Predictive Regression and Robust Hypothesis Testing: Predictability Hidden by Anomalous Observations

Lorenzo Camponovo   Olivier Scaillet   Fabio Trojani

Discussion by Grigory Vilkov

23-27 July 2012

European Summer Symposium in Financial Markets in Gerzensee
Fact 1: predictive model estimation:

\[
\begin{align*}
y_t &= \alpha + \beta x_{t-1} + u_t \\
x_t &= \mu + \rho x_{t-1} + v_t
\end{align*}
\]

especially with \( \rho \approx 1 \),

→ small sample bias

Fact 2: plenty of methods to get test stats:

→ Bias correction methods
→ Local-to-unity asymptotic methods
→ Bootstrap
→ Subsampling tests

Fact 3: non of the above work when we have outliers!
{as also shown in controlled MC environment}

⇒ Houston, we have a problem
Sensitivity to outliers

1. Simulations: common methods are **not resistant to outliers**

2. **Quantile breakdown point (QBP):** largest fraction of anomalous observations for bootstrap/subsampling to be reliable

3. QBP and non-robust statistic: single anomalous observation \(\rightarrow\) arbitrary large effect

4. QBP for robust statistic (QBP \(>\) \(\frac{1}{n}\)): bootstrap/subsampling QBP ↓

5. **Bootstrap/subsampling with robust statistics not resistant to outlier!**
Develop robust (and fast) bootstrap and subsampling

1. Consider robust M-estimators defined as solution $\hat{\theta}_n$ to

$$
\psi_n(X(n), \hat{\theta}_n) := \frac{1}{n} \sum_{i=1}^{n} g(X_i, \hat{\theta}_n) = 0
$$

with bounded $g(\cdot)$.

2. Bootstrap/resampling: $\psi_k(X_{(n,m)}^{K*}, \hat{\theta}_{(n,m)}^{K*}) = 0$, $\forall$ sample $X_{(n,m)}^{K*}$

Instead: use Taylor expansion of moment conditions:

$$
\hat{\theta}_n - \theta_0 = -[\nabla_{\theta} \psi_n(X(n), \theta_0)]^{-1} \psi_n(X(n), \theta_0) + \text{err}
$$

→ fast resampling distribution that uses
full sample $- [\nabla_{\theta} \psi_n(X(n), \theta_n)]^{-1}$ and resampled $\psi_k(X_{(n,m)}^{K*}, \theta_n)$
3 Distribution explodes iff (i) $\nabla_\theta \psi_n(\cdot)$ singular, or (ii) $g(\cdot)$ not bounded

4 Infer **breakdown point** of M-robust estimator as $\min(b, b_{\nabla \psi})$, where $b_{\nabla \psi}$ is the breakdown point of the matrix

Finally: robust predictive regression and hypothesis testing

1 Select a robust estimator with QBP $> 1/n$, e.g., Huber estimator $\hat{\theta}_n^R$, s.t. the **estimating function** is given by

$$
g_c(y_t, w_{t-1}, \theta) = (y_t - \theta' w_{t-1}) w_{t-1} \cdot \min\left(1, \frac{c}{\| (y_t - \theta' w_{t-1}) w_{t-1} \|} \right).
$$
Now iteratively find optimal $c$ ($\approx$ degree of robustness) and block size $m$

$$(m, c) := \arg \inf_{(m, c) \in MC} \left\{ \left| t - P^*[\hat{\theta}_n^R \in CI_{t, (m, c)}] \right| \right\}$$

where $P^*$—bootstrap distribution probability

3 Derive nonstudentized and studentized tests using fast resampling distribution, optimized to be robust to anomalous observations.
P/D predicting returns
P/D and VRP predicting returns

Discussion by Grigory Vilkov

Always improves the statistics for rejecting the non-predictability

Figure 1: $\ln(R_t) = \alpha + \beta \log(D_{t-1}/P_{t-1}) + \epsilon_t$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias-Corrected</td>
<td>0.0292**(**)</td>
<td>0.0167**(**)</td>
<td>0.0191**(**)</td>
<td>0.0156</td>
</tr>
<tr>
<td>Bonferroni</td>
<td>0.0236**(**)</td>
<td>0.0134**(**)</td>
<td>0.0117(*)</td>
<td>0.0112</td>
</tr>
<tr>
<td>Bootstrap</td>
<td>0.0430**(**)</td>
<td>0.0175(*)</td>
<td>0.0306**(**)</td>
<td>0.0355**(**)</td>
</tr>
<tr>
<td>Subsampling</td>
<td>0.0430**(**)</td>
<td>0.0175</td>
<td>0.0306**(**)</td>
<td>0.0355**(**)</td>
</tr>
<tr>
<td>R.Bootstrap</td>
<td>0.0405**(**)</td>
<td>0.0174**(**)</td>
<td>0.0245**(**)</td>
<td>0.0378**(**)</td>
</tr>
<tr>
<td>R.Subsampling</td>
<td>0.0405**(**)</td>
<td>0.0174**(**)</td>
<td>0.0245**(**)</td>
<td>0.0378**(**)</td>
</tr>
</tbody>
</table>

Impressions? Surprises?

Always improves the statistics for rejecting the non-predictability
On a positive side:

Theory

- Resolve small sample bias problem in the presence of outliers
- Theoretically sound and well-developed
- Diagnostic tools: QBP—“explosion” of the bootstrap distribution

Model

- Gets “the best data” for your model, i.e., treats your model well, and
- ⇒Always improves the model fit (see previous slide!)
Data

- Model-based formal way to treat “anomalous” observations
- Very useful method of identifying model-based “outliers”

Figure 2: Huber weights for \( \ln(R_t) = \alpha + \beta \log(D_{t-1}/P_{t-1}) + \epsilon_t \)

4.2% of observations are anomalous
Houston, where are we? III

Major concern: intelligently squeeze the data into the model

- Add two degrees of freedom in the estimation procedure: “anomalous” observation weight and block size
- Selecting the parameters by shrinking the outstanding observations so that with estimated parameters are realistic according to the model
- What if the data are correct and the model is wrong?
Suggestions/ extensions

Model selection/ treatment

- Nested models, non-nested models: how to choose the right one?
  → handle with extreme care: danger of selecting a wrong model
  → restrict the “data-changing” ability of the approach
  → discuss/add model selection tools

- Conditional models, instrumental variables?

- Cross-sectional models, monotonicity tests?

Technical/ diagnostics

- Taylor series expansion of the estimator: any effect?
  → try full implementation with a simple estimator

- Comparison to “traditional” robust trimmed/winsorised estimators
  → what is the equivalent “traditional” procedure?

- Sensitivity to degree of robustness and block size
Interpretation

- Comparison to shrinkage/Bayesian methods: noise–information balance

Misc

- Desperately need some code to download and test

4 Another practical problems of interest

- Estimate the mean? Less noisy over time?
Conclusion

Not fully persuaded yet, but looks good!

Good luck with the paper!