Discussion of:

“Correcting Misspecified Stochastic Discount Factors”

by Raman Uppal, Paolo Zaffaroni, Irina Zviadadze

Christian Julliard

London School of Economics

The big picture

An old idea:

- Asset returns are an adapted process to the SDF one.

$\Rightarrow$ I. use returns to learn about the latent SDF.

II. given a candidate SDF, use asset returns to estimate what the SDF is "missing"

This paper: focus on II assuming a linear factor structure for asset returns.

In a nutshell: construct a linear correction (α-SDF) using the implied cross-sectional pricing errors of a candidate SDF (the β-SDF).

HJ1: min variance SDF

Hansen-Jagannathan (1991 JPE):

Definition (Canonical *HJ*-bound)

For each $\mathbb{E}[M_t] = \bar{M}$, the minimum variance SDF is

$$M_t^*\left(\bar{M}\right) \equiv \underset{\left\{M_t\left(\bar{M}\right)\right\}_{t=1}^{T}}{\arg\min} \sqrt{Var\left(M_t\left(\bar{M}\right)\right)} \text{ s.t. } \mathbf{0} = \mathbb{E}\left[\mathbf{R}_t^e M_t\left(\bar{M}\right)\right] \qquad (1)$$

where $\mathbf{R}_t^e \in \mathbb{R}^N$.
The solution to the above is $M_t^*\left(\bar{M}\right) = \bar{M} + \left(\mathbf{R}_t^e - \mathbb{E}[\mathbf{R}_t^e]\right)' \beta_{\bar{M}}$, where $\beta_{\bar{M}} = Cov\left(\mathbf{R}_t^e\right)^{-1}\left(-\bar{M}\mathbb{E}[\mathbf{R}_t^e]\right)$, and any candidate SDF M_t must satisfy $Var\left(M_t\left(\bar{M}\right)\right) \geq Var\left(M_t^*\left(\bar{M}\right)\right)$.

⇒ $M_t^* \equiv$ projection of true SDF on space of payoffs.

- Dimensionality of order $N^2 \rightarrow$ problem with large N.

Note: violates non-negativity restriction (HJ1 provides a restricted one, but computationally more complex).

HJ2: min variance correction to the SDF

Hansen-Jagannathan (1997 JF):

Definition (*HJ*-correction)

Given a candidate SDF M_t:

$$d_{HJ}^2 \equiv \min_{q \in L^2} \mathbb{E}\left[\left(M_t - q_t\right)^2\right] \quad \text{s.t. } \mathbf{0} = \mathbb{E}\left[q_t \mathbf{R}_t^e\right].$$

⇒ HJ2 looks for the minimum (in a least square sense) linear adjustment that makes $M_t - \theta' \mathbf{R}_t^e$ an admissible SDF (where θ arises from the linear projection of M_t on the space of returns).

- Again, dimensionality of order N^2 and violates non-negativity restriction.

Changing measures

- Consider the vector of Euler equations

$$\mathbf{0} = \mathbb{E}\left[\underbrace{m(\theta, t)\,\psi_t}_{M_t}\mathbf{R}_t^e\right] \equiv \int m(\theta, t)\,\psi_t \mathbf{R}_t^e dP$$

where $m(\theta, t)$ is a known function of observable data, P is the physical probability measure, and ψ_t is an unobservable component.

- Under very weak regularity conditions, we have

$$\mathbf{0} = \int m(\theta, t)\,\frac{\psi_t}{\bar{\psi}}\mathbf{R}_t^e\,dP = \int m(\theta, t)\,\mathbf{R}_t^e\,d\Psi = \mathbb{E}^{\Psi}\left[m(\theta, t)\,\mathbf{R}_t^e\right]$$

where $\bar{x} := \mathbb{E}[x_t]$, and $\frac{\psi_t}{\bar{\psi}} = \frac{d\Psi}{dP}$ is the Radon-Nikodym derivative.

Note:
- If $m(\theta, t)$ is a constant, $\Psi \equiv Q$
- if not, ψ is a multiplicative correction of the candidate sdf m

GJT: minimum entropy SDF and correction

Ghosh, Julliard, Taylor (2016 RFS):

Definition (*GJT*-SDF)

Given an $m(\theta, t)$ and a returns data, estimate the Ψ measure as

$$\hat{\Psi} = \arg\min_{\Psi} D(\Psi || P) \equiv \arg\min_{\Psi} \int \frac{d\Psi}{dP} \ln \frac{d\Psi}{dP} dP \text{ s.t. } \mathbf{0} = \int m(\theta, t) \mathbf{R}_t^e d\Psi$$

⇒ KLIC minimization under the asset pricing restriction.

- Since relative entropy is not symmetric, we can also use $D(P||\Psi)$.
- Dimensionality of order N, and guarantees non-negativity.
- ML interpretation and properties.
- HJ1/HJ2 as approximations/particular cases.
- ψ adds minimum amount of additional info needed to to price assets.

Note: correction is not necessarily orthogonal to m. And not, in the data, for consumption based models.

Out-of-sample (GJT 2018): prices assets (equities, commodities, currencies...) better than usual factors, and delivers the maximum Sharpe ratio returns (e.g. better than $1/N$, momentum+value etc.).

ψ correction à la HJ1

Definition (Volatility bound for ψ_t)

For each $E[\psi_t] = \bar{\psi}$, the minimum variance ψ_t is

$$\psi_t^*(\bar{\psi}) \equiv \underset{\{\psi_t(\bar{\psi})\}_{t=1}^T}{\arg\min} \sqrt{Var(\psi_t(\bar{\psi}))} \text{ s.t. } \mathbf{0} = \mathbb{E}\left[\mathbf{R}_t^e m(\theta, t)\psi_t(\bar{\psi})\right].$$

The solution of the above minimization for a given $m(\theta, t)$ is

$$\psi_t^*(\bar{\psi}) = \bar{\psi} + \left(\mathbf{R}_t^e m(\theta, t) - \mathbb{E}\left[\mathbf{R}_t^e m(\theta, t)\right]\right)' \beta_{\bar{\psi}}$$

where $\beta_{\bar{\psi}} = Var\left(\mathbf{R}_t^e m(\theta, t)\right)^{-1}\left(-\bar{\psi}\mathbb{E}\left[\mathbf{R}_t^e m(\theta, t)\right]\right)$

$\Rightarrow$ $\psi_t^* \equiv$ projection on space of scaled payoffs.

Note: correction is not necessarily orthogonal to m

... just one more factor

Note: in population or in sample, we are always one factor away from perfect pricing (e.g. MacKinlay (1995))

Example: consider a model with, as observable factors, k assets with excess returns in vector $z_{p,t}$

$$\mathbf{R}_t^e = \alpha + B z_{p,t} + \epsilon_t$$

$$\mathbb{E}\epsilon_t = 0, \quad Var[\epsilon_t] = \underset{N \times N}{\Sigma}, \quad \mathbb{E}z_{p,t} = \mu_p, \quad Var[z_{p,t}] = \underset{K \times K}{\Omega}, \quad cov[z_{p,t}, \epsilon_t] = 0$$

The efficient portfolio of the residual assets is then characterized by

- weights: $\Sigma^{-1}\alpha$
- return: $R_{h,t}^e \propto \epsilon_t' \Sigma^{-1}\alpha$ i.e. $\perp z_{p,t}$
- squared SR: $s_h^2 \equiv \alpha' \Sigma^{-1} \alpha$

$\Rightarrow$ adding this portfolio as a factor we achieve perfect pricing.

This paper: add the linear correction $\epsilon_t' \Sigma^{-1}\alpha$ to the SDF (α-SDF).
Note: dimensionality is still N^2, and still violates non-negativity... but good properties as $N \to \infty$ (under some sort of weighted square integrability of the betas assumption)... but paper never use the limiting results in estimation...

So, what's new here?

$\Rightarrow$ structure on α:

$$\alpha = a + A\lambda_{miss} \;\; s.t. \;\; a'\Sigma^{-1}a \leq \delta < \infty \;\; \forall N$$

where A and λ_{miss} are, respectively, the loading and risk premia of the missing factors (identified by the diverging eigenvalues of Σ as $N \to \infty$).

But: that's the "extended APT" of Uppal and Zaffaroni (2017)... say it!

So, what's really new here? ... I think: how to estimate the missing factors... But:

1. that's only in Appendix E!
2. that's only via Q-MLE (i.e. consistent but not efficient)
3. N^2 order of dimensionality (no use of limit results)
4. that's (presented at least) only for the case in which observable and latent factors are orthogonal...

Note: the latter is never the case in popular structural models (e.g. viewing LRR, habits, heterogeneous agent models etc. as corrections to the C-CAPM), and not at all what GJT find (for consumption models at least)... and cannot correct spurious factors with orthogonal additions.

Suggestions

I. needs clear product differentiation – verbatim identical theorems across (not cited) different papers ain't cool.
II. too much time dedicated to known results – focus on what is new! How to estimate the "extended APT" is non-trivial, and worth a paper.
III. The case of missing factors orthogonal to the observed ones is uninteresting – focus on the relevant case.
IV. The constraint $a'\Sigma^{-1}a \leq \delta < \infty$ is crucial – de facto, that's what delivers the identification. For too low δ you are imposing perfect pricing for any T (everything is latent factors). How to chose δ? Cross-validation? Shanken and Barillas (2018 JF)? Sub-sampling in N?
V. Need a clearer case to support the approach. The critique of HJ1, HJ2 etc. based on N^2 dimensionality is a red herring since, in the current estimation procedure, you have the same problem...
VI. Do asset pricing out-of-sample to show validity of the method... but that's a function of $\hat{\Sigma}^{-1}$...
VII. Ideally, we should learn from the data something about the behavior of the true SDF – use your method to dig!

In summary

(+) The paper is after an important question

(+) Worth writing an empirical "how to" paper about the "extended APT" (or merge papers?)

(+) Large N property is very good (but typo in Theorems or Lemmata?)… use it for estimation!

(-) currently, it takes quite some effort to find what is new

(-) the dimensionality argument in favor of the approach is, at this stage, misleading

(-) Needs a criterion for choosing δ

(=) an interesting paper, with a lot of upside potential $\Rightarrow$ looking forward to the next draft!

Note: it could all be presented in a much simpler fashion: start from the "magic" $\Sigma^{-1}\alpha$ portfolio, cite the "extended APT" as a way of putting economic restrictions on it, and write a clear "how to" paper with a salient empirical application.