Asset Management II

Week 6

Daniele Bianchi[1]

whitesphd.com

[1]School of Economics and Finance
Queen Mary, University of London

Summary

This week we expand the use of machine learning methods for the predictability of aggregate stock. In particular, we are going to cover three main papers in the academic literature which opens up the possibility of using both linear and non-linear methods to improve returns forecasting.

Contents

The "Hype"?

Machines taking over hedge funds despite lack of evidence they outperform humans

CNBC, July 12, 2017

Return on AI

Hedge funds embrace machine learning—up to a point

Economist, Dec 2017

Unshackled algorithms

Machine-learning promises to shake up large swathes of finance

Economist, May 2017

The U.S. Stock Market Belongs to Bots

Bloomberg, July 15, 2017

The Quant Fund Robot Takeover Has Been Postponed

Bloomberg, Aug 9, 2017

Same Reporter

3 weeks later

Why Big Data Tools?

- Fama Nobel Lecture: Two pillars of empirical asset pricing:
    1. Describe/understand differences in risk premia across assets.
    2. Describe/understand dynamics of the market equity risk premium.
- What is risk premia? It is a conditional expectation of future realized excess return.
- Why Big Data tools?
    1. Tools specialize in "prediction" tasks. Ideal for risk premium measurement.
    2. Staggering list of return predictors: Close cousins, highly correlated! Need to select variables, and reduce dimensionality of the problem.

Goyal and Welch, Today.

Goyal and Welch with Modern Tools: Summary

- Use Goyal and Welch (2008) data.
- Out of sample performance of: LASSO, Random Forest and compare to the linear model in Goyal and Welch (2008).
- Summary:
    1. LASSO improves OOS performance of 7 of the 14 single variable models.
    2. LASSO and Forest improve the OOS performance of a "multiple variable" model.

Data

- Use the monthly data, which is updated (See Amit Goyal's website). The sample period is December 1927 to December 2015.
- The equity premium is calculated as the simple return (including dividends) on the S&P 500 index minus the prevailing Treasury-bill rate. The equity premium has an annualized mean of 7.59% and an annualized standard deviation of 0.19.
- We use 14 predictors from Goyal and Welch (2008).

Replication: Goyal and Welch (2008)

		IS	OOS	
	Variable	$\bar{R}^2$	$\bar{R}^2$	ΔRMSE
de	Dividend payout ratio	-0.12	-0.65	-0.01
svar	Stock variance	-0.12	-0.91	-0.02
dfr	Default return spread	0.00	-0.38	-0.01
lty	Long term yield	0.05	-0.51	-0.01
ltr	Long term return	0.10	-0.54	-0.01
infl	Inflation	0.12	0.28	0.01
tms	Term spread	0.11	0.08	0.00
tbl	Treasury-bill rate	0.20	0.05	0.00
dfy	Default yield spread	0.21	-0.60	-0.01
dp	Dividend price ratio	0.30	-0.35	0.00
ep	Earning price ratio	0.30	-1.37	-0.03
bm	Book to market	0.69	-2.79	-0.06
ntis	Net equity expansion	0.37	-1.06	-0.02
dy	Dividend yield	0.42	-1.02	-0.02

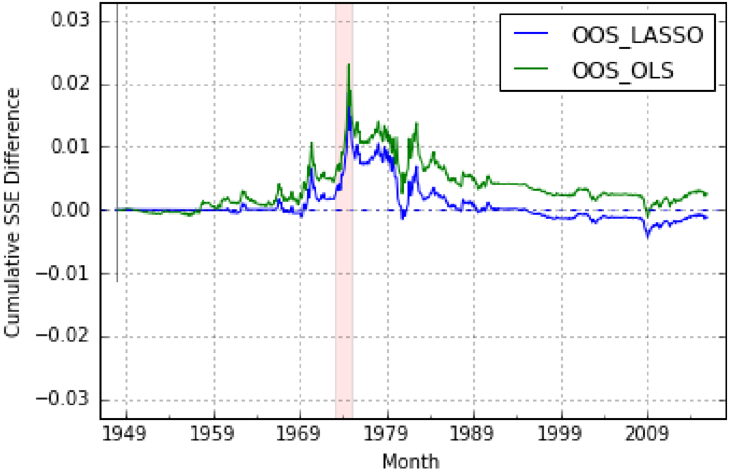
Empirical Results: LASSO with single predictors

- Compare the OOS performance of LASSO with OLS.
- In each model, use only one predictor from the 14 predictors.
- For each model, both an expanding window and a 240-month rolling window are used for comparison. The first prediction starts in January 1948.
  - LASSO improves the OOS performance of 7 out of the 14 predictors in Welch and Goyal (2008), i.e. **tbl**, **ntis**, **ep**, **bm**, **dfy**, **svar** and **tms**.
  - Only present results for **tbl** in the slides.

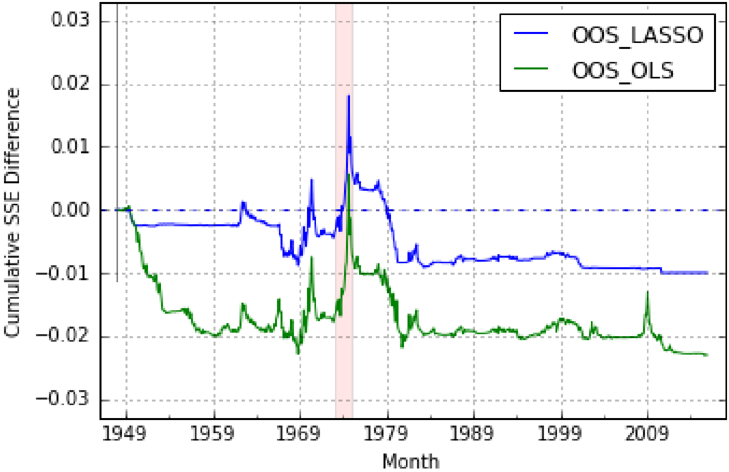
Empirical Results: LASSO with single predictors

- LASSO dynamically chooses λ based on historical information. When λ is large enough, LASSO generates close results to the historical mean model. When λ is close to zero, LASSO generates similar results to unconstrained OLS.
- LASSO prediction improves when a rolling window is used instead of an expanding window. The rolling window takes recent information into estimation, while the expanding window also accounts for historical information.

Predictor: tbl, Sample: Expanding Window



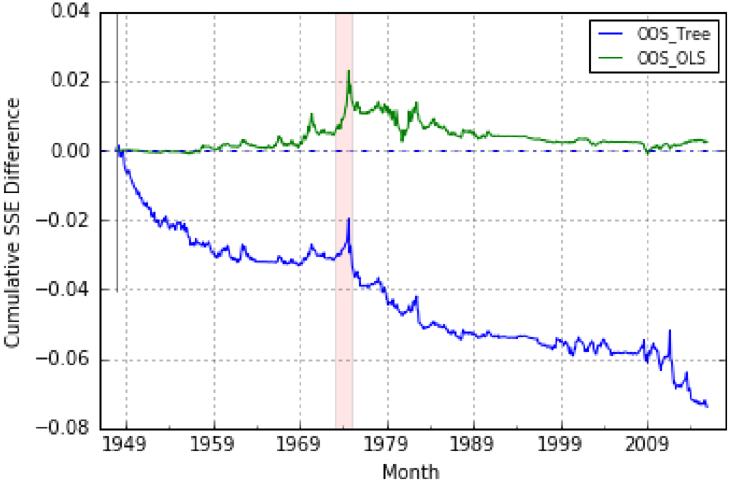
Predictor: tbl, Sample: Rolling Window



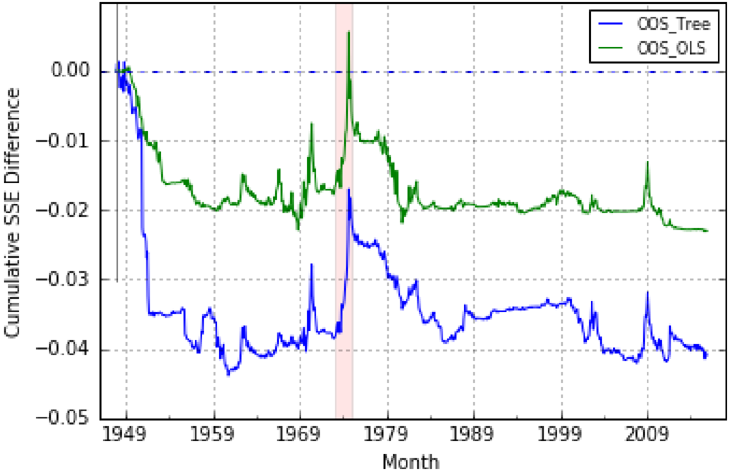
Empirical Results: Tree with Single Predictor

- The Decision Tree consistently underperforms OLS for all 14 single variable models, with either rolling windows or expanding windows.

Predictor: tbl, Sample: Expanding Window



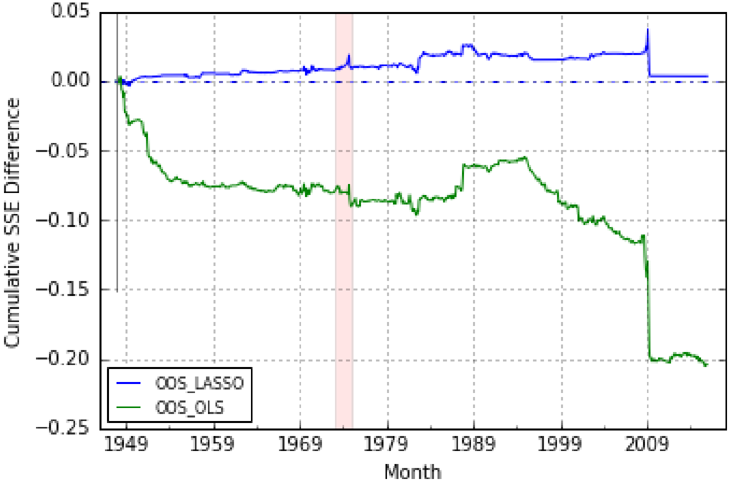
Predictor: tbl, Sample: Rolling Window



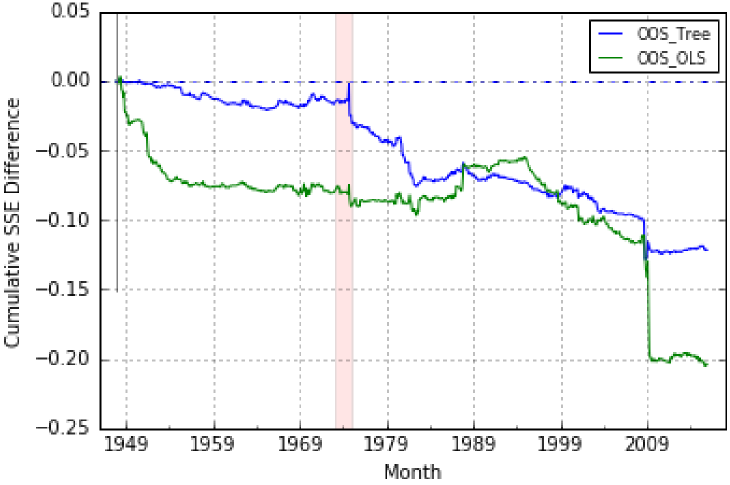
Multiple Predictor Models

- In the second part of our empirical results, we compare the OOS performance of OLS, LASSO, Decision Tree, and Random Forest using a kitchen sink model with all 14 predictor variables.
- A 240-month rolling window is used and the first prediction starts from January 1948.
- The main results are as follows:
  - Random Forest and LASSO outperform OLS, and LASSO provides the best OOS predictions.
  - Decision Tree performs worse than Random Forest. Since Random Forest consists of a large number of separately grown trees, it usually provides better prediction performance than Decision Tree.
  - The performance of Random Forest can be further improved with LASSO selected features.

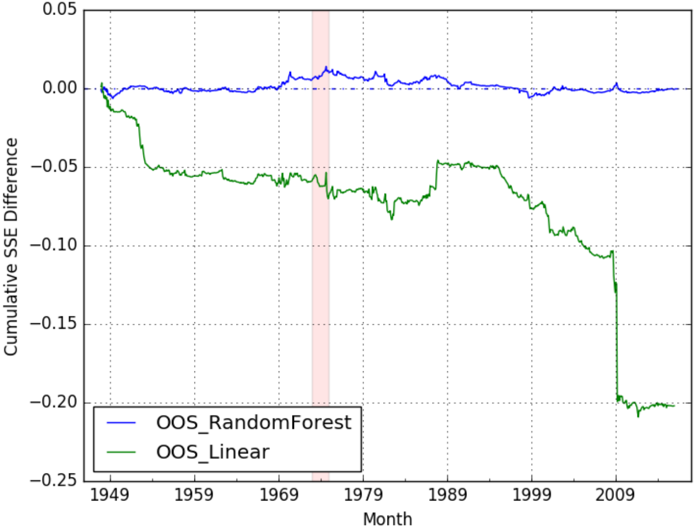
LASSO with Multiple Predictors



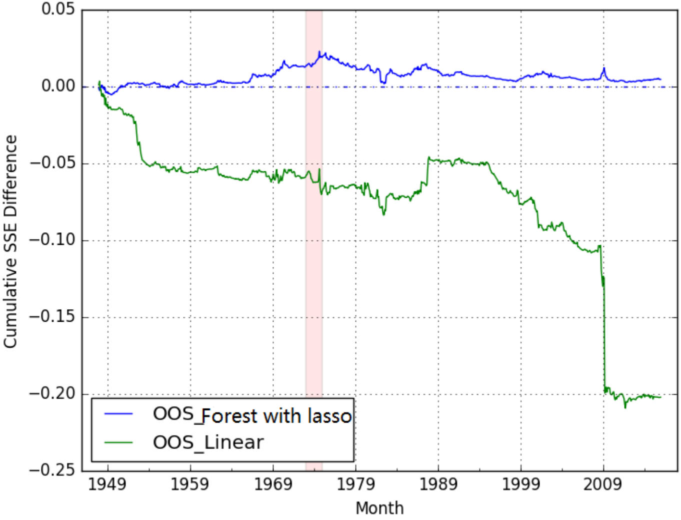
Decision Tree with Multiple Predictors



Random Forest with Multiple Predictors



Random Forest with LASSO selected features



Gu, Kelly and Xiu (2018)

Gu, Kelly and Xiu (2018)

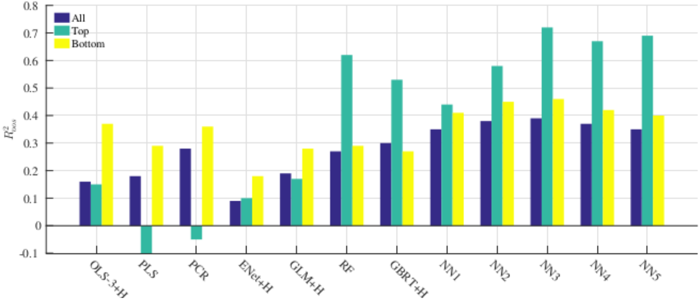
- Discover both "functional form" and "conditioning variables".
- around 100 stock characteristics
- around 10 Macroeconomic predictors (Goyal and Welch)
- Monthly returns on (1) individual stocks and (2) stock portfolios.

Main Findings (Gu, Kelly and Xiu (2018)

- Exploiting non-linearities substantially improve predictions.
- Non-linear methods become even more important when predicting portfolios.
- Gains from big data tools are economically large.
- Most successful predictors: price trends, liquidity and volatility.

Individual Stock Return Prediction: Monthly data

	OLS	OLS-3	PLS	PCR	ENet	GLM	RF	GBRT	NN1	NN2	NN3	NN4	NN5
All	-4.60	0.16	0.18	0.28	0.09	0.19	0.27	0.30	0.35	0.38	0.39	0.37	0.35
Top 1000	-14.21	0.15	-0.10	-0.05	0.10	0.17	0.62	0.53	0.44	0.58	0.72	0.67	0.69
Bottom 1000	-2.13	0.37	0.29	0.36	0.18	0.28	0.29	0.27	0.41	0.45	0.46	0.42	0.40



"Bottom up" Prediction of Portfolio Returns

- Predict portfolio returns by aggregating individual stock return predictions.

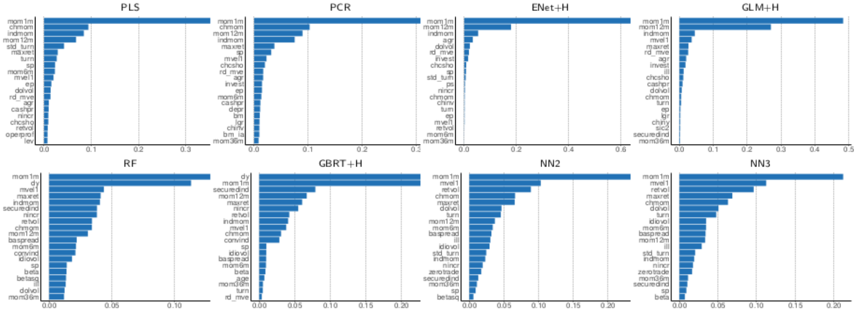
- Given weights $\omega^p_{i,t}$, in portfolio p, and individual return predictions $\hat{r}_{i,t+1}$,

$$\hat{r}^p_{i,t+1} = \sum_{i=1}^{N} \omega^p_{i,t} \times \hat{r}_{i,t+1}$$

Predicting Pre-specified Portfolios

	OLS-3	PLS	PCR	ENet	GLM	RF	GBRT	NN1	NN2	NN3	NN4	NN5
S&P 500	-0.11	-0.86	-2.62	-0.38	0.86	1.39	1.13	0.84	0.96	1.80	1.46	1.60
Big Growth	0.41	0.75	-0.77	-1.55	0.73	0.99	0.80	0.70	0.32	1.67	1.42	1.40
Big Value	-1.05	-1.88	-3.14	-0.03	0.70	1.41	1.04	0.78	1.20	1.57	1.17	1.42
Small Growth	0.35	1.54	0.72	-0.03	0.95	0.54	0.62	1.68	1.26	1.48	1.53	1.44
Small Value	-0.06	0.40	-0.12	-0.57	0.02	0.71	0.90	0.00	0.47	0.46	0.41	0.53
Big Conservative	-0.24	-0.17	-1.97	0.19	0.69	0.96	0.78	1.08	0.67	1.68	1.46	1.56
Big Aggressive	-0.12	-0.77	-2.00	-0.91	0.68	1.83	1.45	1.14	1.65	1.87	1.55	1.69
Small Conservative	0.02	0.75	0.48	-0.46	0.55	0.59	0.60	0.94	0.91	0.93	0.99	0.88
Small Aggressive	0.14	0.97	0.06	-0.54	0.19	0.86	1.04	0.25	0.66	0.75	0.67	0.79
Big Robust	-0.58	-0.22	-2.89	-0.27	1.54	1.41	0.70	0.60	0.84	1.14	1.05	1.21
Big Weak	-0.24	-1.47	-1.95	-0.40	-0.26	0.67	0.83	0.24	0.60	1.21	0.95	1.07
Small Robust	-0.77	0.77	0.18	-0.32	0.41	0.27	-0.06	-0.06	-0.02	0.06	0.13	0.15
Small Weak	0.02	0.32	-0.28	-0.25	0.17	0.90	1.31	0.84	0.85	1.09	0.96	1.08
Big Up	-1.53	-2.54	-3.93	-0.21	0.40	1.12	0.68	0.46	0.85	1.28	0.99	1.05
Big Down	-0.10	-1.20	-2.05	-0.26	0.36	1.09	0.77	0.48	0.89	1.34	1.17	1.36
Small Up	-0.79	0.42	-0.36	-0.33	-0.33	0.31	0.40	0.23	0.60	0.67	0.55	0.61
Small Down	0.40	1.16	0.47	-0.46	0.62	0.93	1.20	0.80	0.97	0.97	0.97	0.96

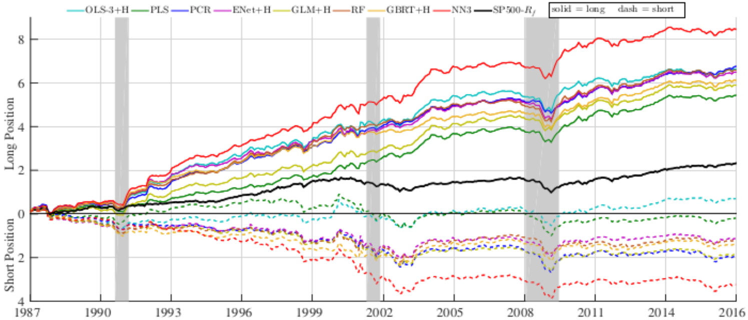
Variable Importance



Long-Short Portfolios

	OLS-3+H				PLS				PCR			
	Pred	Avg	Std	SR	Pred	Avg	Std	SR	Pred	Avg	Std	SR
Low	-0.41	0.19	6.52	0.10	-0.85	-0.05	6.73	-0.03	-0.92	-0.49	7.02	-0.24
2	-0.08	0.40	5.28	0.26	-0.26	0.31	6.02	0.18	-0.29	0.15	6.20	0.08
9	1.44	0.76	7.10	0.37	1.50	1.10	5.34	0.71	1.45	1.34	5.44	0.85
High	1.81	1.84	8.52	0.75	2.15	1.51	5.79	0.90	2.09	1.91	6.08	1.09
H-L	2.22	1.65	6.41	0.89	3.01	1.56	4.45	1.22	3.01	2.40	4.65	1.79
	Enet+H				GLM+H				RF			
Low	0.06	-0.31	7.01	-0.15	-0.43	-0.51	6.94	-0.25	0.26	-0.33	7.13	-0.16
2	0.33	0.39	6.03	0.22	0.02	0.33	6.08	0.19	0.41	0.32	5.68	0.19
9	1.34	1.03	5.92	0.60	1.37	1.29	5.59	0.80	0.89	1.17	5.64	0.72
High	1.65	1.80	7.31	0.86	1.84	1.64	6.34	0.90	1.07	1.83	6.78	0.93
H-L	1.59	2.12	5.48	1.34	2.27	2.15	4.39	1.70	0.81	2.16	5.34	1.40
	GBRT+H				NN1				NN2			
Low	-0.03	-0.38	6.67	-0.20	-0.47	-0.80	7.47	-0.37	-0.37	-0.82	7.97	-0.36
2	0.16	0.43	5.66	0.26	0.14	0.21	6.24	0.12	0.19	0.17	6.44	0.09
9	0.81	1.11	5.26	0.73	1.64	1.21	5.49	0.76	1.40	1.13	5.46	0.72
High	1.02	1.70	6.57	0.90	2.46	2.13	7.30	1.01	2.32	2.36	8.03	1.02
H-L	1.04	2.08	4.25	1.70	2.93	2.93	4.81	2.11	2.69	3.18	4.90	2.25
	NN3				NN4				NN5			
Low	-0.39	-0.96	7.77	-0.43	-0.28	-0.90	7.87	-0.40	-0.21	-0.76	7.93	-0.33
2	0.17	0.13	6.42	0.07	0.25	0.18	6.57	0.09	0.25	0.24	6.58	0.13
9	1.44	1.16	5.50	0.73	1.32	1.22	5.60	0.75	1.30	1.24	5.54	0.77
High	2.30	2.23	7.78	0.99	2.28	2.35	7.95	1.02	2.19	2.21	7.78	0.98
H-L	2.69	3.19	4.77	2.32	2.56	3.25	4.79	2.35	2.39	2.97	5.05	2.03

Cumulative Returns on Long-Short Portfolios



Bianchi, Buchner and Tamoni (2020)

Fixed-income

- Treasury bonds (or guilds) are debt papers issued by governments. Typically one assumes that they do not possess any default risk.
- Corporate bonds are similar in structure to treasuries with **one** important difference: default risk.
- Difference in yields between treasuries and corporate bonds are referred to as the "credit spread"
- The probability of default varies as a function of credit-worthiness / quality. E.g: Greece and United States. E.g: Apple vs. Vimal Inc.
- Note: The probability of default varies over time.

Money-market securities

- Short-term (less than 1 year) debt
- Issued by governments or firms
- Examples:
    - Treasury bills
    - Repurchase Agreements (REPOs)
    - Certificate of deposit
    - Commercial paper

T-Bills

- Issued by the Government: no risk of default
- Short-term securities: little inflation risk
- Highly liquid securities: Large financial institutions trade on the market
- Sold at a discount to face value: “zero-coupon” bonds

Predictability in Bond Returns

- Zero-coupon bond with maturity $t+n$, and a payoff of one dollar.
- Log price: $p_t^{(n)}$, and yield: $y_t^{(n)} = -\frac{1}{n} p_t^{(n)}$.
- Excess returns: $xr_{t+1}^{(n)} = p_{t+1}^{(n-1)} - p_t^{(n)} - y_t$
- Then, log returns on bonds is:

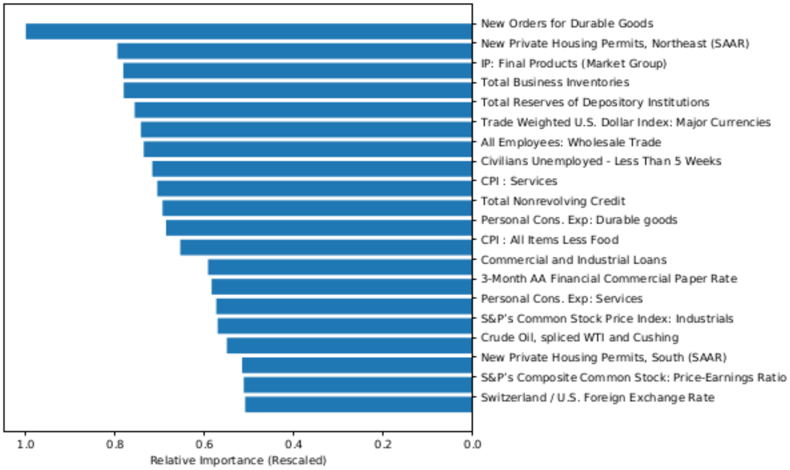
$$xr_{t+1}^{(n)} = -(n-1)\left(y_{t+1}^{(n-1)} - y_t^{(n)}\right) + \left(y_t^{(n)} - y_t^{(1)}\right)$$

- After controlling for the slope ($\left(y_{t+1}^{(n-1)} - y_t^{(n)}\right)$), any variable that forecasts change in bond yield from t to $t+1$ will also forecast the log returns to bonds.

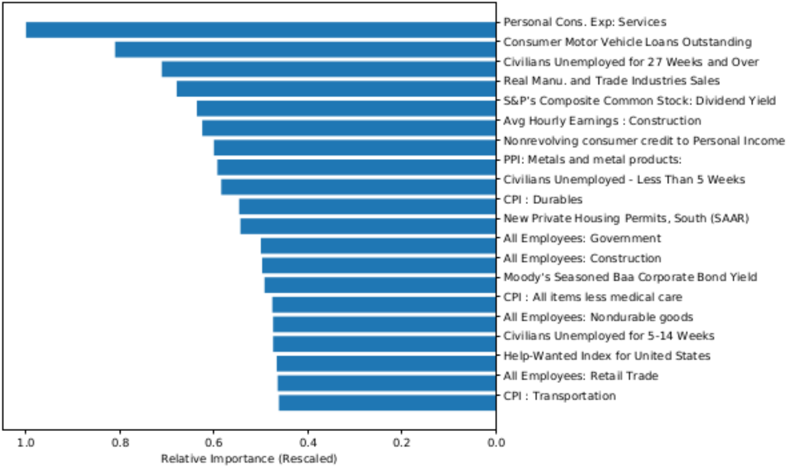
Predicting Annual Holding-period Returns

Models	R^2_{oos}					
	$xr^{(2)}_{t+1}$	$xr^{(3)}_{t+1}$	$xr^{(4)}_{t+1}$	$xr^{(5)}_{t+1}$	$xr^{(7)}_{t+1}$	$xr^{(10)}_{t+1}$
Panel A: PCA and PLS						
PCA - first 8 PC'S	-9.8%	-2.9%	0.3%	3.0%	3.3%	4.5%
PCA as in Ludvigson and Ng (2009)	-3.4%	0.2%	1.6%	1.6%	-1.4%	-4.7%
PLS - 8 components	-40.7%	-19.7%	-12.0%	-8.2%	-2.7%	3.4%
Panel B: Penalized Linear Regressions						
Ridge (using CP factor)	-45.3%	-23.6%	-16.7%	-13.2%	-3.1%	5.3%
Lasso (using CP factor)	6.4%	11.2%	12.9%	14.4%	19.6%	23.7%
Elastic Net (using CP factor)	6.4%	11.0%	14.3%	15.7%	21.7%	29.1%
Ridge (using fwd rates directly)	-52.2%	-28.7%	-22.7%	-18.3%	-13.1%	-3.5%
Lasso (using fwd rates directly)	11.0%	12.0%	12.3%	16.4%	19.9%	23.6%
Elastic Net (using fwd rates directly)	10.2%	14.2%	16.0%	13.2%	19.9%	23.6%
Panel C: Regression Trees and Neural Networks						
Gradient Boosted Tree	13.1%	15.9%	18.1%	23.8%	22.5%	25.5%
Random Forest	26.7%	21.5%	22.0%	24.4%	20.0%	25.0%
Extreme Tree	23.0%	23.4%	22.3%	23.7%	29.9%	29.6%
NN 1 Layer (32 nodes), fwd rates direct	6.0%	13.6%	17.8%	22.0%	22.5%	26.5%
NN 2 Layer (32, 16 nodes), fwd rates direct	16.6%	21.5%	24.9%	28.0%	29.4%	32.3%
NN 3 Layer (32, 16, 8 nodes), fwd rates direct	24.8%	26.3%	29.7%	32.0%	31.7%	33.7%
NN 1 Layer (32 nodes), fwd rates net (1 layer: 3 nodes)	8.4%	19.0%	23.8%	25.6%	27.2%	29.5%
NN 2 Layer (32,16, nodes), fwd rates net (1 layer: 3 nodes)	12.1%	15.7%	20.0%	23.5%	25.4%	28.1%
NN 3 Layer (32,16, 8 nodes), fwd rates net (1 layer: 3 nodes)	7.6%	16.3%	20.2%	23.7%	25.0%	28.1%
NN 1 Layer Group Ensem (1 node per group), fwd rates direct	12.6%	17.3%	21.6%	24.2%	25.9%	29.6%
NN 1 Layer Group Ensem (1 node per group), fwd rates net (1 layer: 3 nodes)	20.0%	25.6%	29.5%	31.2%	33.6%	36.3%
NN 2 Layer Group Ensem (2,1 nodes per group / hidden layer), fwd rates net (2 layer: 3 nodes)	17.3%	23.6%	27.8%	29.8%	31.0%	33.0%
NN 3 Layer Group Ensem (3, 2, 1 nodes per group / hidden layer), fwd rates net (3 layer: 3 nodes)	13.6%	20.0%	23.7%	26.1%	27.5%	30.7%

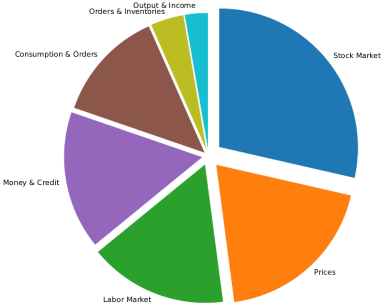
Variable Importance: 2-year maturity



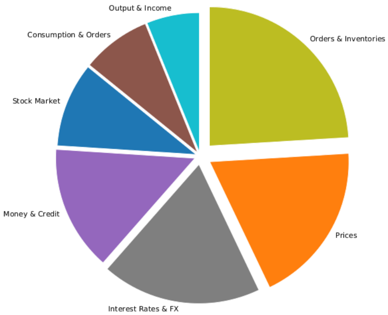
Variable Importance: 10-year maturity



Variable Importance



(c) 2-year maturity, groups of variables



(d) 10-year maturity, groups of variables

Wrap up

- Big Data techniques have revolutionized predictability for both stock and bond return premia.
- Not merely a "hype", but asset managers can indeed improve their portfolios with big data techniques.
- Large hedge funds, and financial intermediaries use these techniques today.
- The importance of these techniques is only growing!