Predicting Local Returns with Macroeconomic Variables

George M Korniotis* †

University of Notre Dame

Department of Finance

April 25, 2005

Abstract

Predictability of the return on the market portfolio is a well established fact. This study shows that predictability is a more general phenomenon and it extends to return indices of the U.S. states. At the state level, the consumption trend deviation of Lettau and Ludvigson, and the collateral ratio of Lustig and Van Nieuwerburgh can predict short-term and long-term state-level returns. The state-level unemployment rate is also a significant predictor, as suggested by the incomplete markets model of Constantinides and Duffie.

JEL codes: C34, E21, G12

*The paper was benefited by comments by Elena Andreou, Andros Kourtelllos, Alok Kumar and the seminar participants at the Department of Economics at the University of Cyprus. I am responsible for all remaining errors.

†E-mail: Korniotis.1@nd.edu. Mail: Department of Finance, 102 Mendoza College of Business, Notre Dame, IN 46556. Telephone: (574) 631-9322.
This paper investigates stock market predictability at the U.S. state level. As Cochrane (1999) points out, predictability of the aggregate stock market return is one of the new facts in finance. This aggregate return is the return on a market portfolio. A natural question is whether predictability of returns extends from the market portfolio to other less aggregated portfolios. If predictability is a robust new fact in finance, it should emerge in other portfolios that include a subset of the assets from the market portfolio. This is the goal of the current paper: investigate if predictability is present using returns of state-level portfolios. This is the first paper to test for predictability at the state level.

In order to predict state-level returns, I use macroeconomic variables. Recent papers argue that using financial ratios, like the dividend-price ratio, to predict returns poses many econometric challenges. Financial ratios are endogenous because they include the stock price, which also affects returns. The endogeneity problem is mitigated when the researcher exploits our understanding of the macro economy to find predictors, which do not include the stock price. In particular, I use the cay of Lettau and Ludvigson (2001), the hy of Lustig and Van Nieuwerburgh (2004), and the state-level unemployment rate. The importance of these three predictors can be explained by consumption-based asset pricing models.

Lettau and Ludvigson (2001) show that the deviation of U.S. consumption from its long-run trend with U.S. wealth can predict the market return. This deviation called cay captures the expectations of individuals about the future path of returns. A positive cay implies that consumption today is above its long-run level. Then, it has to be the case that on average people expect the returns on their wealth to be higher in the future.

Lustig and Van Nieuwerburgh (2004) identify the collateral ratio hy, the log ratio of home wealth to income, as another macroeconomic predictor of the aggregate stock return. As hy decreases, the collateral becomes relatively more scarce, the consumption growth of an individual becomes more perceptible to idiosyncratic income shocks, the variance of consumption growth increases, and individuals ask a higher return at the market level. The significance of the collateral mechanism is stronger in an incomplete market setting. Also, the collateral ratio can be interpreted as a measure of financial distress.

The work of Campbell and Yogo (2004), Hansen and Tuypens (2004), Hjalmarsson (2004) and Mark and Sul (2004) are excellent discussions of the difficulties in uncovering stock-market predictability with aggregate U.S. data using financial ratios that include the stock price.
A novel feature of the study is to recognize that there is no full risk-sharing among the U.S. states, which indicates that the state-level unemployment rate can be a useful predictor. Asdrubali, Sorensen and Yosha (1996) show that only 75% of the shocks on gross state product are smoothed across the states. Similarly, Athanasoulis and Van Wincoop (2001) find that financial markets and federal fiscal policy reduce the uncertainty of state income only by 44%. Markets are therefore incomplete across the U.S. states. Constantinides and Duffie (1996) show that in an incomplete markets setting, where idiosyncratic shocks are not diversifiable, agents ask for a hefty premium to hold stocks. People fear stocks because they tend to do badly during recessions, which are periods with high unemployment. Korniotis (2005) shows that an asset pricing model which includes incomplete markets at the state level can rationalize a higher equity premium, than the complete markets model.

The local return of state $i$ includes the stocks all the firms with their headquarters in state $i$ for the period from 1982(Q1) to 1999(Q1). The local index of state $i$ is affected by the conditions in the national stock market; its correlation with the market return is 0.82.\footnote{This number is reported in Table 1, Panel A.} To predict this component of local returns I use the $cay$ of Lettau and Ludvigson (2001). The local return should also be affected by the local macroeconomic conditions in state $i$. This component of local returns can be captured by the state-level collateral ratio and by the state-level unemployment rate. The prediction regressions also include a state-level dividend-price ratio since it is one of the most successful financial predictors.

The empirical analysis estimates one-quarter ahead forecasting regressions for the local return $r$, the excess of the local return over the Treasury Bill return, $r - r_f$, and the idiosyncratic component of the local return. The last one is defined as the excess of the local return over the market return, $r - r_m$. As in Fama and French (1988), regressions on long-term returns are also considered. Two types of such regressions are estimated: one on the long-term return over the Treasury Bill return, $r_K - r_{Kf}$, and one on the long-term idiosyncratic component of the state return, $r_K - r_{Km}$.

The paper demonstrates that state-level returns are predictable. In particular, the one-quarter-ahead forecasting regressions reveal that the dividend-price ratio becomes insignificant in the presence of the $cay$ residual and the unemployment rate. The trend deviation $cay$ predicts the market component of the local return since it is significant in the regressions for $r$ and $r - r_f$, but it is insignificant in the regression
for $r - r_m$. The collateral ratio is also an important predictor. It is most significant in predicting the idiosyncratic component of the local returns, because it can capture financial distress at the state level. Finally, the state unemployment rate is statistically significant in all three regressions.

The long-run regressions reveal that state returns are the most predictable at the 3 year horizon. Similar to the one-quarter-ahead forecasting regressions, the dividend-price ratio is insignificant because most of its information is captured by the residual $cay$ and the state unemployment rate. It is found that the trend deviation $cay$ can only predict $r_K - r_{Kf}$, and the unemployment rate is an important predictor only for $r_K - r_{Km}$. The collateral ratio though is statistically significant for both the $r_K - r_{Kf}$ and $r_K - r_{Km}$ returns.

The rest of the paper is organized as follows. Section 1 includes more details on the predicting variables. Section 2 describes the data set. Section 3 presents the one-quarter-ahead forecasting regressions. Section 4 includes the results for the regressions with long-term returns. Finally, Section 5 concludes the discussion.

1. The Forecasting Variables

The main goal of the paper is to investigate predictability at the state level using macroeconomic variables. Nevertheless, the dividend-price ratio is included in the analysis because it is one of the most successful financial ratios in predicting returns. Its presence in the forecasting regression will reveal if the macroeconomic variables are good predictors over and above the predictability that might come from the dividend-price ratio.

The dividend-price ratio

In U.S. data one of the most successful variables in forecasting long-horizon returns is the log dividend-price ratio introduced by Campbell and Shiller (1988). Their analysis uses the definition of the stock return $r_{t+1}$,

$$r_{t+1} \equiv \log (P_{t+1} + D_{t+1}) - \log (P_t),$$

where $P$ and $D$ denote the price and the dividend of the stock. By log-linearizing $r_{t+1}$, they derive the
following accounting identity between the log dividend-price ratio and future returns and dividends:

\[ d_t - p_t = \text{constant} + \mathbb{E}_t \left\{ \sum_{j=1}^{\infty} \rho^j (-\Delta d_{t+1+j} + r_{t+1+j}) \right\}, \]

where \( d = \log D, \ p = \log P, \ \rho \) is the discount rate evaluated at the steady state and it equals \( 1/(1 + \exp(d - p)) \), and \( \Delta d \) is the growth of dividends. Given the dividend growth, if future returns are expected to grow, then the dividend-price ratio should increase today. Therefore, there should be a positive relation between \( d_t - p_t \) and future returns \( r_{t+1+j} \). The dividend-price ratio suffers from endogeneity problems because it includes the log of the stock prices, \( \log (P_t) \), which is also part of the definition of the log return \( r_{t+1} \).

The dividend-price ratio is stationary, even though it is highly persistent, and it can forecast long-horizon returns very well. As Cochrane (2005) notes, its importance builds with horizon as it captures decade-to-decade movements as well as business-cycle movements. Therefore, in long-horizon forecasting regressions, the \( R^2 \) of the regression and the significance of the dividend-price ratio increase with the forecasting horizons.3

The consumption trend deviation cay

Lettau and Ludvigson (2001) obtain the deviation of consumption from its long-run trend with total wealth using the budget constraint of a representative agent:

\[ W_{t+1} = (1 + R_{w,t+1})(W_t - C_t), \]

where \( W \) represents total wealth including both financial and human wealth, \( R_w \) is the return on total wealth, and \( C \) is consumption. They log-linearize this budget constraint around the steady state and solve it forward. Then, by taking expectations conditional on information known at period \( t \), they show that the consumption-wealth ratio has the following form:

\[ c_t - w_t = \mathbb{E}_t \sum_{j=1}^{\infty} (\rho_w)^j [r_{w,t+j} - \Delta c_{t+j}], \]

where $\rho_w$ is the discount rate evaluated at the steady state and it equals $(W - C)/W$, $c = \log C$, $w = \log W$ and $r = (1 + R)$. This relation dictates that consumption and wealth are cointegrated since their difference, $c - w$, is equal to a discounted sum of stationary variables, namely the returns on wealth $r_w$ and consumption growth $\Delta c_t$.

Human wealth is unobservable and it is approximated with labor income. Also, total wealth $W$ is broken down to financial assets $A$ and labor income $Y$\footnote{The financial assets $A$, and labor income $Y$ in $w = \log(A + Y)$, are separated via a log-linear approximation around the steady state. Therefore, the consumption deviation $(c - w)$ is approximately equal to $[c - \varpi a - (1 - \varpi) y]$, where $\varpi = A/(A + Y)$ evaluated at the steady state.}. Under reasonable assumptions Lettau and Ludvigson (2001) show that the deviation of log-consumption $c$ from log-financial assets $a$ and log-labor income $y$, called $cay$, is also stationary. As with the difference $c - w$, the $cay$ also contains expectations of future returns and future consumption growth. Lettau and Ludvigson (2001) demonstrate that the $cay$ is a good predictor of asset growth because it can forecast asset returns. Most importantly, they find that in quarterly regressions the $cay$ at quarter $q$ can predict the market and the excess returns at quarter $q + 1$ very well. Not only that, the $cay$ can predict market returns at longer horizons.

**The collateral ratio $hy$**

Lustig and Van Nieuwerburgh (2004) build a consumption asset pricing model with liquidity constraints and collateralized borrowing, where the housing wealth is the collateral asset. In this model the ratio of housing to human wealth, called collateral ratio $hy$, changes the conditional distribution of consumption growth. As the collateral ratio decreases, it becomes more difficult to use housing as collateral to borrow and insulate consumption from negative labor income shocks. Therefore, the probability that the liquidity constraints might bind increases, forcing the variance of consumption growth to increase. In other words, “when the collateral ratio is low, the dispersion of consumption growth across households is more sensitive to aggregate consumption growth shocks, and this raises the market price of aggregate risk” (Lustig and Van Nieuwerburgh, 2004).

The model produces a market price of aggregate risk, which varies with the housing market. Thus, by conditioning on the housing collateral ratio, one can predict the conditional volatility of consumption and predict market returns. Indeed, Lustig and Van Nieuwerburgh (2004) show that the collateral ratio can forecast the return on the U.S. stock market, mainly at lower frequencies.
Incomplete Markets and Unemployment

Constantinides and Duffie (1996) recognize that stocks perform poorly in economic recessions. But recessions are time-periods where the investor is likely to become unemployed and experience a large drop in income. Therefore she will ask for a hefty premium to hold stocks over the Treasury Bill.\(^5\) Constantinides and Duffie (1996) establish this result within the framework of incomplete markets where agents face permanent idiosyncratic income shocks, which are undiversifiable. The work of Asdrubali, Sorensen and Yosha (1996) and Athanasoulis and Van Wincoop (2001) establishes that at the U.S. state-level a large portion of the shocks to state income are not smoothed out by financial markets.

Therefore, the level of the state unemployment rate can serve as a predicting variable. When unemployment is high, individuals are more likely to be laid off and experience a decrease in their income that cannot be diversified. In this case, they will ask a high return on stocks.

2. Data

The paper uses quarterly U.S. state-level data for the period 1982(Q1) to 1999(Q1). All the time series are transformed into real terms using regional inflation rates from the Bureau of Labor Statistics. The base year for the inflation rates is 1992(Q1). All the time series, but the returns, are in per capita terms. State population estimates are from the Census Population Survey.

The Returns

The return data include the common stocks with CRSP share code of 10 or 11. The quarterly local return is a value-weighted quarterly return of firms located in each state\(^6\). The firm’s location is the location of its headquarter, and not where the firm is incorporated.\(^7\) The quarterly return of each firm is computed by compounding its monthly returns. The log-real local return at quarter \(t\) is denoted by \(r_t\). The quarterly market return is the value-weighted return of all the stocks listed at NYSE, AMEX

\(^5\)For details see Constantinides and Duffie (1996). Also, the discussions in Constantinides (2002) and Cochrane (2005) are useful.

\(^6\)Due to missing values across the different variables Alaska, Arizona, Washington DC, New Mexico and West Virginia are excluded from the analysis.

\(^7\)Loughran and Schultz (2004) and Pirinsky and Wang (2004) also use the location of a firm’s headquarter as the location of the firm itself.
and NASDAQ compounded from monthly returns. Similarly, the quarterly risk-free return is calculated by compounding the monthly returns of the 1-month Treasury bill. The log-real market and risk-free returns at quarter \( t \) are denoted by \( r_{m t} \) and \( r_{ft} \).

**The Predicting Variables**

The local returns are predicted using a state-level collateral ratio, the \( c_{ay} \) of Lettau and Ludvigson (2001), the state unemployment rate and the state-level dividend-price ratio. The data for the \( c_{ay} \) are downloaded from Ludvigson’s web site and the unemployment rates are from the BLS. The local dividend-price ratio is calculated using the same selection criteria as the local returns. Its log-value is denoted by \( d_t - p_t \).

The state-level collateral ratio, \( h_y \) uses log state labor income, \( y \), from the Bureau of Economic Analysis (BEA). The time-series of the value of log state housing, \( h \), is from Case, Quigley and Shiller\(^8\) (2001). As in Lustig and Van Nieuwerburgh (2004), I test whether per capita log-real labor income and home equity are cointegrated at the state level. First, the trace and maximum eigenvalue cointegration tests of Johansen and Juselius (1990) are conducted. These tests are calculated for each state. For all states and at the 5% significance level, both statistics rejected the null hypothesis of no cointegrating relation, and accepted the null hypothesis of one cointegrating relation.\(^9\)

Next, the cointegrating vector is estimated by the panel dynamic least-squares (PDSL) of Mark and Sul (2003). A fixed effect (state-specific mean) is allowed in the cointegrating vector to accommodate heterogeneity across the U.S. states. The state-level collateral ratio, \( h_y \), is defined using the estimated cointegrating vector. In mean-free form, it is equal to

\[
h_{y_t} = h_t - 0.808 \cdot y_t,
\]

(0.448)

where 0.808 is the PDSL estimate from the cointegrating vector between \( h \) and \( y \), and 0.448 is the standard error of this estimate.

---

\(^8\)The data for the state-level value of housing equity were provided by Robert Shiller, whom I thank.

\(^9\)The auxiliary regressions include two first differences. Also, the cointegrating vector contains a constant term. Detailed tables for the trace and maximum eigenvalue statistics are available from the author upon request.
Lustig and Van Nieuwerburgh (2004) argue that as the collateral ratio decreases, collateral becomes more scarce, consumption growth becomes more perceptible to idiosyncratic income shocks and the variance of consumption growth increases. This intuition implies that states, which on average have a lower collateral ratio, should have a more volatile consumption growth $\Delta c$, given the variance of labor income growth $\Delta y$. This hypothesis is tested by estimating a least-square cross-sectional regression of the state consumption growth variances on the median\(^{10}\) state collateral ratios $hy$ and on the state income growth variances.

State-level consumption, $c$, is approximated with sales at all retail establishments.\(^ {11}\) Retail sales are from the Regional Financial Associates (Zandi and Goebert, 1997). State-consumption growth is calculated with real per-capita retail sales. The least-square regression takes the following form:

$$\text{Var}(\Delta c) = -0.003 \cdot \text{Median}(hy) + 0.171 \cdot \text{Var}(\Delta y), \quad R^2 = 0.20$$

where the numbers in parenthesis are the $t$-statistics. The regression can explain 20% of the cross-sectional variation in the variance of the state consumption growth. Most importantly, the estimates confirm that even after controlling for the variance of income growth, states with lower collateral ratio do have more volatile consumption growth. Hence the collateral mechanism, that Lustig and Van Nieuwerburgh (2004) show to fan out the distribution of aggregate consumption growth, is present in state-level data.

**Summary Statistics**

Table 1 presents summary statistics for the returns and the predictors. The correlation matrix in Panel A of Table 1 shows that local return and the market return are highly correlated; their correlation coefficient is 0.79. This correlation indicates that a significant part of the volatility of the local return is due to the market return. Therefore, in predicting the local return, one should include a predictor, like the $cay$, that can capture the market component of the local return.

\(^{10}\)I use the median $hy$ because group specific averages have been subtracted from $h$ and $y$ to accommodate heterogeneity across the U.S. states. In this case, the average of the mean-free $hy$ is zero, whereas the median is different from zero.

\(^{11}\)This assumption is also made in Case, Shiller and Quigley (2001), Ostergaard, Sorensen and Yosha (2002), and Asdrubali, Sorensen and Yosha (1996).
The standard deviation of the local return is 0.10 which is higher than that of the market return. This is true for the excess of the local return over the Treasury Bill return. Local returns are therefore more volatile and in principle it will be more difficult to predict them compared to predicting the market return.

The local dividend-price ratio has properties similar to dividend-price ratio related to the market return. Averaging across states, its mean value is -3.76, its standard deviation is 0.44 and its autocorrelation is 0.92\textsuperscript{12} and they are presented in Panel B of Table 1. The autocorrelation coefficients for \( hy \) and \( cay \) are 0.84 and 0.92 respectively. These two variables are less volatile than the dividend-price ratio; the standard deviation for \( hy \) and \( cay \) is 0.10 and 0.01 respectively. The most volatile predicting variable is the state unemployment rate; across states its mean value is 6.18\% and its standard deviation is 1.67. Its autocorrelation is 0.97.

The dividend-price ratio is positively correlated with \( cay \) and the unemployment rate and it is almost orthogonal to \( hy \). The collateral ratio is only slightly correlated with the \( cay \) and the unemployment rate; their correlation coefficients are -0.15 and 0.09, respectively. The correlation between the \( cay \) and the unemployment rate is also low. These correlations are reported in the lower right portion of Panel A in Table 1. Hence, a significant part of the information in the dividend-price ratio is captured by the \( cay \) and the unemployment rate, whereas the other predicting variables have independent information. Nevertheless, the dividend-price ratio will be included in the forecasting regressions to let the estimation results indicate whether the other predictors make it statistically insignificant.

\textbf{3. One-Quarter-Ahead Forecasting Regressions}

In this section, one-quarter ahead forecasting regressions are estimated on the local return \( r \), on the local return over the Treasury Bill return \( r - r_f \) and on the idiosyncratic component of the local return, \( r - r_m \). To control for any serial correlation in the return data, a lag-dependent variable is included in the regressions. By pooling the observations of all states, the forecasting models are formulated as dynamic

\textsuperscript{12}These value are very close to those reported by Lettau and Ludvigson (2001) for the dividend-price ratio related to the market return.
panel regressions with fixed effects:

\[ y_{it} = \alpha_i + \gamma y_{i,t-1} + x_{i,t-1} \beta + \varepsilon_{it}, \tag{2} \]

where \( i \) is the index for the U.S. states, \( t \) is the index for the time periods and \( \alpha_i \) is state-specific mean. The variable \( y_{it} \) is either \( r_{it} \), or \( r_{it} - r_{ft} \), or \( r_{it} - r_{mt} \). The row vector \( x_{i,t-1} \) contains the predictors \( d_{i,t-1} - p_{i,t-1}, \) \( \text{cay}_{t-1}, \) \( h y_{i,t-1} \) and the state unemployment rate. Note that the \( \text{cay} \) is the same for all states. The variable \( \varepsilon_{it} \) is the regression error-term. By pooling the observations, one increases the power of the statistical inference because one is using more observations to estimate the coefficients, calculate their standard errors and \( t\)-\textit{statistics}. In this case the pooled estimates of \( \beta \) and \( \gamma \) measure the average importance across states of the forecasting variables in predicting returns.

The dynamic panel models are estimated with the corrected least-square dummy-variable (\textit{LSDV}) approach of Hahn and Kuersteiner (2002). The estimation results are presented in Table 2. In Panel A of Table 2, I include the estimates and their \( t\)-\textit{statistics}. In Panel B of Table 2, I report the expected percentage change in returns when the forecasting variables increase by one standard deviation.

In general the forecasting ability of the panel regressions in Table 2 is not as good as the forecasting ability of regression on the market return. For example, the \( R^2 \) for the regression on the local return \( r_t \) is only 0.019, which is lower than 0.1, the \( R^2 \) reported in Lettau and Ludvigson (2001) for similar regressions for the market return. This is not surprising because the state panel regressions are asked to forecast a return, which is more volatile than the market return. As reported in Table 1, Panel B the average standard deviation across states for \( r \) is 0.10, and the standard deviation for \( r_m \) is 0.03. The explanatory power of the forecasting regressions is the lowest for the idiosyncratic component of the local return, \( r - r_m \); the \( R^2 \) is only 0.006.

The low explanatory power of the panel regressions is not surprising. The pooled panel regression assumes that the slope coefficients \( \beta \) and \( \gamma \) are the same across all state returns ignoring any possible form of heterogeneity. To test this hypothesis I allow for \( \beta \) and \( \gamma \) to differ across states, and I estimate \textit{OLS}

\footnote{The \textit{t\text{-statistics}} take into account that the estimated residuals \( \varepsilon_{it} \) are heteroscedastic across states. I do not correct for serial autocorrelation in the estimated residuals because they are IID. This conclusion is reached by calculating two Durbin-Watson statistics for each state: one between \( \varepsilon_{it} \) and \( \varepsilon_{i,t-1} \), and another one between \( \varepsilon_{it} \) and \( \varepsilon_{i,t-2} \). The first one, \( DW_1 \), measures first order serial correlation and the second one, \( DW_2 \), measures second order serial autocorrelation. Then I averaged \( DW_1 \) and \( DW_2 \) across states. In all regression these averages are almost equal to 2.}
regressions for each state return. I calculate the $R^2$ of each regression and then I average them.\footnote{Detailed results are available from the author upon request.} The average $R^2$ for the regression on $r$, $r - r_f$, and $r - r_m$ are 0.09, 0.08 and 0.13 respectively. As expected, they are higher than those of the pooled panel regression and similar to those reported in Lettau and Ludvigson (2001).

The three regressions in Table 2 include the past value of the state returns. These lag dependent variables are statistically insignificant, even at the 15% significant level. Therefore, there is no evidence of short-term momentum in the returns.

**The dividend-price ratio**

In the case of the dividend-price ratio, the accounting identity (1) dictates that the null, $H_0$, and alternative, $H_A$, hypothesis of the predictability test are the following:

$$H_0 : \beta = 0, \quad H_A : \beta > 0.$$  

The $t$-statistics of this test are reported in Table 2 and they show that the dividend-price ratio is statistically insignificant in forecasting the one-quarter ahead local returns. This result is not surprising because in the literature this variable typically forecasts market returns at horizons in excess of a couple of years.\footnote{For example, see Campbell, Lo and MacKinley (1997).} Furthermore, the $cay$ and the state unemployment rate have similar information as the dividend-price ratio, making the latter insignificant. Recall that the dividend-price ratio is positively correlated with the $cay$ and the state unemployment rate; their correlation coefficients are 0.36 and 0.49 respectively. See Table 1, Panel B.

Nevertheless, the estimates on the dividend-price ratio are positive as predicted by the accounting identity (1). If the dividend-price ratio increase by one standard deviation, then it is predicted that in the next quarter local returns should increase by 0.3%, local returns over the Treasury Bill return should increase by 0.21% and the idiosyncratic component of the local returns should increase by 0.08%. These numbers are reported in Panel B of Table 2.
The consumption trend deviation $cay$

The work of Lettau and Ludvigson (2001) shows that in the case of the $cay$, the null, $H_0$, and alternative, $H_A$, hypothesis of the predictability test are:

$$H_0 : \beta = 0, \quad H_A : \beta > 0.$$ 

The $t$-statistics of this test are reported in Table 2, and they show that the $cay$ residual is statistically significant in forecasting the raw local return and the local return over the Treasury Bill. As expected, it cannot forecast the idiosyncratic component of the local return since it primarily includes information about the aggregate stock market. In the regressions for $r$ and $r - r_f$ its coefficient is positive, which is in accordance to the theoretical prediction. “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 income in an attempt to insulate future consumption from lower returns, and vice versa” (Lettau and Ludvigson, 2001, p. 829-830). Since the $cay$ mainly includes aggregate information, its impact on returns is lower than the one estimated by Lettau and Ludvigson (2001) for the market return. These authors find that a one-standard-deviation increase in $cay$ implies a 2.2% increase in one-quarter-ahead market return. However, for the one-quarter-ahead local return the implied increase drops to 0.81%.

The collateral ratio $hy$

The consumption asset pricing model of Lustig and Van Nieuwerburgh (2004) shows that for the collateral ratio $hy$, the null, $H_0$, and alternative, $H_A$, hypothesis of the predictability test are:

$$H_0 : \beta = 0, \quad H_A : \beta < 0.$$ 

The $t$-statistics of this test are reported in Table 2, and they show that the state collateral ratio $hy$ is statistically significant in all three regressions. However, it is more important in the regression for the idiosyncratic component of the local return–its $t$-statistic increases from -1.69 in the $r$ regression to -1.99 in the $r - r_m$ regression. The collateral ratio $hy$ reflects the scarcity of housing collateral at the state level,
which is affected by local macroeconomic conditions. It should therefore be important in forecasting the state-specific component of the local return.

In all regressions the estimated coefficient on $h_y$ is negative as predicted by the model of Lustig and Van Nieuwerburgh (2004). When the state-level $h_y$ decreases, the variance of state consumption growth increases and individuals ask for a higher return.\footnote{This result is established in Section 3.} In particular, if the state $h_y$ decreases by one-standard-deviation, then the one-quarter-ahead local return should increase by 0.27\%, the local return over the Treasury Bill return should increase by 0.28\%, and the idiosyncratic component of the local return should increase by 0.2\%.

**The State Unemployment Rate**

The work of Constantinides and Duffie (1996) suggests that for the state unemployment rate, the null, $H_0$, and alternative, $H_A$, hypothesis of the predictability test are:

$$H_0 : \beta = 0, \quad H_A : \beta > 0.$$  

The $t$-statistics of this test are reported in Table 2, and they show that the state unemployment rate is statistically significant in all forecasting regressions.

The coefficient estimates are positive as the theory of Constantinides and Duffie (1996) predicts. In an economy where idiosyncratic income shocks are not diversifiable, individuals are skeptical about holding stocks, which perform poorly during recessions. Therefore, during recessions, which are also the time-periods where people face the highest probability of being laid off, they ask a hefty premium to hold stocks. The estimates confirm this story; when unemployment increases by one-standard-deviation, the one-quarter-ahead local return is predicted to increase by 0.62\%, the local return over the Treasury Bill return is predicted to increase by 0.52\%, and the idiosyncratic component of the local return should increase by 0.32\%.

To summarize, the one-quarter ahead forecasting regressions reveal that the lagged value of the returns has no predictive power. Similarly, the dividend-price ratio becomes insignificant in the presence of the
cay residual and the unemployment rate. The trend deviation cay can predict the market component of the local return since it is significant in the regressions for \( r \) and \( r - r_f \), but it is insignificant in the regression for \( r - r_m \). The collateral ratio is also an important predictor, especially in predicting the idiosyncratic component of the local returns. Finally, the state unemployment rate is statistically significant in all three regressions.

4. Long-horizon Forecasts

In this section, long-run forecasting regressions are estimated following Fama and French (1988). The goal is to test whether we can predict the long-term component of returns. This component is approximated with the \( K \)-horizon return, which is defined as the sum of \( K \) short-term returns. For example, the \( K \)-horizon raw return, denoted by \( r_{K,t} \), equals the sum \((r_t + r_{t+1} + ... + r_{t+K})\). Following the literature, I estimate long-run regressions on the local return in excess of the Treasury Bill return and the market return. By pooling the observations of all states, the \( K \)-horizon forecasting models are formulated as panel regressions with fixed effects:

\[
y_{K,it} = \alpha_{K,i} + x_{i,t-1}\beta_K + \varepsilon_{K,it}, \tag{3}
\]

where \( i \) is the index for the U.S. states, \( t \) is the index for the time periods and \( \alpha_{K,i} \) is state-specific mean. The variable \( y_{K,it} \) is either the \( r_{K,it} - r_{K,ft} \), or \( r_{K,it} - r_{K,mt} \). The row vector \( x_{i,t-1} \) contains the predictors \( d_{i,t-1} - p_{i,t-1}, \text{cay}_{i,t-1}, \text{hy}_{i,t-1} \) and the state unemployment rate. The panel models are estimated with the least-square dummy-variable estimator.\(^{17}\)

The estimation results are presented in Tables 3 and 4. Table 3 includes the estimates and the \( t \)-statistics for the excess long-term return over the Treasury Bill return, \( r_K - r_{K,f} \). Table 4 includes the estimates and the \( t \)-statistics for the state-specific component of the long-run return, \( r_K - r_{K,m} \). The estimated residuals in regression (3) are serially autocorrelated due to the fact that \( y_{K,it} \) and \( y_{K,it+\tau} \), \( \tau = \{1, 2, ..., K\} \), use overlapping observations. Following the work of Hansen and Tuypens (2004), and Hjalmarsson (2004) the \( t \)-statistics are scaled by \( 1/\sqrt{K} \) to reflect this autocorrelation.

\(^{17}\)To apply the least-square dummy-variable estimator, one has to subtract group averages from the data. Then, one applies ordinary least-squares on the mean-free data.
Previous studies, including Fama and French (1988), Lettau and Ludvigson (2001) and Lustig and Van Nieuwerburgh (2004), find that the longer the horizon $K$ is, the better we can predict long-term returns; the $R^2$ of the forecasting regressions increases with the horizon $K$. As Cochrane (1999) explains, predictability builds with time because the predicting variables are persistence. These results also hold with state data. The explanatory power of the long-run regressions in Tables 3 and 4 is higher than the explanatory power of the one-quarter-ahead forecasting regressions in Table 2. For example, when the horizon $K$ is 3 years, the $R^2$ for $r_K - r_{Kf}$ and $r_K - r_{Km}$ is 0.255 and 0.027 respectively. Whereas the $R^2$ for $r - r_f$ and $r - r_m$ decreases to 0.017 and 0.006 respectively.

The estimations also reveal that it is more difficult to predict the idiosyncratic component of the long-run return; for all horizons the $R^2$ for the $r_K - r_{Kf}$ regression is always higher than the one for the $r_K - r_{Km}$ regression. Further predictability peaks at the 3 year horizon for both types for forecasting regressions. See the $R^2$’s in last columns in Tables 3 and 4.

As in the case of the one-quarter-ahead forecasting regressions, I explore whether the explanatory power of the long-horizon regressions increases when the assumption of common $\beta_K$’s across state returns in is relaxed. Thus, I allow for the $\beta_K$’s to be heterogeneous across states, and I estimate regression (3) for each state return. I calculate the $R^2$ of each regression and then I average them. As expected across all horizons these $R^2$’s are higher than the respective ones from the pooled regression. For example, when $K = 1$, the average $R^2$ for the regression on $r_K - r_{Kf}$ and $r_K - r_{Km}$ are 0.13 and 0.18 respectively. For $K = 12$, they increase to 0.51 and 0.48 for the regression on $r_K - r_{Kf}$ and $r_K - r_{Km}$ respectively.

The dividend-price ratio

Cochrane (1997) finds that the dividend-price ratio can predict excess returns at the 5 year horizon. Across almost all horizons however, the $t$-statistics in Tables 3 and 4 of the one-sided test

$$H_0 : \beta = 0, \quad H_A : \beta > 0,$$

reveal that the dividend-price ratio has no predictive power for either $r_K - r_{Kf}$ or $r_K - r_{Km}$. This results is expected because the cay and the unemployment rate include most of the information in the dividend-price ratio.
The consumption trend deviation \( cay \)

In the case of the \( cay \), I calculate the \( t \)-statistics of the following one-sided test:

\[
H_0 : \beta = 0, \quad H_A : \beta > 0.
\]

For all values of \( K \), the \( t \)-statistics show that the \( cay \) is statistically significant in forecasting \( r_K - r_{Kf} \), and it is statistically insignificant for predicting \( r_K - r_{Km} \). This result confirms that the \( cay \) can better predict the component of the local return that covaries with the market return.\(^{19}\) When predicting the excess long-term return over the Treasury Bill return, the significance of the \( cay \) peaks at the 3 year horizon; up to that point its \( t \)-statistic is increasing and from that point onwards it is decreasing. Further across all horizons, Table 3 shows that the estimates on the \( cay \) are always positive, as predicted by the theory. Individual, how like to smooth consumption, will increase consumption above its long-run trend, when returns are expected to be high in the future.

The collateral ratio \( hy \)

In the case of the \( hy \), I calculate the \( t \)-statistics of the following one-sided test:

\[
H_0 : \beta = 0, \quad H_A : \beta < 0.
\]

Across all horizons, the \( t \)-statistics indicate that the collateral ratio \( hy \) is statistically important in predicting \( r_K - r_{Kf} \) and \( r_K - r_{Km} \). It is more significant in short-term horizons and it becomes insignificant after the 3 year horizon; the \( t \)-statistics in Tables 3 and 4 decrease with \( K \) and after the 3 year horizon they become smaller than 1.50. In line with the theory, its estimated coefficient is always positive. The collateral ratio is a measure of financial distress. When it is low, individuals fear negative idiosyncratic

\(^{19}\)Recall that Lettau and Ludvigson (2001) show that the \( cay \) is an important predictor for the market return in excess of the Treasury Bill return.
income shocks, because they do not have enough collateral to take out loans. In this case, they ask for a high return to hold stocks.

**The State Unemployment Rate**

The state unemployment rate is a good predictor for the state-specific component of the long-term returns. The *t-statistics* in Table 4 for the one-sided test

\[ H_0 : \beta = 0, \quad H_A : \beta > 0, \]

reveal that it is statistically significant up to the 2 year horizon. However, in the \( r_K - r_{Kf} \) regressions, it is statistically significant at the 1 quarter horizon, a result found in the one-quarter ahead forecasting regressions in Table 2. It also becomes significant between the 3 year and the 4 year horizons, which are the horizons with the highest \( R^2 \). For the other horizons it is statistically insignificant. See Table 3.

Across all the forecasting horizons in Tables 3 and 4, the estimated coefficient on the unemployment rate are positive, in accordance to the model of Constantinides and Duffie (1996). In a world of incomplete risk-sharing, agents fear persistent idiosyncratic income shocks such as being laid off. Therefore, when the unemployment rate is high, they require a high return on stocks.

To summarize, the long-run regressions reveal that state returns are the most predictable at the 3 year horizon. As in the case of the one-quarter ahead forecasting regression, the dividend-price ratio is insignificant because most of the information it contains is captured by the residual \( cay \) and the state unemployment rate. It is found that the trend deviation \( cay \) can only predict \( r_K - r_{Kf} \), whereas the unemployment rate is an important predictor only for \( r_K - r_{Km} \). The collateral ratio though is statistically significant for both the \( r_K - r_{Kf} \) and \( r_K - r_{Km} \) returns.

**Caveats**

Campbell and Yogo (2003) stress that statistical inference is problematic when predicting returns with highly persistent variables, which have innovations that are correlated with the error \( \varepsilon_K \) in regression (3). They show that large-sample theory provides a poor approximation to the finite-sample distribution of the test-statistics, when testing for predictability. Mark and Sul (2004) establish similar results, which
are also in line with the work of Stambaugh (1999). Even if the predictors are highly persistent, Mark and Sul (2004) show that the large sample distortions are minimal when the correlation between the error $\varepsilon_K$, of the forecasting regression (3) for $K = 1$, and the innovations in the predictors is low.

In the state data all the predictors are persistent. The autocorrelation coefficients, reported in Panel B of Table 1, for the dividend-price ratio, the $cay$, the $hy$ and the unemployment rate are 0.92, 0.84, 0.92 and 0.97, respectively. Even if they are persistent, these variables are stationary. For all the variables I run the panel unit root tests in Choi (2001). These tests reject the hypothesis a unit root.\textsuperscript{20} Of course, it is not surprising that the $cay$ and the $hy$ are stationary; by construction they are cointegrating residuals. As a consistency check I also conduct the Said and Dickey (1984) residual-based cointegration test on the estimated residuals of all the forecasting regressions. If the residuals are cointegrated then the regressions are not spurious, even if the predictors have a near-unit root. In all cases, the spurious regression hypothesis is rejected at the 1% significant level.\textsuperscript{21}

Fortunately, the predictors are not highly correlated with the innovation $\varepsilon_K$ for $K = 1$. First, the correlation coefficients between the current values of the dividend-price ratio, the $cay$, the $hy$ and the unemployment rate with the local return, are around 0.12, 0.11, -0.03 and 0.12 respectively. See Table 1, Panel B. Second, I calculate the contemporaneous correlation between the estimated residuals of the regression (3) for $K = 1$ and the innovations of the predictors.\textsuperscript{22} In the case of the regression on $r_K - r_{Kf}$ for $K = 1$, they are equal to -0.16, -0.16, 0.03, and 0.06 for dividend-price ratio, the $cay$, the $hy$ and the unemployment rate respectively. In the case of the regression on $r_K - r_{Km}$ for $K = 1$, they are equal to -0.09, 0.01, 0.01, and 0.06 for dividend-price ratio, the $cay$, the $hy$ and the unemployment rate respectively. Since these correlations are not too high, one can use the large-sample distributions of the test-statistics, even if the forecasting variables are persistent.

\textsuperscript{20}The tables for the panel unit root tests are available from the author upon request.
\textsuperscript{21}The tables for the Said and Dickey (1984) tests are available from the author upon request.
\textsuperscript{22}For the dividend-price ratio and the state unemployment rate, I estimate panel AR(1) regression using the corrected least-squares of Hanh and Kuersteiner (2002). Then, from the residuals of AR(1) regressions I calculate the innovations for the dividend-price ratio and the state unemployment rate. For the $cay$ and the $hy$ I do not estimate any regressions because they are, by definition, innovations.
5. Conclusion

This paper shows that predictability is a robust phenomenon that extends from the return of the market portfolio to the returns of state-level portfolios. In order to predict state-level returns, I use macroeconomic variables. They include the deviation of U.S. consumption from its long-run trend developed by Lettau and Ludvigson (2001). Another macroeconomic predictor is the collateral ratio of Lustig and Van Nieuwerburgh (2004).

A novel feature of the study is to recognize that there is no full risk-sharing among the U.S. states, which indicates that the state-level unemployment rate can be a useful predictor. Asdrubali, Sorensen and Yosha (1996) show that only 75% of the shocks on gross state product are smoothed across the states. Constantinides and Duffie (1996) show that in an incomplete markets setting, where idiosyncratic shocks are not diversifiable, agents ask for a hefty premium to hold stocks. People fear stocks because they tend to do badly during recessions, which are periods with high unemployment.

The local return of state $i$ includes the stocks all the firms with their headquarters in state $i$ for the period from 1982(Q1) to 1999(Q1). The empirical analysis estimates one-quarter-ahead forecasting regressions for the local return $r$, the excess of the local return over the Treasury Bill return, $r - r_f$, and the idiosyncratic component of the local return. This is defined as the excess of the local return over the market return, $r - r_m$. In line with Fama and French (1988), regressions on long-term returns $r_K - r_{Kf}$ and $r_K - r_{Km}$ are also estimated.

The estimation results establish that state-level returns are predictable using macroeconomic variables. In particular, the one-quarter ahead forecasting regressions reveal that the dividend-price ratio becomes insignificant in the presence of the $cay$ residual and the unemployment rate. The trend deviation $cay$ predicts the market component of the local return since it is significant in the regressions for $r$ and $r - r_f$, but it is insignificant in the regression for $r - r_m$. The collateral ratio is also important especially in predicting the idiosyncratic component of the local returns because it can capture financial distress at the state level. Finally, the state unemployment rate is statistically significant in all three regressions.

The long-run regressions reveal that state returns are the most predictable at the 3 year horizon. Similar to the case of the one-quarter ahead forecasting regression, the dividend-price ratio is insignificant.
because most of its information is captured by the residual \( cay \) and the state unemployment rate. It is found that the trend deviation \( cay \) can only predict \( r_K - r_{Kf} \), and the unemployment rate is an important predictor only for \( r_K - r_{Km} \). The collateral ratio though is statistically significant for both the \( r_K - r_{Kf} \) and \( r_K - r_{Km} \) returns.
Predicting Local Returns with Macroeconomic Variables

References


The data are quarterly U.S. state data for 1982(Q1) to 1999(Q1). All time series are expressed in real terms using regional inflation rates from the BLS. The base year is 1992(Q1). The quarterly local return is a value-weighted return of firms located in each state. The firm’s location is the location of its headquarter. I included all firms listed in CRSP. The quarterly return of each firm is computed by compounding its monthly returns. The log-real local return is denoted by \( r \). Similarly, the quarterly risk-free return is calculated by compounding monthly returns of the 1-month Treasury bill. The log-real market and risk-free returns are denoted by \( r_{mt} \) and \( r_f \). The \( cay \) is from Lettau and Ludvigson (2001). The state-level collateral ratio, \( hy \), uses state log labor income, \( y \), from the Bureau of Economic Analysis (BEA). The time-series of log state housing, \( h \), is from Case, Quigley and Shiller (2001). As in Lustig and Van Nieuwerburgh (2004), the log collateral ratio is equal to \( h - 0.808 \cdot y \) where 0.808 is the \( PDLs \) estimate of Mark and Sul (2003) from the cointegrating vector between \( h \) and \( y \). The state unemployment rates are from the BLS and they are expressed in percentages. The summary statistics presented in the table are averages across states. For example, the reported standard deviation of \( r_t \), is calculated by averaging the standard deviations of all states.

### Table 1

**Summary Statistics for Returns and Predictors**

<table>
<thead>
<tr>
<th>Returns</th>
<th>( r_t )</th>
<th>( r_{mt} )</th>
<th>( r_t - r_f )</th>
<th>( r_t - r_{mt} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_t )</td>
<td>1</td>
<td>0.82</td>
<td>1.00</td>
<td>0.98</td>
</tr>
<tr>
<td>( r_{mt} )</td>
<td>1</td>
<td>0.81</td>
<td>0.69</td>
<td>1</td>
</tr>
<tr>
<td>( r_t - r_f )</td>
<td>1</td>
<td>0.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r_t - r_{mt} )</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Predictors</th>
<th>( d_{t-1} - p_{t-1} )</th>
<th>( cay_{t-1} )</th>
<th>( hy_{t-1} )</th>
<th>( unemp\ rate_{t-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( d_{t-1} - p_{t-1} )</td>
<td>1</td>
<td>0.36</td>
<td>-0.01</td>
<td>0.49</td>
</tr>
<tr>
<td>( cay_{t-1} )</td>
<td>1</td>
<td>-0.15</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>( hy_{t-1} )</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( unemp\ rate_{t-1} )</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel A: Correlation Matrix with Lagged Values of Predictor

Panel B: Summary Statistics and Correlations with Current Values of Predictors

<table>
<thead>
<tr>
<th></th>
<th>( r_t )</th>
<th>( r_{mt} )</th>
<th>( r_t - r_f )</th>
<th>( r_t - r_{mt} )</th>
<th>( d_{t-1} - p_t )</th>
<th>( cay_t )</th>
<th>( hy_t )</th>
<th>( unemp\ rate_{t-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.05</td>
<td>0.03</td>
<td>0.04</td>
<td>0.01</td>
<td>-3.76</td>
<td>0.62</td>
<td>1.97</td>
<td>6.18</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.10</td>
<td>0.03</td>
<td>0.09</td>
<td>0.06</td>
<td>0.44</td>
<td>0.01</td>
<td>0.10</td>
<td>1.67</td>
</tr>
<tr>
<td>Autocorrelation</td>
<td>0.00</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.92</td>
<td>0.84</td>
<td>0.92</td>
<td>0.97</td>
</tr>
<tr>
<td>Correlation with ( d_{t-1} - p_t )</td>
<td>0.12</td>
<td>0.07</td>
<td>0.10</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation with ( cay_t )</td>
<td>0.11</td>
<td>0.07</td>
<td>0.11</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation with ( hy_t )</td>
<td>-0.03</td>
<td>-0.01</td>
<td>-0.04</td>
<td>-0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation with ( unemp\ rate_t )</td>
<td>0.12</td>
<td>0.09</td>
<td>0.10</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2
One-Quarter Ahead Forecasting Regressions

The Table presents the estimation results of the forecasting regressions. The bold-faced numbers are the estimates, and the numbers underneath them are their \textit{t-statistic}. The forecasting regression takes the following form: \( y_{it} = \alpha_i + \gamma y_{i,t-1} + x_{i,t-1} \beta + \epsilon_{it} \), where \( i \) is the index for the U.S. states, \( t \) is the index for the time periods and \( \alpha_i \) is state-specific mean. The variable \( y_{it} \) is either \( r_{it} \), or \( r_{it} - r_{ft} \), or \( r_{it} - r_{mt} \). The row vector \( x_{i,t-1} \) contains the predictors \( d_{i,t-1} - p_{i,t-1} \), \( cay_{i,t-1} \), \( hy_{i,t-1} \) and the state unemployment rate. The data are quarterly U.S. state data for 1982(Q1) to 1999(Q1). All time series are expressed in real terms using regional inflation rates from the BLS. The base year is 1992(Q1). The quarterly local return is a value-weighted return of firms located in each states. The firm’s location is the location of its headquarter. I included all firms listed in CRSP. The quarterly return of each firm is computed by compounding its monthly returns. The log-real local return is denoted by \( r \). Similarly, the quarterly risk-free return is calculated by compounding monthly returns of the 1-month Treasury bill. The log-real market and risk-free returns are denoted by \( r_m \) and \( r_f \). The \textit{cay} is from Lettau and Ludvigson (2001). The state-level collateral ratio, \( hy \) uses state log labor income, \( y \), from the Bureau of Economic Analysis (BEA). The time-series of log state housing, \( h \), is from Case, Quigley and Shiller (2001). As in Lustig and Van Nieuwerburgh (2004), the log collateral ratio is equal to \( h - 0.808 \cdot y \) where 0.808 is the PDLs estimate of Mark and Sul (2003) from the cointegrating vector between \( h \) and \( y \). The state unemployment rates are from the BLS. The forecasting regressions are estimated using the corrected LSDV of Hahn and Kuersteiner (2002).

<table>
<thead>
<tr>
<th>Panel A : Estimates of Forecasting Panel Regressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{Local Return} )</td>
</tr>
<tr>
<td>( \text{Local Return over T-Bill} )</td>
</tr>
<tr>
<td>( \text{Local Return over Market Return} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B : % Change in Returns if Predictors Increase by 1 Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{Local Return} )</td>
</tr>
<tr>
<td>( \text{Local Return over T-Bill} )</td>
</tr>
<tr>
<td>( \text{Local Return over Market Return} )</td>
</tr>
</tbody>
</table>
Table 3
Long-Run Forecasting Regressions of State Returns over the Treasury Bill Return

The Table presents the estimation results of the long-run forecasting regressions for the state-return over the Treasury Bill Return. The bold-faced numbers are the estimates, and the numbers underneath them are their t-statistic. The forecasting regression takes the following form: $r_{K,it} - r_{K,ft} = \alpha_{K,i} + x_{i,t-1} \beta_{K} + \varepsilon_{it}$, where $i$ is the index for the U.S. states, $t$ is the index for the time periods and $\alpha_i$ is state-specific mean. The variable $r_{K,it} - r_{K,ft}$ equals $r_{it} - r_{ft} + r_{i,t+1} - r_{f,t+1} + \ldots + r_{i,t+K} - r_{f,t+K}$. The row vector $x_{i,t-1}$ contains the predictors $d_{i,t-1} - p_{i,t-1}$, $cay_{t-1}$, $hy_{i,t-1}$ and the state unemployment rate. The data are quarterly U.S. state data for 1982(Q1) to 1999(Q1). All time series are expressed in real terms using regional inflation rates from the BLS. The base year is 1992(Q1). The quarterly local return is a value-weighted return of firms located in each states. The firm’s location is the location of its headquarters. I included all firms listed in CRSP. The quarterly return of each firm is computed by compounding its monthly returns. The log-real local return is denoted by $r$.

Similarly, the quarterly risk-free return is calculated by compounding monthly returns of the 1-month Treasury bill. The log-real market and risk-free returns are denoted by $r_m$ and $r_f$. The $cay$ is from Lettau and Ludvigson (2001). The state-level collateral ratio, $hy$ uses state log labor income, $y$, from the Bureau of Economic Analysis (BEA). The time-series of log state housing, $h$, is from Case, Quigley and Shiller (2001). As in Lustig and Van Nieuwerburgh (2004), the log collateral ratio is equal to $h - 0.808 \cdot y$ where 0.808 is the PDLs estimate of Mark and Sul (2003) from the cointegrating vector between $h$ and $y$. The state unemployment rates are from the BLS. The forecasting regressions are estimated using the LSDV estimator.

<table>
<thead>
<tr>
<th>Horizon $K$</th>
<th>Forecasting Variables</th>
<th>R$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$d - p$</td>
<td>$cay$</td>
</tr>
<tr>
<td>1</td>
<td>0.013</td>
<td>0.978</td>
</tr>
<tr>
<td></td>
<td>2.47</td>
<td>4.47</td>
</tr>
<tr>
<td>2</td>
<td>0.018</td>
<td>1.326</td>
</tr>
<tr>
<td></td>
<td>2.04</td>
<td>3.62</td>
</tr>
<tr>
<td>3</td>
<td>0.019</td>
<td>2.439</td>
</tr>
<tr>
<td></td>
<td>1.55</td>
<td>4.50</td>
</tr>
<tr>
<td>4</td>
<td>0.015</td>
<td>3.817</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>5.14</td>
</tr>
<tr>
<td>5</td>
<td>0.012</td>
<td>5.713</td>
</tr>
<tr>
<td></td>
<td>0.65</td>
<td>6.26</td>
</tr>
<tr>
<td>6</td>
<td>0.009</td>
<td>7.232</td>
</tr>
<tr>
<td></td>
<td>0.45</td>
<td>6.75</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Horizon $K$</th>
<th>Forecasting Variables</th>
<th>R$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$d - p$</td>
<td>$cay$</td>
</tr>
<tr>
<td>7</td>
<td>0.010</td>
<td>8.885</td>
</tr>
<tr>
<td></td>
<td>0.43</td>
<td>7.36</td>
</tr>
<tr>
<td>8</td>
<td>0.015</td>
<td>9.991</td>
</tr>
<tr>
<td></td>
<td>0.58</td>
<td>7.64</td>
</tr>
<tr>
<td>12</td>
<td>0.039</td>
<td>13.658</td>
</tr>
<tr>
<td></td>
<td>1.08</td>
<td>7.90</td>
</tr>
<tr>
<td>16</td>
<td>0.013</td>
<td>13.038</td>
</tr>
<tr>
<td></td>
<td>0.26</td>
<td>5.98</td>
</tr>
<tr>
<td>20</td>
<td>-0.007</td>
<td>10.362</td>
</tr>
<tr>
<td></td>
<td>-0.10</td>
<td>3.35</td>
</tr>
<tr>
<td>24</td>
<td>-0.053</td>
<td>10.430</td>
</tr>
<tr>
<td></td>
<td>-0.67</td>
<td>3.04</td>
</tr>
</tbody>
</table>
Table 4
Long-Run Forecasting Regressions of State Returns over the Market Return

The Table presents the estimation results of the long-run forecasting regressions for the state-return over the Treasury Bill Return. The bold-faced numbers are the estimates, and the numbers underneath them are their t-statistic. The forecasting regression takes the following form: \( r_{K,it} - r_{K,mt} = \alpha_{K,i} + x_{i,t-1}\beta_K + \varepsilon_{it} \), where \( i \) is the index for the U.S. states, \( t \) is the index for the time periods and \( \alpha_i \) is state-specific mean. The variable \( r_{K,it} - r_{K,mt} \) equals \( r_{it} - r_{mt} + r_{i,t+1} - r_{m,t+1} + \ldots + r_{i,t+K} - r_{m,t+K} \). The row vector \( x_{i,t-1} \) contains the predictors \( d_{i,t-1} - p_{i,t-1}, \text{cay}_{t-1}, h y_{i,t-1} \) and the state unemployment rate. The data are quarterly U.S. state data for 1982(Q1) to 1999(Q1). All time series are expressed in real terms using regional inflation rates from the BLS. The base year is 1992(Q1). The quarterly local return is a value-weighted return of firms located in each state. The firm’s location is the location of its headquarter. I included all firms listed in CRSP. The quarterly return of each firm is computed by compounding its monthly returns. The log-real local return is denoted by \( r \). Similarly, the quarterly risk-free return is calculated by compounding monthly returns of the 1-month Treasury bill. The log-real market and risk-free returns are denoted by \( r_m \) and \( r_f \). The cay is from Lettau and Ludvigson (2001). The state-level collateral ratio, \( h y \) uses state log labor income, \( y \), from the Bureau of Economic Analysis (BEA). The time-series of log state housing, \( h \), is from Case, Quigley and Shiller (2001). As in Lustig and Van Nieuwerburgh (2004), the log collateral ratio is equal to \( h = 0.808 \cdot y \) where 0.808 is the PDLs estimate of Mark and Sul (2003) from the cointegrating vector between \( h \) and \( y \). The state unemployment rates are from the BLS. The forecasting regressions are estimated using the LSDV estimator.

<table>
<thead>
<tr>
<th>Horizon K</th>
<th>Forecasting Variables</th>
<th>d - p</th>
<th>cay</th>
<th>hy</th>
<th>unemp rate</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>0.004</td>
<td>-0.055</td>
<td>-0.046</td>
<td>0.004</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.11</td>
<td>-0.40</td>
<td>-3.29</td>
<td>3.71</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.004</td>
<td>0.015</td>
<td>-0.066</td>
<td>0.005</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.75</td>
<td>0.06</td>
<td>-2.76</td>
<td>2.89</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.007</td>
<td>0.128</td>
<td>-0.085</td>
<td>0.005</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.86</td>
<td>0.36</td>
<td>-2.52</td>
<td>2.37</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.012</td>
<td>0.184</td>
<td>-0.106</td>
<td>0.006</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.08</td>
<td>0.35</td>
<td>-2.44</td>
<td>1.89</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.014</td>
<td>0.391</td>
<td>-0.124</td>
<td>0.006</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.09</td>
<td>0.59</td>
<td>-2.32</td>
<td>1.64</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0.018</td>
<td>0.385</td>
<td>-0.136</td>
<td>0.007</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.12</td>
<td>0.47</td>
<td>-2.16</td>
<td>1.53</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>0.022</td>
<td>0.665</td>
<td>-0.150</td>
<td>0.008</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.19</td>
<td>0.68</td>
<td>-2.06</td>
<td>1.52</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>0.028</td>
<td>1.007</td>
<td>-0.159</td>
<td>0.008</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.33</td>
<td>0.92</td>
<td>-1.92</td>
<td>1.45</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>0.043</td>
<td>1.147</td>
<td>-0.168</td>
<td>0.008</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.35</td>
<td>0.75</td>
<td>-1.43</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>0.055</td>
<td>0.335</td>
<td>-0.107</td>
<td>0.003</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.24</td>
<td>0.17</td>
<td>-0.70</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>0.046</td>
<td>-0.008</td>
<td>0.010</td>
<td>0.002</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.83</td>
<td>0.00</td>
<td>0.05</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
<td>0.042</td>
<td>-1.557</td>
<td>0.060</td>
<td>0.000</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.62</td>
<td>-0.53</td>
<td>0.28</td>
<td>0.01</td>
<td></td>
</tr>
</tbody>
</table>