Market-Beta

Ivo Welch

May 2019

Notice to PhD Students

This presentation is intended to teach how ~~not~~ to commit suicide on the job market.

Notice to PhD Students

This presentation is intended to teach how ~~not~~ to commit suicide on the job market.

Because, what you need is

I am not saying it's right.
I am saying I am impressed.

(PS: On the job market, it will be cleverness, not necessarily mathiness, that matters.)

Why Not Job Market?

- My brownbag paper is way too simple,
- ...and it is not about
  - new data,
  - big data,
  - new small data,
  - and/or clever quasi-experimental identification.

Why Not Job Market?

- My brownbag paper is way too simple,
- ...and it is not about
  - new data,
  - big data,
  - new small data,
  - and/or clever quasi-experimental identification.

So why bother?

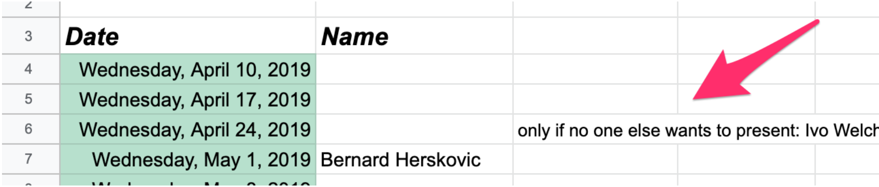
- It’s actually very useful, and
- you will actually want to use this in your lifetime,
- ...and it saved your Wednesday lunch.

3	***Date***	***Name***	
4	Wednesday, April 10, 2019		
5	Wednesday, April 17, 2019		
6	Wednesday, April 24, 2019		only if no one else wants to present: Ivo Welch
7	Wednesday, May 1, 2019	Bernard Herskovic	

- so no complaints, please.

So why bother?

- It's actually very useful, and
- you will actually want to use this in your lifetime,
- ...and it saved your Wednesday lunch.



3	***Date***	***Name***	
4	Wednesday, April 10, 2019		
5	Wednesday, April 17, 2019		
6	Wednesday, April 24, 2019		only if no one else wants to present: Ivo Welch
7	Wednesday, May 1, 2019	Bernard Herskovic	

- so no complaints, please.

Breaks Chernov Rule

- not “serious” research
- instead: this is a “Tinker With Data” paper

Breaks Chernov Rule

- not “serious” research
- instead: this is a “Tinker With Data” paper

Motivation

Why still bother with “boring” old market-beta?

- Market-beta is interesting even w/o CAPM
  - Measure of risk contribution to diversified portfolios.
  - Hedging against bear markets
  - Down-Beta Theories (as in Ang+ or Lettau+)
  - Betting against Beta (as in Frazzini-Pedersen)
  - Pragmatic: used in regulation, etc.
- How should we estimate beta?
  - And can it make a difference?

Performance Metric

I will judge beta quality by prediction.

- future ols(/other) 1-mo or 1-yr market-beta **estimates**
- never future average returns.

- PS: If two proxies are drawn with noise from true value, the expected R^2 of each proxy with the true value is the squareroot of the R^2 of one proxy with the other proxy.

Performance Metric

I will judge beta quality by prediction.

- future ols(/other) 1-mo or 1-yr market-beta **estimates**
- never future average returns.

- PS: If two proxies are drawn with noise from true value, the expected R^2 of each proxy with the true value is the squareroot of the R^2 of one proxy with the other proxy.

Best Beta Estimator Known: Vasicek

- random-effects estimator = bayesian shrinkage
- Run OLS Regressions
- Calculate x-sect means and sds of betas
- For each stock,

$$b_{vck} = w \cdot b_{xs} + (1 - w) \cdot b_{ts},$$

where $w = \sigma_{ts}/(\sigma_{ts} + \sigma_{xs})$.

Other Important Choices

- Always use daily stock returns
- about 1-3 years of data.
- Never use industry beta for individual stocks.
  - Indeed, less noisy;
  - but just like using “1” — low predictive power.
- vasicek and its derivatives
  - (random-effects and/or bayesian justification if no drift.)
  - Levi-Welch linear de-bias.

more alternatives below: Dimson, Frazzini-Pedersen

Vasicek Disadvantages

- "Pseudo Optimal"
  - "optimal design" was never suited to problem:
  - vasicek is designed for measurement error,
  - not for underlying beta drift
  - (ergo 12–24 months windows)
- good R^2, but badly biased
  - levi-welch (2017) suggests empirical de-biasing
  - requires another stage
- spooky entangled estimates
- requires multi-step ts and xs procedure

Better and Simpler Estimator

Standard Bayesian Use of Prior

$$\begin{tikzcd}
\text{Data} \arrow[dr] & \\
& \text{Posterior} \\
\text{Priors} \arrow[ur] &
\end{tikzcd}$$

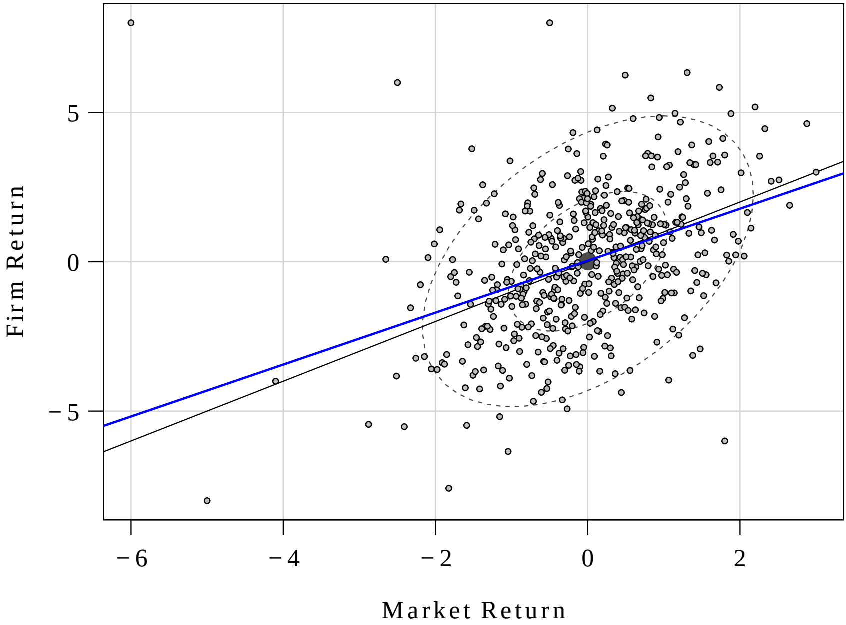
- Involves arguments about reasonable priors
- Often painful—days babysitting, not minutes.
- Usually primarily in dedicated papers

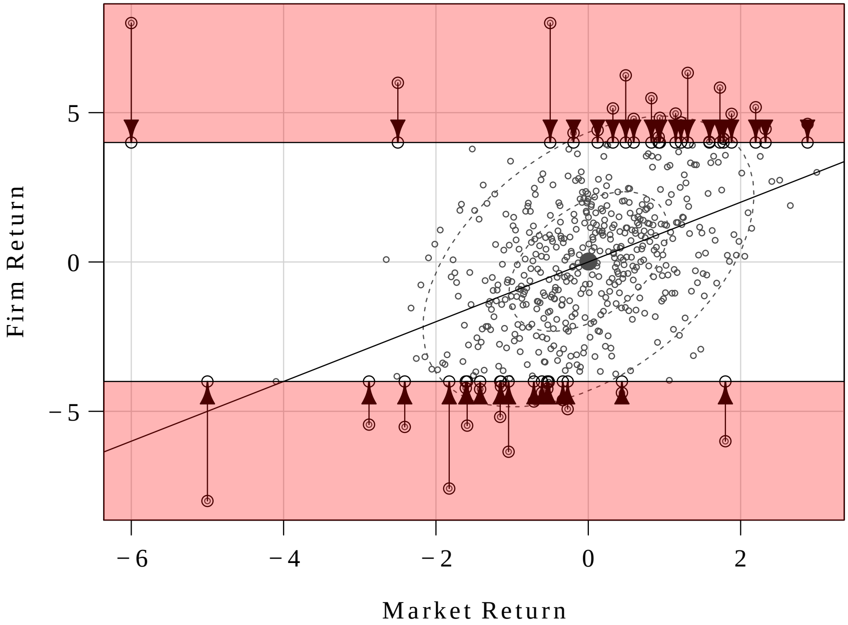
(Ab-)Use of Prior

$$\begin{tikzcd} \text{Data} \arrow[dr] & \\ & \text{Posterior} \\ \text{Priors} \arrow[uu] & \end{tikzcd}$$

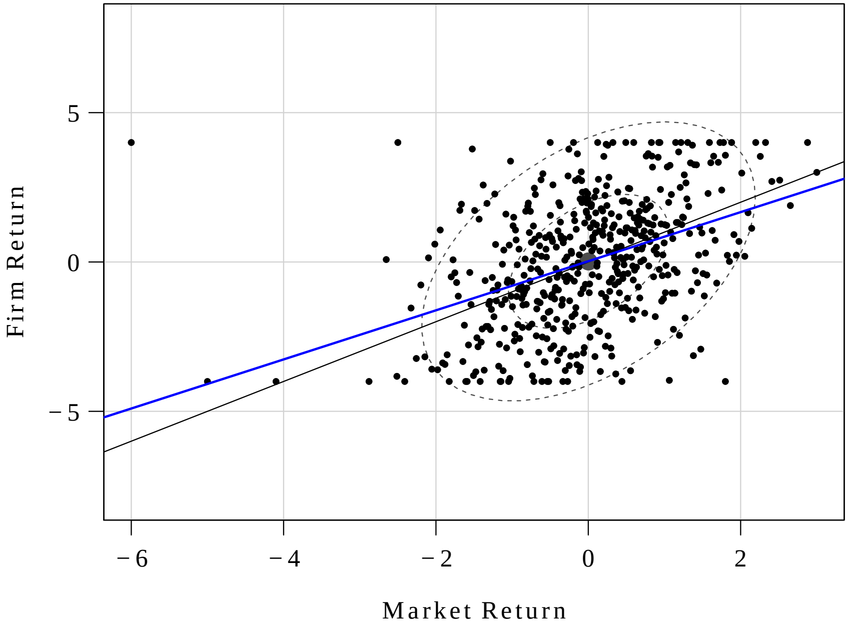
- Still involves arguments about reasonable priors
- Easy to use. minutes, not days.

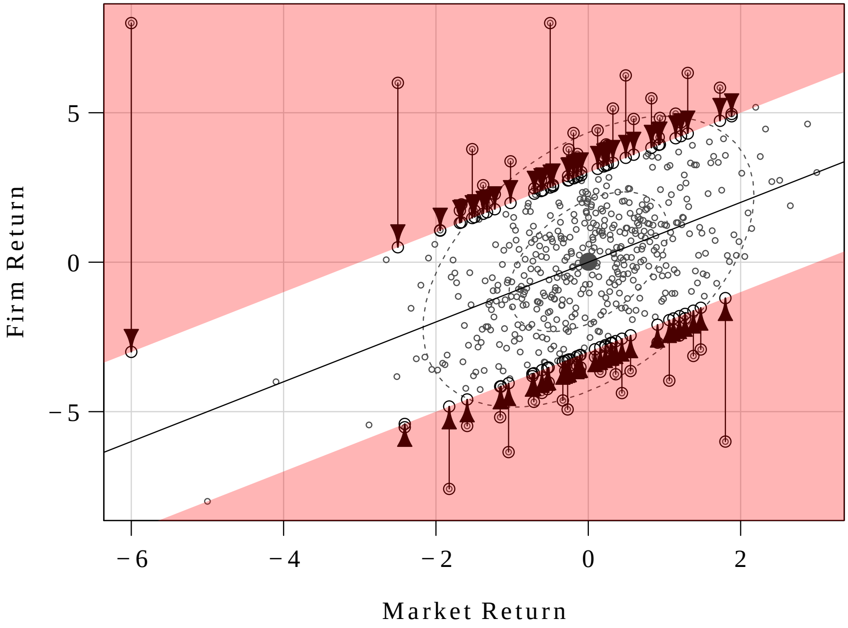
- Likely novel method.

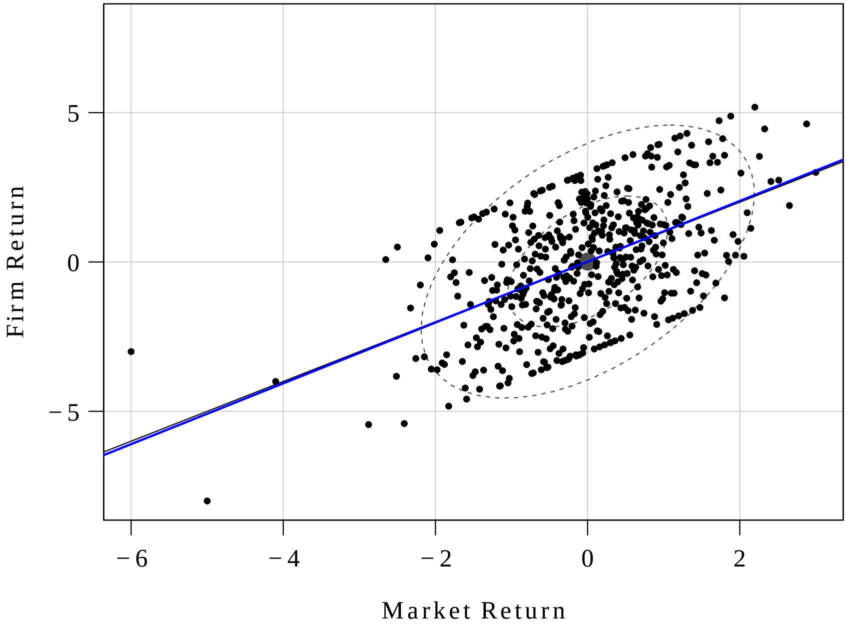

5
0
−5
−6
−4
−2
0
2
Firm Return
Market Return

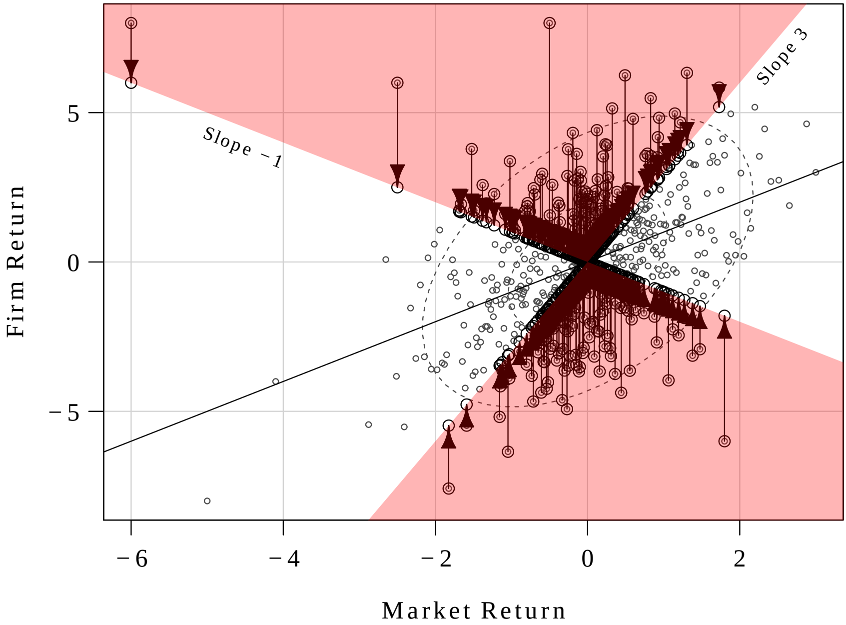


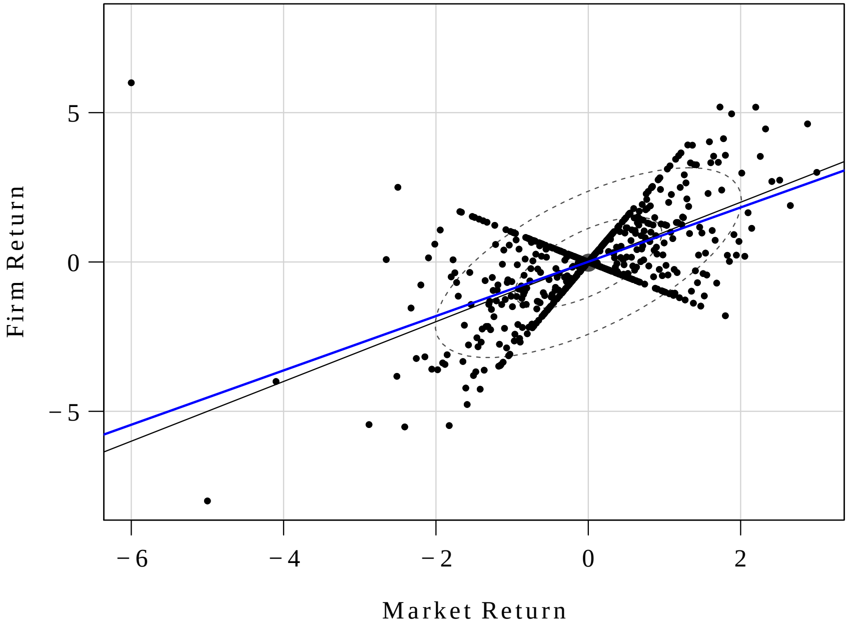
5
0
−5
−6
−4
−2
0
2
Firm Return
Market Return


5
0
−5
−6
−4
−2
0
2
Firm Return
Market Return


5
0
−5
Firm Return
−6
−4
−2
0
2
Market Return


Firm Return
Market Return
5
0
−5
−6
−4
−2
0
2


Slope 3
Slope −1
5
0
−5
−6
−4
−2
0
2
Firm Return
Market Return


Firm Return
Market Return
5
0
−5
−6
−4
−2
0
2

- non-Bayesian use of prior
- with wide priors, not very costly, even if panel is true OLS w/o outliers.

- Note: Even if $\hat{b} = 0.8$, the prior is still effective on individual points.
  - Bayesian OLS-type prior would work on overall $\hat{b}$ estimate.
  - If final is near 1.0, Bayesian method says "just fine."
  - Here, a prior(–1,3) still influences points, and thus even estimates close to 1.0. $\rightarrow$ can move a $\hat{b}$ away from 1.0.
  - PS: could use Bayesian with priors on mixed distributions, plain + outliers. Would work, too, but far more painful.

beta slope winsorized (bsw)

1. 12–24 mos of **daily** stock returns
2. winsorize all returns ($\Delta_s = 2$):

$$rsw_{i,t} \in \; 1.0 + \left[-\Delta_s, \Delta_s \right] \cdot r_{m,t} \, .$$

3. estimate ols market-model

$$rsw_{i,t} = a_i + bsw_i \cdot r_{m,t}$$

(just a reuse of the model with a reasonable prior. note: model-specific.)

why $\Delta_s = 2$?

- fewer than 1% of betas exceed –1 and +3
- fewer than 0.03% repeat in consecutive years
  - (greater than $1\% \cdot 1\%$, but not by much.
- beyond, no monotonicity between b_t and $E(b_{t+1})$
- not philosophical, but also not highly searched:
  - you could also use [–0.5, 2.5] or [–3, 5].
  - lower Δ_s forces too much towards 1.
  - higher Δ_s forces nada.

does it matter?

are betas even different?

$$\text{rmsd}\,(\,\text{bols}_{\text{D}}\,,\,\text{bsw}\,) \approx 0.37$$

$$\text{rmsd}\,(\,\text{bvck}_{\text{D}}\,,\,\text{bsw}\,) \approx 0.20$$

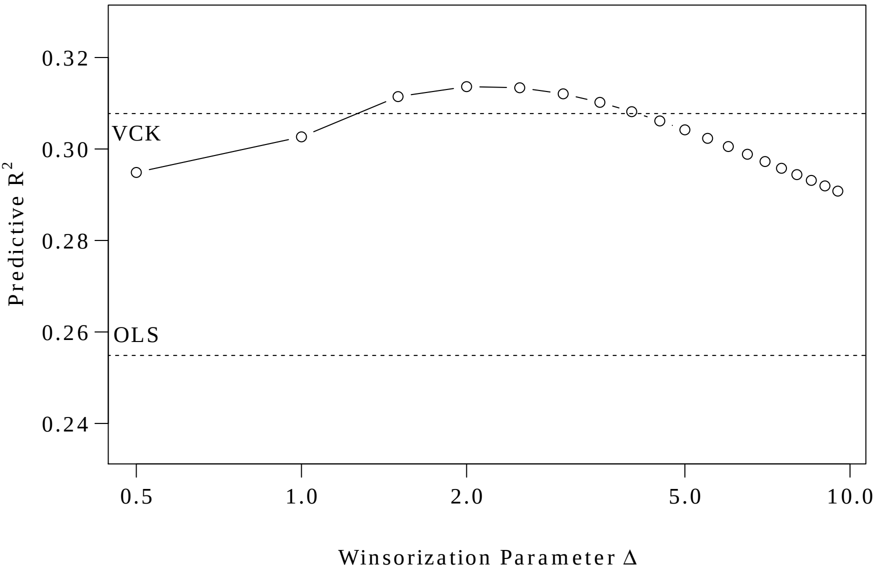
$$\text{rmsd}\,(\,\text{bols}_{\text{M}}\,,\,\text{bsw}\,) \approx 0.60$$

“gamma” panel reg for bols_{t+1}

Dependent: future 1-year **ols** beta from daily returns, same set.

	γ_0	se(γ_0)	γ_1	se(γ_1)	R^2
(bols)	0.34	.004	0.54	.005	25.5%
(bvck)	0.19	.002	0.74	.002	30.8%
... (blw)	–0.01	.003	0.98	.003	same
level (blw)	0.27	.002	0.70	.003	29.7%
band (bbw)	0.04	.002	0.93	.003	30.9%
slope (bsw)	0.01	.002	0.96	.003	**31.4%**
slope + v	–0.01	.003	1.00	.003	31.5%

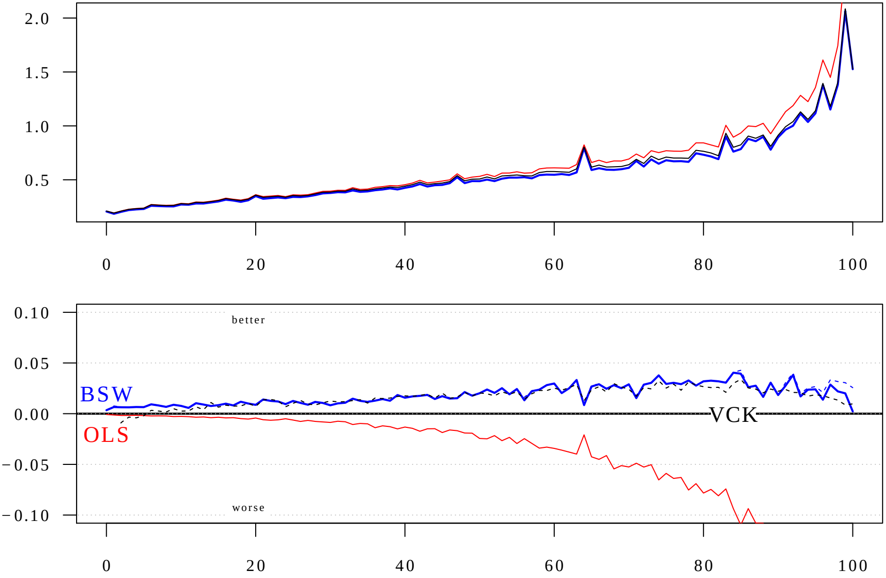
True R^2 is squareroot. $\sqrt{.3} \approx 0.55$.


0.32
0.30
0.28
0.26
0.24
Predictive R²
VCK
OLS
0.5
1.0
2.0
5.0
10.0
Winsorization Parameter Δ

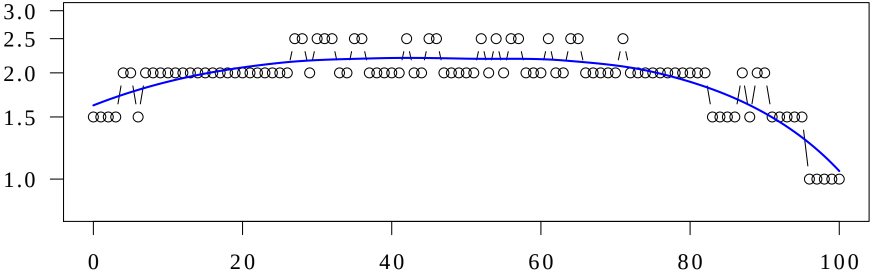
Nothing Sensitive or Edgy

- very stable by year.
- very stable by ols beta.
- no meaningful improvement by varying Δ_S.
  - even by own lagged beta, beta-sd, marketcap, trading volume, volatility, etc.

RMSE by ols se(b) Percentile



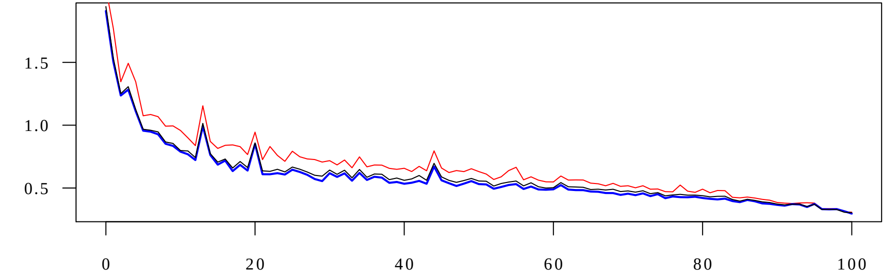
Δ^* by ols se(b) Percentile

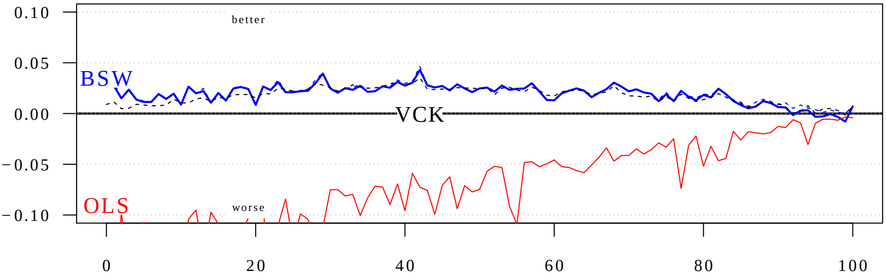


possible improvements: obtain ols b se rank, then

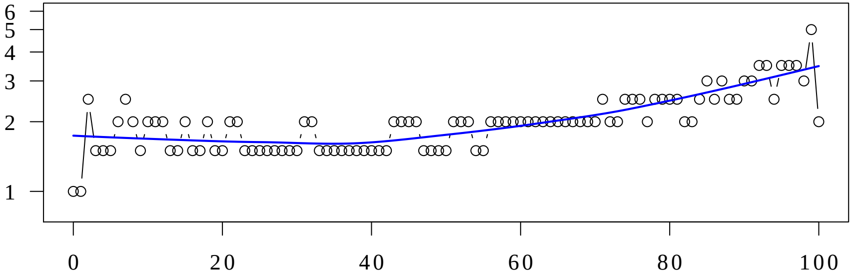
- more winsorization ($\Delta_S = 1.5$) for $> 85\%$ and 1 for $> 95\%$.
- but stable from 5 to 80 and look at absolute improvement,
- only the 95%+ do.
- $\Rightarrow$ 1st-stage firm-specific deltas won't help much, on avg.

RMSE by Market Cap Percentile





Δ^* by market cap percentile



possible improvements: obtain mcap rank, then

- more winsorization ($\Delta_s = 1.5$) for small-caps (rank < 40%),
- less winsorization ($\Delta_s = 3$) for big-caps (rank > 80%).

Another 2% R^2 Improvement

- retain 1-pass simplicity of use
- WLS market-model, w=f(age)

steep exponential decline:

	Now	3 mo	6mo	1yr	2yr
WLS w	100%	80%	50%	10%	2%

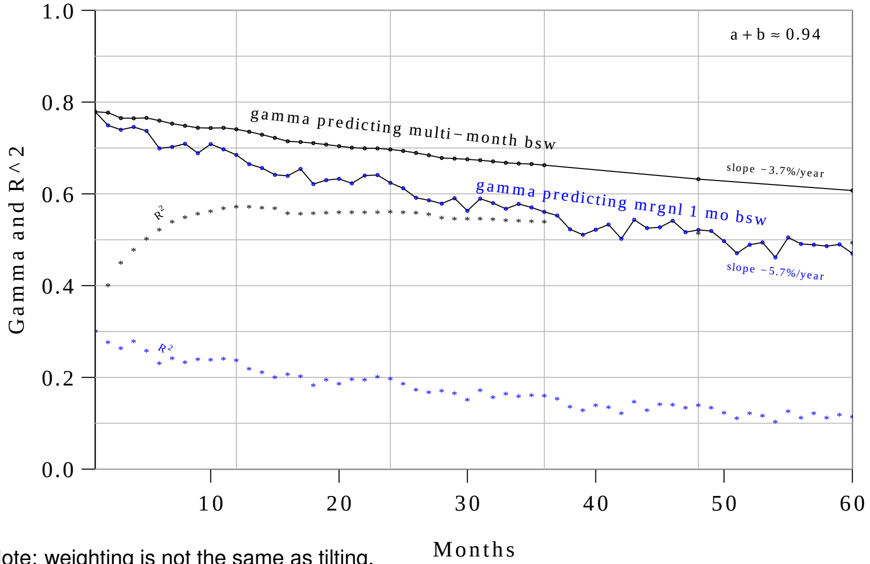
- PS: this WLS decay allowed for a kink
- want approx formula? $\approx$ exp[–2(Δdays)/252]

- trivially easy in time
- no marginal loss of observations

PS: Mean Reversion on Monthly BSW

- not fixing outliers suggests faster mean reversion of beta
- need to estimate mean reversion of betas after fixing outliers

PS: Mean Reversion on Monthly BSW



Note: weighting is not the same as tilting.

Note: dependent ne OLS, but BSW. Indep is WLS.BSW. Daily Stock Returns.

another 1% R^2 improvement

- no longer simple, 1-pass, no-obs-loss
- add one extra variable reflecting firm-size or dollar trading volume.
  - big firms have bigger market-betas (yes!),
  - but use requires first-stage regression,
  - and marketcap requires merging, data loss, etc.
- I could find no other useful accounting compustat or crsp derived variable or ratio.

Dimson + Frazzini-Pedersen

care for		R^2 with x_t being only			
$y_{t+1} \downarrow$	ols	bsw	vck	dim	fp
ols	38%	44%	43%	28%	27%

(Monthly-overlaps)

Dimson + Frazzini-Pedersen

care for	R^2 with x_t being only				
$y_{t+1} \downarrow$	ols	bsw	vck	dim	fp
ols	38%	44%	43%	28%	27%
vck		51%	50%		
bsw		57%	$\Rightarrow$ R^2 to β_{true} should be $\approx$ 75%		
dim		30%		22%	
fp		30%			21%

$\rightarrow$ what should you use if you care (but why?) about future dimson or fp estimates?
(Monthly-overlaps)

if you are interested in future Dimson beta,

$\rightarrow$ use current bsw

never use current Dimson beta as estimator

if you are interested in future Frazzini-Pedersen beta.

$\rightarrow$ use current bsw

never use current FP beta as estimator

did they ever try to validate their measures?

Monthly-Frequency Return Data?

- even long-window monthly betas are miserable predictors of anything (like R^2 of $< 15\%$, not 40%).
- daily predicts monthly better than monthly itself.
- $\rightarrow$ use daily frequency even if interested in future monthly market betas.

Future

Can some of this be generalized?

- To what extent can we use our prior information to manipulate the incoming data first,
- and then run plain classical procedures,
- because Bayesian methods are so painful that only dedicated B papers are using them.
- (e.g., stick fitted values w/ se [as weights?] from 1st-stage OLS into 2nd-stage OLS?)

Conclusion

- novel slope winsorization method afaik, with use of prior in different way,
- novel application of winsorization method in important context of market-beta estimation.
- only simple use of prior. no 1st stage needed.
- superb ease of use. pto.

So Why Not?

```
wins.rel <- function( r, rmin, rmax ) {
   rl <- ifelse( (rmin<rmax), rmin, rmax )
   ru <- ifelse( (rmin<rmax), rmax, rmin )
   ifelse( r<rl, rl, ifelse(r>ru, ru, r) )
}

delta <- 2
wri <- wins.rel( ri, (1-delta)*rm, (1+delta)*rm )


beta <- function(...) coef(lm(...))[2]

bsw <- beta( wri ~ rm )

wbsw <- beta( wri ~ rm, w=exp(-2*(length(ri):1)/256) )
                ## note age = reverse-time weights
```