

Semiparametric SDF Estimators for Pooled, Non-Traded Cash Flows

GOR AG Analytics Workshop, Frankfurt (Deutsche Bahn)

Christian Tausch, Alexander Bohnert

AssetMetrix GmbH, Hochschule München

March 6, 2026

Guiding Question

A black and white photograph of a man with short hair and a mustache, wearing clear-rimmed glasses and a dark leather jacket over a chain necklace. He is looking directly at the camera with a neutral expression. In the background, there is a bonsai tree on the left, a person sitting at a table with food in the center, and another person's back to the camera on the right. The overall atmosphere is professional and sophisticated.

TRADE REPUBLIC

Log in

Get the app

QR code

Invest in Private Markets

Source: <https://traderepublic.com/en-de/private-markets>

Guiding Question

Grow your money responsibly. Explore all benefits & risks >

TRADE REPUBLIC 

Log in



Champions



Balance



Return



Go beyond public markets

Invest in **hidden champions** that are not listed on the stock exchange with ELTIFs.

Balance your portfolio. Private markets are not connected to public market volatility.

Your return depends on the selected investments of the fund managers.

Past performance is not a reliable indicator for future results. Investments may result in partial or total loss.

Guiding Question

Grow your money responsibly. Explore all benefits & risks >

TRADE REPUBLIC 

Log in



Champions



Balance



Return



Go beyond public markets

Invest in hidden champions that are not listed on the stock exchange with ELTIFs.

Balance your portfolio. Private markets are not connected to public market volatility.

Your return depends on the selected investments of the fund managers.

Past performance is not a reliable indicator for future results. Investments may result in partial or total loss.

Guiding Question

Are private markets really not connected to public market volatility?

How can we estimate the true public-market exposure of private equity funds?

- Target: recover latent market exposure from non-traded PE cash flow/NAV data.
- Challenge: stale NAVs, asynchronous cash flows, and overlapping fund lives.
- Plan: naive approach first, then semiparametric SDF estimation.

Useful for:

- Risk-adjusted benchmarks for private equity funds.
- Holistic risk management for public and private portfolios.
- More realistic strategic asset allocation decisions.

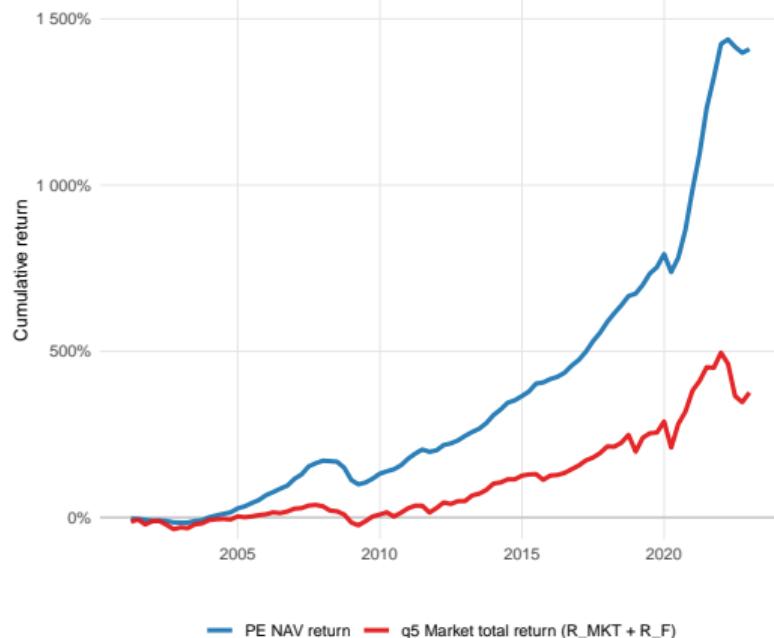
Agenda

- ① Motivation: NAV-return autocorrelation
- ② Estimation framework
- ③ Simulation evidence
- ④ Empirical PE results
- ⑤ Takeaways

Motivation: Public Returns vs Private Cash Flows

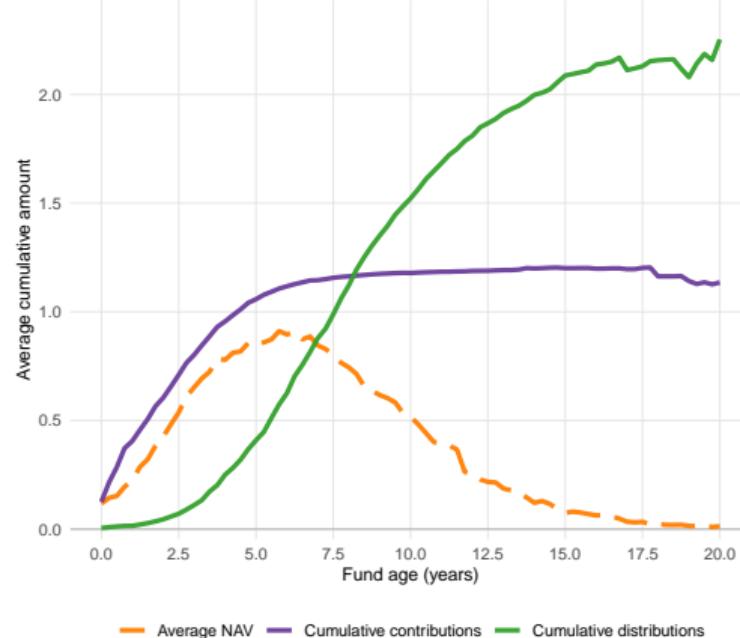
Public Return View

Cumulative returns: PE NAV index vs q5 market total return (includes RF)

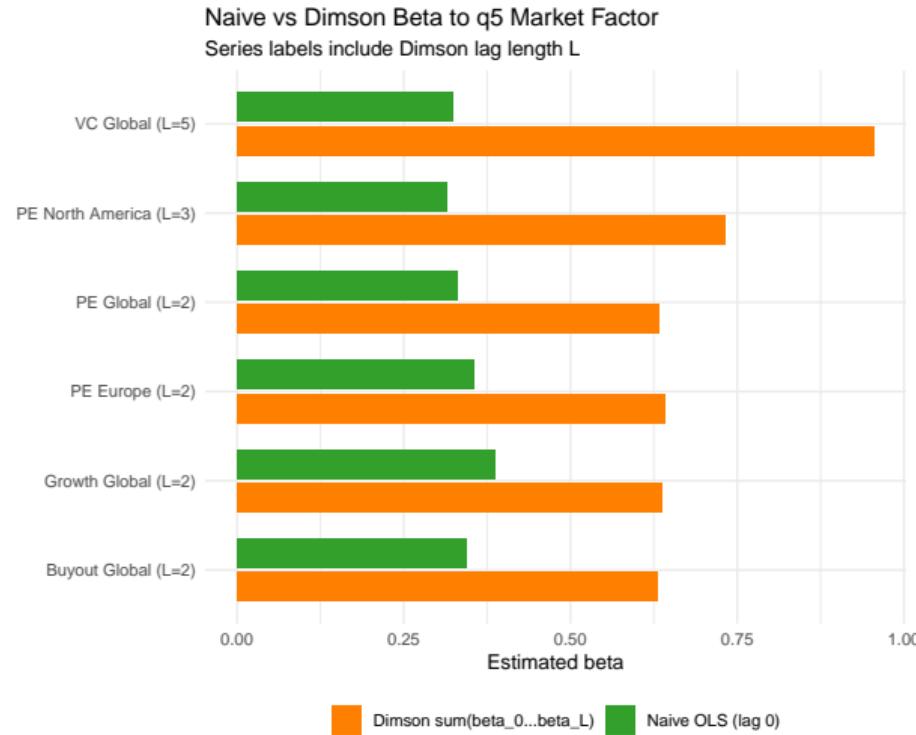
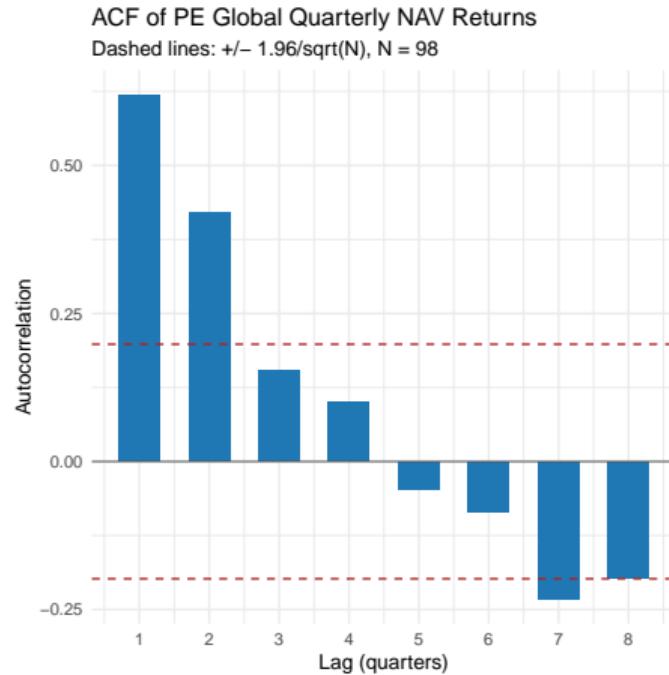


Private Cash-Flow View

Average cumulative cash flows and NAV over fund age for PE (n=37)



Motivation: [Dimson, 1979] Regression from NAV Returns



Data: Preqin quarterly index levels and q5 monthly R_{MKT} aggregated to quarters (overlap through 2022Q4).

Estimator: Starting Point [Driessen et al., 2012]

For each fund i , [Driessen et al., 2012] imposes a zero-NPV pricing condition at inception date $\tau_i^{(0)}$:

$$\epsilon_i^{\text{DLP}}(\theta) = \sum_t \Psi_{\tau_i^{(0)}, t}(\theta) CF_{i,t}$$

Cross-sectional estimator:

$$\hat{\theta}_{\text{DLP}} = \arg \min_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^n (\epsilon_i^{\text{DLP}}(\theta))^2$$

Mapping to our notation

- DLP12 is the special case $\mathcal{T}_i = \{\tau_i^{(0)}\}$, so $|\mathcal{T}_i| = 1$.
- Then $\bar{\epsilon}_i(\theta) = \epsilon_i^{\text{DLP}}(\theta)$ and the LMD objective nests DLP12.
- This paper extends DLP12 by allowing multiple discount dates per fund ($|\mathcal{T}_i| > 1$).

Estimator: Least Mean Distance (LMD)

For fund i and discounting date τ :

$$\epsilon_{\tau,i}(\theta) = \sum_t \Psi_{\tau,t}(\theta) CF_{i,t}$$

Average over selected discounting dates \mathcal{T}_i :

$$\bar{\epsilon}_i(\theta) = \frac{1}{|\mathcal{T}_i|} \sum_{\tau \in \mathcal{T}_i} \epsilon_{\tau,i}(\theta)$$

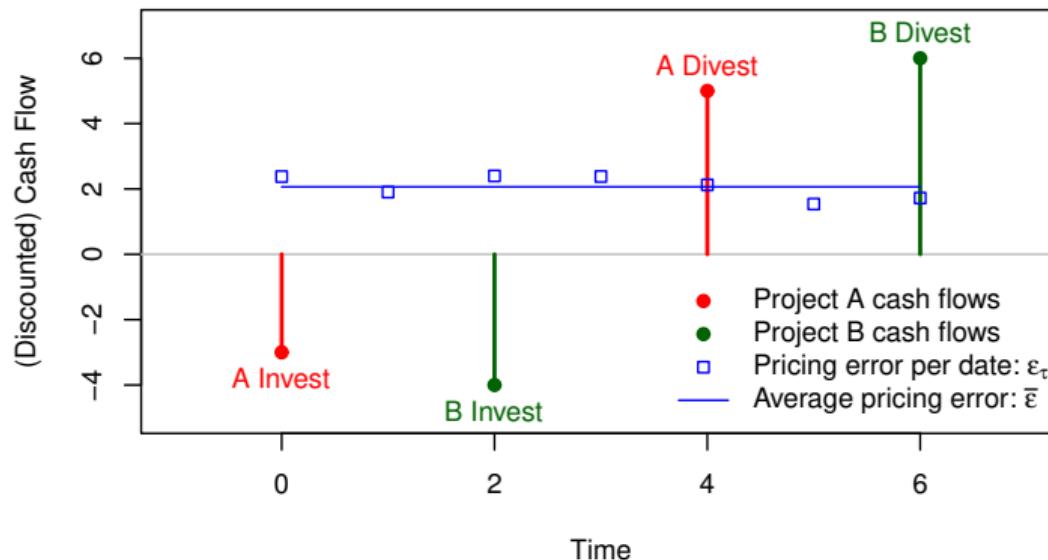
Estimate parameters by nonlinear least-mean-distance:

$$\hat{\theta} = \arg \max_{\theta \in \Theta} \left(-\frac{1}{n} \sum_{i=1}^n L(\bar{\epsilon}_i(\theta)) \right), \quad L(x) = x^2$$

Variables: i fund/portfolio index, t cash-flow time, τ discounting date, $CF_{i,t}$ net cash flow, $\Psi_{\tau,t}(\theta)$ SDF ratio, θ parameter vector, \mathcal{T}_i discount-date set.

Estimator: Net Present Value (NPV) vs Net Future Value (NFV)

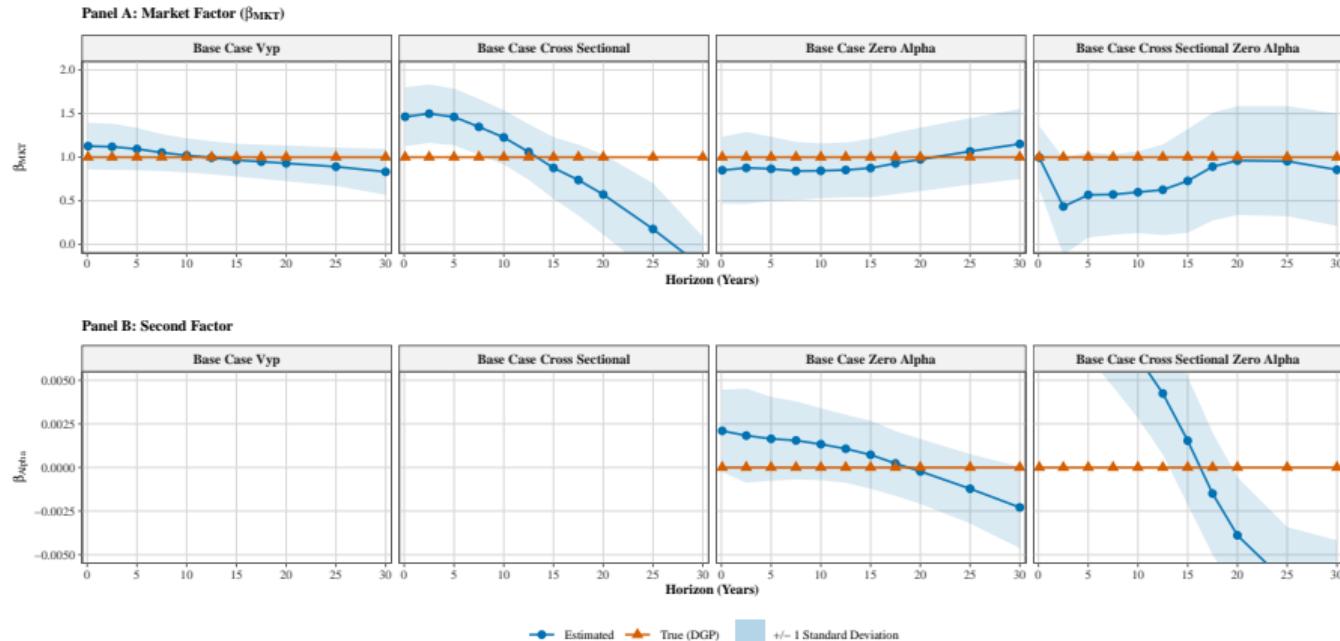
- NPV-only discounting (fund inception) is theoretically unbiased.
- NFV: Adding future-value dates introduces a timing risk **bias** term.
- **But:** Finite-sample performance can improve when averaging across dates.



Simulation: What Is Tested?

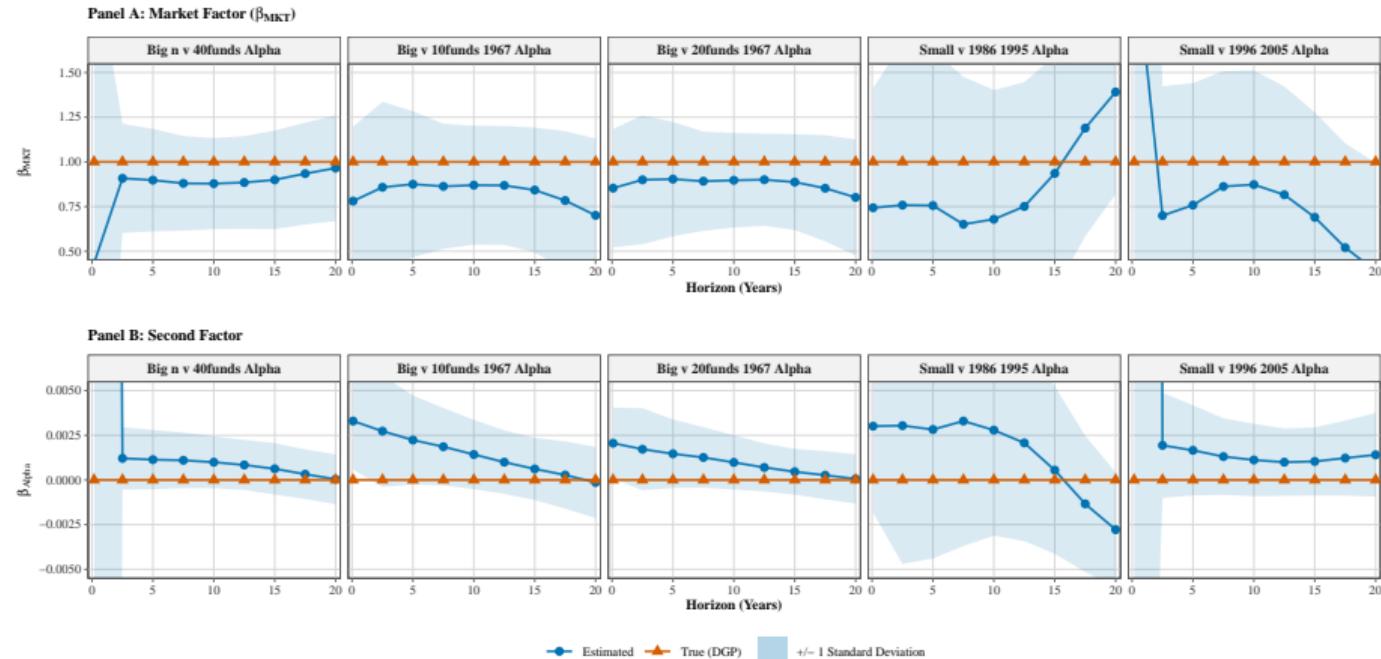
- **Base:** 20 vintages, synthetic fund cash flows generated under known SDF structure.
- **Focus on horizon choice:** size of discounting set \mathcal{T} .
- **Compare units:** individual funds vs vintage-year portfolios (VYP).
- **Compare sample geometry:** more vintages (V) vs larger within-vintage size (n/V).
- **Compare model forms:** simple linear vs exponential affine SDF.

Simulation: 1. Single Funds vs Portfolio Formation



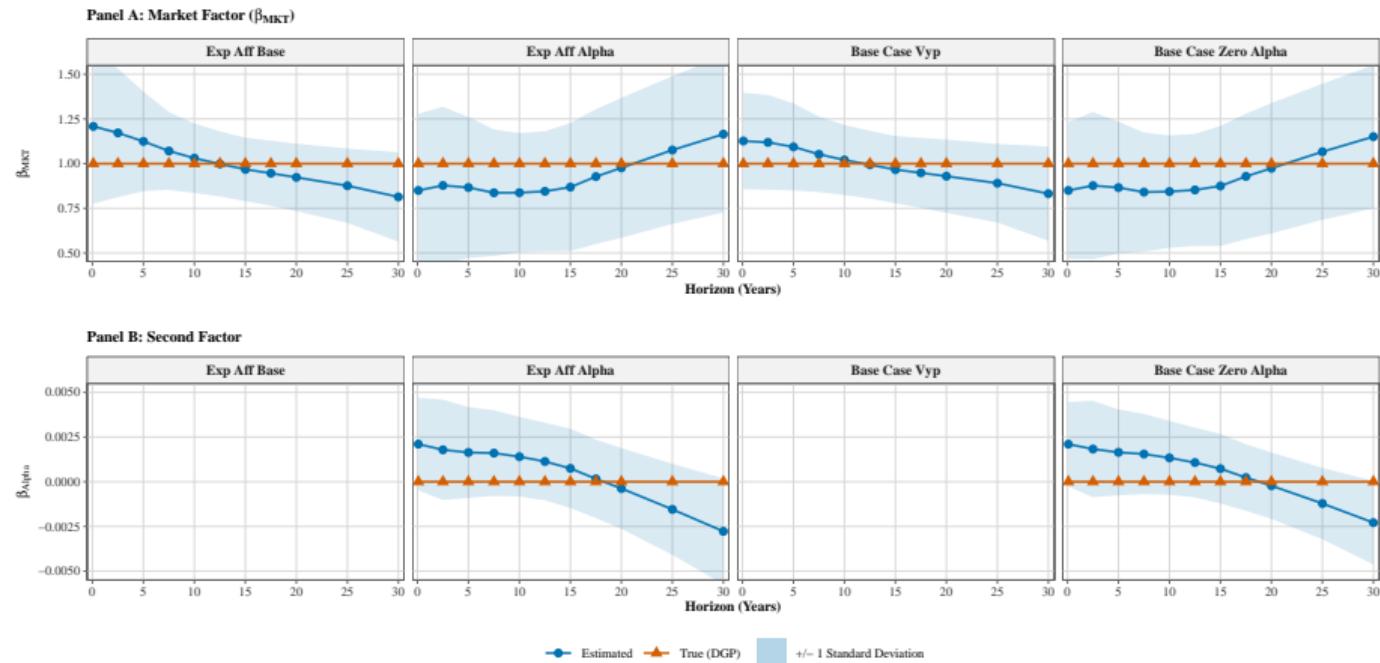
Takeaway: vintage-year portfolios materially reduce bias and variance versus single-fund estimation.

Simulation: 2. More Vintages or More Funds per Vintage?



Takeaway: increasing funds per vintage is more powerful for variance reduction than only extending the time span.

Simulation: 3. Linear vs Exponential Affine SDF



Takeaway: no robust finite-sample superiority of exponential affine specification in this setting.
[Korteweg and Nagel, 2016]

Simulation: Synthesis

- ① Use portfolio aggregation to stabilize estimation.
- ② Prioritize richer cross-sections per vintage when possible (more funds per moment).
- ③ Prefer parsimonious factor structure in current PE data regime.
- ④ Horizon choice (size of T) is a first-order control for finite-sample performance.

Interpretation: practical estimator quality is dominated by finite-sample bias-variance tradeoffs.

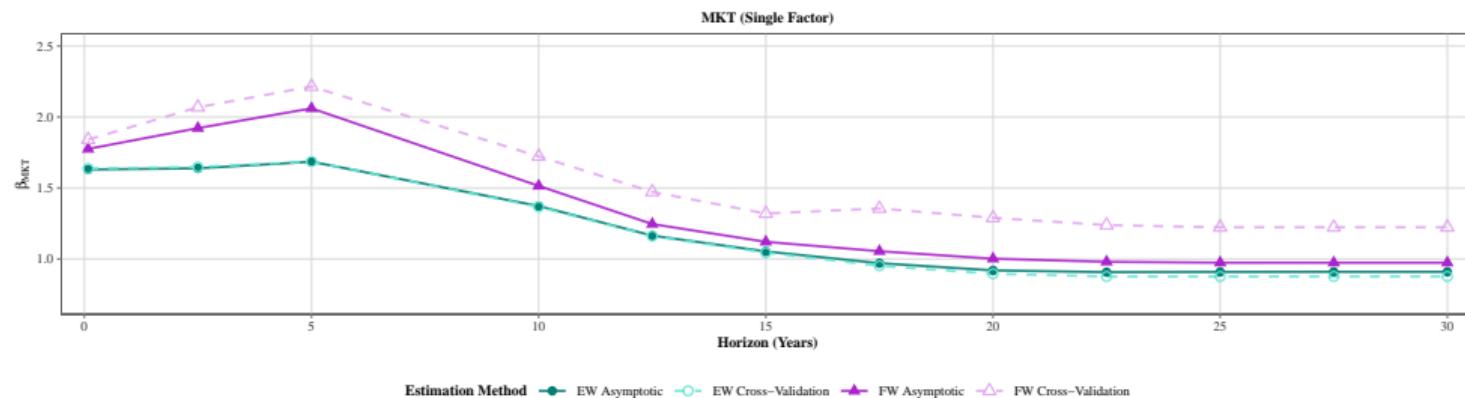
Empirical: Preqin PE Data & q^5 Public Factors

- Data snapshot: 2745 PE funds, vintages 1983–2019.
- Primary unit for estimation: vintage-year portfolios (equal- and value-weighted).
- Factors: q^5 family; focus on MKT and simple two-factor extensions.
- Horizon selected from simulation guidance: 15-year baseline.
- Benchmark for comparison: NAV-based naive/Dimson market-exposure estimates from the motivation section.

Singe-factor models: Dimson beta as lower bound.

Two-factor models: Apply machine-learning methods to form “stronger learner.”

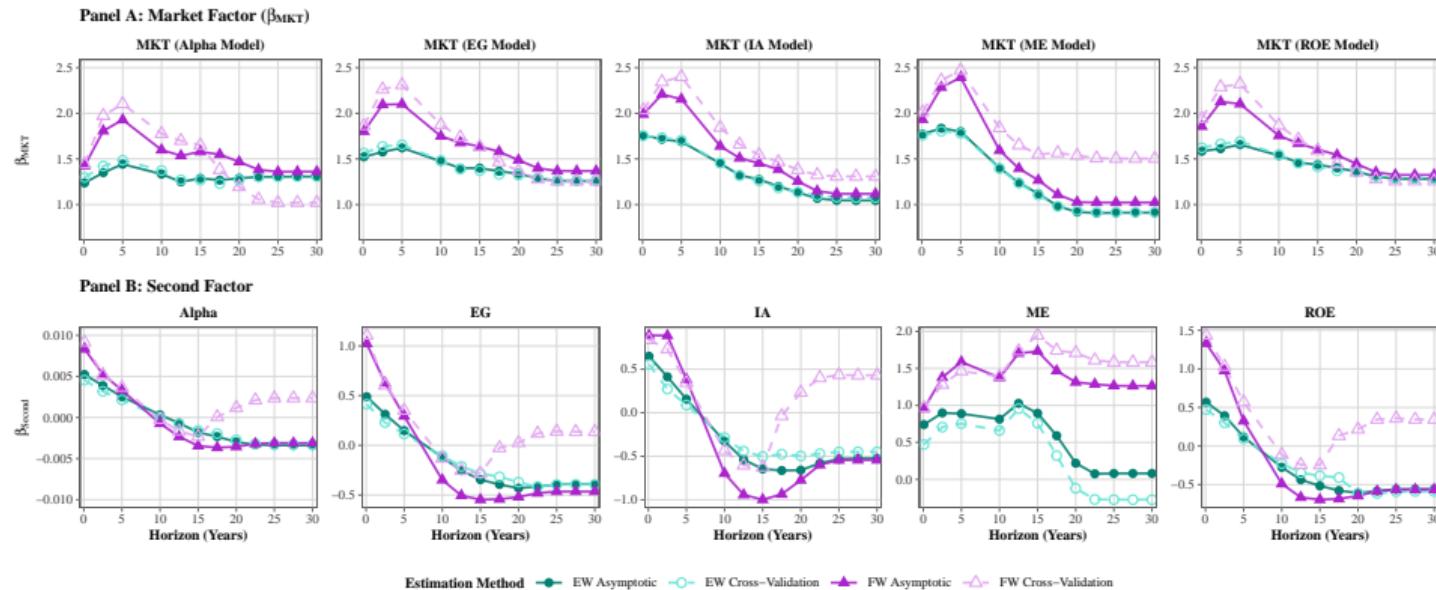
Empirical: 1. Single-Factor MKT Model



Reading:

- Short-horizon MKT betas are high and decline with horizon.
- Betas stabilize near 1 at long horizons.
- CV inference is much more stable than asymptotic t -statistics.
- Relative to naive contemporaneous NAV betas, SDF estimates are materially closer to full market exposure.

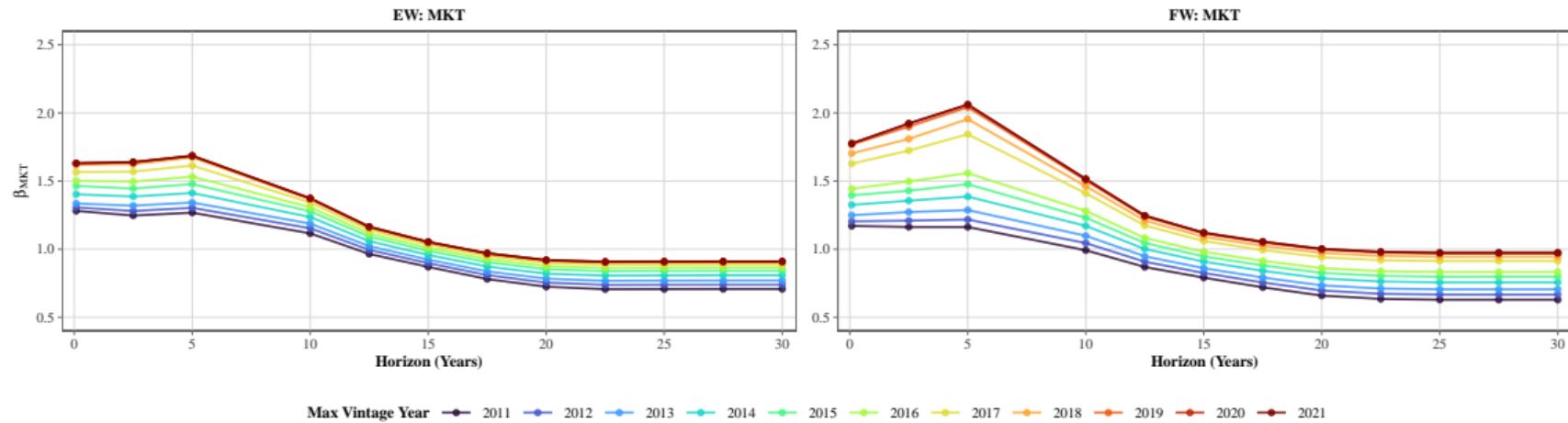
Empirical: 2. Two-Factor Models



Reading:

- Second-factor loadings vary strongly with horizon and weighting scheme.
- Most non-MKT factors do not show robust incremental signal.

Empirical: 3. Vintage Cutoffs for Single-Factor Model

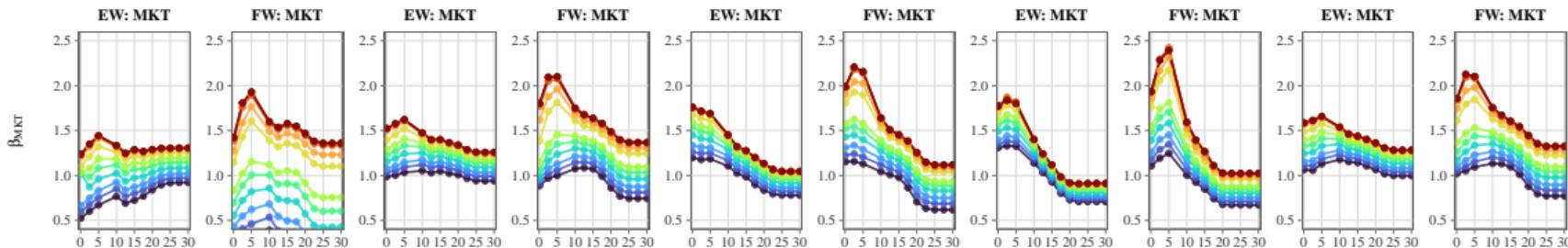


Reading: including newer vintages tends to increase estimated market exposure, with noisier asymptotic uncertainty.

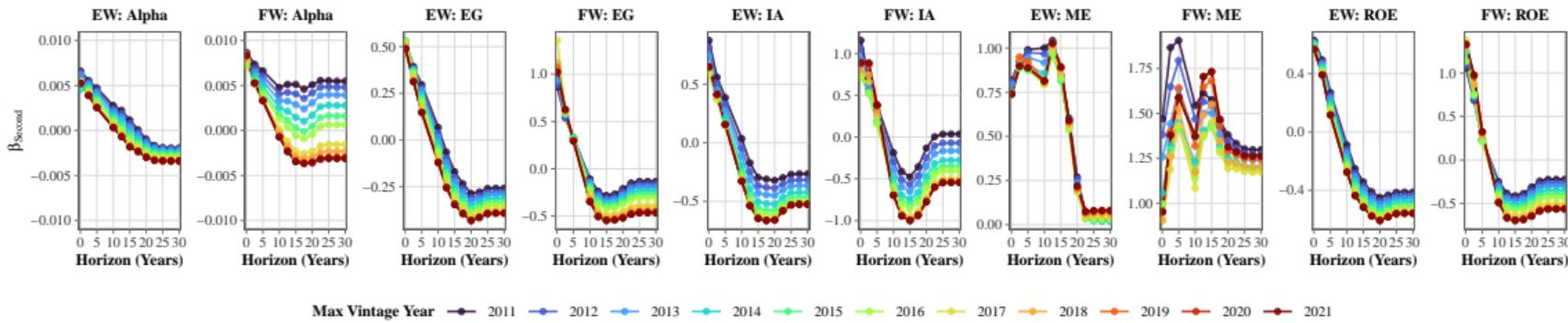
Dimson beta as lower bound!

Empirical: 4. Vintage Cutoffs for Two-Factor Models

Panel A: Market Factor (β_{MKT})



Panel B: Second Factor



Max Vintage Year ● 2011 ● 2012 ● 2013 ● 2014 ● 2015 ● 2016 ● 2017 ● 2018 ● 2019 ● 2020 ● 2021

Conclusion: What This Means for Practice and Research

For practitioners

- Start with parsimonious SDFs (MKT-first), then add complexity cautiously.
- Treat asymptotic significance alone as insufficient in sparse PE samples.
- Use dependence-aware validation (e.g., hv -block CV) as a default diagnostic.

For researchers

- Finite-sample design choices can dominate asymptotic elegance.
- Data architecture (portfolio formation, horizon design) is part of identification.

Conclusion: Main Takeaways

- ① A semiparametric LMD framework can price pooled non-traded cash flows directly.
- ② The central empirical issue is finite-sample stability, not asymptotic theory alone.
- ③ Portfolio aggregation and horizon design are key levers for usable inference.
- ④ Current evidence supports single-factor MKT models as the robust baseline for PE.
- ⑤ Naive contemporaneous NAV betas underestimate exposure; lag-aware methods partially recover it.
- ⑥ Outlook: Multi-factor models can be stabilized by machine-learning techniques.

Questions and Discussion

Literature

-  Dimson, E. (1979).
Risk measurement when shares are subject to infrequent trading.
Journal of Financial Economics, 7(2):197–226.
-  Driessen, J., Lin, T.-C., and Phalippou, L. (2012).
A new method to estimate risk and return of nontraded assets from cash flows: the case of private equity.
Journal of Financial and Quantitative Analysis, 47(3):511–535.
-  Korteweg, A. and Nagel, S. (2016).
Risk-adjusting the returns to venture capital.
Journal of Finance, 71(3):1437–1470.
-  Tausch, C. and Pietz, M. (2024).
Machine learning private equity returns.
The Journal of Finance and Data Science, 10:100141.

Backup: Comparison to DLP12 and KN16

	DLP12	KN16	This paper
Estimator	Cross-sectional NLS	Time-series GMM (public SDF)	Nonlinear LMD
Cash flows priced	PE fund cash flows	Public replicating portfolios	PE fund cash flows
Discount dates	Inception only	Inception only	Flexible via \mathcal{T}_i
Asymptotics	Infill	$V \rightarrow \infty$	Increasing domain
Inference	Bootstrap	SHAC	SHAC + CV focus

Motivation: Why Measuring Risk Is Hard in Private Markets

- PE funds generate **cash flow sequences**, not continuously traded returns.
- Fund lives overlap across vintages, creating dependence beyond standard panel assumptions.
- Fund valuation relies on reported NAVs, which can be stale/smoothed.
- Standard return-based factor models are not directly applicable.

Implication: we need a cash-flow-native SDF estimator with robust dependence-aware inference.

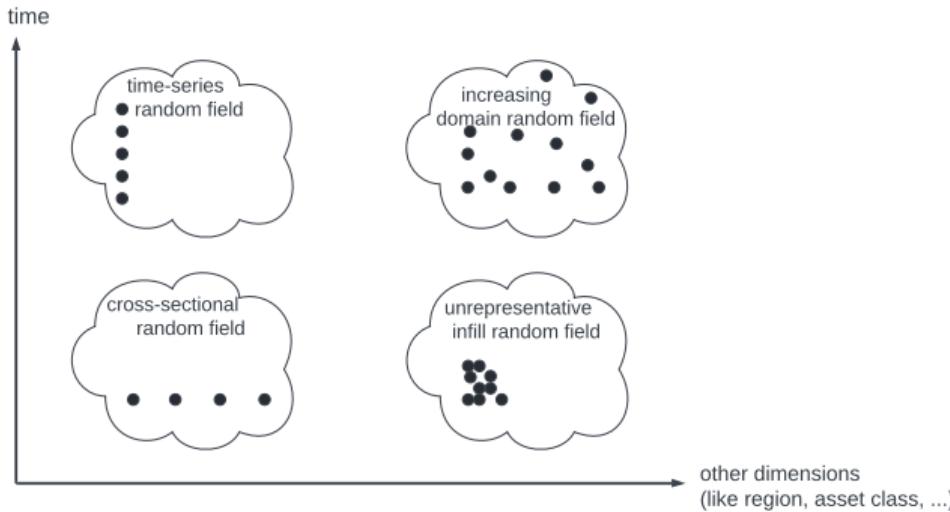
Backup: Significant ACF Lags and Dimson Lag Choice

Series	Significant lags in 1–8	Consecutive from lag 1	Dimson lag L
PE Global	3	2	2
PE North America	5	3	3
PE Europe	3	2	2
Buyout Global	3	2	2
Growth Global	2	2	2
VC Global	5	5	5

Naive benchmark result: contemporaneous beta is low ($\approx 0.32\text{--}0.39$), while Dimson beta increases to $\approx 0.63\text{--}0.96$.

Dependence Structure: Random Field View

- Cross-sectional unit: fund or vintage-year portfolio.
- Dependence driven by economic proximity (here: vintage-year distance).
- Asymptotics: increasing domain ($V \rightarrow \infty$), bounded units per vintage.

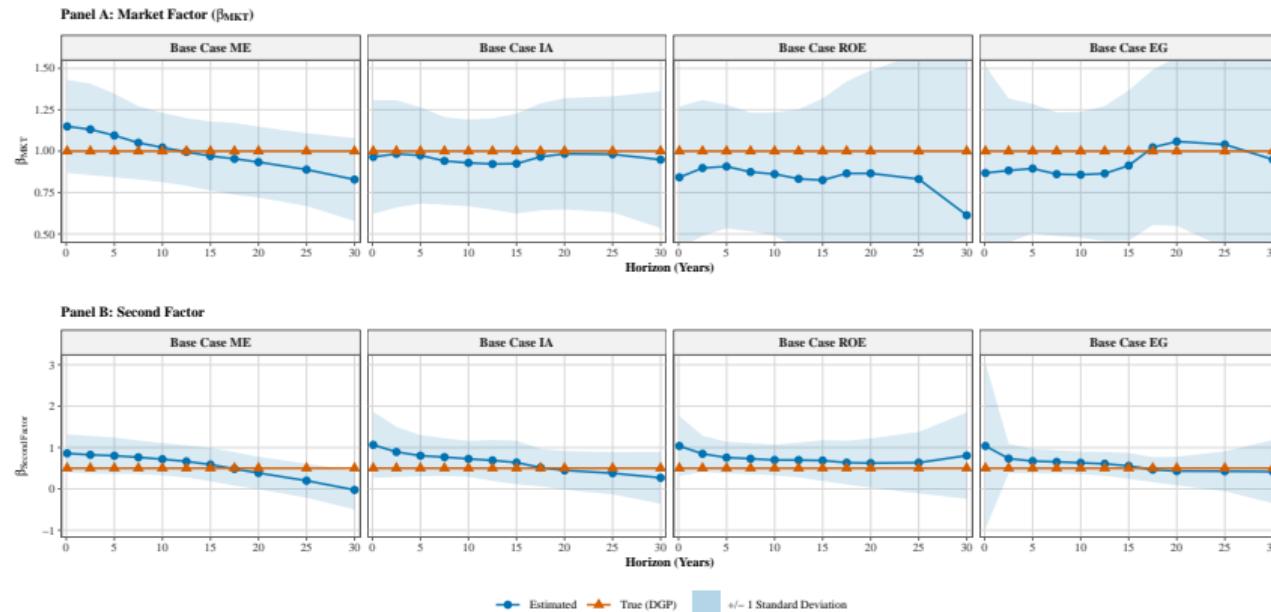


Estimator: Inference Strategy

- Asymptotic covariance: sandwich form $\Sigma = H^{-1}\Lambda H^{-1}$.
- Long-run dependence handled by SHAC (spatial HAC) with vintage-distance kernel.
- Small-sample reliability checked via ***hv-block cross-validation***.

Reason: asymptotic approximations are fragile with only 20–40 vintage portfolios.

Simulation 4: Two-Factor q-Factor Models



Takeaway: two-factor estimates are horizon-sensitive and high-variance; multivariate identification is fragile.

Conclusion: Machine-Learning Ensembles

- Two-factor models are messy.
- Idea: Combine multiple weak learners
- Additionally estimate error term [Tausch and Pietz, 2024]