Università Commerciale
Luigi Bocconi

Lecture 2: Forecasting stock returns (part 2)

Prof. Massimo Guidolin

Advanced Financial Econometrics III

Winter/Spring 2019

Overview

- Predictability in a Bayesian regression framework
- Time variation in predictability
- The economic value of time-varying predictability
- Time-varying predictability and the business cycle
- Decomposing the sources of predictability

The Key Point

- Not only the evidence in favor of time-varying predictive coefficients is strong (in a Bayesian set up), but this also leads to precisely estimated and persistent economic value

- Dangl and Halling (2012, JFE) evaluate predictive regressions that explicitly consider the time-variation of coefficients in a comprehensive Bayesian framework
  - Time-variation and OOS predictability carry a business cycle meaning
  - Several reasons coefficients might vary over time, e.g., due to changes in regulatory conditions, market sentiment, in monetary policies, in the institutional framework, or in macroeconomic interrelations
- For monthly S&P 500 returns, they report statistical and economic evidence of OOS predictability relative to using the historical mean
- Predictive regression coefficients follow a random walk, which gives the maximum degree of instability; in an appendix they verify that RW thetas outperform stationary AR models in OOS terms
- A Bayesian model averaging (BMA) approach allows to address both parameter uncertainty and the choice (inclusion) of predictors

The Econometric Framework

- The benchmark models with constant coefficients used in the paper are equal to ordinary least squares (OLS) regressions with an extending window
  - In an appendix they show that their methodology also dominates rolling-window OLS regressions
- For each degree of time-variation of coefficients, estimate 2^k-1 dynamic linear models from all possible combinations of predictive variables

$$r_{t+1} = X_t'\theta_t + v_{t+1}, \quad v \sim \mathrm{N}(0,V) \quad \text{(observation equation)}$$

$$\theta_t = \theta_{t-1} + \omega_t, \quad \omega \sim \mathrm{N}(0,W_t) \quad \text{(system equation)}$$

  - They are aware that this implies that coefficients follow a random walk and that they might drift to arbitrarily high or low values, hence causing returns to be non-stationary
  - Must impose some structure on the system equation but their results reveal that any deviation from the assumption of no predictability in the shocks on coefficients reduces the predictive power
  - If the system variance matrix W_t equals zero, the regression coefficients are constant over time: model nests the specification of constant regression coefficients

The Econometric Framework

- For the time $t = 0$ specification of the prior information, use a natural conjugate g-prior specification:

$$V \mid D_0 \sim IG[\tfrac{1}{2}, \tfrac{1}{2}S_0],$$

$$\theta_0 \mid D_0, \quad V \sim N[0, gS_0(X'X)^{-1}],$$

where

$$S_0 = \frac{1}{N-1} r'(I - X(X'X)^{-1}X')r.$$

- This is a diffused prior centered around the null-hypothesis of no-predictability
- g serves as the scaling factor that determines the confidence assigned to the null-hypothesis of no-predictability
- The forecast of the t+1 return (i.e., the predictive density) is found by integrating the conditional density of r_{t+1} over the support of θ and V and it turns out to be a Student t-distribution (see Hamilton, 1994)
- They add some ad-hoc structure to W_t to capture the fact that in periods of low system variance, the estimation error of the coefficient vector θ tends toward zero as the sample size increases
- A simple way to capture this relation is to assume that W_t is proportional to the variance/covariance matrix of θ with scaling factor is assumed to be $(1-\delta)/\delta$ with $0 < \delta_i \leq 1$
- West and Harrison (1997) call δ a discount factor: $\delta = 1 \Leftrightarrow$ to $W_t = 0$, i.e., to the assumption that the regression coefficients are constant

The Econometric Framework

- They consider values on a grid for δ ($\delta \in \{0.96, 0.97, 0.98, 0.99, 1.00\}$)
- A Bayesian updating approach integrates (sums) over the range of the discount factor
- Considering a number of *d* different discrete values of δ leads to a total of d $d(2^k - 1)$ possible dynamic linear models

- Use a Bayesian model selection to assign posterior probability weights across individual models that differ in selected variables
- The posterior probability of each of the $d(2^k - 1)$ models is updated month by month according to Bayes rule; i.e., based on the realized likelihood of the model's return prediction
- Dangl and Halling use the same data as Goyal and Welch (2008)
  - Exclude "investment to capital ratio", "percent equity issuing", and "consumption, wealth, income ratio", not available at a monthly frequency; "Dividend-to-price ratio" almost perfectly correlated to DY
  - Considerable emphasis on the dividend yield and the cross-sectional premium, the beta spread between high- and low-beta stocks according to Polk, Thompson, and Vuolteenaho (2006)

Main Findings: Predictive Accuracy

- As for the prior probabilities of individual models and model families, they start with an uninformed prior giving equal weight to each individual model (i.e., $1/(d \cdot (2^k - 1))$) + robustness checks
- Predictive regressions with time-varying coefficients predict market returns significantly better than the unconditional mean and perform significantly better than regressions with constant coefficients

Sample period: 1947+

Predictive model	Models incl. TVar-Coeff.		Models excl. TVar-Coeff.		Comp
	Diff. in MSPE	*p*-Value	Diff. in MSPE	*p*-Value	*p*-Value
BMA-Model	0.0192	$\boxed{0.00}$	0.0046	0.08	$\boxed{0.00}$
ep	0.0055	0.02	0.0026	0.04	0.12
svar	0.0004	0.41	−0.0005	0.95	0.30
bmr	0.0031	0.21	0.0031	0.14	0.05
tbl	0.0055	0.02	0.0033	0.08	0.01
ltr	0.0050	0.01	0.0018	0.14	0.01
dfy	0.0035	0.02	−0.0001	0.64	0.02
inf	0.0027	0.09	0.0009	0.16	0.18
dy	0.0044	0.01	0.0036	0.02	0.08
dpayr	0.0027	0.17	−0.0001	0.54	0.12
csp	0.0094	0.01	0.0052	0.02	0.01
ntis	0.0043	0.01	0.0020	0.10	0.04
lty	0.0035	0.08	0.0025	0.13	0.06
dfr	0.0008	0.16	−0.0005	0.81	0.02
MOST-Model	0.0054	0.10	0.0054	0.10	.
MEDIAN-Model	0.0220	0.00	0.0028	0.22	0.00

Clark-West (2006) test

Main Findings: Economic Value from Mean-Variance

- BMA shifts weights between models depending on historical performance
- MOST and MEDIAN models fix a selection of variables ⇒ any performance differences from these models btw including and excluding time-varying coefficients is related to the influence of time-varying coefficients

- They test whether the OOS predictability of monthly S&P 500 returns may support an investor who rationally uses the predictions for portfolio optimization
- An investor with a single-period horizon and mean-variance preferences determine monthly realized MV utility where we use daily returns to estimate monthly variance

- *BMA-Model incl. (excl.) TVar-Coeff.*: This model represents the Bayesian model average across all individual models including (excluding) models with time-varying coefficients.
- *Univariate Models incl. (excl.) TVar-Coeff.*: These models consider only one predictive variable at a time. In the cases in which we include time-varying coefficients, we still use Bayesian model averaging to average across models with different assumptions of the degree of time-variation of the coefficient.
- *MOST-Model incl. (excl.) TVar-Coeff.*: The MOST-Models represent the individual models that receive most posterior probability – among all individual models including (excluding) models with time-varying coefficients – at the end of the month before the evaluation period starts. Then we keep this model specification (the variable selection and degree of time-variation of the coefficients) constant during the evaluation period, but we update the coefficient estimates.
- *MEDIAN-Model incl. (excl.) TVar-Coeff.*: The MEDIAN-Model is determined in the following way. At the end of the month before the evaluation period starts, we identify all predictive variables that receive more than 50% posterior probability in the BMA-Model incl. TVar-Coeff. (the 50%-threshold is basically an ad hoc way to determine variables that show decent predictive performance). Then we focus on the model that includes these predictive variables. The *MEDIAN-Model incl. (excl.) TVar-Coeff.* is the model that includes these predictive variables and has *time-varying* (*constant*) coefficients.

Main Findings: Economic Value from Mean-Variance

Predictive model	Models incl. TVar-Coeff.				Models excl. TVar-Coeff.			
	1947+	1965+	1976+	1988+	1947+	1965+	1976+	1988+
BMA-Model	**2.57***	**5.75*****	**4.82*****	**1.76**	−1.97	0.16	−1.87	−5.79
ep	−0.61	1.07	−1.36	−1.77	−1.60	−0.22	−1.40	−1.79
svar	**−0.77**	**0.35**	−1.67	−2.95	−1.69	−0.46	−1.80	−2.97
bmr	−1.74	−0.71	−3.67	−3.27	−3.13	−2.07	−3.83	−6.41
tbl	−0.55	**0.63**	**−2.72**	**−4.04**	−1.47	0.14	−3.30	−4.66
ltr	**0.38**	1.60	−0.73	−2.80	−0.96	1.14	−0.82	−2.89
dfy	**0.01**	**1.36**	−1.01	−2.27	−1.46	−0.46	−1.67	−3.12
inf	−0.23	0.74	−2.05	−3.13	−0.75	0.51	−1.98	−3.14
dy	−0.77	0.74	−2.27	−4.84	−1.47	0.38	−2.43	−5.08
dpayr	**−0.82**	**0.36**	−2.47	−4.56	−1.81	−0.26	−2.58	−4.72
csp	**0.88**	**2.90***	**1.22**	**2.94**	−1.11	1.20	−1.71	−1.98
ntis	**−0.13**	1.16	**−1.13**	−3.15	−1.66	−0.05	−1.34	−3.22
lty	−1.03	0.37	−2.79	−3.82	−1.33	0.31	−2.84	−3.88
dfr	**−0.82**	**0.51**	−1.38	−3.69	−1.87	−0.46	−1.69	−3.41
MOST-Model	−2.30	**3.55***	**4.24****	**2.93**	−2.30	−2.44	−4.06	−4.36
MEDIAN-Model	**2.68***	**4.04****	**4.92****	**2.26**	−2.87	−3.06	−3.97	−3.31

Bold utility gains in columns 3–5 indicate that the models including time-varying coefficients perform significantly better than the models excluding time-varying coefficients at least at the 10% level.

- The BMA and the MEDIAN model show best performance, i.e., consistently positive and large utility gains, under time-varying coefficients
- These utility gains are statistically significant during all evaluation periods except the 1988+ period

Business Cycle Variation of OOS Predictability

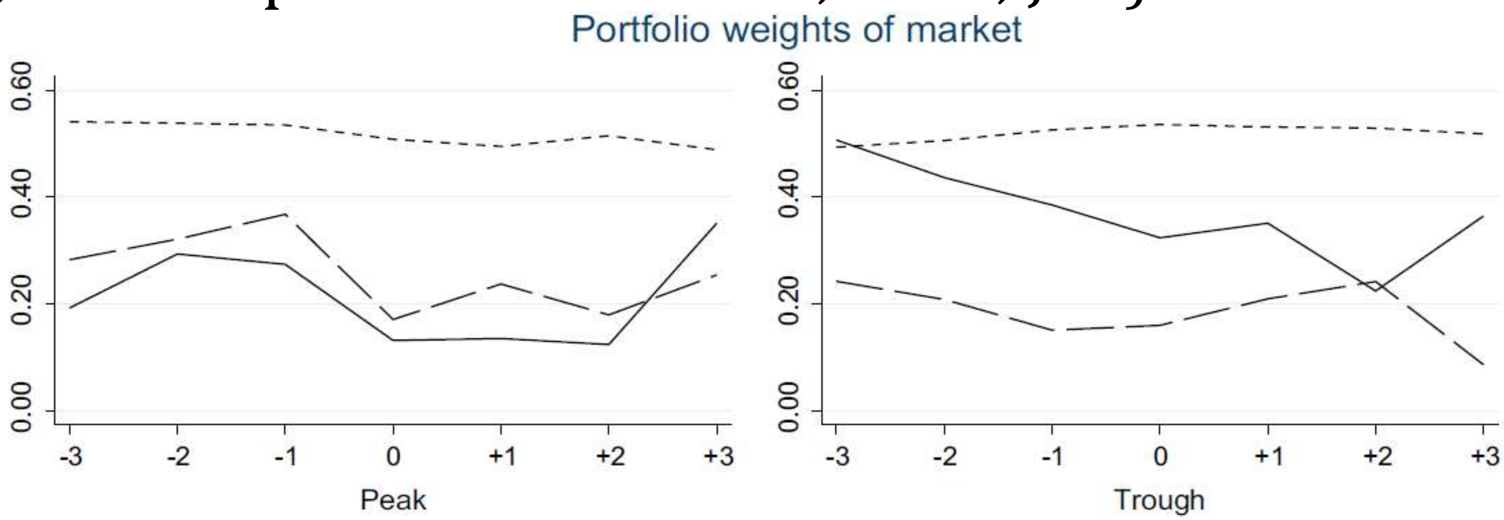
- Consistent with other papers (Henkel, Martin, and Nardari, 2011, JFE; Rapach, Strauss, and Zhou, 2010, RFS), they find significantly stronger evidence for predictability during recessions

Time period	Models incl. TVar-Coeff.		Models excl. TVar-Coeff.		Model comparison	
	Diff. MSPE	Util. gain	Diff. MSPE	Util. gain	Diff. MSPE	Util. gain
Rec.	0.0600***	24.967***	0.0168**	9.378*	0.0478***	15.589***
Exp.	0.0121**	1.622	0.0052	−1.990	0.0152***	3.613**
Diff.	−0.0479***	−23.345***	−.0116	−11.369**		
Late exp.	0.0327***	16.597*	0.0148	3.841	0.0184*	12.756
Early rec.	0.1156***	47.849***	0.0707***	30.543**	0.0401**	17.306**
Late rec.	0.0490**	6.421	−0.0092	−6.810	0.0752***	13.230
Early exp.	0.0075	−2.273	−0.0039	−8.321*	0.0266***	6.048

- Utility gains vs. no-predictability benchmark are huge in recessions
  - The no-predictability benchmark is overly optimistic about the equity premium and thus suffers from severe losses during recessions
- There is clear dominance of models with time-varying coefficients prevails during both recessions and expansions
- Significant levels of OOS predictability during expansions, but only for models including time-varying coefficients

Business Cycle Variation of OOS Predictability

- The asset allocation strategy of an investor relying on the BMA-Model incl. TVar-Coeff. seems to time the market very well
  - The investor withdraws from the market quickly at the beginning of a recession, and then moves back in towards the end of it
  - In contrast, an investor using predictions from the BMA-Model excl. TVar-Coeff. pulls out of the market after a peak but completely fails to move into it again towards the end of the recession
- The model is consistent with implications of asset pricing models that use time-varying risk aversion to generate time-varying risk premiums (e.g., see Campbell and Cochrane, 1999, JPE)
- Predictability reflects business cycle risk vs. market inefficiency



- Therefore, predictability is not driven away over time

Where Does Predictive Power Come From?

- Decompose prediction uncertainty into four components:

① The unexplained variance in the predictive relation

② The estimation uncertainty in coefficients

③ Model uncertainty w.r.t. the choice of predictive variables

④ Model uncertainty w.r.t. to time-variation in coefficients

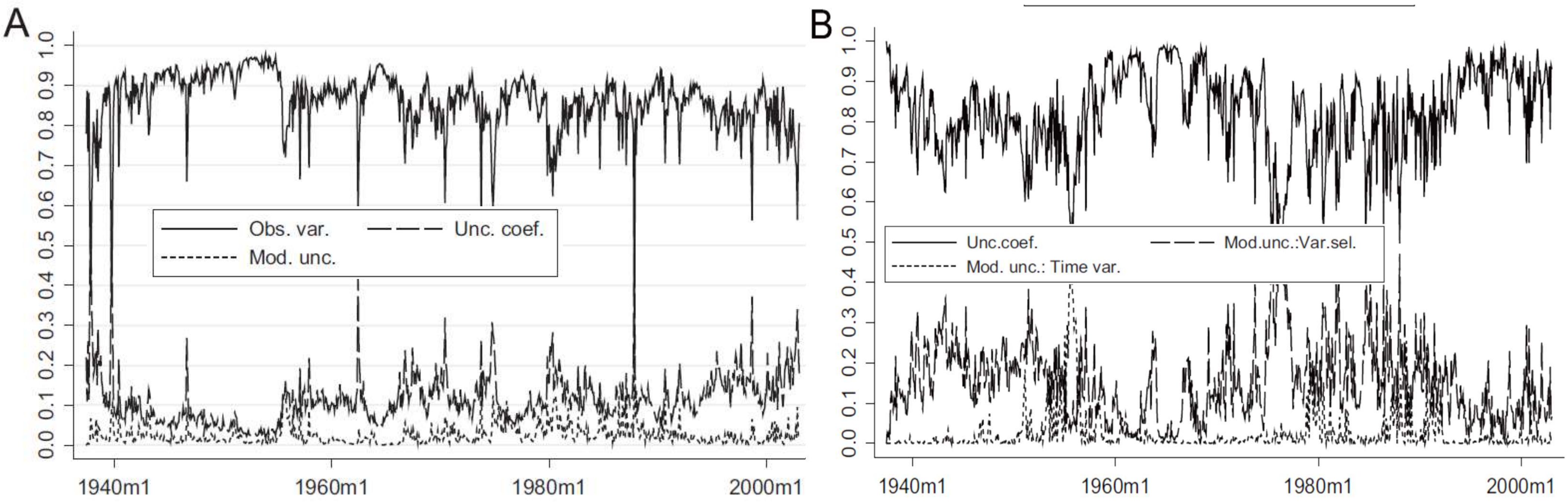
  - Starting with the decomposition with respect to different values of δ, we write $\mathrm{Var}(r) = \mathrm{E}_\delta(\mathrm{Var}(r|\delta)) + \mathrm{Var}_\delta(\mathrm{E}(r|\delta))$
  - The term $\mathrm{E}_\delta(\mathrm{Var}(r|\delta))$ represents the first three terms
  - In a second step, the term $\mathrm{E}_\delta(\mathrm{Var}(r|\delta))$ can be further decomposed into

$$\mathrm{Var}(r|\delta) = \mathrm{E}_M(\mathrm{Var}(r|M,\delta)) + \mathrm{Var}_M(\mathrm{E}(r|M,\delta))$$

$$\begin{aligned}
\mathrm{Var}(r_{t+1}) = & \sum_j \left[\sum_i (S_t | M_i, \delta_j | D_t) P(M_i | \delta_j, D_t) \right] P(\delta_j | D_t) \\
& + \sum_j \left[\sum_i (X_t' R_t X_t | M_i, \delta_j, D_t) P(M_i | \delta_j, D_t) \right] P(\delta_j | D_t) \\
& + \sum_j \left[\sum_i (\hat{r}^j_{t+1,i} - \hat{r}^j_{t+1})^2 P(M_i | \delta_j, D_t) \right] P(\delta_j | D_t) \\
& + \sum_j (\hat{r}^j_{t+1} - \hat{r}_{t+1})^2 P(\delta_j | D_t).
\end{aligned}$$

- The dominant source of uncertainty is observational variance

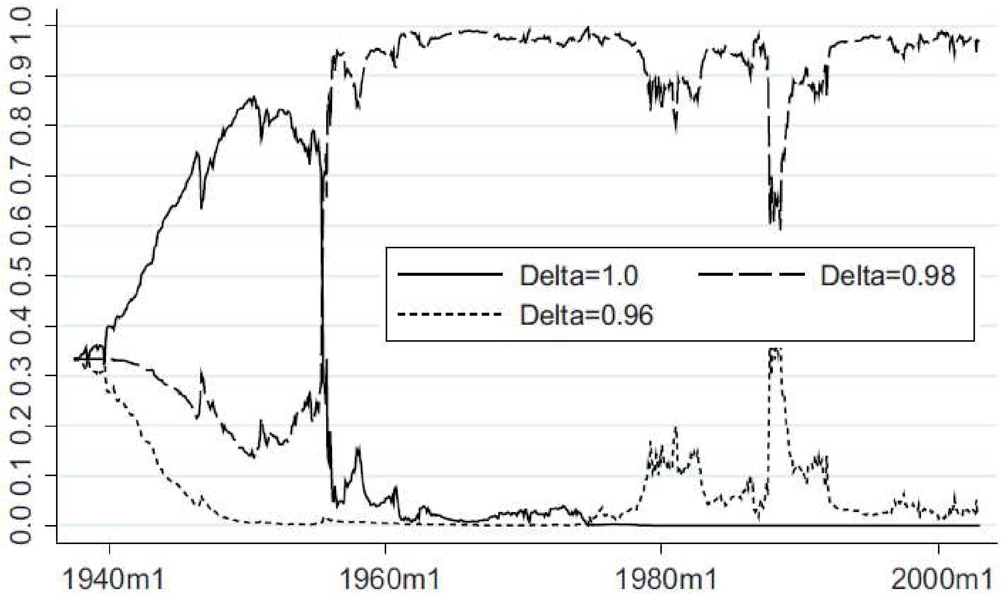
Where Does Predictive Power Come From?



- The estimation uncertainty in coefficients captures more than half of the remaining variance
- In periods of stress, model uncertainty peaks
- Uncertainty about the correct degree of time-variation (δ) is, in general, relatively low except for individual periods
- Constant-coefficient models (i.e., $\delta = 1$) perform well over the first 15 years but lose support from the data in and after 1955

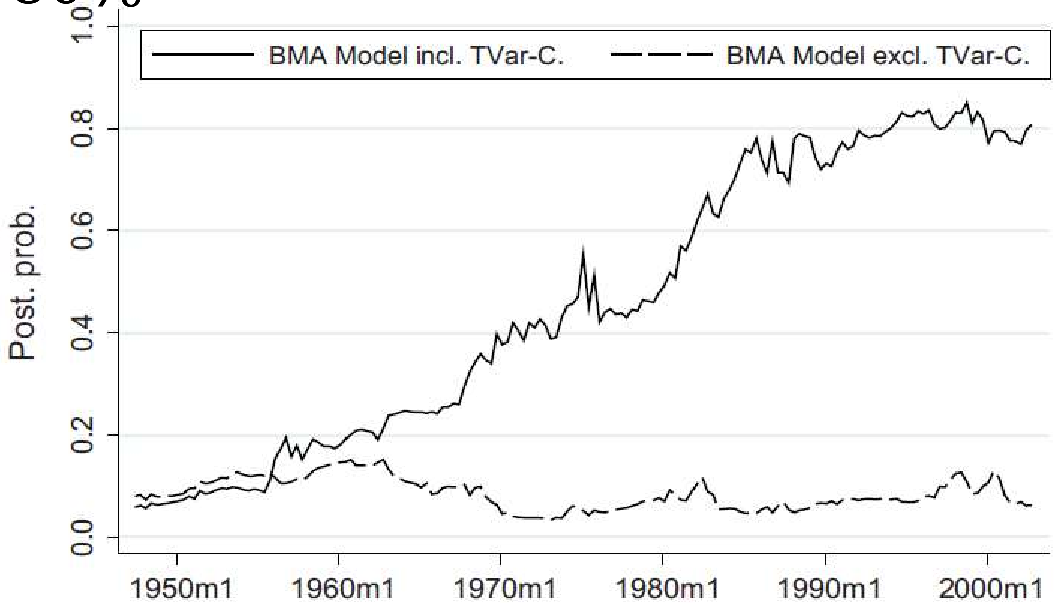
Where Does Predictive Power Come From?

- In the case of the BMA-Model excluding time-varying coefficients, the posterior prob. assigned to the Top-10 models does not account for more than 7% at the end of the sample and never exceeds 16%



- The posterior prob. assigned to the Top-10 models of the BMA Model including time-varying coefficients increases to more than 80%

Predictive variable	Models incl. TVar-Coeff.			
	1964.12	1975.12	1987.12	2003.1
dy	0.94	0.80	0.93	0.91
ep	0.25	0.16	0.46	0.77
dpayr	0.38	0.21	0.58	0.78
svar	0.17	0.07	0.01	0.01
csp	0.99	1.00	1.00	1.00
bmr	1.00	1.00	1.00	1.00
ntis	0.25	0.08	0.13	0.40
tbl	0.37	0.26	0.95	0.71
lty	0.41	0.70	0.08	0.37
ltr	0.27	0.21	0.19	0.12
dfy	0.12	0.06	0.01	0.18
dfr	0.06	0.02	0.13	0.03
inf	0.31	0.19	0.07	0.03



Some Robustness Checks: Predictive Performance

Sample period: 1947+

Predictive model	Models incl. TVar-Coeff. Diff. in MSPE	Models incl. TVar-Coeff. *p*-Value	Models excl. TVar-Coeff. Diff. in MSPE	Models excl. TVar-Coeff. *p*-Value	Comp *p*-Value
BMA-Model	0.0192	0.00	0.0046	0.08	0.00
ep	0.0055	0.02	0.0026	0.04	0.12
svar	0.0004	0.41	−0.0005	0.95	0.30
bmr	0.0031	0.21	0.0031	0.14	0.05
tbl	0.0055	0.02	0.0033	0.08	0.01
ltr	0.0050	0.01	0.0018	0.14	0.01
dfy	0.0035	0.02	−0.0001	0.64	0.02
inf	0.0027	0.09	0.0009	0.16	0.18
dy	0.0044	0.01	0.0036	0.02	0.08
dpayr	0.0027	0.17	−0.0001	0.54	0.12
csp	0.0094	0.01	0.0052	0.02	0.01
ntis	0.0043	0.01	0.0020	0.10	0.04
lty	0.0035	0.08	0.0025	0.13	0.06
dfr	0.0008	0.16	−0.0005	0.81	0.02
MOST-Model	0.0054	0.10	0.0054	0.10	.
MEDIAN-Model	0.0220	0.00	0.0028	0.22	0.00

- To cut execution times, I have used the 7 stronger predictors as reported in the original paper based on OLS performance: DY, EP, Payout ratio, LTY, LTR, DFY, DFR, lagged inflation

	MSPE for Historical Mean	Best OLS Predictor	Best TVP Bayesian Predictor	Bayesian Model Average	Recession BMA	Expansion BMA	Bayesian Model Average with TVP	Recession BMA with TVP	Expansion BMA with TVP
Baseline: 1967-2017	0.0019788	0.0037962 (LTR) CW p-val: 0.061	0.0039134 (Payout) CW p-val: 0.251	0.0019204 CW p-val: 0.062	0.0032261 CW p-val: 0.015	0.0016796 CW p-val: 0.732	0.0019147 CW p-val: 0.140	0.0034535 CW p-val: 0.237	0.0016309 CW p-val: 0.196

Some Robustness Checks: Predictive Performance

	MSPE for Historical Mean	Best OLS Predictor	Best TVP Bayesian Predictor	Bayesian Model Average	Recession BMA	Expansion BMA	Bayesian Model Average with TVP	Recession BMA with TVP	Expansion BMA with TVP
Baseline: 1967-2017	0.0019788	0.0037962 (LTR) CW p-val: 0.061	0.0039134 (Payout) CW p-val: 0.251	0.0019204 CW p-val: 0.062	0.0032261 CW p-val: 0.015	0.0016796 CW p-val: 0.732	0.0019147 CW p-val: 0.140	0.0034535 CW p-val: 0.237	0.0016309 CW p-val: 0.196
Longer: 1947-2017	0.0017373	0.002947 (log DY) CW p-val: 0.047	0.0048197 (LTY) CW p-val: 0.156	0.0016411 CW p-val: 0.073	0.0028954 CW p-val: 0.577	0.0013954 CW p-val: 0.018	0.0019147 CW p-val: 0.140	0.0034535 CW p-val: 0.237	0.0016309 CW p-val: 0.196
γ =1 (CRRA)	0.0019788	0.0037962 (LTR) CW p-val: 0.061	0.0039134 (Payout) CW p-val: 0.251	0.0019204 CW p-val: 0.062	0.0032261 CW p-val: 0.015	0.0016796 CW p-val: 0.732	0.0016305 CW p-val: 0.51734	0.0027881 CW p-val: 0.15612	0.0015252 CW p-val: 0.63235
γ =5 (CRRA)	0.0019788	0.0037962 (LTR) CW p-val: 0.061	0.0039134 (Payout) CW p-val: 0.251	0.0019204 CW p-val: 0.062	0.0032261 CW p-val: 0.015	0.0016796 CW p-val: 0.732	0.0019147 CW p-val: 0.140	0.0034535 CW p-val: 0.237	0.0016309 CW p-val: 0.196
δ =0.9; 0.925; 0.95; 0.975; 1	0.0019788	0.0037962 (LTR) CW p-val: 0.061	0.0039134 (Payout) CW p-val: 0.251	0.0019204 CW p-val: 0.062	0.0032261 CW p-val: 0.015	0.0016796 CW p-val: 0.732	0.0019508 CW p-val: 0.220	0.003433 CW p-val: 0.071	0.0016775 CW p-val: 0.649
g =20	0.0019788	0.0037962 (LTR) CW p-val: 0.061	0.0039302 (Payout) CW p-val: 0.250	0.0019204 CW p-val: 0.062	0.0032261 CW p-val: 0.015	0.0016796 CW p-val: 0.732	0.0019484 CW p-val: 0.298	0.0033566 CW p-val: 0.104	0.0016887 CW p-val: 0.680
g =90	0.0019788	0.0037962 (LTR) CW p-val: 0.061	0.0039302 (Payout) CW p-val: 0.250	0.0019204 CW p-val: 0.062	0.0032261 CW p-val: 0.015	0.0016796 CW p-val: 0.732	0.0018508 CW p-val: 0.111	0.0036371 CW p-val: 0.456	0.0015213 CW p-val: 0.088
Rolling window for variance = 60 months	0.0017373	0.002947 (log DY) CW p-val: 0.047	0.0048197 (LTY) CW p-val: 0.156	0.0016411 CW p-val: 0.073	0.0028954 CW p-val: 0.577	0.0013954 CW p-val: 0.018	0.0019147 CW p-val: 0.140	0.0034535 CW p-val: 0.237	0.0016309 CW p-val: 0.196
Absolute Loss Function	0.033734	0.51749 (LTY) CW p-val: 0.396	0.68653 (Payout) CW p-val: 0.299	0.029759 CW p-val: 0.690	0.039091 CW p-val: 0.224	0.028038 CW p-val: 0.923	0.025886 CW p-val: 0.931	0.035573 CW p-val: 0.426	0.0241 CW p-val: 0.978
Recursive, 30-month Rolling Mean Benchmark	0.0020397	0.017105 (LTR) CW p-val: 0.001	0.012875 (DYF) CW p-val: 0.002	0.0019204 CW p-val: 0.062	0.0032261 CW p-val: 0.015	0.0016796 CW p-val: 0.732	0.001933 CW p-val: 0.225	0.003441 CW p-val: 0.225	0.0016554 CW p-val: 0.401

- Results seem to be rather fragile to changing the predictors (but one may argue these are too few) and the sample period
- It remains generally the case that there is more predictability in recessions than in expansions
- BMA TVP tends to outperform BMA but by a smaller margin

Some Robustness Checks: Predictive Performance

	Realized Utility for Historical Mean	Best OLS Realized Utility	Best TVP Bayesian Realized Utility	Bayesian Model Average	Recession BMA	Expansion BMA	Bayesian Model Average with TVP	Recession BMA with TVP	Expansion BMA with TVP
Baseline: 1967-2017	0.0041844	0.25082 (LTY)	0.22292 (LTY)	-0.042064	0.61261	-0.16603	-0.11285	0.84016	-0.20083
Longer: 1947-2017	0.0064946	0.1127 (LTY)	0.11591 (LTY)	-0.0060828	0.13832	-0.03524	-0.071446	1.0565	-0.29515
γ =1	0.0077718	0.1307 (LTR)	0.081402 (LTR)	-0.055846	0.41226	-0.14259	0.0014212	0.68061	-0.12427
γ =5	0.0042964	0.21031 (LTY)	0.16834 (LTY)	-0.046929	0.36937	-0.12624	0.039997	0.66553	-0.078972
δ =0.9; 0.925; 0.95; 0.975; 1	0.0041844	0.25082 (LTY)	0.22292 (LTY)	-0.042064	0.61261	-0.16603	0.058014	0.6354	-0.051926
g =20	0.0041844	0.25082 (LTY)	0.23032 (LTY)	-0.042064	0.61261	-0.16603	0.12807	1.6117	-0.15057
g =90	0.0041844	0.25082 (LTY)	0.288 (LTY)	-0.042064	0.61261	-0.16603	0.099238	0.30234	0.061787
Rolling window for variance = 60 months	0.0048769	0.17737 (LTY)	0.14398 (LTY)	-0.048673	0.39793	-0.13259	0.077633	0.92575	-0.081444
Absolute Loss Function	0.0041844	0.25082 (LTY)	0.22292 (LTY)	-0.042064	0.61261	-0.16603	-0.11285	0.84016	-0.20083
Recursive, 30-month Mean Benchmark	0.005337	0.13557 (LTY)	0.12259 (LTY)	-0.042064	0.61261	-0.16603	0.14147	1.2116	-0.060281

- Results concerning economic value are more in line with the findings of the paper, even though also individual predictors seem to generate risk-adjusted profits
- It is very hard to generate profits during expansions while risk-adjusted returns are massive during recessions
- The *g* parameter seem to affect economic value considerably

The Power of Forecast Combinations

- Rapach, Strauss, and Zhou (2010, RFS) argue that model uncertainty and instability seriously impair the forecasting ability of individual predictive regression models
- They recommend combining individual forecasts (from 15 variable) because it delivers statistically and economically significant OOS gains relative to the historical average, consistently over time
- Two sources: (i) combining incorporates information from numerous economic variables while substantially reducing forecast volatility; (ii) combination forecasts are linked to the real economy
- To see the intuition behind forecast combination, consider two predictive regression model forecasts: one based on the dividend yield and the other on the term spreads
  - By themselves, they could capture different components of business conditions, and give a number of "false signals" and/or imply an implausible equity risk premium during certain periods
  - If individual forecasts are weakly correlated, an average of the two forecasts should be less volatile and more reliable

The Power of Forecast Combinations

- Equity risk premium forecasts based on combinations are plausible, with distinct local maxima (minima) of the combination forecasts occurring very near NBER-dated business-cycle troughs (peaks)
  - Individual regressions ⇒ forecasts with implausibly large fluctuations, while the historical average ⇒ a forecast that is very "smooth," thereby ignoring fluctuations corresponding to business-cycle fluctuations
- Combination forecasts of the equity premium are also significantly correlated with future growth in a number of macroeconomic variables, including real GDP, real profits, and real net cash flows
- Once more, the OOS gains corresponding to combination forecasts are especially evident during bad growth periods
- The same combinations also generate consistent significant OOS gains when forming combination forecasts of real GDP, real profit, and real net cash flow growth
  - Cochrane (2007, RFS) emphasizes that equity premium forecasts are more plausibly related to macroeconomic risk if equity premium predictors can also forecast business cycles

Forecast Combination Schemes

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

- With one exception (the mean combination forecast) all of the combinations allow the weights to change at each *t*; however, it is typically desirable to have relatively stable combining weights

- The first class uses simple averaging schemes: mean, median, and trimmed mean)
  - The trimmed mean combination sets $\omega_{i,t} = 0$ for the individual forecasts with the smallest and largest values and $\omega_{i,t} = 1/(N - 2)$ for the remaining individual forecasts
- The second class of methods is based on Stock and Watson (2004, JEL), where the weights are functions of the historical performance of the individual models over the holdout OOS period
- Discounted MSPE:

Discount factor

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

Forecast Combination Schemes

- When θ = 1, there is no discounting ⇒ optimal combinations by Bates and Granger (1969, ORQ) for the case where the individual forecasts are uncorrelated (OLS is said not to perform well)
- When θ < 1, greater weight is attached to the recent forecast accuracy of the individual models; paper considers θ = 0.9, fast forgetting!

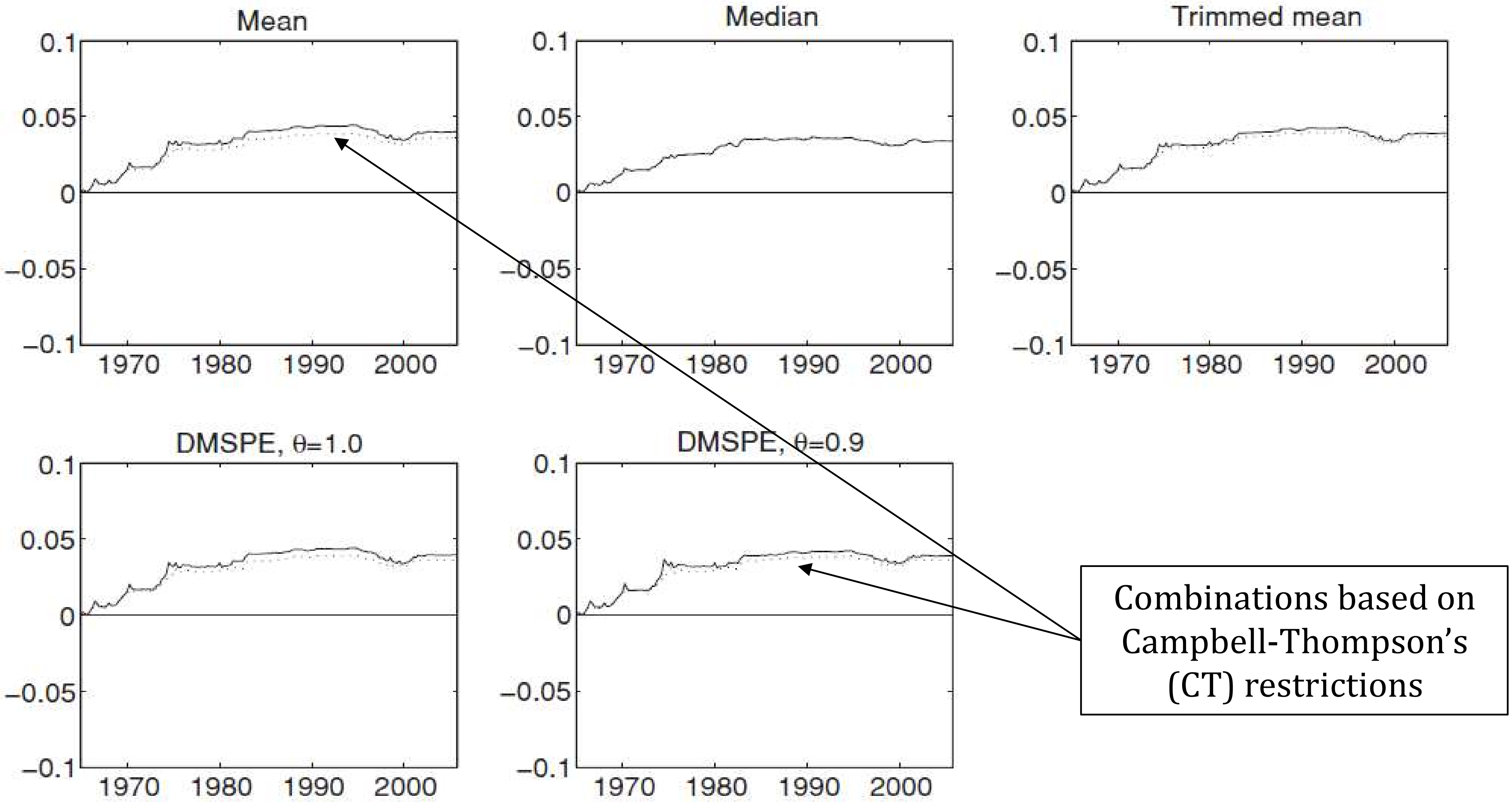
- Forecast evaluation is performed using Campbell and Thompson's (2008) OOS R^2:

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

- The OOS R-square measures the reduction in MSPE relative to the historical average forecast
- They also calculate realized utility gains for a mean-variance investor on a real-time basis
  - Investor estimates the variance using a 10-year rolling window
  - Following Campbell and Thompson (2008, RFS), constrain the weight on stocks to lie between 0% and 150% (inclusive) each month
  - The CER can be interpreted as the management fee that an investor would be willing to pay to have access to a model or combination

Main Results

- The data are standard from Goyal's web site, but because these are quarterly they also use the Investment-to-capital ratio, I/K, the ratio of aggregate (private nonresidential fixed) investment to aggregate capital for the entire economy (as in Cochrane, 1991, JF)



Main Results

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						

- Only I/K individually "opposes" some resistance to forecast combos
- Tests are implemented using Clark-West (2007, JoE) MSPE-difference test which adjusts standard Diebold Mariano's for nested models

Main Results

- While they allow for unequal and time-varying weights, the DMSPE combinations select weights relatively close to the 1/N rule over time
- Some evidence of deterioration of predictive power after the 1970s, but the combinations prove rather resilient

- Via encompassing tests, they demonstrate that combining incurporates useful information from a variety of economic variables

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

  - If $\lambda = 0$, then model *i* forecast encompasses model *j* forecast, as model j does not contain any useful information—beyond that already contained in model *I*
  - If $\lambda > 0$, then model *i* does not encompass model *j*, so model *j* does contain information useful beyond that already contained in model *i*

- Harvey, Leybourne, and Newbold (1998, JBES) develop a statistic to test the null hypothesis that model *i* forecast encompasses model *j* forecast (H_0: $\lambda = 0$) against the (one-sided) alternative $H_1 : \lambda > 0$

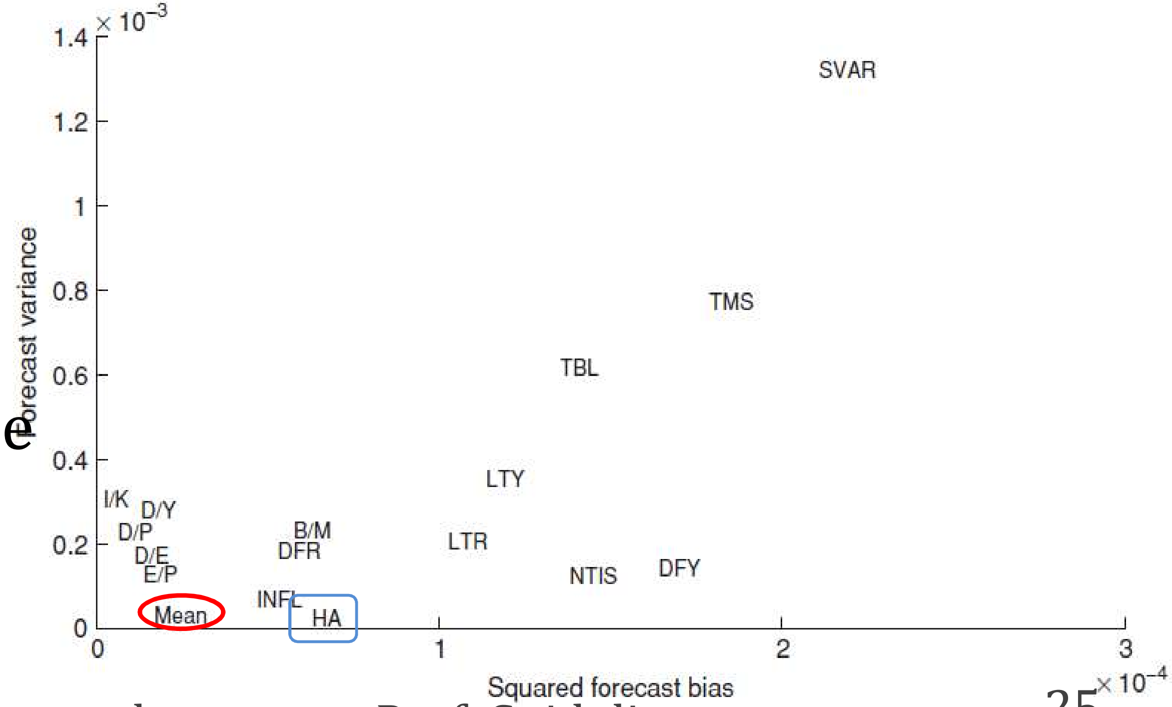
$$MHLN = [(q-q_0-1)/(q-q_0)][\hat{V}(\bar{d})^{-1/2}]\bar{d} \qquad \bar{d} = [1/(q-q_0)]\sum_{k=q_0+1}^{q} d_{R+k}$$

$$d_{t+1} = (\hat{u}_{i,t+1} - \hat{u}_{j,t+1})\hat{u}_{i,t+1}$$

Main Results

- HLN recommend using the MHLN statistic and the t_{q-q0-1} distribution to assess statistical significance

- Individual predictive regression model forecasts are frequently unable to encompass forecasts from other individual models
- The combinations are able to encompass the forecasts from the individual predictive regressions and other combining methods
- Combinations work because they reduce forecast variance and stabilizes the individual forecasts, thereby improving performance in terms of an MSPE metric, as long as the combination forecasts do not have substantial biases
- The mean combination has a lower forecast variance vs. all individual predictive regression models and also modest bias



Main Results

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***

- Combination forecasts of the equity premium are significantly correlated with growth rates in three macroeconomic variable

	Forecast horizon: one quarter				Forecast horizon: four quarters			
Combining method	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, $\theta = 1.0$	3.54***	1.71	1.56	6.26***	7.87***	2.32	3.15	11.46***
DMSPE, $\theta = 0.9$	3.49***	1.60	1.36	6.33***	5.96***	4.71*	0.27	8.27***

- The OOS gains come from extreme periods, especially bad times