上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy

Zhenliang Xiong

December 6, 2024

https://doi.org/10.1093/rfs/hhp063

饮 水 思 源 • 爱 国 荣 校

Contents

Research Background, Methodology, and Main Conclusions

1

SJTU

1.1 Research Background

- Welch and Goyal (2008) found that many economic variables with in-sample predictive ability for equity premiums fail to provide consistent out-of-sample gains relative to the historical average.
- Goyal and Welch (2003) found that the dividend-price ratio is not a robust out-of-sample predictor for U.S. equity premiums.
- Fama and French (1989), among others, showed that changes in economic conditions detected by variables such as dividend yield and term spread may signal fluctuations in equity risk premiums.

- **Summary:**
- Relying solely on dividend yield or term spread can capture different components of business conditions, and a given economic variable may send many "false signals" and/or imply implausible equity risk premiums in certain periods. This suggests the need for improved forecasting methods to better establish the empirical reliability of equity premium predictability.

1.2 Research Methodology and Main Conclusions

- If univariate forecasts based on dividend yield and term spread exhibit weak correlation, the average of the two forecasts (a simple type of forecast combination) should have lower volatility and track variations in equity risk premiums more reliably. The study argues this reasoning can be extended to forecasting factors composed of multiple individual economic variables.
- Structural instabilities caused by regime shifts, policy shocks, technological advances, and investor learning lead to a highly uncertain, complex, and evolving expected stock return data-generating process, which is difficult to approximate with univariate predictive regression models. By combining univariate regression model forecasts, the study finds that economic variables are generally valuable and consistently outperform the historical average forecast of equity premiums.
- Similar to how portfolio diversification reduces variance, forecast combinations integrate information from multiple forecasting models, outperforming univariate models in both information content and stability. Using encompassing tests, the study finds that forecast combinations significantly reduce forecast variance and forecast bias, thereby consistently outperforming the historical average forecast in mean square prediction error (MSPE).

1.2 Research Methodology and Main Conclusions

- Fama and French (1989) and Cochrane (1999, 2007) argue that increased risk aversion during economic downturns requires higher risk premiums, thereby generating predictability in equity premiums. The study argues that forecasts generated by historical averages are overly "smooth," thus ignoring risk premium fluctuations corresponding to business cycle volatility.
- The study uses three macroeconomic variables—real GDP, real profits, and real net cash flows—whose future growth is significantly correlated with combination forecasts of equity premiums to define "good," "normal," and "bad" growth periods. It finds that during poor growth periods, out-of-sample gains corresponding to combination forecasts of equity premiums are especially pronounced.
- The study further demonstrates that when forming combination forecasts for real GDP, real profits, and real net cash flow growth, the same set of 15 economic variables used to form equity premium combination forecasts also produces consistently significant out-of-sample gains, indicating that the usefulness of combination forecasts in equity premium prediction partly stems from their ability to forecast the real economy.

Research Model

2

2.1 Predictive Regression Model

- Using rolling estimation windows to generate out-of-sample forecasts of equity premiums (Welch and Goyal, 2008):

$$r_{t+1} = \alpha_i + \beta_i x_{i,t} + \varepsilon_{t+1}$$

- r_{t+1} is the excess return on the stock market index over the risk-free rate, and $x_{i,\,t}$ represents predictor variables related to forecasting ability.
- The full sample of observations T is divided into an in-sample segment comprising the first m observations and an out-of-sample segment consisting of the last q observations; the rolling regression model then becomes:

$$\hat{r}_{i,m+1} = \hat{\alpha}_{i,m} + \hat{\beta}_{i,m} x_{i,m}$$

- $\alpha_{i,m}^{hat}$ and $\beta_{i,m}^{hat}$ are the ordinary least squares (OLS) estimates of α_i and β_i , respectively.

2.1 Predictive Regression Model

- According to Campbell and Thompson (2008) and Welch and Goyal (2008), the natural benchmark forecasting model is the historical average of equity premiums, corresponding to constant expected equity premiums:

$$\bar{r}_{t+1} = \sum_{j=1}^{t} r_j$$

2.2 Combination Model Forecasts

- Mamaysky, Spiegel, and Zhang (2007) found that combining the forecasts of an OLS model with those of a Kalman filter model (as published by the same authors in 2008) can significantly increase the number of mutual funds with predictable out-of-sample alpha.
- The study constructs a weighted average combination model of predictors using ex-ante combination weights ($\omega_{i,t}$) based on monthly data, with $q_0$observations set aside as an initial holdout period in the out-of-sample phase.

$$\hat{r}_{c,t+1} = \sum_{i=1}^{N} \omega_{i,t} \hat{r}_{i,t+1},$$

2.2 Combination Model Forecasts

- Models are divided into two categories based on how combination weights are determined. The first category uses simple averaging schemes: mean, median, and trimmed mean. The second category of combination methods is based on Stock and Watson (2004), where combination weights formed at time t are functions of the historical forecasting performance of individual models during the hold-out sample period. The second category discounted mean square prediction error (DMSPE) combination method adopts the following weights:

$$\omega_{i,t} = \phi_{i,t}^{-1} \Big/ \sum_{j=1}^{N} \phi_{j,t}^{-1}, \qquad \phi_{i,t} = \sum_{s=m}^{t-1} \theta^{t-1-s} (r_{s+1} - \hat{r}_{i,s+1})^2,$$

- θ is the discount factor. The DMSPE method assigns greater weight to univariate predictive regression model forecasts with lower MSPE values during the hold-out sample period. When $\theta < 1$, recent forecast accuracy of univariate models receives greater weight. The study considers the second category of combination methods with $\theta = 1.0/0.9$, for a total of five combination methods.

2.3 Forecast Performance Evaluation Methods

- The study uses the out-of-sample R^2 statistic (R^2_{OS}) proposed by Campbell and Thompson (2008) to compare the performance of univariate predictions/combination forecasts with the historical mean forecast.

$$R^2_{OS} = 1 - \frac{\sum_{k=q_0+1}^{q} (r_{m+k} - \hat{r}_{m+k})^2}{\sum_{k=q_0+1}^{q} (r_{m+k} - \bar{r}_{m+k})^2}$$

- R^2_{OS} statistic measures the reduction in MSPE (mean squared prediction error) of the predictive regression model or combination forecast relative to the historical mean benchmark.
- Campbell and Thompson (2008) argue that even very small positive values of R^2_{OS}—such as 0.5% for monthly data or 1% for quarterly data—can indicate an economically meaningful degree of return predictability, as they correspond to an increase in annual portfolio returns for a mean-variance investor.

2.3 Forecast Performance Evaluation Methods

- Since R_{OS}^2 does not account for the risk borne during the out-of-sample period, the study also calculates the realized utility gains over time for a mean-variance investor.
- First, the average utility γ, is computed for a mean-variance investor with relative risk aversion, who allocates assets each month using stock premium forecasts based on the historical mean. It is assumed that the investor estimates variance using a ten-year rolling window of quarterly returns. Using stock premium forecasts from the historical mean and from univariate/combination models, the mean-variance investor will allocate portfolio weights $\omega_{0,t}$ and $\omega_{j,t}$ to stocks for period t+1 at the end of period t, The realized average utility over the out-of-sample period for the investor is denoted as v_0^{hat} and v_j^{hat}:

$$w_{0,t} = \left(\frac{1}{\gamma}\right)\left(\frac{\bar{r}_{t+1}}{\hat{\sigma}_{t+1}^2}\right) \qquad \hat{\nu}_0 = \hat{\mu}_0 - \left(\frac{1}{2}\right)\gamma\hat{\sigma}_0^2 \qquad w_{j,t} = \left(\frac{1}{\gamma}\right)\left(\frac{\hat{r}_{t+1}}{\hat{\sigma}_{t+1}^2}\right) \qquad \hat{\nu}_j = \hat{\mu}_j - \left(\frac{1}{2}\right)\gamma\hat{\sigma}_j^2$$

- Where ${\sigma^{hat}}_{t+1}^2$ is the rolling estimated variance of stock returns, and μ_0^{hat}/μ_j^{hat} and ${\sigma^{hat}}_0^2/{\sigma^{hat}}_j^2$ are the sample mean and variance.

2.4 Detailed Variables and Data Sources

- The study considers 15 variables from Welch and Goyal (2008), with quarterly data spanning January 1947 to April 2005. Detailed variables are:

- Dividend-price ratio (log) D/P: The difference between the log of dividends paid by the S&P 500 index and the log of stock price (S&P 500 index), with dividends using a one-year moving sum;
- Dividend yield (log) D/Y: The difference between the log of dividends and the log of lagged stock price;
- Earnings-price ratio (log) E/P: The difference between the log of S&P 500 earnings and the log of stock price, with earnings using a one-year moving sum;
- Dividend payout ratio (log) D/E: The difference between the log of dividends and the log of earnings;
- Stock variance SVAR: Sum of squared daily returns of the S&P 500 index;
- Book-to-market ratio B/M: Ratio of Dow Jones Industrial Average market value to book value;

2.4 Detailed Variables and Data Sources

- Net equity expansion NTIS: Ratio of 12-month net issuance of NYSE-listed stocks to year-end total market value of NYSE stocks;
- Treasury bill rate TBL: Three-month Treasury bill rate (secondary market);
- Long-term yield LTY: Long-term government bond YTM;
- Long-term return LTR: Long-term government bond return rate;
- Term spread TMS: Difference between long-term yield and Treasury bill rate;
- Default yield spread DFY: Difference between BAA and AAA-rated corporate bond yields;
- Default return spread DFR: Difference between long-term corporate bond and long-term government bond yields;
- Inflation INFL: Calculated from CPI (all urban consumers). According to Welch and Goyal (2008), since inflation rate data is published the next month, one-period-ahead data is used;
- Investment-to-capital ratio I/K: Ratio of total (private non-residential fixed) investment to total capital for the entire economy (Cochrane, 1991).

2.5 Out-of-Sample Forecast Periods

- The study also considers three different out-of-sample forecast evaluation periods, with the first two periods corresponding to the periods analyzed by Welch and Goyal (2008):

- The "long" out-of-sample period covering January 1965 - April 2005;
- The recent out-of-sample period covering the last 30 years of the full sample, January 1976 - April 2005, because the out-of-sample forecasting ability of many economic variables deteriorated significantly after the mid-1970s oil crisis;
- The recent out-of-sample period covering the last six years of the full sample, January 2000 - April 2005, enabling analysis of predictor performance during the recent market period characterized by the bursting of the "tech bubble."

Empirical Results

3

SJTU

3.1 Out-of-Sample Forecast Results: 1965.01-2005.04

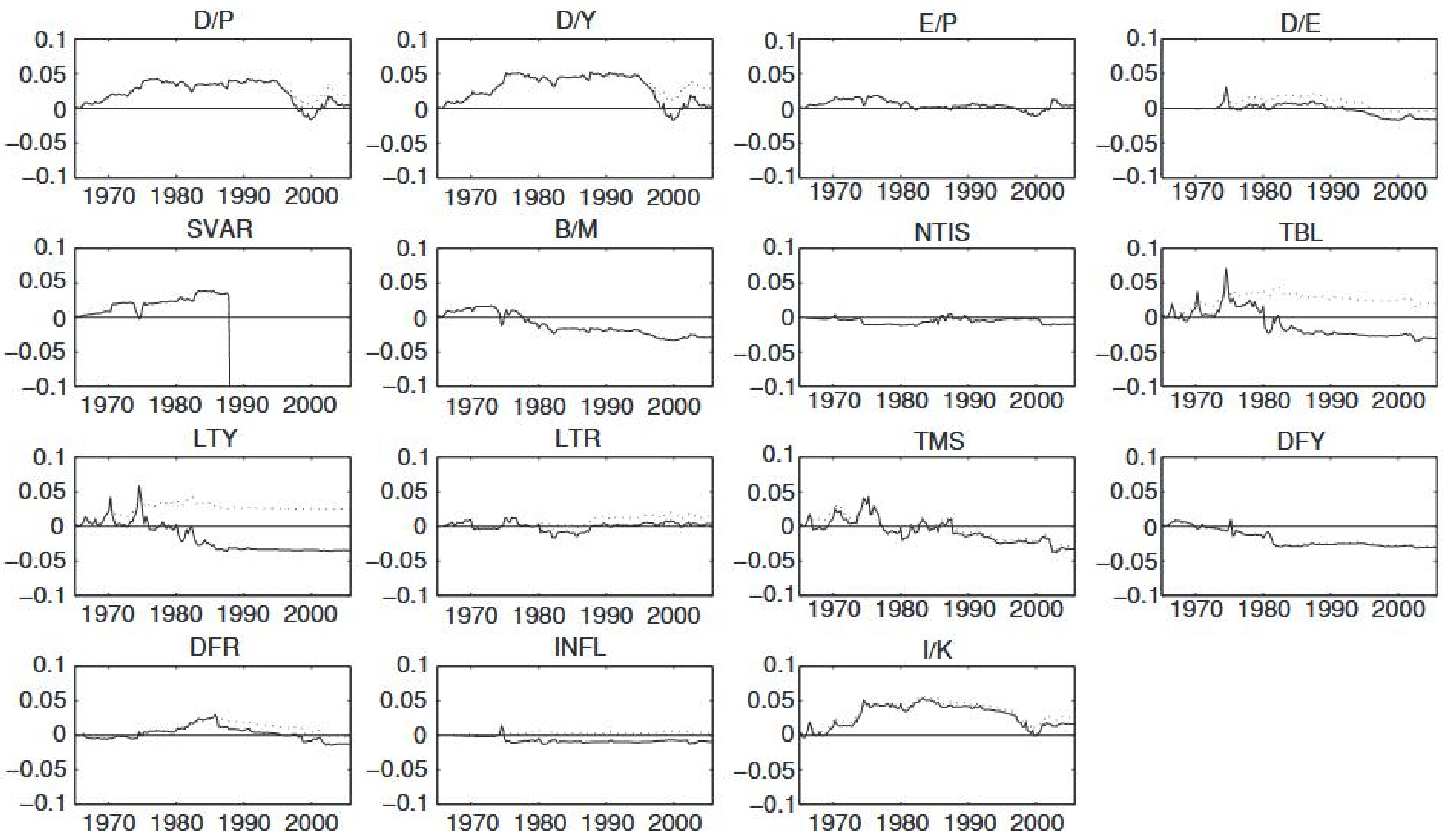

Figure 1
Cumulative square prediction error for the historical average benchmark forecasting model minus the cumulative square prediction error for the individual predictive regression forecasting model, 1965:1-2005:4
The dotted (solid) line corresponds to individual model forecasts that (do not) impose Campbell and Thompson (2008) restrictions.

- Figure 1 shows a time series plot of the difference between the cumulative squared prediction error of the historical average benchmark forecast and the cumulative squared prediction error of the forecast in the univariate predictive regression model.

- The solid lines indicate that none of the 15 individual economic variables consistently outperforms the historical average. This validates Welch and Goyal (2008): in forecasting equity premiums, it is difficult to identify individual predictors that reliably beat the historical average.

3.1 Out-of-Sample Forecast Results: 1965.01-2005.04

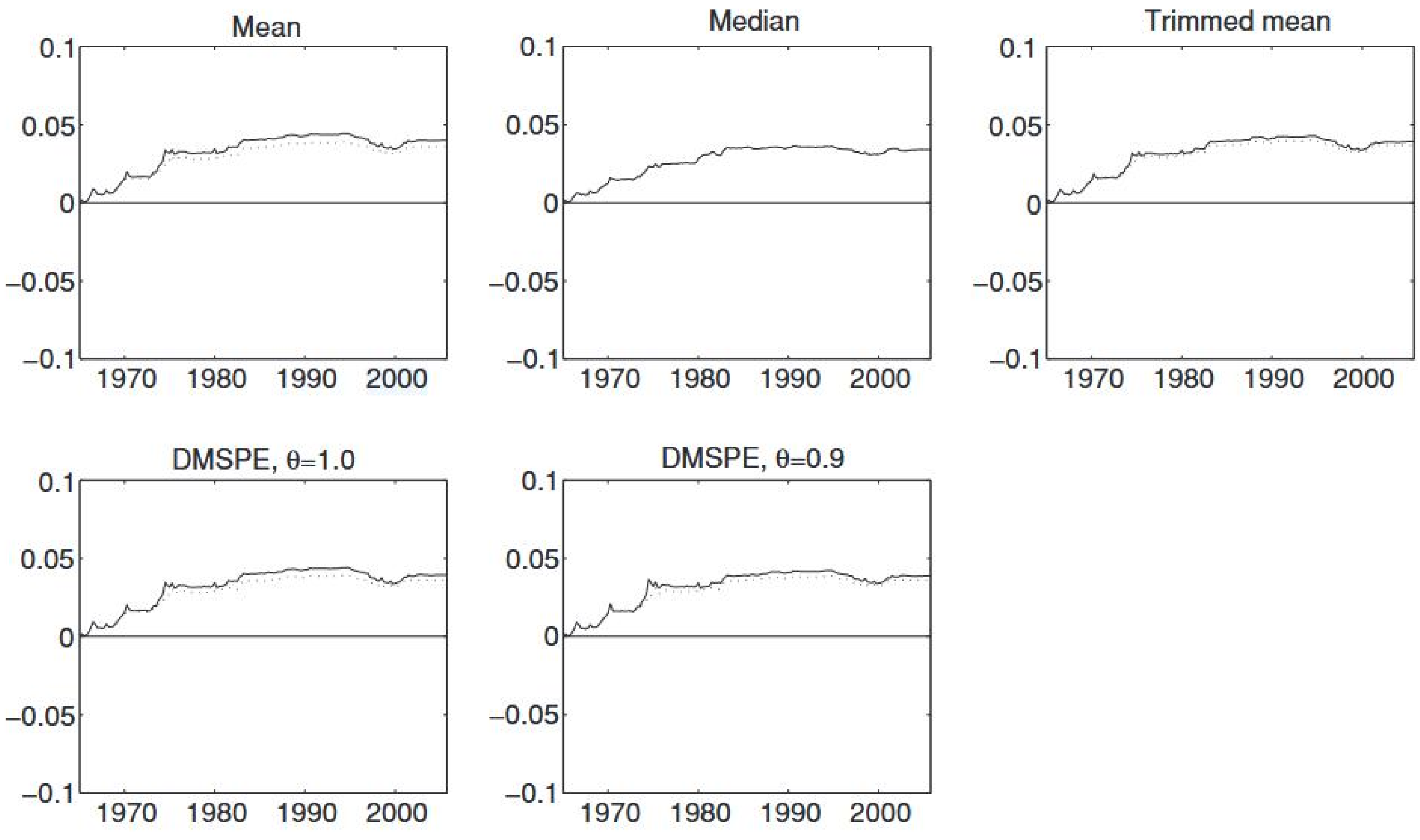

Figure 2
Cumulative square prediction error for the historical average benchmark forecasting model minus the cumulative square prediction error for the combination forecasting model, 1965:1-2005:4
The dotted (solid) line corresponds to combination forecasts based on individual model forecasts that (do not) impose Campbell and Thompson (2008) restrictions.

- The solid lines in Figure 2 plot the difference between the cumulative squared prediction error of the historical average forecast and the cumulative squared prediction error of combination forecasts. The slope of the curves in Figure 2 is predominantly positive, indicating that combination forecasts provide much more consistent out-of-sample gains over time than univariate predictive regression models.
- Figure 2 avoids the frequent, often persistent, and substantial declines in the curves, demonstrating that forecast combinations are an effective strategy for equity premium prediction compared to univariate predictive regression models.

3.1 Out-of-Sample Forecast Results: 1965.01–2005.04

Table 1
Equity premium out-of-sample forecasting results for individual forecasts and combining methods

Individual predictive regression model forecasts						Combination forecasts		
Predictor	R^2_{OS} (%)	Δ (%)	Predictor	R^2_{OS} (%)	Δ (%)	Combining method	R^2_{OS} (%)	Δ (%)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. 1965:1–2005:4 out-of-sample period								
D/P	0.34*	0.55	*LTY*	−3.09	2.29	Mean	3.58***	2.34
D/Y	0.25*	1.41	*LTR*	0.33	1.30	Median	3.04***	1.03
E/P	0.36	0.64	*TMS*	−2.96	5.14	Trimmed mean	3.51***	2.11
D/E	−1.42	0.58	*DFY*	−2.72	−0.83	DMSPE, $\theta = 1.0$	3.54***	2.41
SVAR	−12.97	0.13	*DFR*	−1.10	0.57	DMSPE, $\theta = 0.9$	3.49***	2.59
B/M	−2.60	−0.58	*INFL*	−0.84	1.39			
NTIS	−0.91	0.08	*I/K*	1.44**	2.80	Mean, CT	3.23***	1.25
TBL	−2.78	2.60						
Panel B. 1976:1–2005:4 out-of-sample period								
D/P	−5.08	−0.70	*LTY*	−5.59	−0.89	Mean	1.19*	0.57
D/Y	−6.22	−0.54	*LTR*	−0.27	1.43	Median	1.51**	0.53
E/P	−1.70	0.75	*TMS*	−7.24	2.08	Trimmed mean	1.23*	0.59
D/E	−2.26	−1.65	*DFY*	−2.48	−1.18	DMSPE, $\theta = 1.0$	1.11*	0.54
SVAR	−22.47	0.06	*DFR*	−2.14	−0.64	DMSPE, $\theta = 0.9$	1.01*	0.46
B/M	−4.72	−1.27	*INFL*	−0.08	0.45			
NTIS	0.10	0.60	*I/K*	−3.47	−0.85	Mean, CT	1.20*	0.55
TBL	−7.31	−0.82						
Panel C. 2000:1–2005:4 out-of-sample period								
D/P	10.32*	12.96	*LTY*	−0.32	0.24	Mean	3.04**	2.31
D/Y	10.40*	12.98	*LTR*	−1.72	2.57	Median	1.56*	0.28
E/P	8.02*	9.53	*TMS*	−4.98	4.23	Trimmed mean	2.98**	2.12
D/E	0.56	0.50	*DFY*	−0.53	−1.52	DMSPE, $\theta = 1.0$	2.56**	1.65
SVAR	−5.62	−1.64	*DFR*	−2.10	1.76	DMSPE, $\theta = 0.9$	2.66**	1.97
B/M	2.32	3.09	*INFL*	−1.42	0.57			
NTIS	−4.09	1.33	*I/K*	8.96**	9.13	Mean, CT	2.43**	1.32
TBL	−2.50	−0.20						

- The second and fifth columns of Panel A in Table 1 show that only 5 out of the 15 individual predictor variables have a positive R^2_{OS}, with 4 of them less than or equal to 0.36%. I/K is the only predictor with an R^2_{OS} greater than 0.36% (1.44%). At the 10% significance level, 3 out of the positive R^2_{OS} statistics are significantly greater than zero, while only the R^2_{OS} for I/K is significant at the 5% level.
- The average utility gains shown in the third and sixth columns of Panel A generally provide stronger evidence for out-of-sample predictability, since 13 out of the 15 predictors deliver positive utility gains relative to the historical mean.

3.1 Out-of-Sample Forecast Results: 1965.01–2005.04

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

Table 1
Equity premium out-of-sample forecasting results for individual forecasts and combining methods

Individual predictive regression model forecasts						Combination forecasts		
Predictor	R^2_{OS} (%)	Δ (%)	Predictor	R^2_{OS} (%)	Δ (%)	Combining method	R^2_{OS} (%)	Δ (%)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. 1965:1–2005:4 out-of-sample period								
D/P	0.34*	0.55	*LTY*	−3.09	2.29	Mean	3.58***	2.34
D/Y	0.25*	1.41	*LTR*	0.33	1.30	Median	3.04***	1.03
E/P	0.36	0.64	*TMS*	−2.96	5.14	Trimmed mean	3.51***	2.11
D/E	−1.42	0.58	*DFY*	−2.72	−0.83	DMSPE, θ = 1.0	3.54***	2.41
SVAR	−12.97	0.13	*DFR*	−1.10	0.57	DMSPE, θ = 0.9	3.49***	2.59
B/M	−2.60	−0.58	*INFL*	−0.84	1.39			
NTIS	−0.91	0.08	*I/K*	1.44**	2.80	Mean, CT	3.23***	1.25
TBL	−2.78	2.60						
Panel B. 1976:1–2005:4 out-of-sample period								
D/P	−5.08	−0.70	*LTY*	−5.59	−0.89	Mean	1.19*	0.57
D/Y	−6.22	−0.54	*LTR*	−0.27	1.43	Median	1.51**	0.53
E/P	−1.70	0.75	*TMS*	−7.24	2.08	Trimmed mean	1.23*	0.59
D/E	−2.26	−1.65	*DFY*	−2.48	−1.18	DMSPE, θ = 1.0	1.11*	0.54
SVAR	−22.47	0.06	*DFR*	−2.14	−0.64	DMSPE, θ = 0.9	1.01*	0.46
B/M	−4.72	−1.27	*INFL*	−0.08	0.45			
NTIS	0.10	0.60	*I/K*	−3.47	−0.85	Mean, CT	1.20*	0.55
TBL	−7.31	−0.82						
Panel C. 2000:1–2005:4 out-of-sample period								
D/P	10.32*	12.96	*LTY*	−0.32	0.24	Mean	3.04**	2.31
D/Y	10.40*	12.98	*LTR*	−1.72	2.57	Median	1.56*	0.28
E/P	8.02*	9.53	*TMS*	−4.98	4.23	Trimmed mean	2.98**	2.12
D/E	0.56	0.50	*DFY*	−0.53	−1.52	DMSPE, θ = 1.0	2.56**	1.65
SVAR	−5.62	−1.64	*DFR*	−2.10	1.76	DMSPE, θ = 0.9	2.66**	1.97
B/M	2.32	3.09	*INFL*	−1.42	0.57			
NTIS	−4.09	1.33	*I/K*	8.96**	9.13	Mean, CT	2.43**	1.32
TBL	−2.50	−0.20						

- The seventh to ninth columns of Panel A in Table 1 show that all combination forecasts have R^2_{OS} statistics greater than 3% and exceed the R^2_{OS} of the single predictor variable I/K, with 4 out of 5 combination methods at or above 3.49%. Moreover, all R^2_{OS} statistics are significant at the 1% level.
- Utility gains associated with combination forecasts are also considerable: in Panel A, four out of five weighting combination methods yield utility gains well above 2%. Except for the median-weighting combination, the other combination methods produce very similar out-of-sample predictive results.

Note: Mean, CT is the combination forecast result under Campbell and Thompson (2008) mean combination method

3.1 Out-of-Sample Forecast Results: 1965.01–2005.04

上海交通大学 SHANGHAI JIAO TONG UNIVERSITY

Table 1
Equity premium out-of-sample forecasting results for individual forecasts and combining methods

Individual predictive regression model forecasts						Combination forecasts		
Predictor	R^2_{OS} (%)	Δ (%)	Predictor	R^2_{OS} (%)	Δ (%)	Combining method	R^2_{OS} (%)	Δ (%)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. 1965:1–2005:4 out-of-sample period								
D/P	0.34*	0.55	*LTY*	−3.09	2.29	Mean	3.58***	2.34
D/Y	0.25*	1.41	*LTR*	0.33	1.30	Median	3.04***	1.03
E/P	0.36	0.64	*TMS*	−2.96	5.14	Trimmed mean	3.51***	2.11
D/E	−1.42	0.58	*DFY*	−2.72	−0.83	DMSPE, θ = 1.0	3.54***	2.41
SVAR	−12.97	0.13	*DFR*	−1.10	0.57	DMSPE, θ = 0.9	3.49***	2.59
B/M	−2.60	−0.58	*INFL*	−0.84	1.39			
NTIS	−0.91	0.08	*I/K*	1.44**	2.80	Mean, CT	3.23***	1.25
TBL	−2.78	2.60						
Panel B. 1976:1–2005:4 out-of-sample period								
D/P	−5.08	−0.70	*LTY*	−5.59	−0.89	Mean	1.19*	0.57
D/Y	−6.22	−0.54	*LTR*	−0.27	1.43	Median	1.51**	0.53
E/P	−1.70	0.75	*TMS*	−7.24	2.08	Trimmed mean	1.23*	0.59
D/E	−2.26	−1.65	*DFY*	−2.48	−1.18	DMSPE, θ = 1.0	1.11*	0.54
SVAR	−22.47	0.06	*DFR*	−2.14	−0.64	DMSPE, θ = 0.9	1.01*	0.46
B/M	−4.72	−1.27	*INFL*	−0.08	0.45			
NTIS	0.10	0.60	*I/K*	−3.47	−0.85	Mean, CT	1.20*	0.55
TBL	−7.31	−0.82						
Panel C. 2000:1–2005:4 out-of-sample period								
D/P	10.32*	12.96	*LTY*	−0.32	0.24	Mean	3.04**	2.31
D/Y	10.40*	12.98	*LTR*	−1.72	2.57	Median	1.56*	0.28
E/P	8.02*	9.53	*TMS*	−4.98	4.23	Trimmed mean	2.98**	2.12
D/E	0.56	0.50	*DFY*	−0.53	−1.52	DMSPE, θ = 1.0	2.56**	1.65
SVAR	−5.62	−1.64	*DFR*	−2.10	1.76	DMSPE, θ = 0.9	2.66**	1.97
B/M	2.32	3.09	*INFL*	−1.42	0.57			
NTIS	−4.09	1.33	*I/K*	8.96**	9.13	Mean, CT	2.43**	1.32
TBL	−2.50	−0.20						

- The last two columns of Panel C show that combination forecasts are generally much higher than the historical mean. For the weighting combination methods in column 8 of Panel C, all R^2_{OS} statistics are positive, with 4 out of 5 greater than or equal to 2.56%, and all weighting combination methods are statistically significant at conventional levels (most at the 5% level).
- The utility gains from combination methods in column 9 of Panel C are also all positive, with most close to or above 2%.

Note: Mean, CT is the combination forecast result under Campbell and Thompson (2008) mean combination method

3.2 Forecast Encompassing Test - Introduction

- Chong and Hendry (1986), Fair and Shiller (1990) forecast encompassing: Consider an optimal composite forecast ${r^{hat}}^{*}_{t+1}$ comprised of a convex combination of the forecasts from model i and model j.

$$\hat{r}^{*}_{t+1} = (1-\lambda)\hat{r}_{i,t+1} + \lambda\hat{r}_{j,t+1}, \quad 0 \le \lambda \le 1$$

- If $\lambda = 0$, the forecast of model i encompasses the forecast of model j
- If $\lambda > 0$, the forecast of model i does not encompass the forecast of model j

- Essentially, if we reject the null hypothesis of encompassing, then it is useful to combine forecasts from models i and j, rather than relying solely on the forecast from model i.

3.2 Forecast Encompassing Test - MHLN Statistic

- MHLN-modified HLN statistic by Harvey, Leybourne, and Newbold (1998):
- Define: $d_{t+1} = (\hat{u}_{i,t+1} - \hat{u}_{j,t+1})\hat{u}_{i,t+1}$

  $\hat{u}_{i,t+1} = r_{t+1} - \hat{r}_{i,t+1} \qquad \hat{u}_{j,t+1} = r_{t+1} - \hat{r}_{j,t+1}$
- Let: $\bar{d} = \frac{1}{q-q_0} \sum_{k=q_0+1}^{q} d_{R+k}$

$$MHLN = [(q - q_0 - 1)/(q - q_0)][\hat{V}(\bar{d})^{-1/2}]\bar{d}$$

- Where $\hat{V}(\bar{d}) = (q - q_0)^{-1}\hat{\phi}_0 \quad \hat{\phi}_0 = (q - q_0)^{-1} \sum_{k=q_0+1}^{q} (d_{R+k} - \bar{d})^2$
- H0: $\lambda = 0$ (the forecast of model i encompasses the forecast of model j)
- H1: $\lambda > 0$ (the forecast of model i does not encompass the forecast of model j)

3.2 Forecast Encompassing Test: 1965.01-2005.04

Table 2
Forecast encompassing test results, *MHLN* statistic *p*-values, 1965:1–2005:4

	D/P	*D/Y*	*E/P*	*D/E*	*SVAR*	*B/M*	*NTIS*	*TBL*	*LTY*	*LTR*	*TMS*	*DFY*	*DFR*	*INFL*	*I/K*	Mean	Med.	TM	D(1.0)	D(0.9)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
D/P		0.32	0.23	0.03	0.13	0.03	0.02	0.00	0.01	0.03	0.00	0.01	0.02	0.04	0.03	0.38	0.33	0.41	0.38	0.36
D/Y	0.38		0.21	0.03	0.12	0.04	0.02	0.00	0.01	0.03	0.00	0.01	0.02	0.04	0.03	0.33	0.28	0.36	0.32	0.31
E/P	0.23	0.19		0.04	0.15	0.02	0.06	0.01	0.01	0.09	0.00	0.03	0.06	0.10	0.05	0.63	0.61	0.66	0.62	0.59
D/E	0.08	0.07	0.14		0.15	0.04	0.10	0.02	0.04	0.20	0.01	0.07	0.14	0.31	0.26	0.85	0.76	0.85	0.84	0.85
SVAR	0.01	0.02	0.02	0.02		0.00	0.01	0.00	0.01	0.02	0.00	0.00	0.01	0.03	0.01	0.26	0.11	0.14	0.26	0.24
B/M	0.52	0.41	0.81	0.11	0.17		0.16	0.01	0.02	0.22	0.02	0.16	0.19	0.23	0.12	0.79	0.82	0.81	0.78	0.75
NTIS	0.07	0.07	0.20	0.06	0.16	0.03		0.01	0.01	0.22	0.01	0.06	0.11	0.20	0.06	0.76	0.74	0.76	0.74	0.73
TBL	0.01	0.01	0.02	0.04	0.08	0.01	0.01		0.15	0.04	0.01	0.01	0.02	0.04	0.22	0.20	0.13	0.18	0.20	0.20
LTY	0.04	0.04	0.07	0.11	0.10	0.02	0.05	0.24		0.08	0.00	0.02	0.05	0.11	0.23	0.38	0.29	0.37	0.38	0.39
LTR	0.03	0.03	0.07	0.07	0.14	0.01	0.06	0.01	0.02		0.01	0.02	0.04	0.12	0.06	0.52	0.45	0.51	0.51	0.50
TMS	0.01	0.01	0.02	0.02	0.10	0.02	0.02	0.01	0.01	0.06		0.02	0.03	0.04	0.14	0.25	0.17	0.22	0.25	0.25
DFY	0.18	0.16	0.42	0.17	0.18	0.17	0.26	0.01	0.02	0.37	0.02		0.25	0.32	0.16	0.87	0.89	0.89	0.86	0.86
DFR	0.08	0.06	0.18	0.07	0.15	0.05	0.06	0.01	0.02	0.08	0.01	0.06		0.17	0.08	0.84	0.81	0.85	0.83	0.82
INFL	0.09	0.07	0.22	0.12	0.16	0.06	0.18	0.01	0.03	0.24	0.01	0.09	0.18		0.16	0.94	0.92	0.94	0.93	0.94
I/K	0.01	0.01	0.02	0.01	0.11	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.01	0.01		0.27	0.16	0.25	0.27	0.26
Mean	0.01	0.01	0.02	0.01	0.13	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.01	0.01	0.02		0.16	0.24	0.31	0.37
Med.	0.02	0.02	0.03	0.01	0.14	0.00	0.00	0.01	0.01	0.04	0.00	0.00	0.01	0.01	0.03	0.63		0.68	0.59	0.55
TM	0.01	0.01	0.02	0.01	0.14	0.00	0.00	0.01	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.48	0.19		0.41	0.39
D(1.0)	0.01	0.01	0.02	0.00	0.13	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.67	0.18	0.32		0.40
D(0.9)	0.01	0.01	0.03	0.00	0.13	0.01	0.00	0.00	0.01	0.03	0.00	0.00	0.01	0.01	0.02	0.59	0.23	0.43	0.56	

This table reports *p*-values for the Harvey, Leybourne, and Newbold (1998) *MHLN* statistic. The statistic corresponds to a one-sided (upper-tail) test of the null hypothesis that the forecast given in the column heading encompasses the forecast given in the row heading against the alternative hypothesis that the forecast given in the column heading does not encompass the forecast given in the row heading. The table uses the following abbreviations for the combination forecasts: Med. = Median; TM = trimmed mean; D(1.0) = DMSPE, $\theta = 1.0$; D(0.9) = DMSPE, $\theta = 0.9$.

- The p-value results of the out-of-sample forecast MHLN statistic in Table 2 indicate that the forecast of each economic variable does not encompass forecasts of at least three of the remaining variables; combination forecasts can encompass forecasts from univariate predictive regression models and other combination methods.

3.3 Forecast Stability - Correlation Matrix of Combination Forecasts and Univariate Forecasts

Table 3
Correlation matrix for equity premium forecasts based on individual predictive regression models, 1965:1–2005:4

	D/P	*D/Y*	*E/P*	*D/E*	*SVAR*	*B/M*	*NTIS*	*TBL*	*LTY*	*LTR*	*TMS*	*DFY*	*DFR*	*INFL*	*I/K*
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
D/P	1.00	0.95	0.88	−0.43	0.05	0.78	0.06	−0.80	−0.84	−0.06	−0.16	0.37	0.01	−0.53	−0.29
D/Y	0.95	1.00	0.83	−0.39	−0.07	0.73	0.13	−0.71	−0.76	0.04	−0.07	0.45	0.02	−0.51	−0.20
E/P	0.88	0.83	1.00	−0.68	0.04	0.80	0.13	−0.80	−0.78	−0.05	−0.25	0.34	0.00	−0.54	−0.44
D/E	−0.43	−0.39	−0.68	1.00	−0.04	−0.58	−0.06	0.40	0.42	0.12	0.15	−0.03	0.18	0.71	0.30
SVAR	0.05	−0.07	0.04	−0.04	1.00	0.03	−0.02	−0.07	−0.08	0.07	−0.02	0.03	0.01	−0.09	0.01
B/M	0.78	0.73	0.80	−0.58	0.03	1.00	−0.19	−0.70	−0.83	−0.04	−0.08	0.50	0.02	−0.55	−0.31
NTIS	0.06	0.13	0.13	−0.06	−0.02	−0.19	1.00	−0.02	0.11	0.14	−0.07	−0.04	−0.05	0.10	−0.19
TBL	−0.80	−0.71	−0.80	0.40	−0.07	−0.70	−0.02	1.00	0.88	0.15	0.58	−0.36	0.01	0.39	0.52
LTY	−0.84	−0.76	−0.78	0.42	−0.08	−0.83	0.11	0.88	1.00	0.11	0.19	−0.50	−0.03	0.49	0.26
LTR	−0.06	0.04	−0.05	0.12	0.07	−0.04	0.14	0.15	0.11	1.00	0.13	0.16	−0.32	0.05	0.05
TMS	−0.16	−0.07	−0.25	0.15	−0.02	−0.08	−0.07	0.58	0.19	0.13	1.00	0.12	0.12	0.00	0.65
DFY	0.37	0.45	0.34	−0.03	0.03	0.50	−0.04	−0.36	−0.50	0.16	0.12	1.00	0.18	−0.10	0.05
DFR	0.01	0.02	0.00	0.18	0.01	0.02	−0.05	0.01	−0.03	−0.32	0.12	0.18	1.00	0.28	0.06
INFL	−0.53	−0.51	−0.54	0.71	−0.09	−0.55	0.10	0.39	0.49	0.05	0.00	−0.10	0.28	1.00	0.01
I/K	−0.29	−0.20	−0.44	0.30	0.01	−0.31	−0.19	0.52	0.26	0.05	0.65	0.05	0.06	0.01	1.00

This table reports correlation coefficients for the individual predictive regression model forecasts given in the row and column headings.

- Table 3 shows the correlation matrix constituting combination forecasts and univariate regression model forecasts. The correlations between forecasts generated by various valuation ratios are relatively large. However, the correlations of other variables are very small, some of which are negative. This suggests that combination forecasts may reduce the variance of combination forecasts relative to each univariate predictive regression model forecast.

3.3 Forecast Stability - Univariate Regression Model Forecasts During Out-of-Sample Period

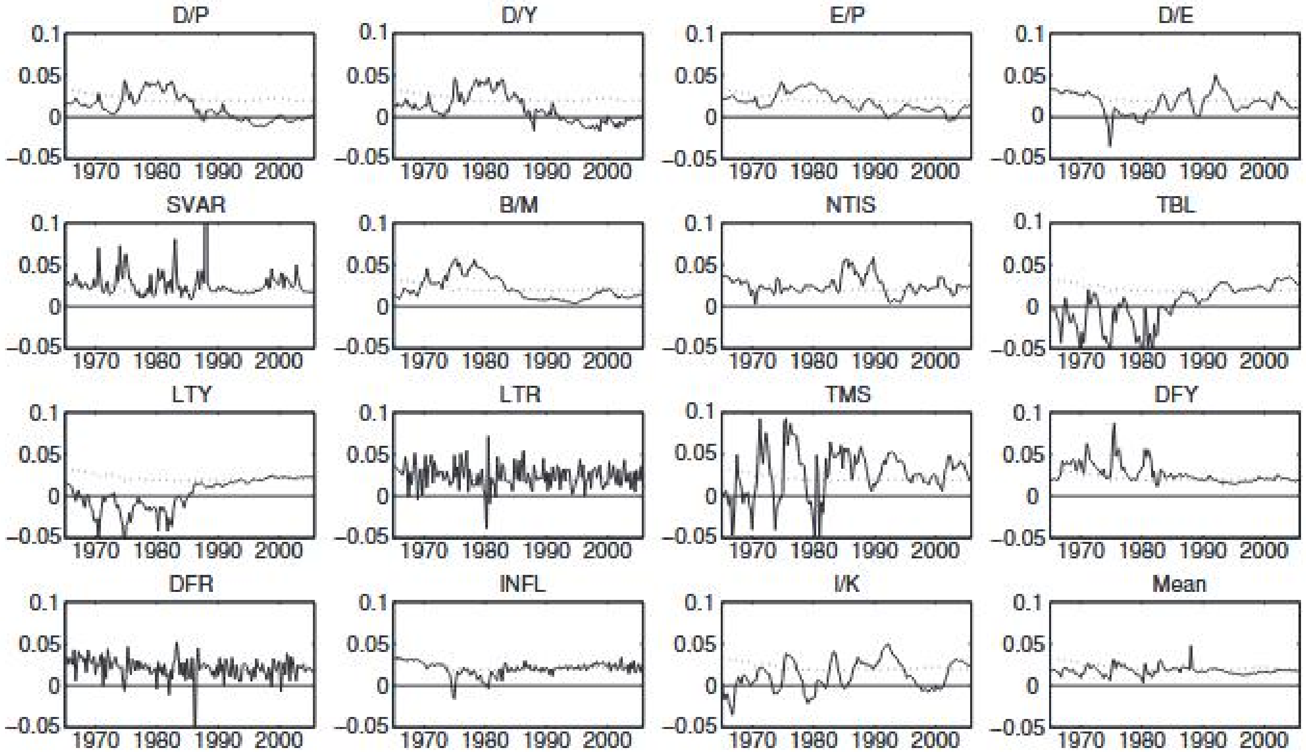

Figure 3
Equity premium forecasts for individual models and the mean combining method, 1965:1–2005:4
The solid (dotted) line corresponds to the forecasting model given in the panel heading (historical average forecasting model).

- Figure 3 confirms that forecast combinations reduce forecast variability. Forecasts from univariate regression models are typically highly variable, implying either negative expected equity premiums or unrealistically large values.

- Overall, univariate forecasts appear to contain substantial "noise" and give too many false signals, thereby undermining forecast effectiveness. In contrast, mean combination forecasts are more stable than univariate forecasts and exhibit more reasonable fluctuations in their magnitude.

3.3 Forecast Stability - Scatter Plot of Forecast Variance and Squared Forecast Bias

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

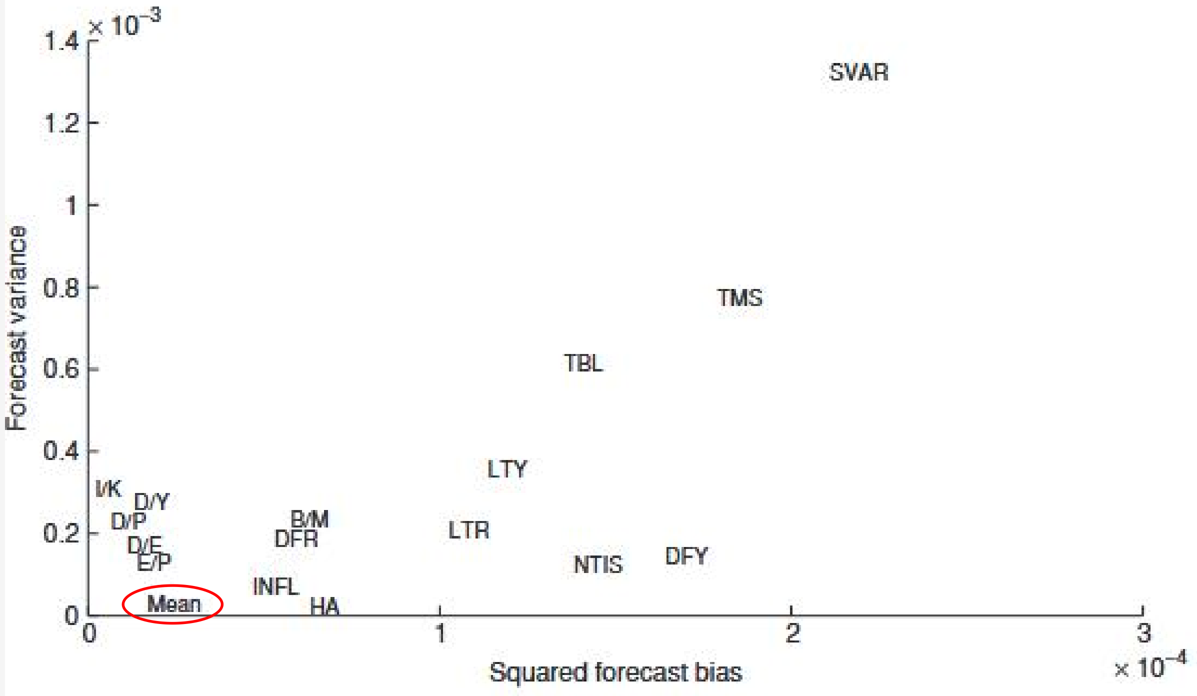

Figure 4

Scatterplot of forecast variances and squared forecast biases, 1965:1–2005:4

HA (Mean) corresponds to the historical average (mean combination) forecast. The other points correspond to the individual predictive regression model forecasts.

Note: Since scatter plots of other combination methods are close to the mean combination method, only Mean forecast variance and squared bias scatter plots are shown

- The forecast variance of the mean combination method is lower than that of all univariate predictive regression models, and its squared forecast bias is close to the minimum squared bias among the univariate models. This also explains why combination forecasts yield R^2_{OS} values greater than those produced by any univariate predictor.

- The variance of combination forecasts is close to that of the smoothed historical mean forecast, thereby reducing the noise present in individual regression model forecasts; furthermore, combination forecasts incorporate information from a wide range of economic variables, information that is ignored by the historical mean, resulting in substantially smaller forecast bias than the historical mean forecast.

SJTU

3.4 Why Does the "Kitchen Sink" Model Have Lower Forecasting Ability?

- The "kitchen sink" model is a multivariate regression model composed of multiple individual variables mixed together:

$$r_{t+1} = \alpha^{KS} + \beta_1^{KS} x_{1,t} + \cdots + \beta_N^{KS} x_{N,t} + \varepsilon_{t+1}$$

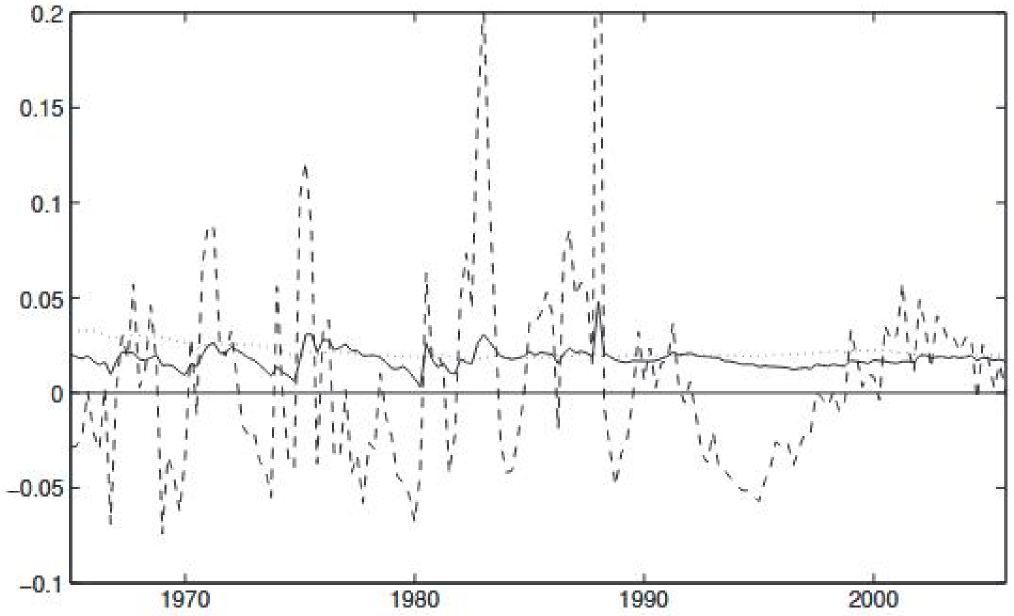

Figure 5
Equity premium forecasts for the mean combining method, historical average, and kitchen sink model, 1965:1–2005:4
The solid (dotted, dashed) line corresponds to the mean combining method (historical average, kitchen sink model) forecast.

- The volatility of kitchen sink model forecasts (horizontal dashed line) is much greater than that of mean combination forecasts (solid line), and even greater than the volatility of univariate regression model forecasts.
- The data-generating process for equity premiums is persistent and complex, with individual variables providing accurate signals in some periods but many false signals in other periods. This confounds the estimation of the unrestricted kitchen sink model and severely impairs its forecasting ability.

Real Economy Mechanism Discussion

4

SJTU

4.1 How Business Cycles Affect Combination Forecast Volatility

- Figure 6 depicts the mean combination forecast of equity premiums, along with vertical lines representing NBER-dated business cycle peaks and troughs.

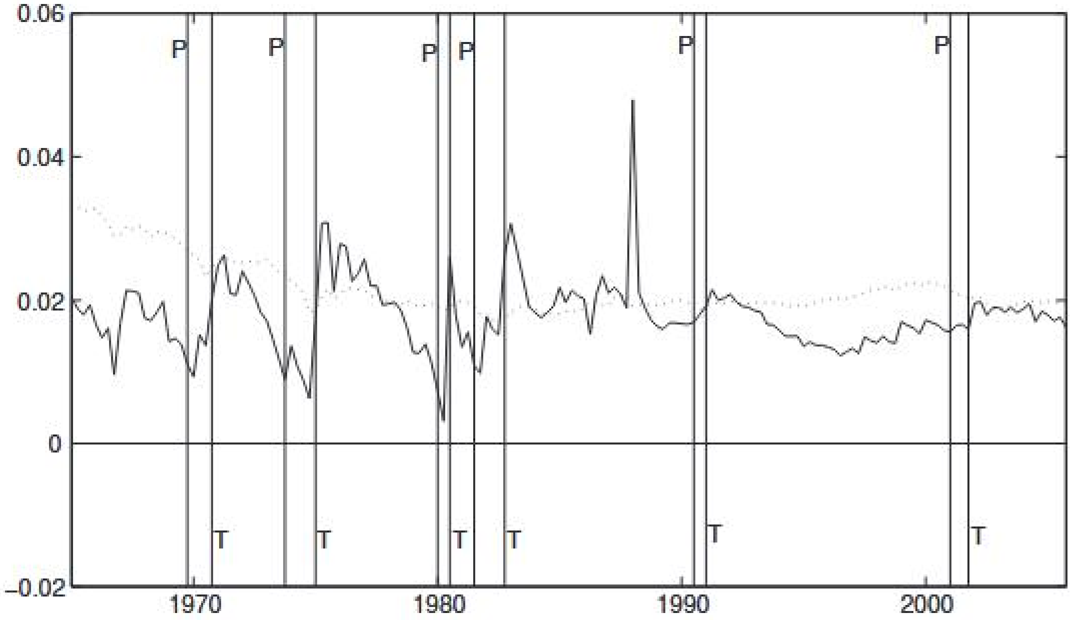

Figure 6
Equity premium forecasts for the mean combining method and NBER-dated business-cycle turning points, 1965:1–2005:4
The solid (dotted) line delineates the mean combination (historical average) forecast. Vertical lines indicate NBER-dated business-cycle peaks (P) and troughs (T).

- Pronounced upward spikes in combination forecasts occur at or shortly after troughs associated with four relatively deep recessions in the 1970s-early 1980s. Generally, equity premium forecasts decline during expansions and rise sharply during recessions.
- Equity premium forecasts generated by combination methods are closely related to NBER business cycle phases, and the behavior of the forecasts is consistent with Fama and French's (1989) and Cochrane's (1999, 2007) explanations of equity premium predictability.

4.1 How Business Cycles Affect Combination Forecast Volatility

- Many individual predictive regression models produce equity premium forecasts ranging approximately from 0.06 to 0.10, implying annual equity risk premiums between 24% and 40%, which seems implausibly large. Moreover, many individual model forecasts exhibit negative premiums in some periods, sometimes dropping to -20% annually.
- From an economic perspective, although forecasts from individual predictive regression models frequently exhibit economically implausible volatility, historical average forecasts based on constant expected equity premium models appear overly smooth. The "problem" with historical average forecasts is that they ignore business cycle fluctuations and thus fail to incorporate meaningful macroeconomic information.
- Combination methods produce a highly reasonable time-varying out-of-sample measure of equity risk premiums by stabilizing individual predictive regression model forecasts and better linking them to the business cycle. Figures 3 and 6 show that combination methods include relevant macroeconomic information omitted by historical average forecasts while avoiding the implausible volatility in equity risk premiums associated with individual predictive regression models.

4.2 Growth Forecasts in Different Growth Periods

Table 4
Correlations between equity premium forecasts and growth rates in three macroeconomic variables, 1965:1–2005:4

Combining method	Real GDP growth	Real profit growth	Real net cash flow growth
(1)	(2)	(3)	(4)
Mean	0.28***	0.35***	0.34***
Median	0.17**	0.24***	0.23***
Trimmed mean	0.31***	0.36***	0.35***
DMSPE, $\theta = 1.0$	0.28***	0.35***	0.34***
DMSPE, $\theta = 0.9$	0.34***	0.36***	0.36***

This table reports correlation coefficients for the equity premium combination forecast given in the row heading and macroeconomic variable growth rate given in the column heading. ** and *** indicate significance at the 5% and 1% levels, respectively.

- Table 4 shows that combination forecasts are positively and significantly correlated with growth rates of three macroeconomic variables: real GDP, real profits, and real net cash flows, with most correlations approaching or exceeding 0.30. These correlations complement the evidence in Figure 6 and further indicate that combination forecasts are related to the real economy.

4.2 Growth Forecasts in Different Growth Periods

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

Table 5
R^2_{OS} statistics for out-of-sample equity premium combination forecasts during good, normal, and bad growth periods, 1965:1–2005:4

Combining method	Forecast horizon: one quarter				Forecast horizon: four quarters			
	Overall	Good	Normal	Bad	Overall	Good	Normal	Bad
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Sorting on real GDP growth								
Mean	3.58***	1.82	1.71	6.17***	8.19***	3.07	3.63*	11.58***
Median	3.04***	2.67**	0.39	5.02***	6.99***	12.74***	6.35**	5.23***
Trimmed mean	3.51***	2.25*	1.24	5.94***	8.13***	5.41*	4.01*	10.63***
DMSPE, θ = 1.0	3.54***	1.71	1.56	6.26***	7.87***	2.32	3.15	11.46***
DMSPE, θ = 0.9	3.49***	1.60	1.36	6.33***	5.96***	4.71*	0.27	8.27***
Panel B. Sorting on real profit growth								
Mean	3.58***	2.87*	−1.03	7.94***	8.19***	0.93	4.89*	14.72***
Median	3.04***	2.56**	0.21	5.74***	6.99***	1.14	8.00**	10.18***
Trimmed mean	3.51***	2.85*	−0.67	7.47***	8.13***	1.74	5.83**	13.55***
DMSPE, θ = 1.0	3.54***	2.74*	−1.21	8.08***	7.87***	0.16	4.41	14.78***
DMSPE, θ = 0.9	3.49***	2.51	−1.56	8.40***	5.96***	−4.28	2.00	14.70***
Panel C. Sorting on real net cash flow growth								
Mean	3.58***	5.44**	2.17*	4.63**	8.19***	3.29*	8.81***	11.42***
Median	3.04***	4.12***	1.80**	4.25**	6.99***	4.99***	6.17**	9.48***
Trimmed mean	3.51***	5.01**	2.36**	4.47**	8.13***	4.39**	9.13***	10.04***
DMSPE, θ = 1.0	3.54***	5.51**	2.13*	4.52**	7.87***	2.97*	8.50***	11.09***
DMSPE, θ = 0.9	3.49***	5.88**	1.84*	4.15*	5.96***	0.53	6.66**	9.56***

Note: In addition to the basic single-quarter combination predictive regression model, the study also extends r_{t+1} to a four-period sum, obtaining cumulative forecasts over four quarters as specified by the following model:

$$r_{t+1:t+4} = \alpha_i + \beta_i x_{i,t} + \varepsilon_{t+1:t+4},$$

- The study divides the sample into three periods based on the ranking of growth in real GDP, real profits, and real net cash flows, and extends the calculation of the R^2_{OS} statistic for each period using the combination forecast model. The results are shown in Table 5.
- Out-of-sample gains from combination forecasts are generally concentrated in extreme periods, especially during low-growth periods. Whether considering single-quarter or four-quarter windows, R^2_{OS} in low-growth periods is always higher compared to normal growth periods; within a single quarter, R^2_{OS}during high-growth periods is higher than during normal growth, whereas evidence of increased out-of-sample gains in high-growth four-quarter periods is limited.
- Conclusion: During extreme (especially low-growth) periods, combination forecasts deliver heightened out-of-sample gains over the historical mean forecast.

4.3 ARDL - 15 Economic Variables Combination Forecasts for Macroeconomic Growth

- Stock and Watson (2003) showed that composite forecasts of output growth consistently outperform autoregressive (AR) benchmark models. This provides a possible explanation for the out-of-sample gains associated with combination forecasts of equity premiums: individual economic variables fail to consistently generate out-of-sample gains in equity premium forecasting because they generate unstable gains in forecasting macroeconomic volatility; in contrast, forecast combinations produce consistent out-of-sample gains for equity premium forecasting because they also produce stable gains in forecasting macroeconomic volatility.

4.3 ARDL - 15 Economic Variables Combination Forecasts for Macroeconomic Growth

- ARDL Model: $y_{t+1} = \zeta_i + \eta_i y_t + \lambda_i x_{i,t} + v_{t+1}$
- y_{t+1} denotes the growth rate of the three macroeconomic variables at t+1, and is similarly extended to the cumulative sum over four periods.

Table 6
Macroeconomic variable out-of-sample forecasting results for individual models and combining methods

Predictor	Forecast horizon: one quarter						Forecast horizon: four quarters					
	1965:1–2005:4 out-of-sample period			1976:1-2005:4 out-of-sample period			1965:1–2005:4 out-of-sample period			1976:1–2005:4 out-of-sample period		
	Real GDP growth	Real profit growth	Real net cash flow growth	Real GDP growth	Real profit growth	Real net cash flow growth	Real GDP growth	Real profit growth	Real net cash flow growth	Real GDP growth	Real profit growth	Real net cash flow growth
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel A. Individual predictive regression model forecasts												
D/P	−0.95	−1.34	−0.79	−1.63	−1.02	−2.61	−3.98	−4.95	−4.84	−3.68	−7.39	−5.02
D/Y	−1.22	−1.15	−0.83	−0.83	−1.52	−1.11	−4.53	−4.56	−3.86	−5.17	−9.95	−2.85
E/P	−0.29	−1.08	0.62	−2.62	−0.07	−1.62	−3.73	−5.31	−0.55	−4.98	−1.68	−0.19
D/E	−0.93	−1.63	−3.66	−6.27	−2.11	−4.50	−0.70	−9.97	−9.31	−11.45	−4.02	−12.19
$SVAR$	−36.57	−7.35	−38.45	−69.76	−6.15	−58.36	−21.90	−15.86	−31.07	−43.96	−19.36	−37.46
B/M	−1.67	−1.82	−1.41	−6.52	−1.50	−2.20	−9.02	−9.09	−5.46	−8.71	−18.78	−6.01
$NTIS$	−0.56	−2.76	−2.20	0.87*	−2.09	−0.67	−6.38	−8.10	−8.95	−4.55	−12.01	−1.79
TBL	−0.93	−0.39	0.17*	−1.59	−2.32	1.63*	0.03**	−11.35	2.23**	−2.43	−14.67	10.61**
LTY	−2.28	−2.70	−1.80	−2.59	−1.70	−0.58	−10.06	−15.68	−9.80	−11.60	−8.78	−1.97
LTR	−19.75	−12.25	−11.18	−19.70	−1.31	−7.09	0.08	−1.39	−2.65	0.78*	−8.02	−11.40
TMS	−5.34	−1.90	−2.34	−9.74	−12.45	−6.63	−9.16	−10.90	−24.18	−33.00	−40.18	−40.72
DFY	−4.77	−2.98	−3.70	−1.80	−2.82	−2.28	−13.94	−2.97	−17.30	−14.45	−5.13	−21.99
DFR	0.55**	−4.08	−2.11	−3.70	−4.23	−3.88	2.13**	0.81**	4.62***	0.69**	2.62**	4.78***
$INFL$	6.32***	−2.71	3.20**	4.64**	0.46	4.18**	15.00***	−8.83	6.10**	11.09**	−6.54	10.13**
I/K	−10.93	5.34***	−10.02	−0.38	3.58***	0.14*	−38.79	17.17***	−67.10	1.21**	15.96***	2.84**
Panel B. Combination forecasts												
Mean	4.48***	2.65**	3.08**	3.47*	2.02*	1.65	10.08**	7.51***	12.05***	11.52**	7.25***	10.85***
Median	4.63***	0.56	2.86**	4.32*	0.78	2.45*	4.45*	0.00	6.37***	4.90*	0.55	7.45***
Trimmed mean	4.28**	2.49**	3.62***	3.29*	1.78*	2.69**	8.74**	4.88**	10.58***	8.96*	4.78***	10.00***
DMSPE, $\theta = 1.0$	4.52***	2.76**	3.11**	3.62*	2.15*	1.70	11.27**	8.36***	13.01***	13.98**	7.99***	11.56***
DMSPE, $\theta = 0.9$	4.35**	2.95**	2.80**	3.43*	1.77	1.45	11.29**	5.23**	10.59***	11.79*	5.39***	10.01***

This table reports the modified Campbell and Thompson (2008) R^2_{OS} statistic (in percent) comparing forecasts from the competing forecasting model given in the row heading to the AR benchmark forecasting model. Statistical significance for the R^2_{OS} statistic is based on the p-value for the Clark and West (2007) out-of-sample *MSPE-adjusted* statistic; the statistic corresponds to a one-sided test of the null hypothesis that the competing forecasting model given in the row heading has equal expected square prediction error relative to the AR benchmark forecasting model against the alternative hypothesis that the competing forecasting model has a lower expected square prediction error than the AR benchmark forecasting model. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

- Panel A shows that single economic variables generally do not outperform the AR benchmark model, and in many cases their R^2_{OS} values are negative or close to zero.
- Panel B demonstrates that combination forecasts almost always generate significant and substantial out-of-sample gains for all three macroeconomic variables. This indicates that combining economic variables can substantially improve the ability to forecast macroeconomic volatility, which in turn enhances the out-of-sample predictability of equity premiums.

4.4 Structural Breaks in Macroeconomic Relationships

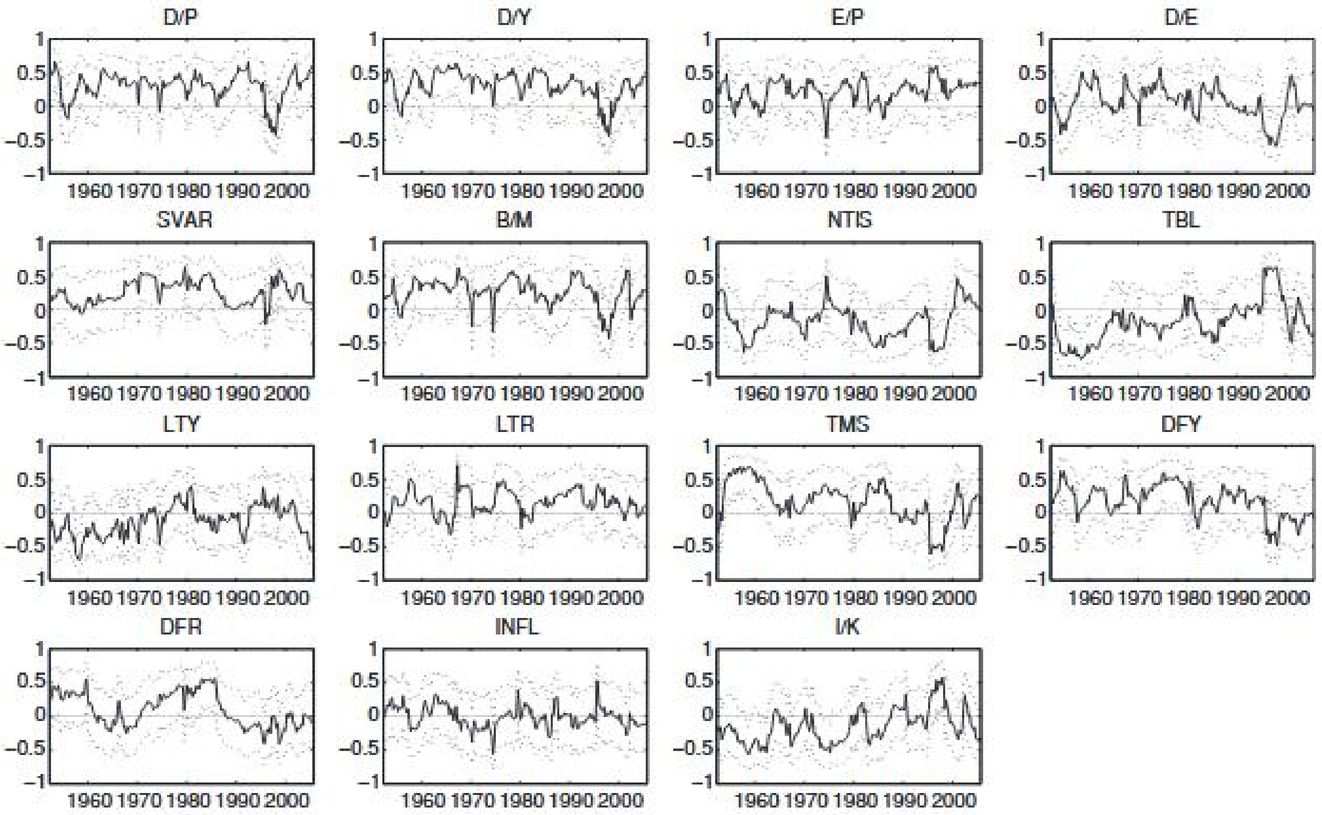

Figure 7
Correlations between the equity premium and individual predictors based on 10-year rolling windows
The date on the horizontal axis gives the end date of the 10-year period. Dotted lines indicate 95% confidence intervals.

- Figure 7 depicts the changing nature of relationships between equity premiums and individual economic variables during ten-year moving window periods between March 1947 and April 2005. Correlations fluctuate substantially during the postwar period, and there are many instances where correlations shift from being significant in some periods to being insignificant in other periods.
- Conclusion: Figure 7 adds to recent empirical evidence on structural breaks in individual equity premium predictive regression models, indicating important structural instabilities in the relationship between equity premiums and the 15 economic variables, and that forecast combinations can improve the performance of individual forecasting models under structural breaks. (Hendry and Clements 2004, Timmermann 2006)

4.4 Structural Breaks in Macroeconomic Relationships

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

- Real GDP Growth Model: $y_{t+1} = \zeta_i + \lambda_i x_{i,t} + \upsilon_{t+1}$,
- Bai and Perron (1998) U Dmax and W Dmax (10%) statistics are used to test for breakpoints in the model, and the $F(l+1/l)$ statistic is checked to determine the number of breaks. Table 7 reports the U Dmax and W Dmax (10%) statistics and estimates the break dates based on economic variables.

Table 7
Bai and Perron (1998) multiple structural break test results for real GDP growth predictive regression models and Chow test results for corresponding equity premium predictive regression models, 1947:3–2005:4

	Bai and Perron (1998) statistics		Bai and Perron (1998) break dates						Chow test
Predictor (1)	*U Dmax* (2)	*W Dmax*(10%) (3)	1st break (4)	2nd break (5)	3rd break (6)	4th break (7)	5th break (8)	6th break (9)	χ^2-statistic (10)
D/P	15.34**	23.54*	1953:2	1959:2	1966:1	1975:1	1985:3	1999:4	44.87***
D/Y	12.16*	18.71*	1953:2	1966:1	1975:1	1985:3	1999:4	–	44.20***
E/P	70.61***	95.35*	1953:2	1959:2	1975:1	1985:3	–	–	16.41**
D/E	14.10**	15.50*	1984:1	–	–	–	–	–	2.83
SVAR	123.89***	136.24*	1955:4	1962:1	1968:2	1982:4	–	–	26.84***
B/M	18.93***	26.26*	1953:2	1974:3	1984:2	–	–	–	13.36**
NTIS	15.65**	25.20*	1956:4	1962:4	1970:4	1982:4	1990:1	–	16.32*
TBL	11.17**	19.04*	1958:3	–	–	–	–	–	7.46**
LTY	18.33***	28.09*	1958:2	1982:4	–	–	–	–	6.58
LTR	34.79***	38.25*	1966:1	–	–	–	–	–	5.04*
TMS	18.48***	20.05*	1953:1	1959:4	1966:1	1980:2	–	–	42.55*
DFY	9.59	14.03*	1970:4	–	–	–	–	–	3.17
DFR	16.48***	20.15*	1984:3	–	–	–	–	–	5.28*
INFL	15.93**	23.53*	1953:1	1961:1	1975:1	1981:1	1999:4	–	12.90
I/K	22.56***	27.98*	1957:3	1963:2	–	–	–	–	4.96

This table reports Bai and Perron (1998) multiple structural break test results for real GDP growth predictive regression models. Columns (2) and (3) report the *U Dmax* and *W Dmax*(10%) statistics, respectively, corresponding to a one-sided (upper-tail) test of the null hypothesis of zero breaks against the alternative hypothesis of one to eight breaks. Columns (4)–(9) report the break dates estimated by the Bai and Perron (1998) procedure. Column (10) reports the Chow test χ^2-statistic for corresponding equity premium predictive regression models, where the break dates tested in the Chow test are the Bai and Perron (1998) break dates for the corresponding real GDP growth predictive regression model. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

- For 14 of the 15 individual models, the U Dmax and W Dmax (10%) statistics in the second and third columns are significant at conventional levels.
- Many breaks occur near the mid-1970s, corresponding to oil shocks, and in the mid-1980s, shortly after changes in Federal Reserve operating procedures.
- About half of the models experienced breaks in the early to mid-1950s, close to the Treasury-Federal Reserve Accord, and accompanying the transition from the wartime economy.

4.4 Structural Breaks in Macroeconomic Relationships

- The tenth column of Table 7 reports the χ^2 statistic corresponding to the application of the Chow test to individual equity premium predictive regression models. The χ^2 is significant for equity premium predictive regression models based on 10 out of 14 economic variables, which also present significant U Dmax and W Dmax (10%) statistics for real GDP growth predictive regression models.

	Bai and Perron (1998) statistics		Bai and Perron (1998) break dates						Chow test
Predictor (1)	$UDmax$ (2)	$WDmax$(10%) (3)	1st break (4)	2nd break (5)	3rd break (6)	4th break (7)	5th break (8)	6th break (9)	χ^2-statistic (10)
D/P	15.34**	23.54*	1953:2	1959:2	1966:1	1975:1	1985:3	1999:4	44.87***
D/Y	12.16*	18.71*	1953:2	1966:1	1975:1	1985:3	1999:4	–	44.20***
E/P	70.61***	95.35*	1953:2	1959:2	1975:1	1985:3	–	–	16.41**
D/E	14.10**	15.50*	1984:1	–	–	–	–	–	2.83
$SVAR$	123.89***	136.24*	1955:4	1962:1	1968:2	1982:4	–	–	26.84***
B/M	18.93***	26.26*	1953:2	1974:3	1984:2	–	–	–	13.36**
$NTIS$	15.65**	25.20*	1956:4	1962:4	1970:4	1982:4	1990:1	–	16.32*
TBL	11.17**	19.04*	1958:3	–	–	–	–	–	7.46**
LTY	18.33***	28.09*	1958:2	1982:4	–	–	–	–	6.58
LTR	34.79***	38.25*	1966:1	–	–	–	–	–	5.04*
TMS	18.48***	20.05*	1953:1	1959:4	1966:1	1980:2	–	–	42.55*
DFY	9.59	14.03*	1970:4	–	–	–	–	–	3.17
DFR	16.48***	20.15*	1984:3	–	–	–	–	–	5.28*
$INFL$	15.93**	23.53*	1953:1	1961:1	1975:1	1981:1	1999:4	–	12.90
I/K	22.56***	27.98*	1957:3	1963:2	–	–	–	–	4.96

- Conclusion: Major structural breaks in macroeconomic relationships frequently correspond to significant simultaneous breaks in equity premium predictive regression models, further demonstrating that combination forecasts of equity premiums are linked to the real economy.