Household Risky Ratio and Expected Stock Returns[1]

Fan A. Zhang[2]

Harvard University

First draft: December 2012

This draft: August 27th, 2013

Abstract

The *household risky ratio*, the ratio of high risk assets over low risk assets directly owned by households, is a strong negative predictor of the equity premium on the US stock market between the years of 1951 and 2012, the range of data available. The predictability is robust to definition of the asset classes, first versus second half of sample, and the finite-sample bias of Stambaugh (1999). The predictability is stronger than, and not subsumed by popular predictors like price-earnings ratios, yield spread, equity share of issues, or consumption-wealth ratios. The main predictive power is decomposed into three nearly equal parts: 1) the household tilt of risky assets, which is novel and generally orthogonal to known predictors; 2) a valuation ratio component; and 3) an issuance component of high risk versus low risk assets.

[1] I am grateful to Malcolm P. Baker, John Y. Campbell, Robin Greenwood, and David I. Laibson for conversations that helped shape this paper, and to Timothy McQuade, Charles G. Nathanson and David C. Yang for comments.
[2] Fan Zhang: Department of Economics, Littauer Center, Harvard University, Cambridge, MA 02138. Email fanzhang@fas harvard.edu.

I. Introduction

Over the last few decades, a key asset pricing finding is that the equity premium varies over time. Predictors of the time varying equity premium include price-earnings ratios (Fama French 1988) (Campbell Shiller 1988), yield spreads (Fama Schwert 1977, Keim Stambaugh 1986, Campbell 1987), consumption-wealth ratios (cay) (Lettau Ludvigson 2001), or equity share of issuances (Baker Wurgler 2000).

The equity premium can be seen as a type of inverse price on the equity market (Gordon 1962). In understanding time variation of prices, it seems natural to look at time-variation of quantities. As the difference between expected returns on high risk versus low risk assets, the equity premium motivates the inspection of the quantities associated with high risk versus low risk assets.

Baker and Wurgler (2000) do exactly this, looking at quantities of equity versus debt issued by the corporate sector as a predictor of future equity market returns and find positive results. In particular, the paper finds that when equity prices are high, and future expected returns are low, corporations rationally respond by issuing a higher ratio of equities versus bonds. The results of Baker and Wurgler (2000) suggest that the corporate sector absorbs and optimizes to the demand shocks.

A natural question to ask is, from where do the demand shocks arise? One sector to examine are households, which might be a source of demand shocks due to either behavioral theories like extrapolative beliefs (Greenwood Shleifer 2013) or to consumption-CAPM theories like Campbell Cochrane (1999). Therefore, I then look at the household sector's direct holdings of high risk versus low risk assets.

I draw my data from the Federal Reserve Flow of Funds, which enumerates all the instruments held by the household sector. I find that the four major holdings which compose the vast majority of household direct holdings are equities, risky mutual funds, credits, and deposits. The first two sum up to high risk assets and the bottom sum up to low risk assets.

I then define the *household risky ratio* as the ratio between the two: the household sector's direct holdings of high risk assets over the household sector's direct holdings of low risk assets. I find the household risky ratio is a negative predictor future equity premium between 1951 and 2012, which is the full sample of available data from the Flow of Funds. The negative relationship is obvious in simple binnings of the predictor variable (Figure 1), and persists for years beyond the binning date (Figure 2).

In a univariate setting, the predictability is strong: the R^2 is 5% quarterly and higher than any single of the following popular predictors: the Campbell Shiller (1988) cyclically adjusted price-earnings ratio (CAPE), the yield spread, the equity share of issues, and the cay (Table 3). The predictive power of the household risky ratio is also not subsumed by the above popular predictors: the coefficient remains essentially the same in a multivariate regression alongside the above predictors (Table 4).

The predictive power of the household risky ratio is robust to a range of variations. It is robust to construction from Flow of Funds data. Dropping any single component of the four mentioned above does not eliminate the predictive power (Table 5). Adding boundary components, components that arguably could be categorized as either inside or outside household discretionary assets, also does not affect the results. It is robust to subsamples: the coefficient in the first half of the sample is almost identical to that of the second half of the sample (Table 6). The effect cannot be explained by Modigilani-Miller effects (Section 5).

Finally it is robust to the small sample correction of Stambaugh (1999). The household risky ratio does experience Stambaugh bias since it has a high autocorrelation (quarterly ρ= .96) and cross correlation with returns (ρ = .86). However, correction for Stambaugh bias via the methods of Kendall (1954) and Kothari Shanken (1997) show that bias to be limited to about 20% of the effect size. This is confirmed by the Lewellen (2004) bounds test and the Campbell Yogo

(2006) correction method (Table 7). The magnitude of this 20% drop does not change whether the predictor is economically or statistically significant, and is less than other popularly accepted predictors like dividend-price ratios.

I also run overlapping regressions, taking care to address econometric issues of such regressions, and show that the household risky ratio has substantial long-horizon predictability. The R^2 is as high as 22% at the one year level and 40% at the three year level. However, for all other regressions I still set the baseline prediction-horizon to be equal to the sampling-period of one quarter. This follows in the theory of Harri and Brorsen (2009) and avoids the complicated corrections and lack of transparency of overlapping regressions.

This paper then investigates the source of the predictive power of the household risky ratio. The household risky ratio can be seen as the sum of three parts. First define an auxiliary variable called the *total risky ratio*, the ratio of economy-wide high risk assets over economy-wide low risk assets. Part one is then the *household tilt,* the ratio of the *household risky ratio* over the *total risky ratio*; it measures how much households tilt towards high risk assets above and beyond the rest of the economy. The household tilt variable is seen to be generally orthogonal to known popular predictors (Table 9) and hence is novel.

Parts two and three explain the total risky ratio. The second part is the *valuation ratio part* of the total risky ratio, and captures the extent to which the total risky ratio can be seen as a price of risky assets, and thus a tracker of historical price changes. For example, if the value of high risk assets doubles overnight, the total risky ratio also doubles. The valuation ratio component captures this doubling. The valuation ratio part is then defined as the projection of the total risky ratio onto the known valuation ratio CAPE.

The final part must mechanically be what remains from the above. In particular, it is the residual from the above projection of the total risky ratio onto the known valuation ratio CAPE. Since asset classes can only grow though capital gains and issuances, and capital gains are captured by CAPE above, the residual can be thought of theoretically as the component of the household risky ratio that arises from issuances. In fact, empirically the final part indeed strongly interacts with Baker-Wurgler issuances (Table 9). For these reasons, I term it the *issuance part*.

These empirical results can be related to the theory of demand systems for assets (Brainard and Tobin 1968). Households can be seen as exogenously demanding more risky assets. This causes the quantity of household high risk assets to increase, while at the same time increasing the price of high risk assets and decreasing forward returns. This justifies the observed negative correlation between the household risky ratio and forward returns. The demand system model can also be seen in recent work by Baker Wurgler (2000), Baker Greenwood Wurgler (2003), and Greenwood Vayanos (2010).

These empirical results have welfare implications. Since the household sector buys equities when future returns are lower than average, the return experienced by households will be less than a constant-ratio standard. I calculate that households receive a Sharpe ratio of .267 from holding risky assets instead of .311 under a constant-ratio standard. Both returns are far below the .533 Sharpe ratio attainable by perfectly conditioning on predictive variables.

Section 2 presents the data, definitions, and various constructions of the household risky ratio variable. Section 3 presents the predictive power of the variable for the equity premium. Section 4 decomposes the predictive power of the household risky ratio into components. Section 5 discusses some theory and implications of the results. Section 6 discusses extensions. Section 7 concludes.

Section 2: Data

The main data series used to construct the household risky ratio and related variables is the *Federal Reserve Flow of Funds Financial Accounts of the United States*, or simply the Flow of Funds. This data series reports quarterly data starting the fourth quarter of 1951, and continues to present day with data released around 60 days after the end of the quarter. All data is collected from the Flow of Funds using the data download program from the quarterly series (Q series) measuring non-seasonally-adjusted levels (FL series).

The Flow of Funds is organized mainly along two dimensions. The first dimension is the sector, defined as a partitioning of players in the economy. Sectors include households and nonprofit organizations, nonfinancial businesses, state and local governments, and soforth (see Appendix Table A1). The second dimension is instrument type, which can be seen as a class of assets or liabilities: home mortgages (liability), total loans (liability), treasury securities (asset), and security credit (asset) are just a few examples. For both the sector and instrument dimension, there are varying levels of aggregation.

To construct the household risky ratio, the numerator and denominator was first separately constructed from data within the household and nonprofit organizations sector. The numerator consists of the sum of equities directly held by households (FL153064105.Q) and the risky mutual funds directly held by households (FL153064205.Q). This corresponds roughly to the total amount of high risk assets that households hold directly with discretion, defined as assets that households can readily and liquidly trade to reflect their preferences and beliefs. The approximation is even better in variation-space instead of levels-space. The denominator consists of household directly held credit market instruments (FL154000025.Q), and deposits which includes low risk money market mutual funds (FL154004005.Q). This corresponds roughly to the amount of low risk assets that households hold directly with discretion.

The household risky ratio then can be seen as approximately the ratio of high risk to low risk assets directly and discretionarily held by households. Two items are to be discussed: what exactly does approximate mean, and what's the significance of assets that are directly and discretionarily held?

The approximation is actually quite good. Within the class of household discretionary and direct low risk assets for example, credits and deposits are by far the largest instruments. The next largest instruments like security credit or miscellaneous assets compose only less than 5%

of credits and deposits, and similarly for risky assets. This is improved even more once one considers that approximation in levels is not as important as approximation in variation. For high risk assets even though some local government retirement plans are not measured, the variation of the unmeasured portion likely moves with the measured portion – which is to say with the same variation as the equities market. This robustness and point that the variation is what matters is demonstrated in the later discussion of Table 6 which shows that predictability barely drops even as large parts of the above four components are dropped. Overall then, the approximation of the numerator and denominator to household discretionary and direct holdings is quite good.

The second issue to discuss is why the paper limits to discretionary and direct holdings of the household. In some respects, this is a fundamental issue. After all, the entire economy is indirectly controlled by the household sector, so even the definition of the household sector by the Flow of Funds assumes some boundary. This paper is simply assuming a narrower boundary.

The major theoretical reason motivating the limit to discretionary and direct holdings is that only assets that are readily and liquidly traded by households can be seen to strongly and immediately reflect their preferences and beliefs. To the extent that the household risky ratio has predictive power through capturing preferences and beliefs, this property is essentially. For example, take the largest asset classes not included: real estate, private businesses, and life insurance reserves. The first is moderately illiquid with transaction costs of 6% plus time, preparation, and other opportunity costs. It is also a bundled good that reflects preferences for internal space, external location, commute distance, and school district. Thus, sales and purchases of housing might be seen as a much noisier measure of household risk preferences then, say, holding the S&P index. This is similarly true for private businesses which suffer from heavy adverse selection issues in sales. Life insurance reserves likewise are often managed by a portfolio manager and so less effectively measure household preferences.

The above two points provide theoretical motivation for the construction of the household risky ratio variable. Even without the above two reasons, the household risky ratio can be taken at face value as an empirical construct with predictive power. Concerns of data snooping could be allayed by the fact that the 4.16 t-statistic in univariate regressions corresponds to a Bonferroni correction of testing one thousand independent variables, whereas the Fed Flow of funds has much fewer, especially independent variables. Further assurance can be realized in the robustness of the ratio to construction (Table 6).

The other primitive variable constructed from the Flow of Funds is the total risky ratio, which is the exact analogue series of the household risky ratio but for all sectors combined instead of just the household sector.

Compared to other broad economic series, the Federal Reserve Flow of Funds is minimally affected by look-ahead bias. Generally the series is release quickly after a quarter's end: around 60 days on average, and almost never exceeding a quarter. Thus, the one quarter lag used in the baseline is more than sufficient to account for this. The Flow of Funds is revised from time to time. However, revisions are generally limited to one to two years back, and this paper did not find any revisions of the four broad series used to define the household risky ratio.

Figure 1 presents the amount of household high risk versus low risk assets on a log scale from 1951-2012. The high risk series, as expected, is more variable over time. The high and low risk series seem to be trending up linearly, perhaps each following a first order autoregressive process with unit root and drift. The two series looks co-integrated, which is justifiable by theory: in the long run they might be expected to be growing at the same rate as the entire economy.

Figure 2 shows the household risky ratio, the first series above divided by the second series. Peaks and trends in this series are apparent. The ratio slowly but steadily climbs from 1951 to about 1969, which might signify a trend where equities began to be seen as less a speculative asset from the aftermath of the great depression, into an asset that everyone could own. As will be discussed later, the rise in price may have caused an increase in quantity if households have extrapolative beliefs, and the two effects reinforced each other.

Households then experienced two large price drops in 1970 and 1974, causing them to be more cautious about stocks for many decades ahead, up until the mid-90s. It is exactly during this time that equities saw their highest returns. In the late 90s, as the Internet bubble got underway, households against started shifting into high risk assets, again to see significant wealth loss as the bubble burst.

Other data sources have been compiled from sources as standard as possible. The ten-year cyclically adjusted price-earnings ratio (CAPE) is calculated as per Campbell Shiller (1988), and the data is collected from Robert Shiller's website. CAY is defined and supplied by Lettau Ludvigson (2001). Equity shares of issuances is from Baker Wurgler (2000). Total stock market returns is the value-weighted series from CRSP. The long rate is the GS10 rate collected from Robert Shiller's website. The short rate is the 1-month treasury bill rate provided by Ibbotson and Associates, Inc. and provided on Kenneth French's website. The equity premium is calculated as the difference between the stock market returns minus the short rate expressed in percentage points. The yield spread is calculated as the difference between long rates and short rates.

Section 3: Predictive Power

This section outlines the predictive power of the household risky ratio. It begins with informal graphical analyses and then formalizes these analyses through regressions. Finally, this section checks robustness through multivariate regressions and additional tests.

A. Baseline Predictive Power

Figure 3 shows 1-year and 3-year value-weighted CRSP forward equity premia sorted by the lowest to highest quintile of the household risky ratio variable. The drop in equity premia from the highest to lowest quintile is generally monotonic, demonstrating a somewhat continuous relationship between the two variables. The magnitude of the effect is larger for 3-years, which is not surprising given the persistence of the household risky ratio predictor variable.

Figure 4 shows forward annual risk premia with respect to the lowest, middle, and highest tercile of household risky ratio. Note that there is no relationship between the terciles and returns before the tercile formation time. For the first 9 years, the lowest tercile has significantly greater returns than then highest tercile. The years of predictability seems to accumulate most at the beginning but persist for many years. One clear explanation for this is simply the fact that the predictor variable is quite persistent, so predictability a few years out can simply reflect this persistence. In this way Figure 4 can also be seen as a cross correlogram between returns and the household risky ratio.

At a glance, it seems clear that there is some negative relationship between the household risky ratio and forward returns. The relationship seems to persist for quite a number of years afterwards.

B. Univariate Regressions

I next show the power of the household risky ratio predictor in a univariate regression setting. I predict one-quarter forward returns using a variety of predictors. All predictors are lagged one quarter, leaving a one quarter minimum gap between the measurement time of the predictors and the start of the return period predicted. As standard in the literature, this gap ensures that the prediction is completely in the future period, and that the conditioning data is available for practitioners. This is especially important because some predictors like the household risky ratio and the Shiller CAPE contain equity prices; common noise in equity price measurement would enter on both sides of the regression and swamp the magnitude of return predictability.

The general regression framework for return predictions following the discussion above is then

$$R_{t,t+1} = \alpha + \beta X_{t-1} + \epsilon_t$$

Table 2 shows the results of this regression against a set of regressors. First, notice that the household risky ratio without additional lags is highly significant with a t-statistic of 4.16. As a univariate regression, the significance of this statistic under these relatively transparent settings is very high. Even more transparently, consider that the adjusted R^2 at .054 and the correlation is 0.232. These figures are all substantially higher than almost all popular predictors in the literature.

To examine the degree of predictability at different points in time, Table 2 includes results for the household risky ratio at one quarter and one year of additional lag. The coefficient noticeably decays from -2.49 to -2.40 to -1.94, but magnitude of the predictor even at an additional year out is only 22% less than originally. The t-statistics and R^2 also remain highly significant but noticeably decay. As will be seen later, this is largely caused by persistence in the household risky ratio, and not the independent ability of different lags of the household risky ratio at predicting future returns.

The next few rows of Table 2 examine the univariate regressions of other popular predictors of future equity premium. The paper chose four of the most popular variables in the literature as benchmarks, following Campbell Thompson (2008): the equity share of issues from Baker Wurgler (2000), the Shiller 10-year cyclically adjusted PE (CAPE) from Campbell Shiller (1988), the 10 year minus 1 month yield spread, and the cay proxy for consumption/wealth by Lettau Ludvigson (2001).

The coefficients for all these four other predictors are significantly less than that for the household risky ratio. The Lettau-Ludvigson cay is most significant with a t-statistic of 2.51 and an R^2 of .022. In second is CAPE with a t-statistic of 1.68 which is marginally significant. The lower significance of CAPE in these regressions compared to the literature can mainly be attributed to the shorter data series. For example, one of the most convincing works for the predictive power of CAPE is Campbell Shiller (1988) which uses data back to 1871 versus this paper's 1951, and a specialty of Robert Shiller's research is general marshaling past data to gain higher significance.

The equity-share of issuance then has a t-statistic of around 1.34. There are quite a few differences between this analysis and Baker Wurgler (2000) that may give rise to lower significance. Baker Wurgler (2000) note their effect has most significance and their test logically has most power in equal-weighted CRSP, while this paper uses the value-weighted CRSP equity premium. Baker Wurgler (2000) also uses different prediction periods from this paper, with predicted periods of one-year. The significance does increase if the duration of the issuance is increased from one quarter to four quarters. Finally, the time period matters again: the regressions in this paper are from beginning of the Flow of Fund data (1951) forward whereas Baker Wurgler (2000) used data from 1929. The yield spread in this analysis barely has univariate predictability at all.

Overall the result of this analysis is that the household risky ratio is a powerful predictor of future equity premium. This is not only the case in Table 2 which compares four popular variables under this setting, but also by direct examination of t-statistics and R^2 of the household risky ratio predictor, and comparing against analogous values across a wide variety of popular predictors in the setting of their seminal papers.

C. Default Inferences for Regressions

This subsection discusses and justifies the inference assumptions used by default for the regressions above and the rest of the paper. In particular, it discusses the focus on the one-quarter prediction horizon, as well as the choice to use Newey-West standard errors with five periods of lag.

By default in this entire paper, the prediction period is chosen to be one quarter due to that being the unit of observation in the *Federal Reserve Flow of Funds*. The later section on long-term predictions discusses more in depth the results of different periods of predictions, and econometric implications of having a prediction period that is longer than the data period. As a general preview, the results are more or less the same with quarter returns, one year returns, or three year returns, but the one quarter returns are most natural and least dependent on model corrections.

Also by default in this entire paper, regressions are done with Newey-West with five periods of lags for robustness. The Newey-West procedure encompasses the Eicker Huber White (EHW) heteroskedastic robust error correction (White 1980). The household risky ratio regression is not especially hetereoskedastic, but inferences are often affected 10% or more using the EHW method. Often the EHW standard errors are *less* than OLS due to return variance being concentrated near the center of the predictor.

Newey-West also additionally takes into account time series correlation of error terms assuming a triangle (Bartlett) kernel for time series correlation structure. The household risky ratio predictor has somewhat autocorrelated errors, and inferences are often affected 40% or more using Newey-West procedures instead of EHW. Often the Newey-West standard errors are less than EHW due to negative residual correlation. The negative residual correlation can be verified by adding lagged returns in the household risky ratio regression. However, the paper does not put lags into the baseline specification because the negative correlation is weak. In a sample regression with ten lags, only one coefficient exceeds 1.5 and none are significant at the 5% Bonferroni-corrected level.

The lag period was chosen using the procedure suggested by Newey (1993), and set at $\frac{3}{4}T^{\frac{1}{3}} \approx 5$, with T=240. In reality lags between one and ten periods generate almost the same result everywhere. While the paper does not show simple EHW standard errors or OLS standard errors, neither of these two less-robust alternate procedures generate a univariate t-statistic of less than 3.5 for the household risky ratio. More generally, econometric surveys (Harri Brorsen 2009) have shown that Newey-West corrects many substantially critical errors, and almost never give estimates substantially worse than other estimators in the space of heteroskedastic and autocorrelation robust estimators or less, including EHW and OLS. For both general and situation specific reasons this paper then universally uses Newey-West with five periods of lag as default.

D. Multivariate Regressions.

The univariate regressions are useful in examining the prediction power of each variable, as well as cross comparisons. However, to more fully understand the structure of the household risky ratio, the source of interactions with other variables, the source of the prediction power, it is imperative to examine multivariate regressions in Table 3. As above, the paper uses a one-quarter prediction period with a one period lag for the prediction, and Newey-West standard errors with 5 lags by default.

Table 3 begins its multivariate regression by regressing returns on household risky ratio and one year of additional lag. This regression gives insight into the time series structure of the predictability of household risky ratio. Do lags of household risky ratio have independent predictive power for future returns, as is the case for cay? This might arise if there is period-specific noise in measurement of the household risky ratio. Do the two lags knock each other out, as might be expected if the predictability of the household risky ratio is due to low frequency components? Looking at Figure 2 it might be tempting to suspect that predictability is driven completely by a few periods in which the household risky ratio was particularly high or low. Also, this regression lets us see whether it's the level of household risky ratio or its change that matters more, and whether longer lags might have more predictive power for some reason.

Looking at Table 3 regression (2), and comparing against the univariate regression (1), a few things stand out. First, the coefficient on the household risky ratio is near identical, going from -2.49 to -2.86. However, the t-statistic drops significantly from 4.16 to 2.69. Second, the coefficient on the one-year-lagged household risky ratio is near zero, with a t-statistic of 0.27. That the t-statistics are lower in regression (2) does suggest that household risky ratio and lagged household risky ratio knock each other out somewhat. Some portion of the predictability then can be due to a low frequency component of the household risky ratio. However, the significance of the leading household risky ratio term and its being equal to the univariate coefficient suggests that high frequency changes on the order of less than a year matter, and have the same effect on future returns as low-frequency changes. The high frequency variation in the household risky ratio is not simply noise.

The R^2 of regression (1) compared to (2) is the same, in line with the fact that the marginal t-statistic of the one-year-lagged household risky ratio is near zero. Adding an extra term does not increase predictability, so there is no independent prediction power by lags. (This is verified but not shown with other periods of lags). In a Markov-chain sense, the most recent household risky ratio could be seen as a state variable or a sufficient statistic for the household risky ratio process. This also means it is assuredly the level of household risky ratio that matters, versus changes, in contrast to other predictors like the equity shares of issuance (Baker Wurgler 2000).

The next regressions run in Table 3, regressions (3) through (6), examine bivariate regressions of the household risky ratio against various other predictors commonly used to forecast equity premia. The paper looks to see if any predictors knock the other out – a sign that one predictor is subsumed by the other better, less noisier, predictor. To the extent both predictors still maintain size and significance, it is a sign that the two predictors have somewhat independent predictability.

Regression (3) looks at the household risky ratio versus the equity share of issuances. These two variables have some relationship because companies issuing equity versus debt should be expected to mechanically increase or decrease the household risky ratio over time. Also, in the next section it will be seen that one component of the household risky ratio can be thought of as issuances. Some confounds might be expected, but instead the comparison of regression (3) against the univarate regressions show that significance and magnitude of both regressors increases slightly.

The household risky ratio coefficient increases from -2.49 to -2.52, significance increasing from -4.16 to -4.33. The equity share of issues increases much more from -7.32 (t-statistic of -1.34) to -10.49 (t-statistic of -2.13). The adjusted R^2 also increases from .054 and .003 in the univariate regression to .062 together. This suggests not only that the two predictors have independent predictive power, but also that the prediction is strengthened when the two are put together. This might occur if companies issue stock for two reasons: one is a mechanical response to the demand side of the economy, captured by the household risky ratio; another could be CFOs knowing to market time beyond demand signals.

Regression (4) looks at the bivariate regression of the household risky ratio against the Campbell-Shiller CAPE. One immediately notes that the coefficient on the household risky ratio actually becomes significantly higher in magnitude, jumping from -2.49 to -3.76. The coefficient on CAPE is also significant but now in the opposite direction: moving from -.13 to .18. This shows that not only is the household risky ratio not subsumed by the popular valuation ratio predictor CAPE, but that it actually may be a better valuation ratio than CAPE. Another interpretation might be that the household risky ratio captures more predictive components than CAPE, causing CAPE to act as a negative control. The decomposition of the household risky ratio in the next section lends some credence to this idea.

Campbell-Shiller (1988a) gives an accounting relationship between PE ratios, earnings growth, and expected returns. Given this opposite response in CAPE along with the household risky ratio, it may be interesting to see whether CAPE along with the household risky ratio can finally predict earnings changes.

Regression (5) runs the household risky ratio against the term spread. Neither variables are affected significantly. The term spread is not particularly significant, having both univariate and bivariate t-statistics of less than one. This seems consistent with a story in which the term spread is not a particularly powerful variable. Regression (6) runs the household risky ratio with cay. Both variables are reduced somewhat in magnitude and significance. The bivariate R^2 is also somewhat less that the summed R^2, showing that two variables are picking up on some common predictability. This is not surprising given Lettau Ludvigson (2001) theory of cay being a consumption wealth valuation ratio.

Regression (7) uses all variables and covariates, except the lagged household risky ratio which was seen before to be near collinear with the household risky ratio itself. Because many of the popular predictors tend to be different variations on each other, the significance and size of predictability is generally less than in the bivariate and univariate regressions. This is true of the household risky ratio as well. But what is striking is that the household risky ratio variable is still quite significant ($p<.05$). Also, across all the regressions from (1) through (7), the magnitude of the coefficient on the household risky ratio is more or less the same at around -2.5.

This remarkable consistency across the bivariate and multivariate regressions show that the household risky ratio is not subsumed by other popular predictors, or even the space spanned by other popular predictors. Finally regression (8) standardizes each predictor to have unit variance. The coefficients then are identically proportional to the square of the correlation and square root of R^2, and give a sense of the strength of each predictor on the same scale.

The two main takeaways from Table 3 is both the empirical robustness of the household risky ratio as a predictor, and the point that there is something novel about the household risky ratio. Further evidence will be given below that there is particular significance to the household portion of the household risky ratio.

E. Robustness

One concern in regressions over large timespans is parameter stability. How much of the predictability at persists from the start of the sample to the end of the sample? Alternately, is predictability completely isolated to one portion of the sample, or even worse, a few years? If the household risky ratio process or the returns process has underlying long persistence, then normal inferences could be incorrect. All of the above issues can be heuristically addressed by doing a split regression on the first and second half of the data (Table 4). Such a procedure is model free, and while not the most powerful, provides a clear glance at whether the predictive power is stable. The paper finds that the lineup between the first and second half of the sample is remarkable: the coefficients are quite close: -2.94 versus -2.15. This gives evidence that the result isn't driven just by a few years, and the inferences are not spuriously caused by persistence.

There may be concern over what exactly is classified as components of the numerator and denominator of the household risky ratio. In particular, there might be concerns that one component is the main driver of the result, and so the household risky ratio is not robust to definition. As categorized by the Flow of Funds, there are four major categories in the paper's definition of the household risky ratio. The high risk component is composed of equities and mutual funds. The low risk portion consists of credit market securities and deposits. Table 5 uses a jackknife-like methodology to test robustness of the household risky ratio. The paper runs regressions removing one component at a time, and sees whether predictability drops. It seems that predictability remains in almost all case. The t-statistics generally remain significant, and the coefficient varies from -1.21 to -3.22. The robustness of regressions to removing components shows that the household risky ratio is robust to definition.

F. Small Sample Bias of Stambaugh (1999)

Stambaugh (1999) and Nelson and Kim (1993) note that in regressions in which predictors X_t are autocorrelated, there is a possibility for a small sample bias. In particular, using notation of Kothari Shanken (1997), suppose returns r_{t+1} is being predicted with a univariate autocorrelated variable x_t.[3]

$$r_{t+1} = \alpha + \beta x_t + u_{t+1} \ (1)$$

$$x_{t+1} = \gamma + \phi x_t + v_{t+1} \ (2)$$

[3] The theory of Stambaugh bias is naturally presented without a one period gap between the predictors and predicted. This does not pose a problem for us as we simply define x_t := Risky_t-1

Correlations between in the error term u_{t+1} and v_{t+1} will cause a small sample bias. In the language of Lewellen (2004), the source of this bias is the well-known downward bias of OLS in estimating (2) equation for positive ϕ, in particular $E[\hat{\phi}_{OLS} - \phi] < 0$. This transmits over to β in through the correlation. In particular:

$$\hat{\beta} = \beta + (x'x)^{-1}x'u$$

$$\hat{\phi} = \phi + (x'x)^{-1}x'v$$

Which translates into, as Stambaugh (1999) notes, defining $\gamma \coloneqq \frac{Cov(u,v)}{Var(v)}$

$$E[\hat{\beta} - \beta] = \gamma E[\hat{\phi} - \phi] \neq 0$$

Since the household risky ratio does include a significant price component, it is very close to the original canonical example used by Stambaugh (1999) to demonstrate the bias. In particular, the predictor is quite autocorrelated with $\hat{\phi}_{OLS} = .956$. The error terms are of (1) and (2) are also notably correlated at $\widehat{corr}(u, v) = .869$. All correlations are taken at the one-quarter level.

This paper corrects for the error below using a few methodologies: Kendall (1954), Kothari and Shanken (1997), and a more recent approach by Campbell and Yogo (2006). The correction shows that the household risky ratio is indeed subject to small sample bias about 20% of the magnitude of the coefficient. This bias is not enough to affect qualitatively the results in the paper, especially many of the baseline t-statistics are above 3 or 4. The paper follows the literature and does not by default correct for Stambaugh bias in order to maintain regressions that are multivariate, heteroskedastic and autocorrelation consistent, and standardized to well-known methodologies. However, it is important to keep this relative magnitude of 20% in mind when reading other tables, which affects not only the household risky ratio, but other popular price predictors that have a valuation ratio component like CAPE.

The first correction is a point-estimate correction suggested by Kendall (1954), who notes that analytically the bias in ϕ, $E[\hat{\phi} - \phi] \approx \frac{1+T\hat{\phi}}{T-3}$ which holds generally for T>50 as in this case, giving a value of -0.01598. Then, $E[\hat{\beta} - \beta] = \gamma E[\hat{\phi} - \phi] \approx \frac{\widehat{cov}(u,v)}{\widehat{var}(v)}\left(-\frac{1+3\hat{\phi}}{T}\right)$ again for T>50 can be estimated, in this case 30.00 x -.01598 = -.479. Thus the Kendall correction reduces the OLS estimate of β from -2.27 to -1.79, a move of about 21.1%. The Kendall method does not change the standard errors, leading to a t-statistic of 2.75 and p-value of .0065. The reduction in significance comes solely from the move in the point estimate.

Kothari-Shanken (1997) extends the Kendall (1954) method through bootstrap re-estimation of the standard errors. In particular, Kothari-Shanken simulates the data series by taking the Kendall-adjusted values for ϕ, β and drawing error terms u, v. For each bootstrapped series, a beta is estimated. The distribution of bootstrapped betas then simulates the sampling distribution of β and lets us do inferences. The process results in a p-value of 0.004.

The Kothari-Shaken methodology does not generate wider confidence intervals from the OLS ones. In fact, the confidence interval is actually lower than OLS due to the same reason mentioned previously of EHW standard errors being less than OLS: the middle of the predicted variable has more residual variance than the edges. Like the Kendall methodology the loss in t-statistic comes mostly from the point estimate moving closer to zero versus the standard error increasing. The Stambaugh bias again is about 20% of the OLS coefficient magnitude.

Campbell and Yogo (2006) propose a different Stambaugh bias correction methodology, motivated by the fundamentals of statistical hypothesis testing. The idea behind Campbell Yogo, in relation to its use in this paper, is as follows. First, Campbell and Yogo arrive at an intuitively powerful and rigorous Q-test that accounts for Stambaugh bias. The Q-test depends on knowing ϕ above, which in reality must be estimated rendering the original test infeasible. Campbell Yogo cleverly patch this problem by using the Bonferroni procedure to merge the infeasible Q-test with an estimate of ϕ through the DF-GLS estimator.

The rationale for the Q-test is as follows. Suppose the autocorrelation ϕ of the predictor x_t was known beforehand. The Neyman-Pearson lemma proposes a likelihood ratio test (LRT) as the most powerful test for hypothesis of $\beta = \beta_0$ vs $\beta = \beta_1$. The LRT can be conditioned on an ancillary statistic and be considered inside the space of invariant tests, tests that do not change in response to invariant changes to x_t or r_{t+1}. Then this LRT is also uniformly most powerful (UMP) for β inside the above space of tests. Call this UMP test statistic the Q-statistic, which can be estimated easily in an OLS regression setting. The test based on cutoffs of the Q-statistic is the Q-test.

Kendall provides an intuitive first order estimate of the Stambaugh bias and Kothari Shanken (1997) provide intuitive first order inference corrections. Neither however have foundations in rigorous statistical testing, which is what is provided by Campbell Yogo (2006). Further, the construction procedure of the Campbell Yogo (2006) is designed towards ensuring high power while being rigorous. As such, this paper considers the Campbell-Yogo correction to be the gold standard for comparison.

The test is implemented via Campbell Yogo (2005). The DF-GLS statistic on the household risky ratio with one period if lag is calculated at -2.287 (Table 6), corresponding to a 97.5% confidence interval for c of [-21.87, 1.47] from Table 1 of Campbell Yogo (2005). This leads to a 97.5% CI over phi of [.910, 1.006]. This leads to point estimates of $\hat{\beta}(\phi)$ of -3.65 and -.76 respectively. For each of the two above point estimate, this paper takes another 97.5% confidence interval around the estimates to arrive at a 95% Bonferroni estimate for β: [-4.32, -.03]. An approximate 90% CI would then be [-4.06, -.30].

As noted in Campbell, due to the Bonferroni methodology, the actual confidence interval coverage is guaranteed to be 95% or above. Thus, the estimation above is conservative. Even with the most conservative methodology, it is seen that the household risky ratio is still significant. As a baseline, Campbell Yogo (2006) calculate 90% confidence intervals for D/P, E/P, T-bill rate, and Yield spread, and find many of their variables have 90% confidence intervals that cross 0, and almost all are very close to 0 with respect to the length of the confidence interval. It seems likely that if 95% confidence intervals were calculated, all of the other above univariate predictors would cross zero. Keep in mind that much of the Campbell Yogo (2006) data set extends back to 1926 or even 1880. With this as a baseline, the [-4.32,-.03] confidence interval is remarkably strong for the household risky ratio.

Finally, to verify significance, this paper does one last check in the style of Lewellen (2004). It seems reasonable to assume that the household risky ratio is not an explosive process. Looking at Figure 2, the household risky ratio looks likely even stationary as household high risk assets and low risk assets ought to be cointegrated with the size of the economy. Although it is possible it has a unit root, it makes little sense for the household risky ratio to be explosive. Assuming $\phi \leq 1$, the first step to re-run is the Campbell Yogo (2006) bounds. No longer is the 97.5% confidence interval for ϕ [.910, 1.006], but rather it is instead [.910,1]. This leads to a 95% confidence interval for β of [-4.06,-0.217].

Following Lewellen (2004) more directly, consider again $E\left[\hat{\beta} - \beta\right] = \gamma E\left[\hat{\phi} - \phi\right]$. Rearranging this leads to the estimator that eliminates the bias: $\hat{\beta}_{adj} = \hat{\beta}_{OLS} - \gamma(\hat{\phi} - \phi)$. Assuming the largest correction possible (being overly conservative) under the above assumptions sets $\phi = 1$, yielding $\hat{\beta}_{adj} = -.92$. This is significantly less than the OLS estimate or any other estimates really because it is a lower bound in magnitude. The advantage of the Lewellen method is that the standard errors are much reduced, from .65 in OLS to .32 here. This yields a much smaller relative confidence interval than Campbell Yogo (2006) and gives a larger t-statistic of 2.88.

The Lewellen (2004) method does not provide any unbiased point estimates of β, but as an advantage, it does have more rigorous foundations than Kendall (1954) or Kothari Shanken (1997), and since the process studied here has an autoregressive root close to unity, it is more powerful than Campbell Yogo (2006).

G. Persistence and Long-Horizon Predictions

Table 7 analyzes the persistence properties of the household risky ratio predictor. The analysis of this is similar to that of the persistence of D/P. From a theoretical perspective, the household risky ratio should not have a unit root: as the economy expands it seems that the fraction of high risk assets versus low risk assets ought to vary around some constant. There certainly is no theoretical reason to believe it has a greater-than-unit root: no model seems to justify an explosive process. From an empirical perspective, if it is believed that expected excess returns do not have a unit root, then it makes no sense to predict it with a series with a unit root. Table 7 Panel A shows the DF-GLS test on the household risky ratio: while a unit root cannot be ruled out, this may be due to the power of the test with only 240 quarters as input. Table & Panel B shows the DF-GLS test on changes of household risky ratio: a unit root is clearly rejected. Thus, the household risky ratio is shown to be integrated maximally of order one.

In this paper, the predictive regressions are generally run on one-quarter forward returns, which also is the frequency of the data as provided by the Federal Reserve flow of funds. The literature often considers longer-horizon returns (Baker Wurgler 2000, Campbell Shiller 1988) to increase the R^2 of the regression as well as to demonstrate the predictive properties over the long term. This section runs longer-dated returns using the household risky ratio and discusses the longer-dated results in the context of econometric theory.

A main advantage of long-horizon returns is that the coefficients estimated are directly interpretable as the long-term impact of the predictor variable. If the exact estimate of interest is the size of the one-year or three-year return as a result of a fixed change in the household risky ratio, then the most natural regression to run would be a long-horizon regression.

Another purported reason for using overlapping regressions is the higher R^2 of the regressions. However, especially for persistent predictors, the R^2 increase is mechanical. Valkanov (2003), Hjalmarsson (2006), and Boudoukh et al (2008) show that increasing the prediction period does not increase the power of tests. In fact, in monte-carlo simulations, the estimated coefficients of one period disaggregated returns and N-period aggregated returns are often correlated by .99 or more.

Other reasons often cited for using overlapping regressors include errors in the predictor variable (Cochrane Piazzesi 2005), missing observations, and higher prediction accuracy. However, Harri and Brorsen (2009) show that for a vast majority of statistical reasons given for using overlapping regressions, better statistical predictors are possible. They also show that the usual inference correction methods for overlapping regressions are quite problematic, which is borne out in this paper in Table 8.

Turning to this paper's specific results, in table 8, regressions (1) and (2) report disaggregated returns, returns with one-quarter prediction periods, replicating the univariate and multivariate results from before. Regressions (3), (4), and (5) run OLS no overlap (OLSNO) regressions with prediction horizons of one year or three years. OLSNO represents the process in which intermediate observations are dropped and OLS inferences are used. OLSNO is not perfectly efficient due to the dropped observations, but it preserves the validity of the inferences.

Regressions (3) runs a prediction with a one-year horizon. The coefficient and R^2 is exactly four times that of the quarterly regression. However, note that the power of the regression in terms of t-statistic does not increase, in line with the theory of Harri and Brorsen (2009). Similarly, regression (4), which predicts a three-year horizon has about double the coefficient and R^2 as regression (3), has slightly lower t-statistics. That the coefficient and R^2 does not triple, and that the power goes down, is all indicative of the fact that the horizon of the prediction power does start breaking down sometime between one and three years. Three years is also sufficient to see substantial decay in the persistence in the household risky ratio predictor.

Regression (5) runs a one-year-horizon multivariate regression. As expected from theory, the coefficients and R^2 of the one-year regression is about four times the one-quarter numbers. The loss in efficiency from the dropped observations is also clearer here as the power tends to be somewhat lower.

Regressions (6), (7), and (8) replicate regressions (3), (4), and (5) respectively, except instead of dropping intermediate observations and using OLS, it uses Newey-West regressions with 4+k periods of lag, where k is the horizon of the prediction expressed in quarters. The coefficients are roughly the same as the OLSNO regressions since Newey-West uses OLS for point estimates; the small deviations are due to the fact that intermediate observations are no longer dropped. The t-statistics however all seem higher than OLSNO. This is due to two reasons: the efficiency is higher since no observations are dropped, and less desirably, the Newey-West standard errors are biased downwards (Harri and Brorsen 2009). This is evident in the fact that the t-statistic actually increases between one-year and three-year regressions, whereas theory dictates that it should decrease.

To conclude, table 8 demonstrates the extension of the household risky ratio predictor to longer-horizons. The R^2 does increase from .054 for one quarter to .223 for one year and .401 for three years, which as an aside is quite high for known predictors. While the R^2 and coefficient increases, the power of the tests and t-statistics are not better. Sometimes power is worse due to dropped observations in the OLSNO regressions. The Newey-West regressions illustrate the complexity of correcting the inferences of overlapping regressions correctly. To maximize power then and minimize inference assumptions, this paper chooses to use disaggregated quarterly regressions, in line with the recent theory of Harri and Brorsen (2009)

Section 4: Empirical Decomposition of the Household Risky Ratio

This section analyzes the empirical sources of predictability that the household risky ratio has for forward returns. This section decomposes the household risky ratio into components related to household tilt, valuation ratios, and issuances. Then, I run univariate and multivariate regressions against each of these components to see what the source of predictability is.

A. Define the components theoretically

In the section below, I will decompose the household risky ratio into the following three components, with terms to be defined below:

$$\log(Risky\ Ratio_t) = \log(Household\ Tilt_t) + Valuation\ Ratio\ Part_t + Issuance\ Part_t$$

$$Valuation\ Ratio\ Part_t \perp Issuance\ Part_t$$

Consider again the household risky ratio, the ratio of household holdings of high-risk assets to low-risk assets. The first component of this ratio is the portion that is specific to the household, which is termed the *household tilt*. This can be defined as the portion of the household risky ratio that is not common to the economy. To define the household tilt, this paper then first defines an economy-wide risky ratio, termed the *total risky ratio* as follows:

$$Total\ Risky\ Ratio = \frac{Total\ Risky\ Assets}{Total\ Assets}$$

And calculate the difference between the total risky ratio and household risky ratio as:

$$Household\ Tilt = \frac{Household\ Risky\ Ratio}{Total\ Risky\ Ratio}$$

Note that this also gives:

$$Household\ Tilt = \frac{\%\ Risky\ Assets\ Held\ by\ Households}{\%\ NonRisky\ Assets\ Held\ by\ Households}$$

And

$$\log(Household\ Risky\ Ratio) = \log(Total\ Risky\ Ratio) + \log(Household\ Tilt)$$

$Household\ tilt$, is then a measure of the household's holding of high risk over low-risk assets relative to the entire economy. Thus, $household\ tilt$ is cleansed of all economy-wide effects, effects that will be discussed later on that should affect high risk and low risk assets in general, having nothing to do with households in particular.

In this section, preference is also given to the log version of the decomposition due to the additivity of the components. Other sections use the raw non-log risky ratio for simplicity and transparency of construction. As shown in Appendix Table A2, the predictive properties of the log household risky ratio are almost identical to the household risky ratio. This is a natural consequence of the fact that the variation of household risky ratio is small relative to the neighborhood in which the log function is locally linear, and so log household risky ratio is nearly an affine transform of the household risky ratio.

The second component of the household risky ratio, encompassed entirely within the log total risky ratio, is the idea of a valuation ratio or past price changes. If the market value of all high risk assets homogenously doubles overnight then the total risky ratio would double as well. Note that this effect is completely subsumed in total risky ratio part of household risky ratio: there is nothing special about households. Many authors in the literature (Poterba Summers 1988, Fama French 1988) have demonstrated that past prices changes, especially at the three to four year level, are good negative predictors of future returns. Looking at Figure 4 the predictability of different horizons, which can be seen as a type of cross correlogram, it is seen that indeed much of predictability of the household risky ratio becomes strong within the last three to four years of lags. To the extent then that the total risky ratio is capturing price changes, the total risky ratio has predictive power.

More fundamental than price changes are valuation ratios. Past returns are only weak predictors of future returns, and really the fundamental predictability comes from valuation ratios like D/P, CAPE, and B/M (Cochrane 2008). When run in a multivariate regression, the predictive power of past price changes is almost always subsumed by valuation ratios. As exposited by Campbell Shiller (2005), the predictive power comes from the fact that if earnings growth is difficult to predict, which is empirically the case, then CAPE changes must be attributed to future changes in expected return.

A similar mechanism is at work thinking of the total risky ratio as a valuation ratio. Consider assets in the economy to be paying off a safe stream securitized into the present through low risk assets, and a high risk stream securitized into the present through a high risk asset. Then as the required risk premium on high-risk assets decreases, this will push up the valuation of the high risk assets with respect to low risk assets, increasing the total risky ratio. Thus, the total risky ratio acts as a valuation ratio in predicting future returns in the same way that CAPE does.

To capture this portion of the household risky ratio, termed the *valuation ratio part,* this paper projects the total risky ratio onto a standard valuation ratio: the CAPE of Campbell Shiller (1988):

$$\log(Total\ Risky_t) = \beta_0 + \beta_1 CAPE_t + \beta_2 CAPE_t^2 + \epsilon_t$$

$$Valuation\ Ratio\ Part := \log(\widehat{Total\ Risky_t})$$

Note then risky valuation as defined here is just an affine transform of CAPE and its square.

To arrive at the final component of the predictive power of household risky ratio, consider that the valuation ratio analysis above considers only the part of total risky ratio with a fixed quantity of assets. In reality, issuances and redemptions play a role in both the composition of the total risky ratio and its predictive power. In particular, Baker Wurgler (2000) studies corporation issuance of assets. They show that when corporations issue a higher fraction of equities in a given period, there are lower future real returns on the equity market. After this seminal work, many other papers confirm that issuance, especially by corporations, are biased towards assets that are overvalued and will return less in the future.

Then in the projection above, a final component *risky issuance* can be defined:

$$\log(Total\ Risky_t) = \beta_0 + \beta_1 CAPE_t + \beta_2 CAPE_t^2 + \epsilon_t$$

$$Issuance\ Part := \epsilon_t$$

The *issuance part* reinforces the *valuation ratio part* even further. During good times, high-risk assets are overvalued and thus already have a large total valuation. Exactly during this time, corporations are issuing even more high-risk assets that are overvalued. The high-risk asset valuation measure then receives two reinforcing shocks during this time: a positive price shock and a positive quantity shock.

Putting all three components together, the accounting relationship below holds:

$$\log(Household\ Risky\ Ratio_t)$$
$$= \log(Household\ Tilt_t) + Valuation\ Ratio\ Part_t + Issuance\ Part_t$$

$$Valuation\ Ratio\ Part_t \perp Issuance\ Part_t$$

Thus, the log household risky ratio can be thought of as composed of three components: a household tilt variable, a valuation ratio part, and an issuance part. The last two are guaranteed to be orthogonal by mechanical construction. All three variables turn out to be decently orthogonal empirically as well.

B. Estimation of Components

In this subsection, I estimate the predictive powers of the components as defined above, which allows me to decompose the predictive power of the log household risky ratio. Table 9 displays the results of predicting future equity premium using the log household risky ratio and its decomposition above: the log household tilt, the valuation ratio part, and the issuance part.

Regression (1) replicates the baseline univariate regression, this time with the log household risky ratio. Regression (2) runs the same regression with log household tilt variable. The predictive power of the log household tilt variable is quite significant (t=-2.32) and has decent predictive power (R^2 = .018). The predictive power of log household tilt is actually quite high on an absolute scale (compare to Table 2), and is about a third of the of the entire log household risky ratio. Also, note that the coefficient on the household risky ratio is -6.47 while the coefficient on log household tilt is -9.26. Thus the log household tilt is a relatively stronger driver component of the log household risky ratio. Unlike the entire log household risky ratio, the log household tilt component is much more orthogonal to the other predictors, both in theory and also can be seen by the lack of interaction between regression (2) and (7). Thus, the household tilt variable is that theoretically novel part of the household risky ratio predictor. It is the component of household risky ratio's power that cannot be explained empirically by previous variables.

The next regression (3), looks as the predictive power of the valuation ratio part of the household risky ratio. The t-statistic is moderately significant (t=1.79), and the R^2 is decent at .0113, about a fifth of the predictive power of the log household risky ratio. The coefficient is -3.46, which is about half that of the log household risky ratio. This all suggests that the valuation ratio part is a relatively weaker driver of the predictive power of the household risky ratio. Also by construction, the valuation ratio is a linear combination of CAPE and $CAPE^2$, so all the predictive properties of CAPE carry over. Therefore, even though the t-statistics are not large, it is known from myriad of studies with longer term data (Campbell Shiller 1988) and rolling predictions (Campbell Thompson 2008) that the valuation ratio part must be a good predictor of the equity premium.

Regression (4) looks at the predictive power of issuance part. This variable is significant with a t-stat of 2.39, and an R^2 of .0164, again about a third of the predictive power of the log household risky ratio. The coefficient at -7.68 is also slightly higher than that on the log household risky ratio, demonstrating this is an important component of the predictive power of the household risky ratio. By construction, it is also orthogonal to the valuation ratio part. In the above dimensions, the issuance part is quite similar to the log household tilt variable.

However, unlike log household tilt, the issuance part is not orthogonal to known predictors. Regression (9) adds the equity share of issuances to a regression with the issuance part. Both the magnitude and t-statistic of the issuance part are halved by equity share of issuances, showing the close interaction of the two terms. Also, by adding all other covariates in regression (10), it is seen that the predictive power of the issuance part is nearly driven to zero. Thus, the issuance part can be seen to be picking up the predictive power of already known predictor variables, chiefly equity shares of issuance.

The valuation ratio part by definition is a function of CAPE. However, compare univariate regression (3) with a multivariate regression with all covariates besides CAPE, regression (8). It is seen that adding covariates does not reduce the predictive power of risky valuation. If anything, it is increased somewhat. This is to be expected from the fact that none of the other covariates are valuation ratios.

Finally, consider regressions (6) and (7), which are multivariate regressions with a full set of covariates. Regression (7) only contains the log household tilt portion of the household risky ratio, while regression (8) contains the entire log household risky ratio, but note that the adjusted R^2 is nearly identical, as well as the t-statistics. This provides evidence once again that the marginal predictive power of the log household risky ratio on top of the four covariates is isolated mainly to the log household tilt.

In summary, it is possible to decompose the log household risky ratio into three components: log household tilt, a valuation ratio part which is a function of CAPE, and the remaining issuance part. In a univariate sense, all three components are important and significant. With respect to known predictors, the predictive power of the log household risky ratio arises marginally from the log household tilt.

Section 5 Discussion

A. Ruling Out the Modigliani Miller Explanation

One possible explanation of the household risky ratio predicting future lower equity premium could be simple Modigliani Miller, as pointed out by Baker Wurgler (2000). In particular, Modigliani Miller posits that weighted average cost of capital is the same regardless of how corporations fund themselves. Therefore, as the amount of high risk assets in the economy increase with respect to low risk assets, the high risk assets effectively become less risky and command a lower return. Thus the household risky ratio could predict negative future equity returns just as an accounting artifact.

However a rough calibration shows that Modigliani Miller cannot explain anywhere near the size of the effect observed. Between the lowest and highest terciles, the household risky ratio doubles. Modigliani Miller would predict a halving in excess returns. However, in reality excess returns drop about ten times. This order of magnitude difference is similar found in the equity share of corporate issuances in Baker Wurgler (2000).

B. Welfare Effects

The empirical result that the household risky ratio negatively predicts the equity premium does not take a stand on whether this predictability reflects a rational risk factor or a misoptimization. Assume the latter case. Then what is the Sharpe ratio lost to the household sector from sizing out of the equity market exactly when it is performing well? What is the loss compared to a constant-fraction-hold benchmark, or the optimal conditioning on predictors benchmark? If it is also assumed that households have log utility and no exogenous sources of time-varying utility then return loss can also be calculated.

I calculate three investment possibilities. The first is the actual returns and Sharpe ratio realized by the household sector assuming the fraction of high risk assets is the fraction households invest in the market index at any given moment, and the fraction of low risk assets in the fraction households invest in treasury bills in any given moment. This gives a Sharpe ratio of .267 and annualized log excess returns of 3.26% assuming log utility. The next possibility assumes that households always hold 70% equities and 30% treasury bills, the unconditional average amount of high risk assets versus low risk assets held by the households. This results in a Sharpe ratio of .311 or excess log returns of 3.76%. This represents an increase of 50 bps over the baseline, or 14% increase over the base amount.

If households instead were to scale in optimally assuming log utility, then formula (14) of Campbell Thompson gives an increase of about 3x the current baseline rate.

$$2.9 = \left(\frac{R^2}{1 - R^2}\right)\left(\frac{1 + S^2}{S^2}\right)$$

With $S^2 = (.31)^2 = 0.0967$. $R^2 = .206$.

This represents a return of 11% assuming log utility and a Sharpe ratio of .533. Compared to actual outcomes, in case of optimal conditioning on predictors, the returns are more than triple, and the Sharpe ratios are more than double. While 11% may seem high, this is not out of line with the strength of the predictor. Also, the actual amount that could be realized by a real agent estimating out of sample would be less than this optimal ideal (Welch Goyal 2008, Pastor Stambaugh 2009). Thus, mistiming represents a substantial welfare loss for the household sector.

Section 6: Extensions

A. A theory for demand systems

This paper has examined the empirical phenomenon of the prediction power of the household risky ratio and decomposed the prediction power into various empirical components. Here, I explore some ideas behind why the household risky ratio might have such predictive power in theory.

One approach to the theory is to see high risk assets and low risks assets as being cleared in markets or demand systems with different players, each with a net demand curve. Market clearing happens at the price and quantity that sets total net demand to zero. Such demand system view goes back to the seminal work of Brainard and Tobin (1968).

In this view, the household sector has a demand curve for high risk assets. The demand curve receives shocks that are exogenous, the source of which I will examine below. After a positive demand shock, households demand a higher quantity of the high risk asset for the same price. Assuming the remaining sectors' demand curves remain constant, this translates to higher prices on the high risk assets, or lower future expected returns. In this way, higher quantities and valuations of high risk assets held by the household translate to a negative relationship to future returns.

Recent literature like Baker Wurgler (2000) speaks to this demand system view of high risk versus low risk assets. They show that corporations' supply curve of high risk versus low risk (equity versus debt) rationally responds to prices. When equity prices are high, and hence future expected returns are low, corporations supply more equities. Baker Greenwood Wurgler (2003) show a similar phenomenon between maturities in the debt market: when the yield of a maturity is particularly low, and hence the price particularly high, corporations tend to issue at that maturity. Greenwood Vayanos (2008) is more along the lines of this paper, in showing that government issuances of bonds are supply shocks: when the government exogenously issues excess bonds at a certain maturity prices go down, and yields are higher in the future.

It is important to note that in such a system a demand shock generally should cause both price and quantity responses, but the relative amount of each must be determined by the elasticity of the supply curve. That the response is not purely in price as evidenced here and in Baker Wurgler (2000) shows that securities are not in perfectly inelastic supply. A theoretical basis for this is apparent in that corporations can always start new real projects that are funded. That the response is not purely in quantities suggests that the supply is not perfectly elastic. Corporations need time to put new projects online, and new projects have aggregate diminishing returns in the economy as in the model of Solow (1956).

The fact that Baker Wurgler (2000) and this paper examine similar sets of markets explains why in the tables the issuance share of equities is so related to the issuance component of the household risky ratio. They examine flows by corporations, a subset of the supply side, while this paper examines stocks by households, a subset of the demand side.

On a first order then, some theoretical sense can be made out of the empirical phenomenon present in this paper by using a theory of demand systems and price pressure.

B. *What Drives Demand Shocks?*

Following the above idea that variation in household risky ratio is caused by demand shocks, this motivates the question of what causes the demand shocks.

One possibility is rational time variation in risk premiums and preferences. The underlying cause could be the same as that behind business cycles. A proximal model might include Campbell Cochrane (1999) habits. In particular, as stock returns receive a positive shock, excess consumption increases and effective risk aversion decreases. This justifies both a quantity shift from low risk to high risk assets as well as price increase as future required returns decrease. To test this model, a model of habit and surplus consumption could be calibrated, and habit can be correlated with the household risky ratio to see if a relationship exists.

Another strand of thought might explain the demand shocks as arising from sentiment. It is known that perfect optimization often does not describe individuals and even the firm (Laibson et al 1998; Zhang 2013). This may be seen as untestable as if the causes are fundamentally from outside the economy, then no immediate predictors are available, besides perhaps survey evidence, to validate such model. However, many behavioral models accept economic factors as the driver of psychology. For example Greenwood Shleifer (2013) posit that people have extrapolative beliefs: they believe that stocks will go up more following a move upwards. This case is testable as past prices can be used to see whether they relate to the household risky ratio.

C. *Further Data*

A central point of this paper is the empirical predictive power of the household risky ratio. A clear and transparent way to extend the empirical power would be to extend the duration of the data series as much as possible, in the style of Robert Shiller. Currently the results are based on more than 60 years of data, so if the data series were doubled, the series would extend back to around 1890. A first-order advantage of such a dataset extension is that it serves as a true out-of-sample test of the hypothesis above, since this paper is uncontaminated by observation of data before 1951. The true out-of-sample test can be used to validate the predictor in a way immune to any claims of data snooping, as well as test the stability of the coefficient estimated. The predictive properties can be understood much better by using the entire expanded series with rolling out-of-sample predictors.

Of course, data expansion has limitations. As is generally the case, data further back in history are noisier due to less advanced data collection technologies and more data that has been lost through time. Even common price series such as equity returns and price-earnings ratios become significantly lower quality before the 1920s. Quantity data such as that used in this paper would be even rarer. Further, for dates far enough back, one must question even the existence of equity markets accessible to households. As a raw method of increasing power, historical data extension is less fruitful, especially with the already high t-statistics observed here.

Another possible extension is to extend the predictability results here to other datasets. If the value of household holding of risky assets predicts future returns on risky assets, then it might stand to reason that the household holding of other asset classes might negatively predict future returns on those areas as well. Preliminarily, this seems to hold with government bonds and corporate bonds. In this way, the ideas and evidence presented in this paper can be developed into a general theory of household tilt.

Finally, this paper can be seen as an extension of the usual price predictors into quantity space. One way to generalize this more is to look at how all the price and quantity variables flow into each other economy-wide. High risk and low risk assets are just two components of what is the household sector's savings stock. The savings stock is affected by investment flows, which is known to be highly procyclical. Tracing investment back to output gives the GDP as the source of this flow. In this space itself, GDP depends on factor prices paid to labor and capital, the latter of which household total assets forms a component. On the capital markets side, there are corporations who are issuing the securities and financial assets being used in the household risky ratio. These corporations translate funds raised to real investment and real projects with payoffs and risk profiles.

Financial economics often centers around theories of prices in the economy and how they relate to each other, especially theories involving the efficient market. Household high risk and low risk assets then is a first step of a journey towards looking at more quantity-type predictors and looking at the entire economic system to provide macroeconomic foundations for finance.

Section 7: Conclusion

This paper shows that the ratio of high risk assets to low risk assets held by the household sector, termed the household risky ratio, is a negative predictor of future equity premium. The predictive power is robust and strong: the univariate t-statistics are above 4, and the annualized R^2 is above .20. The predictive power remains even after variation in construction of the variable, first/second half of the time series, and adjustment for the bias of Stambaugh (1999). The predictive power also is not subsumed by popular predictors like CAPE, equity shares of issuances, term spread, and the cay.

The paper empirically decomposes the predictability into three roughly orthogonal components. First is a household tilt component representing the preference of households for high risk assets above and beyond the entire economy. The second is a valuation ratio part that is a function of the Campbell-Shiller CAPE. And the third is an issuance part that is the residual from the decomposition above. All three components play important roles in the predictability of the household risky ratio: the R^2 is divided generally evenly between them and the coefficient size is the same order of magnitude. The second and third components reflect known predictors in the literature, while the first, the household tilt, seems orthogonal to known predictors.

This paper adds to the literature understanding time variation in equity premia by looking at Federal Reserve Flow of Funds data. It follows the footsteps of Baker Wurgler (2000) in going beyond price predictors to quantity data. Additionally, this paper looks at economy-wide household sector quantity data in a first step at connecting the variation in equity premium to economic fundamentals.

Appendix:

Figure 1: Household High Risk and Low Risk Assets, 1951-2012. High risky (blue) and low risk (red) household holdings are from the Federal Reserve Flow of Funds data. High risk assets are defined as equity held directly or risky mutual funds. Low risk assets include credit instruments and deposits held directly. All dollars figures are nominal.

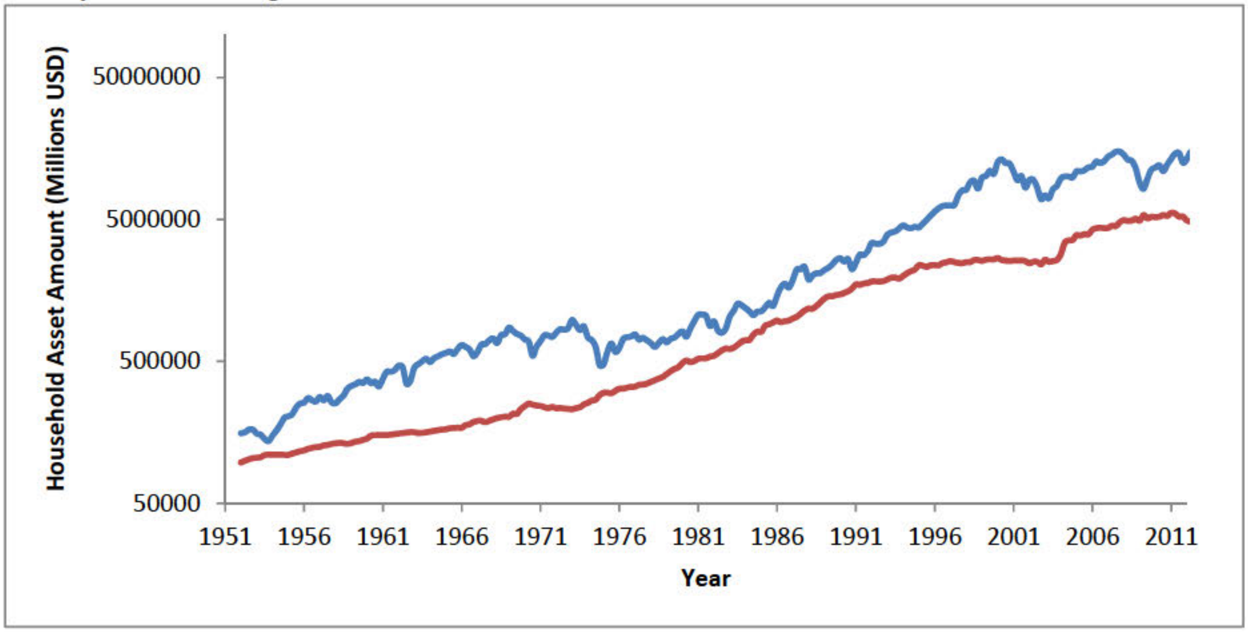

Figure 2: The household risky ratio 1951-2012. The household risky ratio, defined as the value of high risk household assets divided by low risk household assets from the Federal Reserve Flow of Funds data. High risk assets are defined as equity held directly or risky mutual funds. Low risk assets include credit instruments and deposits held directly.

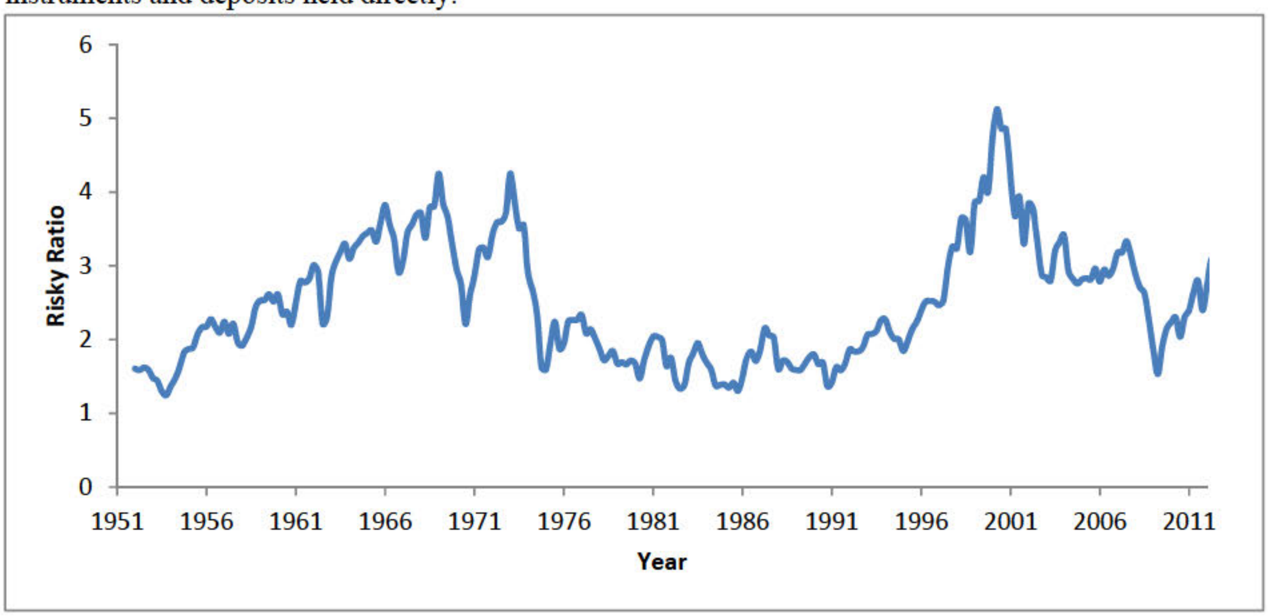

Table 1
Summary Statistics.

The summary statistics below covers all 243 quarters from the beginning of the *Federal Reserve Flow of Funds* series in 1951 to 2012. All data on sectors and their holdings are collected from the Federal Reserve Flow of Funds: Household equities is the series on households and nonprofit organizations, corporate equities, asset (FL153064105.Q); Household Equity Mutual Funds is the series on households and nonprofit organizations, mutual fund shares, asset (FL153064205.Q); Household Credit Markets is the series on households and nonprofit organizations, total currency and deposits including money market fund shares, asset (FL154000025.Q); Household Deposits is the series on households and nonprofit organizations, credit market instruments, assets (FL154004005.Q).

Household risky ratio is calculated as the sum of the first two series above divided by the sum of the second two series. Total risky ratio is calculated as the analogously except for the series corresponding to all sectors. Change in household risky ratio is the current period household risky ratio minus a one quarter lag. Log household risky ratio is the natural log of household risky ratio. Log total risky ratio is the natural log of the total risky ratio.

In future tables, R denotes the excess return calculated as the difference between the value-weighted CRSP minus the one-month Treasury bill rate from Ibbotson's in percentage points. S denotes the equity share in new issues as defined in Baker Wurgler (2000). CAPE denotes the ten year cyclically adjusted price to earnings ratio, defined as in Campbell Shiller (1988); Term Spread denotes the yield premium of ten year (from Robert Shiller's website) over one-month treasury bill rates. CAY is the consumption wealth ratio proxy defined by Lettau Ludvigson (2001). Risky valuation is the projection of CAPE on the household risky ratio. Risky Issuance is the residual from this projection.

	1951-2012		1951-1981		1982-2012	
Variable	Mean	SD	Mean	SD	Mean	SD
Panel A: Federal Reserve Flow of Funds Data Series						
Household Equities ($trillions)	4.289	5.356	.567	.258	8.009	5.445
Household Equity Mutual Funds ($tn)	2.441	3.443	.104	.090	4.776	3.579
Household Credit Market ($trillions)	1.437	1.593	.214	.099	2.525	1.459
Household Deposits ($trillions)	1.552	1.592	.226	.117	2.646	1.461
Panel B: Other Primary Data Series						
Excess Return R (%)	1.829	8.399	1.652	8.138	2.001	8.676
Lettau-Ludvigson CAY	.000	0.017	-0.004	0.01	0.004	0.021
Term Spread	2.275	1.231	1.794	0.941	2.743	1.302
Baker-Wurgler Equity Share of Issuance	0.186	0.095	0.224	0.082	0.143	0.090
Campbell-Shiller CAPE	18.887	7.534	15.735	4.597	21.963	8.525
Panel C: Derived Data Series						
Household risky ratio	2.509	0.815	2.540	.761	2.470	.866
Change in Household risky ratio	0.007	0.24	0.001	0.239	0.012	0.243
Log Household risky ratio	0.869	0.321	0.888	0.305	0.850	0.337
Total risky ratio	0.501	0.162	0.492	0.141	0.509	0.181
Log Total risky ratio	-0.747	0.340	-0.754	0.311	-0.740	0.367
Risky Valuation	-0.747	0.301	-0.865	0.243	-0.632	0.310
Risky Issuance	.000	0.157	0.111	0.096	-0.108	0.126

Figure 3:
Mean 1 and 3 year forward returns by household risky ratio, 1951-2012.

Mean 1-year forward (red) and 3-year forward (blue) value-weighted CRSP excess return by quintile of household risky ratio. The risky-ratio variable (228 observations) were ranked and then binned into quintiles. Quintile 1 below contains the lowest fifth of the value of household risky ratios, and quintile 5 contains the highest fifth. Then the excess return for the next 1 and 3 years on the value-weighted CRSP is calculated and plotted.

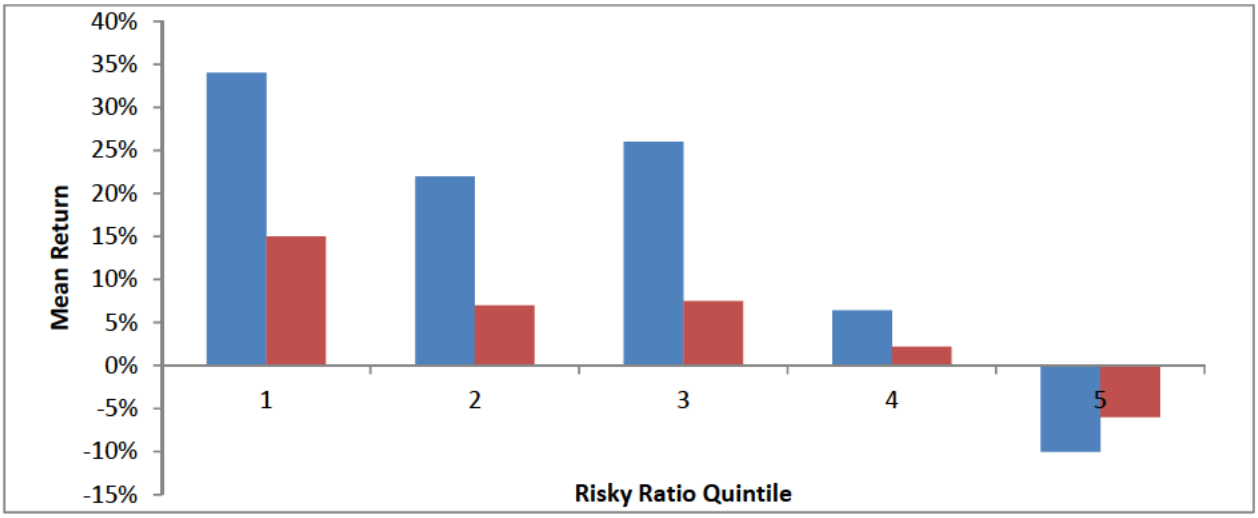

Figure 4:
Mean past and future annual equity returns from *t-2* to *t+12* by household risky ratio.

Each quarter is binned by terciles of household risky ratio – low (red), medium (grey), and high (green). Annual excess valued-weighted CRSP returns from 2-years past (t-2) to 12 years in the future (t+12) for each bin are plotted in the bar chart below.

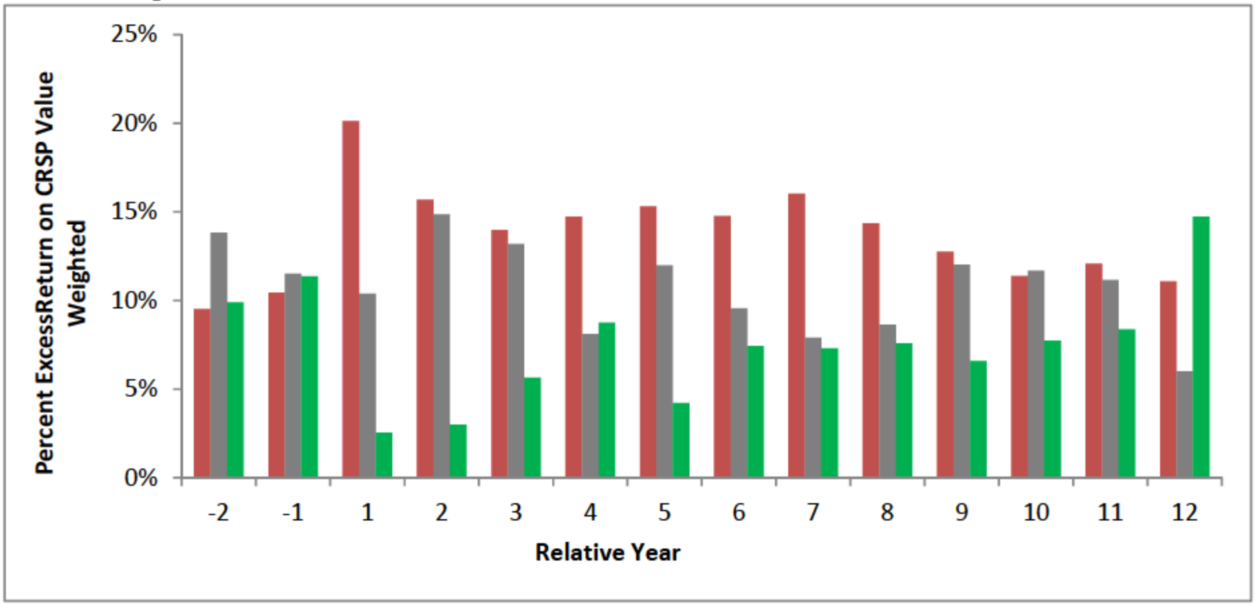

Table 2:
Univariate OLS regressions for predicting five-year-ahead market returns.

One-quarter ahead valued-weighted CRSP excess returns are regressed on a variety of predictors:

$$R_{t,t+1} = \alpha + \beta X_{t-1} + \epsilon_t$$

Where $R_{t,t+1}$ denotes the excess returns on the value-weighted CRSP for one quarter ahead and X_{t-1} variously denotes the household risky ratio, defined as household high risk assets divided by household low risk assets; the household risky ratio lagged an extra quarter; the household risky ratio lagged an extra year; equity shares of issues (Baker Wurgler 2000); Campbell Shiller (1988) 10-year cyclically adjusted price to earnings ratio (CAPE); the 10-year 1-month government obligation yield spread; and the consumption-wealth proxy CAY (Lettau Ludvigson 2001). t-statistics are heteroskedastic and autocorrelation robust (Newey-West) with 5 periods of lag.

Predictor	β	$t(\beta)$	α	$t(a)$	R^2
Household risky ratio	-2.49	[-4.16]	8.05	[5.42]	.054
Household risky ratio (lagged extra quarter)	-2.40	[-4.35]	7.84	[5.79]	.050
Household risky ratio (lagged extra year)	-1.94	[-2.94]	6.72	[4.20]	.032
Baker-Wurgler Equity Share	-7.32	[-1.34]	3.02	[2.83]	.003
Campbell-Shiller CAPE	-.13	[-1.68]	4.36	[2.81]	.011
Term Spread	.32	[0.87]	1.08	[0.99]	.000
Lettau-Ludvigson CAY	79.73	[2.51]	1.82	[3.49]	.022

Table 3:
Multivariate OLS Regressions for Predicting Excess Returns.

OLS regressions of one quarter ahead excess returns on the value-weighted CRSP on one period lagged multiple predictors:

$$R_{t,t+1} = \alpha + \beta_1 Risky_{t-1} + \beta_2 Risky_{t-2} + \beta_3 S_{t-1} + \beta_4 CAPE_{t-1} + \beta_5 Term_{t-1} + \beta_6 CAY_{t-1} + \epsilon_t$$

Where $R_{t,t+1}$ denotes the excess return of the value-weighted CRSP for one quarter forward; $Risky_{t-1}$ denotes the household risky ratio calculated by dividing household high risk assets over low risk assets, as collected from the Federal Reserve Flow of Funds; $Risky_{t-5}$ is the household risky ratio lagged one extra year; S_{t-1} denotes the equity share in new issues as defined in Baker Wurgler (2000); $CAPE_{t-1}$ denotes the ten year cyclically adjusted price to earnings ratio, defined as per Campbell Shiller (1988); $Term_{t-1}$ denotes the yield premium of ten year over one month federal government obligations; CAY_{t-1} is the consumption wealth ratio proxy defined by Lettau Ludvigson (2001); t-statistics are shown in brackets using Newey-West heteroskedastic and autocorrelation robust standard errors with 5 periods of lags. Regression (8) normalizes all predictors to have unit variance, while all other regressions use non-normalized predictors. N=226.

	Predictors Not Normalized							Predictors Normalized
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept								
$Risky_{t-1}$	-2.49	-2.86	-2.52	-3.76	-2.63	-2.21	-2.39	-1.94
	[-4.16]	[-2.69]	[-4.33]	[-4.20]	[-3.95]	[-3.52]	[-1.97]	[-1.97]
$Risky_{t-5}$		.32						
		[0.27]						
S_{t-1}			-10.49				-7.52	-.71
			[-2.13]				[-1.16]	[-1.16]
$CAPE_{t-1}$				.18			.02	.12
				[1.75]			[0.11]	[0.11]
$Term_{t-1}$					-.28		-.31	-.38
					[-0.73]		[-0.72]	[-0.72]
CAY_{t-1}						49.42	55.94	.95
						[1.51]	[1.24]	[1.24]
R^2	0.054	0.054	0.062	0.060	0.052	0.060	0.060	0.060

Table 4:
Robustness to First and Second Half

OLS regressions of one quarter ahead value-weighted CRSP excess returns on one period lagged household risky ratio.:

$$R_{t,t+1} = \alpha + \beta_1 Risky_{t-1} + \epsilon_t$$

Where $R_{t,t+1}$ denotes the excess return on the value-weighted CRSP for one quarter forward; $Risky_{t-1}$ denotes the household risky ratio calculated by dividing household high risk assets over low risk assets, as collected from the Federal Reserve Flow of Funds. The regression period is for the first half of the sample, the second half of the sample, and then the entire sample. *t*-statistics are shown in brackets using Newey-West heteroskedastic and autocorrelation robust standard errors with 5 years of lags. N=241

	First Half	Second Half	Entire Sample
$Risky$	-2.94	-2.15	-2.49
	[-3.34]	[-2.64]	[-4.16]
R^2	.066	.038	.054

Table 5:
Construction of the Data:

OLS regressions of one quarter ahead value weighted CRSP excess returns on one period lagged of the household risky ratio, defined in various ways:

$$R_{t,t+1} = \alpha + \beta_1 X_{t-1} + \epsilon_t$$

Where $R_{t,t+1}$ denotes the excess return on CRSP for one quarter forward; X_{t-1} denotes various household risky ratios calculated by dividing various household high risk assets over various low risk assets, as collected from the Federal Reserve Flow of Funds. The numerator may contain equities or mutual funds, and the denominator may contain credit markets or deposits. An X in the table below denotes the inclusion of each variable in the definition of X_t for that regression. *t*-statistics are shown in brackets using Newey-West heteroskedastic and autocorrelation robust standard errors with 5 periods of lags. N=241

	(1)	(2)	(3)	(4)	(5)
X_{t-1}	-2.49	-1.21	-1.05	-3.22	-2.34
	[-4.16]	[-3.40]	[-2.97]	[-3.48]	[-1.58]
Numerator has Equities	X	X		X	X
Numerator has Mutual Fund	X		X	X	X
Denominator has Credit Market	X	X	X		X
Denominator has Deposit	X	X	X	X	
Adj R^2	.054	.041	.034	.035	.008

Table 6:
Univariate regression under different inference assumptions.

The univariate regression:

$$R_{t,t+1} = \alpha + \beta Risky_t + \epsilon_t$$

Where $R_{t,t+1}$ denotes the excess return of value-weighted CRSP for one quarter forward; $Risky_t$ denotes the household risky ratio calculated by dividing household high risk assets over low risk assets, as collected from the Federal Reserve Flow of Funds. The OLS method uses standard inferences. The Newey-West method uses heteroskedastic and autocorrelation corrected standard errors with 5 years of lag. The Kendall (1954) correction is a point estimate adjustment. The Kothari-Shanken inferences use the bootstrap methodology to correct for small sample bias as outlined in Stambaugh (1999). Campbell-Yogo (2006) is a hypothesis test whose implementation is outlined in Campbell Yogo (2005). The Lewellen (2004) provides not an unbiased estimate but an upper bound for β and the p-value, and a lower-bound for the t-statistic.

	Newey-West	OLS	Kendall (1954)	Kothari-Shanken	Campbell-Yogo	Lewellen Bound
β on $Risky_t$	-2.27	-2.27	-1.79	-1.89		-.92
95% CI $\underline{\beta}$	-3.43	-3.55	-3.07	-2.89	-4.32	-1.55
95% CI $\bar{\beta}$	-1.11	-.993	-0.51	-1.12	-.03	-.29
t-statistic	-3.85	-3.50	-2.75			-2.88
p value	.0001	.0006	.0065	.0040	.0500	.0062

Table 7:
Dickey-Fuller GLS Test of Unit Root in the Household risky ratio.

The table displays the Dickey-Fuller GLS (DF-GLS) test for a unit root in the household risky ratio (Panel A) and changes in household risky ratio (Panel B). The household risky ratio is the ratio of household high risk assets over household low risk assets as reported by the *Federal Reserve Flow of Funds* data. For both the household risky ration and changes in the household risky ratio, the table displays for testing 1 lag and 14 lags (the default maximum tested by the Stata software used) of the DF-GLS test statistic, along with the 1%, 5%, and 10% hypothesis test cutoff values of the test-statistic. N=228.

Number of Lags	DF-GLS Statistic	Rejected?	1% Critical Value	5% Critical Value	10% Critical Value
	Panel A: Household risky ratio				
14	-1.578	No	-3.48	-2.813	-2.534
1	-2.287	No	-3.48	-2.919	-2.630
	Panel B: Changes in Household risky ratio				
14	-3.243 **	Yes	-3.48	-2.811	-2.532
1	-4.841***	Yes	-3.48	-2.922	-2.633

Table 8
Long Horizon Regressions

OLS regressions of k-period ahead value-weighted CRSP excess returns on multiple predictors lagged one quarter:

$$R_{t,t+k} = \alpha + \beta_1 Risky_{t-1} + \beta_2 S_{t-1} + \beta_3 CAPE_{t-1} + \beta_4 Term_{t-1} + \beta_5 CAY_{t-1} + \epsilon_t$$

Where $R_{t,t+k}$ denotes the excess return of the CRSP value-weighted holding return for k quarters forward as reported by CRSP; $Risky_{t-1}$ denotes the household risky ratio calculated by dividing household high risk assets over low risk assets, as collected from the Federal Reserve Flow of Funds; S_{t-1} denotes the equity share in new issues as defined in Baker Wurgler (2000); $CAPE_{t-1}$ denotes the ten year cyclically adjusted price to earnings ratio, defined as per Campbell Shiller (1988); $Term_{t-1}$ denotes the yield premium of ten year over one month federal government obligations; CAY_{t-1} is the consumption wealth ratio proxy defined by Lettau Ludvigson (2001).

Regressions are divided into three categories. For Disaggregated Returns, the prediction period k is always one quarter and so by construction there is both no overlap and no data dropped. t-statistics are shown in brackets using Newey-West heteroskedastic and autocorrelation robust standard errors with 5 periods of lags. For OLS No Overlap, the prediction period k is always ranges between one year (k=4) and three years (k=12). There is no overlap but data is dropped. t-statistics are shown in brackets using Eicker-Huber-White standard errors. For Overlap Corrected via Newey-West, there is no data dropped but overlap. *t-statistics* are calculated by Newey-West standard errors with 4+k periods of lags, a generalization of the number of lags used in the Disaggregated Returns method.

	Disaggregated Returns		OLS No Overlap – Intermediate Data Dropped			Overlap Corrected via Newey-West		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Period	1-Qtr	1-Qtr	1-Year	3-Year	1-Year	1-Year	3-Year	1-Year
$Risky_t$	-2.49	-2.39	-10.22	-23.93	-8.58	-9.21	-18.42	-9.25
	[-4.16]	[-1.97]	[-4.31]	[-3.36]	[-1.48]	[-4.84]	[-5.91]	[-2.40]
S_t		-7.52			-29.75			-26.76
		[-1.16]			[-1.03]			[-1.69]
$CAPE_t$		.02			-.22			.07
		[0.11]			[-.39]			[0.17]
$Term_t$		-.31			-1.74			-.44
		[-0.72]			[-1.05]			[1.06]
CAY_t		55.94			265.62			153.04
		[1.24]			[1.33]			[1.08]
R^2	0.054	0.060	0.223	0.401	0.031	0.192	0.365	0.236

Table 9: Comparison of the Household risky ratio, Total risky ratio, and Household Bias as Return Predictors:

OLS regressions of one-quarter forward value-weighted CRSP excess returns on multiple predictors lagged one quarter:

$$R_{t,t+1} = \alpha + \beta_1 \ln(Risky_{t-1}) + \beta_2 \ln(Household\ Tilt_{t-1}) + \beta_3 Valuation\ Ratio\ Part_{t-1} + \beta_4 Issuance\ Part_{t-1} + \beta_5 S_{t-1} + \beta_6 CAPE_{t-1} + \beta_7 Term_{t-1} + \beta_8 CAY_{t-1} + \epsilon_t$$

Where $R_{t,t+1}$ denotes the excess return on CRSP for one quarter forward; $\ln(Risky_{t-1})$ denotes the log household risky ratio calculated by dividing household high risk assets over low risk assets, as collected from the Federal Reserve Flow of Funds; $\ln(Household\ tilt_{t-1})$ denotes the log of the household tilt, the quotient between the household risky ratio and *total risky ratio* as calculated by dividing economy-wide high risk assets over economy-wide low risk assets, as collected from the Federal Reserve Flow of Funds; $Valuation\ Ratio\ Part_{t-1}$ is the linear projection of the total risky variable above onto CAPE and $CAPE^2$; $Issuance\ Part_{t-1}$ is the residual from the above projection; S_{t-1} denotes the equity share in new issues as defined in Baker Wurgler (2000); $CAPE_t$ denotes the ten year cyclically adjusted price to earnings ratio, defined as per Campbell-Shiller (1988); $Term_{t-1}$ denotes the yield premium of ten year over one month federal government obligations; CAY_t is the consumption wealth ratio proxy defined by Lettau Ludvigson (2001). *t*-statistics are shown in brackets using Newey-West heteroskedastic and autocorrelation robust standard errors with 5 periods lags. N=228.

	No Covariates					Covariates				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\ln(Risky_{t-1})$	-6.47					-6.44				
	[-4.12]					[-2.11]				
$\ln(Household\ Tilt_{t-1})$		-9.26			-14.44		-9.49			
		[-2.31]			[-3.39]		[-2.13]			
$Risky\ Valuation_{t-1}$			-3.46		-6.17			-5.94		
			[-1.79]		[-3.14]			[-2.75]		
$Risky\ Issuance_{t-1}$				-7.68	-6.50				-.13	-4.89
				[-2.39]	[-2.25]				[-.03]	[-1.32]
S_{t-1}						-7.73	-10.45	-13.50	-11.39	-9.70
						[-1.20]	[-1.73]	[-2.15]	[-1.62]	[-1.59]
$CAPE_{t-1}$						.02	-.24		-.20	
						[.15]	[-2.63]		[-2.24]	
$Term_{t-1}$						-.36	-.00	-.32	-.25	
						[-.84]	[.01]	[-.71]	[-.52]	
CAY_{t-1}						53.98	70.01	96.10	94.15	
						[1.26]	[1.67]	[2.39]	[2.20]	
R^2	.0571	.0172	.0113	.0164	.0664	.0638	.0635	.0568	.0458	.0236

Appendix A

As a preview of the data collection process, an excerpt of the *Federal Reserve Flow of Funds* is shown below. By gross assets, note that the largest sectors are domestic non-finance ($77 trillion) as well as domestic finance ($70 trillion), followed by households and nonprofit organizations ($54 trillion). This ordering changes significantly when accounting for net assets – assets minus liabilities. In this case, due to low relative leverage, households are by far the largest sector ($41 trillion), aligning with the idea that households are the ultimate owners of all assets. The domestic nonfinancial sector is moderately leveraged, leading them to be second in net assets ($27 trillion). The financial sector is relatively miniscule in net size due to their extreme leverage ($4 trillion).

Figure A1: Flow of Funds Excerpt: Size of Sectors

The following excerpt show the first three rows of table Z.1 of the March 7, 2013 *Federal Reserve Flow of Funds*. The rows show the financial assets and liabilities of all sectors defined by the *Flow of Funds*. Within each sector "A" specifies assets and "L" specifies liabilities.

Flow of Funds Matrix for 2012

Z.1, March 7, 2013

(Billions of dollars; All Sectors -- Assets and Liabilities)

		Households and Nonprofit Organizations		Nonfinancial Business		State and Local Governments		Federal Government	
		A	L	A	L	A	L	A	L
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	Total financial assets	54390.5	--	19560.3	--	2095.4	--	1416.8	--
2	Total liabilities and equity	--	13453.1	--	44080.0	--	3732.7	--	13468.7
3	Total liabilities	--	13453.1	--	19820.5	--	3732.7	--	13468.7

Domestic Nonfinancial Sectors		Domestic Financial Sectors		Rest of the World		All Sectors		Instrument Discrep-ancy
A	L	A	L	A	L	A	L	
(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
77462.9	--	69869.8	--	19384.4	--	166717.2	--	-7418.8
--	74734.5	--	70743.7	--	13820.2	--	159298.4	--
--	50475.0	--	65763.9	--	9074.5	--	125313.4	--

Is the household risky ratio robust to definition? It seems equally natural to define the household risky ratio as the inverse or log. Here I take the other reasonable transforms of the household risky ratio and regress against those. I find that the R^2 is nearly identical, and the significance seems similar as well. This is not surprising because the household risky ratio does not vary a lot with respect to its baseline level. Therefore, under both the log and inverse transformations, the functions are locally linear and preservation of the regression as seen in Table A3 is expected.

Table A2: Regression of forward CRSP return on various transformations of the household risky ratio.

OLS regressions of one quarter future equity market CRSP returns on multiple transformations of the household risky ratio:

$$R_{t,t+1} = \alpha + \beta_1 X_{t-1} + \epsilon_t$$

Where $R_{t,t+1}$ denotes the one quarter forward excess return as reported by CRSP. X_{t-1} variously denotes $Risky_{t-1}$ the household risky ratio calculated by dividing household high risk assets over low risk assets, as collected from the Federal Reserve Flow of Funds; or X_{t-1} denotes 1/ $Risky_{t-1}$ the multiplicative inverse of $Risky_{t-1}$; or X_{t-1} denotes ln $Risky_{t-1}$ denotes the natural log of $Risky_{t-1}$. *t*-statistics are shown in brackets using Newey-West heteroskedastic and autocorrelation robust standard errors with 5 quarters lags. N=228

	(1)	(2)	(3)
Intercept	8.05	-4.87	7.43
	[5.42]	[-2.68]	[5.53]
$Risky_{t-1}$	-2.49		
	[-4.16]		
$\frac{1}{Risky_{t-1}}$		15.13	
		[4.12]	
$\ln(Risky_{t-1})$			-6.47
			[-4.12]
R^2	0.054	0.059	.057

Table A3 shows that the household risky ratio has a median value of about 2.3, and varies the majority of the time between 1.7 and 3.3. The low variance with respect the baseline value explains the linearity seen in Table A2.

Table A3. Values of Percentiles of the Household risky ratio:

The table displays values of select percentiles of the household risky ratio variable. The household risky ratio is the ratio of household high risk over household low risk assets as reported by the *Federal Reserve Flow of Funds* data. Percentiles are selected to be approximately 1/6 of the distribution. 95% confident intervals are given by the Thiel-Sen method.

Percentile	Value	95% Conf Interval
17%	1.685	(1.62 , 1.77)
33%	1.996	(1.87 , 2.09)
50%	2.304	(2.20 , 2.52)
67%	2.866	(2.73 , 3.03)
83%	3.352	(3.23 , 3.55)

References

Baker, M., Greenwood, R., & Wurgler, J. (2003). The maturity of debt issues and predictable variation in bond returns. *Journal of Financial Economics*, *70*(2), 261-291.

Baker, M., & Wurgler, J. (2000). The equity share in new issues and aggregate stock returns. *The Journal of Finance*, *55*(5), 2219-2257.

Boudoukh, J., Richardson, M., & Whitelaw, R. F. (2008). The myth of long-horizon predictability. *Review of Financial Studies*, *21*(4), 1577-1605.

Brainard, W. C., & Tobin, J. (1968). Pitfalls in financial model building. *The American Economic Review*, *58*(2), 99-122.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of financial economics*, *18*(2), 373-399.

Campbell, J. Y., & Cochrane, J. H. (1999). By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *The Journal of Political Economy*, *107*(2), 205-251.

Campbell, J. Y., & Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of financial studies*, *1*(3), 195-228.

Campbell, J. Y., & Shiller, R. J. (1988). Stock prices, earnings, and expected dividends. *The Journal of Finance*, *43*(3), 661-676.

Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average?. *Review of Financial Studies*, *21*(4), 1509-1531.

Campbell, J. Y., & Yogo, M. (2005). Implementing the econometric methods in “Efficient tests of stock return predictability”. *Unpublished working paper. University of Pennsylvania*.

Campbell, J. Y., & Yogo, M. (2006). Efficient tests of stock return predictability.*Journal of financial economics*, *81*(1), 27-60.

Cochrane, J. H. (2008). The dog that did not bark: A defense of return predictability. *Review of Financial Studies*, *21*(4), 1533-1575.

Cochrane, J. H., & Piazzesi, M. (2005). Bond risk premia. *The American economic review*, *95*(1), 138-160.

Fama, E. F., & French, K. R. (1988). Permanent and temporary components of stock prices. *The Journal of Political Economy*, 246-273.

Fama, E. F., & Schwert, G. W. (1977). Asset returns and inflation. *Journal of financial economics*, *5*(2), 115-146.

Gordon, M. J. (1962). *The investment, financing, and valuation of the corporation*. Greenwood Press.

Greenwood, R., & Shleifer, A. (2013). *Expectations of returns and expected returns* (No. w18686). National Bureau of Economic Research.

Greenwood, R., & Vayanos, D. (2010). Price pressure in the government bond market. *The American Economic Review*, *100*(2), 585-590.

Laibson, D. I., Repetto, A., Tobacman, J., Hall, R. E., Gale, W. G., & Akerlof, G. A. (1998). Self-control and saving for retirement. *Brookings Papers on Economic Activity*, *1998*(1), 91-196.

Harri, A., & Brorsen, B.W. (2009). The Overlapping Data Problem. *Quantitative and Qualitative Analysis in Social Sciences,* 2009(3),78-115.

Hjalmarsson, E. (2006). On the predictability of global stock returns. *rapport nr.: Working Papers in Economics*, (161).

Keim, D. B., & Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. *Journal of financial Economics*, *17*(2), 357-390.

Kendall, M. G. (1954). Note on bias in the estimation of autocorrelation.*Biometrika*, *41*(3-4), 403-404.

Kothari, S. P., & Shanken, J. (1997). Book-to-market, dividend yield, and expected market returns: A time-series analysis. *Journal of Financial Economics*, *44*(2), 169-203.

Lettau, M., & Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. *the Journal of Finance*, *56*(3), 815-849.

Lewellen, J. (2004). Predicting returns with financial ratios. *Journal of Financial Economics*, *74*(2), 209-235.

Nelson, C. R., & Kim, M. J. (1993). Predictable stock returns: The role of small sample bias. *The Journal of Finance*, *48*(2), 641-661.

Newey, W. K. (1993). 16 Efficient estimation of models with conditional moment restrictions. *Handbook of statistics*, *11*, 419-454.

Newey, W. K., & West, K. D. (1986). A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix.

Pástor, Ľ., & Stambaugh, R. F. (2009). Predictive systems: Living with imperfect predictors. *The Journal of Finance*, *64*(4), 1583-1628.

Solow, R.M. (1956). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics,* 70(1), 65-94.

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics*, *54*(3), 375-421.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 817-838.

Valkanov, R. (2003). Long-horizon regressions: Theoretical results and applications. *Journal of Financial Economics*, *68*(2), 201-232.

Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, *21*(4), 1455-1508.

Zhang, Fan A., (2013): “Household risky ratio and the Equity Premium,” *Harvard University Working Paper.*