The Household Equity Share and Expected Market Returns*

David C. Yang Fan Zhang

August 2016

Abstract

The "Household Equity Share" ($HEShare$), the share of the household sector's equity and credit assets allocated to equities, is a negative predictor of excess returns on the US stock market. This predictability is robust to the definition of the asset classes, first versus second half of sample, and adjusting for finite sample bias. $HEShare$ outperforms popular forecasters of market returns, including the cyclically adjusted price-earnings ratio, the equity share in new issuances, the consumption-wealth ratio, the term spread, and the Treasury bill yield. Our results suggest that household holdings of financial assets play an important role in setting asset prices. At times, $HEShare$ predicts negative expected excess returns on the market, which suggests that the negative predictability is due to behavioral reasons.

*Contact: `david.yang@uci.edu`. Yang is at the University of California, Irvine, Merage School of Business. Zhang is at PrepScholar Education. This paper is based on a chapter from Zhang's dissertation at Harvard. We thank Malcolm Baker, John Campbell, Robin Greenwood, Scott Joslin, David Laibson, and seminar participants at Harvard, UC Irvine, and the L.A. Finance Day for feedback. The NSF Graduate Research Fellowship provided financial support for this research.

1 Introduction

The household sector is a major holder of U.S. equities. In the Federal Reserve's 2015 Financial Accounts of the United States, the household sector owned roughly 60% of corporate equities outstanding through direct ownership and through equity mutual funds.[1] These two types of holdings are liquid and can readily respond to shifts in household beliefs and preferences. If the other sectors of the economy do not fully offset demand shocks by households, then increases in household holdings of equities will increase equity prices and reduce future equity returns. This paper analyzes this idea empirically and finds evidence that when households own relatively more stock, future excess returns on the stock market are lower.

We create a measure called the Household Equity Share ($HEShare$), which is the share of the household sector's equity and credit assets allocated to equities. $HEShare$ is a fraction that ranges from 0 (the household sector owns credit assets and no equity assets) to 1 (the household sector owns equity assets and no credit assets). We calculate this measure using data from the Federal Reserve's Financial Accounts of the United States.

The U.S. $HEShare$ is a negative predictor of excess U.S. market returns. A one percentage point increase in $HEShare$ forecasts a 0.25% decline in the quarterly (i.e. before annualizing) excess market return, with a t-statistic exceeding 4 in magnitude. In standard deviation terms, a one standard deviation increase in $HEShare$ forecasts a 2% decline in the expected quarterly market excess return. This is a large decline, given that the mean quarterly market excess return in our sample is 1.68%.

This predictability is robust to alternate specifications, including further lagging $HEShare$, splitting the subsample into first-half and second-half, and alternate definitions of "equity assets" and "credit assets." Since $HEShare$ is persistent variable, we further test for the finite sample bias of Nelson and Kim (1993) and Stambaugh (1999). This bias affects our point estimates by about 10%, but the adjusted coefficients remain highly statistically significant

[1] This statistic is only an approximation because the Federal Reserve does not break out the household sector's holdings of corporate equities into publicly traded vs closely held stock (Federal Reserve Financial Accounts of the United States, Table L.223). However, at the aggregate level, the vast majority (85%) of corporate equity is publicly traded. French (2008) augments the Federal Reserve data with other data to adjust for closely held stock as well as other issues in directly applying the Federal Reserve data. He estimates that roughly 50% of US publicly traded common equity is held by direct holdings of households and by mutual funds during his sample period of 1980-2007.

with an adjusted t-statistic exceeding 3.6 in magnitude.

We compare our Household Equity Share variable with other known predictors of excess market returns, including the cyclically adjusted price-equity ratio (Campbell and Shiller, 1988b), the equity share in new issuances (Baker and Wurgler, 2000), the consumption-wealth ratio (Lettau and Ludvigson, 2001), the term spread and the Treasury bill rate (Fama and Schwert, 1977, Campbell, 1987, Fama and French, 1989). The predictive power of $HEShare$ persists, controlling for those variables. In bivariate regressions, none of the other predictors meaningfully affects the economic magnitude or statistical significance of $HEShare$'s return forecasting. When we control for both the cyclically adjusted price-earnings ratio and the equity share of new issues, the marginal effect of $HEShare$ declines, but remains economically and statistically significant. Furthermore, in terms of adjusted r-squared, $HEShare$ outperforms these other forecasting variables in univariate forecasting regression. Whereas the other variables have univariate adjusted r-squared values ranging from 0.5% to 3.0% when forecasting quarterly excess market returns, the univariate adjusted r-squared of $HEShare$ is 5.1% for quarterly excess market returns (14.9% for annual returns).

Our Household Equity Share is related to, but distinct from, the "equity share in new issues" variable of Baker and Wurgler (2000). Both variables take the form of $e/(e+c)$ where e is a variable related to equity securities and c is a variable related to credit securities. Baker and Wurgler (2000) study equity and debt issuances (i.e. flow variables), so e is gross firm issuance of equities and c is gross firm issuance of credit securities. In contrast, we study equity and debt holdings (i.e. level variables), so e is household equity assets and c is household credit assets. Our results are robust to controlling for the Baker and Wurgler (2000) equity share in new issues.

Both behavioral and rational theories can predict that household demand shocks lead to return predictability. On the behavioral side, the financial press often portrays households as "naive money", susceptible to behavioral biases that lead households to hold stocks at the "wrong time". Papers such as Bacchetta, Mertens, and Van Wincoop (2009), Case, Shiller, and Thompson (2012), and Greenwood and Shleifer (2014) find evidence that survey measures of investor beliefs negatively forecast future asset returns. Research has also shown that "naive money" affects individual stock returns in the cross section, for example Frazzini and Lamont (2008); we study the time series predictability of aggregate returns. On the rational side, many consumption based asset pricing models focus on a representative agent, so they do not have a household sector per se, but models such as Campbell and Cochrane (1999) can be adapted to produce rational return predictability from the household equity share.

In its current form, this paper aims to robustly document an empirical fact, and so it does not take a strong stance on the rational vs behavioral debate. However, we do find that, at times, $HEShare$ predicts negative expected excess returns on the market. This suggests the negative predictability may be due to behavioral reasons, since most rational models instead predict a positive equity risk premium. For a rational model to predict a negative equity risk premium, the model would typically need the stock market to hedge aggregate consumption.

Section 2 contains a simple model to formalize the framework. Section 3 describes our data and the construction of $HEShare$. Section 4 analyzes the ability of $HEShare$ to forecast market excess returns and contains various robustness tests. Section 5 compares $HEShare$ with other forecasters of excess market returns. Section 6 concludes.

2 Model

Our model has two time periods, two assets, and two types of investors. The time periods are denoted 0 and 1. The two assets are a risky asset ("stock", i.e. the aggregate stock market) and a risk-free asset. At time 1, the stock pays a single terminal dividend $F + \epsilon$, where $\epsilon \sim N(0, \sigma^2)$. There is a total supply of Q for stock. We normalize the net risk-free rate to 0, by assuming the risk-free asset is elastically supplied at that rate.

The two types of investors are households, denoted with subscript H, and non-households, denoted with subscript N. Investors have constant absolute risk aversion (CARA) utility, with risk tolerance τ_H for the households and τ_N for non-households. Each investor has unit endowment and there are a measure w_H of households and w_N of non-households. Non-households have correct beliefs and hence they demand $x_N = \tau_N \cdot (F - P)$ units of stock. In contrast, households beliefs are potentially biased by sentiment S_H and hence they demand $x_H = \tau_H \cdot (F + S_H - P)$ units of stock. When $S_H > 0$, households are optimistic and when

$S_H < 0$ households are pessimistic.

Solving for the equilibrium, we find that the equilibrium price is:

$$P^* = F + \frac{\tau_H w_H S_H - Q}{\tau_H w_H + \tau_N w_N} \tag{1}$$

and the expected return on the stock is therefore:

$$F - P^* = \frac{Q - w_H \tau_H S_H}{\tau_H w_H + \tau_N w_N} \tag{2}$$

Households have unit endowment so the share of their individual wealth allocated to stocks is $P^* x_H^*$. We study the relationship between the household equity share and expected returns.

Proposition 1. *As sentiment S_H increases, households increase the share allocated to stocks (and decrease the share allocated to the risk-free asset) and expected returns $F - P^*$ decrease. For large S_H, expected returns become negative.*

Intuitively, as sentiment S_H increases, households become more optimistic about stocks. As a result, households hold more stocks and less of the risk-free asset. Non-households have correct beliefs about the value of stock, and they do respond to the household's incorrect beliefs. However, the non-households cannot fully offset the increased demand because they have finite risk tolerance. As a result, in equilibrium, prices rise and expected returns fall. Since non-households can only partially offset the households' optimism, when households are extremely optimistic, prices can be high enough that expected returns are negative.

Proposition 2. *Suppose households have correct beliefs ($S_H = 0$). Then, as risk-tolerance τ_H increases, households increase the share allocated to stock (and decrease the share allocated to the risk-free asset) and expected returns $F - P^*$ decrease. However, expected returns $F - P^*$ cannot be negative.*

Intuitively, as household risk tolerance τ_H increases, households become less concerned with the volatility from holding stock. Hence, households allocate more to stock and less to the risk-free asset. This shift in the demand curve raises the stock price and lowers expected returns. However, because both households and non-households have correct beliefs when

$S_H = 0$, the stock price never rises to the point where expected returns are negative. Most rational theories similarly predict positive expected excess returns on the aggregate stock market, which is a difference between the rational and behavioral explanations.

3 Data and Defining the Household Equity Share

Our main data source is the Federal Reserve's Z.1 Statistical Release, "Financial Accounts of the United States". Before 2013, this Federal Reserve report was known as the "Flow of Funds Accounts of the United States". Released quarterly, the Financial Accounts reports balance sheet information for different sectors of the economy: households and nonprofit institutions serving households, nonfinancial businesses, etc. For each sector, the Financial Accounts reports assets (e.g. treasury securities owned by the household sector) and liabilities (e.g. mortgage borrowing by the household sector).

Our main explanatory variable "Household Equity Share" (HEShare) is the share of the household sector's equity and credit assets allocated to equities:

$$HEShare := \frac{\text{Household Equity Assets}}{\text{Household Equity Assets} + \text{Household Credit Assets}} \tag{3}$$

Hence, when $HEShare = 0$, the household sector holds credit assets, but no equity assets. When $HEShare = 1$, the household sector holds equity assets, but no credit assets.

"Household equity assets" are the sum of equities held by households (series: FL153064105.Q) and equity mutual funds held by households (series: FL153064245.Q). "Household credit assets" are the sum of debt securities held by households (series: FL154022005.Q), loans held by households (series: FL154023005.Q), and bond mutual funds held by households (series: FL153064235.Q). We refer to these as "credit assets" to clearly denote that they are assets, not liabilities, from the perspective of the households. Debt securities are primarily investments in municipal securities, corporate and foreign bonds, and Treasuries. Loans are primarily "other loans and advances", which "includes cash accounts at brokers and dealers and syndicated loans to nonfinancial corporate business by nonprofits and domestic hedge funds." We use both debt securities and loans held by households because the Federal Reserve grouped them together under the heading “Credit Market Instruments” heading before 2015. Our results are robust to the definition of household equity assets and household credit assets, including dropping mutual fund assets and loans.

Strictly speaking, the “households” we study are the sector known as “households and nonprofit institutions serving households”. However, the Federal Reserve uses this grouping as a major sector of the US economy and also informally refers to this sector as “households” in the text of the Financial Accounts of the United States and so we do as well.

One potential question is why we focus on equity and credit holdings either directly owned or owned through a mutual fund by the household sector. The household sector indirectly owns everything in the economy, so where to draw the boundary is a fundamental question for any study of household investment decisions. We focus on direct and mutual fund holdings because these are readily and liquidly traded by households. As a result, these assets respond most strongly to household preferences and beliefs.

By studying household equity and credit assets, we are excluding the following household financial assets: deposits, equity in noncorporate businesses, pension entitlements, and life insurance reserves. Deposits are generally not held for investment, but rather for transactional needs of households. In standard portfolio choice models, investors that want to avoid risk will hold the risk-free asset, not deposits; US Treasuries are included in our definition of household credit assets. Equity in noncorporate businesses are significantly illiquid. Pensions entitlements and life insurance reserves are not easily redeemable and so these assets do not respond to shifts in household risk preference the way mutual fund assets do. We also exclude household nonfinancial assets, which is primarily owner-occupied housing. Homes are a relatively illiquid asset with transaction costs of 6% plus weeks of selling time, preparation, and other opportunity costs. Moreover, homes are a bundled good that reflect preferences for internal amenities and location (e.g. commute times, school districts, etc.). Thus, real estate holdings are a noisy measure of household risk preferences, especially relative to equities that are traded for future returns.

Other data come from standard sources. For stock market returns, we use the returns on the value-weighted market index from the Center for Research in Security Prices (CRSP). We use the return on the 90-day US Treasury bill from CRSP as the risk-free return. The long government bond rate is the 10-year constant maturity Treasury bond rate from the Federal Reserve. The cyclically-adjusted price earnings ratio (Campbell and Shiller, 1988b) is from Robert Shiller's website. The equity share of new issues (Baker and Wurgler, 2000) are from Jeffrey Wurgler's website; as his dataset stops in 2007, we update it using the Federal Reserve's "New Security Issues, US Corporations". The consumption-wealth ratio CAY (Lettau and Ludvigson, 2001) is from Martin Lettau's website. We compute excess market returns as the difference between the stock market returns minus the risk-free return. We compute the term spread as the difference between 10-year Treasury yield and the 90-day Treasury yield. After merging, our dataset spans 1953q2 to 2015q3.

3.1 Descriptive Statistics

Table 1 displays the summary statistics of our dataset. The Household Equity Share has an average value of 0.66, meaning that households allocate about two-thirds of their equity and credit assets to equities. Figure 2 displays the time series of the Household Equity Share (solid line using the left scale) along with excess returns on the value-weighted market index over the next five years (dotted line using the right scale, which has an inverted axis). Since $HEShare$ is a negative predictor of future market returns, we invert the right scale to make the relationship between $HEShare$ and future excess market returns easier to see visually. Households hold a relatively high fraction of equities in the late 1970s and late 1990s, periods that are associated with low excess market returns going forward. In the 1950s and late 2000s, households held a lower fraction of equities and excess market returns were strong going forward. The period of the 1960s and 1980s offers a more mixed relationship. This graph casually displays the relationship of the Household Equity Share and future stock market returns. In the next section, we analyze this relationship more formally with quarterly return data.

Next, we examine the components that makeup household equity assets and household credit assets. On average, the bulk (84%) of household equity assets are directly held corporate equities. Similarly, on average, the bulk (70%) of household credit assets are household holdings of debt securities. Figure 1 plots the time series of the components of household equity assets and the components of household credit assets. In both figures, we can see the rise of equity and bond mutual funds from a tiny fraction in the 1950s (note that both figures use a log scale) to a substantial fraction in the present day. Household holdings of loans roughly grows at the same rate as household holdings of debt securities.

Finally, Figure 3 shows a correlation heatmap of the return forecasting variables. The Household Equity Share is positively correlated with the cyclically adjusted price-earnings ratio (CAPE) and negatively correlated with the consumption-wealth ratio (CAY) and the term spread (TermSpread). It has mild positive correlation with the equity share of new issuances (EquityIssue) and mild negative correlation with the Treasury bill rate (TBill).

4 Forecasting Excess Market Returns using the Household Equity Share

We use a regression framework to formally test the ability of the Household Equity Share ($HEShare$) to forecast excess market returns. Let $R^e = R^{mkt} - R^f$ be quarterly (i.e. before annualizing) excess returns of the CRSP value-weighted market index. We use the two quarter ahead market excess return R^e_{t+2} to avoid concerns about when the Federal Reserve releases the Financial Accounts of the United States. Hence, we use data from, say, 2014Q1 to forecast excess returns in 2014q3. Furthermore, equity prices are a part of the Household Equity Share and a part of measures like the cyclically adjusted price-earnings ratio. If we did not skip a period, measurement error in equity prices today P_t could induce artificial predictability, since P_t is a part of both R^e_{t+1} and $HEShare_t$ and $CAPE_t$. Since we use quarterly returns, return periods do not overlap.

4.1 Univariate Regressions

We first run the univariate regression:

$$R^e_{t+2} = b_0 + \lambda \cdot HEShare_t + \epsilon_{t+2} \tag{4}$$

For inference, we use Newey and West (1987) standard errors with five periods of lags. For the lag-length, we use the rule of thumb of $\frac{3}{4} \cdot T^{(1/3)}$ with $T = 249$ quarters rounded to the nearest integer, as suggested by Newey and West (1994). Varying the lag length from one to ten quarters yields similar results. The Newey-West procedure includes the correction for heteroskedasticity (White, 1980) and further accounts for autocorrelation of error terms by using a triangle/Bartlett kernel for the time series correlation structure.

Table 2 displays our univariate regression results, with standard errors in parentheses. Regression (1) shows the univariate return forecasting ability of $HEShare$. We estimate $\hat{\lambda} = -0.25$, implying that a one percentage point increase in $HEShare$ is associated with a 0.25% decline in quarterly excess returns on the market. We can also restate the economic magnitude in terms of standard deviations. $HEShare$ has a standard deviation of 8%, so a one standard deviation increase in $HEShare$ is associated with a 2% decline in the expected quarterly market excess return. As the mean quarterly market excess return in our sample is 1.68%, this is a large decline. The t-statistic exceeds 4 in magnitude and the adjusted r-squared exceeds 5%.

Our effect is robust to further lagging $HEShare$. In Regression (2), Regression (3), and Regression (4), we use lags of $HEShare$ ranging from two to six quarter lags. For example, Regression (2) uses $HEShare_{t-2}$ to forecast R^e_{t+2}. Each successive two quarter lag lowers the economic magnitude of our effect by roughly 0.05. So, in the baseline regression with no lags, $\hat{\lambda}_{nolag} = -0.25$ whereas $\hat{\lambda}_{lag2q} = -0.20$, $\hat{\lambda}_{lag4q} = -0.16$, and $\hat{\lambda}_{lag6q} = -0.10$. The statistical significance declines with successive lags, with four quarters of lags $HEShare_{t-4}$ still being a statistically significant forecaster of excess market returns at the 5% p-value level.

We check if particular subsamples drive our result. Regression (5) and Regression (6) split our sample into the first-half and second-half. Our effect remains significant at the 1% level, suggesting that the effect is robust over time. We can also inspect for outliers visually, using the scatter plot of Figure 4. In that figure, we use the convention of time t for excess returns and time $t-2$ for $HEShare$. We use this timing convention for the figure so that the time labels in the graph refer to returns as opposed to $HEShare$, e.g. 2008q4 returns (which most people associate with strongly negative returns) rather than 2008q4 $HEShare$. From visual inspection, we see that potential negative outliers are 1974q4, 1987q4, 2008q4, and 2009q3. However, removing those potential outliers in fact slightly strengthens the negative relationship.

4.2 Forecasting with Alternate Definitions of $HEShare$

We next show that alternative definitions of the Household Equity Share lead to similar forecasting results. Recall from Section 3 that we defined the Household Equity Share as:

$$HEShare := \frac{\text{Household Equity Assets}}{\text{Household Equity Assets} + \text{Household Credit Assets}} \tag{5}$$

"Household equity assets" are the sum of equities held by households and equity mutual funds held by households. "Household credit assets" are the sum of debt securities held by households, loans held by households, and bond mutual funds held by households. Debt securities are primarily investments in municipal securities, corporate and foreign bonds, and Treasuries. Loans are primarily "other loans and advances", which "includes cash accounts at brokers and dealers and syndicated loans to nonfinancial corporate business by nonprofits and domestic hedge funds." As explained in Section 3, we use both debt securities and loans held by households because the Federal Reserve grouped them together under the heading "credit market instruments" heading before 2015.

In Table 3, we consider alternate definitions for the Household Equity Share. Regression (1) displays the baseline regression results using the above definition. Regression (2) excludes household holdings of equity and bond mutual funds. This test is important because household holdings of mutual funds have rising meaningfully since 1980 (Figure 1). Regression (3) excludes household holdings of loan assets. Regression (4) excludes both household holdings of equity and bond mutual funds and household holdings of loan assets. These alternate specifications all lead to very similar estimates in both economic magnitude $\hat{\lambda} \approx -0.25$ and statistical significance $|t| \approx 4$.

4.3 Adjusting for Finite Sample Bias

In return forecasting regressions, persistence in the predictor variable can create finite sample bias, see Nelson and Kim (1993) and Stambaugh (1999). Extending our paper's notation, we describe the finite sample bias as follows:

$$R_{t+2} = b_0 + \lambda \cdot HEShare_t + \epsilon_{t+2} \quad (6)$$

$$HEShare_{t+2} = c_0 + \phi \cdot HEShare_t + \eta_{t+2} \quad (7)$$

Equation 6 is the predictive regression we have run thus far. We maintain the skip a quarter convention in both equations, forecasting time $t+2$ variables using time t variables. The variable ϕ from Equation 7 measures the persistence of the predictor variable $HEShare$. The finite sample bias is that $\hat{\lambda}$ has the following bias:

$$E[\hat{\lambda} - \lambda] = \frac{\sigma_{\epsilon\eta}}{\sigma_\eta} E[\hat{\phi} - \phi] \quad (8)$$

Kendall (1954) emphasizes that $\hat{\phi}$ has a bias of approximately $E[\hat{\phi} - \phi] \approx -\frac{1+3\phi}{T}$, so the bias in estimating the persistence of $HEShare$ translates into a bias of estimating the predictor power of $HEShare$ on future excess returns.

Stambaugh (1999) suggests an approximate way to correct for this bias is to adjust the point estimate $\hat{\lambda}$ using Equation 8. In our data, we estimate: $\hat{\phi} = 0.911$ and $\frac{\hat{\sigma}_{\epsilon\eta}}{\hat{\sigma}_\eta} = 1.75$. Therefore, our bias is roughly $E[\hat{\lambda} - \lambda] = 1.75 \cdot (-0.015) = -0.026$. This calculation suggests that our estimates in Table 2 are biased by roughly 10%. This correction affects our results in Table 2, but $HEShare$ is still a highly statistically significant predictor of future excess returns because the unadjusted t-statistics exceed 4 in magnitude. For example, in our baseline regression of Table 2 Regression (1), the Stambaugh (1999) correction implies that $\hat{\lambda}_{adj} = -0.253 + 0.026 = -0.227$ and $t = -3.64$, which is still highly statistically significant with $p = 0.0003$.

Furthermore, we can establish an upper bound (lower bound, in magnitude) on λ if we assume that $HEShare$ does not have a unit root. Because $HEShare$ is a fraction, it must be between 0 and 1. Hence, it is reasonable to assume it does not grow explosively, i.e. with a unit root. Lewellen (2004) observes the conditional relationship:

$$E[\hat{\lambda} - \lambda | \hat{\phi}] = \frac{\sigma_{\epsilon\eta}}{\sigma_\eta} \cdot (\hat{\phi} - \phi)$$

While we do not know ϕ, if the predictor does not have a unit root, then the bias is greatest when $\phi = 1$. This observation establishes an upper bound (lower bound, in magnitude) of $\hat{\lambda}_{adj} = -0.253 - 1.75 \cdot (0.911 - 1) = -0.0972$. Lewellen (2004) also establishes the standard error for his bias-adjusted estimator, which in our application equals 0.0495. Therefore, we can establish an upper bound (lower bound in magnitude) on the t-statistic of $t = -1.96$ and an upper bound on the p-value of 0.051. As the bound uses the worst-case bias, it suggests that the true value of λ is in fact negative and $HEShare$ does negatively predict future excess market returns.

4.4 Long-Horizon Regressions

Here, we examine the effect of forecasting returns at longer horizons. Let $R^e_{t+1\rightarrow t+k+1}$ denote the annualized market excess return from $t+1$ to $t+k+1$. We continue to skip a quarter between $HEShare_t$ and the forecasted returns. We examine returns one year ahead ($k = 4$ quarters), three years ahead ($k = 12$ quarters) and five years ahead ($k = 20$ quarters). We run the regression:

$$R^e_{t+1\rightarrow t+k+1} = b_0 + \lambda_k \cdot HEShare_t + \epsilon_{t+k+1} \quad (9)$$

In this regression, the dependent variable now overlaps, which creates serial correlation in the errors. To estimate standard errors, we use two methods: The first method is to adjust for the overlapping returns by using Newey-West standard errors with $k+5$ quarters of lags, so that the lag length increases with the return horizon k. Adjusting for the overlapping returns econometrically has the benefit of using all the observations in our dataset. However, this type of adjustment can sometimes lead to spurious long-run predictability (Ang and Bekaert, 2007). Hence, we supplement it with a second method of dropping data to eliminate the overlapping returns. For example, for the one year ahead regression, we only keep the January observation and discard the other observations that year. This method is econometrically inefficient since it discards data, but it has the benefit of directly avoiding the overlapping returns concern.

Table 5 displays the results. Regressions (1), (3), (5) show the effect on forecasting returns one year ahead (4 quarters), three years ahead (12 quarters) and five years ahead (20 quarters), using Newey-West standard errors with $k+5$ quarters of lags. Regressions (2), (4), and (6) show the effect of estimating returns the same time horizons ahead, but instead estimate standard errors using non-overlapping returns. As expected, the non-overlapping return regressions show slightly lower statistical significance, since we are discarding data by lowering the data frequency.

First, we examine Table 5 Regressions (1) and (2), which forecasts one year ahead returns. Both methods yield similar coefficients $\hat{\lambda}_{4qtr,overlap} = -0.886$ and $\hat{\lambda}_{4qtr,noOverlap} = -0.797$. This coefficient size is roughly 3-4x the quarterly coefficient $\hat{\lambda}_{1qtr} = -0.253$ (Table 2 Regression (1)). We also observe that the adjusted r-squared rises with the horizon with $AdjR2_{4qtr,overlap} = 0.149$ and $AdjR2_{4qtr,noOverlap} = 0.121$.

Both of these effects are closely related to the quarterly frequency return regressions in Table 2, since long-horizon returns are an accumulation of short-horizon returns and $HEShare$ is a persistent variable. Whether or not long-horizon regressions have more statistical power than short-horizon regressions is a debate that we do not re-visit here, see Campbell (2001), Valkanov (2003), Boudoukh, Richardson, and Whitelaw (2008), Cochrane (2008). Regardless of its statistical properties, the long-horizon regression has the advantage of being directly interpretable if our interest in the longer term. For example, papers on other return forecasters sometimes focus on annual returns. The results in Table 5 Regressions (1) allow us to directly state that a 1% increase in $HEShare$ forecasts a 0.88% decline in average excess returns over the following year and that $HEShare$ explains 14.9% of the variation in annual returns, which allows for easy comparison.

Next, we examine the effect of further lengthening the return horizon. The economic magnitude declines to $\hat{\lambda}_{12qtr,overlap} = -0.601$ at the three year horizon (Table 5 Regression (3)) and to $\hat{\lambda}_{20qtr,overlap} = -0.534$ at the five year horizon (Table 5 Regression (5)); the results of using the non-overlapping returns are highly similar so we focus the discussion in this paragraph around the overlapping return regressions. The declining coefficient shows that the forecasting ability of $HEShare$ declines with the horizon. This decline in forecasting ability is the same as the observation in Table 2 Regressions (2), (3), and (4) that as we increase the time gap between $HEShare$ and the forecasted returns (e.g. using 2015q4's $HEShare$ to forecast the quarterly returns in 2016q2 versus the quarterly returns in 2016q4), the economic magnitude and statistical significance of the return forecastability falls. The adjusted r-squareds also continue to increase with the return horizon with $AdjR2_{12qtr,overlap} = 0.279$ and with $AdjR2_{20qtr,noOverlap} = 0.407$.

5 Comparison with Other Return Forecasters

In this section, we compare $HEShare$ with known forecasters of excess market returns. In particular, we consider $CAPE$, the ten-year cyclically adjusted price-to-earnings-ratio (Campbell and Shiller, 1988b); $EquityIssue$, the equity share in new issuances (Baker and Wurgler, 2000); CAY, the consumption-wealth ratio (Lettau and Ludvigson, 2001); $TermSpread$, the yield spread between the 10-year US Treasury and 90-day US Treasury; and $TBill$, the 90-day US Treasury yield.

We first compare the univariate return forecasting regressions:

$$R^e_{t+2} = b_0 + b_1 \cdot X_t + \epsilon_{t+2} \tag{10}$$

where X_t is a return forecaster, e.g. $HEShare$, $CAPE$, etc. Table 6 contains the results of the univariate comparison with other return forecasters. Each column in this table displays two regressions: one with the regressors unadjusted (b_1) and one with the regressors normalized to have unit variance (b_1^{norm}). We examine the different return forecasters along three dimensions: economic magnitude, statistical significance, and r-squared. Amongst the return forecasters we consider, $HEShare$ performs the best along all three dimensions and CAY performs the next best. The economic magnitude of $HEShare$ is $|b^{norm}_{1,HEShare}| = 0.020$ versus $|b^{norm}_{1,CAY}| = 0.015$ for CAY. These coefficients imply that a one standard deviation change in $HEShare$ forecasts a 2% change in mean excess quarterly returns. In contrast, a one standard deviation change in CAY forecasts a 1.5% change in mean quarterly excess returns. The statistical significance of $HEShare$ is $|t_{HEShare}| = 4.0$ versus $|t_{CAY}| = 2.9$ for CAY. The r-squared for $HEShare$ is $R^2_{HEShare} = 5.1\%$ quarterly versus $R^2_{CAY} = 3.0\%$ quarterly.

We next examine the effect of controlling for other return forecasters using the regression:

$$R^e_{t+2} = b_0 + \lambda \cdot HEShare_t + b_1 \cdot X_t + \epsilon_{t+2} \tag{11}$$

Due to space constraints, we split the results into Table 7 (bivariate regression results) and Table 8 (multivariate regression results). In both tables, Regression (1) displays the baseline univariate results.

Table 7 Regression (2) controls for $CAPE$, the cyclically adjusted price-earnings ratio (Campbell and Shiller, 1988b), which is a known negative forecaster of equity returns. By controlling for $CAPE$, we address the potential concern that movements in $HEShare$ may reflect movements in valuation ratios, which are known to forecast excess market returns. From the regression, we see that controlling for CAPE does not affect the coefficient on $HEShare$ $\hat{\lambda} = -0.25$ and the statistical significance only declines marginally. We use CAPE, instead of dividend yield (Fama and French, 1988, Campbell and Shiller, 1988a), because CAPE is unaffected by the trend of corporations to favor buybacks, as opposed to dividends, in recent years. In an undisplayed regression, we confirm that controlling for dividend yield gives similar results.

Table 7 Regression (3) controls for $EquityIssue$, the equity share of new issues (Baker and Wurgler, 2000), which is a known negative forecaster of equity returns. While $HEShare$ is a level variable, $EquityIssue$ is a flow variable that measures the proportion of equity and debt issuances that went to equities. Despite this level versus flow difference, we could potentially be concerned that perhaps the forecasting power of $HEShare$ comes from households purchasing the equity that corporations are issuing. From Regression (3), we see that is not the case. Adding $EquityIssue$ only changes the coefficient on $HEShare$ slightly to $\hat{\lambda} = -0.24$ and the statistical significance remains $t = -3.6$. $EquityIssue$ itself is statistically insignificant in this bivariate regression. This contrasts with the univariate regression of future excess market returns on $EquityIssue$ alone , which yields $t_{EquityIssue} = -1.89$ (Table 6, Regression (3)).

Table 7 Regression (4) controls for CAY, the consumption-wealth ratio (Lettau and Ludvigson, 2001). Controlling for CAY causes a modest decline in the economic magnitude effect of $HEShare$ ($\hat{\lambda} = -0.21$), but $HEShare$ remains statistically significant ($t = -2.8$). CAY is statistically insignificant in this bivariate regression. This contrasts with the univariate regression of future excess market returns on CAY alone ($b_{1,CAY} = 0.78$ and $t_{CAY} = 2.95$ from Table 6 Regression (4)). $HEShare$ appears to absorb the forecasting ability of CAY, so that the marginal effect of CAY is statistically insignificant after controlling for $HEShare$.

Table 7 Regression (5) controls for $TermSpread$, the yield spread the 10-year US Treasury and 90-day US Treasury (Campbell, 1987, Fama and French, 1989). Table 7 Regression (6) controls for $TBill$, the yield on the 90-day US Treasury Bill (Fama and Schwert, 1977, Campbell, 1987). Controlling for $TermSpread$ and $TBill$ marginally increases the effect of $HEShare$ ($\hat{\lambda} = -0.26, -0.27$). $HEShare$ remains statistically significant at the 1% level across both specifications. In this sample, $TermSpread$ and $TBill$ alone are not statistically significant forecasters of excess market returns, see Table 6 Regressions (5) and (6). However, when combined with $HEShare$, the marginal effect of $TBill$ becomes statistically significant with $t = -2.0$.

We can also examine the adjusted r-squareds. The univariate regression has an adjusted r-squared of 5.1%. We see that adding $EquityIssue$ or $TBill$ improves the adjusted r-squared to 5.4% and 6.1%, respectively. However, adding $CAPE$, CAY, or $TermSpread$ decreases the adjusted r-squared.

In Table 8, we examine multivariate comparisons with the other return forecasters. First, we examine Table 8 Regression (2), which controls for $CAPE$ and $EquityIssue$ together. When we do so, the coefficient on $HEShare$ declines in economic magnitude from $\hat{\lambda} = -0.253$ to $\hat{\lambda} = -0.16$ and in statistical significance from $t = -4.0$ to $t = -2.15$. Furthermore, the marginal effect of $EquityIssue$ becomes significant at the 5% p-value level in this multivariate regression. Hence, part of $HEShare$'s predictive power comes from the joint combination of $CAPE$ and $EquityIssue$. However, even after these controls, $HEShare$ remains statistically significant at the 5% p-value level. Next, in Table 8 Regression (3), we examine the combination of CAY, $TermSpread$, and $TBill$. We find that $HEShare$ remains statistically significant with a similar coefficient $\hat{\lambda} = -0.28$ to its univariate forecasting in Table 8 Regression (1).

When we control for all five variables (Table 8 Regression (4)), we see a different effect. First, the coefficient on $HEShare$ is $\hat{\lambda} = -0.16$ when we control for all five return forecasters. Given the results of Table 8 Regression (2), we conclude that most of this decline in economic magnitude is due to adding $CAPE$ and $EquityIssue$. Second, the statistical significance of $\hat{\lambda}$ falls to the 10% p-value level. This decline is likely due to adding too many controls, given we only have 249 quarters of data. We see that the point estimate with all the controls ($\hat{\lambda} = -0.16$) is similar to the point estimate with just $CAPE$ and $EquityIssue$ ($\hat{\lambda} = -0.17$). However, once we also control for CAY, $TermSpread$, and $TBill$, the standard error increases by enough to make $\hat{\lambda}$ only statistically significant at the 10% level.

6 Conclusion

This paper shows that when households tilt their portfolios toward equities, future market excess returns are lower on average. We define the Household Equity Share, which is the share of the household sector's equity and credit assets allocated to equities, and show that it is a robust negative predictor of the excess returns on the aggregate stock market. The univariate t-statistic exceeds 4.0 in magnitude and the quarterly adjusted r-squared exceeds 5%. The predictive power remains even after varying the definition of the Household Equity Share, splitting the sample into first-half/second-half, and adjusting for finite sample bias due to a persistent return forecaster. The predictive power also is not subsumed by popular predictors, including the cyclically adjusted price earnings ratio, equity shares of new issuances, the consumption-wealth ratio, the term spread, and the Treasury bill yield. These results suggest a close relationship between household holdings of financial assets and asset prices. At times, $HEShare$ predicts negative expected excess returns on the market, which suggests that the negative predictability may be due to behavioral reasons.

A Proofs

The following two lemmas are useful to state upfront.

Lemma 3. *Comparative statics for P^*: As S_H increases, P^* increases. As τ_H increases, P^* increases if the sentiment S_H is not too pessimistic (specifically $S_H > \frac{-Q}{\tau_N w_N}$).*

Lemma 4. *Comparative statics for x_H^*: As S_H increases, x_H^* increases. As τ_H increases, x_H^* increases if the sentiment S_H is not too pessimistic (specifically $S_H > \frac{-Q}{\tau_N w_N}$).*

Proof of Lemma 3: We have that $P^* = F + \frac{\tau_H w_H S_H - Q}{\tau_H w_H + \tau_N w_N}$. Therefore,

$$\frac{\partial P^*}{\partial S_H} = \frac{w_H \tau_H}{\tau_H w_H + \tau_N w_N} > 0$$

and

$$\frac{\partial P^*}{\partial \tau_H} = \frac{(Q + S_H \tau_N w_N) w_H}{(\tau_H w_H + \tau_N w_N)^2}$$

If $S_H > \frac{-Q}{\tau_N w_N}$, then $\frac{\partial P^*}{\partial \tau_H} > 0$.

Proof of Lemma 4: We have that $x_H^* = \tau_H \cdot (F + S_H - P^*)$.

Therefore,

$$\begin{aligned}\frac{\partial x_H^*}{\partial S_H} &= \tau_H (1 - \frac{\partial P^*}{\partial S_H}) \\ &= \frac{\tau_H \tau_N w_N}{\tau_H w_H + \tau_N w_N} > 0\end{aligned}$$

and

$$\frac{\partial x_H^*}{\partial \tau_H} = \tau_N w_N \frac{Q + S_H \tau_N w_N}{(\tau_H w_H + \tau_N w_N)^2} > 0$$

If $S_H > \frac{-Q}{\tau_N w_N}$, then $\frac{\partial x_H^*}{\partial \tau_H} > 0$.

Proof of Proposition 1: The fraction of household wealth allocated to stocks is $P^* x_H^*$. Applying Lemma 3 and 4, we can conclude:

$$\frac{\partial (P^* x_H^*)}{\partial S_H} = x_H^* \frac{\partial P^*}{\partial S_H} + P^* \frac{\partial x_H^*}{\partial S_H} > 0$$

Also, since $\frac{\partial P^*}{\partial S_H} > 0$, we have $\frac{\partial(F-P^*)}{\partial S_H} < 0$. Finally, when $S_H > \frac{Q}{w_H \tau_H}$ we have that $F - P^* < 0$.

Proof of Proposition 2: This proof is similar to the proof of proposition 1. Throughout this proof, we assume $S_H = 0$, as in the proposition statement.

The fraction of household wealth allocated to stocks is $P^* x_H^*$. Applying Lemma 3 and 4, we can conclude:

$$\frac{\partial(P^* x_H^*)}{\partial \tau_H} = x_H^* \frac{\partial P^*}{\partial \tau_H} + P^* \frac{\partial x_H^*}{\partial \tau_H} > 0$$

Also, since $\frac{\partial P^*}{\partial \tau_H} > 0$, we have $\frac{\partial(F-P^*)}{\partial \tau_H} < 0$. When $S_H = 0$, then

$$F - P^* = \frac{Q}{\tau_H w_H + \tau_N w_N} > 0$$

so expected returns must be positive.

References

Ang, Andrew, and Geert Bekaert, 2007, Stock return predictability: Is it there?, *Review of Financial Studies* 20, 651–707.

Bacchetta, Philippe, Elmar Mertens, and Eric Van Wincoop, 2009, Predictability in financial markets: What do survey expectations tell us?, *Journal of International Money and Finance* 28, 406–426.

Baker, Malcolm, and Jeffrey Wurgler, 2000, The equity share in new issues and aggregate stock returns, *The Journal of Finance* 55, 2219–2257.

Boudoukh, Jacob, Matthew Richardson, and Robert F Whitelaw, 2008, The myth of long-horizon predictability, *Review of Financial Studies* 21, 1577–1605.

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

———, 2001, Why long horizons? a study of power against persistent alternatives, *Journal of Empirical Finance* 8, 459–491.

———, 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.

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

———, 1988b, Stock prices, earnings, and expected dividends, *The Journal of Finance* 43, 661–676.

Case, Karl E, Robert J Shiller, and Anne Thompson, 2012, What have they been thinking?: Home buyer behavior in hot and cold markets, *Brookings Papers on Economic Activity*.

Cochrane, John H, 2008, The dog that did not bark: A defense of return predictability, *Review of Financial Studies* 21, 1533–1575.

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

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

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

Frazzini, Andrea, and Owen A Lamont, 2008, Dumb money: Mutual fund flows and the cross-section of stock returns, *Journal of Financial Economics* 88, 299–322.

French, Kenneth R, 2008, Presidential address: The cost of active investing, *The Journal of Finance* 63, 1537–1573.

Greenwood, Robin, and Andrei Shleifer, 2014, Expectations of returns and expected returns, *Review of Financial Studies* 27, 714–746.

Kendall, Maurice G, 1954, Note on bias in the estimation of autocorrelation, *Biometrika* 41, 403–404.

Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns, *The Journal of Finance* 56, 815–849.

Lewellen, Jonathan, 2004, Predicting returns with financial ratios, *Journal of Financial Economics* 74, 209–235.

Nelson, Charles R, and Myung J Kim, 1993, Predictable stock returns: The role of small sample bias, *The Journal of Finance* 48, 641–661.

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

——— , 1994, Automatic lag selection in covariance matrix estimation, *The Review of Economic Studies* 61, 631–653.

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

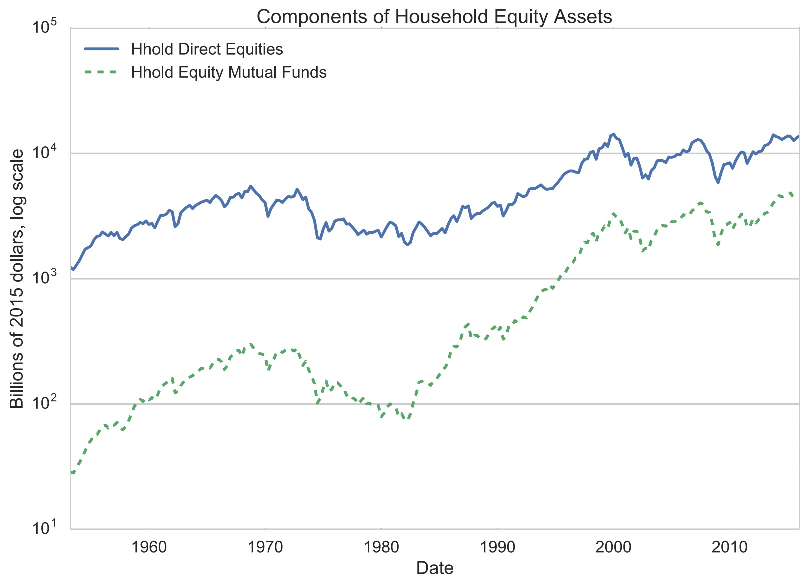
Valkanov, Rossen, 2003, Long-horizon regressions: Theoretical results and applications, *Journal of Financial Economics* 68, 201–232.

White, Halbert, 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* pp. 817–838.

List of Figures

1 Components of Household Equity Assets and Household Credit Assets 24
2 Time Series Plot of $HEShare$ and Future 5-Year Excess Market Returns . . . 25
3 Correlation Heatmap between Return Forecasting Variables 26
4 Scatter Plot and Binscatter of Excess Market Returns vs Household Equity Share 27

Figure 1: Components of Household Equity Assets and Household Credit Assets
The Household Equity Share ($HEShare$) is defined as (Household Equity Assets)/(Household Equity Assets + Household Credit Assets). The two figures here show the breakdown of the components of Household Equity Assets and the components of Household Credit Assets.
(a) The figure below plots the components of Household Equity Assets = Household Directly Holdings of Equities ("Hhold Direct Equities") + Household Holdings of Equity Mutual Funds ("Hhold Equity Mutual Funds").



(b) The figure below plots the components of Household Credit Assets = Household Holdings of Debt Securities ("Hhold Debt Securities") + Household Holdings of Loans ("Hhold Loans") + Household Holdings of Bond Mutual Funds ("Hhold Bond Mutual Funds"). Until 2015, the Federal Reserve grouped "debt securities" and "loans" together under "credit market instruments".

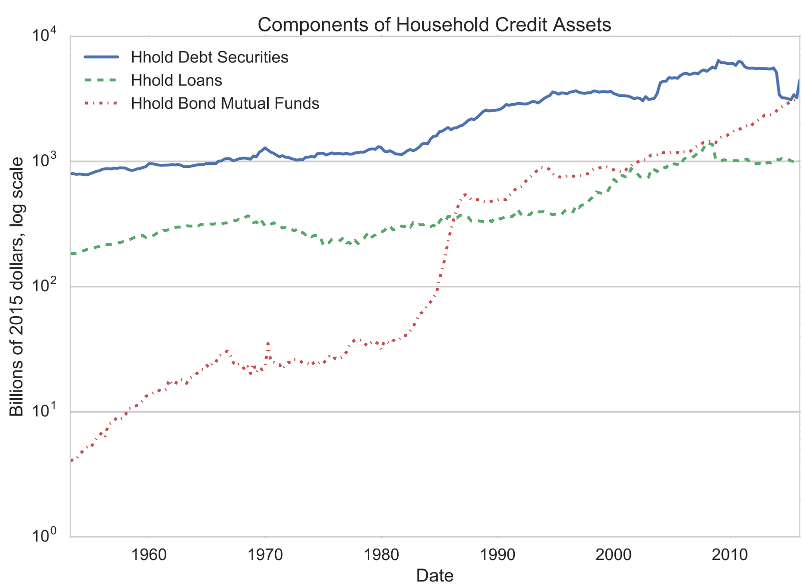


Figure 2: Time Series Plot of $HEShare$ and Future 5-Year Excess Market Returns
This figure plots the time series our main explanatory $HEShare$ and future excess market returns. The Household Equity Share ($HEShare$) is the share of the household sector's equity and credit assets allocated to equities, calculated using data from the Federal Reserve's Financial Accounts of the United States. The blue solid line (left scale) plots $HEShare$. The red dotted line (right scale, inverted axis) plots future 5-year excess market returns, which is defined as the annualized returns of the CRSP value-weighted market index less the return on the 90-day Treasury bill. Since higher $HEShare$ forecasts lower future returns, we invert the axis for the excess market returns to make the relationship easier to see visually.

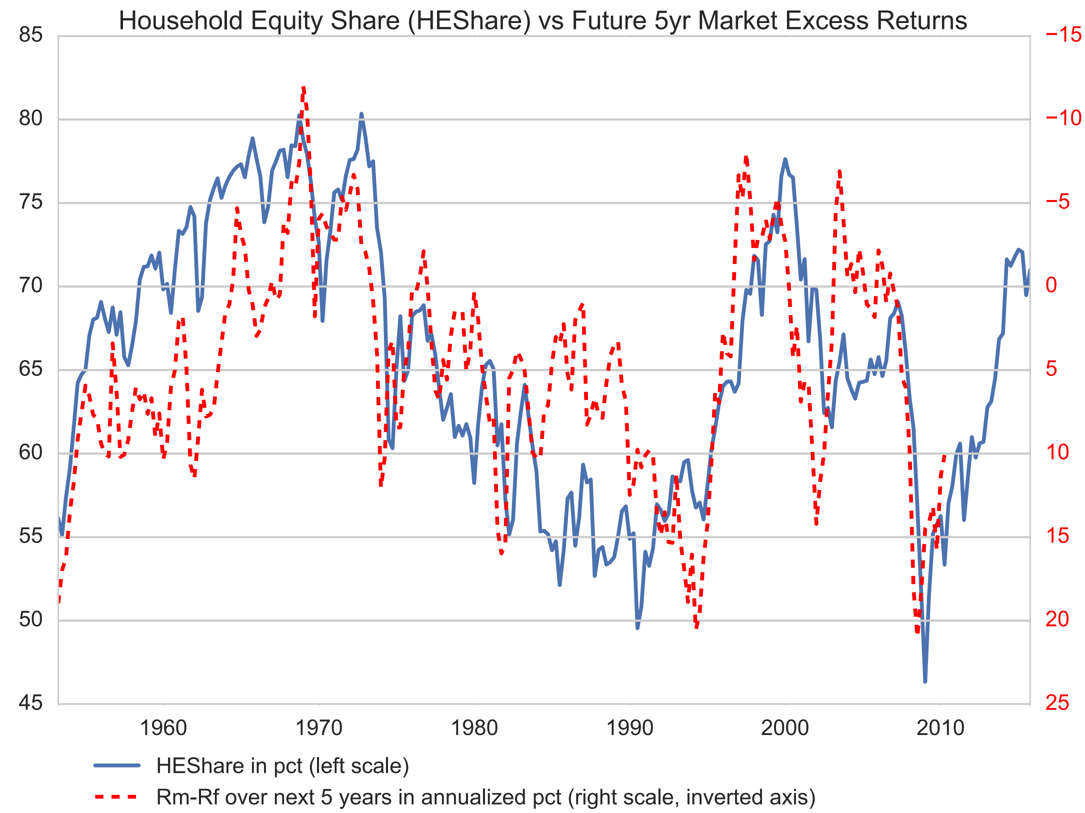


Figure 3: Correlation Heatmap between Return Forecasting Variables

This figure shows a heatmap of the correlations between the return forecasting variables we examine. In the heatmap, white corresponds to a correlation of zero, deep purple corresponds to a correlation of +1, and deep orange corresponds to a correlation of -1. The Household Equity Share ($HEShare$) is the share of the household sector's equity and credit assets allocated to equities, calculated using data from the Federal Reserve's Financial Accounts of the United States. $CAPE$ is the ten-year cyclically adjusted price-to-earnings-ratio (Campbell and Shiller, 1988b). $EquityIssue$ is the equity share in new issuances (Baker and Wurgler, 2000). CAY is the consumption-wealth ratio (Lettau and Ludvigson, 2001). $TermSpread$ is the yield spread between 10-year US Treasury and 90-day US Treasury. $TBill$ is the 90-day US Treasury yield.

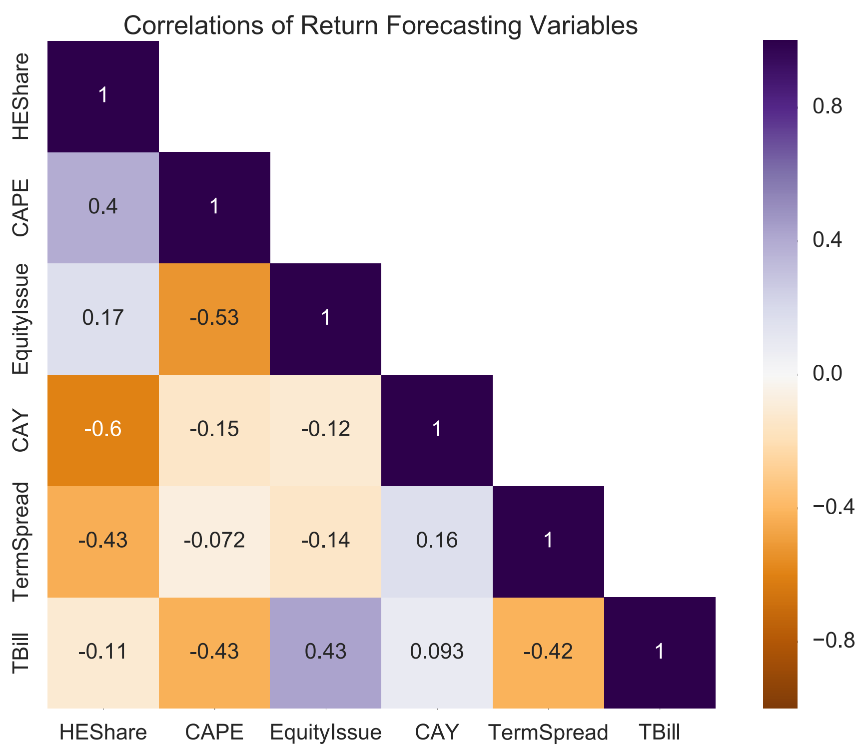
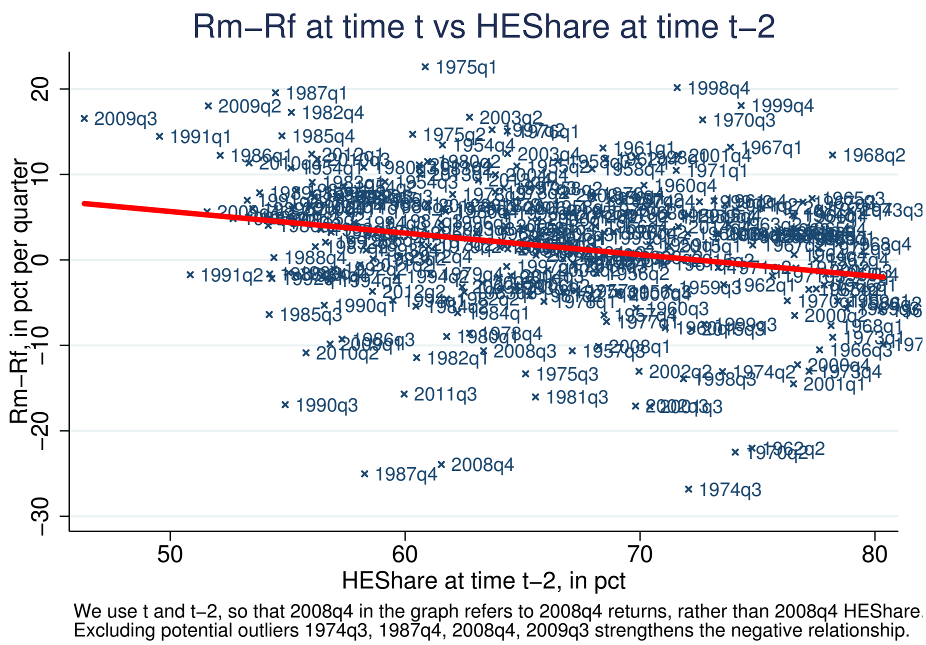
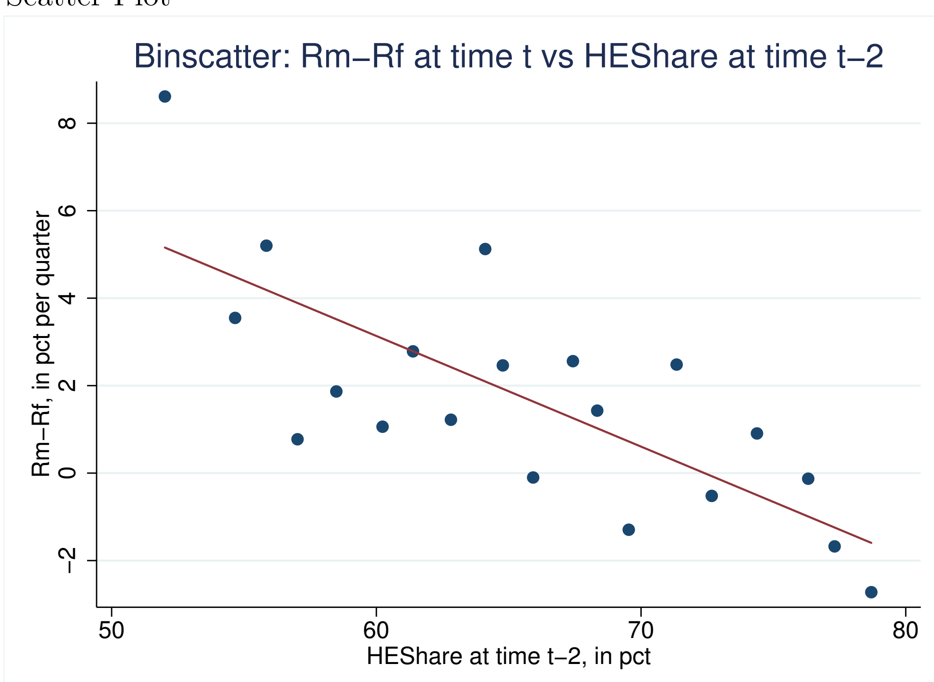


Figure 4: Scatter Plot and Binscatter of Excess Market Returns vs Household Equity Share
This figure shows the scatter plot and binned scatter plot of excess market returns at time t versus $HEShare$ at time $t-2$. The binned scatter plot splits $HEShare$ into 20 equal-sized bins, then plots the mean of $HEShare$ and the mean of excess market returns of each bin. We denote the returns as t (as opposed to $t+2$ as in the regression tables), so that the time labels in the graph refer to returns as opposed to $HEShare$, e.g. so the 2008q4 data label refers to returns in 2008q4. The Household Equity Share ($HEShare$) is the share of the household sector's equity and credit assets allocated to equities.

(a) Scatter Plot



(b) Binned Scatter Plot



List of Tables

1 Summary Statistics 29
2 Univariate Regressions with the Household Equity Share 30
3 Alternate Definitions of the Household Equity Share 31
4 Adjusting for Finite Sample Bias 32
5 Long Horizon Regressions 33
6 Univariate Comparison with Other Forecasters of Excess Market Returns . . . 34
7 Bivariate Comparison with Other Forecasters of Excess Market Returns 35
8 Multivariate Comparison with Other Forecasters of Excess Market Returns . . 36

Table 1: Summary Statistics

This table displays summary statistics of our main variables. The Household Equity Share ("HEShare") is the share of the household sector's equity and credit assets allocated to equities, calculated using data from the Federal Reserve's Financial Accounts of the United States. "Hhold Direct Equities" measures corporate equities directly owned by households. "Hhold Equity Mutual Funds" measures equity mutual funds holdings of households. Before 2015, the Federal Reserve grouped household asset holdings of debt securities ("Hhold Debt Securities") and household asset holdings of loans ("Hhold Loans") together under the heading of household holdings of "credit market instruments". "Hhold Bond Mutual Funds" measures bond mutual funds holdings of households. $CAPE$ is the ten-year cyclically adjusted price-to-earnings-ratio (Campbell and Shiller, 1988b). $EquityIssue$ is the equity share in new issuances (Baker and Wurgler, 2000). CAY is the consumption-wealth ratio (Lettau and Ludvigson, 2001) and is multiplied by 100 to make it visible at two decimal places. The Term Spread is the yield spread between the 10-year US Treasury and 90-day US Treasury. Rm is the return on the value-weighted CRSP market index; the units are percent per quarter as our data have quarterly frequency. Rf is the return on the 90-day Treasury Bill; the units are also percent per quarter. Dollar figures are inflation-adjusted to 2015 dollars. $N = 249$.

	mean	std	25%	50%	75%
Hhold Equity Share (HEShare)	0.66	0.08	0.59	0.65	0.72
Hhold Direct Equities ($ billions)	5,513.22	3,507.07	2,730.30	4,125.44	8,265.58
Hhold Equity Mutual Funds ($ billions)	1,066.70	1,320.77	134.55	272.17	2,144.73
Hhold Debt Securities ($ billions)	2,410.21	1,634.00	1,052.01	1,592.16	3,481.60
Hhold Loans ($ billions)	483.76	306.81	280.80	333.94	713.59
Hhold Bond Mutual Funds ($ billions)	573.94	735.04	23.92	82.46	893.12
CAPE	19.36	7.53	14.36	18.84	23.29
EquityIssue	0.18	0.08	0.11	0.16	0.22
CAY (x100)	-0.06	1.98	-1.26	0.05	1.42
Term Spread (pct per year)	1.48	1.20	0.57	1.48	2.42
90-day TBill Yield (pct per year)	4.48	3.10	2.43	4.46	6.00
Rm (pct per quarter)	2.89	8.30	-1.60	3.75	8.08
Rf (pct per quarter)	1.21	0.84	0.64	1.18	1.61

Table 2: Univariate Regressions with the Household Equity Share

This table displays our univariate regression results. R^e_{t+2} is the quarterly excess return of the value-weighted market index, two quarters ahead; we skip a quarter to avoid look-forward bias and return periods do not overlap. The Household Equity Share ($HEShare$) is the share of the household sector's equity and credit assets allocated to equities, calculated using data from the Federal Reserve's Financial Accounts of the United States. Regression (1) is the baseline regression. Regression (2) lags $HEShare$ by two quarters, i.e. using $HEShare_{t-2}$ to forecast R^e_{t+2}. Regression (3) and Regression (4) lag $HEShare$ by four and six quarters, respectively. Regression (5) uses the first-half of the sample. Regression (6) uses the second-half of the sample.

$$R^e_{t+2} = b_0 + \lambda \cdot HEShare_t + \epsilon_{t+2}$$

Data frequency is quarterly, 1953q2 to 2015q3. Newey-West standard errors with five quarters of lags.

	(1)	(2)	(3)	(4)	(5)	(6)
HEShare	-0.253***				-0.366***	-0.270***
	(0.062)				(0.111)	(0.101)
HEShare, lag 2q		-0.202***				
		(0.064)				
HEShare, lag 4q			-0.161**			
			(0.065)			
HEShare, lag 6q				-0.097		
				(0.066)		
Intercept	0.183***	0.149***	0.122***	0.079*	0.269***	0.186***
	(0.041)	(0.041)	(0.041)	(0.042)	(0.078)	(0.060)
Adj. R2	0.051	0.031	0.018	0.004	0.079	0.040
N	249	247	245	243	124	125
Dataset	All	All	All	All	First-Half	Second-Half

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3: Alternate Definitions of the Household Equity Share

This table shows the results of using alternate definitions of the Household Equity Share. The Household Equity Share ($HEShare$) is the share of the household sector's equity and credit assets allocated to equities, calculated using data from the Federal Reserve's Financial Accounts of the United States. Regression (1) uses our baseline definition from Section 3. Regression (2) excludes household holdings of equity and bond mutual funds. Regression (3) excludes household holdings of loan assets. Regression (4) combines the two exclusions. R^e_{t+2} is the quarterly excess return of the value-weighted market index, two quarters ahead; we skip a quarter to avoid look-forward bias and return periods do not overlap.

$$R^e_{t+2} = b_0 + \lambda \cdot HEShare_t + \epsilon_{t+2}$$

Data frequency is quarterly, 1953q2 to 2015q3. Newey-West standard errors with five quarters of lags.

	(1)	(2)	(3)	(4)
HEShare	-0.253***			
	(0.062)			
HEShare, no MF		-0.264***		
		(0.062)		
HEShare, no Loan			-0.231***	
			(0.060)	
HEShare, no LoanOrMF				-0.250***
				(0.060)
Intercept	0.183***	0.190***	0.177***	0.192***
	(0.041)	(0.041)	(0.041)	(0.041)
Adj. R2	0.051	0.052	0.046	0.050
N	249	249	249	249

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Adjusting for Finite Sample Bias

This table shows the effect of correcting for finite sample bias due to a persistent return forecaster, as emphasized by Nelson and Kim (1993) and Stambaugh (1999). R^e_{t+2} is the quarterly excess return of the value-weighted market index, two quarters ahead; we skip a quarter to avoid look-forward bias and return periods do not overlap. The Household Equity Share ($HEShare$) is the share of the household sector's equity and credit assets allocated to equities, calculated using data from the Federal Reserve's Financial Accounts of the United States. Regression (1) uses Newey and West (1987) heteroskedastic and autocorrelation adjusted standard errors with five quarters of lags. The Stambaugh correction in Regression (2) adjusts the point estimate, using Kendall (1954) and Stambaugh (1999). The Lewellen bound in Regression (3) is not an estimate of the coefficient, but rather establishes an upper bound (lower bound, in magnitude) for the coefficient and t-statistic, which also establishes an upper bound on the p-value, using Lewellen (2004).

$$R^e_{t+2} = b_0 + \lambda \cdot HEShare_t + b_1 \cdot X_t + \epsilon_{t+2}$$

Data frequency is quarterly, 1953q2 to 2015q3.

	(1) Newey-West	(2) Stambaugh correction	(3) Lewellen bound
$\hat{\lambda}$	-0.253	-0.227	-0.0972
Standard Error	0.0623	0.0623	0.0495
t-statistic	-4.06	-3.64	-1.96
p-value	0.00006	0.0003	0.051

Table 5: Long Horizon Regressions

This table shows forecasting ability of $HEShare$ on long horizon returns. $R^e_{t+1\to t+k+1}$ is the annualized market excess return from $t+1$ to $t+k+1$; we continue to skip a quarter between $HEShare_t$ and the forecasted returns. Regressions (1), (3), (5) show the effect on forecasting returns one year ahead (4 quarters), three years ahead (12 quarters) and five years ahead (20 quarters). We adjust for overlapping returns using Newey-West standard errors with $k+5$ quarters of lags. Regressions (2), (4), (6) also show the effect on forecasting returns one year ahead (4 quarters), three years ahead (12 quarters) and five years ahead (20 quarters). For these regressions, we drop data so return periods do not overlap.

$$R^e_{t+1\to t+k+1} = b_0 + \lambda \cdot HEShare_t + \epsilon_{t+k+1}$$

Underlying data spans 1953q2 to 2015q3.

	(1)	(2)	(3)	(4)	(5)	(6)
HEShare	-0.886***	-0.797***	-0.601***	-0.633***	-0.534***	-0.567***
	(0.223)	(0.235)	(0.158)	(0.208)	(0.118)	(0.133)
Intercept	0.650***	0.590***	0.447***	0.475***	0.399***	0.415***
	(0.146)	(0.158)	(0.104)	(0.136)	(0.078)	(0.091)
Adj. R2	0.149	0.121	0.279	0.248	0.407	0.451
N	246	61	238	20	230	12
Horizon	1yr	1yr	3yr	3yr	5yr	5yr
Overlap	Y	N	Y	N	Y	N

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6: Univariate Comparison with Other Forecasters of Excess Market Returns

This table compares $HEShare$ with other forecasters of excess market returns. R^e_{t+2} is the quarterly excess return of the value-weighted market index, two quarters ahead; we skip a quarter to avoid look-forward bias and return periods do not overlap. The Household Equity Share ($HEShare$) is the share of the household sector's equity and credit assets allocated to equities, calculated using data from the Federal Reserve's Financial Accounts of the United States. $CAPE$ is the ten-year cyclically adjusted price-to-earnings-ratio (Campbell and Shiller, 1988b). $EquityIssue$ is the equity share in new issuances (Baker and Wurgler, 2000). CAY is the consumption-wealth ratio (Lettau and Ludvigson, 2001). $TermSpread$ is the yield spread between the 10-year US Treasury and 90-day US Treasury. $TBill$ is the 90-day US Treasury yield.

$$R^e_{t+2} = b_0 + b_1 \cdot X_t + \epsilon_{t+2}$$

Data frequency is quarterly, 1953q2 to 2015q3. Newey-West standard errors with five quarters of lags. Each column in the table displays two regressions: one with the regressors unadjusted (b_1) and one with the regressors normalized to have unit variance (b_1^{norm}).

X_t	(1) HEShare	(2) CAPE	(3) EquityIssue	(4) CAY	(5) TermSpread	(6) TBill
b_1	-0.253***	-0.001	-0.125*	0.780***	0.673	-0.254
b_1^{norm}	-0.020***	-0.009	-0.010*	0.015***	0.008	-0.008
t-statistic	-4.06	-1.43	-1.89	2.95	1.56	-1.45
Adj R2	0.051	0.007	0.011	0.030	0.005	0.005
N	249	249	249	249	249	249

Table 7: Bivariate Comparison with Other Forecasters of Excess Market Returns
This table shows the result of controlling for other known predictors of excess market returns. R^e_{t+2} is the quarterly excess return of the value-weighted market index, two quarters ahead; we skip a quarter to avoid look-forward bias and return periods do not overlap. The Household Equity Share ($HEShare$) is the share of the household sector's equity and credit assets allocated to equities, calculated using data from the Federal Reserve's Financial Accounts of the United States. $CAPE$ is the ten-year cyclically adjusted price-to-earnings-ratio (Campbell and Shiller, 1988b). $EquityIssue$ is the equity share in new issuances (Baker and Wurgler, 2000). CAY is the consumption-wealth ratio (Lettau and Ludvigson, 2001). $TermSpread$ is the yield spread between the 10-year US Treasury and 90-day US Treasury. $TBill$ is the 90-day US Treasury yield.

$$R^e_{t+2} = b_0 + \lambda \cdot HEShare_t + b_1 \cdot X_t + \epsilon_{t+2}$$

Data frequency is quarterly, 1953q2 to 2015q3. Newey-West standard errors with five quarters of lags.

	(1)	(2)	(3)	(4)	(5)	(6)
HEShare	-0.253***	-0.247***	-0.237***	-0.210***	-0.256***	-0.267***
	(0.062)	(0.068)	(0.064)	(0.074)	(0.068)	(0.060)
CAPE		-0.000				
		(0.001)				
EquityIssue			-0.086			
			(0.061)			
CAY				0.286		
				(0.320)		
TermSpread					-0.048	
					(0.440)	
TBill						-0.325**
						(0.158)
Intercept	0.183***	0.182***	0.188***	0.155***	0.186***	0.207***
	(0.041)	(0.040)	(0.041)	(0.048)	(0.047)	(0.041)
Adj. R2	0.051	0.047	0.054	0.050	0.047	0.061
N	249	249	249	249	249	249

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Multivariate Comparison with Other Forecasters of Excess Market Returns
This table is a multivariate version of Table 7. It is on a separate page due to space constraints; see Table 7 for variable definitions. Regression (1) is our baseline regression. Regression (2), Regression (3), and Regression (4) adding multivariate controls of known forecasters of excess market returns.

$$R^e_{t+2} = b_0 + \lambda \cdot HEShare_t + b_1 \cdot X_t + \epsilon_{t+2}$$

Data frequency is quarterly, 1953q2 to 2015q3. Newey-West standard errors with five quarters of lags.

	(1)	(2)	(3)	(4)
HEShare	-0.253***	-0.166**	-0.278***	-0.164*
	(0.062)	(0.077)	(0.084)	(0.096)
CAPE		-0.001		-0.002
		(0.001)		(0.001)
EquityIssue		-0.168**		-0.118
		(0.082)		(0.085)
CAY			0.250	0.352
			(0.323)	(0.316)
TermSpread			-0.655	-0.605
			(0.478)	(0.489)
TBill			-0.448***	-0.482**
			(0.172)	(0.190)
Intercept	0.183***	0.184***	0.230***	0.214***
	(0.041)	(0.041)	(0.061)	(0.062)
Adj. R2	0.051	0.059	0.062	0.069
N	249	249	249	249

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1