Testing for Episodic Predictability in Stock Returns

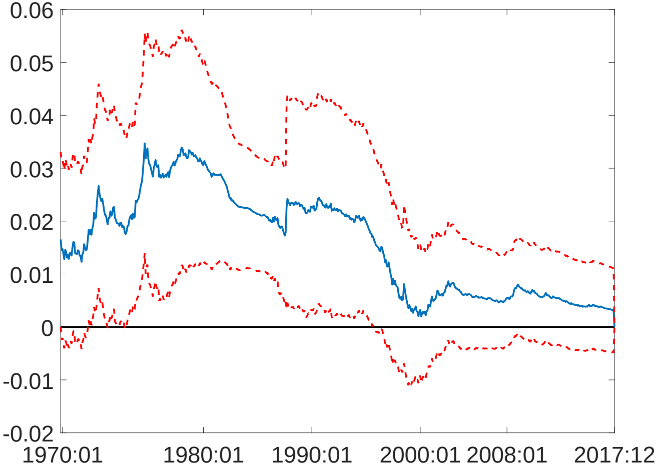
Robert Taylor, University of Essex

joint work with Matei Demetrescu, Iliyan Georgiev and Paulo Rodrigues

ESEM Manchester, August 28th 2019

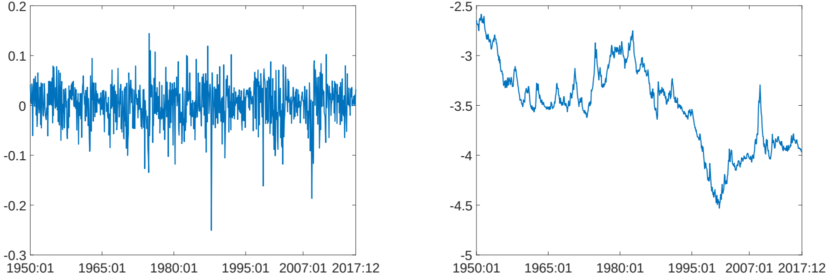
This research is funded by the ESRC under award number ES/R00496X/1.

Is there any predictability in the equity premium?



Dividend yield: Forward Recursive IV regression estimates and pointwise CIs, 1950-2017 (Goyal/Welch 2008 updated monthly data).

... what about the persistence of the predictor?



The equity premium looks very mean reverting etc (almost noise), but the dividend yield looks strongly persistent (usual ADF test has p-value of 0.41).

Outline

1. Background
2. The Tests
3. Finite Sample Simulations
4. Equity Premium Predictability
5. Concluding Remarks

Moving on to ...

1. Background
2. The Tests
3. Finite Sample Simulations
4. Equity Premium Predictability
5. Concluding Remarks

Trouble in paradise

Consider the predictive regression

$$y_t = \beta_0 + \beta_1 \, x_{t-1} + u_t$$

where

$$x_t = \rho \, x_{t-1} + v_t,$$

with $(u_t, v_t)' \sim iid(0, \mathbf{\Sigma})$ where

$$\mathbf{\Sigma} = \mathrm{E}\left(\begin{pmatrix} u_t \\ v_t \end{pmatrix} \begin{pmatrix} u_t & v_t \end{pmatrix} \right) = \begin{pmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{pmatrix}.$$

Null hypothesis: x_{t-1} does not predict y_t, i.e.

$$H_0 : \quad \beta_1 = 0 \, .$$

Yet, even in this simplest setup...

Endogeneity and (high) persistence

Should

- the shocks u_t and v_t correlate (so that $\phi := \sigma_{uv}/\sigma_u\sigma_v \neq 0$; for the EP-DY data above this correlation is estimated to be $\hat{\phi} = -0.98$), and
- the regressor x_t be autocorrelated,

one speaks of endogeneity. (A bit of a misnomer.)

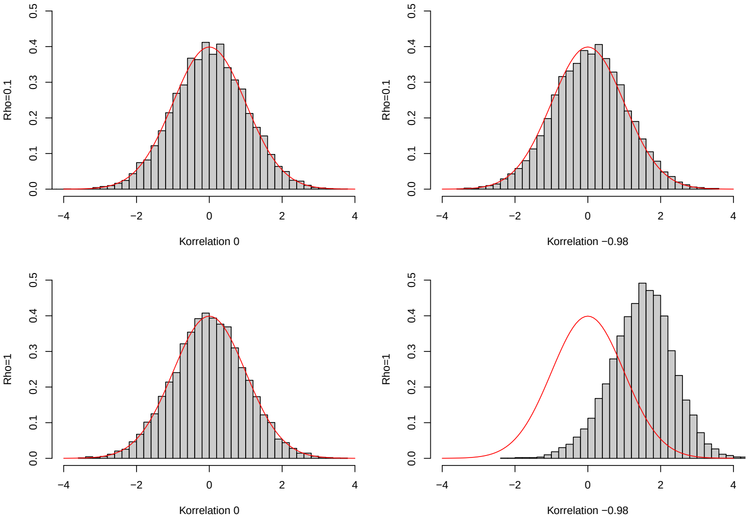
Under endogeneity and high persistence (near integration, $\rho = 1 - c/T$),

- the OLS estimator is 2nd order biased and
- the *t*-statistic has a non-normal limiting distribution.

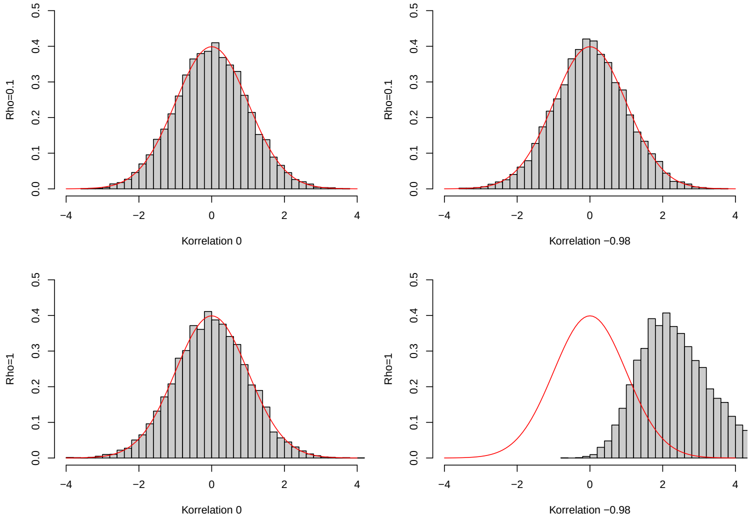
See Elliott/Stock (1994), Stambaugh (1999), Campbell/Yogo (2006) etc.

No problem when regressors are stationary or weakly persistent.

Trouble brewing - OLS t-statistics, $T = 305$



More trouble with variance breaks - volatility of both shocks 3 times higher in the first 20% of the sample



Feasible solutions for the full sample

If ρ were known, one could employ GLS estimation. For unknown ρ:

- Bayes methods - Elliott/Stock (1994)
- Bonferroni - Campbell/Yogo (2006), but see Phillips (2012)
- Restricted log-likelihood - Jansson/Moreira (2006), Chen/Deo (2009)
- Almost optimal tests - Elliott *et al.* (2015)
- Variable addition - Toda/Yamamoto (1995), Dolado/Lütkepohl (1996)
- Generic IV estimation - Breitung/Demetrescu (2015)
- IVX method of Kostakis *et al.* (2015)

(How) Can we extend these for analysing various sequences of subsamples?

Moving on to ...

Episodic Predictability

Consider the time-varying DGP

$$y_t = \beta_0 + \beta_{1,t} x_{t-1} + u_t, \qquad t = 1, \ldots, T \tag{1}$$

where x_t, $t = 0, ..., T$, is given by the DGP

$$x_t = \mu_x + \xi_t, \qquad t = 0, \ldots, T \tag{2a}$$
$$\xi_t = \rho\, \xi_{t-1} + v_t, \qquad t = 1, \ldots, T \tag{2b}$$

with $\xi_0 = O_p(1)$. (Extensions for multiple predictors straightforward.)

Uncertain Persistence

Assumption 1

One of the following two conditions is assumed to hold:

1. ***Weakly persistent predictors:*** *The autoregressive parameter ρ in (2) is fixed and bounded away from unity, $|\rho| < 1$.*
2. ***Strongly persistent predictors:*** *The autoregressive parameter ρ in (2) is local-to-unity with $\rho := 1 - \frac{c}{T}$ where c is a fixed non-negative constant.*

The literature uses both equally often, but the asymptotics are different.

Time-Varying Features I

Assumption 2

In the context of (1) and (2), let $\beta_{1,t} := n_T^{-1} b\left(t/T\right)$, *where* $b(\cdot)$ *is a piecewise Lipschitz-continuous real function on* $[0,1]$, *with* $n_T = \sqrt{T}$ *under Assumption 1.1, and* $n_T = T$ *under Assumption 1.2.*

We may reformulate the null hypothesis stated previously as

$$H_0: \text{ The function } b(\tau) \text{ is identically zero for all } \tau \in [0,1]. \tag{3}$$

We can now also formally specify the alternative hypothesis as,

$$H_{1,b(\cdot)}: b(\cdot) \text{ is non-zero over at least one non-empty subinterval of } [0,1]. \tag{4}$$

Time-Varying Features II

- Under the null hypothesis, H_0, y_t is not predictable by x_{t-1} in any subsample.
- Under the alternative hypothesis, $H_{1,b(\cdot)}$, there exists at least one subset of the sample observations (this need not be a strict subset, so it could contain all of the sample observations) comprising contiguous observations and for which $\beta_{1,t} \neq 0$. A predictive episode, often termed a *pocket of predictability*. The size of this subset is proportional to the sample size T.

Time-Varying Features III

Assumption 3

Let $\begin{pmatrix} u_t \\ v_t \end{pmatrix} := \begin{pmatrix} 1 & 0 \\ 0 & B(L) \end{pmatrix} \mathbf{H}(t/T) \begin{pmatrix} a_t \\ e_t \end{pmatrix}, \quad \begin{pmatrix} a_t \\ e_t \end{pmatrix} \sim \mathit{WN}(\mathbf{0}, \mathbf{I}_2),$
where:

1. $\zeta_t := (a_t, e_t)'$ *is a uniformly* L_4*-bounded MD sequence which is such that* $\sup_t \mathrm{E}\left|\mathrm{E}\left(\zeta_t\zeta_t' - \mathbf{I}_2 | \zeta_{t-m}, \zeta_{t-m-1}, \ldots\right)\right| \to 0$ *as* $m \to \infty$*;*
2. $\mathbf{H}(\cdot) := \begin{pmatrix} h_{11}(\cdot) & h_{12}(\cdot) \\ h_{21}(\cdot) & h_{22}(\cdot) \end{pmatrix}$ *is a matrix of piecewise Lipschitz-continuous bounded functions on* $(-\infty, 1]$*, which is of full rank at all but a finite number of points;*
3. $B(L)$*, where* L *denotes the usual lag operator, is an invertible lag polynomial with* $b_0 = 1$ *and 1-summable coefficients,* $\sum_{j\geq 0} j\,|b_j| < \infty$*, for which* $\omega := \sum_{j\geq 0} b_j > 0$*.*

IV estimation

Notice that

- IV estimation may be adapted to subsample analyses (e.g. Andrews, 1993), while
- other approaches are not easily implemented for subsamples.

Breitung and Demetrescu (2015) discuss full-sample IV estimation using a vector $\boldsymbol{z}_t$ of instruments

... and recommend 2SLS-based testing:

- a so-called type-I instrument, $z_{I,t}$, is less persistent than x_t,
- a so-called type-II instrument, $z_{II,t}$, is persistent yet exogenous.

Full sample IV statistics

Breitung and and Demetrescu (2015) consider the full-sample 2SLS t-ratio:

$$t_{\beta_1} := \frac{\boldsymbol{A}_T' \mathbf{B}_T^{-1} \boldsymbol{C}_T}{\sqrt{\boldsymbol{A}_T' \mathbf{B}_T^{-1} \mathbf{D}_T \mathbf{B}_T^{-1} \boldsymbol{A}_T}} \tag{5}$$

where

$$\begin{aligned} \boldsymbol{A}_T &:= \sum_{t=1}^{T} \hat{x}_{t-1} \hat{\boldsymbol{z}}_{t-1}, \ \mathbf{B}_T := \sum_{t=1}^{T} \hat{\boldsymbol{z}}_{t-1} \hat{\boldsymbol{z}}_{t-1}', \\ \boldsymbol{C}_T &:= \sum_{t=1}^{T} \hat{\boldsymbol{z}}_{t-1} \hat{y}_t, \quad \text{and} \quad \mathbf{D}_T := \sum_{t=1}^{T} \hat{\boldsymbol{z}}_{t-1} \hat{\boldsymbol{z}}_{t-1}' \hat{u}_t^2, \end{aligned}$$

and $\hat{y}_t$, $\hat{x}_{t-1}$ and $\hat{\boldsymbol{z}}_{t-1}$ are demeaned versions of y_t, x_{t-1} and $\boldsymbol{z}_{t-1}$

Breitung and and Demetrescu (2015) show that $(t_{\beta_1})^2$ has a limiting $\chi^2(1)$ distribution, even with time-varying volatility (regularity conditions assumed).

Max subsample statistics

Let $t_{\beta_1}(\tau_1, \tau_2)$ be the 2SLS t statistic computed for the subsample $t = [\tau_1 T] + 1, \ldots, [\tau_2 T]$, with Eicker-White standard errors.

Consider the sequences of statistics

- forward recursive, $\mathcal{T}^f := \max_{\tau_L \leq \tau \leq 1} \left\{ (t_{\beta_1}(0, \tau))^2 \right\}$
- backward recursive, $\mathcal{T}^b := \max_{0 \leq \tau \leq \tau_U} \left\{ (t_{\beta_1}(\tau, 1))^2 \right\}$
- rolling, $\mathcal{T}^r := \max_{0 \leq \tau \leq 1 - \Delta\tau} \left\{ (t_{\beta_1}(\tau, \tau + \Delta\tau))^2 \right\}$
- double recursive, $\mathcal{T}^d := \max_{\substack{0 \leq \tau_1, \tau_2 \leq 1 \\ \tau_2 - \tau_1 \geq \Delta\tau}} \left\{ (t_{\beta_1}(\tau_1, \tau_2))^2 \right\}$.

We'll need specific regularity conditions to deal with them.

Type-I instruments

Assumption 4

1. $\mathrm{E}\left(\zeta_t|\zeta_{t-1},\zeta_{t-2},\ldots,z_{I,t-1},z_{I,t-2},\ldots\right)=0$, $\exists\delta_I\geq 0$ *s.t.* $T^{-\delta_I}z_{I,t}$ *unif.* L_4*-bd.* $\sup_{\tau\in[0,1]}\left|\frac{1}{T^{1+\delta_I}}\sum_{t=1}^{\lfloor\tau T\rfloor}z_{I,t-1}\right|\xrightarrow{p}0$,$\sup_{\tau\in[0,1]}\left|\frac{1}{T^{1+\delta_I}}\sum_{t=1}^{\lfloor\tau T\rfloor}z_{I,t-1}u_t^2\right|=O_p(1)$
2. *Under weak persistence:*

   2.1 $\frac{1}{T^{1+\delta_I}}\sum_{t=1}^{\lfloor\tau T\rfloor}z_{I,t-1}\xi_{t-1}\Rightarrow K_{z_Ix}\left(\tau\right)$, *where* $K_{z_Ix}\left(\tau\right)>0$

   2.2 $\frac{1}{T^{1+2\delta_I}}\sum_{t=1}^{\lfloor\tau T\rfloor}z_{I,t-1}^2\xrightarrow{p}K_{z_I^2}\left(\tau\right)$ *strictly increasing*

   2.3 $\frac{1}{T^{1/2+\delta_I}}\sum_{t=1}^{\lfloor\tau T\rfloor}z_{I,t-1}u_t\Rightarrow G_I(\tau)$ *Gaussian with* $\eta_I(\tau):=\frac{[G_I](\tau)}{[G_I](1)}$, $[G_I](\tau)$ *denotes the quadratic variation process of* $G_I(\tau)$

   2.4 $\frac{1}{T^{1+2\delta_I}}\sum_{t=1}^{\lfloor\tau T\rfloor}z_{I,t-1}^2u_t^2\xrightarrow{p}\left[G_I\right]\left(\tau\right)$
3. *Under strong persistence*

   3.1 $\sup_{\tau\in[0,1]}\left|\frac{1}{T^{3/2+\delta_I}}\sum_{t=1}^{\lfloor\tau T\rfloor}z_{I,t-1}\xi_{t-1}\right|\xrightarrow{p}0$

   3.2 $\frac{1}{T^{1+2\delta_I}}\sum_{t=1}^{\lfloor\tau T\rfloor}z_{I,t-1}^2\xrightarrow{p}K_{z_I^2}\left(\tau\right)$ *strictly increasing*

   3.3 $\sup_{\tau\in[0,1]}\left|\frac{1}{T^{1/2+\delta_I}}\sum_{t=1}^{\lfloor\tau T\rfloor}z_{I,t-1}u_t\right|=O_p(1)$

   3.4 $\frac{1}{T^{1+2\delta_I}}\sum_{t=1}^{T}z_{I,t-1}^2u_t^2=O_p(1)$

Type-II and interactions

Assumption 5

The variable $z_{II,t}$ *is deterministic and, for some function* $Z(\tau)$*, Hölder-continuous of order* $\alpha > 1/2$*, and some* $\delta_{II} \geq 0$*, satisfies* $T^{-\delta_{II}} z_{II,\lfloor \tau T \rfloor} \Rightarrow Z(\tau)$ *in* $\mathcal{D}$ *where* $Z(\cdot)$ *is such that, for all* $0 \leq \tau_1 < \tau_2 \leq 1$*,* $\int_{\tau_1}^{\tau_2} \tilde{Z}^2_{\tau_1,\tau_2}(s)\mathrm{d}s \neq 0$ *with*
$\tilde{Z}_{\tau_1,\tau_2}(s) := Z(s) - \frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} Z(s)\mathrm{d}s$*.*

Assumption 6

For instruments $z_{I,t}$ *and* $z_{II,t}$ *satisfying the conditions of Assumptions 4 and 5, respectively, it is also required that:*

1. $\sup_{\tau \in [0,1]} \left| \frac{1}{T^{1+\delta_I+\delta_{II}}} \sum_{t=1}^{\lfloor \tau T \rfloor} z_{I,t-1} z_{II,t-1} \right| \xrightarrow{p} 0$ *and*
2. $\sup_{\tau \in [0,1]} \left| \frac{1}{T^{1+\delta_I+\delta_{II}}} \sum_{t=1}^{\lfloor \tau T \rfloor} z_{I,t-1} z_{II,t-1} u_t^2 \right| = O_p(1)$*.*

Instruments in practice

Use a vector of two instruments per regressor; e.g.:

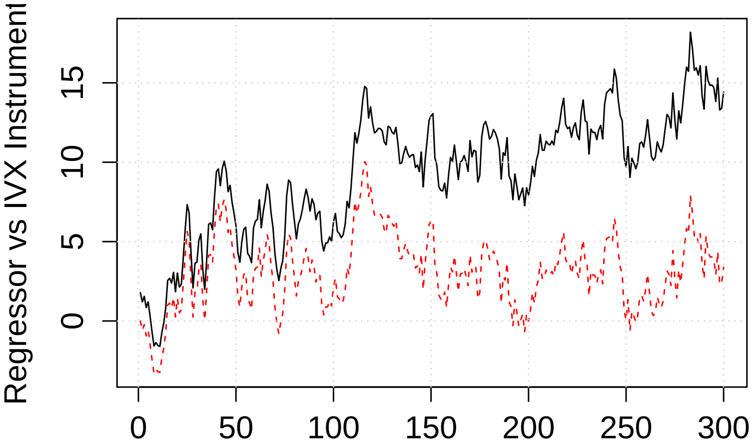
$$z_{I,t} := \sum_{j=0}^{t-1} \varrho^j \Delta x_{t-j} = (1 - \varrho L)_+^{-1} \Delta x_t$$

where $\varrho := 1 - \frac{a}{T^\gamma}$ with $\gamma \in (0, 1)$ and $a \geq 0$, and

$$z_{II,t} := \sin\left(\frac{\pi t}{2T}\right).$$

- $z_{I,t}$ is the IVX instrument (Kostakis *et al.*, 2015) – other choices possible
- Both satisfy the generic assumptions above

The IVX trick applied to a random walk - using $a = 1$, $\gamma = 0.95$ as in Kostakis *et al.*



Still trouble in paradise I

Distributions under local alternatives available in paper.

Corollary 1

Under the null hypothesis H_0 *of* (3) *and Assumption 1.1,*

$$\mathcal{T}^f \Rightarrow \sup_{\tau \in [\tau_L, 1]} \frac{(G_I(\tau))^2}{[G_I](1)\, \eta_I(\tau)} := \mathcal{T}_\infty^{f,I}$$

$$\mathcal{T}^b \Rightarrow \sup_{\tau \in [0, \tau_U]} \frac{(G_I(1) - G_I(\tau))^2}{[G_I](1)\,(1 - \eta_I(\tau))} := \mathcal{T}_\infty^{b,I}$$

$$\mathcal{T}^d \Rightarrow \sup_{\substack{0 \le \tau_1, \tau_2 \le 1 \\ \tau_2 - \tau_1 \ge \Delta\tau}} \frac{(G_I(\tau_2) - G_I(\tau_1))^2}{[G_I](1)\,(\eta_I(\tau_2) - \eta_I(\tau_1))} := \mathcal{T}_\infty^{d,I}$$

$$\mathcal{T}^r \Rightarrow \sup_{0 \le \tau \le 1 - \Delta\tau} \frac{(G_I(\tau + \Delta\tau) - G_I(\tau))^2}{[G_I](1)\,(\eta_I(\tau + \Delta\tau) - \eta_I(\tau))} := \mathcal{T}_\infty^{r,I}$$

as $T \to \infty$, *where* $\eta_I(\tau) := [G_I](\tau)/[G_I](1)$ *and* G_I *is the limiting process given in the Assumption for Type-I instruments.*

Still trouble in paradise II

Corollary 2

Under the null hypothesis H_0 of (3) and Assumption 1.2,

$$\mathcal{T}^f \Rightarrow \sup_{\tau_L \le \tau \le 1} \frac{\left(\int_0^\tau \tilde{Z}_{0,\tau}(s)\mathrm{d}U(s)\right)^2}{\int_0^\tau \tilde{Z}_{0,\tau}^2(s)\mathrm{d}[U](s)} := \mathcal{T}_\infty^{f,II}$$

$$\mathcal{T}^b \Rightarrow \sup_{0 \le \tau \le \tau_U} \frac{\left(\int_\tau^1 \tilde{Z}_{\tau,1}(s)\mathrm{d}U(s)\right)^2}{\int_\tau^1 \tilde{Z}_{\tau,1}^2(s)\mathrm{d}[U](s)} := \mathcal{T}_\infty^{b,II}$$

$$\mathcal{T}^r \Rightarrow \sup_{0 \le \tau \le 1-\Delta_\tau} \frac{\left(\int_\tau^{\tau+\Delta\tau} \tilde{Z}_{\tau,\tau+\Delta_\tau}(s)\mathrm{d}U(s)\right)^2}{\int_\tau^{\tau+\Delta_\tau} \tilde{Z}_{\tau,\tau+\Delta_\tau}^2(s)\mathrm{d}[U](s)} := \mathcal{T}_\infty^{r,II}$$

$$\mathcal{T}^d \Rightarrow \sup_{\substack{0 \le \tau_1,\tau_2 \le 1 \\ \tau_2-\tau_1 \ge \Delta_\tau}} \frac{\left(\int_{\tau_1}^{\tau_2} \tilde{Z}_{\tau_1,\tau_2}(s)\mathrm{d}U(s)\right)^2}{\int_{\tau_1}^{\tau_2} \tilde{Z}_{\tau_1,\tau_2}^2(s)\mathrm{d}[U](s)} := \mathcal{T}_\infty^{d,II}$$

with $\tilde{Z}_{\tau_1,\tau_2}(s) := Z(s) - \frac{1}{\tau_2-\tau_1}\int_{\tau_1}^{\tau_2} Z(s)\mathrm{d}s$ *and* $U(s)$ *a time-transformed Brownian motion.*

The main take-aways from these rather uninformative looking functionals are that the limiting null distributions of the subsample maxima statistics:

1. Do not depend on the magnitude of ρ under Assumption 1.1 or the mean-reversion parameter c under Assumption 1.2 (in both cases asymptotic local power does though).
2. All depend in general on any heteroskedasticity present, despite being based on White standard errors.
3. Have different functional forms depending on whether x_t is near-integrated or stable.

The last two features pose significant problems for conducting inference that are not encountered with tests based on the full sample IV-combination statistic, $t^2_{\beta_1}$ of (5). However, as we will see next, these issues can be solved by using a fixed regressor wild bootstrap implementation of the subsample tests.

A fixed-regressor wild bootstrap

1. Construct the wild bootstrap innovations $y_t^* := \hat{y}_t R_t$, where $\hat{y}_t := y_t - \frac{1}{T}\sum_{t=1}^{T} y_t$ are the demeaned sample observations on y_t, and R_t, $t = 1, ..., T$, are $IID\mathcal{N}(0,1)$, independent of the data.
2. Using the bootstrap sample data $\left(y_t^*, x_{t-1}, \boldsymbol{z}_{t-1}'\right)'$, in place of the original sample data $\left(y_t, x_{t-1}, \boldsymbol{z}_{t-1}'\right)'$, construct the bootstrap analogues of the statistics $\mathcal{T}^s$, $s = f, b, d, r$.

   Denote these bootstrap statistics as $\mathcal{T}^{s*}$, $s = f, b, d, r$.
3. Define the bootstrap p-values as $P_T^{s,*} := 1 - G_T^{s,*}\left(\mathcal{T}^s\right)$, $s = f, b, d, r$, with $G_T^{s,*}(\cdot)$ denoting the conditional (on the original data) cumulative distribution function (cdf) of $\mathcal{T}^{s*}$, $s = f, b, d, r$. (In practice, simulate $G_T^{s,*}(\cdot)$ in the usual way.)
4. The wild bootstrap test of the null hypothesis H_0 of (3) at level α based on $\mathcal{T}^s$ rejects if $P_T^{s,*} \leq \alpha$, $s = f, b, d, r$.

... which is first-order asymptotically valid

Proposition 1

Under either the null hypothesis H_0 *of* (3) *or the local alternative* $H_{1,b(\cdot)}$ *of* (4)*:*

(i) *Under Assumption 1.1, as* $T \to \infty$*, it holds that* $\mathcal{T}^{f*} \overset{w}{\Rightarrow}_p \mathcal{T}_\infty^{f,I}$, $\mathcal{T}^{b*} \overset{w}{\Rightarrow}_p \mathcal{T}_\infty^{b,I}$, $\mathcal{T}^{r*} \overset{w}{\Rightarrow}_p \mathcal{T}_\infty^{r,I}$*, and* $\mathcal{T}^{d*} \overset{w}{\Rightarrow}_p \mathcal{T}_\infty^{d,I}$.

(ii) *Under Assumption 1.2 as* $T \to \infty$*, it holds that,* $\mathcal{T}^{f*} \overset{w}{\Rightarrow}_p \mathcal{T}_\infty^{f,II}$, $\mathcal{T}^{b*} \overset{w}{\Rightarrow}_p \mathcal{T}_\infty^{b,II}$, $\mathcal{T}^{r*} \overset{w}{\Rightarrow}_p \mathcal{T}_\infty^{r,II}$*, and* $\mathcal{T}^{d*} \overset{w}{\Rightarrow}_p \mathcal{T}_\infty^{d,II}$.

Moving on to ...

1. Background
2. The Tests
3. Finite Sample Simulations
4. Equity Premium Predictability
5. Concluding Remarks

Monte Carlo design

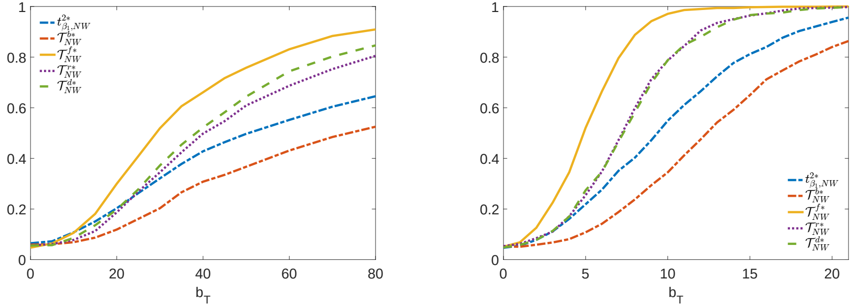
- $\rho := 1 - c/T$ with $c \in \{0,\ 2.5,\ 5,\ 10,\ 20,\ 0.5T\}$ ($c = 0.5T$ corresponds to stable $\rho = 0.5$)
- $(u_t, v_t)'$ is zero-mean IID bivariate Gaussian with covariance matrix $\mathbf{\Sigma}_t := \begin{bmatrix} \sigma_{ut}^2 & \phi\sigma_{ut}\sigma_{vt} \\ \phi\sigma_{ut}\sigma_{vt} & \sigma_{vt}^2 \end{bmatrix}$ and $\phi = -0.9$
- IVX with $a = 1$, $\gamma = 0.95$, and finite-sample correction (see Kostakis *et al.*, 2015)
- Also examine statistics with ordinary standard errors (Non-White)
- $\tau_L = 1/4$ and $\tau_U = 3/4$ for forward and backward recursive statistics, $\Delta\tau = 1/3$ for rolling and double recursive statistics
- 5000 MC replications, $B = 399$

Some size results

c	$t^{2*}_{\beta_1}$	t^{2*}_{NW}	$t^{2}_{\beta_1}$	t^{2}_{NW}	$\mathcal{T}^{f*}$	$\mathcal{T}^{b*}$	$\mathcal{T}^{f*}_{NW}$	$\mathcal{T}^{b*}_{NW}$	$\mathcal{T}^{r*}$	$\mathcal{T}^{r*}_{NW}$	$\mathcal{T}^{d*}$	$\mathcal{T}^{d*}_{NW}$
DGP1: $T = 250$, $\phi = -0.90$ and $\sigma^2_{ut} = \sigma^2_{vt} = 1$												
0.0	0.069	0.073	0.069	0.074	0.037	0.035	0.047	0.055	0.011	0.053	0.018	0.055
2.5	0.055	0.056	0.055	0.057	0.030	0.040	0.037	0.058	0.011	0.052	0.020	0.062
5.0	0.053	0.052	0.051	0.050	0.034	0.045	0.039	0.055	0.012	0.051	0.021	0.064
10.0	0.057	0.055	0.056	0.058	0.038	0.048	0.044	0.059	0.017	0.055	0.029	0.065
20.0	0.060	0.060	0.057	0.056	0.046	0.051	0.050	0.062	0.029	0.061	0.034	0.067
$0.5T$	0.055	0.053	0.051	0.052	0.059	0.067	0.057	0.060	0.073	0.058	0.067	0.055
DGP2: $T = 250$, $\phi = -0.90$ and $\sigma^2_{ut} = \sigma^2_{vt} = 1\mathbb{I}(t \leq \lfloor 0.5T \rfloor) + 4\mathbb{I}(t > \lfloor 0.5T \rfloor)$												
0.0	0.057	0.057	0.044	0.071	0.032	0.031	0.026	0.046	0.009	0.030	0.018	0.021
2.5	0.052	0.053	0.044	0.076	0.035	0.032	0.027	0.051	0.011	0.028	0.012	0.031
5.0	0.054	0.053	0.046	0.076	0.037	0.037	0.033	0.054	0.013	0.032	0.014	0.038
10.0	0.054	0.051	0.046	0.077	0.042	0.043	0.037	0.055	0.018	0.039	0.019	0.051
20.0	0.054	0.052	0.049	0.080	0.047	0.046	0.042	0.054	0.032	0.045	0.036	0.055
$0.5T$	0.058	0.055	0.053	0.097	0.067	0.059	0.054	0.057	0.063	0.056	0.073	0.058
DGP3: $T = 250$, $\phi = -0.90$ and $\sigma^2_{ut} = \sigma^2_{vt} = 1\mathbb{I}(t \leq \lfloor 0.5T \rfloor) + \frac{1}{4}\mathbb{I}(t > \lfloor 0.5T \rfloor)$												
0.0	0.072	0.078	0.125	0.118	0.029	0.036	0.049	0.076	0.007	0.072	0.009	0.084
2.5	0.046	0.048	0.059	0.068	0.022	0.046	0.038	0.071	0.008	0.071	0.005	0.077
5.0	0.049	0.048	0.055	0.068	0.027	0.049	0.040	0.064	0.010	0.068	0.006	0.075
10.0	0.058	0.051	0.052	0.073	0.033	0.049	0.041	0.056	0.018	0.061	0.013	0.062
20.0	0.053	0.053	0.049	0.077	0.040	0.052	0.050	0.050	0.028	0.055	0.036	0.061
$0.5T$	0.053	0.052	0.047	0.091	0.058	0.067	0.056	0.056	0.064	0.056	0.070	0.061

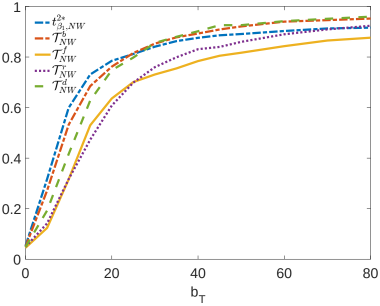
Power plots: early predictability pocket

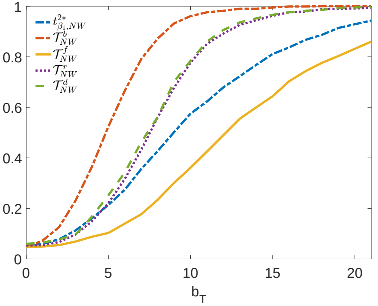
- For $c = 0$ (left), set $\beta_{1t} = \frac{b_T}{T}\mathbb{I}(t \leq T/5)$ with $b_T \in \{0, 5, \ldots, 80\}$
- For $c = 0.5T$ (right), set $\beta_{1t} = \frac{b_T}{\sqrt{T}}\mathbb{I}(t \leq T/5)$ with $b_T \in \{0, 1, \ldots, 21\}$



Power plots: late predictability pocket

- For $c = 0$ (left), set $\beta_{1t} = \frac{b_T}{T}\mathbb{I}(t \geq 4T/5)$ with $b_T \in \{0, 5, \ldots, 80\}$
- For $c = 0.5T$ (right), set $\beta_{1t} = \frac{b_T}{\sqrt{T}}\mathbb{I}(t \geq 4T/5)$ with $b_T \in \{0, 1, \ldots, 21\}$





Moving on to ...

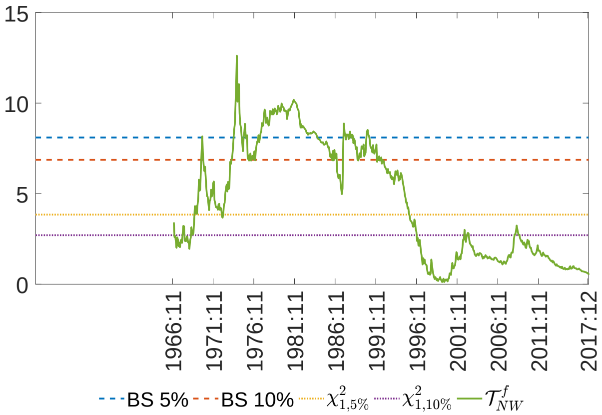
Welch and Goyal data, sample 1950:01-2017:12

Detailed description of the variables used can be found on Amit Goyal's web page.

- The dependent variable, y_t, is the equity premium, EP_t which corresponds to the total rate of return of the S&P 500 index minus a short-term (risk-free) interest rate.
- DP_t is the dividend price ratio - difference between the log of dividends and the log of prices.
- DY_t is the dividend yield - difference between the log of dividends and the log of lagged prices.
- E/P_t is the earnings price ratio - difference between log of earnings and log of prices.
- DE_t is the dividend payout ratio - difference between log of dividends and log of earnings.
- $RVOL_t$ is the equity risk premium volatility.
- $SVAR_t$ is the stock variance computed as sum of squared daily returns on the S&P 500.
- BM_t is the book to market ratio - ratio of book value to market value for the Dow Jones Industrial Average.
- $NTIS_t$ is the net equity expansion - ratio of twelve-month moving sums of net issues by NYSE listed stocks divided by the total market capitalization of NYSE stocks.
- tbl_t is the treasury bill rate.
- lty_t is the long-term government bond yield.
- ltr_t is the long-term government bond rate of return.
- tms_t is the term spread - difference between the long-term yield on government bonds and the treasury bill rate.
- dfy_t is the default yield spread - difference between BAA- and AAA- rated corporate bond yields.
- dfr_t is the default return spread - difference between the return on long-term corporate bonds and returns on the long-term government bonds.
- $INFL_t$ is inflation - the consumer price index (all urban consumers).

Welch and Goyal data 1950:01-2017:12 (FRW bootstrap p-values)

	$t^2_{\beta_1}$	$t^2_{\beta_1,NW}$	$\mathcal{T}^f$	$\mathcal{T}^f_{NW}$	$\mathcal{T}^b$	$\mathcal{T}^b_{NW}$	$\mathcal{T}^r$	$\mathcal{T}^r_{NW}$	$\mathcal{T}^d$	$\mathcal{T}^d_{NW}$	LM_x	$supF_x$
DP	0.472	0.457	10.088	**11.480**	5.608	6.885	6.882	8.280	10.284	12.519	**2.229**	**131.915**
	(0.473)	(0.487)	(0.100)	**(0.006)**	(0.333)	(0.184)	(0.726)	(0.108)	(0.998)	(0.194)	**(0.000)**	**(0.000)**
DY	0.581	0.568	**15.565**	**12.616**	6.241	7.849	11.143	9.318	10.252	**11.891**	**0.295**	**11.178**
	(0.424)	(0.420)	**(0.032)**	**(0.006)**	(0.254)	(0.133)	(0.484)	(0.177)	(0.072)	**(0.029)**	**(0.028)**	**(0.038)**
EP	0.335	0.451	8.459	**9.189**	2.163	4.744	8.583	9.419	8.583	11.742	0.209	**38.229**
	(0.565)	(0.570)	(0.319)	**(0.041)**	(0.621)	(0.409)	(0.375)	(0.152)	(0.990)	(0.231)	(0.116)	**(0.000)**
DE	0.291	0.490	**12.553**	**17.087**	0.291	0.490	13.399	**20.112**	15.145	**21.400**	0.192	5.031
	(0.615)	(0.603)	**(0.022)**	**(0.013)**	(0.872)	(0.860)	(0.077)	**(0.007)**	(0.099)	**(0.010)**	(0.210)	(0.355)
RVOL	1.809	2.288	3.765	4.432	2.657	3.200	4.230	6.187	4.624	6.692	0.124	6.614
	(0.134)	(0.136)	(0.230)	(0.316)	(0.263)	(0.221)	(0.403)	(0.223)	(0.455)	(0.278)	(0.525)	(0.192)
BM	0.037	0.042	7.150	7.125	5.959	7.321	7.612	8.299	7.612	8.299	**0.342**	7.555
	(0.844)	(0.847)	(0.214)	(0.227)	(0.299)	(0.155)	(0.781)	(0.294)	(0.989)	(0.544)	**(0.012)**	(0.148)
NTIS	0.041	0.059	5.634	5.235	1.648	2.600	9.383	10.543	9.679	10.874	0.375	8.102
	(0.843)	(0.838)	(0.148)	(0.293)	(0.622)	(0.560)	(0.061)	(0.070)	(0.101)	(0.114)	(0.070)	(0.180)
TBL	**7.001**	**9.408**	**11.763**	**15.989**	7.001	**9.408**	8.203	**11.618**	11.764	**16.588**	0.131	12.348
	(0.006)	**(0.009)**	**(0.006)**	**(0.001)**	(0.089)	**(0.039)**	(0.307)	**(0.049)**	(0.267)	**(0.029)**	(0.174)	(0.132)
LTY	**4.178**	**5.524**	**11.036**	**13.410**	6.529	6.547	8.224	10.070	11.135	**14.308**	0.103	7.985
	(0.035)	**(0.032)**	**(0.007)**	**(0.002)**	(0.135)	(0.147)	(0.419)	(0.087)	(0.487)	**(0.044)**	(0.303)	(0.182)
LTR	**4.172**	**5.724**	**8.479**	**10.313**	4.438	6.021	**8.145**	9.702	8.902	10.709	0.163	6.407
	(0.030)	**(0.032)**	**(0.025)**	**(0.038)**	(0.115)	(0.147)	**(0.047)**	(0.079)	(0.061)	(0.094)	(0.341)	(0.325)
TMS	2.726	3.075	**15.839**	**17.534**	4.769	5.133	**12.142**	**12.611**	**15.839**	**17.534**	0.350	**10.877**
	(0.091)	(0.094)	**(0.000)**	**(0.001)**	(0.140)	(0.141)	**(0.046)**	**(0.010)**	**(0.042)**	**(0.003)**	(0.060)	**(0.026)**
DFY	0.011	0.020	2.777	3.079	0.996	2.872	7.431	**19.169**	8.805	**19.473**	0.063	9.773
	(0.903)	(0.910)	(0.593)	(0.615)	(0.681)	(0.570)	(0.247)	**(0.035)**	(0.216)	**(0.046)**	(0.788)	(0.106)
DFR	0.768	1.535	4.487	5.475	2.407	5.398	11.218	7.685	11.218	13.141	0.156	6.134
	(0.472)	(0.443)	(0.444)	(0.389)	(0.338)	(0.335)	(0.228)	(0.550)	(0.295)	(0.336)	(0.537)	(0.377)
INFL	3.042	4.890	9.083	**16.834**	3.999	6.138	11.705	11.516	12.195	**16.839**	0.326	**11.517**
	(0.067)	(0.060)	(0.289)	**(0.004)**	(0.259)	(0.147)	(0.479)	(0.053)	(0.722)	**(0.026)**	(0.102)	**(0.032)**



Dividend yield: Forward Recursive IV t-statistics with marginal and bootstrap 10% and 5% critical values.

Moving on to ...

Summing up

- Inference in predictive regressions with regressors of uncertain persistence is challenging.
- Time-varying coefficients and variances add complexity.
- IV-based subsample inference can help ...
- ... in conjunction with a fixed regressor wild bootstrap.
- Evidence of predictability still not overwhelming, but improved compared to full sample methods.