tay's as good as *cay*: Reply[1]

Martin Lettau Sydney C. Ludvigson

Published in
Finance Research Letters,
March, 2(1), 2005

February 18, 2005

[1] Lettau: Department of Finance, Stern School of Business, New York University; Ludvigson: Department of Economics, New York University. We thank John Y. Campbell, John Cochrane, Lars Hansen and Jessica Wachter for helpful comments.

In a recent comment on our published work (Lettau and Ludvigson 2001), Michael Brennan and Yihong Xia (2002) advance the following argument: A “mechanistic” variable *tay*, where *t* is a linear time trend, forecasts stock returns. Since “*t* has no foresight,” the argument goes, the predictive power of this variable must be attributable to what they call “look-ahead bias.” The authors assert that *cay* is subject to the same look-ahead bias (generated because we use the full sample to estimate the cointegrating parameters in *cay*), implying that its forecasting power must be spurious.

In this response, we explain why this critique is misplaced. Our disagreement is based on the following observations. First, cointegration theory implies that the forecasting power of *cay* for future returns cannot be spurious (attributable to any kind of “bias”) merely because the full sample has been used to estimate the cointegrating coefficients. Instead, precisely the opposite is true: the predictive power of a cointegrating residual can be evaluated only if all available information in the sample is used in estimating the cointegrating coefficients. Second, *tay* likely contains more economic content that the authors realize, and its forecasting power is perfectly consistent with the interpretation of *cay* provided in our *Journal of Finance* article. Finally, the authors assert that out-of-sample tests better support their arguments than in sample tests. This ignores recent theoretical work, which finds that in-sample tests are more credible than out-of-sample tests for assessing forecasting power.

I. Is the Predictive Power of *cay* Attributable to “Look-ahead Bias”?

We begin with the authors’ main claim, namely that the forecasting power of *cay* for future returns is spurious, a feature of “look-ahead bias,” which Brennan and Xia argue stems from our use of the full sample to estimate the cointegrating coefficients in *cay*. Is it plausible that the forecasting power we have uncovered is spurious and entirely attributable to our use of the full sample to estimate of the cointegrating coefficients in *cay*?

The answer to this question is “no,” for two reasons. First, there is direct evidence that the use of all available data to estimate the cointegrating residuals in *cay* has nothing to do with its strong forecasting power. In Lettau and Ludvigson (2004a) we show how the predictability of *cay* can be indirectly assessed without any prior estimation of cointegrating parameters. The resulting forecastability of returns is just as strong using this technique as it is using the technique employed by Lettau and Ludvigson (2001). Here we follow the approach in Lettau and Ludvigson (2004a) to directly investigate the alleged look-ahead bias (please see that paper for details). Table 1 reports output from regressions of excess stock returns on log consumption, log-asset wealth and log labor income using the data from our original *Journal of Finance* paper. Univariate forecasting regressions using *cay* can be restated in terms of the multivariate regression of Table 1. Since this latter regression does not include the cointegration residual, no future data is used and there is no scope for a look-ahead bias. The R-squared statistics are almost identical to those from regressions of excess returns on the cointegrating residual reported in our original *Journal of Finance* paper. This demonstrates that the forecasting power of *cay* cannot be attributable to a look-ahead bias.

Second, econometric theory also tells us that the forecasting power we have uncovered cannot be attributable to our use of the full sample to estimate of the cointegrating coefficients in *cay*. If a set of variables are thought to be cointegrated over a particular sample, econometric theory dictates that the full sample be used to estimate the cointegrating coefficients. If instead a researcher uses subsamples or rolling estimates of these coefficients, the resulting estimated cointegrating residual is likely to be riddled with sampling error, and therefore unlikely to produce reliable estimates of the true cointegrating residual. It follows that the forecasting power of the latter cannot be assessed using subsamples or rolling estimates of the cointegrating coefficients. By contrast, it is well known that once a sufficiently large sample is available to undertake estimation, the cointegrating parameters are superconsistent and may be treated as known in subsequent estimation, i.e., in estimation of the error-correction representation or of some other forecasting regression.[2] It is therefore a direct implication of cointegration theory there can be no bias in a cointegrating residual merely because the cointegrating parameters have been estimated using all the available data in the sample over which the variables are believed to be cointegrated. On the contrary, using only part of the sample eliminates relevant information that is required to accurately measure the cointegrating residual, and to test its predictive power in subsequent estimation.

Nor can the budget constraint interpretation of *cay* be evaluated unless all available data is used to estimate the cointegrating parameters. Under that interpretation, the parameters in *cay* are steady state wealth shares, which, if the theory is true, are clearly known to the agent in equilibrium. Only the use of all available data to estimate the parameters in *cay* comes closest to delivering superconsistent estimates and therefore revealing the true parameters that would have been known to the representative consumer when making investment decisions. Out of sample tests are a problem from this perspective because they eliminate information necessary to estimate cointegrating coefficients and *cay* accurately.

We repeat this point for emphasis:

Any forecasting procedure that does not use the full sample to estimate the parameters of the common trend in cay is an inappropriate test of the equilibrium model, because it necessarily eliminates a large part of the sample which is required to uncover the hypothetical wealth shares that would have been known to the representative investor if the model were true.

Brennan and Xia argue that the Granger representation theorem (GRT hereafter) does not provide a rationale for the predictive power of *cay*. It does. Moreover, neither the GRT nor our budget constraint interpretation implies that *cay* should forecast real interest rates, as Brennan and Xia claim. The GRT says *cay* must forecast something: either consumption growth, labor income growth, or some component of asset wealth growth (driven by asset returns). That is, the GRT say that the vector **B** in Brennan and Xia's equation (5) cannot be a zero vector.[3] Nevertheless, the GRT does *not* say that *cay* must forecast all of these variables, or all components of wealth growth, such as real interest rates, bond returns or other non-stock market financial returns. The cointegrating residual will forecast only those parts of asset wealth that have significant transitory components. The preponderance of evidence suggests that it is the stock market component of asset wealth that has a large transitory component and is therefore forecastable (Lettau and Ludvigson 2004b). Thus, since *cay* must forecast something, the finding that it does not forecast real interest rates, other non-stock market returns, consumption growth, or labor income growth merely serves to strengthen our argument that it forecasts stock returns.

This analysis illustrates the problem with Brennan and Xia's reasoning: Brennan and Xia agree that, (1) *c*, *a*, and *y* are cointegrated, a testable implication that is supported by the data. They also assert (2) that the forecasting power of *cay* for the growth in stock market wealth is spurious, attributable to look-ahead bias. Surely, given their emphasis on the superiority of out-of-sample prediction, Brennan and Xia do not mean to argue that *cay* forecasts the growth in *c*, *y* or the non-stock market component of *a* (such as short term interest rates), since the future values of those variables do not even display any *in-sample* relation with *cay*. Consequently it seems clear that

---

[2] Cointegrating parameter estimates converge at a rate proportional to the sample size *T*.

[3] The marginal significance of *cay* in Brennan and Xia's equation (6) for wealth growth is attributable to the fact that they do not correctly implement the error-correction representation in (5) for wealth growth. When correctly implemented, *cay* has strong one-quarter ahead forecasting power for Δa_t in the error correction representation (Lettau and Ludvigson 2001a, 2004b).

Brennan and Xia implicitly agree with, (3) *cay* has no forecasting power for the growth in *c*, *y*, or non-stock market returns. But if both (1) and (3) are true, the only remaining possibility is that (2) is false, implying that *cay* forecasts the growth in the stock market component of *a*, consistent with the empirical evidence we present.

When we wrote of "look-ahead bias" in our original paper, we were implicitly concerned with the problem of a *practitioner* seeking to make forecasts. For all the reasons given above, "look-ahead bias" turns out to have been an unfortunate word choice for describing this situation, since we have seen that there can be no "bias" in a forecast that uses a cointegration residual as a predictive variable merely because all available information has been used to estimate the cointegrating parameters. Instead, bias arises if information is ignored. The use of the full sample only becomes an issue if one seeks to assess whether a practitioner, operating in real time in the earlier parts of our sample, and who had no knowledge of the true cointegrating parameters or wealth shares, could have detected predictability using a *poor estimate* of *cay*. Without the benefit of the full sample, such a practitioner may well have trouble obtaining an accurate estimate of *cay* and successfully forecasting any variable. (But notice that this would not apply to a practitioner today, since the required data is now available.) This is because out-of-sample tests in which the parameters of *cay* are recursively reestimated over short subsamples of the data necessarily throw away information that is required to form accurate estimates of *cay*. As a consequence, it is a *prediction of the model* that forecasting tests in which the parameters of *cay* are estimated by throwing away relevant information will never display as much predictive power as those in which the parameters in *cay* are estimated superconsistently and, in effect, set at their theoretically correct values.[4] This should not be viewed as an indictment of the ability of *cay* to forecast, but rather as an indictment of the ability of a practitioner to properly measure *cay* and exploit its genuine forecasting power. After all, when the parameters in *cay* are accurately estimated using the full sample, *cay* displays much stronger out-of-sample forecasting power than when it is poorly measured over short subsamples of the data. Far from refuting our model, this evidence actually provides confirmation of it. It is a truism that the output of any analysis is only as good as the data entered as an input; if incorrect data are entered, the result is a wrong answer. This applies here: throwing away useful information produces bad estimates of *cay*, which results in inferior forecasts of returns.

II. *tay* Contains More Economic Content Than You Think

Brennan and Xia claim that the forecasting power of *cay* is spurious because an alternative variable, *tay*, which replaces consumption with calendar time, also forecasts returns. Their reasoning appears to be the following. Since calendar time is devoid of economic content, *tay* is devoid of economic content, implying that its forecasting power must be spurious. Moreover, since we have merely replaced consumption with an "inanimate" variable, calendar time, which has no intrinsic foresight, we may conclude that the predictive power of *cay* must also be spurious.

We find this reasoning difficult to follow. First, if *tay* really were a mechanistic variable devoid of economic content, it is unclear how that fact alone would convey any information about the true forecasting power of *cay*, which, by contrast, follows from a budget-constraint identity that must be a part of any sensible economic model. But there are good reasons also to question the premise of the argument. Does merely replacing *c* with *t* really make *tay* an inanimate variable devoid of economic content? Put differently, if the framework we study were true, would it be surprising to find that *tay* forecasts returns?

[4] Our own Monte Carlo analysis suggests that just over 40 years of data are required for the parameters in *cay* to converge to their true values.

The answer to both of these questions is again "no." Although statistical evidence suggests that consumption contains a stochastic trend, like many macroeconomic time-series, a large fraction of its variability is governed by a simple deterministic time trend. This observation is especially true for aggregate consumption, whose time-series variation appears to be dominated by permanent shocks and well described by a combination of deterministic and stochastic trends. Accordingly, calendar time, that is, a deterministic time trend, is a proxy for consumption, and *tay* is a proxy for *cay*. If returns are predictable (for any reason), then by virtue of an economic identity, *cay* is a good candidate to pick up the predictable component. If *cay* forecasts returns, then any variable *xay*, where *x* proxies for consumption, is also likely to forecast returns.

Brennan and Xia argue that using consumption in *cay* is akin to "trend fitting," without ever asking why that might be so. An immediate theoretical explanation is that *tay* contains a lot more economic content than the authors realize, and the use of *t* in place of *c* does not demonstrate a lack of foresight on the part of consumers. On the contrary, it is consistent with extraordinary foresight: for forward-looking, permanent income investors who want to consume only the permanent components of wealth and income, consumption *should* define the trend in these variables. Only those investors with little ability to anticipate future returns would have consumption that is highly correlated with the transitory variation in their wealth and poorly described by a trending series. This does not require that investors have perfect foresight, only that they do not make systematic forecasting errors. Such permanent income logic was first emphasized by Cochrane (1994), who used the same reasoning to argue that consumption should define the trend in GDP.

In summary, the forecasting ability of *tay* is perfectly consistent with forward looking behavior and optimization. Whereas calendar time *per se* has no foresight, the fact that consumption is well described by a trending process is consistent with tremendous foresight on the part of consumers. Moreover, the replacing *c* with *t* is not informative about why *cay* has such strong forecasting power for returns. If consumption and income were not close to trending variables, deviations of asset wealth from its common trend with consumption and labor income would not uncover the transitory component of wealth, and *cay* would not forecast returns.

III. In-Sample versus Out-of-Sample Forecasting

Brennan and Xia claim that out-of-sample tests directly address the question of whether the in-sample performance is due to look-ahead bias, and the concern that the good in-sample results are spurious. There appears to be no econometric basis for these claims. On the contrary, Inoue and Kilian (2002) demonstrate that in-sample and out-of-sample tests of predictability are asymptotically equally reliable under the null of no predictability.[5] They conclude that with or without data mining, "the conventional wisdom that in-sample tests are biased in favor of detecting spurious predictability cannot be supported by theory." Given that in-sample tests display no greater size distortions than do out-of-sample tests, the choice between in-sample and out-of-sample prediction is reduced to the question of which test is more powerful. Inoue and Kilian address this question by considering a sequence of local alternatives, and a variety of out-of-sample procedures. They show that for most local alternatives and out-of-sample design choices, in-sample tests are more powerful than out-of-sample tests, even asymptotically. (It is known that they are more powerful in small samples.) In addition, they find that the one-sided *t*-test, commonly employed in asset pricing applications and also relevant for our application, is uniformly more powerful than all out-of-sample tests they consider. Often the power of out-of-sample tests is only half that of the in-sample, one-sided *t*-test. As Inoue and Kilian point out, these results dispel the notion that out-of-sample tests are more

[5] A test is defined to be unreliable if its effective size exceeds its nominal size.

convincing than in-sample tests, and they conclude that in-sample tests of predictability will "typically be more credible than results of out-of-sample tests." It is important to understand the implications of these results: the low power of out-of-sample tests means that they are unlikely to reveal true forecasting power, *even* the forecasting power a practitioner might have had in real time!

We close with two additional points. First, Brennan and Xia report that *cay* has no out-of-sample forecasting power for returns. These findings are contradicted by our own evidence (Lettau and Ludvigson, 2001a) and that more recently of Guo (2004). Second, cointegration tests do not reject the hypothesis that the regressor *tay* is nonstationary and contains a unit root. This necessarily implies that *tay* is highly persistent, and indeed it has a first-order autocorrelation coefficient of 0.95. By comparison, cointegration tests for *cay* reject the null that *cay* contains a unit root, and the variable is correspondingly less persistent than *tay*, having an autocorrelation of 0.83. This extreme degree of persistence of *tay* poses a problem for predictive regressions. There is a growing body of both theoretical and empirical work, which shows that coefficient estimates, *t*-ratios, and *R*-squared statistics can all be severely biased upward in predictive regressions where the dependent variable is so highly persistent that cointegration tests suggest that the variable may contain a unit root. For example, simulation evidence presented in Ferson et. al., (2003) shows that regressors with autocorrelation coefficients as large as that of *tay* suffer from a serious spurious regression problem if the true process for expected returns is also persistent. Upward biases in both the *R*-squared statistics and *t*-ratios can be dramatic for regressors that have an autocorrelation coefficient as high as 0.95. By contrast, Ferson et. al., find that the *t*-statistics and *R*-squared statistics are well behaved for regressors with autocorrelation coefficients on the order of 0.83, as is the case for *cay*. Thus, the spurious regression concern to which *tay* is subject is not applicable to *cay*. Apart from making the statistical results using *tay* questionable, it makes no sense to forecast returns with a unit root variable. We conclude that the forecasting findings based on *cay* are statistically more reliable than those using its proxy, *tay*.

References

Brennan, Michael and Yihong Xia (2002), "*tay*'s as good as *cay*," *Finance Research Letters* 2(1), 1-15.

Ferson, Wayne, Sergei Sarkissian and Timothy Simons (2003), "Spurious Regressions in Financial Economics?" *Journal of Finance* 58, 1393-1414.

Guo, Hui, (2003), "On the Out-of-Sample Predictability of Stock Market Returns," *Journal of Business*, forthcoming.

Inoue, Atsushi and Lutz Kilian (2002), "In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?" Unpublished paper, Department of Economics, University of Michigan.

Lettau, Martin and Sydney Ludvigson (2001), "Consumption, Aggregate Wealth, and Expected Stock Returns," *Journal of Finance*, 56(3): 815-850.

Lettau, Martin and Sydney Ludvigson (2004a), "Expected Returns and Expected Dividend Growth," *Journal of Financial Economics*, forthcoming.

Lettau, Martin and Sydney Ludvigson (2004b), "Understanding Trend and Cycle in Asset Values: Bulls, Bears, and the Wealth Effect on Consumption," *American Economic Review*, 94 (1), 279-299.

Newey, Whitney K. and Kenneth D. West (1987), "A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica*, 55, 703-708.

Table 1: Multivariate Long-horizon Regressions

Variables	Horizon h (in years) 1	2	4	8	12	24
$\sum_{i=1}^{h}(r_{t+i} - r_{f,t+i}) = k + \beta_c(h)c_t + \beta_a(h)a_t + \beta_y(h)y_t + \epsilon_{t,t+h}$						
c_t	**2.23**	**4.18**	**7.25**	**9.10**	**9.05**	**13.80**
	(4.45)	(5.29)	(4.63)	(4.30)	(4.54)	(4.22)
a_t	**-0.61**	**-1.05**	**-1.86**	**-2.15**	**-1.39**	**-2.85**
	(-3.31)	(-4.21)	(-3.55)	(-3.17)	(-2.01)	(-2.47)
y_t	**-1.53**	**-2.68**	**-4.65**	**-6.02**	**- 6.63**	**-9.04**
	(-4.61)	(-5.21)	(-4.59)	(-4.32)	(-4.94)	(-4.60)
	[0.09]	[0.13]	[0.20]	[0.19]	[0.21]	[0.20]

Notes: This tables reports results from h-period regression of CRSP-VW returns in excess of a 3-month Treasury-bill rate on c_t, y_t and a_t: $\sum_{i=1}^{h}(r_{t+i} - r_{f,t+i}) = k + \beta_c(h)c_t + \beta_a(h)a_t + \beta_y(h)y_t + \epsilon_{t,t+h}$. For each regression, the table reports OLS coefficient estimates, t-statistics (in parentheses), and adjusted R^2 statistics (in square brackets). The t-statistics are Newey-West (1987) corrected and take into account that c_t, y_t and a_t are cointegrated. Coefficients that are significant at the 5% level are highlighted in bold face. Under the maintained hypothesis of cointegration, if there is a relation between the (stationary) left-hand side variable to be forecast and some stationary linear combination of the regressors c_t, a_t, and y_t, this regression can freely estimate the non-zero coefficients $\beta_c(h), \beta_a(h)$ and $\beta_y(h)$ that generate such a relation. This forecasting regression is nonspurious and the limiting distributions of the coefficient estimates are normal, implying that the regression will produce valid R^2 and t-statistics. Because this procedure does not require any first-stage estimation of cointegration parameters, it is clear that the forecasting outcome, in particular its coefficients and R^2 statistics, cannot be influenced by such a prior analysis. Please see Lettau and Ludvigson (2004a) for details.

All data are quarterly in real, per capita terms. Consumption, c_t, is measured as expenditures on nondurables and services, excluding shoes and clothing. Labor income, y_t, is measured as wages and salaries + transfer payments + other labor income - personal contributions for social insurance - Taxes. Taxes are defined as [wages and salaries/(wages and salaries + proprietors' income with IVA and Ccadj + rental income + personal dividends + personal interest income)] times personal tax and nontax payments, where IVA is inventory valuation and Ccadj is capital consumption adjustments. Our source for these variables is the Bureau of Economic Analysis. Asset wealth, a_t, is total household net worth in billions of dollars, measured at the beginning of the period. Our source for asset wealth is the Board of Governors of the Federal Reserve System. Please see Lettau and Ludvigson (2001) for a more detailed description of the data. The sample is quarterly and spans the period 1952Q4 to 1998Q3.