Discussion of

Market Expectations in the Cross Section of Present Values

by Kelly, B., and S. Pruitt

Stefan Nagel

Stanford University, NBER, CEPR

January 2013

The problem tackled by this paper

- Market return

$$r_{mt+1} = \mu_t + \eta_{t+1}$$

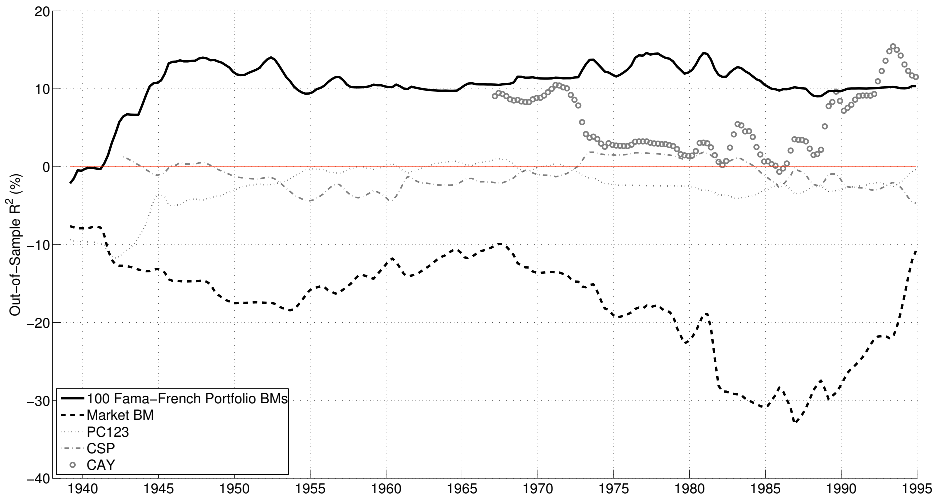
- Present-value decomposition of market M/B ratio

$$x_{mt} \approx a - b_\mu \mu_t + b_g g_t$$

  with VAR(1) dynamics of book ROE, g_t, and expected return, μ_t
- Presence of g_t obscures "signal" μ_t
- This paper: Cross-sectional information helps filter out μ_t to improve predictive regressions
- Much of my discussion focuses on Kelly and Pruitt (2012, "3PRF", WP), the paper that supplies the methodology

The results: U.S.

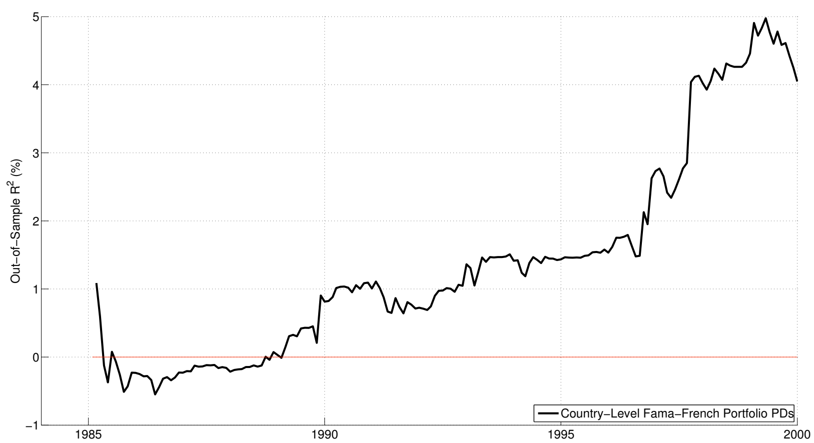
Figure 1: Out-of-Sample R^2 by Sample Split Date, One Year Returns



Notes: Out-of-sample R^2 across sample split dates. Forecasts are based on a single PLS factor from 100 book-to-market ratios of size and value-sorted portfolios, the aggregate book-to-market ratio, the cross-section premium of Polk et al. (2006), the consumption-wealth ratio of Lettau and Ludvigson (2001), and the first three principal components of the 100 book-to-market ratio cross section.

The results: International

Figure 3: Out-of-Sample R^2 by Sample Split Date, One Month International Returns



Notes: Out-of-sample percentage R^2 across sample split dates for forecasts of one month international stock returns using a single PLS factor from 42 price-dividend ratios of high value and low value portfolios across 21 countries (Fama and French (1998)). See Section III.A.5 for list of countries.

How does it work? Example

- Define: $g_t \equiv$ ROE component orthogonal to expected returns
- Assume: Three-asset cross-section where M/B ratios load on μ_t and g_t with cross-sectionally **uncorrelated** loadings

$$
\begin{aligned}
x_{1t} &= 2\mu_t - 1.5g_t & \tilde{x}_{1t} &= \mu_t - 0.5g_t \\
x_{2t} &= \mu_t & \tilde{x}_{2t} &= g_t \\
x_{3t} &= -1.5g_t & \tilde{x}_{3t} &= -\mu_t - 0.5g_t
\end{aligned}
$$

- First-stage t.s. regressions of x_{it} on r_{mt+1} (large T): slopes proportional to

$$
\begin{aligned}
\phi_1 &= 2 & \tilde{\phi}_1 &= 1 \\
\phi_2 &= 1 & \tilde{\phi}_2 &= 0 \\
\phi_3 &= 0 & \tilde{\phi}_3 &= -1
\end{aligned}
$$

- Second-stage c.s. regressions of $\tilde{x}_{it}$ on $\tilde{\phi}_i$ each t: slopes

$$F_t = \text{const.} \times [(\mu_t - 0.5g_t) + 0 - (-\mu_t - 0.5g_t)] = \text{const.} \times \mu_t$$

How does it work? Crucial assumption

- Assumption that loadings on μ_t and g_t are c.s. **uncorrelated** is important: Second stage regression slopes F_t are then...
  - "Long" in assets with positive M/B loadings on μ_t
  - "Short" in assets with negative M/B loadings on μ_t
  - "Long" and "short" in assets with M/B that load similarly on g_t: g_t exposure cancels out
  - As $N \to \infty$, sample c.s. correlation closer to zero population c.s. correlation: μ_t consistently estimated
- But what if loadings on μ_t and g_t are c.s. **correlated**?
  - Lucky: For stock market application in JF paper, correlation seems to be close to zero
  - But this may not be true in other applications
  - Remedy: Use proxies for g_t in addition to μ_t proxy
  - But that means we have to take a stand on **all** of the systematic factors driving M/B ratios

Concern: Large N, small T

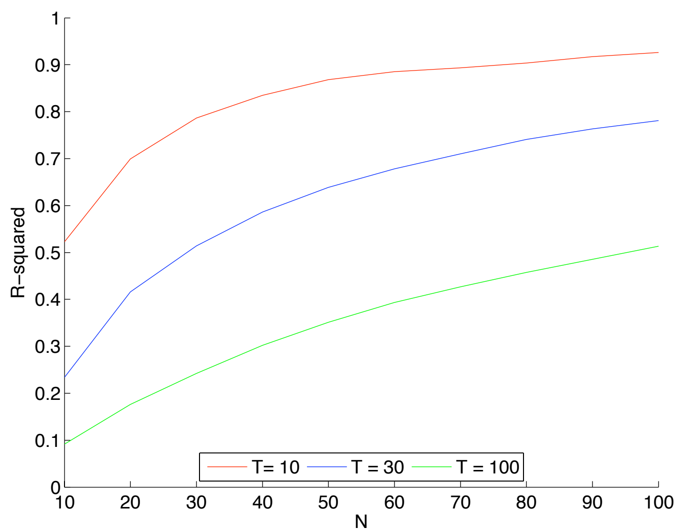
- What are the properties of the estimator under the null hypothesis of no predictability?
- Simulation: No-predictability & pure-noise M/B null

$$r_{mt+1} = \eta_{t+1}$$
$$x_t = \epsilon_t$$

where

$$\begin{pmatrix} \eta_{t+1} \\ \epsilon_t \end{pmatrix} \sim \mathcal{N}(0, I_{N+1})$$

Large N, small T: Spurious fit as N grows



Mean third-stage R-squared under no-predictability null (100 simulations for each (N, T) pair)

Large N, small T: Example with $T = 3$ and $N = 9$

Pure noise returns

Pure noise M/B

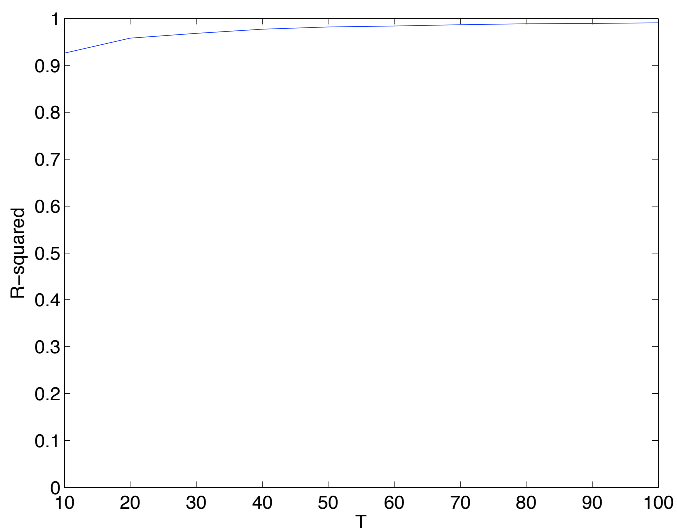
First stage (sorted)

Second stage

Third stage (fitted)

r_{mt+1}			
-1.35	0.91	-1.49	
x_{it}			$\hat{\phi}_i$
-0.52	-1.26	0.39	-5.56
-0.82	-0.44	-0.22	-1.73
-0.20	1.07	2.51	-0.49
-0.93	-2.62	0.34	-0.20
0.40	-0.06	-1.14	-0.02
1.94	0.43	2.08	1.18
-1.89	-0.96	-1.83	1.83
0.44	-0.82	1.51	10.88
-0.68	-0.37	-0.07	18.28
$\hat{F}_t$			
0.0044	0.0100	0.0040	
$\hat{y}_{t+1}$			
-0.98	0.80	-1.75	
R^2			
0.94			

N and T grow simultaneously: Spurious fit with $N = T^2$



Mean third-stage R-squared under no-predictability null with $N = T^2$
(100 simulations for each (N, T) pair)

Spurious predictability

- Thus, when N is not small relative to T, there is a bias that overstates predictability
- It seems that this is not just a small-sample bias: also asymptotic bias if N grows sufficiently fast relative to T
  - Theorem 1 in Kelly and Pruitt (2012, "3PRF") seems to need additional assumption: N cannot grow too fast relative to T
- Not a concern for the empirical results in the JF paper, as out-of-sample tests are not affected by this bias
- But concern underscores importance of out-of-sample testing
- Calls for further study of small-sample and asymptotic properties of this estimator

Summing up

- Nice idea
- Impressive empirical results: Strong out-of-sample predictability of stock market returns
- Some further work necessary on the properties of the 3PRF estimator
  - Correlated factor loadings
  - Small-T/large-N behavior
  - Asymptotic behavior when N grows fast relative to T
- These concerns do not affect out-of-sample tests: results in the JF paper are robust