Expected Macroeconomic Conditions and Market Risk Premium: Evidence from a Term Structure of Macroeconomic Forecasts

Yizhe Deng [a 1] Yunqi Wang [b 2] Ti Zhou [b *]

First version: Jan, 2020
This version: Jan, 2022

[†]We would like to give special thanks to Jonas Eriksen (discussant) whose comments greatly improved the paper. We also thank Jian Chen, Weikai Li, Guohao Tang, Jun Tu, Chu Zhang, Ran Zhang, and participants at Tianjing University, Xiamen University, 2020 China International Risk Forum, the 18th International Symposium on Financial System Engineering and Risk Management, 4th China Finance Scholar Forum, 2021 Northern Finance Association Conference, 2021 FMA Annual Meeting for their helpful comments and suggestions. All remaining errors are ours.

[a]Hong Kong University of Science and Technology, HK SAR.
[b]Department of Finance, Southern University of Science and Technology
[1]E-mail address: ydengao@connect.ust.hk
[2]E-mail address: wangyq2019@mail.sustech.edu.cn
[*]Corresponding author at: Department of Finance, Southern University of Science and Technology, Xueyuan Rd. 1088, Shenzhen 518000, China. Tel.: +86-755-88018610. E-mail address: zhout@sustech.edu.cn

Expected Macroeconomic Conditions and Market Risk Premium: Evidence from a Term Structure of Macroeconomic Forecasts

ABSTRACT

We construct an *aligned macro risk index* using survey forecasts of future economic activities with the purpose of tracking the equity premium. The index aggregates diverse pieces of forward-looking information about output, housing, inflation, and labor market conditions. Empirically, it significantly predicts stock market returns, and outperforms popular forecasting variables and well-recognized macroeconomic predictors. It can also produce sizable out-of-sample utility gains to investors in real time. We present rich evidence demonstrating that the index tracks the rising equity premium induced by heightened risk and risk aversion during recessions, and its predictive power mainly stems from a discount rate channel. We also show that a long-term macro risk index based on the term structures of survey forecasts has stronger predictive ability for long-term returns, revealing that ex-ante perceived fundamental risks and equity premium agree in horizon. Taken together, our findings establish a sound relation between the time variation of equity premium and the cyclical risks in macroeconomy.

***JEL classification*: C53; G11; G12; G17**

***Keywords*: Macroeconomic expectation, Return predictability, Business-cycle risk, Discount rate, Term structure, SPF**

1 Introduction

Stock market returns are time-varying and correlated with business cycle (Fama and French, 1989). At least as early as Fama (1991), researchers have recognized that the predictability of market returns can be consistent with time-varying risk premia in an efficient frictionless market. Over past two decades, a number of equilibrium asset pricing models have been developed to rationalize such predictability. Despite of differences in mechanisms, a common implication of these models is that the equity premium should be relatively higher during adverse economic times, such as recessions, when average investors have higher marginal utility, and should be relatively lower during good times (Cochrane, 2011, 2017). How to measure ex-ante good and bad economic times and to establish a robust relation between macroeconomic conditions and equity premium remains empirical challenging, however. Whereas numerous return predictors were proposed in the literature (Lettau and Ludvigson, 2010; Rapach and Zhou, 2013), a majority of them lose predictive power during recent periods; moreover, most of them fail to consistently beat the historical average forecast in out-of-sample tests (Goyal, Welch, and Zafirov, 2021; Welch and Goyal, 2008).

In this paper, we construct a macro risk index using the consensus forecasts on future macroeconomic activities from the Survey of Professional Forecasters (SPF) with the purpose of tracking the equity premium. An intriguing feature of the SPF is that it covers broad aspects of the macroeconomy, which enables us to explore the dynamic and complex relation between the equity premium and macroeconomic quantities. Specifically, we collect the forecasts for the current quarter, i.e., "nowcasts", on seven aspects of the macroeconomy, including the growth rates of real gross domestic product (GDP) and industrial production, recession probability, unemployment rate, corporate profit growth, housing starts growth, and inflation. We apply the Partial Least Square method (Kelly and Pruitt, 2013, 2015; Wold, 1966) to condense the rich pieces of information in these forecasts into a single predictive factor for the equity premium, which we refer to *the aligned macro risk index*, denoted as $\mathrm{MR}^{\mathrm{PLS}}$. Kelly and Pruitt (2015) show theoretically that PLS is an efficient dimension reduction technique which can extract the common variation of forecasting variables that is most aligned with the target of interest, and at the same time filters out irrelevant common noises.

Our macro risk index has three key advantages. First, it is built upon survey forecasts that capture investor *ex-ante* expectations about macroeconomic conditions in a model-free manner (Amato and Swanson, 2001; Croushore and Stark, 2001). The usage of survey data is free from restricted specifications about expectational dynamics (Choi and Robertson, 2020; Manski, 2004), and helps bypass the timing issue and data revision problem associated with conventional macroeconomic variable.[1] Second, the macro risk index is constructed without

[1]Ghysels, Horan, and Moench (2018) show that the OOS predictive power of macro variables on treasury bond returns is substantially weakened when vintage data rather than final revised macro data are used.

using asset prices. This largely eliminates the concern that the equity premium predictability, if any, arises from time-varying aggregate mispricings. Third, the information used in prevailing empirical predictive analyses is often a small subset of the investors' (Ludvigson and Ng, 2007). In regard to this multifaceted features of macroeconomy, our aligned index summarizes the information about various aspects of macroeconomic conditions, which could potentially contribute to more robust predictive power.

Over the sample period of 1969Q1 through 2019Q4, the common factor extracted by PLS from the SPF macro variables forecasts significantly predicts market excess returns up to 12 quarters. At the quarterly forecast horizon, the aligned macro risk index generates in-sample and statistically significant out-of-sample (OOS) R^2s of 5.75% and 3.12%, respectively. The predictability of it remains significant after controlling for the 16 economic predictors in Welch and Goyal (2008). Furthermore, the OOS forecast encompassing test (Harvey, Leybourne, and Newbold, 1998) confirms that the aligned macro risk index contains incremental information relative to popular predictors. It is worth mention that the SPF also provides a term structure of macro variable forecasts covering one to four quarters into the future besides the current quarter, and we use them to construct a long-term macro risk index. We find that the long-term macro risk index exhibits more substantial predictive power for long-term market returns relative to the one only relying on nowcasts. Its in- and out-of-sample R^2s are 8.65% and 2.84% at the two-quarter horizon and 27.09% and 17.91% at the 12-quarter horizon. Our findings suggest that forecasts of the future macroeconomic activities can predict future market returns and more importantly, the information content of the macro forecasts and the return predictability agree in horizon.

Moreover, we follow Campbell and Thompson (2008) and Rapach, Strauss, and Zhou (2010) to assess the economic significance delivered by the macro risk index's forecasting ability via an asset allocation analysis. Under the risk aversion levels of three and five, the gains in annualized certainty equivalent return produced by the macro risk index are 324 and 195 basis points, respectively, relative to the historical mean benchmark. These economic gains far exceed not only those generated by conventional predictors but those by the buy-and-hold strategy.

To understand the predictive power of the aligned macro risk index, we first show that the SPF data can predict the growth rates of real consumption, real GDP, real labor income, and industrial production up to eight quarters. This indicates that the forecasts on macroeconomic activities reflect the rational belief of professional forecasters and contain forward-looking information about future business conditions. Second, we highlight the efficacy of PLS in incorporating information from a variety of macro variables forecasts. Although there is a few SPF variables that are significant for in-sample predicting, their OOS forecasting performance is rather unstable. In contrast, the PLS macro risk index is a linear combination of individual SPF variables where the weights are adjusted by the covariance with the forecast target. Since it effectively aggregates the information from a set of SPF variables, the PLS index mitigates the impact of model uncertainty and structural breaks in individual predictors. For comparison, we also use the first principal component (PC) of the seven SPF variables as a measure of macro risk index. By its econometric design, the first PC is a combination of the seven SPF variables that explains the most covariance among them, whereas it ignores the information about the forecast target. This explains why the PLS macro risk index outperforms the PC approach by a large margin in practice. Similarly, because the PLS method assigns dynamically-adjusted weight to each predictor, it achieves better performance than an equal-weighted forecasts combination approach (Rapach, Strauss, and Zhou, 2010) in OOS forecasting.

Economically, why does the aligned macro risk index predict stock returns? We present plentiful evidence pointing out that the index is closely tied to the underlying economic conditions and well characterizes the macro risk varying over business cycles. By plotting the PLS estimated factor over the OOS period, we observe a prominent counter-cyclicality of the macro risk index. The subsequent subsample analysis reveals that its predictive power concentrates around economic recessions which is confirmed by OOS R^2 and utility gains. Therefore, the main finding that the aligned macro risk index positively predicts market returns can be consistent with a class of equilibrium asset pricing models. It has been shown both theoretically and empirically that equity premia tend to increase in bad economic times and fall in good times (Campbell and Cochrane, 1999; Gabaix, 2012). In particular, Campbell and Cochrane (1999) state that business cycle variation induces counter-cyclical investor's preference, thereby giving rise to counter-cyclical expected return. Since the aligned macro risk index also exhibits a counter-cyclical pattern, it likely captures the heightened risk and investor's preference during economic recessions.

If our index does reflects cyclical macroeconomic risk, we would also expect it to predict the returns on stock portfolios that are sensitive to business-cycle fluctuations. Indeed, this is true. The predictive ability of the aligned macro risk index is more significant for returns on small size firms and cyclical industry portfolios (such as durable goods) than for large size firms and defensive industries (such as healthcare equipment and utilities). Additionally, a decomposition of market return illustrates that the source of macro risk index's predictive power mainly stems from a discount rate channel. This is in accord with our expectation since the aligned macro risk index better anticipates the changes in discount rate due to the variation of macroeconomic risk than conventional predictors do.

We are aware of the existence of other macro-based return predictors such as the output gap, a production-based macroeconomic variable, of Cooper and Priestley (2009) and the cyclical consumption of Atanasov, Møller, and Priestley (2020) that is built on the consumption expenditures on nondurables and services. Empirically, we find that the predictive power of aligned macro risk index remains significant after controlling for these macro variables and even subsumes their predictability. As emphasized by Cochrane (2017), exploiting multiple state variables that drive the marginal utility, and hence, the expected returns is of necessity to better account for time variation in the equity premium. Particularly, Piazzesi, Schneider, and Tuzel (2007) examine the impact of housing services on asset prices in a general equilibrium model. The findings of Chen and Zhang (2011) suggest empirical link between the stock market and the labor market. Intuitively, in economic bad times, when investors face increasing risk about losing their jobs and houses, expected stock returns need to be high in order to induce investors to invest. Since our macro risk index also incorporates information about the housing and labor markets, this explains why it better captures business-cycle-related risk premiums than those standalone macro indicators that only reflect a single aspect of the macroeconomy do.

While we underscore the connection between our results and the rational-based equilibrium models, the finding of this paper lends support to the notion that beliefs play an important role in asset pricing models (Brunnermeier, Farhi, Koijen, Krishnamurthy, Ludvigson, Lustig, Nagel, and Piazzesi, 2021). More specifically, we link the beliefs about future macroeconomic conditions, including GDP, unemployment, and housing to the variation of expected market returns. Nevertheless, one may cast a doubt on the predictability of the aligned macro risk index that whether it arises from the sentiment or belief heterogeneity reflected by the SPF survey expectations. To address this concern, we consider additional control variables such as the aligned sentiment index of Huang, Jiang, Tu, and Zhou (2015) and disagreement index of Huang, Li, and Wang (2021), and demonstrate that the channel of our predictive power is distinct from theirs.

Consistent with what economic theory predicts, the paper empirically establishes the link between the stock market prices and business cycle fluctuations. More importantly, we stress that the variation in the equity premium is consistent with the rational response of investors to changing aggregate macroeconomic risks. The remainder of this paper proceeds as follows. Section 2 introduces the construction of the aligned macro risk index. Section 3 describes the SPF forecasts data. Section 4 reports the in- and out-of-sample forecasting results for market excess return. Section 5 assesses the economic significance of return predictability. Section 6 explores the source of the predictive power of the aligned macro risk index and Section 7 concludes the article.

2 Methodology

In this section, we introduce the two econometric methods for constructing the index that measures aggregate macro risk.

2.1 Predicting equity premium using macro variables

We assume a linear relation between the expected one-step ahead market excess return and the macroeconomy risk such that

$$\mathbb{E}_t(R_{t+1}) = \alpha + \beta M_t, \tag{1}$$

where M_t measures the macroeconomic condition at time t. Since the realized market return is equal to its conditional expectation plus a shock, we can write the market excess return at time $t+1$ as

$$R_{t+1} = \alpha + \beta M_t + \epsilon_{t+1}, \tag{2}$$

where ϵ_{t+1} is a shock that is unrelated to M_t.

Though the true factor M_t is not directly observable, the investor may use various economic fundamentals to forecast the market since each of them subsumes a fraction of the information about M_t. In fact, the economic condition is highly complex and has multi-dimensional features such as the GDP and consumption, and we hardly can characterize the overall economy using a single indicator. Therefore, we rely on seven macroeconomic fundamentals, including the real GDP, industrial production, recession probability, unemployment, corporate profit, housing starts, and inflation, to summarize the economic condition. Because the macroeconomic data are usually unavailable to investors in real time due to publication lags, we use the real-time survey forecasts on the above-mentioned seven macro variables from the SPF as macroeconomic condition proxies to gather information about M_t.

Let $s_t = (s_{1,t}, s_{2,t}, ..., s_{N,t})'$ denote the vector of survey forecasts on N ($N = 7$) macro variables. We assume a linear factor model for $s_{i,t}$ that follows

$$s_{i,t} = \delta_{i,0} + \delta_{i,1} M_t + \delta_{i,2} E_t + \eta_{i,t}, \quad i = 1, ..., N \tag{3}$$

where M_t is the latent factor that measures macroeconomic condition, E_t denotes the measurement error that is common to all survey forecasts and irrelevant to market returns, and $\eta_{i,t}$ is the shock to variable i exclusively. Since all individual survey forecasts carry information about M_t, a naive way to predict market return is to use them together by running the following regression,

$$R_{t+1} = \gamma_0 + \sum_{i=1}^{N} \gamma_i s_{i,t} + e_{t+1}, \tag{4}$$

which is also known as the kitchen sink model. By doing this, however, we are unable to identify the true factor M_t, and more importantly, such multivariate setting unavoidably inherits the noise from each predictor, thereby generating forecasts with numerous false signals. Besides, the kitchen sink model usually suffers from in-sample overfitting and unreliable out-of-sample performance (Welch and Goyal, 2008). These shortcomings indicate the need for a more parsimonious predictive mode that helps to recover the latent factor M_t from a pool of macroeconomic condition proxies while filtering out their common error component E_t and the idiosyncratic noise $\eta_{i,t}$.

2.2 Two econometric methods

We now briefly introduce the two econometric methods considered to recover the latent factor M_t.

Principal component regression

The first method is the principal component regression (PCR) which is a commonly used dimension reduction approach in predicting stock and bond returns (Eriksen, 2017; Ludvigson and Ng, 2007, 2009; Neely, Rapach, Tu, and Zhou, 2014). PCR achieves dimension reduction by consolidating a set of predictors into a few principal components (PCs). Each PC, being orthogonal to others, is a linear combination of the original predictors where the weights are designed to capture the maximum fraction of total variations in predictors. In our study, we rely on the first PC which explains the most covariance among $\{s_{i,t}\}$ to predict market returns. Though PCR substantially reduces the dimension of predictors, it is apparent that the PC contains variations of both M_t, the return-relevant part, and E_t, the return-irrelevant part. Namely, when E_t accounts for the major variation of $\{s_{i,t}\}$, the information content of the first PC will be mostly dominated by E_t. In such case, the PC approach fails to extract an effective proxy for M_t and is unable to provide significant return forecast. Regarding on this issue, we further consider the partial least squares (PLS) (Kelly and Pruitt, 2015; Wold, 1966)[2] to extract the macro risk factor, M_t, from the cross-section and eliminate the irrelevant component E_t at the same time.

Partial least squares regression

PLS is a target-driven approach that it condenses the cross-section according to the covariance between predictors and target. By its econometric design, PLS is able to tease out M_t from the common variations in $\{s_{i,t}\}$ and remove E_t. Accordingly, the extracted PLS factor is a more efficient use of the information about M_t subsumed by individual predictors than PC does.

Specifically, the PLS factor can be estimated through a two-pass OLS regression. In the first stage, we run a time-series regression for each survey variable $s_{i,t}$ on the future market

[2] Kelly and Pruitt (2015) extend the PLS method initially proposed by Wold (1966) into a more general setting and derive associated asymptotic properties. As Kelly and Pruitt's method can be implemented via three consecutive OLS regressions, it is also named as "the three-pass regression filter".

excess return R_{t+1} individually,

$$s_{i,t} = \delta_{i,0} + \delta_i R_{t+1} + \xi_{i,t}, \quad \text{for } i = 1, ..., N \tag{5}$$

where R_{t+1} is time $t+1$ market excess return and δ_i is the factor loading that describes the sensitivity of $s_{i,t}$ to movements in R_{t+1}. Since we use R_{t+1} as the instrumental variable and it is mainly driven by M_t by assumptions, δ_i approximately gauges to what degree does the macroeconomic condition proxy $s_{i,t}$ depend on the true macroeconomic condition M_t. Namely, the larger the absolute value of δ_i is, the more is $s_{i,t}$ exposed to the movements of M_t and thereby, the more important the role it will play in estimating the latent factor.

In the second stage, we treat the factor loadings estimated in the first stage as independent variables and run a cross-sectional regression of s_t on the union of them, $\hat{\delta} = (\hat{\delta}_1, ..., \hat{\delta}_N)'$, for each period t,

$$s_t = \theta_t + \hat{\delta}\, \mathrm{MR}_t^{\mathrm{PLS}} + v_t, \quad \text{for } t = 1, ..., T. \tag{6}$$

By estimating the slope coefficient of the above regression model, we obtain the *the aligned macro risk index* based on the seven SPF survey variables, denoted as $\mathrm{MR}^{\mathrm{PLS}}$. The idea here is to use the exposure of each proxy $s_{i,t}$ toward target to identify the return-relevant factor that is common to all proxies and remove components including E_t and $\eta_{i,t}$ that are unrelated to return forecasting.

Given the full-sample data spanning from time 1 through time T, the $T \times 1$ vector of $\mathrm{MR}^{\mathrm{PLS}} = (\mathrm{MR}_1^{\mathrm{PLS}}, ..., \mathrm{MR}_T^{\mathrm{PLS}})'$ can be calculated in one step using the following closed form expression

$$\mathrm{MR}^{\mathrm{PLS}} = S\, \underbrace{\mathbf{J_N S' J_T R (R' J_T S J_N S' J_T R)^{-1} R' J_T R}}_{W\text{: weight}}, \tag{7}$$

where $R = (R_2, ..., R_{T+1})'$ denotes the vector of the market excess returns, $S = (s_1, ..., s_T)'$ denotes the $T \times N$ matrix of the lagged proxies, $J_H \equiv I_H - \frac{1}{H}\iota_H \iota_H'$, I_H is an identity matrix with dimension H, and ι_H is a H $\times$ 1 vector of ones ($H \in \{N, T\}$). That is, the estimated latent factor $\mathrm{MR}^{\mathrm{PLS}}$ can be expressed as a linear combination of the proxies S whose weights, W, are determined by their covariance with future market excess returns. We use this analytical expression latter to explore the contribution of each proxy to the aligned macro risk index.

3 Data and summary statistics

3.1 Data from the Survey of Professional Forecasters

To construct the aligned macro risk index, we use the survey forecasts data from the SPF, which is one of the oldest macroeconomic survey in United States.[3] Specifically, we collect professional forecasts on seven macro variables that have the longest history since the initiation (1968Q4) of the SPF:

- growth rate forecast for the chain-weighted real GDP (GDP_e),
- growth rate forecast for the industrial production index (Indprod_e),
- forecast for the probability of the chain-weighted real GDP level fall below the level of preceding quarter (Recess_e),
- forecast for the civilian unemployment rate (Unemp_e),
- growth rate forecast for the quarterly nominal corporate profits after tax (Cprof_e),
- growth rate forecast for the housing starts (Housing_e),
- growth rate forecast for the chain-weighted GDP price index (Inflation_e).

In each quarter, the survey forecasts include a "nowcast" for the current quarter as well as a term structure of forecasts over horizons ranging from one to four quarters ahead. In our analysis, we first consider a collection of nowcasts on the seven macro variables, denoted as **SPF7**, to forecast quarterly market excess return. We then add the term structure of forecasts up to three quarters ahead to obtain a larger variable set, which we refer to as **SPF7TS**, in the study of long-term return predictability.[4] The data span from 1968Q4 to 2019Q4 with a total of 205 quarters.

The forecasts for the GDP, the industrial production index, corporate profits after tax, housing starts, and the GDP price index are annualized *Quarter-over-Quarter* growth rate forecasts computed as

$$g_{i,t,j} = 100 \times \left[\left(\frac{M_{i,t,t+j}}{M_{i,t,t+j-1}} \right)^4 - 1 \right],$$

where $j = 0, 1, ..., 3$, $i = \{\text{GDP, Indprod, Cprof, Housing, Infl}\}$, and $M_{i,t,t+j}$ denotes quarter t consensus forecast, which is defined as the average of forecasts made by individual forecasters,

[3]We obtain the historical data from the Federal Reserve Bank of Philadelphia, `https://www.philadelphiafed.org/surveys-and-data`. According to the SPF, the forecasters are able to access the Bureau of Economic Analysis's advance estimate about historical quarters when they receive the survey's questionnaires. The result is typically released at the end of the mid month of each quarter. Nonetheless, due to a few exceptions with delayed releases, we carefully treat all surveys as available only at the end of each quarter, as in Huang, Li, and Wang (2021).

[4]We remove the four-quarter ahead survey forecasts for their missing observations in early years.

on the level of variable i at quarter $t+j$.[5] For the forecasts on the unemployment rate and the recession probability, we directly use the level data.

Panel A in Table I reports the descriptive statistics of the seven macro variables forecasts, including the sample mean, standard deviation, minimum, maximum, skewness, kurtosis, median, and the estimated first-order autoregressive (AR(1)) coefficient. First, the average nowcast of the GDP growth is around 2.32%, roughly consistent with the real GDP growth 2.79% (not tabulated) during the sample period of 1968 through 2019. Second, compared to the current-quarter forecasts on the other five macro variables, Recess_e and Housing_e exhibit substantial volatility. Notably, the nowcasts on housing starts display the smallest autocorrealtion (0.43), whereas Unemp_e and Infl_e are quite persistent, both of which have an AR(1) coefficient of 0.96. Third, as shown in the eighth column, the relatively strong correlations of Unemp_e and Housing_e with the future market risk premium (0.16 and 0.19, respectively) imply that they appear likely to contain useful information in forecasting market return. Turning to long-horizon forecasts, we note that the term structures of GDP_e, Indprod_e, Cprof_e, and Housing_e are upward sloping, while that of Recess_e is slightly downward sloping. The average values of Unemp_e and Infl_e slightly change indicating that their term structures are relatively flat. Besides, all survey forecasts become less volatile as the forecast horizon increases, confirmed by their decreasing standard deviations.

[**Insert Table I and Figure 1 here**]

To provide additional perspective on the dynamics across different survey variables, we plot the term structures of the seven macro variables forecasts. As shown by Figure 1, we find that the growth forecasts on the real GDP, industrial production, and corporate profits are pro-cyclical, whereas the forecasts on the recession probabilities and unemployment rates are counter-cyclical. The latter two typically fall to the local minimum at the business cycle peak and spike up at the trough. The long swings in the inflation forecasts persist even beyond the measured business cycles. Next, except for Unemp_e and Infl_e, the other variable forecasts are more volatile and involve considerable temporal fluctuations. On the other hand, the forecasts on unemployment rates are smooth and slowly evolve across NBER business-cycle phases. In particular, after the Oil Shock in 1970s, there is an apparent downward trend in the expected inflation during the recent three decades. Finally, though the 2008 global financial crisis is the most prolonged recession among the 7 recessions in our sample, the growth forecasts for the real GDP, industrial production, housing starts, and corporate profits appear to be the lowest in the mid 1970s (Oil Shock recession) and early 1980 recessions.

[5] We consider alternative ways to construct the growth forecasts, including the approach of Eriksen (2017) who uses the median of individual growth forecasts. We find that the predictive performance is indeed the strongest when the growth forecast is calculated in a *Quarter-over-Quarter* manner using consensus forecast. The empirical results for the alternative constructing approaches are available upon request.

3.2 Market return data and other predictors

We collect the monthly return on the CRSP value-weighted index and the one-month T-bill rate.[6] The monthly CRSP returns and the T-bill rates are then compounded to quarterly market returns and risk-free rates, respectively, to match the frequency of the SPF survey data. We subtract the risk-free rate from the market return to measure the realized market risk premium. According to Panel B in Table I, the average quarterly market risk premium is 1.65% associated with standard deviation of 8.64 in our sample, producing a Sharpe ratio of 0.19 (not tabulated). The distribution of the market excess returns is negatively skewed and leptokurtic in line with the literature. Compared to the survey forecasts, the quarterly market return is much less persistent displaying an autocorrelation closed to zero (0.05).

In addition to survey forecasts on the seven macro variables, we consider 16 economic and financial predictors studied by Welch and Goyal (2008), including the logarithm of the dividendCprice ratio (DP), the logarithm of the dividend yield (DY), the logarithm of the earningsCprice ratio (EP), the logarithm of the dividendCpayout ratio (DE), stock variance (SVAR), the book-to-market ratio (BM), net equity expansion (NTIS), the three-month Treasury bill rate (TBL), the long-term government bond yield (LTY), the return on long-term government bonds (LTR), the term spread (TMS), the default yield spread (DFY), the default return spread (DFR), the growth rate of Consumer Price Index (INFL), the consumption to wealth ratio (CAY), and the investment to capital ratio (IK). We obtain the data of these predictors from Amit Goyals website and more detailed variable definitions can be found in the Online Appendix.[7]

4 Equity premium prediction

In this section, we present the empirical results of predicting market excess return using the aligned macro risk index based on the SPF data. We begin with an in-sample analysis and then move to OOS tests. We also assess the return predictability over longer forecast horizons and over different subsample periods. Finally, we provide results on international predictability.

[6]We obtain the data from Kenneth French's webpage `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`. We thank Kenneth French for making the data available.

[7]We thank Amit Goyal for making the data available in his webpage `http://www.hec.unil.ch/agoyal`.

4.1 In-sample return predictability

We estimate the following one-step ahead predictive regression model using the whole sample data from 1969Q1 to 2019Q4,

$$R_{t+1} = a + \beta X_t + \epsilon_{t+1} \tag{8}$$

where R_{t+1} is the quarterly market excess return (annualized).[8] The predictor X_t in Eq. (8) can be one of the variables from SPF7, the first PC of SPF7, or the aligned macro risk index. We try to answer the question: does the predictive power of the SPF forecasts arise from their information content, or from the econometric method, or both? For comparison, we also consider the 16 economic predictors from Welch and Goyal (2008) that we refer to as Econ.

[**Insert Table II here**]

Panel A of Table II provides the in-sample estimation results for the univariate regression model, including the OLS estimates, Newey-West t-statistics, and in-sample R^2 statistics. To ease the interpretation of the regression slope, each predictor is standardized to have zero mean and unit variance. By looking at the top block of Panel A, we note that only two individual survey forecasts display significant predictability for the quarterly market excess return. First, Unemp_e positively predicts the future market return, with an R^2 of 2.53%. Its regression slope is statistical significant at the 5% level according to the Newey-West t-statistic. Chen and Zhang (2011) theoretically show that the lagged aggregate payroll growth should negatively predict the market excess return under the time-to-build production technology. Our finding is supportive for their predictions to the extent that the unemployment rate is negatively correlated with the payroll growth. Second, Housing_e displays the most significant predictability for the market at the 1% level among the seven forecasts, along with an R^2 over 3%. As illustrated by Figure 1, Housing_e often spikes up when the business drops to a trough, coinciding with the latter rebound of the stock market. This indicates the forward-looking feature of the survey forecasts and underlines a cross-market linkage between the investment in the real estate and the equity markets. While the rest of the survey variables are not significant, their regression slope signs are consistent with the theoretical expectations.[9]

Looking at the middle block, we note that MR^{PLS} is a positive predictor for the market risk premium, whereas SPF7^{PC} behaves just the opposite. This is possibly due to the

[8] We also examine the predictability for log market excess returns whose results are displayed in Table OA.4 of the Online Appendix.

[9] The negative coefficient of Infl_e is in line with (Fama and Schwert, 1977) which document that the expected stock return is negative correlated with the expected inflation (Fama and Schwert, 1977). The negative coefficients of GDP_e, Indprod_e, and Cprof_e corroborate the evidence that the expected business condition is negative related to the equity premium (Campbell and Diebold, 2009).

differences between the weights assigned by PCR and the weights assigned by PLS to each individual survey forecast. Indeed, we latter show that SPF7^{PC} loads more on those negative predictors such as GDP_e and Indprod_e, while MR^{PLS} loads more on Unemp_e and Housing_e. SPF7^{PC}, however, fails to produce significant forecasts for the market return, with an R^2 merely of 0.04% that is inferior to all individual survey forecasts. In contrast, MR^{PLS} generates a sizable R^2 of 5.75% and has a slope coefficient of 0.083 that is significant at the 1% level. Such remarkable difference between the PCR and the PLS approaches reveals that a large fraction of the common variations in survey forecasts is unrelated to the equity premium. We also evaluate a multivariate regression model that employs all individual survey forecasts (SPF7^{KS}). As shown by the eighth row, this kitchen sink model achieves an R^2 of 6.85% that is the highest value of all R^2s reported in Panel A. To some extent, this sets a ceiling on the in-sample predictability of the seven survey variables, while the PLS macro risk index shows a comparable predictive ability to that.

The last 18 rows of Panel A present the estimation results for the 16 predictors from Econ and the two dimension reduction approaches based on Econ. Among the 16 economic predictors, only four (LTR, TMS, CAY, and IK) generate statistically significant predictive ability for the quaterly market return at the 10% level or better. In particular, barely no individual predictor produces R^2 higher than 2% except for IK (2.91%), and DP is insignificant with an R^2 less than 1%. These findings are consistent with Welch and Goyal (2008) who show that a series of standard predictors fail to predict the equity premium in-sample after the Oil Shock in mid 1970s. Nonetheless, Econ^{PLS} positively predicts the market with a regression slope of 0.085 that is significant at the 1% level, and its corresponding in-sample R^2 exceeds 6% (6.04%) that is higher than that of MR^{PLS}. However, when we apply the look-ahead bias-free approach to construct the PLS factor as in Huang, Jiang, Tu, and Zhou (2015), we find that the sign of the regression slope of $\text{Econ}^{\text{PLS}}_{\text{Bias-free}}$ reverses, being negative, albeit significant. In contrast, the relation between $\text{MR}^{\text{PLS}}_{\text{Bias-free}}$ and market return remains stable and strong, with a slope of 0.070 that is significant at the 1% level. The results for PLS forecast without look-ahead bias can be found in the Online Appendix.

Controlling for economic and survey variables

In the Online Appendix, we show that MR^{PLS} are closely related to the predictor variables that track business-cycle fluctuations such as TMS and IK, and are weakly related to market-based valuation ratios such as DP and EP. To investigate whether the aligned macro risk index contributes incremental information to extant predictors, we run the following bivariate regression,

$$R_{t+1} = \alpha + \beta\, \text{MR}_t^{\text{PLS}} + \psi Ctrl_t + \epsilon_{t+1}, \tag{9}$$

where $Ctrl$ is one of the variables listed in the first column of Table II other than MR^{PLS}.

Panel B of Table II presents the in-sample estimation results for the above bivariate regression model. According to the first seven rows of Panel B in which the results for individual survey variables are displayed, the regression slopes of $\mathrm{MR}^{\mathrm{PLS}}$ are all significant at the 1% level except for the case including $\mathrm{Housing}_e$. Note that adding the individual survey variable merely increases the R^2 for a small amount relative to the univariate model that employs $\mathrm{MR}^{\mathrm{PLS}}$ alone. This reflects that the PLS factor efficiently collects the forecasting information subsumed by individual predictors. Moreover, the last 18 rows of Panel B show that neither the magnitude nor the statistical significance of the regression slope of $\mathrm{MR}^{\mathrm{PLS}}$ changes much after controlling for the economic variables. The coefficients associated with $\mathrm{MR}^{\mathrm{PLS}}$ are all significant at the 1% level with the only exception being the case including $\mathrm{Econ}^{\mathrm{PLS}}$. In that case, the regression slope of $\mathrm{MR}^{\mathrm{PLS}}$ drops to 0.052 while significant at the 10% level, whereas $\mathrm{Econ}^{\mathrm{PLS}}$ becomes statistically insignificant. Therefore, the bivariate regression test reveals that $\mathrm{MR}^{\mathrm{PLS}}$ provides incremental information to the extant economic predictors in forecasting the market return.

In short, the in-sample analysis confirms that $\mathrm{MR}^{\mathrm{PLS}}$ significantly predicts quarterly market excess returns over the sample period from 1969 to 2019. On the one hand, each survey forecast variable, Unemp_e and $\mathrm{Housing}_e$ in particular, contains pieces of useful information about the market since $\mathrm{SPF7}^{\mathrm{KS}}$ yields the second largest value of all in-sample R^2s. On the other hand, compared with $\mathrm{SPF7}^{\mathrm{PC}}$, the superior performance of $\mathrm{MR}^{\mathrm{PLS}}$ could be partially attributed to the great efficiency of PLS in condensing multivariate information.

4.2 Out-of-sample return predictability

The previous in-sample analysis demonstrates that the proposed aligned macro risk index, $\mathrm{MR}^{\mathrm{PLS}}$, displays substantial ability in explaining the market risk premium variation given the full-sample data. Nonetheless, Welch and Goyal (2008), among others, underline the necessity of OOS tests in examining the stability and reliability of a predictor's predictive performance. Since OOS forecasts avoid the look-ahead bias and over-fitting problems, the associated predictability is more practical relative to in-sample test, thereby more relevant to real-time investors. Thus, in this section, we investigate the OOS forecasting performance of survey forecast variables.

Out-of-sample forecasting and evaluation criterion

To generate OOS forecasts, we estimate the regression coefficients of Eq. (8) recursively using only the information available at the time when forecasts are made. Specifically, we split a full-sample data of T observations into two sets, namely, the initial training set and the OOS evaluation set. Let q be the length of the initial estimation window. For the first OOS forecast, we run the predictive regression (8) of market excess returns $\{R_2, ..., R_q\}$ on the lagged predictor series $\{X_1, ..., X_{q-1}\}$ to obtain OLS estimates $\hat{\alpha}$ and $\hat{\beta}$. So, the time $q+1$ market excess return forecast made at time q is computed as $\hat{\alpha} + \hat{\beta} X_q$. Moving to the next period, we expand the estimation window to include new information at time $q+1$ while fixing the starting date. We repeat the preceding process to compute the time $q+2$ market forecast, and the entire procedure is carried forward till the end of the sample with a total of $T-q$ OOS forecasts. Note that the procedure to construct the OOS macro risk index via PLS shares a similar spirit here that we evaluate Eqs. (5) and (6) only using available information at the time.

Following the literature, we assess how well a predictive model does in forecasting the market OOS based on several statistical criteria. We take the historical mean of the market excess returns as the benchmark forecast since if the market is unpredictable, simply using the average of historical returns will outperform using the forecasts generated by sophisticated predictive models. Let R_{t+1}, $\bar{R}_{t+1}$, and $\hat{R}_{i,t+1}$ denote the realized market excess return, the historical mean benchmark forecast for time $t+1$ market return, and the time $t+1$ market return forecast made by predictive model i, respectively. The first performance metric is the OOS R^2 of Campbell and Thompson (2008), which is defined as one minus the ratio of the mean squared forecast error (MSFE) of predictive model i over the benchmark model's MSFE,

$$R^2_{OS} = 1 - \frac{MSFE_{Model\ i}}{MSFE_{Bench}} = 1 - \frac{\sum_{t=q+1}^{T} (R_t - \hat{R}_{i,t})^2}{\sum_{t=q+1}^{T} (R_t - \bar{R}_t)^2}, \tag{10}$$

where $\bar{R}_t = \frac{1}{t} \sum_{s=1}^{t} R_s$. Evidently, a lower $MSFE_{Model\ i}$ relative to $MSFE_{Bench}$ leads to a positive R^2_{OS}, implying that the predictive model i outperforms the historical mean benchmark in terms of OOS predictive accuracy. Beside, we rely on the MSFE-adjusted statistic of Clark and West (2007) (CW) to test the null hypothesis $R^2_{OS} \leq 0$ against the one-sided alternative hypothesis $R^2_{OS} > 0$. This is equivalent to test the population MSFE of the benchmark is less than or equal to that of the predictive model, against the alternative that the predictive model has a lower population MSFE.

The second performance measure is the difference in cumulative squared forecast error (DCSFE), which is the difference between the CSFE for the historical mean benchmark and the CSFE for a predictive model i,

$$\text{DCSFE}_{i,t+1} = \text{CSFE}_{Bench,t+1} - \text{CSFE}_{Model\ i,t+1}, \tag{11}$$

where

$$\text{CSFE}_{Bench,t+1} = \sum_{s=q+1}^{t+1} (R_s - \bar{R}_s)^2 \quad \text{and} \quad \text{CSFE}_{Model\ i,t+1} = \sum_{s=q+1}^{t+1} (R_s - \hat{R}_{i,s})^2.$$

As argued by Welch and Goyal (2008), a time series plot of $\{\text{DCSFE}_{i,t}\}$ can serve as a visual tool to diagnose the accuracy and stability of a predictive model relative to the no-predictability benchmark. That is, the time series of DCSFE associated with a robust predictive model should stay positive and exhibit an upward trend most of the time.

The third performance measure considered is the forecast encompassing test that helps to rank two competing predictive models according to their information contents (Granger and Newbold, 1973). To this end, we construct an optimal composite forecast using a convex combination of the forecasts generated by model i and model j,

$$\tilde{R}_{t+1} = (1-\lambda)\hat{R}_{i,t+1} + \lambda\hat{R}_{j,t+1}, \qquad 0 \le \lambda \le 1, \tag{12}$$

where $\tilde{R}_{t+1}$ denotes the optimal return forecast. A positive λ indicates that model j provides incremental forecasting information about the market excess return to model i. In contrast, a trivial λ implies that model j fails to contribute any additional information to model i in forming the optimal forecast, thereby being "encompassed" by model i. We gauge the significance of λ based on the Harvey, Leybourne, and Newbold (1998) (HLN) statistic, which tests the null hypothesis $\lambda = 0$ against the one-sided alternative $\lambda > 0$.[10]

Out-of-sample forecasting performance

[**Insert Table III here**]

Table III presents the OOS quarterly market return predictability measured by R^2_{OS} statistic and the estimated weights on the return forecasts generated by $\mathrm{MR}^{\mathrm{PLS}}$ in encompassing tests. We use the first 60 quarters as the initial training period, and the OOS period starts from 1984Q1 to 2019Q4 containing 144 observations in total.[11] The first seven rows in Panel A report the results for the predictive regression models based on individual SPF variables. Among them, $\mathrm{Housing}_e$ outperforms the historical mean benchmark in terms of MSFE at the quarterly horizon, with a positive R^2_{OS} of 1.64% that is significant at the 5% level according to the MSFE-adjusted statistic. Nevertheless, except for $\mathrm{Indprod}_e$ producing a positive but insignificant R^2_{OS} of 0.33%, the rest five SPF variables fail to generate more accurate forecasts for the market than the historical mean benchmark does. In contrast to the strong in-sample performance, the kitchen sink model performs poorly with a negative R^2_{OS} of -0.49% in predicting the market OOS. Rapach, Strauss, and Zhou (2010) point out that the large discrepancy between the in-sample and OOS performance arises from the complexity and fickleness of the data-generating process of equity premium. An unrestricted

[10] We also use the modified Diebold and Mariano (1995) statistic by Harvey, Leybourne, and Newbold (1997) to assess the significance of λ. The results are similar and are available upon request.

[11] According to Welch and Goyal (2008), our choice of the initial estimation window is to make a balance between adequate start-up observations to estimate parameters reliably and a sufficiently long OOS period for evaluating the predictive performance.

multiple regression inherits numerous false signals from its predictors and hence suffers from structural instability.[12]

As shown by the last three rows in Panel A, $\mathrm{MR}^{\mathrm{PLS}}$ continues to deliver superior forecasting performance relative to $\mathrm{SPF7}^{\mathrm{PC}}$. The R^2_{OS} statistic of 3.12% produced by $\mathrm{MR}^{\mathrm{PLS}}$ is both statistically significant (at the 5% level) and economically meaningful (Campbell and Thompson, 2008). Given the quarterly market Sharpe ratio being 0.19, a mean-variance investor could increase her portfolio return by a proportional factor of 0.86, nearly doubling her payoff, if she switch from using the historical mean forecast to the forecast made by $\mathrm{MR}^{\mathrm{PLS}}$. Though the mean combination forecast based on the seven SPF variables ($\mathrm{SPF7}^{\mathrm{FC}}$) also surpasses the benchmark forecast, its R^2_{OS} statistic (0.86%) is much lower than that of $\mathrm{MR}^{\mathrm{PLS}}$.

[**Insert Figure 2 here**]

To further investigate the consistency of the OOS forecasting performance over time, we follow Welch and Goyal (2008) to compute the DCSFE that is the difference between the CSFE for a predictive model and the CSFE for the benchmark model, as in Eq. (11). The top panel of Figure 2 depicts the time-series plots of DCSFEs for the univariate models based on the seven SPF variables. Only $\mathrm{Housing}_e$ exhibits an upward swing most of the time, especially during the 2008 global financial crisis. In contrast, all the remaining SPF variables, such as Cprof_e and Recess_e, fail to consistently outperform the historical mean benchmark. It is worth noting that the predictive accuracy of Unemp_e visibly deteriorated over the period of 1995 to 2000 and while turning to be remarkably well after the recent millennium, signaling a model uncertainty issue here.

The DCSFEs for the three dimension reduction methods are shown in the bottom panel of Figure 2. Both $\mathrm{SPF7}^{\mathrm{FC}}$ and $\mathrm{MR}^{\mathrm{PLS}}$ end up with lower CSFE relative to the historical mean benchmark, and the slopes of their DCSFE series are mostly positive over the whole sample period, illustrating that the good OOS predictive performance is not confined to some special periods. Intuitively, by integrating the forecasting information contained in a set of SPF variables, $\mathrm{SPF7}^{\mathrm{FC}}$ and $\mathrm{MR}^{\mathrm{PLS}}$ greatly mitigate the impact of the false signals in each individual predictor, thereby leading to more stable performance. Nevertheless, the curve of $\mathrm{MR}^{\mathrm{PLS}}$ is predominantly higher than that of $\mathrm{SPF7}^{\mathrm{FC}}$. Compared to the forecasts combination method, PLS is more efficient in condensing a large amount of information by extracting a common factor from a pool of predictors that conveys the most relevant information in forecasting the target. Accordingly, $\mathrm{MR}^{\mathrm{PLS}}$ eliminates the return-irrelevant components of each SPF variable and substantially reduces the noises that are unavoidably existed in the forecasts made by $\mathrm{SPF7}^{\mathrm{FC}}$. Interestingly, the predictability of $\mathrm{MR}^{\mathrm{PLS}}$ appears

[12]Another explanation is that from a statistical perspective, an unrestricted multiple regression incorporating a pool of correlated predictors would generate highly volatile forecasts, resulting in a large MSFE.

to be strong during the recession periods, such as the 2008 global financial crisis, and displays a counter-cyclical pattern consistent with the findings of Rapach, Strauss, and Zhou (2010) and Henkel, Martin, and Nardari (2011). We will conduct a thorough analysis on this phenomenon latter in Section 6.

The results from the above two performance measures, R^2_{OS} and DCSFE, highlight the outstanding OOS performance of $\mathrm{MR}^{\mathrm{PLS}}$ in forecasting the quarterly market risk premium, while we are also interested in the information content in $\mathrm{MR}^{\mathrm{PLS}}$ relative to the other forecasting models based on the seven SPF variables. Following Rapach, Strauss, and Zhou (2010), we conduct a forecast encompassing test. As indicated by the third column of Table III, all the other models (with the exception of $\mathrm{Housing}_e$) fail to encompass $\mathrm{MR}^{\mathrm{PLS}}$. Though we cannot reject the null hypothesis that $\mathrm{Housing}_e$ encompasses $\mathrm{MR}^{\mathrm{PLS}}$ at the conventional significant level, the fairly large encompassing coefficient λ (0.88) suggests that the forecast generated by $\mathrm{MR}^{\mathrm{PLS}}$ dominates that by $\mathrm{Housing}_e$ in constructing the optimal return forecast. In addition, the closed-to-one encompassing coefficients in the cases of $\mathrm{SPF7}^{\mathrm{PC}}$ (0.92) and $\mathrm{SPF7}^{\mathrm{FC}}$ (0.93) confirm that $\mathrm{MR}^{\mathrm{PLS}}$ has already incorporated most of the relevant forecasting information subsumed by SPF variables. The dominant role played in constructing the optimal return forecast also helps to explain the prominent forecasting performance of $\mathrm{MR}^{\mathrm{PLS}}$ as reported in the second column.

Panel B of Table III presents the OOS forecasting results of the predictive models based on economic variables. In line with the finding of Welch and Goyal (2008), most conventional predictors, such as DP and TMS, underperform the benchmark in terms of MSFE, resulting in negative R^2_{OS} statistics as reported in the fifth column. Among the 16 individual economic predictors, only the inflation rate delivers a positive R^2_{OS} of 1.13% that is significant at the 10% level. Neither the dimension reduction method ($\mathrm{Econ}^{\mathrm{PC}}$, $\mathrm{Econ}^{\mathrm{FC}}$, and $\mathrm{Econ}^{\mathrm{PLS}}$) outperforms the historical mean benchmark. Recall that though $\mathrm{Econ}^{\mathrm{PLS}}$ exhibits substantial in-sample predictability for the market with an R^2 over 6% (6.04%), such performance no longer exists in OOS test. Apparently, $\mathrm{MR}^{\mathrm{PLS}}$ outstrips all conventional predictors we considered in OOS return forecasting. The encompassing test results presented in the last column of Panel B further demonstrates that $\mathrm{MR}^{\mathrm{PLS}}$ provides additional forecasting information to the extant predictors. Notably, we can reject the null hypothesis that $\mathrm{MR}^{\mathrm{PLS}}$ is encompassed by another predictive model at the 5% level or better.[13]

To summarize this subsection, we find that $\mathrm{MR}^{\mathrm{PLS}}$ significantly outperforms the historical mean benchmark in OOS forecasting for the quarterly market return, whereas most individual SPF variables and conventional predictors we considered fail to do so. In addition to the previous in-sample evidence, the superior OOS performance of $\mathrm{MR}^{\mathrm{PLS}}$ relative to

[13]The only exception is that we are unable to reject $\mathrm{MR}^{\mathrm{PLS}}$ being encompassed by SVAR at conventional significant level. The extremely negative R^2_{OS} of -59.72% signals a great variation in the OLS forecasts based on SVAR, resulting in excessive standard errors of the weight estimates that render the encompassing test result insignificant.

SPF7$^{\mathrm{PC}}$, SPF7$^{\mathrm{FC}}$, and SPF7$^{\mathrm{KS}}$ reaffirms the efficiency of PLS in finding the relevant factor that drives forecast target.

4.3 Long-horizon prediction

In this subsection, we extend our market return predictability analysis from quarterly to longer forecast horizons. Given that the macroeconomic conditions are persistent,[14] the impact of business conditions on the expected market excess return is likely to be persist as well (Campbell and Diebold, 2009). Because most of the SPF variables display strong autocorrealtions as shown by Table I, we expect that the aligned macro risk index MR$^{\mathrm{PLS}}$ can track the long-run effect of macro risks on the stock market and thus, can predict the market at longer horizons. Moreover, since the duration of business cycle is time-varying (Diebold and Rudebusch, 1990; Filardo and Gordon, 1998), the information content of the term structure of SPF survey forecasts should have a certain advantage in predicting the long-term market return. It is thus of interest to construct a long-term macro risk index (LT-MR$^{\mathrm{PLS}}$) using the variable set SPF7TS, which includes the term structure of SPF forecasts from current quarter up to three quarters ahead, and investigate its forecasting power at longer horizons.[15]

To predict the market excess return over the future two to 12 quarters, we estimate the following multi-step predictive regression model,

$$R_{t+1:t+h} = a + \beta X_t + \epsilon_{t+1:t+h}, \quad h = 2, 4, 8, 12 \text{ quarter} \tag{13}$$

where $R_{t+1:t+h}$ is the market excess return from time $t+1$ to $t+h$ and the forecast horizons $h = 2, 4, 8, 12$ quarter correspond to the next half, one, two and three years. The in-sample results of the above predictive regression in Table IV suggest that both MR$^{\mathrm{PLS}}$ and LT-MR$^{\mathrm{PLS}}$ significantly predict the long-term market excess returns, with R^2s ranging from 7.24% to 27.09% over the four horizons. Their predictive power increases with the horizon, while the R^2 values of LT-MR$^{\mathrm{PLS}}$ are uniformly larger than the MR$^{\mathrm{PLS}}$ counterparts. To robustify our statistical inference about the long-term predictability, we consider several tests concerning the statistical significance of the regression slope estimate, including using a wild bootstrap procedure to determine the empirical p-values as in Huang, Jiang, Tu, and Zhou (2015), using the Hodrick (1992) corrected t-statistic, and applying the Kostakis, Magdalinos, and Stamatogiannis (2015) Wald statistic to test $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$. The results can

[14]For instance, Hamilton (1989) models the dynamics of macroeconomic regimes as a Markov regime-switching process and estimate the one-period transitional probability of staying at the expansion (contraction) regime to be 0.9 (0.75), respectively.

[15]For brevity, we focus on two SPF predictors, MR$^{\mathrm{PLS}}$ and LT-MR$^{\mathrm{PLS}}$, in this subsection. The empirical results of the PC approach, the forecast combination approach, and the kitchen sink model are indeed much weaker than the PLS approach, and are available upon request

be found in the Online Appendix.

[**Insert Table IV here**]

Consistent with the in-sample results, the OOS R^2 statistics in Panel A of Table V show that both $\mathrm{MR}^{\mathrm{PLS}}$ and LT-$\mathrm{MR}^{\mathrm{PLS}}$ outperform the historical mean benchmark at all horizons, with sizable R^2_{OS}s that are significant at least at the 10% level. Similarly, the OOS predictive ability of LT-$\mathrm{MR}^{\mathrm{PLS}}$ dominates that of $\mathrm{MR}^{\mathrm{PLS}}$ at all forecast horizons. The forecast encompassing results in Panel B further confirm that LT-$\mathrm{MR}^{\mathrm{PLS}}$ contains additional forecasting information for long-term market returns relative to $\mathrm{MR}^{\mathrm{PLS}}$. As we expected, the richer information content of the term structure of SPF forecasts about the long-run business conditions gives LT-$\mathrm{MR}^{\mathrm{PLS}}$ an advantage over $\mathrm{MR}^{\mathrm{PLS}}$ in forecasting long-horizon market returns.

[**Insert Table V here**]

To compare the forecasting performances of $\mathrm{MR}^{\mathrm{PLS}}$ and LT-$\mathrm{MR}^{\mathrm{PLS}}$ over time, we plot their DCSFEs at each forecast horizon in Figure 3. Two features follow the plots. First, LT-$\mathrm{MR}^{\mathrm{PLS}}$ (solid line) steadily outperforms $\mathrm{MR}^{\mathrm{PLS}}$ (dashed line) by generating smaller squared forecast error because the former's curve always stays above the latter's across different sample periods. In addition, the dashed line illustrates that the CSFE of $\mathrm{MR}^{\mathrm{PLS}}$ is inferior to that of the benchmark over the periods from the mid 1980s to 2000s, while the solid line is predominantly larger than zero. Second, when the horizon moves from two-quarter to 12-quarter, both lines become much smoother and show a progressively upward trend, reaffirming the more reliable and powerful forecasting powers of $\mathrm{MR}^{\mathrm{PLS}}$ and LT-$\mathrm{MR}^{\mathrm{PLS}}$ at longer horizons.

[**Insert Figure 3 here**]

Turning to the long-horizon forecasting performance of the economic variables whose results are presented in the last 19 rows of Table V, $\mathrm{Econ}^{\mathrm{PLS}}$ fails to outperform the historical mean benchmark in OOS test. Instead, a few individual predictors, including TMS and IK, exhibit significant OOS predictive power for the market. For instance, the R^2_{OS} values of TMS and IK are 14.63% and 16.28%, respectively, at the 12-quarter horizon. Nevertheless, the forecasting performance of LT-$\mathrm{MR}^{\mathrm{PLS}}$ dominates all models based on economic predictors at all horizons. The OOS encompassing test results in Panel B also confirm that LT-$\mathrm{MR}^{\mathrm{PLS}}$ contributes additional information to the extant predictors and undertakes a dominant weight (λ) in the optimal return forecast.

In summary, the results presented in Tables II–V and Figures 2–3 demonstrate that the PLS macro risk index can significantly predict the aggregate stock market returns at the quarterly and longer horizons up to three years ahead and both in- and out-of-sample. Such predictability remains robust after controlling for a host of extant predictors. Our results imply that forecasts of the future macroeconomic conditions help to predict future stock market returns. More importantly, the information content of the macro forecasts and the return predictability agree in horizon. Taken together, the previous finding complements Fama (1990) that the stock market returns predict future real economic activities and portrays a tight relation between the equity premia and the aggregate macro risk at business cycle frequencies.

4.4 Subsample analysis

Our baseline in-sample forecasting results are based on the sample period of 1969 through 2019 and the OOS forecasting starts in 1984. To examine the stability of predictive power of the aligned macro risk index, we additionally consider three subsamples: 1980 to 2019, 1990 to 2019, and 2000 to 2019. Table VI presents the forecasting results for the three subsamples, where we use $\mathrm{MR}^{\mathrm{PLS}}$ as return predictor at the quarterly horizon and use LT-$\mathrm{MR}^{\mathrm{PLS}}$ for longer horizons. First, we find that the results for the post-1980 and post-1990 sample periods are comparable to the full-sample results shown in Tables II and IV. Second, the magnitude of regression estimates and in-sample R^2 values for the post-2000 period are substantially raised and well above the full-sample counterparts. For instance, in the post-2000 period, the R^2s range from 14.40% to 58.04% over the five forecast horizons, whereas their full-sample estimates are usually about half of these values. Third, both $\mathrm{MR}^{\mathrm{PLS}}$ and LT-$\mathrm{MR}^{\mathrm{PLS}}$ significantly outperform the historical mean benchmark in OOS forecasting, no matter whether the forecasting starts in 1980, 1990, or 2000.

[**Insert Table VI here**]

In contrast to Welch and Goyal (2008) and Goyal, Welch, and Zafirov (2021), who argue that a majority of the traditional and even recently uncovered predictors evince weak in-sample significance and perform poorly OOS over the last few decades, the subsample analysis in this subsection shows that the relation between the aligned macro risk index and future market returns is stable. Moreover, we emphasize that the predictive power is not confined to any particular time period and is robust to whatever choice of subsamples.

4.5 International evidence

Besides the SPF data provided by the Federal Reserve Bank of Philadelphia, the European Central Bank (ECB) collects forecasts on the expected rates of inflation, real GDP growth, and unemployment in the euro area from professional forecasters in terms of survey. We refer to this data set as ECB SPF. In this subsection, we use the ECB SPF data to construct a macro risk index for each of the seven countries in Europe we considered, including France, Germany, Italy, the Netherlands, Sweden, Switzerland, and the United Kingdom,[16] and for the aggregate European market. Similarly, we apply the PLS method to construct the macro risk indices based on the current-year and the next-year forecasts on the expected rates of inflation, real GDP growth, and unemployment from the ECB SPF. We follow Rapach, Strauss, and Zhou (2013) to obtain national currency excess returns for the seven countries and use the STOXX Europe 600 Index as a proxy for the European equity market.[17] Then, we run the following predictive regression for each country i (or for Europe),

$$R_{t+1}^{i} = \alpha + \beta\,\mathrm{EMR}_{i,t}^{\mathrm{PLS}} + \epsilon_{t+1}^{i}, \tag{14}$$

where R_{t+1}^{i} is the excess return on the i-th country or on the STOXX Europe 600 Index and $\mathrm{EMR}_{i,t}^{\mathrm{PLS}}$ denotes the extracted PLS macro risk index using R^{i} as the target variable.

[**Insert Table VII here**]

Table VII looks at the predictability of European countries (Europe) using the ECB SPF data. Over the sample period from 1999Q1 to 2019Q4, where 1999Q1 marks the initiation of the ECB SPF data, the macro risk indices for European countries (Europe) can significantly predict future returns on the corresponding market index. All regression estimates are highly significant at the 1% level associated with hefty in-sample R^2 statistics ranging from 7.69% (Germany) to 13.93% (the Netherlands). In addition, the European estimates and R^2 values in this sample period (post-2000) are comparable to those for the U.S. in Table VI. Consistent with our benchmark findings, the results on international predictability suggest that the predictive power derived from using SPF data to construct the macro risk index is not peculiar to the U.S. stock market.

5 Economic value of return predictability

The results presented in Section 4 provide statistical evidence for the market return predictability of the aligned macro risk index, while we are also interested in its asset allocation implications. In this section, we assess the economic significance of the predictability by answering the question that how much additional risk-adjusted value can an investor gain if

[16]Though Sweden, Switzerland, and the United Kingdom are not a member state of the eurozone, they are important industrialized countries with developed market indices, as in Rapach, Strauss, and Zhou (2013).

[17]We use returns on the FTSE MIB index as Italian returns. Excess returns are computed relative to each countrys three-month Treasury bill rate.

she switches from the non-predictability benchmark to a predictive regression forecast? Following Campbell and Thompson (2008) and Rapach, Strauss, and Zhou (2010), we consider a mean-variance investor who can dynamically allocate her wealth between the aggregate stock market and a risk-free bond. The portfolio choice problem faced by the investor at time t follows

$$\max_{w_t} \quad E_t(R_{p,t+1}) - \frac{1}{2}\gamma Var_t(R_{p,t+1}), \tag{15}$$

where γ is her relative risk-aversion coefficient, w_t is the proportion of wealth allocated to the stock market during period $t+1$, and the portfolio return at time $t+1$ is calculated as $R_{p,t+1} = R_{f,t+1} + w_t R_{t+1}$ where $R_{f,t+1}$ is the risk-free rate. The optimal solution to the above maximization problem (15) is:

$$w_t^* = \frac{E_t(R_{t+1})}{\gamma Var_t(R_{t+1})}. \tag{16}$$

If the investor believes that the stock market is unpredictable and treats the prevailing mean as the best forecast for return, the optimal portfolio weight is given by

$$w_{b,t} = \frac{\bar{R}_{t+1}}{\gamma \hat{\sigma}_{t+1}^2}, \tag{17}$$

where $\bar{R}_{t+1}$ is the historical mean excess return benchmark forecast and $\hat{\sigma}_{t+1}^2$ is the market variance forecast. Following Campbell and Thompson (2008), we use the sample variance of market excess returns over the past ten years as the forecast for market variance. Alternatively, if the investor chooses to adopt the return forecast generated by predictive regression model i, she will allocate the following share of her portfolio to the stock market:

$$w_{i,t} = \frac{\hat{R}_{t+1}^i}{\gamma \hat{\sigma}_{t+1}^2} \tag{18}$$

where $\hat{R}_{t+1}^i$ is the OOS forecast of market excess return by model i. Note that the market variance forecast $\hat{\sigma}_{t+1}^2$ is the same for all portfolios, so that the differences among portfolio weights are determined only by the different return forecasts. Similarly to Campbell and Thompson (2008), we impose short sale and maximum leverage constraints to restrict w_t to lie between zero and 1.5.

We use three risk-adjusted measures to evaluate and compare the performance of different portfolios. First, we consider the certain equivalent return (CER) defined as

$$\text{CER}_p = \hat{\mu}_p - \frac{1}{2}\gamma \hat{\sigma}_p^2, \tag{19}$$

where $\hat{\mu}_p$ and $\hat{\sigma}_p^2$ are the mean and variance, respectively, of the portfolio returns over the

OOS evaluation period. Accordingly, the economic value of the return predictability afforded by predictive model i can be gauged by the gains in CER relative to the historical mean benchmark model, that is,

$$\text{CER gain}_i = \text{CER}_i - \text{CER}_b,$$

where CER_i and CER_b correspond to the CERs generated by portfolio (18) and portfolio (17), respectively. We multiply the gains in CER for a quarterly portfolio by four such that the CER gain can be seen as an annual management fee that an investor would be willing to pay to switch from the non-predictable benchmark to a predictive model. We determine the significance of the CER gain using the statistic of Diebold and Mariano (1995).

Second, we report the portfolio's Sharpe ratio which is the most commonly used performance measure. Since the Sharpe ratio is independent of the choice of investor's risk aversion level, we could use it to compare portfolio performance across different γs. We assess the significance of the difference between two Sharpe ratios by the test of Jobson and Korkie (1981) with correction by Memmel (2003). Though the Sharpe ratio is widely adopted, it is subject to a few drawbacks, including being easily manipulated (Goetzmann, Ingersoll, Spiegel, and Welch, 2007) and inadequate to penalize suboptimal portfolio leverage (Kan and Zhou, 2007). Therefore, we further apply the manipulation-proof measure of Goetzmann, Ingersoll, Spiegel, and Welch (2007) as third measure to gauge the portfolio performance:

$$\Theta = \frac{1}{1-\rho} \ln \left(\frac{1}{T} \sum_{t=1}^{T} \left(\frac{1 + R_{p,t+1}}{1 + R_{f,t+1}} \right)^{1-\rho} \right), \tag{20}$$

where ρ denotes the extent to which the risk is penalized and $R_{p,t+1}$ is the portfolio return at time $t+1$. Similarly to CER, the statistic Θ summarizes the portfolio performance over time by a single score that can be interpreted as the portfolio's premium return after adjusting for risk. We then calculate the difference between the Θ of predictive model i and the Θ of the benchmark to quantify the the economic value of return predictability, and we express the term as annualized percentage.

Asset allocatioin results

[**Insert Table VIII here**]

We assume that the investor rebalances her portfolio at the quarterly frequency which coincides with the forecast horizon of the one-step ahead return forecast in subsection 4.2. Consequently, the OOS evaluation period for the asset allocation practice is from 1984Q1 to 2019Q4 with a total of 144 observations. Panels A and B of Table VIII report the portfolio performance measured by the above-mentioned metrics under two different levels of risk aversion, three and five, respectively. The annualized Sharpe ratios reported in the first column suggest that the portfolio employing the forecasts generated by $\mathrm{MR}^{\mathrm{PLS}}$ attains the highest value of all Sharpe ratios at $\gamma = 3$. Its Sharpe ratio (0.64) is significantly higher than that of the historical mean benchmark (0.41) at the 5% level as well as that of the buy-and-hold strategy (0.52). The performance of $\mathrm{MR}^{\mathrm{PLS}}$ becomes even more impressive when we move to the annualized CER gains (in percent) reported in the second column. Consistent with the Sharpe ratio results, $\mathrm{MR}^{\mathrm{PLS}}$ delivers a CER gain of 324 basis points (bps) that outweighs all the other predictive models and the passively managed portfolio (182 bps). Namely, an investor is willing to pay an annual portfolio management fee of 324 bps to exploit the return forecasts based on $\mathrm{MR}^{\mathrm{PLS}}$ instead of using the historical mean forecast. Among the 19 predictive models that are based on economic predictors, only four produce positive CER gains, with the highest being 103 bps by LTR. In addition, the results of the manipulation-proof measure (annualized and in percent) in the third column confirm the robustness of the economic value generated by $\mathrm{MR}^{\mathrm{PLS}}$. We find the size of risk-adjusted returns produced by $\mathrm{MR}^{\mathrm{PLS}}$ barely changed when switching from the CER measure to the manipulation-proof measure. Notably, the gain in Θ of $\mathrm{MR}^{\mathrm{PLS}}$ is about 2.5 to four times larger than those of the other models based on SPF variables and is twice the amount of the buy-and-hold portfolio. In contrast, none of the conventional predictors outperforms the naive buy-and-hold portfolio in terms of the gain in Θ.

The asset allocation results for $\gamma = 5$ reported in Panel B are broadly in line with the results in Panel A. The predictive power of $\mathrm{MR}^{\mathrm{PLS}}$ continues to generate substantial economic value to a mean-variance investor that exceeds all the other predictive models and the passively managed portfolio by a large margin. Moreover, we find that the gains in CER and Θ of the buy-and-hold portfolio greatly decrease from 182 bps and 157 bps, respectively, to 24 bps and -54 bps, respectively, when we change the risk aversion coefficient from three to five. This is mainly because the risk is not so heavily penalized when $\gamma = 3$. In contrast, the economic benefit produced by $\mathrm{MR}^{\mathrm{PLS}}$ remains sizable when $\gamma = 5$, with gains in CER and Θ of 195 bps and 189 bps, respectively.

[**Insert Figure 4 here**]

To provide further insight on the behavior of the portfolio base on $\mathrm{MR}^{\mathrm{PLS}}$, we follow Eriksen (2017) to plot its cumulative CER gain along with the cumulative CER gain of the buy-and-hold portfolio in Figure 4. Panels A and B illustrate that $\mathrm{MR}^{\mathrm{PLS}}$ consistently generates more CER than the historical mean benchmark during the whole sample period rather than over some special periods under both risk aversion levels considered. The solid line (portfolio base on $\mathrm{MR}^{\mathrm{PLS}}$) is upward sloping at most of the time and even during the market downturns, such as the 2008 global financial crisis, signaling a prominent market-timing ability. On the other hand, as depicted by Panel B, the utility gain delivered by the buy-and-hold strategy grows remarkably from 1996 to 2000. This passively managed portfolio, however, unavoidably suffers a large drawdown in the subsequent dot-com bubble and during the 2008 crisis, and the cumulative CER gain at the end of 2019 hardly recovers to its level at the end of 1999.

In short, the results from the asset allocation practice demonstrate that the predictive power of $\mathrm{MR}^{\mathrm{PLS}}$ generates substantial economic value for a mean-variance investor in terms of the Sharpe ratio and the gains in CER and manipulation-proof measure. The performance of $\mathrm{MR}^{\mathrm{PLS}}$ clearly stands out and is stronger than those of the other models based on SPF variables, the models based on the 16 economic predictors, and the passive management.

6 Explore the Source of Predictability

In this section, we explore the source of the aligned macro risk index's predictive power from four perspectives, including the information content of the SPF data, inspections on the econometric method, links to the macroeconomy, and the economic channel. Finally, we consider additional control variables that may share a common source of predictability with our macro risk index.

6.1 Predicting macroeconomic activities with SPF data

A branch of study discussing the rationality of professional forecasts on economic variables argues that the forecasts are criticized for inefficiency due to herding behavior, reputation concern, and under- or over-reaction (Bordalo, Gennaioli, Ma, and Shleifer, 2020; Lamont, 2002). Because our aligned macro risk index is formed based on survey forecasts made by professional forecasters, this may raise the question of whether the SPF forecasts reflect rational beliefs about the future business conditions of forecasters. If so, the SPF forecasts first and foremost should predict macroeconomic activities. To this end, we consider four real macroeconomic activity measures, including the real consumption per capita, real GDP, real labor income, and industrial production,[18] and we use the same SPF data as in the return predictability analysis to predict these macro variables. Similarly, we apply the PLS method to condense information from the SPF forecasts for its efficacy in dimension reduction.

[**Insert Table IX here**]

Table IX looks at the predictability of real macroeconomic activities using the SPF forecasts. We report Newey-West t-statistics of coefficient estimates for the one-quarter ahead regression, and report t-statistics based on Hodrick (1992) standard errors at longer forecast

[18]All macro variables are deflated, seasonally adjusted, and continuously compounded growth rates. The data are obtained from the FRED database of St. Louis Fed `https://fred.stlouisfed.org/`.

horizons with overlapping observations. Panel A presents the results of univariate regression models. We find that in the PLS regression, the factors extracted from the SPF forecasts can significantly predict future macroeconomic activities. The predictive ability for the industrial production growth is particularly strong, with R^2s of 23.21% and 23.49% at the quarterly and annual horizons. Nevertheless, the growth rates of these macro variables are known to be highly persistent and can be predicted by financial indicators (Chen, 1991; Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1998; Harvey, 1988). We thus consider a multivariate setting that controls for the one-period-lagged growth rate, TBL, and TMS, where the results are shown in Panel B. After controlling for economic predictors and lagged growth rate, the PLS factors remain significant at the 10% level or better, except for two cases, suggesting that they provide incremental information to financial indicators. Therefore, results from Table IX indicate that the SPF forecasts contain forward-looking information about the economic conditions, and reflect forecasters' rational belief about the macro risk.[19]

6.2 Statistical explanation

According to the results in Section 4, the PLS macro risk index and the first PC of SPF7 display distinct patterns for market return predictability, especially in out-of-sample tests. In particular, $\mathrm{MR}^{\mathrm{PLS}}$ produces an R^2_{OS} of 3.12% that is the highest value of all the R^2_{OS} statistics in Table III, whereas $\mathrm{SPF7}^{\mathrm{PC}}$ underperforms the historical mean and is even inferior to some individual SPF variables. To explain the difference between the predictive power of the PLS factor and that of the PCR factor, Figure 5 presents the weights of $\mathrm{MR}^{\mathrm{PLS}}$ and $\mathrm{SPF7}^{\mathrm{PC}}$ on the seven SPF variables over the OOS evaluation period. It comes to our first observation that the PLS weights vary substantially more than the PCR weights, and the absolute weight ranking of SPF variables frequently changes. In sharp contrast, the PCR weights barely deviate from the full-sample average and their ranking remain the same over time. Given the data-generating process of equity premium is highly complex and constantly evolving (Rapach, Strauss, and Zhou, 2010, p. 845), the relation between the equity premium and the macroeconomic condition variables is implausible to be unchanged over time. As such, the more flexible weights relative to the PCR weights enable $\mathrm{MR}^{\mathrm{PLS}}$ to better capture the changing dynamics of the market.

[**Insert Figure 5 here**]

Figure 5 also illustrates that $\mathrm{MR}^{\mathrm{PLS}}$ assigns large weights to $\mathrm{Housing}_e$ and Unemp_e, while $\mathrm{SPF7}^{\mathrm{PC}}$ loads heavily on GDP_e and Recess_e and stingily on $\mathrm{Housing}_e$ and Unemp_e. Recall

[19]From an intuitive perspective, professional forecasters are more sophisticated and well-trained likely making the SPF forecasts more informative than the consumer survey forecasts. In addition, since the SPF respondents are anonymous, this mitigates the impact of reputation or career concerns.

that Housing_e is the only SPF variable with a significantly positive R^2_{OS} and Unemp_e exhibits outstanding predictive power after the recent millennium (see Figure 2), this further explains why MR^{PLS} outperforms SPF7^{PC} in predicting the market. Thus, the superior forecasting ability of MR^{PLS} can also be attributed to the more sensible factor weights assigned by it relative to SPF7^{PC}, in that by design, the former picks the weights according to the covariance with the forecast target while the latter primarily tracks the variations of the individual SPF variables. Moreover, we examine whether MR^{PLS} and SPF7^{PC} successfully incorporate all the predictive information contained in individual SPF variables via the forecast encompassing test of Harvey, Leybourne, and Newbold (1998). The null hypothesis that MR^{PLS} encompasses one of the individual SPF variables has p-values ranging from 0.37 to 0.70, whereas we can reject the null hypothesis that SPF7^{PC} encompasses Indprod_e or Housing_e at least at the 10% level.[20] This reaffirms that PLS is an efficient approach to condense multivariate information.

[**Insert Figure 6 here**]

To glean further insights into the forecasting performance of different predictive models, Figure 6 plots the market return forecasts by SPF7^{PC}, SPF7^{FC}, MR^{PLS}, and the benchmark model over the OOS period. The solid line depicts the realized market excess return smoothed with four quarter moving average serving as a reference. First note that the MR^{PLS} forecast is more volatile than the SPF7^{PC} forecast, the SPF7^{FC} forecast, and the historical mean forecast. Because the actual returns display even stronger volatility, intuitively, the MR^{PLS} forecast is more likely to capture the great variation in the market return. Second, we note that the forecasted returns based on MR^{PLS} and SPF7^{FC} usually move in the same direction, and a correlation test shows that the correlation between them is about 82%. Nonetheless, we can tell that the MR^{PLS} forecast tracks the actual returns more closely than the SPF7^{FC} forecast does, especially during the period of 2007 through 2019. Since the PLS method dynamically adjusts the weights on individual predictors as shown in Figure 5, it more efficiently summarizes the sparse forecasting information in the SPF variables than an equal-weighted combination of forecasts does. This provides an intuitive explanation why MR^{PLS} outperforms SPF7^{FC} in tracking the expected market return.

6.3 Counter-cyclical forecast performance

Counter-cyclical equity premia and the PLS factor

Dynamic asset pricing models posit that risk-averse investors demand for higher equity premia in recessions, leading to counter-cyclical equity premia[21], an implication well sup-

[20] More detailed results are available upon request.

[21] Leading asset pricing models such as external habit formation model of Campbell and Cochrane (1999), the long run risks model of Bansal and Yaron (2004), and the time-varying disaster risk model of Gabaix

ported by empirical studies (Fama and French, 1989; Golez and Koudijs, 2018; Rapach, Strauss, and Zhou, 2010). In light of this, we would expect the PLS factor $\mathrm{MR}^{\mathrm{PLS}}$ to well predict the market if it also exhibits a cyclical pattern. Figure 7 supports our conjecture. The solid line depicts the PLS factor evaluated over the OOS period and the dash-dotted line depicts the Chicago Fed National Activity Index (CFNAI). First observe that the sold line and the dash-dotted line usually move in an opposite direction. For instance, during the 2008 global financial crisis, the CFNAI dropped to an all-time low whereas the PLS factor reached its sample maximum. Because a positive (negative) value of the CFNAI means the aggregate economic activity is above (below) the long-term trend, the negative correlation between the two curves suggests that the PLS factor is high when business conditions are weak and is low when conditions are strong. Second, the PLS factor decreases near the peaks preceding the 1990, 2001, and 2008 recessions identified by the NBER (shaded area), while it spikes at the troughs of these recessions, displaying salient counter-cyclical dynamics. Therefore, the depicted PLS factor implies that $\mathrm{MR}^{\mathrm{PLS}}$ encompasses a cyclical component of the macroeconomic risk.

[**Insert Figure 7 here**]

Forecast performance over different economic conditions

Recent empirical studies document that the strength of equity premium predictability is time-varying. Specifically, the forecasting ability of conventional predictors, such as the dividend-price ratio, is concentrated around economic recessions (Dangl and Halling, 2012; Henkel, Martin, and Nardari, 2011; Rapach, Strauss, and Zhou, 2010). It is thus of interest to investigate how the predictive performance of $\mathrm{MR}^{\mathrm{PLS}}$ is related to the economic states. To this end, we calculate the contemporaneous correlations of the forecasted returns, difference in squared forecast error (DSFE), and CER gains (CERG), respectively, with the several economic condition measures, including the real GDP growth, real consumption growth, real labor income growth, the CFNAI, and the macro uncertainty index by Jurado, Ludvigson, and Ng (2015) in Panel A of Table X. The significantly negative correlation coefficients in the first row reveal that the equity premium forecasts made by $\mathrm{MR}^{\mathrm{PLS}}$ are large during economic bad times when the economy growth and the CFNAI are low, and are small during good times when the growth and the CFNAI are high, consistent with the counter-cyclicality of the equity premia. Moreover, we can tell from the second and third columns that the forecasting performance of $\mathrm{MR}^{\mathrm{PLS}}$ and associated economic gains relative to the historical mean benchmark are negatively correlated with the economy growth and positively correlated with the macro uncertainty. This indicates that the forecast gains of $\mathrm{MR}^{\mathrm{PLS}}$ tend to be large during economic contractions in which the CFNAI is low and economic uncertainty is high.

(2012) all imply counter-cyclical equity premia.

[**Insert Table X here**]

The results of subsample R^2_{OS} over different business conditions reported in Panel B of Table X shed further light on the time-varying performance of $\mathrm{MR}^{\mathrm{PLS}}$. In particular, $\mathrm{MR}^{\mathrm{PLS}}$ evinces more substantial predictive power for the market excess return during NBER-dated recessions than during expansions, with an R^2_{OS} of 7.29% that is three times larger than its expansion counterpart. A similar scenario is observed when the subsamples are categorised into high- and low-growth periods based on the real GDP growth. Although it has not come to an agreement on the reason for the time-varying predictability, the counter-cyclicality of equity premium provides a plausible explanation, in that the equity premium is more variable in recessions, leading to counter-cyclical predictability.[22] To summarize, the forecasting ability of $\mathrm{MR}^{\mathrm{PLS}}$ is concentrated in periods with recessions consistent with the extant literature.

Predicting portfolios sorted on size and industry

If $\mathrm{MR}^{\mathrm{PLS}}$ captures the cyclical fluctuations in the macroeconomy, we would expect it to predict the returns on stock portfolios that also display cyclical pattern besides the aggregate market premium. To corroborate our conjecture, we explore the forecasting ability of $\mathrm{MR}^{\mathrm{PLS}}$ for portfolios sorted on size and industry SIC codes, whose returns are known to be affected by the business cyclicality (Gomes, Kogan, and Yogo, 2009; Perez-Quiros and Timmermann, 2000).[23] Panel A in Table XI presents the forecasting results for Fama-French 10 size portfolios. We find that $\mathrm{MR}^{\mathrm{PLS}}$ significantly predicts all portfolios sorted on size, with in-sample R^2s over 5%, and the positive R^2_{OS} values in the fourth column confirm the robustness of predictive power in OOS test. Remarkably, the regression slopes shown in the second column are monotonically increasing from large to small firms, signaling an ascending predictability. As argued by Perez-Quiros and Timmermann (2000), small firms are more susceptible than large firms to the changes in the underlying business conditions. Therefore, the stronger forecasting ability of $\mathrm{MR}^{\mathrm{PLS}}$ for small size portfolios echoes previous findings, in which we illustrate that $\mathrm{MR}^{\mathrm{PLS}}$ captures the business cycle variation.

[**Insert Table XI here**]

Besides the size portfolios, we also examine the predictability for industry portfolios. According to Panel B in Table XI, $\mathrm{MR}^{\mathrm{PLS}}$ significantly in-sample predicts the industry returns except for the energy category, while the regression slopes and R^2 values vary across

[22] For instance, in the famous external habit formation model of Campbell and Cochrane (1999), heightened risk aversion drives up the expected return in recessions. Accordingly, predictable variation in stock return tends to be large during recessions and can be tracked by variables that can predict the risk aversion level.

[23] The data of portfolios sorted on size and industry come from Kenneth Frenchs data library.

industries. The slope estimates for cyclical industries, such as durable goods (0.146) and hi-tech equipment (0.111), are usually two to three times larger than that for defensive industries that are less sensitive to business cyclicality, including healthcare equipment (0.050) and utilities (0.047). Particularly, we uncover high level of predictability for returns on durable goods, with sizable in- and out-of-sample R^2s of 8.94% and 6.13%, respectively, whereas that for non-durable goods is comparatively weaker. Gomes, Kogan, and Yogo (2009) point out that the demand for durable goods displays stronger cyclical pattern than that for non-durable goods. Following this, the cash flow of firms that produce durable goods varies more over the business cycle relative to those produce non-durable goods, and so do their expected stock returns. The difference in their predictability reiterates the notion that $\mathrm{MR}^{\mathrm{PLS}}$ well characterizes the cyclical risk of macroeconomy.

6.4 Return decomposition

In this section, we investigate the underlying economic channel from which the predictive power of $\mathrm{MR}^{\mathrm{PLS}}$ stems. According to Campbell and Shiller (1988), the log total market return (r_{t+1}) can be decomposed into the expected return ($E_t[r_{t+1}]$) and two news components, the cash flow news (η_t^{CF}) and the discount rate news (η_t^{DR}):

$$r_{t+1} = \mathbb{E}_t(r_{t+1}) + \underbrace{(\mathbb{E}_{t+1} - \mathbb{E}_t)\left(\sum_{j=0}^{\infty} \rho^j \Delta d_{t+j+1}\right)}_{\text{cash flow news}} - \underbrace{(\mathbb{E}_{t+1} - \mathbb{E}_t)\left(\sum_{j=1}^{\infty} \rho^j r_{t+j+1}\right)}_{\text{discount rate news}} \tag{21}$$

where Δd_{t+j+1} is the log dividend growth at time $t + j + 1$ and ρ is a log-linearization constant ($\rho \in (0, 1)$).[24] Examining the predictability for each of the return components allows us to understand the source of forecasting power of $\mathrm{MR}^{\mathrm{PLS}}$. We estimate the three return components via the vector autoregression (VAR) methodology of Campbell (1991) and Campbell and Ammer (1993), and we run the following predictive regression,

$$y_{t+1} = \alpha_y + \beta_y \, \mathrm{MR}_t^{\mathrm{PLS}} + \epsilon_{t+1}, \tag{22}$$

where y_t is one of the three estimated return components for quarter t, and α_y is set equal to zero when $y = \hat{\eta}^{\mathrm{CF}}$ or $\hat{\eta}^{\mathrm{DR}}$.

[Insert Table XII here]

Panel A of Table XII reports the slope estimates of the above regression when the expected return, the cash flow news, and the discount rate news are estimated based on VAR

[24]See Appendix B for a detailed description on the return decomposition.

models comprising the log market return, DP, and one of the 15 economic predictors (excluding DP). Following Engsted, Pedersen, and Tanggaard (2012), we always include DP in the VAR to properly estimate the cash flow news and discount rate news. Most of the $\beta_{\hat{E}}$ estimates in the second column are statistically significant, revealing that $\mathrm{MR}^{\mathrm{PLS}}$ can predict the expected return conditional on variables in the first column. Only a few of the β_{CF} estimates in the fourth column are significant at the 10% level. In contrast, all the β_{DR} estimates in the sixth column are significant and are usually two to three times larger than the $\beta_{\hat{E}}$ or β_{CF} counterpart. The last row of panel A, where we employ the first three PCs extracted from the 16 economic predictors in the VAR, confirms the robustness of our findings. Consequently, the discount rate channel appears as the primary economic source of the market return predictability of $\mathrm{MR}^{\mathrm{PLS}}$.

The results in Panel A demonstrate that the time variation in expected return missed by the economic predictors, appearing as the discount rate news, is anticipated by the PLS macro risk index. As stressed by Welch and Goyal (2008) and Rapach, Strauss, and Zhou (2010), structural breaks and parameter instability render the forecasting relationships between conventional predictors, the valuation ratios in particular (Lettau and Van Nieuwerburgh, 2008), and future returns unstable, so that these predictors fail to fully track the equity premium variations in out sample. This leads to a natural question: economically, what is the part of variations in expected return that is missed by the conventional predictors and is captured by $\mathrm{MR}^{\mathrm{PLS}}$? Leading equilibrium pricing models link the time variation of the equity premium to the time-varying aggregate risk and/or investor's preference (Cochrane, 2008). In particular, Campbell and Cochrane (1999) elucidate that the counter-cyclical risk aversion gives rise to the counter-cyclical movements in equity premia. Intuitively, $\mathrm{MR}^{\mathrm{PLS}}$ displays a prominent counter-cyclical pattern that enables it to better capture the cyclicality of expected return induced by the counter-cyclical preference than conventional predictors. In this context, we would expect the predictability of the estimated discount rate news to be weakened when we add variables also containing information about the time-varying preference to the VAR.

To verify our conjecture, we consider a VAR comprising the log return, DP, the first three PCs of the 16 economic predictors, and the cyclical consumption of Atanasov, Møller, and Priestley (2020) which is used as a proxy for the surplus consumption ratio in Campbell and Cochrane (1999) that is closely tied to risk aversion. As Panel B indicates, the predictive power of $\mathrm{MR}^{\mathrm{PLS}}$ for discount rate news fades away and becomes insignificant after adding the cyclical consumption into the VAR setting, while the predictability of $\hat{E}_t[r_{t+1}]$ is substantially improved. Consistent with our expectation, the empirical evidence suggests that the market return predictability of $\mathrm{MR}^{\mathrm{PLS}}$ can be partially attributed to its ability to capture the variations in expected return due to changing risk aversion by which conventional predictors omit.

6.5 Relation with other predictive variables

Results from previous analyses demonstrate that the source of predictive power of $\mathrm{MR}^{\mathrm{PLS}}$ stems from its ability to characterize the cyclical macroeconomic risk. Besides the 16 predictors studied by Welch and Goyal (2008), the more recent literature also identifies a number of macro variables the can track business cycle conditions and significantly predict market return. We are then intrigued to examine the information content of our macro risk index relative to other extant macro variables. Specifically, we consider the output gap (OG) of Cooper and Priestley (2009), the cyclical consumption (CC) of Atanasov, Møller, and Priestley (2020), the share of labor income to consumption (s^w) of Santos and Veronesi (2006), the consumption volatility measure (σ_c) of Bansal, Khatchatrian, and Yaron (2005), the price-output ratio (py) of Rangvid (2006), and the quarterly non-housing consumption to total consumption ratio (house) following the construction of Piazzesi, Schneider, and Tuzel (2007).[25] Moreover, the SPF data may contain the sentiment and the extent of disagreement of the professional forecasters, which also display cyclical pattern and are used as market return predictors (Huang, Jiang, Tu, and Zhou, 2015; Huang, Li, and Wang, 2021). We further control the investor sentiment indexes of Huang, Jiang, Tu, and Zhou (2015) and Baker and Wurgler (2006) denoted by $\mathrm{senti}^{\mathrm{HJTZ}}$ and $\mathrm{senti}^{\mathrm{BW}}$, respectively, the Michigan Consumer Sentiment Index ($\mathrm{senti}^{\mathrm{MC}}$), and the disagreement index ($\mathrm{Disag}^{\mathrm{HLW}}$) of Huang, Li, and Wang (2021) to disentangle the source of predictive power of $\mathrm{MR}^{\mathrm{PLS}}$.[26]

[**Insert Table XIII here**]

Panel A of Table XIII shows that OG and CC significantly predict the quarterly market excess return in this sample, with corresponding R^2s of 4.92% and 3.98%. Nonetheless, among all macro variables considered, $\mathrm{MR}^{\mathrm{PLS}}$ appears as the strongest one for predicting the market. Consistent with Huang, Jiang, Tu, and Zhou (2015), $\mathrm{senti}^{\mathrm{HJTZ}}$ exhibits substantial predictive power, with an R^2 of 6.49%, whereas $\mathrm{senti}^{\mathrm{BW}}$ fails to generate significant forecasts for future market returns. Turning to the results for bivariate predictive regression in Panel B, we find that the slope estimates of $\mathrm{MR}^{\mathrm{PLS}}$ remain significant and sizable after controlling for the additional predictors considered, whereas the slopes of OG and CC become statistically insignificant. Also note that the magnitude and significance of the slopes of $\mathrm{MR}^{\mathrm{PLS}}$, $\mathrm{senti}^{\mathrm{HJTZ}}$, and $\mathrm{disag}^{\mathrm{HLW}}$ barely change in the bivariate regression, and the R^2 statistic of the bivariate model based on $\mathrm{MR}^{\mathrm{PLS}}$ and $\mathrm{senti}^{\mathrm{HJTZ}}$ is roughly equal to the sum of the R^2s for $\mathrm{MR}^{\mathrm{PLS}}$ and $\mathrm{senti}^{\mathrm{HJTZ}}$. This reflects that all these predictors contribute useful, yet different types of, information regarding forecasting market return.

[25] We follow Piazzesi, Schneider, and Tuzel (2007) to construct a quarterly variable that measures the expenditure share on non-housing consumption, while their variable is available annually.

[26] We thank Guofu Zhou and Dashan Huang for making the data available.

Furthermore, Panel C of Table XIII presents the contemporaneous correlations between $\mathrm{MR}^{\mathrm{PLS}}$ and the control variables. On the one hand, the high correlations (in absolute value) between $\mathrm{MR}^{\mathrm{PLS}}$, OG, and CC confirm that these macro variables share a commonality in that they capture the cyclical macroeconomic risk. Notwithstanding the similarity in information source, $\mathrm{MR}^{\mathrm{PLS}}$ contains richer information leading to the largest in-sample R^2 (5.75%) among these three variables, and subsumes the predictability of OG and CC as indicated by their insignificant bivariate regression coefficients in Panel B. On the other hand, we observe that $\mathrm{MR}^{\mathrm{PLS}}$ is weakly and negatively correlated to $\mathrm{senti}^{\mathrm{HJTZ}}$ and $\mathrm{Disag}^{\mathrm{HLW}}$, reiterating the distinctness in their sources of predictability. To conclude, the predictive power of $\mathrm{MR}^{\mathrm{PLS}}$ mainly derives from a macroeconomy-based channel rather than an investor sentiment or disagreement channel.

7 Conclusion

In this paper, we construct an aligned macro risk index using the SPF consensus forecasts on output, unemployment, housing starts, and inflation through the PLS method, and show that it exhibits significant predictability for stock market excess returns in both statistical and economic criteria. The predictive power of the aligned macro risk index is robust to the inclusion of a host of popular economic predictive variables and different subsamples considered. More importantly, the aligned macro risk index outperforms other macro-based variables, such as the output gap (Cooper and Priestley, 2009) and the cyclical consumption (Atanasov, Møller, and Priestley, 2020), and subsumes their predictability. We also construct a long-term macro risk index based on term structures of the SPF forecasts, which evinces even stronger forecasting ability for the long-term stock returns. In addition, we extend the study to international equity markets, and find that the macro risk indices formed based on the ECB SPF data are strong in-sample predictors of European countries stock returns.

Statistically, we attribute the success of our macro risk index to the efficiency of PLS in summarizing forecasting-relevant information from a vast pool of information. From an economic perspective, the aligned macro risk index closely tracks the cyclical variation in economic conditions and characterizes the macro risk varying over business cycles. We elucidate the source of its predictability by performing a battery of tests including comparison of predictability over good and bad economic times, predicting characteristics portfolios, and return decomposition. We illustrate that the index displays salient counter-cyclical dynamics reflecting the heightened risk and investor's preference during economic downturns, and its predictive power mainly stems from a discount rate channel, consistent with its ability to better capture the changing equity premium induced by cyclical risk.

In line with theoretical expectations, our empirical findings depict a sound relation between the time variation of expected returns and the changing macroeconomic risk. Moreover, the more substantial market predictability of the aligned macro risk index relative to other macroeconomic business cycle variables emphasizes the importance of considering multiple aspects of the economy when measuring the aggregate macro risk. This may in turns serve as guideposts for future theories of asset pricing that it is of significance to consider consumption risk as well as the non-consumption state variable risks related to housing, production, and labor income in explaining the time-varying equity premium.

References

Amato, Jeffery D, and Norman R Swanson, 2001, The real-time predictive content of money for output, *Journal of Monetary Economics* 48, 3–24.

Atanasov, Victoria, Stig V Møller, and Richard Priestley, 2020, Consumption fluctuations and expected returns, *Journal of Finance* 75, 1677–1713.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645–1680.

Bansal, Ravi, Varoujan Khatchatrian, and Amir Yaron, 2005, Interpretable asset markets?, *European Economic Review* 49, 531–560.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer, 2020, Overreaction in macroeconomic expectations, *American Economic Review* 110, 2748–82.

Brunnermeier, Markus, Emmanuel Farhi, Ralph SJ Koijen, Arvind Krishnamurthy, Sydney C Ludvigson, Hanno Lustig, Stefan Nagel, and Monika Piazzesi, 2021, Perspectives on the Future of Asset Pricing, *Review of financial studies*.

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

Campbell, John Y, and John Ammer, 1993, What moves the stock and bond markets? A variance decomposition for long-term asset returns, *Journal of Finance* 48, 3–37.

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

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

Campbell, John Y, and Samuel B Thompson, 2008, Predicting excess stock returns out of sample: Can anything beat the historical average?, *Review of Financial Studies* 21, 1509–1531.

Campbell, Sean D, and Francis X Diebold, 2009, Stock returns and expected business conditions: Half a century of direct evidence, *Journal of Business & Economic Statistics* 27, 266–278.

Chen, Long, and Lu Zhang, 2011, Do time-varying risk premiums explain labor market performance?, *Journal of Financial Economics* 99, 385–399.

Chen, Nai-Fu, 1991, Financial investment opportunities and the macroeconomy, *Journal of Finance* 46, 529–554.

Choi, James J, and Adriana Z Robertson, 2020, What matters to individual investors? Evidence from the horse's mouth, *Journal of Finance* 75, 1965–2020.

Clark, Todd E, and Kenneth D West, 2007, Approximately normal tests for equal predictive accuracy in nested models, *Journal of Econometrics* 138, 291–311.

Cochrane, John H, 1991, Production-based asset pricing and the link between stock returns and economic fluctuations, *Journal of Finance* 46, 209–237.

Cochrane, John H., 2008, Chapter 7 - Financial Markets and the Real Economy, in Rajnish Mehra, eds.: *Handbook of the Equity Risk Premium* (Elsevier, San Diego).

Cochrane, John H, 2011, Presidential address: Discount rates, *Journal of finance* 66, 1047–1108.

Cochrane, John H, 2017, Macro-finance, *Review of Finance* 21, 945–985.

Cooper, Ilan, and Richard Priestley, 2009, Time-varying risk premiums and the output gap, *Review of Financial Studies* 22, 2801–2833.

Croushore, Dean, and Tom Stark, 2001, A real-time data set for macroeconomists, *Journal of Econometrics* 105, 111–130.

Dangl, Thomas, and Michael Halling, 2012, Predictive regressions with time-varying coefficients, *Journal of Financial Economics* 106, 157–181.

Diebold, Francis X, and Robert S Mariano, 1995, Comparing predictive accuracy, *Journal of Business & Economic Statistics* 13, 253–263.

Diebold, Francis X, and Glenn D Rudebusch, 1990, A nonparametric investigation of duration dependence in the American business cycle, *Journal of Political Economy* 98, 596–616.

Engsted, Tom, Thomas Q Pedersen, and Carsten Tanggaard, 2012, Pitfalls in VAR based return decompositions: A clarification, *Journal of Banking & Finance* 36, 1255–1265.

Eriksen, Jonas N., 2017, Expected business conditions and bond risk premia, *Journal of Financial and Quantitative Analysis* 52, 1667–1703.

Estrella, Arturo, and Gikas A Hardouvelis, 1991, The term structure as a predictor of real economic activity, *Journal of Finance* 46, 555–576.

Estrella, Arturo, and Frederic S Mishkin, 1998, Predicting US recessions: Financial variables as leading indicators, *Review of Economics and Statistics* 80, 45–61.

Fama, Eugene F, 1990, Stock returns, expected returns, and real activity, *Journal of Finance* 45, 1089–1108.

Fama, Eugene F, 1991, Efficient capital markets: II, *Journal of Finance* 46, 1575–1617.

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

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

Filardo, Andrew J, and Stephen F Gordon, 1998, Business cycle durations, *Journal of Econometrics* 85, 99–123.

Gabaix, Xavier, 2012, Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance, *Quarterly Journal of Economics* 127, 645–700.

Ghysels, Eric, Casidhe Horan, and Emanuel Moench, 2018, Forecasting through the Rearview Mirror: Data Revisions and Bond Return Predictability, *Review of Financial Studies* 31, 678–714.

Goetzmann, William, Jonathan Ingersoll, Matthew Spiegel, and Ivo Welch, 2007, Portfolio performance manipulation and manipulation-proof performance measures, *Review of Financial Studies* 20, 1503–1546.

Golez, Benjamin, and Peter Koudijs, 2018, Four centuries of return predictability, *Journal of Financial Economics* 127, 248–263.

Gomes, Joao F, Leonid Kogan, and Motohiro Yogo, 2009, Durability of output and expected stock returns, *Journal of Political Economy* 117, 941–986.

Goyal, Amit, Ivo Welch, and Athanasse Zafirov, 2021, A comprehensive look at the empirical performance of equity premium prediction ii, Working paper, Swiss Finance Institute.

Granger, Clive WJ, and Paul Newbold, 1973, Some comments on the evaluation of economic forecasts, *Applied Economics* 5, 35–47.

Hamilton, James D., 1989, A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica* 57, 357–384.

Harvey, Campbell R, 1988, The real term structure and consumption growth, *Journal of Financial Economics* 22, 305–333.

Harvey, David, Stephen Leybourne, and Paul Newbold, 1997, Testing the equality of prediction mean squared errors, *International Journal of Forecasting* 13, 281–291.

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

Henkel, Sam James, J Spencer Martin, and Federico Nardari, 2011, Time-varying short-horizon predictability, *Journal of Financial Economics* 99, 560–580.

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

Huang, Dashan, Fuwei Jiang, Jun Tu, and Guofu Zhou, 2015, Investor sentiment aligned: A powerful predictor of stock returns, *Review of Financial Studies* 28, 791–837.

Huang, Dashan, Jiangyuan Li, and Liyao Wang, 2021, Are disagreements agreeable? Evidence from information aggregation, *Journal of Financial Economics* 141, 83–101.

Jobson, J Dave, and Bob M Korkie, 1981, Performance hypothesis testing with the Sharpe and Treynor measures, *Journal of Finance* pp. 889–908.

Jurado, Kyle, Sydney C Ludvigson, and Serena Ng, 2015, Measuring uncertainty, *American Economic Review* 105, 1177–1216.

Kan, Raymond, and Guofu Zhou, 2007, Optimal portfolio choice with parameter uncertainty, *Journal of Financial and Quantitative Analysis* 42, 621–656.

Kelly, Bryan, and Seth Pruitt, 2013, Market expectations in the cross-section of present values, *Journal of Finance* 68, 1721–1756.

Kelly, Bryan, and Seth Pruitt, 2015, The three-pass regression filter: A new approach to forecasting using many predictors, *Journal of Econometrics* 186, 294–316.

Kostakis, Alexandros, Tassos Magdalinos, and Michalis P Stamatogiannis, 2015, Robust econometric inference for stock return predictability, *Review of Financial Studies* 28, 1506–1553.

Lamont, Owen A, 2002, Macroeconomic forecasts and microeconomic forecasters, *Journal of Economic Behavior & Organization* 48, 265–280.

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

Lettau, Martin, and Sydney C. Ludvigson, 2010, Chapter 11 - Measuring and modeling variation in the risk-return trade-off, in Yacine At-Sahalia, and Lars Peter Hansen, eds.: *Handbook of Financial Econometrics: Tools and Techniques* (North-Holland, San Diego).

Lettau, Martin, and Stijn Van Nieuwerburgh, 2008, Reconciling the return predictability evidence, *Review of Financial Studies* 21, 1607–1652.

Ludvigson, Sydney C., and Serena Ng, 2007, The empirical riskreturn relation: a factor analysis approach, *Journal of Financial Economics* 83, 171–222.

Ludvigson, Sydney C, and Serena Ng, 2009, Macro factors in bond risk premia, *Review of Financial Studies* 22, 5027–5067.

Lustig, Hanno N, and Stijn G Van Nieuwerburgh, 2005, Housing collateral, consumption insurance, and risk premia: An empirical perspective, *Journal of Finance* 60, 1167–1219.

Manski, Charles F, 2004, Measuring expectations, *Econometrica* 72, 1329–1376.

Memmel, Christoph, 2003, Performance hypothesis testing with the Sharpe ratio, *Finance Letters* 1.

Neely, Christopher, David E. Rapach, Jun Tu, and Guofu Zhou, 2014, Forecasting the Equity Risk Premium: The Role of Technical Indicators, *Management Science* 60, 1772–1791.

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

Perez-Quiros, Gabriel, and Allan Timmermann, 2000, Firm size and cyclical variations in stock returns, *Journal of Finance* 55, 1229–1262.

Piazzesi, Monika, Martin Schneider, and Selale Tuzel, 2007, Housing, consumption and asset pricing, *Journal of Financial Economics* 83, 531–569.

Rangvid, Jesper, 2006, Output and expected returns, *Journal of Financial Economics* 81, 595–624.

Rapach, David, and Guofu Zhou, 2013, Chapter 6 - Forecasting stock returns, in Graham Elliott, and Allan Timmermann, eds.: *Handbook of Economic Forecasting* (Elsevier,).

Rapach, David E, Jack K Strauss, and Guofu Zhou, 2010, Out-of-sample equity premium prediction: Combination forecasts and links to the real economy, *Review of Financial Studies* 23, 821–862.

Rapach, David E, Jack K Strauss, and Guofu Zhou, 2013, International stock return predictability: what is the role of the United States?, *Journal of Finance* 68, 1633–1662.

Santos, Tano, and Pietro Veronesi, 2006, Labor income and predictable stock returns, *Review of Financial Studies* 19, 1–44.

Welch, Ivo, and Amit Goyal, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

Wold, Herman, 1966, Estimation of principal components and related models by iterative least squares, *Multivariate Analysis* pp. 391–420.

Table I: **Descriptive Statistics**

This table presents descriptive statistics of the SPF consensus macroeconomic forecasts and the CRSP value-weighted market excess return. The six statistics reported for each variable are the average (Mean), standard deviation (Std), skewness (Skew), kurtosis (Kurt), median (Med), and the first-order autocorrelation coefficient (AR(1)). The heading $\rho(X_t, R_{t+1})$ refers the Pearson correlation of the variable forecast in the first column of Panel A at time t with the excess return on the CRSP index (R) at time $t+1$. The SPF data include survey forecasts for seven macroeconomic variables: 1) the real GDP (GDP$_e$), 2) the unemployment rate (Unemp$_e$), 3) the probability of a decline in real GDP (Recess$_e$), 4) industrial production index (Indprod$_e$), 5) corporate profits after tax (Cprof$_e$), 6) housing starts (Housing$_e$), 7) the GDP price index (Infl$_e$). The forecasting horizon spans from the current quarter to three-quarter ahead. The consensus forecasts for Unemp$_e$ and Recess$_e$ are in levels, while the forecasts take the form of quarter-over-quarter growth rates (annualized and in percentage) for the remaining five variables. The market risk premium (R) is calculated as the return on the CRSP value-weighted index in excess of the short-term T-bill rate. The sample period is from 1968Q4 to 2019Q4, 205 quarters in total.

(1)	(2) **Mean**	(3) **Std**	(4) **Skew**	(5) **Kurt**	(6) **Med**	(7) **AR(1)**	(8) $\rho(\mathbf{X_t}, \mathbf{R_{t+1}})$
Panel A: SPF Macroeconomic Forecasts							
I. Current Quarter Forecast							
GDP$_e$	2.32	2.11	−1.03	5.32	2.51	0.72	−0.05
Indprod$_e$	2.44	4.28	−1.00	6.00	3.00	0.62	−0.07
Recess$_e$	18.31	22.03	1.97	5.96	9.51	0.74	0.02
Unemp$_e$	6.17	1.63	0.64	2.82	5.87	0.96	0.16
Cprof$_e$	6.21	11.39	−0.01	5.49	5.86	0.59	−0.02
Housing$_e$	0.49	21.08	0.35	3.09	−1.55	0.43	0.19
Infl$_e$	3.48	2.22	1.36	4.09	2.64	0.96	−0.08
II. 1-Quarter Ahead Forecast							
GDP$_e$	2.59	1.64	−0.70	5.20	2.56	0.79	−0.05
Indprod$_e$	3.12	3.01	−0.47	5.28	3.18	0.73	−0.01
Recess$_e$	18.59	15.70	1.86	6.06	12.68	0.79	0.04
Unemp$_e$	6.19	1.60	0.64	2.81	5.84	0.96	0.16
Cprof$_e$	6.45	8.91	−0.02	4.65	6.06	0.62	−0.02
Housing$_e$	5.74	17.87	1.20	5.38	1.90	0.83	0.10
Infl$_e$	3.46	2.04	1.31	3.90	2.67	0.97	−0.06
III. 2-Quarter Ahead Forecast							
GDP$_e$	2.80	1.28	−0.54	6.08	2.74	0.81	−0.05
Indprod$_e$	3.45	2.40	0.03	4.70	3.27	0.79	−0.02
Recess$_e$	17.78	10.23	1.84	6.57	14.86	0.77	0.05
Unemp$_e$	6.16	1.56	0.62	2.79	5.85	0.96	0.16
Cprof$_e$	7.43	7.74	1.20	7.82	6.16	0.57	−0.03
Housing$_e$	8.08	17.04	1.24	4.75	4.25	0.86	0.08
Infl$_e$	3.47	1.93	1.22	3.63	2.81	0.98	−0.05
IV. 3-Quarter Ahead Forecast							
GDP$_e$	2.98	0.96	0.54	3.91	2.86	0.83	−0.10
Indprod$_e$	3.70	1.98	0.69	4.74	3.32	0.84	−0.04
Recess$_e$	17.17	6.27	0.89	3.92	16.43	0.77	0.06
Unemp$_e$	6.12	1.51	0.60	2.77	5.77	0.96	0.16
Cprof$_e$	8.18	6.92	1.84	10.66	6.65	0.57	−0.05
Housing$_e$	8.53	15.24	0.98	3.49	4.66	0.89	0.05
Infl$_e$	3.46	1.85	1.14	3.36	2.80	0.98	−0.04
Panel B: Quarterly Market Risk Premium (%)							
R	1.65	8.64	−0.52	3.68	2.69	0.05	-

Table II: **In-sample Return Predictability: 1969Q1-2019Q4**

This table presents the OLS estimates, Newey-West t-statistics, and R^2 of the in-sample predictive regressions for quarterly market excess return. Panel A reports the results of the univariate predictive regression model,

$$R_{t+1} = \alpha + \beta X_t + \epsilon_{t+1},$$

where R_{t+1} is the excess return on the CRSP value-weighted index (annualized) for quarter $t+1$. The predictive variables X_t include the seven current-quarter survey forecasts (SPF7) as well as a set of 16 financial and economic variables (Econ) from Welch and Goyal (2008). The terms $\mathrm{SPF7^{PC}}$ and $\mathrm{Econ^{PC}}$ denote the first principal component (PC) of SPF7 and Econ, respectively. The terms $\mathrm{MR^{PLS}}$ and $\mathrm{Econ^{PLS}}$ denote the PLS factors extracted from SPF7 and Econ, respectively. The term $\mathrm{SPF7^{KS}}$ refers to the multivariate linear regression (kitchen sink) using SPF7. Panel B reports the results of the multivariate regression model,

$$R_{t+1} = \alpha + \beta\, \mathrm{MR}_t^{\mathrm{PLS}} + \psi Ctrl_t + \epsilon_{t+1}$$

where $Ctrl$ denotes one of the control variables taken from the first column other than $\mathrm{MR^{PLS}}$. The kitchen sink model is omitted for collinearity. Each variable is standardized to have a zero mean and unit variance. The sample period is from 1969Q1 to 2019Q4. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Univariate Regression			**Panel B: Bivariate Regression**				
Variable	β	t-stat	R^2(%)	β(PLS)	t-stat	ψ(Ctrl)	t-stat	R^2(%)
SPF variables								
GDP_e	−0.018	−0.56	0.27	0.096***	3.45	0.027	0.72	6.23
$\mathrm{Indprod}_e$	−0.024	−0.82	0.50	0.095***	3.34	0.023	0.64	6.08
Unemp_e	0.055**	2.28	2.53	0.077***	3.26	0.011	0.45	5.82
Recess_e	0.008	0.23	0.05	0.099***	3.67	−0.035	−0.91	6.59
Cprof_e	−0.008	−0.27	0.06	0.088***	3.70	0.018	0.54	5.98
$\mathrm{Housing}_e$	0.065***	2.58	3.54	0.071**	2.08	0.018	0.48	5.90
Infl_e	−0.026	−0.80	0.57	0.081***	3.52	−0.012	−0.39	5.87
$\mathrm{SPF7^{KS}}$	-	-	6.85	-	-	-	-	-
$\mathrm{SPF7^{PC}}$	−0.007	−0.20	0.04	0.093***	3.66	0.027	0.71	6.26
$\mathbf{MR^{PLS}}$	**0.083*** **	**3.65**	**5.75**	-	-	-	-	-
Economic variables								
DP	0.025	0.94	0.50	0.082***	3.77	0.021	0.83	6.11
DY	0.027	0.99	0.60	0.081***	3.77	0.020	0.79	6.09
EP	0.009	0.28	0.07	0.086***	3.66	0.021	0.69	6.10
DE	0.018	0.61	0.28	0.084***	3.62	−0.002	−0.08	5.75
SVAR	0.017	0.45	0.24	0.082***	3.69	0.007	0.20	5.79
BM	0.006	0.22	0.03	0.083***	3.74	0.009	0.34	5.82
NTIS	−0.023	−0.82	0.46	0.081***	3.61	−0.011	−0.40	5.84
TBL	−0.032	−1.33	0.86	0.080***	3.39	−0.009	−0.36	5.81
LTY	−0.016	−0.68	0.20	0.082***	3.71	−0.008	−0.34	5.80
LTR	0.045*	1.65	1.67	0.080***	3.67	0.039	1.51	7.04
TMS	0.044*	1.65	1.58	0.080***	3.10	0.006	0.21	5.77
DFY	0.033	0.99	0.89	0.082***	3.54	0.003	0.09	5.75
DFR	0.033	1.09	0.92	0.080***	3.50	0.017	0.57	5.97
INFL	−0.041	−1.29	1.39	0.078***	3.47	−0.026	−0.85	6.29
CAY	0.041*	1.77	1.42	0.079***	3.55	0.031	1.37	6.56
IK	−0.059**	−2.39	2.91	0.072***	2.70	−0.020	−0.69	6.00
$\mathrm{Econ^{PC}}$	−0.002	−0.08	0.00	0.084***	3.71	0.005	0.18	5.77
$\mathrm{Econ^{PLS}}$	0.085***	3.03	6.04	0.051*	1.75	0.056	1.59	7.48

Table III: **Out-of-sample Return Predictability: 1984Q1-2019Q4**

This table reports the out-of-sample forecasting performance for quarterly market excess return. The individual predictive variables include the seven current-quarter survey forecasts (SPF7) as well as a set of 16 financial and economic variables (Econ) from Welch and Goyal (2008). The terms SPF7^{FC} and Econ^{FC} refer to the equal-weighted forecast combination method based on individual forecasts generated by SPF7 and Econ, respectively. See the notes to Table II for further details on the variable definitions. We use out-of-sample R^2 statistic, whose significance is determined by the MSFE-adjusted statistics by Clark and West (2007) that tests the null hypothesis $R^2_{OS} \leq 0$ against the alternative one $R^2_{OS} > 0$, to assess the predictability of each model. We also report the results of forecast encompassing tests. The test is conducted by constructing the following optimal composite forecast,

$$\hat{R}_{t+1} = (1-\lambda)\hat{R}^i_{t+1} + \lambda\hat{R}^{\text{MR}^{\text{PLS}}}_{t+1}, \quad 0 \leq \lambda \leq 1$$

where $\hat{R}^i_{t+1}(\hat{R}^{\text{MR}^{\text{PLS}}}_{t+1})$ is the market excess return forecast generated by model i in the first and fourth columns (MR^{PLS}). The null hypothesis is $\lambda = 0$, indicating that model i encompasses MR^{PLS}, against the alternative hypothesis $\lambda > 0$ that model i does not encompass MR^{PLS}. The statistical significance of λ is assessed by the upper-tail p-value for the Harvey, Leybourne, and Newbold (1998) statistic. Panel A and B present results for SPF variables and economic variables, respectively. The OOS analysis is based on the sample period of 1984Q1 through 2019Q4. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: SPF Variables			Panel B: Economic Variables	
Variable	$R^2_{OS}(\%)$	Encompassing λ	Variable	$R^2_{OS}(\%)$	Encompassing λ
GDP_e	−0.05	0.92**	DP	−6.63	0.99***
Indprod_e	0.33	0.86**	DY	−8.00	1.00***
Recess_e	−0.83	1.00**	EP	−4.27	0.97***
Unemp_e	−0.77	1.00***	DE	−3.16	1.00***
Cprof_e	−0.19	0.90**	SVAR	−59.72	0.95
Housing_e	1.64**	0.88	BM	−6.91	1.00***
Infl_e	−0.29	0.92**	NTIS	−2.91	0.79***
			TBL	0.05	0.93**
SPF7^{KS}	−0.49**	1.00**	LTY	−1.19	0.94***
SPF7^{PC}	−0.31	0.92**	LTR	0.34	0.70**
SPF7^{FC}	0.86*	0.93**	TMS	−2.51	1.00***
$\textbf{MR}^{\textbf{PLS}}$	**3.12****	-	DFY	−4.99	1.00***
			DFR	−7.59	1.00**
			INFL	1.13*	0.75**
			CAY	−2.06	0.77***
			IK	−1.91	1.00***
			Econ^{PC}	−1.81	0.87***
			Econ^{FC}	−0.22	0.90**
			Econ^{PLS}	−12.35	1.00***

Table IV: **In-sample Return Predictability: Long horizons**

This table presents the OLS estimates, Newey-West corrected t-statistics with 2*($h-1$) lags (t-NW), and R^2 values of the predictive regression of the form

$$R_{t+1:t+h} = \alpha + \beta X_t + \epsilon_{t+1:t+h},$$

where h denotes the forecast horizon and $R_{t+1:t+h}$ is the h-quarter-ahead excess return on the CRSP value-weighted index (annualized). The terms MR^{PLS} and $\text{LT-MR}^{\text{PLS}}$ refer to the PLS factors extracted from the seven current-quarter survey forecasts and the term structures of the seven survey variables, respectively. The sample period is from 1969Q1 to 2019Q4. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)
	$h=2$	$h=4$	$h=8$	$h=12$
Predictor variable: MR^{PLS}				
β	0.07***	0.07***	0.05***	0.05***
t-NW	3.69	3.73	2.97	3.57
$R^2(\%)$	7.24	13.82	13.54	17.13
Predictor variable: $\text{LT-MR}^{\text{PLS}}$				
β	0.08***	0.07***	0.06***	0.06***
t-NW	4.23	4.68	4.29	5.67
$R^2(\%)$	8.65	14.66	20.91	27.09

Table V: **Out-of-sample Return Predictability: Long horizons**

This table reports the out-of-sample forecasting performance for multiple forecast horizons (h = 2, 4, 8, and 12 quarters). The terms $\mathrm{MR}^{\mathrm{PLS}}$ and LT-$\mathrm{MR}^{\mathrm{PLS}}$ refer to the extracted PLS factors based on the seven current-quarter survey forecasts and the term structures of the seven survey variables, respectively. See the notes to Table II for further details on the variable definitions. Panel A reports the OOS R^2 statistics whose significance is assessed by the MSFE-adjusted statistics of Clark and West (2007) that tests the null hypothesis $R^2_{OS} \leq 0$ against the alternative one $R^2_{OS} > 0$. We account for the serial correlations in overlapping observations by using 2*$(h-1)$ lags for the Newey-West regression when computing the MSFE-adjusted statistics. Panel B reports the results of OOS forecast encompassing tests. The test is conducted by constructing the following optimal composite forecast,

$$\hat{R}_{t+1} = (1-\lambda)\hat{R}^i_{t+1} + \lambda \hat{R}^{\text{LT-MR}^{\text{PLS}}}_{t+1}, \quad 0 \leq \lambda \leq 1$$

where $\hat{R}^i_{t+1}$ ($\hat{R}^{\text{LT-MR}^{\text{PLS}}}_{t+1}$) is the market excess return forecast generated by model i in the first column (LT-$\mathrm{MR}^{\mathrm{PLS}}$). The null hypothesis is $\lambda = 0$, indicating that model i encompasses LT-$\mathrm{MR}^{\mathrm{PLS}}$, against the alternative hypothesis $\lambda > 0$ that model i does not encompass LT-$\mathrm{MR}^{\mathrm{PLS}}$. The statistical significance of λ is assessed by the upper-tail p-value for the Harvey, Leybourne, and Newbold (1998) statistic. The OOS evaluation period is from 1984Q1 to 2019Q4. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

(1)	(6)	(7)	(8)	(9)	(9)	(10)	(11)	(12)
	Panel A: Out-of-sample R^2 (%)				Panel B: Encompassing λ			
Variable	$h=2$	$h=4$	$h=8$	$h=12$	$h=2$	$h=4$	$h=8$	$h=12$
SPF variables								
$\mathrm{MR}^{\mathrm{PLS}}$	2.51**	1.41*	7.67**	10.98**	0.55	0.79**	0.85**	0.91**
LT-$\mathbf{MR}^{\mathbf{PLS}}$	**2.84****	**4.73****	**12.38****	**17.91****	-	-	-	-
Economic variables								
DP	−15.44	−31.75	−47.17	−53.48	0.86***	0.93***	1.00**	1.00**
DY	−13.82	−27.05	−36.91	−44.50	0.88***	0.94***	1.00**	1.00**
EP	−11.48	−20.07	−25.88	−30.69	0.82**	0.82**	0.98**	1.00**
DE	−5.77	−4.09	−8.23	−14.07	0.81**	0.71**	0.92*	1.00*
SVAR	−92.81	−52.93	−40.91	−49.98	0.93	0.90	0.90*	0.99*
BM	−16.92	−32.54	−37.30	−46.37	0.91***	0.94***	1.00***	1.00**
NTIS	−9.65	−23.35	−11.33	−11.36	0.76***	0.77***	0.81**	0.92**
TBL	−0.44	−1.33	−6.62	−13.80	0.65**	0.66**	0.91**	1.00*
LTY	−3.12	−8.77	−14.74	−22.36	0.71**	0.74**	0.89*	0.92*
LTR	1.14**	−0.28	1.28***	0.61**	0.55*	0.61**	0.77*	0.87*
TMS	−1.14	4.91***	9.77***	14.63***	0.68**	0.49**	0.60*	0.60*
DFY	−5.91	−5.51	−5.63	−13.05	0.82***	0.75***	0.95**	1.00**
DFR	−2.56	−2.37	−3.03	−0.53	0.72**	0.70**	0.85**	0.91*
INFL	1.16	−1.35	−6.32	−7.59	0.58*	0.66**	0.96**	1.00**
CAY	−2.85	−2.71*	5.99*	−0.96*	0.64**	0.60**	0.58**	0.66*
IK	−0.59	3.91*	4.90	16.28***	0.69**	0.53**	0.80**	0.57*
$\mathrm{Econ}^{\mathrm{PC}}$	−2.55	−7.89	−15.05	−26.32	0.67**	0.72**	0.94**	1.00**
$\mathrm{Econ}^{\mathrm{FC}}$	0.20	−0.82	−3.06	−5.41	0.62*	0.65**	0.88**	0.98*
$\mathrm{Econ}^{\mathrm{PLS}}$	−19.25	−21.11	−3.30	−6.10	0.93**	0.94**	0.88**	1.00**

Table VI: **Return Predictability for Subsamples**

This table presents both in- and out-of-sample results of predictive regression of the form

$$R_{t+1:t+h} = \alpha + \beta X_t + \epsilon_{t+1:t+h},$$

where h denotes the forecast horizon and $R_{t+1:t+h}$ is the h-quarter-ahead excess return on the CRSP value-weighted index (annualized). We use $\mathrm{MR}^{\mathrm{PLS}}$ as predictor variable (X_t) for the quarterly horizon ($h = 1$) and use $\mathrm{LT\text{-}MR}^{\mathrm{PLS}}$ for longer horizons ($h = 2, 4, 8, 12$). For each regression, we report the slope estimate, Newey-West corrected t-statistic with 2*($h - 1$) lags (t-NW), in-sample R^2, and out-of-sample R^2 whose significance is assessed by the MSFE-adjusted statistics of Clark and West (2007). Subsamples cover periods from 1980Q1 to 2019Q4 (Panel A), 1990Q1 to 2019Q4 (Panel B), and 2000Q1 to 2019Q4 (Panel C). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)	(6)
	$h = 1$	$h = 2$	$h = 4$	$h = 8$	$h = 12$
			Panel A: Post-1980 Period		
β	0.08***	0.08***	0.06***	0.04**	0.04***
t-NW	3.17	4.04	3.54	2.32	3.87
$R^2(\%)$	5.46	10.54	11.62	9.52	15.40
$R^2_{OS}(\%)$	3.38**	3.23**	1.97**	4.04**	20.32***
			Panel B: Post-1990 Period		
β	0.08***	0.06***	0.05***	0.05***	0.05***
t-NW	2.85	2.58	2.66	3.05	3.42
$R^2(\%)$	6.40	6.79	8.29	15.53	17.13
$R^2_{OS}(\%)$	4.29**	2.53**	4.85**	13.67**	19.74**
			Panel C: Post-2000 Period		
β	0.13***	0.09***	0.08***	0.08***	0.09***
t-NW	3.53	2.58	2.78	4.28	7.88
$R^2(\%)$	14.40	14.75	18.56	37.40	58.04
$R^2_{OS}(\%)$	5.69**	3.68**	7.20**	19.99**	36.71***

Table VII: **International Evidence**

This table reports the OLS estimates, Newey-West t-statistics, and R^2 of in-sample predictive regression

$$R^i_{t+1} = \alpha + \beta \, \mathrm{EMR}^{\mathrm{PLS}}_{i,t} + \epsilon^i_{t+1},$$

where R^i_{t+1} is the excess return (annualized) on one of the seven European countries considered, including France, Germany, Italy, the Netherlands, Sweden, Switzerland, the United Kingdom (UK), or on the STOXX Europe 600 Index. We collect survey forecasts on the expected rates of inflation, real GDP growth, and unemployment in the euro area at horizons of the current year and the next year from the European Central Bank Survey of Professional Forecasters (ECB SPF), and $\mathrm{EMR}^{\mathrm{PLS}}_{i,t}$ is the extracted PLS macro risk index from the above-mentioned survey forecasts using R^i as the target variable. The sample period is from 1999Q1 to 2019Q4. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	STOXX	France	Germany	Italy	Netherlands	Sweden	Switzerland	UK
β	0.14***	0.13***	0.13***	0.13***	0.15***	0.14***	0.09***	0.10***
t-stat	3.14	3.27	2.88	2.69	3.22	3.41	2.89	3.39
$R^2(\%)$	11.03	11.33	7.69	9.90	13.93	10.67	9.62	12.42

Table VIII: **Economic Value of Out-of-sample Return Predictability**

This table reports the out-of-sample performance of quarterly asset allocation practice. The investor optimally allocates a portion $\omega_t = \hat{R}_{t+1}/(\gamma\hat{\sigma}^2_{t+1})$ of her wealth to the market index and $1-\omega_t$ to the risk-free asset, where γ is the risk aversion coefficient, $\hat{R}_t$ is the OOS forecast of $t+1$ market index excess return made at time t using the models listed in the first column, and $\hat{\sigma}^2_{t+1}$ is the forecast of $t+1$ market return variance based on calculated as the sample variance of the market excess returns over past ten years. The weight on the market index is constrained to lie between zero and 1.5. The term HAV corresponds to the strategy using the historical average excess return forecast. The term Buy&Hold refers to the passive strategy that holds the market index. See the notes to Table III for further details on the variable definitions. The Sharpe ratio (annualized) is computed as the average portfolio excess return divided by the standard deviation of returns. We apply the Jobson and Korkie (1981) statistic corrected by Memmel (2003) to assess whether the difference between the Sharpe ratio of HAV and the Sharpe ratio of any strategy other than HAV in column one is significant. The certain equivalent return (CER) gain (annualized and in precent) is defined as the difference between the CER delivered by HAV and the CER delivered by any strategy other than HAV in column one, and its statistical significance is determined by the Diebold and Mariano (1995) statistic. The term $\triangle\Theta(\%)$ (annualized and in percent) denotes the difference between the manipulation-proof measure (MPPM) (Goetzmann, Ingersoll, Spiegel, and Welch, 2007) of HAV and the MPPM of any strategy other than HAV in column one. Panels A and B report the portfolio performance under risk aversion level three and five, respectively. Each strategy is quarterly rebalanced. The OOS evaluation period is from 1984Q1 to 2019Q4. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: $\gamma = 3$, $\rho = 3$			**Panel B: $\gamma = 5$, $\rho = 5$**		
Variable	**Sharpe**	**CER gain (%)**	**$\Delta\Theta$ (%)**	**Sharpe**	**CER gain (%)**	**$\Delta\Theta$ (%)**
HAV	0.41	-	-	0.40	-	-
Buy & Hold	0.52*	1.82*	1.57	0.52*	0.24	−0.54
SPF variables						
SPF7^{KS}	0.52	1.72	1.25	0.54	1.34	1.06
SPF7^{PC}	0.39	−0.22	−0.22	0.39	−0.13	−0.13
SPF7^{FC}	0.48**	0.78**	0.76	0.48**	0.48**	0.47
MR^{PLS}	**0.64****	**3.24****	**3.17**	**0.61***	**1.95****	**1.89**
Economic variables						
DP	0.33	−1.40	−1.36	0.33	−0.83	−0.81
DY	0.40	−1.17	−1.13	0.40	−0.69	−0.67
EP	0.48	−0.21	−0.20	0.48	−0.11	−0.11
DE	0.32	−1.08	−1.17	0.32	−0.69	−0.74
SVAR	0.22	−2.91	−3.29	0.20	−2.79	−3.28
BM	0.25	−1.65	−1.65	0.25	−0.97	−0.98
NTIS	0.37	−0.84	−1.24	0.37	−0.95	−0.92
TBL	0.43	0.29	0.10	0.41	0.00	−0.11
LTY	0.37	−0.44	−0.47	0.37	−0.24	−0.25
LTR	0.48	1.03	1.07	0.41	−0.08	0.04
TMS	0.41	−0.17	−0.81	0.34	−1.17	−2.27
DFY	0.29	−1.36	−1.43	0.26	−1.03	−1.09
DFR	0.32	−1.18	−1.33	0.29	−1.07	−1.21
INFL	0.47	0.93	0.96	0.45	0.38	0.38
CAY	0.39	−0.68	−0.93	0.37	−1.41	−1.99
IK	0.44	0.61	0.07	0.44	0.38	0.00
Econ^{PC}	0.37	−0.66	−0.63	0.37	−0.39	−0.37
Econ^{FC}	0.41	−0.06	−0.07	0.42	0.02	0.02
Econ^{PLS}	0.34	−1.83	−2.70	0.33	−2.11	−2.46

Table IX: **Predicting Macroeconomic Activities with SPF Data**

This table reports the in-sample predictive regression estimation results of several real macroeconomic activity measures using the SPF data. The macroeconomic variables include the real consumption per capita, real GDP, real labor income per capita and industrial production, all of which are seasonally adjusted and continuously compounded growth rates. Panel A presents the results for univariate predictive regression setting,

$$y^i_{t+1:t+h} = \alpha + \beta(\text{SPF variables})^{\text{PLS,Macro}}_{i,t} + \epsilon^i_{t+1:t+h},$$

where $y^i_{t+1:t+h}$ is the real macroeconomic activity measure from quarter $t+1$ to quarter $t+h$ and (SPF variables)$^{\text{PLS,Macro}}_{i,t}$ is the PLS factor extracted from the set of SPF variables using y_i as the target variable. We use the seven current-quarter survey forecasts (SPF7) as the PLS regressors at the quarterly horizon ($h=1$), and use the term structures of the seven survey variables (SPF7TS) at longer forecast horizons ($h=4,8$). Panel B presents the results for multivariate regression setting,

$$y^i_{t+1:t+h} = \alpha + \beta(\text{SPF variables})^{\text{PLS,Macro}}_{i,t} + \gamma'\text{Ctrl}_t + \rho y^i_{t-h+1:t} + \epsilon^i_{t+1:t+h},$$

where Ctrl denotes several control variables, including the term spread (TMS) and short-term T-bill yield (TBL), and $y^i_{t-h+1:t}$ is the lagged observation of y_i. At the quarterly horizon, we report the Newey and West (1987) corrected t-statistics (t-NW). At longer forecast horizons, we report the t-statistics that are adjusted by Hodrick (1992) standard errors (t-Hodrick). The sample period is from 1969Q1 to 2019Q4. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Univariate				**Panel B: Multivariate**			
	Real Consumption	Real GDP	Real Labor Income	Industrial Production	Real Consumption	Real GDP	Real Labor Income	Industrial Production
One quarter ($h=1$), variable set SPF7								
β(%)	0.13***	0.29***	0.22***	0.73***	0.07*	0.29***	0.27***	0.12
t-NW	3.37	4.39	3.54	4.62	1.70	3.24	3.23	0.60
R^2(%)	10.36	13.44	9.54	23.21	28.40	17.77	11.51	38.53
Four quarter ($h=4$), variable set SPF7TS								
β(%)	0.59***	1.05***	0.80***	2.28***	0.22	0.98***	0.74***	2.32***
t-Hodrick	3.92	3.77	3.65	3.75	1.32	3.02	3.18	3.55
R^2(%)	12.90	21.15	18.45	23.49	35.37	35.28	19.83	33.19
Eight quarter ($h=8$), variable set SPF7TS								
β(%)	0.59***	1.10**	0.75**	2.20***	0.54**	1.08**	0.79**	2.03**
t-Hodrick	2.61	2.37	2.13	2.65	2.40	2.53	2.33	2.53
R^2(%)	5.92	10.59	6.92	10.18	29.48	36.80	8.75	33.19

Table X: **Relation of Forecasting Performance with Economic Conditions**

Panel A reports the contemporaneous correlations between the out-of-sample return forecasting performance and several macroeconomic condition measures. The term $E_t[R^e_{t+1}]$ denotes the OOS quarter $t+1$ market excess return forecast made at quarter t by $\mathrm{MR}^{\mathrm{PLS}}$. The term DSFE_t refers to the difference between the squared forecast error of $\mathrm{MR}^{\mathrm{PLS}}$ and the squared forecast error of the historical mean benchmark at quarter t. The term CERG_t refers to the gain in the certain equivalent return (CER) produced by the market timing strategy employing return forecasts by $\mathrm{MR}^{\mathrm{PLS}}$ relative to the strategy employing historical mean forecast at quarter t. The macroeconomic condition measures (X_t) include the real GDP growth, real consumption per capita growth, real labor income per capita growth, the Chicago Fed National Activity Index (CFNAI), and the macroeconomic uncertainty index by Jurado, Ludvigson, and Ng (2015). Panel B reports the R^2_{OS} statistics over subsamples. We use the NBER-dated business-cycle phase and the real GDP growth to individually classify good and bad times of the overall economy. We compute the subsample R^2_{OS} statistic as

$$R^2_{OS,c} = 1 - \frac{\sum_{t=1}^{T} I^s_t (R_t - \hat{R}_t)^2}{\sum_{t=1}^{T} I^s_t (R_t - \bar{R}_t)^2}, \quad c = \text{Good, Bad},$$

where I^{Good}_t (I^{Bad}_t) is set equal to one whenever the business condition is in expansion (recession) or the real GDP growth is above (below) the bottom third of sorted observations in quarter t, and zero otherwise. $\bar{R}_t$ refers to the historical mean benchmark forecast and $\hat{R}_t$ refers to the market excess return forecast generated by $\mathrm{MR}^{\mathrm{PLS}}$. The OOS evaluation period is from 1984Q1 to 2019Q4. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Correlations with macroeconomic variables (model: $\mathrm{MR}^{\mathrm{PLS}}$)					
	Real GDP	Real Consumption	Real Labor Income	CFNAI	Macro Uncertainty
$\rho(E_t[R_{t+1}], X_t)$	−0.25***	−0.35***	−0.15*	−0.22***	0.05
$\rho(\mathrm{DSFE}_t, X_t)$	−0.14*	−0.17**	−0.19**	−0.18**	0.17**
$\rho(\mathrm{CERG}_t, X_t)$	−0.15*	−0.16*	−0.12	−0.23***	0.17**
Panel B: Out-of-sample R^2(%) over subsamples					
	NBER Recession	NBER Expansion		Real GDP Bad	Real GDP Good
$\mathrm{MR}^{\mathrm{PLS}}$	7.29***	2.12**		7.23**	0.03

Table XI: **Predicting Characteristics Portfolios**

This table presents the forecasting results for the characteristics portfolios with $\mathrm{MR}^{\mathrm{PLS}}$. We evaluate the following predictive regression both in- and out-of-sample:

$$R^p_{t+1} = \alpha + \beta\,\mathrm{MR}^{\mathrm{PLS}}_t + \epsilon_{t+1},$$

where R^p_{t+1} is the quarterly excess returns on the portfolios sorted on size or industry category. We display the in-sample regression slopes and R^2s and the OOS R^2s. Panels A and B report the results for the 10 size portfolios and the 10 industry portfolios, receptively. The in-sample period is from 1969Q1 to 2019Q4 and the OOS evaluation period is from 1984Q1 to 2019Q4. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Size portfolios				Panel B: Industry portfolios		
	β	R^2 (%)	R^2_{OS} (%)		β	R^2 (%)	R^2_{OS} (%)
Small	0.133***	6.41	2.83***	Nondurable	0.080***	5.47	2.47**
Size2	0.115***	5.19	2.17**	Durable	0.146***	8.94	6.13***
Size3	0.109***	5.29	2.70**	Manufacture	0.086***	5.29	4.20**
Size4	0.107***	5.51	2.73**	Energy	0.044	1.33	−1.23
Size5	0.106***	5.69	3.31**	HiTech	0.111***	5.08	1.54**
Size6	0.104***	6.34	4.66***	Telecom	0.054**	2.40	2.18**
Size7	0.096***	5.38	3.18**	Shops	0.111***	7.07	3.43***
Size8	0.093***	5.56	3.96***	Health	0.050**	1.91	0.59
Size9	0.081***	5.04	3.04**	Utility	0.047**	2.41	−0.50
Large	0.073***	5.21	2.51**	Other	0.086***	4.39	2.87**

Table XII: **Predictive Regression Results for Market Return Components**

This table reports the predictive regression results for market return components that are estimated using the Campbell (1991) and Campbell and Ammer (1993) vector autoregression (VAR) approach. The three estimated components of the CRSP log return are the expected return, cash flow news, and discount rate news. The in-sample predictive regression model follows

$$y_{t+1} = \alpha_y + \beta_y \, \mathrm{MR}_t^{\mathrm{PLS}} + \epsilon_{t+1},$$

where y_t is one of three estimated return components for quarter t and $\mathrm{MR}_t^{\mathrm{PLS}}$ denotes the extracted factors based on the seven current-quarter survey forecasts via PLS. The intercept of the above regression model is set equal to zero when we predict the cash flow news and the discount rate news. We use a VAR approach based on the variables in the first column to measure the return components, where "r" denotes the CRSP log return. The description to the variables in Panel A can be referred to Appendix A. The term "PC3" denotes the first three principal components extracted from the 16 financial and economic variables of Welch and Goyal (2008). The variable "CC" in Panel B denotes the cyclical consumption of Atanasov, Møller, and Priestley (2020). The sample period is from 1969Q1 to 2019Q4. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Expected return		Cash flow news		Discount rate news	
VAR Variables	$\hat{\beta}_{\hat{E}}$	t-stat	$\hat{\beta}_{\mathrm{CF}}$	t-stat	$\hat{\beta}_{\mathrm{DR}}$	t-stat
Panel A: Economic Predictors from Welch and Goyal (2008)						
r, DP	0.167	1.43	0.381	1.62	−1.188***	−3.16
r, DP, DY	0.108	0.80	0.380	1.61	−1.247***	−3.26
r, DP, EP	0.234*	1.94	0.603	1.64	−0.898**	−2.40
r, DP, DE	0.196*	1.67	0.624	1.57	−0.916**	−2.32
r, DP, SVAR	0.271**	2.25	0.394*	1.68	−1.071***	−2.99
r, DP, BM	0.378***	3.02	0.369	1.55	−0.989***	−2.83
r, DP, NTIS	0.233**	1.97	0.240	0.95	−1.262***	−3.15
r, DP, TBL	0.628***	5.02	0.259	0.91	−0.849**	−2.11
r, DP, LTY	0.339***	2.87	0.427	1.63	−0.969**	−2.52
r, DP, LTR	0.226	1.61	0.383	1.64	−1.127***	−2.91
r, DP, TMS	0.525***	4.66	0.298	1.21	−0.913**	−2.40
r, DP, DFY	0.301**	2.37	0.410*	1.74	−1.024***	−2.83
r, DP, DFR	0.287**	2.03	0.359	1.52	−1.090***	−3.15
r, DP, INFL	0.560***	3.26	0.390	1.64	−0.787**	−2.15
r, DP, CAY	0.315**	2.17	0.199	0.87	−1.222***	−2.82
r, DP, IK	0.628***	5.74	0.463**	1.97	−0.645*	−1.66
r, DP, PC3	0.782***	4.64	0.312	1.29	−0.642*	−1.90
Panel B: Including the Cyclical Consumption						
r, DP, PC3, CC	1.106***	6.31	−0.023	−0.10	−0.653	−1.46

Table XIII: **Relation with other predictive variables**

This table presents the results of in-sample predictive regressions for quarterly market excess return based on MR^{PLS} and additional predictive variables considered: the output gap (OG), as in Cooper and Priestley (2009); the cyclical consumption (CC), as in Atanasov, Møller, and Priestley (2020); the share of labor income to consumption (s^w), as in Santos and Veronesi (2006); the consumption volatility (σ_c), as in Bansal, Khatchatrian, and Yaron (2005); the price-output ratio (py), as in Rangvid (2006); the quarterly non-housing consumption to total consumption ratio (house) following the construction of Piazzesi, Schneider, and Tuzel (2007); the PLS investor sentiment ($\text{senti}^{\text{HJTZ}}$), as in Huang, Jiang, Tu, and Zhou (2015); the PC investor sentiment (senti^{BW}), as in Baker and Wurgler (2006); the Michigan Consumer Sentiment Index (senti^{MC}); the PLS disagreement index ($\text{disag}^{\text{HLW}}$), as in Huang, Li, and Wang (2021). Panel A and B report the results for univariate and bivariate regressions, respectively. See the notes to Table II for specifications of regression models. Panel C reports the contemporaneous correlations between MR^{PLS} and the control variables. Each variable is standardized to have a zero mean and unit variance. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. The sample period is from 1969Q1 to 2019Q4, except for Senti^{BW} (1969Q1 to 2018Q4) and $\text{Disag}^{\text{HLW}}$ (1970Q1 to 2018Q4).

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Univariate regression		Panel B: Bivariate regression			Panel C: correlation
Variable	β	$R^2(\%)$	β (MR^{PLS})	ψ (Ctrl)	$R^2(\%)$	$\rho(\text{MR}^{\text{PLS}}, \text{Ctrl})$
MR^{PLS}	0.083***	5.75	-	-	-	-
OG	−0.077***	4.92	0.058**	−0.043	6.78	−0.58
CC	−0.069***	3.98	0.065***	−0.035	6.51	−0.52
s^w	−0.016	0.22	0.083***	−0.012	5.88	−0.04
σ_c	−0.034	0.96	0.088***	−0.043*	7.29	0.11
py	−0.037	1.13	0.080***	−0.027	6.33	−0.13
house	−0.037	1.14	0.079***	−0.015	5.92	−0.28
$\text{senti}^{\text{HJTZ}}$	−0.088***	6.49	0.073***	−0.079***	10.79	−0.13
senti^{BW}	−0.036	1.07	0.083***	−0.034	6.80	−0.02
senti^{MC}	−0.024	0.50	0.087***	0.010	5.82	−0.40
$\text{disag}^{\text{HLW}}$	−0.077***	4.80	0.068***	−0.062**	8.45	−0.21

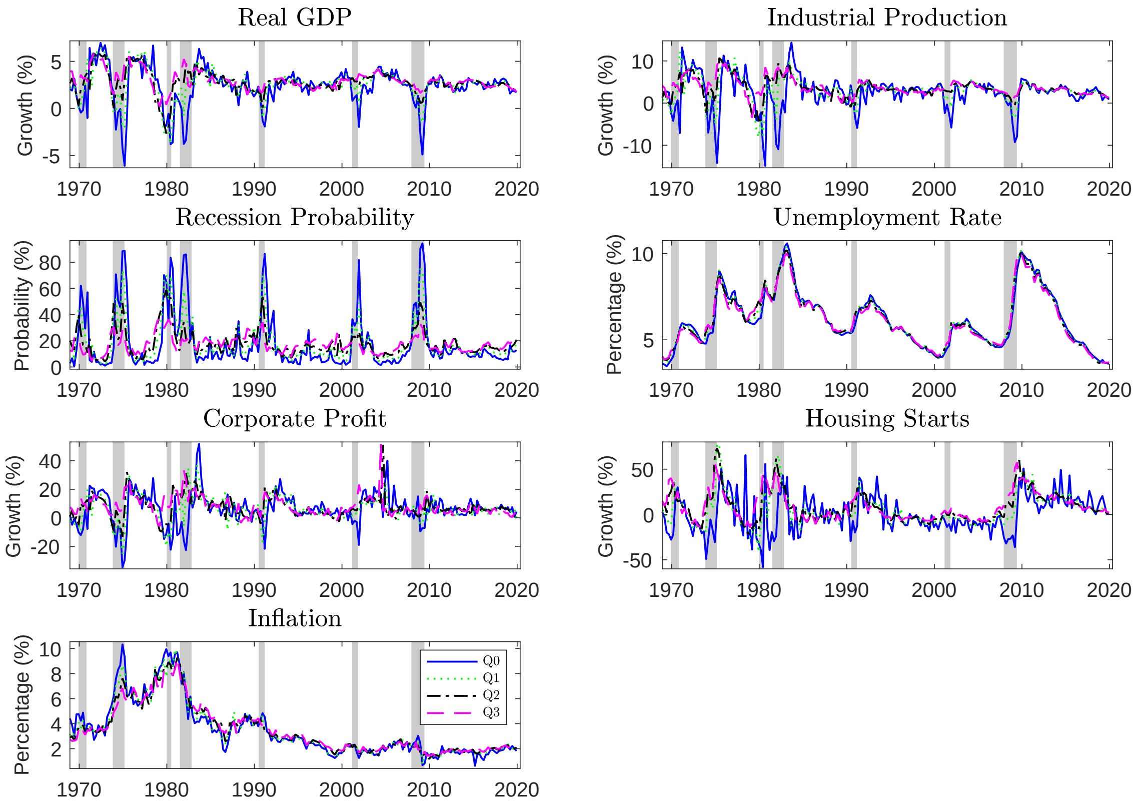

Figure 1: **Consensus Macroeconomic Forecasts from the Survey of Professional Forecasters**

Figure 1 plots the term structure of consensus macroeconomic forecasts for the seven aspects of the macroeconomy, including the Real GDP Growth, Industrial Production Growth, Recession Probability, Unemployment Rate, Corporate Profit Growth, Housing Starts Growth, and Inflation. The term structure of each survey variable is consist of the forecasts of the current quarter and the following three quarters. The solid line depicts the forecast for current quarter. The dotted line depicts one-quarter ahead forecast. The dash-dotted line depicts two-quarter ahead forecast and the dashed line depicts three-quarter ahead forecast. The survey data are from the Survey of Professional Forecasters database. The sample period is from the fourth quarter of 1968 (1968Q4) to the fourth quarter of 2019 (2019Q4). The shaded area corresponds to the NBER recession period.

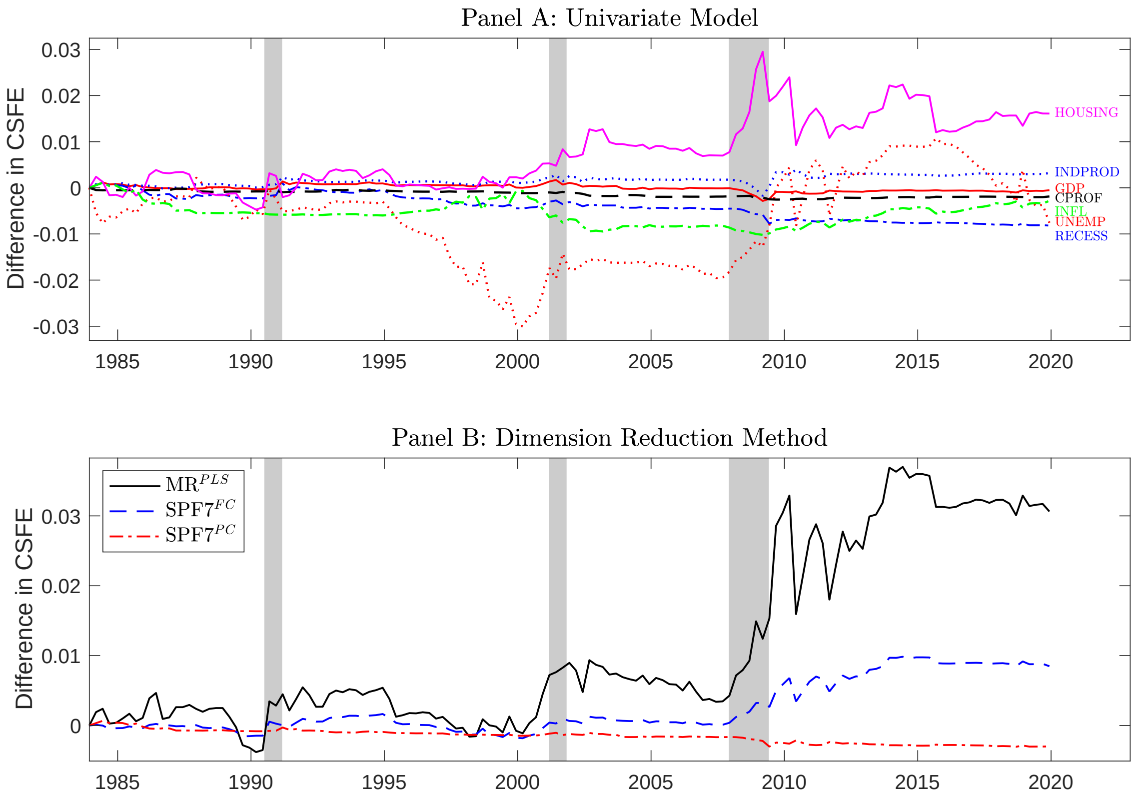

Figure 2: **The differences in cumulative squared forecast errors (CSFE) over quarterly horizon**

Figure 2 depicts the difference between the CSFE for the predictive regression models and the CSFE for the recursive historical mean over the quarterly forecast horizon. Panel A shows the results for the seven univariate regression models based on individual survey variables. Panel B shows the results for the three dimension reduction methods considered in our analysis: the PLS regression ($\mathrm{MR}^{\mathrm{PLS}}$), the combination forecast method ($\mathrm{SPF7}^{\mathrm{FC}}$), and the principal component regression ($\mathrm{SPF7}^{\mathrm{PC}}$). The shaded area corresponds to the NBER recession period. The out-of-sample period is from the fourth quarter of 1984 (1984Q1) to the fourth quarter of 2019 (2019Q4).

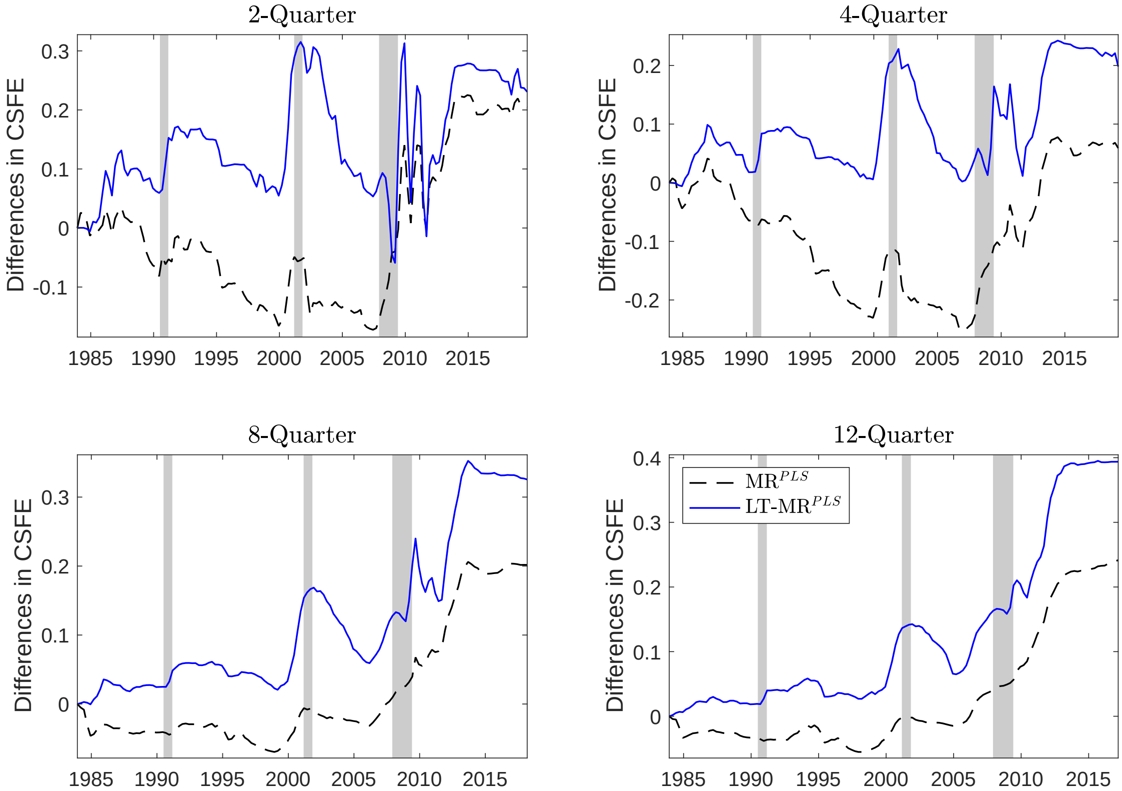

Figure 3: **The differences in cumulative squared forecast errors (CSFE) over long horizons: MR$^{\mathbf{PLS}}$ vs LT-MR$^{\mathbf{PLS}}$**

The dashed line depicts the difference between the CSFE for the PLS regression using the seven current-quarter survey forecasts (MR$^{\text{PLS}}$) and the CSFE for the recursive historical mean over four forecast horizons (two-, four-, eight- and 12-quarter ahead). The solid line depicts the difference between the CSFE for the PLS regression using the term structures of the seven survey variables (LT-MR$^{\text{PLS}}$) and the CSFE for the recursive historical mean. The shaded area corresponds to the NBER recession period. The out-of-sample period is from the fourth quarter of 1984 (1984Q1) to the fourth quarter of 2019 (2019Q4).

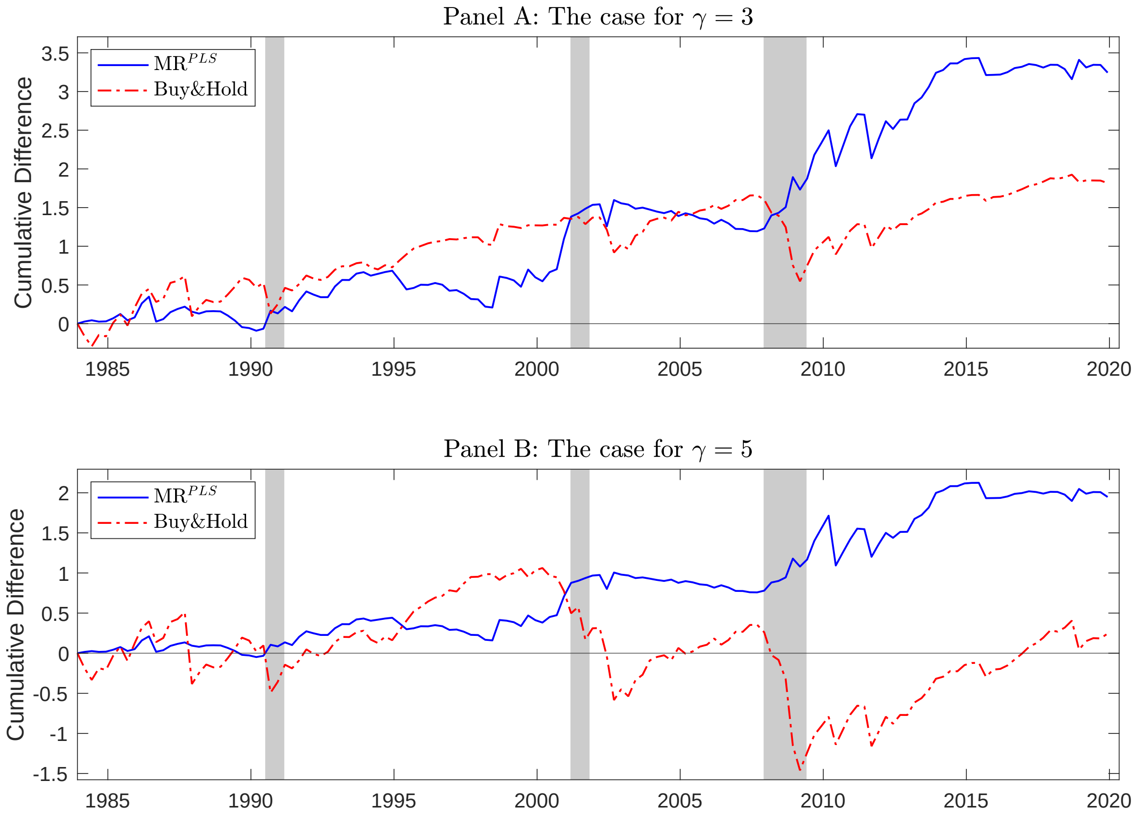

Figure 4: **The differences in cumulative certain equivalent return (CER)**
The solid line depicts the difference between the cumulative CER for the active investment strategy based on the forecasting model MR$^{\text{PLS}}$ and the cumulative CER for the passive strategy based on the recursive historical mean forecast. The dash-dotted line depicts the difference between the cumulative CER for the simple buy-and-hold investment strategy and the cumulative CER for the passive strategy based on the recursive historical mean forecast. Panel A and B report the results for a mean-variance investor with relative risk aversion coefficients of three and five, respectively. The shaded area corresponds to the NBER recession period. The out-of-sample period is from the fourth quarter of 1984 (1984Q1) to the fourth quarter of 2019 (2019Q4).

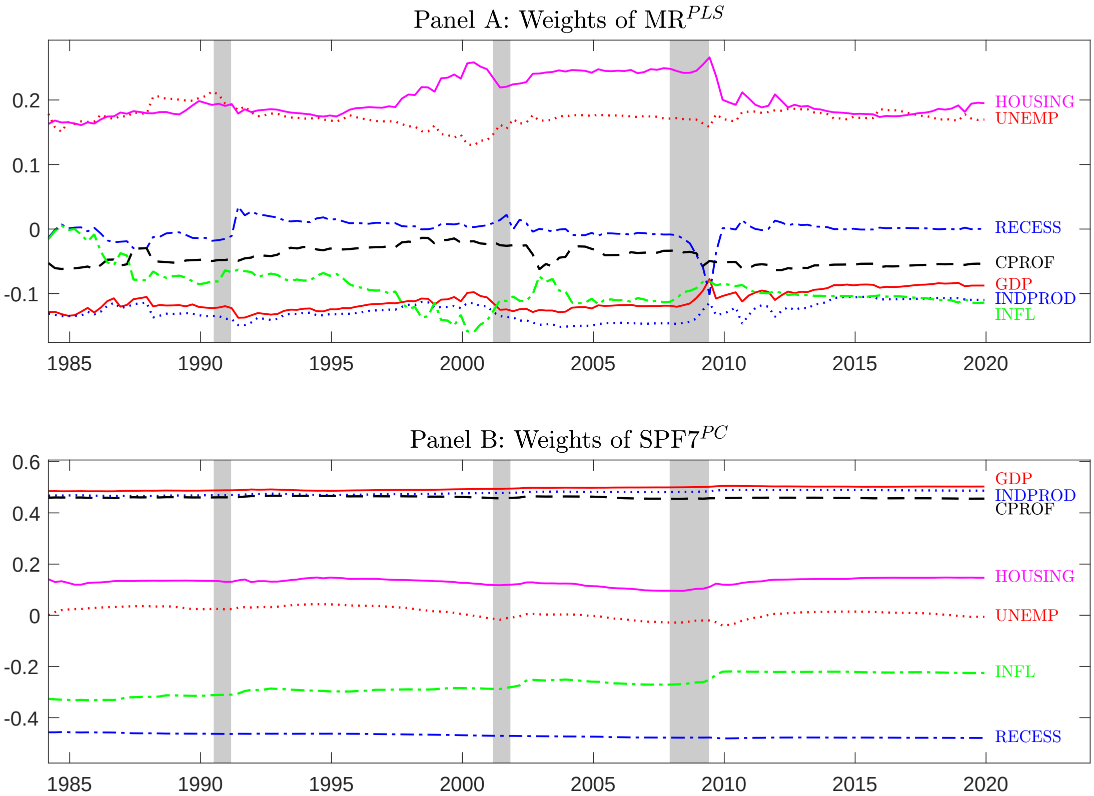

Figure 5: **Weights of $\mathbf{MR^{PLS}}$ on individual survey variables**

Figure 5 depicts the weights of the aligned macro risk index constructed by PLS ($\mathrm{MR^{PLS}}$) using the seven current-quarter survey forecasts over the out-of-sample period, which spans from the fourth quarter of 1984 (1984Q1) to the fourth quarter of 2018 (2018Q4). The seven macroeconomic survey variables include current quarter forecasts on the real GDP growth, industrial production growth, recession probability, unemployment rate, corporate profit growth, housing starts growth, and inflation. The shaded area corresponds to the NBER-dated recession period.

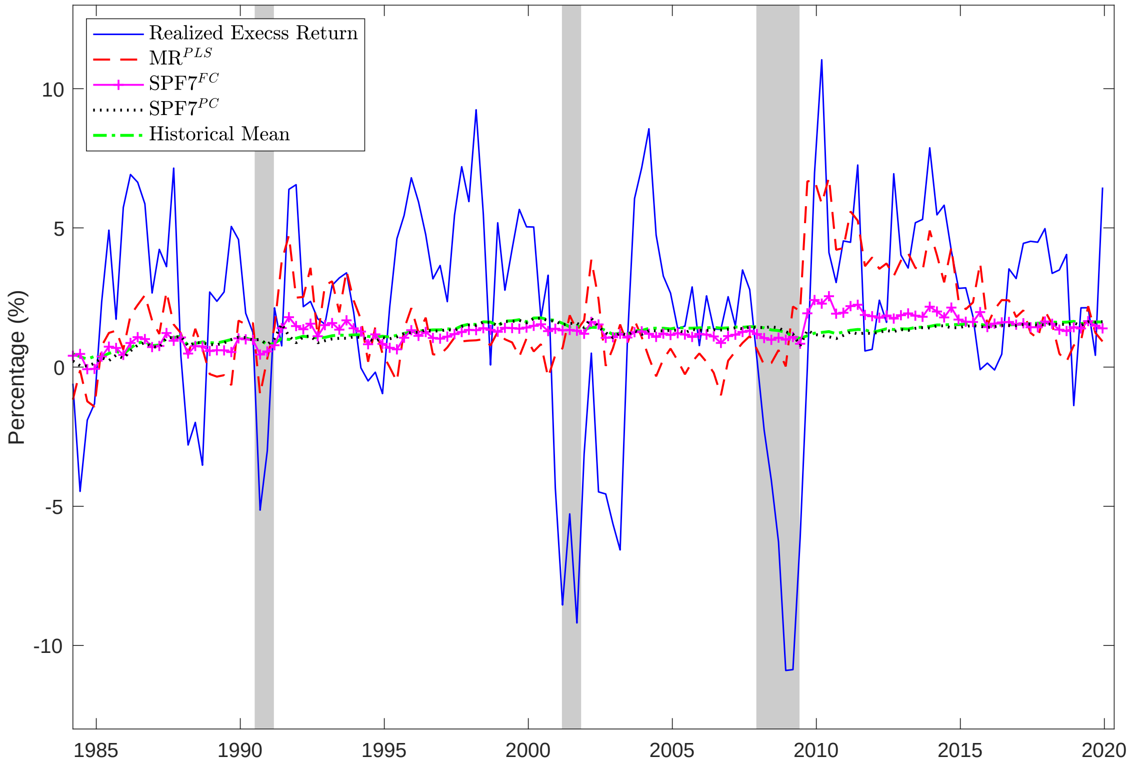

Figure 6: **Out-of-sample Market Risk Premium Forecasts**

Figure 6 depicts the quarterly realized market excess returns, the recursive historical mean returns, and the market excess return forecasts based on the seven current-quarter survey forecasts via the three dimension reduction methods we considered, including the PLS regression (MR^{PLS}), the forecasts combination method (SPF7^{FC}), and the principal component regression (SPF7^{PC}). The dashed green line plots the recursive historical average excess return. The solid blue line refers to the realized excess market return (smoothed over past four quarters). The sample period is from the fourth quarter of 1984 (1984Q1) to the fourth quarter of 2019 (2019Q4). The shaded area corresponds to the NBER recession period.

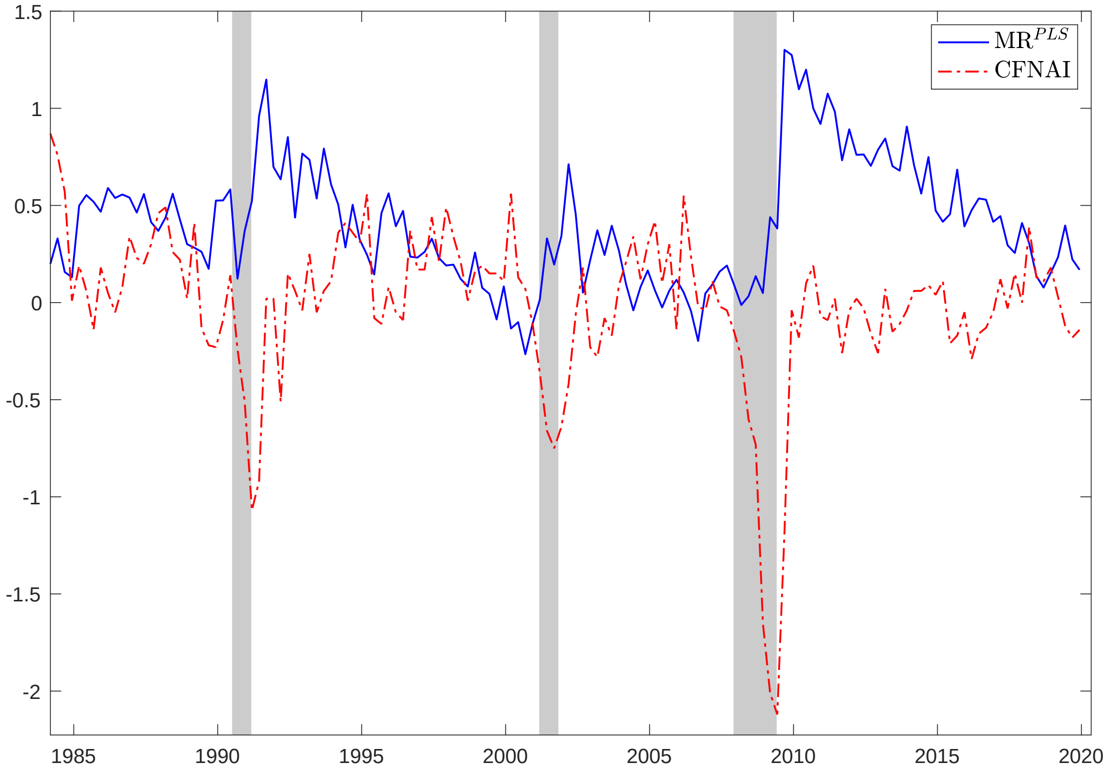

Figure 7: **Time series of MR$^{\mathbf{PLS}}$ and the CFNAI**

The solid line depicts the time series of the aligned macro risk index constructed by PLS using the seven current-quarter survey forecasts (MR$^{\text{PLS}}$) over the out-of-sample period. The dash-dotted line depicts the Chicago Fed National Activity Index. The sample period is from the fourth quarter of 1984 (1984Q1) to the fourth quarter of 2019 (2019Q4). The shaded area corresponds to the NBER-dated recession period.

Appendices

The appendix presents the definition of the 16 financial and economic variables from Goyal and Welch (2008) used in our empirical analyses and a detailed description on return decomposition methodology.

A Variable Definitions

- Dividend Price Ratio (DP): Difference between the log of 1-year moving sum of dividends paid on the S&P 500 index and the log of the S&P 500 index level.
- Dividend Yield (DY): Difference between the logarithm of 1-year moving sum of dividends paid on the S&P 500 index and the log of lagged S&P 500 index level.
- Earnings Price Ratio (EP): The log of earnings minus the log of the S&P 500 index level. Earnings are 12-month moving sums of earnings on the S&P 500 index.
- Dividend Payout Ratio (DE): Difference between the log of dividends and the log of the earnings of the S&P 500 index.
- Stock variance (SVAR): Sum of squared daily S&P 500 index returns.
- Book-to-Market Ratio (BM): The ratio of book equity value to market equity value for the Dow Jones Industrial Average.
- Net Equity Expansion (NTIS): The ratio of 12-month moving sums of net issues by NYSE-listed stocks divided by the total end-of-year market capitalization of NYSE stocks.
- Treasury Bill Rate (TBL): Yield on a 3-month Treasury bill traded in the secondary market.
- Long Term Yield (LTY): Long-term government bond yield.
- Long Term Rate of Returns (LTR): Return on long-term government bonds.
- Term Spread (TMS): Difference in yield between the long-term government bonds and the 3-month Treasury bill.
- Default Yield Spread (DFY): Difference between BAA and AAA-rated corporate bond yields.
- Default Return Spread (DFR): Difference in return between the long-term corporate bonds and long-term government bonds.
- Inflation (INFL): Inflation is the growth rate of Consumer Price Index (All Urban Consumers). Since the inflation data is released in the next month, we use the lagged inflation, following Welch and Goyal (2008).
- Consumption to Wealth Ratio (CAY): The residual from a co-integration regression of the aggregate consumption on aggregate wealth and labor income (Lettau and Ludvigson, 2001).
- Investment to Capital Ratio (IK): The ratio of aggregate (private nonresidential fixed) investment to aggregate capital for the whole economy (Cochrane, 1991).

B Return Decomposition

Denote by P_t and D_t the stock price and the dividend at time t, respectively. We define the log dividend-price ratio as $x_t = log(D_t/P_t) = log(D_t) - log(P_t) = d_t - p_t$. According to Campbell and Shiller (1988), the log-linear approximation of the stock return is given by

$$r_{t+1} = log\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right) \approx k + x_t + \Delta d_{t+1} - \rho x_{t+1}, \tag{B.1}$$

where

$$\rho = \frac{1}{1+e^{\bar{x}}} \in (0,1), \tag{B.2}$$

$$k = -\rho log(\rho) - (1-\rho) log(1-\rho), \tag{B.3}$$

$\bar{x}$ is the mean of x_t, and $\Delta d_{t+1} = d_{t+1} - d_t$. We can rewrite Eq. (B.1) as

$$\begin{aligned} x_t &\approx r_{t+1} - k - \Delta d_{t+1} + \rho x_{t+1} \\ &= r_{t+1} - k - \Delta d_{t+1} + \rho(r_{t+2} - k - \Delta d_{t+2} + \rho x_{t+2}) = ... \\ &= -\frac{k}{1-\rho} - \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} + \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, \end{aligned} \tag{B.4}$$

where in the last step, we impose the no-bubble transversality condition $\lim_{j\to\infty} \rho^j x_{t+j} = 0$. Taking time-$t$ conditional expectation on both sides of Eq. (B.4) yields the dividend-price ratio decomposition of Campbell and Shiller (1988),

$$x_t = -\frac{k}{1-\rho} - \mathbb{E}_t\left(\sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}\right) + \mathbb{E}_t\left(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}\right). \tag{B.5}$$

Using the results from Eqs. (B.1) and (B.5), we obtain the following decomposition of the log stock return innovation:

$$r_{t+1} - \mathbb{E}_t(r_{t+1}) = (\mathbb{E}_{t+1} - \mathbb{E}_t)\left(\sum_{j=0}^{\infty} \rho^j \Delta d_{t+j+1}\right) - (\mathbb{E}_{t+1} - \mathbb{E}_t)\left(\sum_{j=1}^{\infty} \rho^j r_{t+j+1}\right). \tag{B.6}$$

Equation (B.6) indicates that the unexpected log stock return can be decomposed into cash flow news and discount rate news components:

$$\eta^r_{t+1} = \eta^{\mathrm{CF}}_{t+1} - \eta^{\mathrm{DR}}_{t+1}, \tag{B.7}$$

where $\eta^r_{t+1} = r_{t+1} - \mathbb{E}_t(r_{t+1})$, $\eta^{\mathrm{CF}}_{t+1} = (\mathbb{E}_{t+1} - \mathbb{E}_t)\left(\sum_{j=0}^{\infty} \rho^j \Delta d_{t+j+1}\right)$, and $\eta^{\mathrm{DR}}_{t+1} = (\mathbb{E}_{t+1} - \mathbb{E}_t)\left(\sum_{j=1}^{\infty} \rho^j r_{t+j+1}\right)$ denote the innovations to the stock return, cash flow, and discount rate, respectively.

Next, we follow Campbell (1991) and Campbell and Ammer (1993) to use a VAR framework to estimate η^r_{t+1}, η^{CF}_{t+1}, and η^{DR}_{t+1}. Specifically, consider the following VAR(1) model:

$$v_{t+1} = A v_t + u_{t+1}, \tag{B.8}$$

where $v_t = [r_t, x_t, z_t']'$ is an $(n+2)$-vector, z_t is an n-vector of conditioning variables, A is an $(n+2)$-by-$(n+2)$ matrix of VAR slope coefficients, and u_{t+1} is an $(n+2)$-vector of innovations with zero mean.[27] Let

[27] The elements in v_t are demeaned before using, while we use the same notation here for convenience.

$e_1' = [1, 0, ..., 0]'$ be an $(n+2)$-vector, the stock return innovation and discount rate news are given by

$$\eta_{t+1}^{r} = e_1' u_{t+1} \tag{B.9}$$

and

$$\eta_{t+1}^{\mathrm{DR}} = (\mathbb{E}_{t+1} - \mathbb{E}_t)\left(\sum_{j=1}^{\infty} \rho^j e_1' v_{t+1+j}\right) = e_1' \sum_{j=1}^{\infty} \rho^j A^j u_{t+1} = e_1' \rho A (I - \rho A)^{-1} u_{t+1}, \tag{B.10}$$

respectively. Accordingly, the cash flow news is residually defined as

$$\eta_{t+1}^{\mathrm{CF}} = \eta_{t+1}^{r} + \eta_{t+1}^{\mathrm{DR}}. \tag{B.11}$$

Moreover, Eq. (B.8) implies that the expected stock return for time $t+1$ made at time t is

$$\mathbb{E}_t(r_{t+1}) = e_1' A v_t. \tag{B.12}$$

Taken all together, we obtain the decomposition of the log stock return as

$$r_{t+1} = \mathbb{E}_t(r_{t+1}) + \eta_{t+1}^{\mathrm{CF}} - \eta_{t+1}^{\mathrm{DR}}. \tag{B.13}$$

Empirically, we use OLS to estimate A and $\{u_{t+1}\}_{t=1}^{T-1}$ in Eq. (B.8) based on sample observations for $\{v_t\}_{t=1}^{T}$. Denote by $\hat{A}$ and $\hat{u}_t$ the OLS estimates, respectively. In addition, we estimate ρ using the sample mean of x_t, and we denote the estimate by $\hat{\rho}$. Finally, we can plug $\hat{A}$, $\hat{u}_t$, and $\hat{\rho}$ into Eqs. (B.9)–(B.12) to obtain the estimated return decomposition components, $\hat{\mathbb{E}}_t(r_{t+1})$, $\hat{\eta}_{t+1}^{r}$, $\hat{\eta}_{t+1}^{\mathrm{DR}}$, and $\hat{\eta}_{t+1}^{\mathrm{CF}}$ for $t = 1, ..., T-1$.

Online Appendices

The online appendix presents supplementary results to the paper "Expected Macroeconomic Conditions and Market Risk Premium: Evidence from a Term Structure of Macroeconomic Forecasts"".

A Supplementary Tables

Table OA.1: **SPF Variable Correlations**

This table presents correlations for the current-quarter forecasts on the seven macroeconomic variables from the Survey of Professional Forecasters (SPF) database. The seven SPF variables includes 1) gross domestic product growth (GDP_e), 2) the industrial production index growth (Indprod_e), 3) the probability of a decline in real GDP (Recess_e), 4) the unemployment rate (Unemp_e), 5) the corporate profits after tax (Cprof_e), 6) housing starts (Housing_e), and 7) GDP price index growth (Infl_e). The sample period is from 1968Q4 to 2019Q3.

(1) Variable	(2) GDP_e	(3) Indprod_e	(4) Recess_e	(5) Unemp_e	(6) Cprof_e	(7) Housing_e
GDP_e	1.00					
Indprod_e	0.93	1.00				
Recess_e	−0.88	−0.83	1.00			
Unemp_e	−0.05	0.03	0.19	1.00		
Cprof_e	0.80	0.79	−0.68	0.16	1.00	
Housing_e	0.19	0.13	−0.17	0.35	0.23	1.00
Infl_e	−0.30	−0.21	0.39	0.19	−0.28	−0.21

Table OA.2: **Predictor Variable Correlations**

This table presents contemporaneous correlations for the 16 economic predictors from Welch and Goyal (2008), as well as the aligned macro risk index (MR^{PLS}). The sample period is from 1968Q4 to 2019Q3.

(1) Variable	(2) DP	(3) DY	(4) EP	(5) DE	(6) SVAR	(7) BM	(8) NTIS	(9) TBL	(10) LTY	(11) LTR	(12) TMS	(13) DFY	(14) DFR	(15) INFL	(16) CAY	(17) IK
DP	1.00															
DY	0.98	1.00														
EP	0.73	0.72	1.00													
DE	0.25	0.24	−0.48	1.00												
SVAR	−0.02	−0.11	−0.27	0.36	1.00											
BM	0.91	0.89	0.81	0.02	−0.08	1.00										
NTIS	0.16	0.15	0.15	−0.01	−0.17	0.25	1.00									
TBL	0.68	0.68	0.66	−0.06	−0.13	0.69	0.22	1.00								
LTY	0.74	0.74	0.62	0.07	−0.10	0.69	0.26	0.90	1.00							
LTR	0.04	0.04	0.04	0.00	0.27	0.01	−0.09	−0.06	−0.02	1.00						
TMS	−0.15	−0.13	−0.32	0.27	0.10	−0.26	0.00	−0.55	−0.14	0.09	1.00					
DFY	0.47	0.47	0.13	0.41	0.43	0.45	−0.24	0.25	0.35	0.24	0.10	1.00				
DFR	0.00	0.06	−0.14	0.20	−0.12	−0.01	0.06	−0.07	0.00	−0.42	0.16	0.03	1.00			
INFL	0.48	0.48	0.56	−0.18	−0.10	0.57	0.19	0.51	0.43	0.10	−0.32	0.11	−0.07	1.00		
CAY	−0.13	−0.10	−0.18	0.09	0.08	−0.36	−0.11	0.08	0.27	0.14	0.34	−0.05	−0.06	−0.21	1.00	
IK	−0.14	−0.16	0.13	−0.36	−0.04	0.06	0.09	0.42	0.19	−0.05	−0.59	−0.23	−0.18	0.24	−0.12	1.00
MR^{PLS}	0.04	0.08	−0.14	0.25	0.11	−0.04	−0.16	−0.29	−0.10	0.07	0.46	0.36	0.21	−0.19	0.12	−0.54

Table OA.3: **In-sample Return Predictability (look-ahead bias-free PLS forecast): 1984Q1-2019Q4**

This table presents the OLS estimates, Newey-West t-statistics, and R^2 of the in-sample predictive regressions for quarterly market excess return. Panel A reports the results of the univariate predictive regression model,

$$R_{t+1} = \alpha + \beta \mathrm{X}^{\mathrm{PLS}}_{\text{Bias-free},t} + e_{t+1},$$

where R_{t+1} is the excess return on the CRSP value-weighted index (annualized) for quarter $t+1$, X could be one of the variable sets $\{\mathrm{SPF7}, \mathrm{Econ}\}$, and $\mathrm{X}^{\mathrm{PLS}}_{\text{Bias-free}}$ denotes the look-ahead bias-free factor extracted via PLS. Panel B reports the results of the bivariate predictive regression model,

$$R_{t+1} = \alpha + \beta\, \mathrm{MR}^{\mathrm{PLS}}_{\text{Bias-free},t} + \psi Ctrl_t + e_{t+1},$$

where $Ctrl$ denotes one of the control variables taken from the first column other than $\mathrm{MR}^{\mathrm{PLS}}_{\text{Bias-free}}$. The term $\mathrm{Econ}^{\mathrm{PLS}}_{\text{Bias-free}}$ denotes the look-ahead bias-free factor extracted from a set of 16 financial and economic variables (Econ) from Welch and Goyal (2008) via PLS. The terms $\mathrm{Econ}^{\mathrm{PC}}$ and $\mathrm{SPF7}^{\mathrm{PC}}$ denote the first principal component (PC) of the Econ variable set and the SPF7 variable set, respectively. Each predictor is standardized to have a zero mean and unit variance. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Since we use the first 15-year data as training period, the in-sample analysis for the look-ahead bias-free PLS forecast is based on the sample period of 1984Q1 through 2019Q4.

(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Univariate Regression				
		β	t-stat	R^2(%)	
$\mathrm{Econ}^{\mathrm{PLS}}_{\text{Bias-free}}$		−0.072***	−2.82	4.81	
$\mathrm{MR}^{\mathrm{PLS}}_{\text{Bias-free}}$		0.070***	2.81	4.58	
	Panel B: Bivariate Regression				
Variable	β(PLS)	t-stat	ψ(Ctrl)	t-stat	R^2(%)
DP	0.062**	2.28	0.024	0.82	5.05
DY	0.063**	2.31	0.020	0.63	4.88
EP	0.070***	2.74	0.020	0.56	4.96
DE	0.071**	2.41	−0.002	−0.05	4.58
SVAR	0.070***	2.81	−0.002	−0.07	4.58
BM	0.068***	2.70	0.009	0.36	4.64
NTIS	0.070***	2.86	−0.004	−0.12	4.59
TBL	0.073***	2.74	0.009	0.32	4.64
LTY	0.069***	2.77	−0.011	−0.46	4.68
LTR	0.071***	2.82	0.036	1.19	5.79
TMS	0.099***	3.12	−0.054	−1.51	6.52
DFY	0.073***	2.72	−0.013	−0.35	4.72
DFR	0.074***	2.84	−0.012	−0.31	4.69
INFL	0.067***	2.73	−0.034	−1.18	5.65
CAY	0.071***	2.83	0.012	0.62	4.71
IK	0.088**	2.57	0.026	0.66	4.90
$\mathrm{Econ}^{\mathrm{PC}}$	0.069***	2.75	0.012	0.49	4.70
$\mathrm{Econ}^{\mathrm{PLS}}_{\text{Bias-free}}$	0.059**	2.21	−0.035	−1.28	5.58
GDP_e	0.070**	2.39	−0.001	−0.04	4.58
$\mathrm{Indprod}_e$	0.065**	2.45	−0.021	−0.70	4.97
Recess_e	0.092***	2.61	−0.029	−0.79	4.89
Unemp_e	0.075**	2.56	−0.015	−0.40	4.76
Cprof_e	0.067***	2.58	−0.022	−0.91	5.03
$\mathrm{Housing}_e$	0.077**	2.18	−0.010	−0.25	4.63
Infl_e	0.068***	2.75	−0.028	−1.02	5.31
$\mathrm{SPF7}^{\mathrm{PC}}$	0.068**	2.51	−0.009	−0.28	4.64

Table OA.4: **Robustness Checks for the Statistical Inference**

This table presents robustness checks concerning β for the predictive regression model,

$$y_{t+1:t+h} = \alpha + \beta X_t + e_{t+1:t+h},$$

where h denotes the forecast horizon. We use the h-quarter-ahead simple (log) excess return on the CRSP value-weighted index as the forecast target $y_{t+1:t+h}$ in Panels A and B (C and D). We annualize $y_{t+1:t+h}$ to make βs comparable across forecast horizons. We use MR^{PLS} ($\text{LT-MR}^{\text{PLS}}$) as the predictor variable X_t in Panels A and C (B and D), where MR^{PLS} and $\text{LT-MR}^{\text{PLS}}$ refer to the PLS factors extracted from the seven current-quarter survey forecasts and the term structures of the seven survey variables, respectively. For predictive regressions in Panels A and B, we report the OLS estimates and the Newey-West corrected t-statistics with $2^*(h-1)$ lags (t-NW) where the statistical significance is based on one-sided wild bootstrap p-values following Huang, Jiang, Tu, and Zhou (2015). For predictive regressions in Panels A and B, we report the OLS estimates, the Hodrick (1992) corrected t-statistics (t-Hodrick), and the Kostakis, Magdalinos, and Stamatogiannis (2015) Wald statistics (IVX-Wald) where the 10%, 5%, and 1% critical values are 2.71, 3.84, and 6.64, respectively. The sample period is from 1969Q1 to 2019Q4. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)	(6)
	$h=1$	$h=2$	$h=4$	$h=8$	$h=12$
Panel A: MR^{PLS}, *Simple Excess Return*					
β	0.08	0.07	0.07	0.05	0.05
t-NW	3.65***	3.69**	3.73**	2.97	3.57*
Panel B: $\text{LT-MR}^{\text{PLS}}$, *Simple Excess Return*					
β	0.08	0.08	0.07	0.06	0.06
t-NW	3.69**	4.23**	4.68**	4.29*	5.67**
Panel C: MR^{PLS}, *Log Excess Return*					
β	0.09	0.07	0.07	0.05	0.04
t-Hodrick	3.64***	3.13***	2.94***	2.24**	2.26**
IVX-Wald	12.27***	9.56***	9.87***	5.77**	6.08**
Panel D: $\text{LT-MR}^{\text{PLS}}$, *Log Excess Return*					
β	0.09	0.08	0.07	0.06	0.05
t-Hodrick	3.47***	3.13***	2.93***	2.54**	2.42**
IVX-Wald	11.98***	9.35***	8.06***	6.54**	6.56**