cay}_{t-1}$. This again follows from the fact that estimates of the cointegrating parameters converge to their true values at rate T, rather than at the usual rate $\sqrt{T}$. The table reveals two interesting properties of the data on consumption, household wealth, and labor income.

First, estimation of the asset growth equation shows that $\widehat{cay}_{t-1}$ predicts asset growth, implying that deviations in asset wealth from its shared trend with labor income and consumption uncover an important transitory variation in asset holdings. In Section IV below, we show that this variable predicts asset growth because the estimated trend deviation forecasts asset returns, consistent with the theoretical framework discussed above.

---

[9] This lag length was chosen in accordance with findings from Akaike and Schwartz tests. This system is also studied in Ludvigson and Steindel (1999).

A second feature of the data is revealed by inspecting the consumption and labor income growth regressions. Transitory variation in the (log) levels of a series requires forecastability of the growth rates. Both consumption and labor income growth are somewhat predictable by lags of consumption growth, as noted elsewhere (Flavin (1981), Campbell and Mankiw (1989)), but the adjusted R^2 statistics (especially for the labor income equation) are lower than those in the asset regression. More importantly, the magnitude of the coefficient on $\widehat{cay}_{t-1}$ in the asset growth equation is substantially larger than in either the consumption or labor income equation. Furthermore, this error-correction term does not enter at a statistically significant level in the equations for consumption or labor income growth.

These results suggest that deviations from the shared trend in consumption, labor income, and assets are better described as transitory movements in asset wealth than as transitory movements in consumption or labor income. When log consumption deviates from its habitual ratio with log labor income and log assets, it is asset wealth, rather than consumption or labor income, that is forecast to adjust until the equilibrating relationship is restored.

The next step in our analysis is to investigate the role of transitory movements in asset wealth in forecasting asset returns. Before doing so, we discuss the data used in this investigation and examine summary statistics for $\widehat{cay}_t$ and for our financial data.

III. Asset Return Data and Summary Statistics

Our financial data include stock returns, dividends per share, and quarterly earnings per share from the Standard & Poor's (S&P) Composite Index for which quarterly earnings data are available. In addition, we also consider returns on the value-weighted CRSP Index (CRSP-VW). The CRSP Index (which includes the NYSE, AMEX, and Nasdaq) should provide a better proxy for nonhuman components of total asset wealth because it is a much broader measure than is the S&P Index.

Let r_t denote the log real return of the index under consideration and $r_{f,t}$ the log real return on the 30-day Treasury bill (the "risk-free" rate). The log excess return is $r_t - r_{f,t}$. Log price, p, is the natural logarithm of the relevant index. Log dividends, d, are the natural logarithm of the sum of the past four quarters of dividends per share. Log earnings, e, are the natural logarithm of a single quarter's earnings per share. We call $d - p$ the dividend yield and, following Lamont (1998), $d - e$ the payout ratio.

Table II presents summary statistics for the variables mentioned above and for the relative bill rate, the T-bill rate minus its 12-month backward moving average. Campbell (1991) and Hodrick (1992) apply this stochastic detrending method to T bills in order to forecast returns.

The properties of stock returns, $d - p$, $d - e$, and the relative bill rate are well known; thus, we focus our discussion on the estimated trend deviation variable $\widehat{cay}$. This variable is contemporaneously positively correlated with excess stock returns, the dividend–price ratio, and the dividend–earnings

Table II
Summary Statistics

$r_t - r_{f,t}$ is quarterly log excess returns on the S&P Composite Index; $d_t - p_t$ is the log dividend yield; $d_t - e_t$ is the log dividend payout ratio; $RREL_t$ is the relative bill rate; $\widehat{cay}_t$ is $c_t - \hat{\beta}_a a_t - \hat{\beta}_y y_t$, where c_t is consumption, a_t is asset wealth, and y_t is labor income. The statistics are computed for the largest common span of available data for all the variables. The sample period is fourth quarter of 1952 to third quarter of 1998.

	$r_t\text{-}r_{f,t}$	$d_t\text{-}p_t$	$d_t\text{-}e_t$	$RREL_t$	$\widehat{cay}_t$
	Panel A: Correlation Matrix				
$r_t - r_f$	1.00	−0.15	0.08	−0.30	0.28
$d_t - p_t$		1.00	0.01	−0.01	0.43
$d_t - e_t$			1.00	−0.46	0.15
$RREL_t$				1.00	−0.23
$\widehat{cay}_t$					1.00
	Panel B: Univariate Summary Statistics				
Mean	0.016	−3.358	0.680	0.000	0.593
Standard error	0.077	0.277	0.199	0.011	0.011
Autocorrelation	0.12	0.93	0.70	0.71	0.79

ratio. The correlation with the relative bill rate is negative. However, none of the correlations are large (in absolute value). Relative to its mean, $\widehat{cay}$ varies less than $d - p$ and $d - e$.

How does the persistence of $\widehat{cay}_t$ compare to other variables known to forecast excess stock returns? It is well known that the price–dividend ratio is very persistent. The autocorrelation of $\widehat{cay}_t$ is fairly high, but substantially lower than for $d - p$, 0.79 compared to 0.93. Thus, the use of $\widehat{cay}_t$ in the forecasting equations below does not present the inference problems that arise with the very persistent dividend yield.

Figure 1 plots the standardized trend deviation, $\widehat{cay}_t$, and the standardized excess return on the S&P Composite Index over the period spanning the fourth quarter of 1952 to the third quarter of 1998. The figure shows a multitude of episodes during which positive trend deviations preceded large positive excess returns and negative ones preceded large negative returns. Moreover, large swings in the trend deviation tend to precede spikes in excess returns. This pattern is evident during the 1950s and early 1960s when $\widehat{cay}_t$ shot up prior to a sequence of up-ticks in excess returns, during the 1970s when sharp declines in $\widehat{cay}_t$ led the bear markets of those years, and during the 1980s when the trend deviation turned negative prior to the 1987 stock market crash. The trend deviation term also displays some notable cyclicality, typically rising during recessions and falling during booms.

To some extent, the tight link exhibited between these variables has broken down in the most recent period; $\widehat{cay}_t$ became negative in 1995 and declined sharply until the second quarter of 1998, whereas the stock market
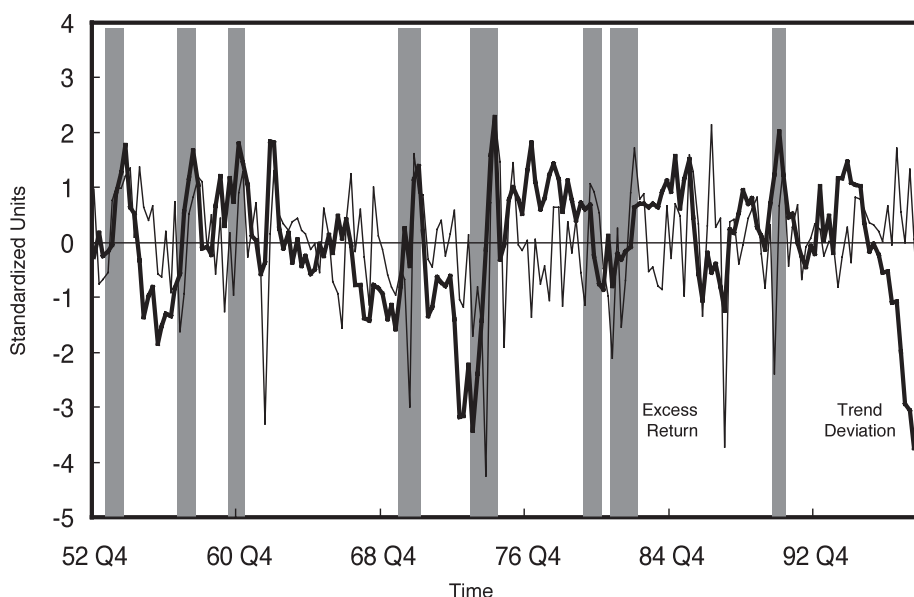
Figure 1. Excess returns and trend deviations. Excess return is the return on the S&P Composite Index less the return on the three-month Treasury bill rate. Trend deviation is the estimated deviation from the shared trend in consumption c, labor income y, and asset wealth a: $\widehat{cay}_t = c_t - \hat{\beta}_a a_t - \hat{\beta}_y y_t$. Both series are normalized to standard deviations of unity. The sample period is fourth quarter of 1952 to third quarter of 1998. Shaded areas denote NBER recessions.

(in a delayed response relative to its historical pattern) did not turn down until the third quarter of 1998.

Perhaps the most striking feature of the Figure 1 is how foreboding are current levels of $\widehat{cay}_t$ for returns in 2000 and beyond. This model is not alone in exhibiting such a bearish projection; the log dividend–price ratio is well below its historical mean, indeed at a postwar low in our S&P data. Of course, we cannot preclude the possibility that a structural shift has occurred in the underlying parameters governing these relationships. Nevertheless, the unusually low values of $\widehat{cay}_t$ in recent data suggests that consumers have factored the expectation of lower future stock returns into today's consumption.

IV. Quarterly Forecasting Regressions

We now move on to assess the forecasting power of detrended wealth for asset returns. Table III shows a typical set of results using the lagged trend deviation, $\widehat{cay}_t$, as a predictive variable. The table reports one-quarter-ahead forecasts of the real return on the S&P Composite Index and on the CRSP-VW Index as well as forecasts for excess returns. In all of the regres-

<div align="center">

Table III

Forecasting Quarterly Stock Returns

</div>

The table reports estimates from *OLS* regressions of stock returns on lagged variables named at the head of a column. All returns are in logs using the S&P 500 Index, except for regressions 4 and 8 (indicated by †), which use CRSP value-weighted returns. The regressors are as follows: *lag* denotes a one-period lag of the independent variables; $\widehat{cay}_t \equiv c_t - \hat{\beta}_a a_t - \hat{\beta}_y y_t$, where c_t is consumption, a_t is asset wealth, and y_t is labor income; $d_t - p_t$ is the log dividend yield; $d_t - e_t$ is the log dividend payout ratio; $RREL_t$ is the relative bill rate; TRM_t is the term spread, the difference between the 10-year Treasury bond yield and the 3-month Treasury bond yield; and DEF_t is the BAA Corporate Bond rate minus the AAA Corporate Bond rate. Newey–West corrected *t*-statistics appear in parentheses below the coefficient estimate. Significant coefficients at the five percent level are highlighted in bold face. Regressions use data from the fourth quarter of 1952 to the third quarter of 1998, except for regression 13 (denoted with a ‡), which begins in the second quarter of 1953, the largest common sample for which all the data are available.

#	Constant (*t*-stat)	*lag* (*t*-stat)	$\widehat{cay}_t$ (*t*-stat)	$d_t - p_t$ (*t*-stat)	$d_t - e_t$ (*t*-stat)	$RREL_t$ (*t*-stat)	TRM_t (*t*-stat)	DEF_t (*t*-stat)	$\bar{R}^2$
				Panel A: Real Returns; 1952:4–1998:3					
1	**0.017** (3.131)	**0.136** (2.221)							0.01
2	**0.029** (4.672)		**2.220** (3.024)						0.09
3	**0.026** (4.645)	0.062 (0.981)	**2.109** (2.806)						0.09
4†	**0.028** (4.889)	−0.007 (−0.157)	**2.513** (4.754)						0.10
				Panel B: Excess Returns; 1952:4–1998:3					
5	**0.014** (2.952)	0.119 (1.976)							0.00
6	**0.024** (4.328)		**2.165** (3.226)						0.09
7	**0.023** (4.345)	0.043 (0.707)	**2.089** (2.988)						0.09
8†	**0.022** (4.612)	−0.038 (−0.483)	**2.528** (4.583)						0.10
				Panel C: Additional Controls; Excess Returns; 1952:4–1998:3					
9	0.100 (1.157)			0.025 (0.965)					0.00
10	−0.012 (−0.135)		**2.274** (4.131)	−0.011 (−0.404)					0.09
11	0.058 (0.656)			0.025 (0.805)	0.060 (1.531)				0.02
12	−0.036 (−0.398)		**2.145** (4.182)	−0.009 (−0.267)	0.043 (1.289)				0.09
13‡	0.038 (0.278)	0.004 (−0.065)	**1.906** (3.197)	0.011 (0.095)	−0.004 (0.270)	**−1.377** (−2.443)	−0.082 (−0.125)	−0.883 (−0.543)	0.10

sions in Table III, we make a Newey–West correction (Newey and West (1987)) to the t-statistics for generalized serial correlation of the residuals. We call these our benchmark regressions.

Focusing on the S&P Composite Index, the first row of each panel of Table III shows that the forecasting power of a regression of returns on one lag of the dependent variable is quite weak. This model predicts only one percent of next quarter's variation in real returns, and a negligible percentage of next quarter's excess return variation. By contrast, the trend deviation explains a substantial fraction of the variation in next quarter's return. For the S&P Composite Index, regressions of real returns and excess returns on their own lags and on one lag of $\widehat{cay}_t$ both produce an adjusted R^2 of nine percent, so that adding last quarter's value of $\widehat{cay}_t$ to the model allows the regression to predict an additional nine percent of the variation in next period's excess return and an extra eight percent of the variation in next period's real return. Moreover, the Newey–West corrected t-statistic for this variable indicates that the coefficient estimate is nonzero with very high probability. These results are little affected by whether the lagged value of the real return is included in the regression as an additional explanatory variable (row 3).

The predictive impact of $\widehat{cay}_t$ on future returns is economically large: The point estimate of the coefficient on $\widehat{cay}_t$ is about 2.2 for real returns. To understand these units, note that the variables comprised in $\widehat{cay}_t$ are in per-capita terms, measured in billions of 1992 dollars, and that $\widehat{cay}_t$ itself has a standard deviation of about 0.01. Thus a one-standard-deviation increase in $\widehat{cay}_t$ leads to a 220 basis points rise in the expected real return on the S&P Index and about the same rise in the excess return, roughly a nine percent increase at an annual rate.[10]

Using the broader CRSP-VW Index as a measure of returns produces slightly higher $\bar{R}^2$ statistics than for the S&P 500 Index (Table III, rows 4 and 8). The t-statistics of the $\widehat{cay}_t$ coefficients increase to over four for both real returns and excess returns. According to the consumption framework presented above, $\widehat{cay}_t$ forecasts expectations of future returns to the market portfolio, so it is not surprising that it forecasts the broader CRSP-VW Index better than the S&P 500 Index.

These results accord well with the economic intuition from the framework presented in Section I. If returns are expected to decline in the future, investors who desire smooth consumption paths will allow consumption to dip temporarily below its long-term relationship with both assets and labor in-

---

[10] Relative to financial variables, macroeconomic variables are reported with a lag of about one month. Results (not reported) show that when the one-quarter lagged value, $\widehat{cay}_t$, is replaced with the two-quarter lagged value, $\widehat{cay}_{t-1}$, as a predictive variable, the coefficient on the two-period lagged variable is also a strongly statistically significant predictor of excess returns, although the point estimate falls to about 1.6. And, although, as would be expected, the $\bar{R}^2$ is somewhat lower than when the two-period lagged value is used, including the two-period lagged value of $\widehat{cay}_t$ into the benchmark equation still allows the regression to pick up an additional five percent of the variation in both next quarter's raw and excess returns. Note that the delay in the data release is less of an issue for the long-horizon forecasts we consider in Section 6.

come in an attempt to insulate future consumption from lower returns, and vice versa. Thus, investors' own optimizing behavior suggests that deviations in the long-term trend among c, a, and y should be positively related to future stock returns, consistent with what we find.

Not only does $\widehat{cay}_t$ covary positively with expected future returns, the variation in $\widehat{cay}_t$ is countercyclical: Its contemporaneous correlation with consumption growth and real GDP growth is -0.07 and -0.12 respectively, a phenomenon illustrated graphically in Figure 1, which shows that $\widehat{cay}_t$ tends to decline during expansions and rise just prior to the onset of a recession. Accordingly, expansions are characterized by increasing consumption, but an even greater rate of increase in assets. These features of the data may be interpreted using the time-varying risk aversion framework of Campbell and Cochrane (1999). In that model, consumption booms are periods during which consumption increases above habit, leading to a decline in risk aversion. The decline in risk aversion leads, in turn, to a greater demand for risky assets and a decrease in expected excess returns, or risk premia. Thus, in that model, booms are times of rising consumption but declining *ratios* of consumption to wealth, consistent with what we find.

How robust are the results? Panel C of Table III reports estimates from forecasting regressions that include a variety of variables shown elsewhere to contain predictive power for excess returns. Shiller (1984), Campbell and Shiller (1988), and Fama and French (1988) all find that the ratios of price to dividends or earnings have predictive power for excess returns. Lamont (1998) finds that the ratio of dividends to earnings has forecasting power at quarterly horizons. Campbell (1991) and Hodrick (1992) find that the relative T-bill rate (the 30-day T-bill rate minus its 12-month moving average) predicts returns, and Fama and French (1989) study the forecasting power of the term spread (the 10-year Treasury bond yield minus the 1-year Treasury bond yield) and the default spread (the difference between the BAA and AAA corporate bond rates). We include these variables in the benchmark equations for the excess return on the S&P Composite Index in Table III, Panel C.[11]

The first row of Table III, Panel C, shows that the dividend yield has virtually no effect on excess returns at a horizon of one quarter; the $\bar{R}^2$ statistic for this regression is negligible. This is not surprising because it is well known that this variable typically performs better at forecast horizons in excess of two years (Campbell (1991), Campbell, Lo, and MacKinlay (1997)). When we include the trend deviation in this regression with the dividend yield (row 10), the $\bar{R}^2$ statistic increases to nine percent, and the point estimate on $\widehat{cay}_t$ is strongly significant.

---

[11] In other tests (not reported), we included the one-period lagged value of consumption growth, labor income growth, and PCE inflation as predictive variables. None of these variables influence the coefficient estimates on $\widehat{cay}_t$, or the incremental $\bar{R}^2$ from including $\widehat{cay}_t$ in the regression. The former two are not statistically significant, and inflation, although individually significant, does not increase the explanatory power of the regression by a measurable amount.

In a recent paper, Lamont (1998) argues that the dividend payout ratio should be a potentially potent predictor of excess returns, a result of the fact that high dividends typically forecast high returns whereas high earnings typically forecast low returns. Row 11 of Table III shows the regression results when both the lagged dividend yield, $d_t - p_t$, and the lagged dividend payout ratio, $d_t - e_t$, are included in the forecasting equation. This regression has more explanatory power than the univariate model for returns used in Table III, Panel C, but the $\bar{R}^2$ statistic is still just two percent. Adding the trend deviation term, $\widehat{cay}_t$, again significantly improves the one-quarter-ahead predictive capacity of the regression; of the three, this variable is the only one with statistically significant explanatory power, and including it increases the $\bar{R}^2$ to nine percent.[12]

The final row of Table III, Panel C, augments the benchmark regression by including the lagged relative bill rate, the lagged term spread, and the lagged default spread along with the lagged dividend yield and the lagged payout ratio. Of these seven explanatory variables, the only ones that have significant marginal predictive power are the relative bill rate and the trend deviation term. Both the trend deviation and the relative bill rate are highly significant, and the estimated coefficient on the relative bill rate has the expected negative sign. Nevertheless, a comparison of rows 12 and 13 makes clear that the relative bill rate, although having marginal explanatory power, does not help explain much of the variation in next quarter's excess return. The $\bar{R}^2$ including this variable (row 12) is 10 percent, just one percent higher than in row 12 where the variable is excluded (the unadjusted R-squared statistics are similar). Furthermore, the coefficient estimate for $\widehat{cay}_t$ is little affected by the inclusion of the relative bill rate or other variables in the forecasting equation; regardless of which specification we consider, the point estimates are always between 1.9 and 2.3 and the t-statistics are above 3. This reveals that $\widehat{cay}_t$ contains information about future asset returns that is not included in other forecasting variables.

V. Out-of-Sample Tests

One possible concern about the forecasting results presented above is the potential for "look-ahead" bias due to the fact that the coefficients in $\widehat{cay}_t$ are estimated using the full sample. This concern may be addressed by performing out-of-sample forecasts where the parameters in $\widehat{cay}_t$ are reestimated every period, using only data available at the time of the forecast. The difficulty with this technique is that consistent estimation of the parameters in $\widehat{cay}_t$ requires a large number of observations, and an out-of-sample proce-

---

[12] Our results on the forecasting power of the dividend payout ratio and the dividend yield differ from those of Lamont (1998), who reports that both variables are strong predictors of the excess return on the S&P Index. The source of this discrepancy in our results is the data sampling period: Lamont's data spanned the period from the first quarter of 1947 to the fourth quarter of 1994, whereas our data runs from the fourth quarter of 1954 to the fourth quarter of 1998.

dure is likely to induce significant sampling error in the coefficient estimates during the early estimation recursions. This would make it more difficult for $\widehat{cay}_t$ to display forecasting power if the theory is true.

With this caveat in mind, we nevertheless compare the mean-squared error from a series of one-quarter-ahead out-of-sample forecasts obtained from a prediction equation that includes $\widehat{cay}_t$ as the sole forecasting variable, to a variety of forecasting equations that do not include $\widehat{cay}_t$. We begin by making nested forecast comparisons. That is, we compare the mean-squared forecasting error from an unrestricted model, which includes $\widehat{cay}_t$, to a restricted, benchmark model, which excludes this variable. Thus the unrestricted model *nests* the benchmark model.

A. *Nested Forecast Comparisons*

Before making this comparison, we must choose an appropriate benchmark. We do so by comparing the mean-squared forecasting error from a regression that includes just the lagged excess return, $r_t - r_{f,t}$, as a predictive variable, to the mean-squared error from regressions that included, in addition to this variable, combinations of the other variables discussed above (the log dividend–price ratio, the log dividend–earnings ratio, the relative bill rate, and the term spread). Including any of these control variables does not improve, and often even degrades, the out-of-sample predictive power of a regression that uses just the lagged dependent variable as a predictor. Accordingly, we take the more parsimonious model, using just the one-period lagged value of the excess return as a benchmark, and refer to this forecasting model as the *autoregressive benchmark*. For comparison, we also specify a model of constant expected returns (where a constant is the sole explanatory variable for excess returns) as the restricted model and refer to this as the *constant expected returns benchmark*. The constant expected returns benchmark is compared to one that includes both a constant and $\widehat{cay}$ as a predictive variable.

The nested comparisons are made by alternately augmenting the benchmark with either the one-period lagged value of $\widehat{cay}_t$, or the two-period lagged value, denoted $\widehat{cay}_{t-1}$. Results using the two-period lagged value are presented because macroeconomic indicators are available with a one-month delay relative to financial variables. Thus, the two-*quarter* lagged value of $\widehat{cay}$ is an overly conservative estimate of the information that would be available to a practitioner who wishes to forecast quarterly returns. (Of course, from an equilibrium point of view, agents know their current consumption and wealth and, hence, cay_t is in the information set at time t.) Lagging $\widehat{cay}$ an additional period comes at a cost because, according to the framework presented in (3), the two-period lagged value of the consumption–wealth ratio should not be expected to have as much predictive power as the one-period lagged value. We perform a few tests with the two-period lagged variable anyway, on the grounds that it may have some predictive power because it is fairly highly autocorrelated, having a first-order autocorrelation coefficient of 0.79 in our sample.

Each model is first estimated using data from the fourth quarter of 1952 to the first quarter of 1968. We use recursive regressions to reestimate both the parameters in $\widehat{cay}_t$ as well as the forecasting model each period, adding one quarter at a time and calculating a series of one-step-ahead forecasts. The forecasts are evaluated by comparing the mean-squared error from the set of one-step-ahead forecasts. We also present results based on a fixed cointegrating vector where the cointegrating parameters are set equal to their values estimated in the full sample. The latter case gives some idea of how the model would perform going forward if a practitioner used the existing estimates of these parameters and faced the same distribution of data. Moreover, this exercise is justified on the grounds that the parameters in the cointegrating relation are theoretically motivated and can be treated as known once a sufficient amount of data is available to obtain superconsistent estimates. Note that both lagging $\widehat{cay}_t$ an additional period and reestimating the cointegrating parameters every period puts the theory at a great disadvantage, because (according to the framework presented in the paper) the two-period lagged value of $\widehat{cay}$ with the additional sampling error induced by reestimation should not be expected to have as much predictive power as the one-period lagged value with parameters set at their theoretically correct values.

Table IV presents the results. Turning first to the comparisons with the autoregressive benchmark, the table shows that, regardless of whether the cointegrating parameters are reestimated, or whether the one- or two-period lagged value of $\widehat{cay}_t$ is used as a predictive variable, the mean-squared forecasting error of the $\widehat{cay}_t$-augmented model is always lower than that of the benchmark autoregressive model. Thus, information on the aggregate consumption–wealth ratio consistently improves forecasts over models that used only lagged returns as a predictive variable.

The bottom panel of Table IV shows that fixing the cointegrating parameters at their full sample values produces the greatest forecasting gains from augmenting the benchmark equation with $\widehat{cay}_t$. This is not surprising and suggests that reestimation of the cointegrating parameters induces greater sampling error into the parameter estimates, making it harder for the augmented models to register an improvement over the benchmark. Nevertheless, including $\widehat{cay}_t$ in the benchmark regression consistently improves forecasts even when the cointegrating vector is reestimated. Furthermore, using the one-period lagged value of $\widehat{cay}$ reduces the mean-squared error by more than a model that uses the two-period lagged value of this variable, but both models post improvements over the benchmark.

The $\widehat{cay}$ forecasting model also has superior forecasting performance relative to a model of constant expected excess returns. Several researchers have found that the dividend yield has no ability to predict out-of-sample relative to a constant expected returns model despite its ability to do so in-sample (e.g., Bossaerts and Hillion (1999), Goyal and Welch (1999)). Rows 3, 4, 7, and 8 of Table IV show that including a measure of $\widehat{cay}$ in the forecasting regression leads to a lower mean-squared forecasting error regardless of whether the cointegrating coefficients are reestimated or whether the one- or two-period lagged value of $\widehat{cay}$ is used.

One-Quarter-Ahead Forecasts of Excess Returns: Nested Comparisons

The table reports the results of one-quarter-ahead, nested forecast comparisons of excess returns on the S&P Composite Index, $r_{t+1} - r_{f,t+1}$. $\widehat{cay}_t$ is $c_t - \hat{\beta}_a a_t - \hat{\beta}_y y_t$, the estimated one-period lagged value of the deviation from the cointegrating relation among consumption c, asset wealth a, and labor income y. Rows 1, 2, 5, and 6 give forecast comparisons of an *unrestricted* model, which includes both the one-period lagged dependent variable $r_t - r_{f,t}$ and a lagged value of $\widehat{cay}$ as explanatory variables, with the autoregressive benchmark (*AR*), *restricted* model, which includes just the lagged dependent variable. Rows 3, 4, 7, and 8 give forecast comparisons of an unrestricted model that includes a constant and a lagged value of $\widehat{cay}$ as the sole explanatory variable, with the constant expected returns model (*const*), which includes just a constant. In rows 1 and 3, the one-period lagged value of $\widehat{cay}_t$ is used as an explanatory variable; rows 2 and 4 replace the one-period lagged value, $\widehat{cay}_t$, with the two-period lagged value, $\widehat{cay}_{t-1}$. In rows 3, 4, 7, and 8, the unrestricted model includes a constant and a lagged value of $\widehat{cay}$ as explanatory variables; the restricted model in these rows includes only a constant. Rows 3 and 7 use the one-period lagged value of $\widehat{cay}$; rows 4 and 8 use the two-period lagged value. MSE_u is the mean-squared forecasting error from the relevant unrestricted model in each row; MSE_r is the mean-squared error from the relevant restricted model. A number less than one indicates that the $\widehat{cay}$ augmented model has lower forecasting error than the restricted model. The column labeled "ENC-NEW" gives the modified Harvey, Leybourne, and Newbold test statistic (Clark and McCracken (1999)); the null hypothesis is that the restricted model encompasses the unrestricted model; the alternative is that the unrestricted model contains information that could be used to significantly improve the restricted model's forecast. The column labeled "MSE *F*" gives the results of the out-of-sample *F* test (McCracken (1999)); the null hypothesis is that the restricted and unrestricted models have equal mean-squared error (MSE); the alternative is that the restricted model has higher MSE. The subcolumn labeled "Statistic" gives the test statistic itself; the subcolumn labeled "Asy. CV" gives the 95th percentile of the asymptotic distribution of the statistic as derived in Clark and McCracken (1999); the subcolumn labeled "BS. CV" gives the 95th percentile of the bootstrapped distribution of the statistic derived under the null. The initial estimation period begins with the fourth quarter of 1954 and ends with the first quarter of 1968. The model is recursively reestimated until the third quarter of 1998.

Row	Comparison	MSE_u/MSE_r	ENC-NEW			MSE *F*		
			Statistic	Asy. CV	BS. CV	Statistic	Asy. CV	BS. CV
		Panel A: Cointegrating Vector Reestimated						
1	$\widehat{cay}_t$ vs. *AR*	0.975	9.15**	2.09	2.15	5.72**	1.52	1.74
2	$\widehat{cay}_{t-1}$ vs. *AR*	0.985	4.02**	2.09	2.21	1.85*	1.52	1.73
3	$\widehat{cay}_t$ vs. *const*	0.984	9.68**	2.09	1.99	2.00*	1.52	1.67
4	$\widehat{cay}_{t-1}$ vs. *const*	0.996	4.88**	2.09	2.23	0.38	1.52	1.71
		Panel B: Fixed Cointegrating Vector						
5	$\widehat{cay}_t$ vs. *AR*	0.922	12.93**	2.09	2.37	9.84**	1.52	1.36
6	$\widehat{cay}_{t-1}$ vs. *AR*	0.955	6.58**	2.09	2.43	5.72**	1.52	1.61
7	$\widehat{cay}_t$ vs. *const*	0.921	14.30**	2.09	2.31	10.52**	1.52	1.48
8	$\widehat{cay}_{t-1}$ vs. *const*	0.957	8.15**	2.09	2.36	5.48**	1.52	1.36

*Significant at the five percent or better level.
**Significant at the one percent or better level.

For each forecast comparison, we also ask whether the superior performance of the $\widehat{cay}_t$-augmented forecasts is statistically significant. Table IV provides two test statistics designed to determine whether the one-step-ahead forecasting performance from the restricted, benchmark model is statistically different from an unrestricted model that includes $\widehat{cay}_t$. Clark and McCracken (1999) derive the (nonstandard) asymptotic distributions for a large number of statistical tests as applied to nested models and numerically generate the asymptotic critical values. The two tests we consider are those found by Clark and McCracken to have the best overall power and size properties. The first, called the ENC-NEW test, is a modified Harvey, Leybourne, and Newbold (1998) test statistic adapted to address the fact that the limiting distribution of this test statistic is nonnormal when the forecasts are nested under the null. The ENC-NEW statistic provides a test of the null hypothesis that the restricted model's forecast "encompasses" all the relevant information for next period's value of the dependent variable, against the alternative that the unrestricted model contains additional information. The second is an out-of-sample F-type test (MSE F test) developed in McCracken (1999). The MSE F test is a test of equal mean-squared forecasting error. The null hypothesis for this test is that the restricted model, which excludes $\widehat{cay}$, has a mean-squared forecasting error that is less than or equal to that of the unrestricted model that includes $\widehat{cay}$; the alternative is that the unrestricted model has lower mean-squared error.

Table IV presents the ENC-NEW and MSE F-test statistics along with their asymptotic 95 percent critical values derived in Clark and McCracken (1999). For comparison, 95 percent bootstrapped critical values of these statistics, generated by performing repeated simulations under the null hypothesis, are also presented. The bootstrapped statistics are derived by first estimating the model under the null hypothesis and then performing bootstrap simulations of the data by drawing randomly (with replacement) from the errors of the appropriate null model (i.e., autoregressive or constant expected return), and a vector autoregressive equation for $\widehat{cay}_t$. Data for returns and $\widehat{cay}_t$ are formed by iterating forward using these randomly chosen errors, and forecasts of returns are computed using the simulated data for both the nested model (true under the null) and the nonnested model (which includes $\widehat{cay}_t$ as a forecasting variable), and the test statistics are computed. This is done 10,000 times. The 95th percentile of this bootstrapped statistic is reported in Table IV.

The use of this bootstrap procedure addresses a concern raised by several recent studies, namely that the researcher is likely to find spurious evidence of return predictability in small samples when the forecasting variable is sufficiently persistent (e.g., Nelson and Kim (1993), Ferson, Sarkissian, and Simin (1999), Stambaugh (1999)). For example, the bootstrapped statistics in rows 3, 4, 7, and 8 of Table IV are determined by the data sample we use under the null hypothesis of no return predictability.

For each set of one-step-ahead forecasts we consider, the ENC-NEW and MSE F test both strongly reject the null that $\widehat{cay}_t$ contains no information about future excess returns that could be used to improve upon the forecast

from the autoregressive benchmark when the test statistics are compared against their asymptotic critical values. The boot-strapped critical values are very similar and do not change these conclusions. The same conclusions are drawn when comparing the *cay*-augmented model to the constant expected returns model. The ENC-NEW test statistic strongly rejects the null that $\widehat{cay}$ has no predictive power for excess returns. For every case except one, the MSE *F*-test also rejects the null of equal forecast accuracy between the constant expected returns benchmark and the $\widehat{cay}$-augmented model. The one exception is the case when $\widehat{cay}$ is both lagged two periods and re-estimated; for this case the MSE *F*-test does not reject the null of equal forecast accuracy, despite the fact that the $\widehat{cay}$-augmented model does display a lower mean-squared error than the constant expected returns benchmark. This finding is not particularly surprising because, as discussed above, if the framework considered here is correct, a model in which $\widehat{cay}$ is both reestimated and lagged twice is placed at a great disadvantage. Furthermore, Clark and McCracken (1999) show that equal forecast accuracy tests such as the MSE *F*-test are less powerful than encompassing tests like the ENC-NEW test. Thus the MSE *F*-test may simply lack the power to detect the predictive content of $\widehat{cay}$ when it is both lagged two periods and reestimated.

In summary, the results presented above indicate that $\widehat{cay}_t$ has displayed statistically significant out-of-sample predictive power for excess stock returns over the postwar period, and contains information that is not included in lagged value of the excess return. Moreover, a model of constant expected returns is rejected in favor of a model of time-varying expected returns when $\widehat{cay}$ is used as the predictive variable.

B. Non-nested Forecast Comparisons

The tests above show that adding information on the aggregate consumption–wealth ratio improves the forecasting performance of either a simple first-order autoregressive benchmark for excess stock returns or a constant expected returns benchmark. It is sometimes the case that a researcher wishes to compare two alternative forecasting models of the same variable that are not nested. For example, we could compare a set of competitor forecasts with those generated by the predictor $\widehat{cay}$. We now consider several such competitor forecasts, and test whether they exhibit useful information absent in forecasts using lagged $\widehat{cay}$. As above, each model is first estimated using data from the fourth quarter of 1952 to the first quarter of 1968. We use recursive regressions to reestimate the model, adding one quarter at a time and calculating a series of one-step-ahead forecasts.

Results from a set of nonnested forecast comparisons is given in Table V. Forecasting equations in which the lagged value of $\widehat{cay}$ is the sole predictive variable are alternately compared with "competitor models" in which either the lagged excess return, lagged dividend yield, lagged dividend payout ra-

One-Quarter-Ahead Forecasts of Excess Returns: Nonnested Comparisons

The table reports the results of one-quarter-ahead, nonnested forecast comparisons. The dependent variable is the excess return on the S&P Composite Index. In each case, two models are compared. Model 1 always uses just lagged $\widehat{cay}$ as a predictive variable; Model 2 uses one of several alternate variables labeled in the second column. The column labeled "MSE_1/MSE_2" reports the ratio of the root-mean-squared forecasting error of Model 1 to Model 2. All of the models include a constant. The model denoted $r - r_f$ uses the lagged dependent variable as a predictive variable; the model denoted $d - p$ uses the lagged log dividend yield as a predictive variable; the model denoted $d - e$ uses the lagged log dividend payout ratio as a predictive variable; the model denoted $RREL$ uses the lagged relative bill rate as a predictive variable. The modified Diebold–Mariano test statistic, altered to test for forecast encompassing (Harvey et al. (1998)), appears in the column labeled "MDM Test." The column labeled "95 Percent CV" reports the 95 percent critical value for this statistic based on a t_{n-1} distribution, where n is the number of out-of-sample forecast observations. The null hypothesis is that Model 2 encompasses Model 1. The initial estimation period begins with the fourth quarter of 1954 and ends with the first quarter of 1968. The model is recursively reestimated until the third quarter of 1998. A * (**) denotes significance at the five (one) percent or better level.

| | | | MDM Test | |
| | | | | |
Row	Model 1 vs. Model 2	MSE_1/MSE_2	Test Statistic	95 Percent CV
	Panel A: Cointegrating Vector Reestimated			
1	$\widehat{cay}$ vs. $r - r_f$	0.972	2.14**	1.658
2	$\widehat{cay}$ vs. $d - p$	0.961	2.72**	1.658
3	$\widehat{cay}$ vs. $d - e$	0.982	2.60**	1.658
4	$\widehat{cay}$ vs. $RREL$	0.992	2.39**	1.658
	Panel B: Fixed Cointegrating Vector			
5	$\widehat{cay}$ vs. $r - r_f$	0.908	3.26**	1.658
6	$\widehat{cay}$ vs. $d - p$	0.897	3.58**	1.658
7	$\widehat{cay}$ vs. $d - e$	0.917	3.47**	1.658
8	$\widehat{cay}$ vs. $RREL$	0.927	2.85**	1.658

tio, or lagged stochastic bill rate is the sole predictive variable.[13] As indicated in the table, compared to every one of these competitor models, the model using lagged $\widehat{cay}$ produces a superior forecasting record: The mean-squared error from the series of one-quarter-ahead forecasts is lower when $\widehat{cay}$ is used as a predictive variable than when any of the other predictive variables is employed.

We can once again ask whether these differences in forecasting performance can be statistically discerned. One possibility would be to test the equality of mean-squared forecasting error. As Harvey et al. (1998) point

---

[13] A constant is always included in each of the forecasting equations.

out, however, a more stringent requirement for concluding that a particular model had superior forecasting ability would be that the competing forecasts embody no useful information that is absent in the preferred forecasts; put another way, we may test whether the competing model encompasses the preferred model. Harvey et al. (1998) provide a formal hypothesis testing procedure for the analysis of competing forecasts from nonnested models. They advocate the use of a Diebold and Mariano (1995) test statistic modified to test for forecast encompassing and to account for finite-sample biases. This modified test statistic is referred to as the MDM test statistic. This test statistic is formed by asking whether the difference in forecast errors between two models is correlated with the forecast error of the model that is encompassing under the null. The statistic is compared with critical values from the t_{n-1} distribution, where n is the number of out-of-sample forecasts. The null hypothesis is that the competitor model, without lagged $\widehat{cay}$, (labeled "Model 2" in Table V) encompasses the model where the predictive variable is $\widehat{cay}$. The alternative hypothesis is that the preferred model using $\widehat{cay}$ contains information that could have improved the forecasts of the competitor models.

All of these encompassing tests indicate that the $\widehat{cay}$ forecasting model contains information that produces superior forecasts to those produced by any of the competitor models. The findings are statistically significant at better than the one percent level in almost every case, regardless of whether the cointegrating parameters are reestimated. Boot strap methods (not reported) produced similar results. In summary, the MDM tests suggest that forecasts using the proxy for fluctuations in the consumption–wealth ratio presented here would be consistently superior to forecasts using other popular forecasting variables.[14]

VI. Long-horizon Forecasts

In this section, we investigate the relative predictive power of the variables we have been studying for returns at longer horizons. The theory behind (4) makes clear that the consumption–wealth ratio, like the dividend-

[14] One possible concern about some nonnested encompassing tests is that they fail to account for estimation error in the parameters of the forecasting regressions. However, West and McCracken (1998) show that, for at least some types of encompassing tests, estimation error in the parameters of the forecasting equation is asymptotically negligible as long as the forecasting scheme is recursive (as is the case here) and the forecast errors display no conditional heteroskedasticity or serial correlation. For the case where the forecast errors are conditionally heteroskedastic, such estimation error may not be negligible, and West and McCracken develop an alternative, regression-based test that properly accounts for this error. Results (not reported) using the West–McCracken regression-based tests led to the same conclusions as the those using the MDM test, for all of the eight forecast comparisons made in Table V except one. That exception was one in which the cointegrating parameters in $\widehat{cay}$ were reestimated and a comparison was being made between a model using those reestimated values of $\widehat{cay}$ with the forecasting power from a model using RREL, corresponding to row 4 in Table V. In this case, the West–McCracken test does not reject the null of equal forecast performance.

price ratio, should track longer-term tendencies in asset markets rather than provide accurate short-term forecasts of booms or crashes. Moreover, Figure 1 illustrates that the last 47 years of data have been characterized by several episodes during which deviations from the common trend in consumption, asset wealth, and labor income have persisted for many periods, solidifying the idea that the consumption–wealth ratio should provide a more accurate signal of longer-term trends in asset returns than of short-term movements.

In principle, $\widehat{cay}_t$ could be a long-horizon forecaster of consumption growth, returns, or both. Table VI presents the results of single-equation regressions of either consumption growth or excess returns, over horizons spanning 1 to 24 quarters, on several lagged forecasting variables. The dependent variable in Panel A is the H-period consumption growth rate $\Delta c_{t+1} + \ldots + \Delta c_{t+H}$; in Panel B, the dependent variable is the H-period log excess return on the S&P Composite Index, $r_{t+1} - r_{f, t+1} + \ldots + r_{t+H} - r_{f, t+H}$. For each regression, the table reports the estimated coefficient on the included explanatory variable(s), the adjusted R^2 statistic, and the Newey–West corrected t-statistic for the hypothesis that the coefficient is zero.

Consistent with the results presented in Table II, Panel A of Table VI shows that $\widehat{cay}_t$ has no forecasting power for future consumption growth at any horizon over our postwar sample. The individual coefficient estimates are not statistically significant and the adjusted R^2 statistics are all very close to zero. This result is analogous to the finding that dividend–price ratios have little forecasting power for future dividend growth (e.g., Campbell (1991); Cochrane (1991, 1994, 1997)). Returning to equation (4), this finding, in conjunction with the observation that the consumption–wealth ratio has varied over time, reinforces the notion that fluctuations in the consumption–wealth ratio should forecast asset returns.

Panel B of Table VI reports results from forecasting of the log excess return on the S&P Composite Index. In contrast to long-horizon regressions where future consumption growth is the dependent variable, row 2 of Panel B shows that $\widehat{cay}_t$ has significant forecasting power for future excess returns. The forecasting power of the trend deviation term is particularly strong at short to intermediate horizons. The predictive power of $\widehat{cay}_t$ is humpshaped and peaks around one year; using this single variable alone achieves an $\bar{R}^2$ of 0.21 for excess returns over a five quarter horizon (not reported in the table).

The remaining rows of Panel B give an indication of the forecasting power of other variables for long-horizon excess returns. Row 3 reports long-horizon regressions using the dividend yield as the sole forecasting variable. These results are similar to those obtained elsewhere (e.g., Fama and French (1988), Campbell et al. (1997)) but somewhat weaker because we use more recent data. For example, at a horizon of one year, the dividend yield displays little predictive power for returns, the $\bar{R}^2$ is negligible and the coefficient estimate is not significantly different from zero. The dividend yield only becomes a significant forecaster at a return horizon of six years. For

Table VI

Long-horizon Regressions

The table reports results from long-horizon regressions of excess returns on lagged variables. H denotes the return horizon in quarters. The dependent variable in Panel A is H-period consumption growth $\Delta c_{t+1} + \ldots + \Delta c_{t+H}$. In Panel B, the dependent variable is the sum of H log excess returns on the S&P Composite Index, $r_{t+1} - r_{f,t+1} + \ldots + r_{t+H} - r_{f,t+H}$. The regressors are one-period lagged values of the deviations from trend $\widehat{cay}_t = c_t - \hat{\beta}_a a_t - \hat{\beta}_y y_t$, the log dividend yield $d_t - p_t$, the dividend earnings ratio $d_t - e_t$, the detrended short-term interest rate $RREL_t$, and combinations thereof. For each regression, the table reports OLS estimates of the regressors, Newey–West corrected t-statistics in parentheses, and adjusted R^2 statistics in square brackets. Significant coefficients at the five percent level are highlighted in bold. The sample period is fourth quarter of 1952 to third quarter 1998.

Row	Regressors	Forecast Horizon H							
		1	2	3	4	8	12	16	24
		Panel A: Consumption Growth							
1	$\widehat{cay}_t$	0.11	0.62	1.23	1.98	2.29	0.33	−1.17	0.21
		(0.33)	(0.87)	(1.09)	(1.33)	(1.13)	(0.14)	(−0.41)	(0.05)
		[0.00]	[0.01]	[0.02]	[0.03]	[0.02]	[0.01]	[0.00]	[0.01]
		Panel B: Excess Stock Returns							
2	$\widehat{cay}_t$	**2.16**	**3.80**	**5.43**	**6.72**	**8.35**	**8.57**	**7.86**	**12.44**
		(3.44)	(3.34)	(3.37)	(3.70)	(3.73)	(3.24)	(2.99)	(3.41)
		[0.09]	[0.12]	[0.16]	[0.18]	[0.16]	[0.15]	[0.11]	[0.16]
3	$d_t - p_t$	0.03	0.06	0.10	0.13	0.24	0.27	0.30	**0.76**
		(1.40)	(1.23)	(1.16)	(1.22)	(1.18)	(1.27)	(1.36)	(3.12)
		[0.01]	[0.02]	[0.03]	[0.04]	[0.06]	[0.07]	[0.07]	[0.25]
4	$d_t - p_t$	0.03	0.07	0.10	0.14	0.26	0.31	**0.40**	**0.95**
		(1.02)	(1.12)	(1.10)	(1.16)	(1.14)	(1.34)	(1.98)	(5.27)
	$d_t - e_t$	0.06	0.07	0.09	0.08	0.13	0.20	**0.35**	**0.68**
		(1.71)	(1.12)	(0.94)	(0.70)	(0.73)	(1.00)	(2.30)	(3.51)
		[0.02]	[0.02]	[0.03]	[0.04]	[0.07]	[0.09]	[0.14]	[0.39]
5	$RREL_t$	**−1.68**	**−2.87**	**−3.91**	**−4.51**	−2.52	−1.55	−3.60	−5.10
		(−3.83)	(−3.38)	(−3.20)	(−2.67)	(−1.53)	(−0.76)	(−1.51)	(−1.51)
		[0.06]	[0.07]	[0.10]	[0.10]	[0.01]	[0.00]	[0.03]	[0.03]
6	$\widehat{cay}_t$	**1.89**	**3.13**	**4.43**	**5.37**	**7.24**	**7.69**	5.63	**5.90**
		(3.45)	(3.01)	(3.03)	(3.25)	(3.01)	(2.20)	(1.81)	(1.91)
	$d_t - p_t$	−0.00	0.02	0.03	0.07	0.17	0.24	0.31	**0.85**
		(−0.08)	(0.34)	(0.43)	(0.67)	(0.84)	(1.04)	(1.49)	(4.95)
	$d_t - e_t$	0.02	−0.02	−0.03	−0.05	0.17	0.24	**0.31**	**0.65**
		(0.47)	(−0.29)	(−0.33)	(−0.49)	(0.44)	(1.16)	(2.06)	(2.86)
	$RREL_t$	**−1.17**	**−2.35**	**−3.23**	**−3.82**	−0.27	1.99	0.12	1.36
		(−2.48)	(−3.05)	(−2.91)	(−2.57)	(−0.16)	(0.81)	(0.05)	(0.48)
		[0.11]	[0.15]	[0.21]	[0.23]	[0.18]	[0.19]	[0.18]	[0.42]

such long horizons, the $\bar{R}^2$ of the dividend yield regressions exceeds the $\bar{R}^2$ in the regression using $\widehat{cay}_t$ as the forecasting variable. For all shorter horizons, however, $\widehat{cay}_t$ is a much more powerful predictor of future stock returns than is the dividend yield. Thus, consistent with existing evidence, the dividend yield is a strong forecaster of long-horizon returns, but has little capacity to forecast short- and medium-horizon returns.

Row 4 shows that adding the dividend payout ratio to this equation produces results that are very similar to those using just the dividend yield. Each variable has an important impact on returns over horizons exceeding eight quarters, but not on returns over shorter horizons. The $\bar{R}^2$ statistic suggests that these variables have their greatest predictive power at horizons of three years or more, explaining about 40 percent of the variation in returns at a six-year horizon.

Row 5 reports results using *RREL* as the sole forecasting variable. *RREL* is statistically significant at horizons up to and including one year, although the $\bar{R}^2$ statistics are found to be lower than those obtained in row 2 using $\widehat{cay}_t$ as the predictive variable.

When we include $\widehat{cay}_t$, the stochastically detrended short rate, the dividend yield, and the payout ratio together in one regression, the forecasting results are qualitatively similar to those for one-quarter-ahead returns. At short and intermediate horizons, $\widehat{cay}_t$ continues to have the most forecasting power; the predictive power of *RREL* is also concentrated at short horizons, and $d - p$ and $d - e$ are significant only at very long horizons. By including all four variables, the regression specification now has forecasting power for returns at every horizon we consider, although the total fraction of variation in long-horizon returns that is predicted remains above that of short-horizon returns. These results underscore the finding that $\widehat{cay}_t$ is the best univariate predictor of returns at short-to-intermediate horizons: At a four-quarter horizon, the $\bar{R}^2$ from the regression using just $\widehat{cay}_t$ is almost as large as that in the last panel of Table VI obtained using all four variables.

What factors might account for the relative strengths and weaknesses of $\widehat{cay}_t$ and $d_t - p_t$ at forecasting returns over different horizons? One way to understand these differences is to note that the discount rates in (9) and (10) differ. In (10), $\rho_a \equiv 1/(1 - \exp(\overline{d - p}))$, where $\exp(\overline{d - p})$ is the average ratio of dividends to prices, about 0.99 at a quarterly rate.[15] By contrast, the discount rate, ρ_w, in (9) is slightly smaller, equal to about 0.97 as suggested by using our estimates of ω and sample mean ratios of C/A and C/Y. Accordingly, changes in expected returns in the far future are discounted a bit more in the equation for cay_t than in the equation for $d_t - p_t$. Thus the dividend–price ratio is a better proxy for returns into the distant future than is the trend deviation term. Even abstracting from this difference in discount rates, however, differences in the forecasting power at different horizons may arise if the time series process for expected asset returns is more persistent than that for expected returns to human capital. In this case, $\widehat{cay}$ would be less persistent than the dividend–price ratio, consistent with the evidence in Table II, and would therefore explain a smaller fraction of the variation in expected returns at longer horizons than would the dividend yield.

The single-equation regressions presented in Table VI provide a simple way to summarize the marginal predictive power of each forecasting variable, as well as the overall explanatory power of the forecasting equation. An

---

[15] These estimates come from Campbell, Lo, and MacKinlay (1997).

alternative approach uses VARs to impute long-horizon statistics rather than estimating them directly. One advantage of this approach is that it avoids small-sample biases that may occur in single-equation techniques that can be especially pronounced when the horizon is large relative to the sample size. The methodology for measuring long-horizon statistics by estimating a VAR has been covered by Kandel and Stambaugh (1989), Campbell (1991), and Hodrick (1992), and we refer the reader to those articles for a description of the approach.

Table VII investigates the predictive power of the full VAR counterpart to the equations analyzed previously for long-horizon returns. For each return horizon we consider, we calculate an implied R^2 using the coefficient estimates of the VAR and the estimated covariance matrix of the VAR residuals. Table VII gives the results from estimating two first-order VARs. The first system is a four-variable VAR that includes the excess return on the S&P Composite Index, the relative bill rate, the log dividend–price ratio, and the log dividend–earnings ratio. The second is a five-variable VAR that adds detrended asset wealth to this system. The bottom row of each panel gives the implied R^2 of a regression of long-horizon excess returns on the other variables in the system. The pattern of the implied R^2 statistics is very similar to those from the single-equation regression, indicating that those results are robust to the vector autoregression approach.

VII. Conclusion

The last decade has brought forth an outpouring of research suggesting the existence of time variation in expected asset returns. It is now widely accepted that excess returns are predictable by variables such as dividend–price ratios, earnings–price ratios, dividend–earnings ratios, and an assortment of other financial indicators. For the most part, these financial variables have been successful at predicting long horizon returns, but less successful at predicting returns at shorter horizons.

In this paper, we investigate conditions under which consumption, labor income, and asset wealth are cointegrated, and deviations from this shared trend summarize investors' expectations of future returns on the market portfolio. We show that these deviations from trend primarily forecast future movements in asset wealth, rather than future movements in consumption or labor income. Deviations in the shared trend among consumption, asset wealth, and labor income pick up fluctuations in the aggregate consumption–wealth ratio.

To develop the empirical implications of this framework, we investigate the power of fluctuations in the log consumption–wealth ratio for forecasting asset returns. We find that these fluctuations contain important predictive elements for stock market returns over short and intermediate horizons. Indeed, of the popular forecasting variables explored to date, we find that this variable is the best univariate predictor of stock returns for horizons up to one year. Combining observations on these trend deviations in asset wealth

<div align="center">

Table VII

**Vector Autoregression of Excess Returns
and Implied Long-horizon R^2**

</div>

The table reports coefficient estimates from vector autoregressions (VARs) of returns, relative bill rate, dividend yield, dividend payout ratio, and the trend deviation term. $r_{t+1} - r_{f,t+1}$ is quarterly log excess returns on the S&P Composite Index; $RREL_t$ is the relative bill rate; $d_t - p_t$ is the log dividend yield; $d_t - e_t$ is the log dividend payout ratio; $\widehat{cay}_t$ is $c_t - \hat{\beta}_a a_t - \hat{\beta}_y y_t$, the estimated trend deviation. Newey–West corrected t statistics appear in parentheses below the coefficient estimate. Significant coefficients at the five percent level are highlighted in bold. H denotes the return horizon in quarters. The final column gives the adjusted R^2 for each equation of the VAR estimated over the full sample. The row labeled "Implied R^2" gives the explanatory power of the VAR for the return at horizon H named in the row above and is calculated from the estimated parameters of the VAR and the estimated covariance matrix of VAR residuals. The sample period is fourth quarter of 1952 to third quarter of 1998.

Dependent Variable	Constant (t-stat)	$r_t - r_{f,t}$ (t-stat)	$RREL_t$ (t-stat)	$d_t - p_t$ (t-stat)	$d_t - e_t$ (t-stat)	$\widehat{cay}_t$ (t-stat)	$\bar{R}^2$
			Panel A: Excluding $\widehat{cay}$				
$r_{t+1} - r_{f,t+1}$	0.000	0.078	**−1.309**	0.033	0.025		0.06
	(0.035)	(1.287)	(−2.700)	(1.307)	(0.682)		
$RREL_{t+1}$	−0.000	0.009	**0.684**	−0.004	−0.006		0.52
	(−0.058)	(1.127)	(10.062)	(−1.201)	(−1.839)		
$d_{t+1} - p_{t+1}$	−0.007	−0.102	**1.347**	**0.972**	−0.038		0.94
	(−1.491)	(−1.535)	(2.784)	(40.597)	(−1.080)		
$d_{t+1} - e_{t+1}$	−0.001	−0.142	**−2.499**	0.012	**0.644**		0.50
	(−0.119)	(−1.107)	(−2.191)	(0.356)	(8.273)		
H	2	3	4	8	12	16	24
Implied R^2	0.13	0.17	0.17	0.23	0.28	0.32	0.39
			Panel B: Including $\widehat{cay}$				
$r_{t+1} - r_{f,t+1}$	−0.000	−0.008	**−1.181**	−0.003	0.016	**1.908**	0.11
	(−0.028)	(−0.126)	(−2.600)	(−0.101)	(0.451)	(3.316)	
$RREL_{t+1}$	−0.000	0.012	**0.679**	−0.002	−0.005	−0.072	0.52
	(−0.037)	(1.515)	(9.435)	(−0.620)	(−1.790)	(−0.952)	
$d_{t+1} - p_{t+1}$	−0.007	−0.047	**1.264**	**0.995**	−0.032	**−1.235**	0.94
	(−1.458)	(−0.698)	(2.687)	(39.723)	(−0.913)	(−2.358)	
$d_{t+1} - e_{t+1}$	−0.001	−0.162	**−2.469**	0.004	**0.642**	0.441	0.50
	(−0.127)	(−1.057)	(−2.184)	(0.091)	(8.178)	(0.411)	
$\widehat{cay}_{t+1}$	−0.000	**−0.049**	−0.062	**0.004**	0.001	**0.862**	0.78
	(−0.303)	(−8.235)	(−1.499)	(2.858)	(0.331)	(21.854)	
H	2	3	4	8	12	16	24
Implied R^2	0.22	0.30	0.30	0.35	0.38	0.40	0.44

with those on the log dividend–price ratio and the log dividend–earnings ratio reveals that stock returns exhibit substantial forecastability at horizons ranging from short to long.

We find that the deviation in the shared trend among consumption, labor income, and wealth forecasts excess stock returns over a risk-free rate just as well as it forecasts real stock returns. This feature of the data suggests that expected excess returns, or risk premia, vary over time, a conclusion

that has been drawn previously from evidence that long-horizon returns are predictable by variables such as the dividend–price ratio. Yet the dividend–price ratio provides only indirect evidence that risk premia vary at *cyclical* frequencies, because its forecasting power is concentrated at longer horizons. By contrast, we find that the deviation in wealth from its shared trend with consumption and labor income has strong predictive power for excess stock returns at business cycle frequencies, providing direct evidence that risk premia vary countercyclically.

Although our findings of out-of-sample predictability are particularly strong relative to those of some other studies, we caution that our results do not imply forecastability in all episodes. It is clear, for example, that the last five years have been marked by highly unusual stock market behavior, as prices relative to any sensible divisor have reached unprecedented levels. Similarly over this period, consumption often remained far below its trend relationship with assets and labor earnings even as returns climbed, thereby weakening the tight quarterly link between the consumption–wealth ratio and one-step-ahead returns. Some observers might interpret such an extraordinary episode as a signal that future stock returns will no longer display long-horizon forecastability, but as the framework investigated in this paper makes clear, such a future would require either a constant consumption–wealth ratio or consumption growth that is forecastable by the consumption–wealth ratio, neither of which have been true historically.

An important policy implication of our results is that large swings in financial assets need not be associated with large subsequent movements in consumption. Recently, this issue has become one of pressing importance as fears rise that substantial market swings will cause consumer spending to fluctuate sharply. The model considered in this paper suggests that the real economy may be less vulnerable to transitory movements in asset values than many analysts presume: The model implies that households smooth out transitory variation in their asset wealth and, with consumption currently well below its traditional ratio to asset wealth and labor income, have already factored the expectation of lower returns into today's consumption.

These findings on the time-series behavior of excess returns can be linked to the large literature on cross-sectional asset pricing. As a start, Lettau and Ludvigson (2001) explore the ability of theoretically based asset pricing models such as the CAPM and the consumption CAPM to explain the cross section of average stock returns when fluctuations in the log consumption–wealth ratio are used as a conditioning variable. More generally, the economic variable explored in this paper provides a fresh opportunity to investigate the determinants of asset risk. The predictive power of the consumption-wealth ratio may be obtained under the relatively unrestrictive assumption that the nonstationary, or trend, component of human capital is well captured by labor income itself. The consumption framework investigated here implies that investors' own behavior should reveal expectations of future returns to aggregate wealth, providing a unique proxy of expected returns to the market portfolio.

Table AI

Phillips–Ouliaris Test for Cointegration

The Dickey–Fuller test statistic has been applied to the fitted residuals from the cointegrating regression of consumption on labor income and wealth. Critical values assume trending series. "Lags" refers to the number of lags of first differences used in the regression of residuals on the lagged residual and lags of first differences of the residual.

Dickey–Fuller t-statistic				Critical Values	
Lag = 1	Lag = 2	Lag = 3	Lag = 4	5% Critical Level	10% Critical Level
−4.282	−4.017	−3.800	−3.636	−3.80	−3.52

Appendix: Tests for Cointegration

This appendix provides a description of our macroeconomic data and describes procedures we use to test for cointegration among consumption, labor income, and household wealth. We furnish the output from these tests in Tables AI and AII.

The consumption data are for nondurables and services excluding shoes and clothing, in 1992 chain weighted dollars. The asset holdings data is the household net worth series provided by the Board of Governors of the Federal Reserve. Labor income is defined as wages and salaries plus transfer payments plus other labor income minus personal contributions for social insurance minus taxes. Taxes are defined as (wages and salaries/(wages and salaries + proprietors income with IVA and Ccadj + rental income + personal dividends + personal interest income)) ∗ (personal tax and non-tax payments), where IVA is inventory evaluation and Ccadj is capital consumption adjustments. Both the net worth variable and the labor income variable are deflated by the PCE chain-type price deflator, 1992 = 100. Our source for all of the consumption and income components, as well as the PCE deflator, is the U.S. Department of Commerce, Bureau of Economic Analysis. Note that, although there is some back-casting of the consumption and income data due to revisions and interpolated estimation by the Bureau of Economic Analysis of some service expenditure components, no information about future stock returns is being used to alter any of the data used to construct $\widehat{cay}$.

A representative agent is assumed to consume aggregate consumption, and have income and asset holdings equal to their aggregate values; thus, we use aggregate data in all of our analyses. Use of aggregate data would likely bias downward the forecasting power of cay_t if there is limited participation in asset markets. Unfortunately, it is not possible to determine what part of aggregate consumption, wealth, and income is attributable to stock holders, and household level data is too limited to carry out an extensive forecasting analysis. Heaton and Lucas (2000) present evidence that, for the subset of the population that has significant stock holdings, proprietary income is an important source of earnings. An alternative measure of

<div align="center">

Table AII

Johansen Cointegration Test with Linear Trend in the Data

</div>

A constant is included in the cointegrating relation. The columns labeled "Test Statistic" give the value for the test named in the row above; "90% CV" gives the 90 percent confidence level of that statistic.

L-Max		Trace		$H_0 = r$
Test Statistic	90% CV	Test Statistic	90% CV	$r =$
Panel A: Lag in VAR Model = 1				
19.49	13.39	24.85	26.70	0
5.28	10.60	5.36	13.31	1
0.08	2.71	0.08	2.71	2
Panel B: Lag in VAR Model = 2				
21.03	13.39	25.17	26.70	0
4.05	10.60	4.14	13.31	1
0.10	2.71	0.10	2.71	2
Panel C: Lag in VAR Model = 3				
16.01	13.39	20.08	26.70	0
4.00	10.60	4.08	13.31	1
0.08	2.71	0.08	2.71	2
Panel D: Lag in VAR Model = 4				
14.80	13.39	19.17	26.70	0
4.09	10.60	4.37	13.31	1
0.27	2.71	0.27	2.71	2

labor income that includes non-farm proprietary income compiled by the Bureau of Economic Analysis, in addition to the wage and salary components we use here, produced results that were almost identical to those using the measure described above.

We report the results of two types of cointegration tests: residual based tests designed to distinguish a system without cointegration from a system with at least one cointegrating relationship, and tests for cointegrating rank designed to estimate the number of cointegrating relationships. The former requires that each individual variable pass a unit root test and are conditional on this pretesting procedure. Dickey–Fuller tests for the presence of a unit root in c, y, and a (not reported) are consistent with the hypothesis of a unit root in those series.

Table AI reports test statistics corresponding to the Phillips–Ouliaris (1990) residual-based cointegration tests. This test is designed to distinguish a system without cointegration from a system with at least one cointegrating relationship. The approach applies the augmented Dickey–Fuller unit root test to the residuals from a regression of consumption on labor income and

household wealth. The table shows both the Dickey–Fuller t statistic and the relevant 5 and 10 percent critical values. In the model without a deterministic trend, the hypothesis of no cointegration is rejected at the five percent level by the augmented Dickey–Fuller test with one, two, or three lags, but is not rejected by the test with four lags. We applied the data-dependent procedure suggested in Campbell and Perron (1991) for choosing the appropriate lag length in an augmented Dickey–Fuller test. This procedure suggested that the appropriate lag length was one, implying that test results favoring cointegration should be accepted.

Next we consider testing procedures suggested by Johansen (1988, 1991) that allow the researcher to estimate the number of cointegrating relationships. This procedure presumes a p-dimensional vector autoregressive model with k lags, where p corresponds to the number of stochastic variables among which the investigator wishes to test for cointegration. For our application, $p = 3$. The Johansen procedure provides two tests for cointegration: Under the null hypothesis, H_0, that there are exactly r cointegrating relations, the "Trace" statistic supplies a likelihood ratio test of H_0 against the alternative, H_A, that there are p cointegrating relations, where p is the total number of variables in the model. A second approach uses the "L-max" statistic to test the null hypothesis of r cointegrating relations against the alternative of $r + 1$ cointegrating relations. The test procedure depends on the number of lags assumed in the vector autoregressive structure. Table AI presents the test results obtained under a number of lag assumptions. The same effective sample (1954:1 to 1998:3) was used in estimating the model under each lag assumption.

The critical values obtained using the Johansen approach also depend on the trend characteristics of the data. We present results allowing for linear trends in data, but assuming that the cointegrating relation has only a constant (Table AII) (see Johansen (1988, 1991) for a more detailed discussion of these trend assumptions). In choosing the appropriate trend model for our data, we are guided by both theoretical considerations and statistical criteria. Theoretical considerations imply that the long-run equilibrium relationship between consumption, labor income, and wealth does not have deterministic trends, although each individual data series may have deterministic trends. Moreover, statistical criteria suggest that modeling a trend in the cointegrating relation is not appropriate: The normalized cointegrating equation under this assumption suggests that the parameters of the cointegrating vector are negative, at odds with any sensible model of consumer behavior. The table also reports the 90 percent critical values for these statistics.

The Johansen L-max test results establish strong evidence of a single cointegrating relation among log consumption, log labor income, and the log of household wealth. Table AII shows that, for every lag specification we consider, we may reject the null of no cointegration against the alternative of one cointegrating vector. In addition, we cannot reject the null hypothesis of one cointegrating relationship against the alternative of two or three. Although the evidence in favor of cointegration is somewhat weaker according

to the Trace statistic (we cannot reject the null of no cointegration against the alternative of three cointegrating relations), this evidence is contradicted by the unit root tests, which suggest that each variable contains a unit root. Moreover, according to the Trace statistic, we may not reject the null of one (or two) cointegrating relations against the alternative of three.

REFERENCES

Blinder, Alan S., and Angus Deaton, 1985, The time series consumption function revisited, *Brookings Papers on Economic Activity* 0(2), 465–511.

Bossaerts, Peter, and Pierre Hillion, 1999, Implementing statistical criteria to select return forecasting models: what do we learn? *Review of Financial Studies* 12, 405–428.

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

Campbell, John Y., and Robert Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–227.

Campbell, John Y., and Gregory Mankiw, 1989, Consumption, income and interest rates: Reinterpreting the time series evidence; in Olivier J. Blanchard and Stanley Fischer, eds.: NBER *Macroeconomics Annual* (MIT Press, Cambridge, MA).

Campbell, John Y., 1991, A variance decomposition for stock returns, *Economic Journal* 101, 157–179.

Campbell, John Y., and Pierre Perron, 1991, Pitfalls and opportunities: What macroeconomists should know about unit roots; in Olivier Blanchard and Stanley Fischer, eds. NBER *Macroeconomics Annual* (MIT Press, Cambridge, MA).

Campbell, John Y., 1996, Understanding risk and return, *Journal of Political Economy* 104, 298–345.

Campbell John Y., Andrew Lo, and Craig MacKinlay, 1997, *The Econometrics of Financial Markets* (Princeton University Press, Princeton, NJ).

Campbell, John Y., and John H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.

Clark, Todd and Michael McCracken, 1999, Tests of equal forecast accuracy and encompassing for nested models, Working paper, Federal Reserve Bank of Kansas City.

Cochrane, John H., 1991, Volatility tests and efficient markets: A review essay, *Journal of Monetary Economics* 127, 463–485.

Cochrane, John H., 1994, Permanent and transitory components of GDP and stock prices, *Quarterly Journal of Economics* 109, 241–265.

Cochrane, John H., 1997, Where is the market going? Uncertain facts and novel theories, *Economic Perspectives,* Federal Reserve Bank of Chicago 11, 1–37.

Constantinides, George, 1990, Habit-formation: A resolution of the equity premium puzzle, *Journal of Political Economy* 98, 519–543.

Diebold, Francis, and Roberto Mariano, 1995, Comparing predictive accuracy, *Journal of Business and Economic Statistics* 12, 253–63.

Fama, Eugene, and Kenneth French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3–27.

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

Ferson, Wayne, Sergei Sarkissian, and Timothy Simin, 1999, Spurious regressions in financial economics? Unpublished paper, University of Washington.

Flavin, Marjorie, 1981, The adjustment of consumption to changing expectations about future income, *Journal of Political Economy* 89, 974–1009.

Galí, Jordi, 1990, Finite horizons, life-cycle savings, and time-series evidence on consumption, *Journal of Monetary Economics* 26, 433–452.

Goyal, Amit, and Ivo Welch, 1999, Predicting the equity premium, Unpublished paper, University of California, Los Angeles.

Hansen, Gary D., 1985, Indivisible labor and the business cycle, *Journal of Monetary Economics* 16, 309–327.

Harvey, David, Stephen Leybourne, and Paul Newbold, 1998, Tests for forecast encompassing, *Journal of Business and Economic Statistics* 16, 254–259.

Heaton, John, and Deborah Lucas, 2000, Portfolio choice and asset prices: The importance of entrepreneurial risk, *Journal of Finance*, 55, 1163–1198.

Hodrick, Robert, 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357–386.

Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3–54.

Johansen, Soren, 1988, Statistical analysis of cointegrating vectors, *Journal of Economic Dynamics and Control* 12, 231–254.

Johansen, Soren, 1991, Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica* 56, 1551–1580.

Kandel, Shmuel, and Robert Stambaugh, 1989, Modeling expected stock returns for long and short horizons, Working paper 42-88, Rodney L. White Center, Wharton School, University of Pennsylvania.

Kydland, Finn, and Edward C. Prescott, 1982, Time to build and aggregate fluctuations, *Econometrica* 50, 343–360.

Lamont, Owen, 1998, Earnings and expected returns, *Journal of Finance* 53, 1563–1587.

Lettau, Martin, and Sydney Ludvigson, 2001, Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying, *Journal of Political Economy*, forthcoming.

Ludvigson, Sydney, and Charles Steindel, 1999, How important is the stock market effect on consumption? *Economic Policy Review,* Federal Reserve Bank of New York 5, 29–51.

McCracken, Michael, 1999, Asymptotics for out-of-sample tests of causality, Unpublished manuscript, Louisiana State University.

Nelson, Charles C., and Myung J. Kim, 1993, Predictable stock returns: The role of small sample bias, *Journal of Finance* 43, 641–661.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Phillips, Peter, and Sam Ouliaris, 1990, Asymptotic properties of residual based tests for cointegration, *Econometrica* 58, 165–193.

Shiller, Robert, 1984, Stock prices and social dynamics, *Brookings Papers on Economic Activity* 2, 457–498.

Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.

Stock, James H., 1987, Asymptotic properties of least squares estimators of cointegrating vectors, *Econometrica* 55, 113–144.

Stock, James H., and Mark Watson, 1993, A simple estimator of cointegrating vectors in higher order integrated systems, *Econometrica* 6, 783–820.

Sundaresan, Suresh, 1989, Intertemporally dependent preferences and the volatility of consumption and wealth, *Review of Financial Studies* 2, 73–89.

West, Kenneth D., and Michael W. McCracken, 1998, Regression-based tests of predictive ability, *International Economic Review* 39, 817–840.