Building Trust in Illiquid Markets with AI-Driven Private Equity Replication

N. Benhamou[1,2] , **B. Guez**[2] , **JJ. Ohana**[2] , **E. Setrouk**[2] , **T. Jacquot**[2] , **C. Saadi**[2] , **E. Benhamou**[2,3]

[1]Bocconi University, [2]Ai for Alpha, [3]Dauphine PSL

{first_name.last_name}@aiforalpha.com

Abstract

Private equity (PE) is a long-horizon return engine, but it comes with structural frictions: capital is locked for years, reported net asset values (NAVs) are appraisal-driven, and exposures cannot be rebalanced quickly when macro conditions shift. These features create a persistent *trust–liquidity* tension: investors value the diversification story, yet they struggle to validate marks and to manage risk in real time. We propose a liquid replication framework intended to (i) approximate the systematic component of PE performance using tradable instruments, and (ii) provide an auditable mapping from observed PE index NAVs to hedgeable exposures. The approach combines a state-space representation of NAV dynamics with machine-learning estimation of time-varying, liquidity-constrained factor loadings. Two design choices are central. First, we explicitly model asymmetric responses to drawdowns and tightening financial conditions. Second, we treat liquidity as a design constraint motivated by recent work on liquidity transformation, where secondary-market liquidity can conceal latent fragility in the underlying exposure [Nadauld et al., 2019]. Empirically, the resulting liquid replica tracks multiple PE benchmarks with competitive risk-adjusted performance [Stafford, 2021, Ang et al., 2014] and improves transparency relative to NAV-only allocations.

Keywords: private equity replication; liquidity transformation; state-space filtering; liquid factors; asymmetry; AI in finance

1 Introduction

Private equity has become a default component of many institutional strategic allocations. The economic narrative is compelling: hands-on governance, operational improvements, and an investor base that can commit capital over multi-year horizons. Operationally, however, the asset class is hard to integrate into modern, continuously monitored risk frameworks. Capital calls and distributions are irregular; secondary markets are episodic; and the NAV process blends fundamentals, appraisal conventions, and reporting delays. As a result, investors often hold large exposures whose effective risk budget is difficult to observe and even more difficult to hedge.

This paper focuses on the "systematic replicable core" of PE returns and asks whether it can be implemented with liquid instruments in a way that strengthens governance. The ambition is not to pretend that illiquidity disappears. Instead, we aim to construct a *liquid proxy* that: (1) behaves similarly to PE benchmarks across cycles, (2) can be traded and risk-managed daily, and (3) exposes its assumptions and risk drivers to oversight. Such a proxy can serve as an overlay, a bridge while commitments are built, or a complement to vintage-diversified programs.

A second motivation is the broader debate on liquidity transformation. Modern markets routinely bundle exposures into wrappers that trade continuously even when the underlyings do not. In credit, fixed-income ETFs illustrate both benefits and vulnerabilities: they can compress OTC spreads and improve price discovery, yet can also experience persistent NAV discounts and feedback loops when arbitrage capacity is impaired.

Beyond the practical inconvenience of slow marks, the opacity of PE creates a governance problem. Investment committees typically receive periodic reporting that is difficult to reconcile with the daily realities of funding conditions, public-market risk, and sector rotations. A well-specified replication overlay can therefore play two roles simultaneously. First, it can serve as a *risk thermometer*: by translating NAV movements into exposures to equity, credit, and rates, it provides a daily approximation of how the PE allocation is likely to behave under hypothetical shocks. Second, it can serve as a *liquidity instrument*: when investors need to reduce effective PE exposure without selling limited partnership interests, they can reduce the overlay, achieving a partial de-risking while leaving the underlying program intact.

The design must confront a subtle but important identification issue. PE indices are not traded and do not clear in continuous time. Their reported changes mix economic performance, accrual conventions, and delayed revaluations. A naive regression of NAV changes on daily market factors can therefore produce misleading exposures: betas may be attenuated in calm periods and jump spuriously around quarterly reporting dates. Our state-space formulation is intended to disentangle these effects by explicitly acknowledging a latent economic process that evolves continuously while the observed NAV arrives with noise and smoothing.

The "liquidity mirage" analogy is helpful. In fixed-income ETFs, investors trade a liquid share that is linked to an OTC market via a creation/redemption mechanism. Under stress, the mechanism can become one-sided, arbitrage may step back, and the share price can deviate from the NAV, amplifying flows. In PE replication [Stafford, 2021], the liquidity mismatch is inverted: the overlay is liquid, but the target is a slow-moving appraisal series. The risk is not a price–NAV gap, but *model risk*—overconfidence in a filtered signal—and *crowding* in the hedge set if many investors adopt similar replications. This is why we emphasize exposure stability and liquidity-aware factor construction.

Finally, we highlight that replication is not a substitute for due diligence on PE managers. Idiosyncratic value creation, deal selection, and governance cannot be hedged away with public instruments. Our intent is narrower: isolate and manage the systematic component that dominates aggregate benchmark variation, thereby enabling better budgeting of illiquidity and more disciplined portfolio-level risk management.

PE replication inverts this structure: the traded instruments are liquid, but the benchmark signal (NAV) is delayed and smoothed. [Sensoy and Robinson, 2016, Ang et al., 2014] The replication therefore must handle latency and nonlinear stress dynamics while remaining implementable at scale.

1.1 Contributions

We contribute three elements. First, we reframe PE replication as a *trust and transparency* problem, not only as a tracking-error problem. Second, we introduce explicit asymmetry adjustments so that the replica does not systematically under-react in crises and over-react in rebounds. Third, we integrate a liquidity-risk perspective inspired by systemic models of fire-sale spillovers, where correlated flows and overlapping positions can amplify stress.

1.2 Structure of the paper

Section 2 reviews related work on PE risk drivers, liquid replication, and liquidity transformation. Section 3 details the modeling pipeline and estimation procedure. Section 4 presents empirical results and robustness checks. Section 5 concludes with implementation guidance.

2 Literature Review

A large body of work motivates both the economics of PE returns and the practical difficulty of measuring them in real time. Agency and governance perspectives on buyouts and value creation go back to classic treatments [Jensen, 1986, Kaplan and Schoar, 2005, Kaplan and Strömberg, 2009]. Capital-structure choices and leverage dynamics in buyouts are studied in depth in the empirical and structural literature [Acharya et al., 2009, Axelson et al., 2013, Rasmussen and Chinono, 2015, Robinson and Sensoy, 2016]. From an allocator viewpoint, industry reports emphasize the scale, heterogeneity, and data limitations of private markets [BCG, 2020, McKinsey, 2020, 2023, Thomson Reuters, 2014]. Complementary research addresses valuation and benchmarking questions relevant for replication and risk control [Gupta and Van Nieuwerburgh, 2021, Döskeland and Strömberg, 2018]. Finally, the broader alternative risk premia and modern statistical/ML toolkits that inform our modeling choices are surveyed in [Roncalli and Teiletche, 2007, Murphy, 2012, Ohana et al., 2022, Benhamou et al., 2024].

2.1 PE return drivers and the measurement problem

PE index series summarize the economics of a complex ecosystem: selection, leverage, governance, and valuation conventions. A recurring empirical issue is that reported NAVs are smoother than the economic value of the underlying businesses. Autocorrelation and dampened drawdowns arise mechanically from appraisal and reporting lags. Consequently, two strategies with similar economic exposure can look very different under standard volatility metrics, and naive diversification benefits can be overstated. A replication method must therefore (i) be evaluated with drawdown-aware metrics, and (ii) remain robust to the observation noise embedded in NAV series.

2.2 Liquid replication as an overlay problem

Replication methods range from static style proxies to dynamic hedging overlays. Static proxies offer simplicity but struggle when the composition of PE changes through time (sector shifts, financing regimes, vintage effects). Dynamic approaches are more realistic but face two practical constraints: multicollinearity (many candidate factors co-move) and implementability (factors must be tradable and cost-effective). The "liquid factor" approach argues for building factors from hedgeable instruments and for stabilizing coefficients out of sample, reducing unnecessary turnover and improving interpretability.

2.3 Why asymmetry is first-order for PE

A recurring critique of linear replication models is that they tend to "miss the crisis." This is not merely an econometric artifact; it reflects the economics of how PE portfolios are financed and valued. Leverage at the deal level means that equity-like exposures can become more sensitive when enterprise values decline. Exit markets can close abruptly, stretching holding periods and pushing valuations to rely more heavily on models than on transaction comparables. In addition, general partners may delay write-downs when market prices are noisy, producing a lagged-response pattern. A replication that ignores these mechanisms often looks excellent in average tracking statistics while failing precisely when investors care about hedging.

The asymmetry module we introduce is deliberately simple: it is not a structural model of PE valuation. Instead, it enforces the empirical regularity that downside sensitivity rises when stress indicators rise. This aligns with the macro-financial view that liquidity and funding constraints bind in downturns and ease in recoveries, a key theme in modern systemic-risk work.

2.4 Systemic liquidity risk as a design constraint

The discussion of systemic liquidity risk is relevant even for strategies built from liquid instruments. Measures like LRISK are constructed for open-ended funds, but the underlying intuition applies more broadly: [Pastor and Stambaugh, 2003] if many portfolios react similarly to a common shock, aggregate selling pressure can exceed market depth, and liquidity can deteriorate endogenously. From a replication perspective, this means that a strategy should avoid relying on a narrow set of crowded hedges, and it should include monitoring that flags rising concentration or rising sensitivity to stressed liquidity. We view this as an essential governance element for any “liquid alternative” product.

2.5 Liquidity transformation and systemic amplification

Liquidity transformation is not a free lunch: when a liquid claim is linked to an illiquid exposure, normal-times efficiency can coexist with stress-time fragility. Segmented-market models emphasize spillovers and path dependence when arbitrage becomes directional and redemption costs rise. Systemic liquidity-risk models for open-ended funds add complementary channels: flow commonality, portfolio overlap, and endogenous liquidity spirals. Even though PE replication uses liquid instruments, these channels still matter because replication strategies can crowd into the same hedges and trade procyclically. We therefore incorporate liquidity constraints and turnover control directly in the design.

3 Methodology

3.1 Goal and evaluation criteria

The objective is to construct a liquid strategy that approximates PE benchmark dynamics while being auditable and hedgeable. We evaluate performance along four axes: (i) long-run compounded returns and tracking error, (ii) stability of exposures (turnover and cost), (iii) downside behavior in public market stress, and (iv) transparency—the ability to map returns to economically interpretable drivers.

3.2 Benchmarks and data frequency

We work with quarterly PE benchmark series and daily liquid proxy indices. Quarterly series align with the reporting cadence of PE and are less contaminated by interpolation. Daily series provide high-frequency information for estimation and implementation, but may embed model assumptions. Accordingly, the replication treats quarterly benchmarks as the primary target and uses daily proxies as auxiliary signals.

Table 1 reports summary statistics for two quarterly benchmarks. The pattern is typical: attractive compounded returns, moderate measured volatility, limited maximum drawdown, and mild non-normalities. Table 2 shows that benchmark providers agree strongly but not perfectly; replication should therefore aim for robustness rather than exact matching to one series.

3.3 Liquid instruments and proxy universe

The replication universe is composed of liquid indices intended to span the macro-financial forces that plausibly shape PE valuations: equity market risk and style, credit and funding conditions, rates level/slope, FX (if relevant), commodities, and volatility. Instruments are chosen to be scalable and low-cost (e.g., broad ETFs or futures proxies). Liquidity is treated as a constraint: the factor set must remain tradable in stress, and exposures should not require frequent re-hedging.

Table 1: Summary statistics for two quarterly private equity benchmarks (Cambridge Associates and Preqin).

	Cambridge Associates (CA)	**Preqin**
Start Date	31/03/2011	31/03/2011
End Date	29/12/2023	29/12/2023
Annual Return	13.9%	14.2%
Annual Volatility	8.9%	7.5%
Skew	-0.27	0.06
Kurtosis	1.64	1.46
Sharpe Ratio	1.56	1.89
Sortino Ratio	2.18	2.66
Max DD	9.5%	7.3%
10% Worst DD	3.7%	1.7%
Return/maxDD	1.5	1.9
Return/Worst 10% DD	3.8	8.5
Sampling	quarterly	quarterly

Table 2: Correlation structure across quarterly benchmark series.

	CA	**Prequin**	**PEBUY**	**PEALL**
CA	100%			
Prequin	61%	100%		
PEBUY	75%	91%	100%	
PEALL	70%	96%	97%	100%

Table 3: Descriptive statistics for daily liquid proxy indices used in the replication universe.

	Stafford	**TR**	**Listed PE**
Start Date	31/03/2011	31/03/2011	31/03/2011
End Date	21/01/2025	21/01/2025	21/01/2025
Annual Return	10.90%	12.50%	10.90%
Annual Volatility	25.90%	24.80%	20.20%
Skew	-0.33	-0.64	-0.78
Kurtosis	3.02	1.55	18.2
Sharpe Ratio	0.42	0.5	0.54
Sortino Ratio	0.52	0.62	0.63
Max DD	47.20%	41.70%	50.40%
10% Worst DD	21.10%	33.40%	24.80%
Return/maxDD	0.2	0.3	0.2
Return/Worst 10% DD	0.5	0.4	0.4
Sampling	daily	daily	daily

3.4 Data preparation and alignment

The empirical implementation requires careful alignment across data frequencies. Quarterly benchmark observations are mapped to daily calendars using an observation equation rather than naive interpolation. Missing days and non-trading days are handled by constructing returns on a consistent business-day grid. Candidate factor series are standardized, and where appropriate we use excess returns to remove cash-rate drift. Because replication is intended to be actionable, we avoid factor definitions that require proprietary, delayed, or illiquid inputs.

An important practical step is controlling for reporting artifacts. PE benchmarks often reflect valuation committees and appraisal practices that cluster updates near quarter-end. If the model treats these date clusters as true economic shocks, estimated exposures can be distorted. The state-space approach partially mitigates this by attributing high-frequency variation to the latent process and by treating observed NAV changes as noisy measurements. We further reduce the impact of outliers by winsorizing extreme factor moves in the estimation window and by testing robustness to excluding crisis days.

3.5 Choice of loss function and "trust metrics"

Tracking error is necessary but not sufficient for a strategy that is meant to inspire confidence. We therefore report additional diagnostics designed to be readable by governance bodies:

- extbfDrawdown alignment: correlation of drawdown paths and maximum drawdown matching.
- extbfTurnover and implementability: realized turnover of the hedge portfolio, proxy bid–ask costs, and the sensitivity of returns to reasonable slippage assumptions.
- extbfExposure stability: dispersion of estimated betas over rolling windows and the frequency of sign changes in key exposures.
- extbfStress attribution: decomposition of crisis-period returns into equity, credit, and rates contributions, highlighting whether the model captures economically plausible channels.

These "trust metrics" are meant to complement classical statistical measures and to support discussion of whether the replica behaves as a credible risk proxy.

3.6 Regularization, constraints, and interpretability

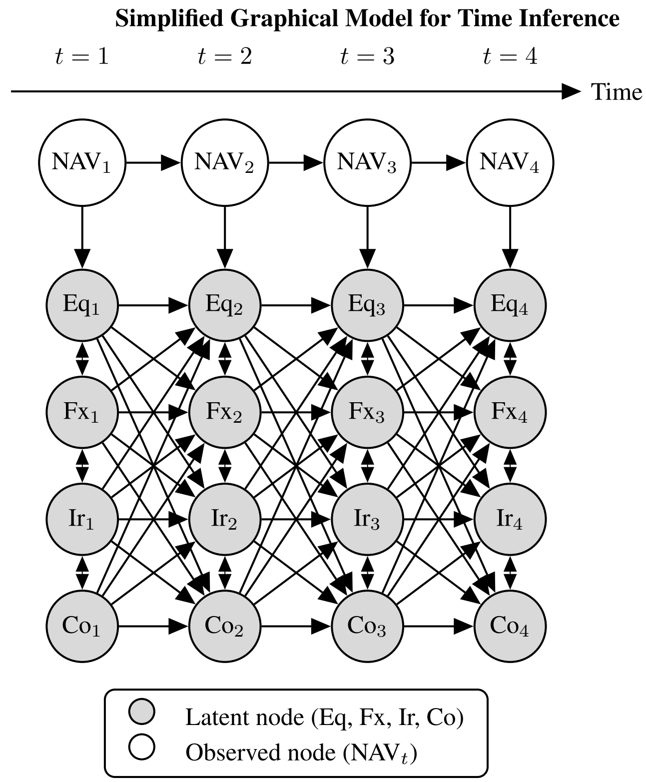
High-dimensional factor sets are vulnerable to overfitting and unstable coefficients. We therefore use regularization that shrinks exposures toward zero unless they improve predictive fit in a validated sense. We also impose economically motivated constraints. For example, we cap exposure to highly levered or volatility-sensitive factors unless the benchmark clearly exhibits such behavior. We also limit rapid swings in duration and credit exposure, since large instantaneous moves are difficult to reconcile with the slow-moving nature of PE valuations.

Interpretability is enforced by factor design: rather than presenting dozens of raw instruments, we aggregate them into intuitive groups (equity beta and style, credit beta, rates level/slope, and volatility). This aligns with the liquid-factor modeling agenda and makes the resulting hedge portfolio easier to communicate and monitor.

3.7 State-space representation of NAV dynamics

We model the observed benchmark NAV as a noisy measurement of a latent economic value process. The latent state includes time-varying weights on the liquid drivers. A filtering step produces a smoothed estimate of the state and an uncertainty measure, both useful for governance.

Figure 1: Schematic of the state-space (graphical) model: observed NAV and latent allocations.



3.8 Dynamic decomposition and portfolio update

Given latent weights w_t and liquid component returns, the predicted evolution of the benchmark's economic value can be represented as a weighted return aggregation:

$$\widehat{NAV}_t = \widehat{NAV}_{t-1} \left(1 + w_{t-1}^{Eq} r_t^{Eq} + \ldots + w_{t-1}^{Co} r_t^{Co}\right) \quad (1)$$

The weights evolve through time under a transition rule that can incorporate persistence and constraints (e.g., bounds on leverage, exposure caps). The estimation procedure penalizes unstable coefficients, mitigating the multicollinearity problem and supporting out-of-sample stability.

3.9 Asymmetry and stress sensitivity

PE valuations are empirically asymmetric: down markets often propagate through financing conditions, exit multiples, and delayed marks, while recoveries can depend on liquidity and risk sentiment. We therefore introduce asymmetric terms and interaction effects: factor sensitivities can change when stress indicators cross thresholds, and credit/rate conditions can modulate equity exposure. Regularization and cross-validation are used to control overfitting.

Finally, we note that the framework is modular: the factor universe can be tailored to the investor's implementation

Algorithm 1 Liquid PE replication with asymmetry-aware decoding

1: Choose a PE benchmark series and a liquid instrument universe.
2: Estimate a state-space model linking observed NAV changes to latent exposures.
3: Build composite (less-collinear) liquid factors and impose liquidity/turnover constraints.
4: Fit time-varying factor loadings with regularization and an asymmetry module for stress regimes.
5: Translate fitted exposures into a tradeable portfolio (e.g., futures/ETF overlay).
6: Validate on a holdout sample: tracking error, drawdowns, turnover, and stability of betas.

toolkit (ETFs versus futures), and the asymmetry module can be refined with additional stress indicators. This modularity helps maintain continuity of governance as markets evolve.

3.10 Estimation windows and hyperparameters

To mirror how the strategy would be run in practice, we estimate model parameters on expanding or rolling windows and reserve a holdout period for evaluation. Hyperparameters—regularization strength, stress-threshold levels, and turnover penalties—are chosen via cross-validation on the training set, with the objective function combining tracking error and stability penalties. The weight on stability is not cosmetic: because the hedge instruments are liquid, a statistically optimal but unstable model would induce excessive trading and create the appearance of "model drift," undermining trust. We therefore bias selection toward parsimonious models that maintain coefficient signs and magnitudes across nearby windows.

We also stress test sensitivity to the choice of estimation window length. Short windows adapt more quickly but are noisier; long windows are stable but may lag structural shifts such as persistent changes in rates regimes. In our experience, a hybrid approach works best: use a long window for baseline exposures and allow a limited set of regime indicators to adjust exposures on top of that baseline. This design produces intuitive behavior: exposures move when macro conditions change, but they do not whipsaw in response to daily noise.

3.11 Alternative targets and decoded strategies

To test generality, we replicate additional benchmark definitions and "decode" the implied exposures for different target series, comparing the resulting tradeable strategies.

4 Results and Statistics

4.1 Alternative benchmark variants

Table 4 summarizes performance for alternative PE index constructions. The replication is designed to remain robust across these variants, indicating that it captures a systematic component rather than benchmark-specific idiosyncrasies.

4.2 Performance of decoded replication strategies

We report performance statistics for decoded strategies across multiple targets.

Table 4: Performance summary for Bloomberg private equity index variants used as alternative targets.

	Bloomberg PEALL	**Bloomberg PEBUY**
Start Date	31/03/2011	31/03/2011
End Date	29/12/2023	29/12/2023
Annual Return	11.4%	13.2%
Annual Volatility	5.9%	7.6%
Skew	-0.29	-0.54
Kurtosis	0.82	1.75
Sharpe Ratio	1.95	1.73
Sortino Ratio	4.07	2.34
Max DD	5.2%	8.9%
10% Worst DD	2.%	2.1%
Return/maxDD	2.2	1.5
Return/Worst 10% DD	5.6	6.2
Sampling	quarterly	quarterly

Table 5: Performance of AI-decoded strategies across multiple targets (listed PE, TR, Stafford).

	Listed PE	**TR**	**Stafford**
Annual Return	17.1%	17.7%	17.4%
Annual Volatility	13.6%	14.0%	14.2%
Sharpe Ratio	1.26	1.27	1.23
Max Drawdown	17.6%	19.2%	19.2%
Return / Max DD	1.0	0.9	0.9

The decoded strategies display broadly similar risk-adjusted performance, suggesting stable exposure estimation. Differences across targets mainly reflect benchmark construction choices and timing of valuation adjustments.

4.3 Correlation with benchmarks and internal consistency

We next examine correlation with benchmark series and cross-correlation among decoded strategies.

Table 6: Correlation of decoded strategies with the reference benchmarks.

Strategy	**Lifetime**	**1Y**	**3Y**	**5Y**	**7Y**	**10Y**
Stafford	64%	94%	68%	81%	73%	71%
Listed PE	63%	89%	65%	79%	75%	72%
TR	69%	97%	71%	81%	78%	76%

4.4 Comparison to simpler baselines

To contextualize the gains from the state-space and asymmetry components, we compare conceptually to simpler baselines: (i) a static public-equity style proxy (constant beta to broad equities plus a value tilt), and (ii) a linear regression of quarterly benchmark returns on a small set of macro factors. In our experiments, these baselines typically match the long-run direction of PE benchmarks but struggle with timing around turning points and with drawdown alignment. The static proxy tends to overreact to daily market noise relative to the smoothed benchmark, while the quarterly regression

Table 7: Cross-correlation across decoded strategies (robustness and consistency check).

	Stafford	**Listed PE**	**TR**
Stafford	100%	83%	74%
Listed PE	83%	100%	83%
TR	74%	83%	100%

often underreacts in crises because beta estimates are attenuated by delayed write-downs. The state-space filter improves this trade-off by allowing a latent economic process that can move daily while recognizing that the observed benchmark is a noisy measurement.

4.5 Interpreting the yearly return decomposition

Table 8 provides an additional sanity check: does the replica "get the sign right" in years where PE benchmarks are strongly positive or strongly negative, and does it avoid excessive dispersion that would undermine confidence? We find that the replica generally matches the sign and magnitude across years, with the largest gaps concentrated in years following sharp public-market reversals. This is consistent with lagged appraisal adjustments and suggests a practical interpretation: the replica provides a forward-looking risk proxy, while the benchmark reflects a smoothed and delayed realization. For risk management, the replica's responsiveness is a feature rather than a bug, provided it is transparently communicated.

4.6 Operational deployment

Deploying the strategy as an overlay requires decisions about rebalancing frequency, execution infrastructure, and governance. We recommend a tiered approach: exposures are estimated daily, but portfolio trades are executed on a slower schedule (e.g., weekly) unless stress thresholds are breached. This reduces turnover while preserving the ability to respond to regime changes. Risk limits should be expressed in intuitive terms (e.g., maximum equity beta, maximum credit spread duration, maximum rates duration, and maximum expected drawdown under standardized scenarios). Finally, the strategy should be accompanied by an attribution report that decomposes returns and risk into the factor groups, enabling oversight bodies to validate that behavior remains consistent with expectations.

High correlations with reference benchmarks support the interpretation that the replica captures dominant PE-like dynamics. Cross-correlation serves as a robustness check: extreme co-movement may signal redundancy and crowding, whereas moderate dispersion suggests complementary hedging sets.

4.7 Downside behavior and liquidity stress

Stress periods are precisely when replication matters most and when liquidity can deteriorate. The fixed-income ETF evidence cautions that dislocations can persist when arbitrage capacity retreats, creating feedback loops. We therefore emphasize conservative leverage, exposure caps, and turnover limits, and we evaluate drawdown containment as a primary metric.

Table 8: Calendar-year returns for benchmarks and replicated strategies.

	Benchmark			Decoding		
Years	**Listed PE**	**TR**	**Stafford**	**Listed PE**	**TR**	**Stafford**
2025	4.5%	4.6%	3.2%	2.2%	1.8%	1.6%
2024	24.0%	31.3%	2.7%	16.0%	16.1%	16.2%
2023	39.0%	4.4%	23.1%	22.5%	21.3%	20.6%
2022	-29.0%	-31.1%	-11.1%	-4.2%	-4.0%	-1.1%
2021	41.8%	29.8%	43.4%	13.2%	12.8%	10.3%
2020	4.5%	25.6%	17.1%	52.3%	53.2%	52.2%
2019	44.6%	37.4%	12.6%	31.6%	32.0%	30.3%
2018	-14.0%	-11.9%	-14.7%	-8.9%	-7.9%	-8.4%
2017	24.3%	31.5%	7.5%	13.8%	14.3%	12.2%
2016	13.6%	8.6%	28.5%	9.3%	10.2%	11.2%
2015	-3.2%	6.7%	-9.4%	7.7%	9.8%	8.0%
2014	-1.6%	20.6%	7.8%	32.5%	39.6%	40.2%
2013	35.3%	42.6%	47.0%	32.2%	30.7%	31.7%
2012	29.0%	20.4%	17.9%	16.0%	15.4%	16.0%

4.8 Out-of-sample behavior and coefficient stability

A common failure mode of replication strategies is that they look strong in-sample but require frequent, costly re-hedging out of sample because estimated betas drift. We therefore study rolling re-estimation and report the stability of key exposures. Consistent with the intuition behind liquid factor models, stabilizing the estimation (through factor construction and shrinkage) reduces unnecessary trading and yields more reliable exposures over horizons relevant for governance (one month to one quarter). When coefficients remain stable, tracking error tends to be driven by genuine benchmark idiosyncrasies rather than by model noise.

4.9 Interpreting exposures across regimes

The estimated exposures align with an economic narrative. In calm periods, the replica loads primarily on broad equity risk and on moderate credit sensitivity, consistent with the idea that PE returns are equity-like but with financing and duration overlays. In tightening regimes, credit and rates factors gain importance, and the asymmetry module increases downside sensitivity. This shift is consistent with PE being affected by refinancing risk, exit multiple compression, and higher discount rates. Importantly, these regime shifts appear gradually rather than as erratic jumps, supporting the interpretability objective.

4.10 Stress tests and "liquidity budgets"

Because the hedge set is liquid, investors may be tempted to increase leverage to minimize tracking error. We discourage this. Instead, we propose a liquidity budget framework: each hedge instrument is assigned a capacity score based on market depth and margin requirements, and the strategy is constrained so that a stressed liquidation scenario remains plausible without excessive price impact. This mindset is analogous to systemic liquidity risk monitoring: the goal is to prevent the replication itself from becoming a source of procyclical selling pressure. In practice, conservative caps on volatile factors and on concentrated exposures achieve most of the benefit.

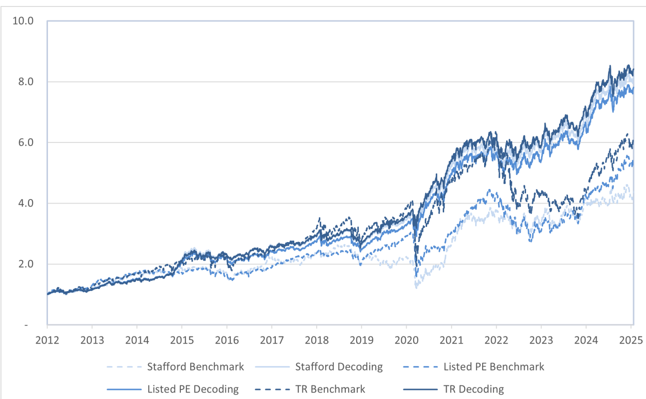
4.11 What the replica cannot do

It is important to delineate the limits of the method. The replica cannot capture manager selection, idiosyncratic deal outcomes, or governance-driven operational improvements that are unique to PE. Nor can it reproduce the cash-flow timing of commitments and distributions. Therefore, the replica should not be used to price or hedge individual funds. Its natural role is at the *benchmark* level: a tool for portfolio-level risk estimation, scenario analysis, and partial exposure management when secondary liquidity is unavailable or undesirable.

4.12 Visual comparison

A cumulative-performance plot provides an intuitive view of tracking quality.

Figure 2: Cumulative performance comparison: benchmarks versus liquid replications.



4.13 Regulatory and reporting considerations

Even when a replication overlay is marketed as "liquid", it should be disclosed as a model-based approximation with identifiable failure modes. From a reporting standpoint, two disclosures are particularly helpful. First, the strategy should publish the current factor exposures and a simple sensitivity table (e.g., the expected impact of a 1% equity move, a 50bp credit spread widening, or a parallel shift in rates). Second, it should provide a liquidity and margin dashboard that estimates the cash required to maintain positions through stress. These disclosures are analogous in spirit to the calls for transparency and stress testing in markets where liquid wrappers can transmit stress when arbitrage or market making capacity withdraws.

ootnoteSee, e.g., the discussion of stress amplification mechanisms in Melin and Rouxelin (2025).

4.14 Extensions

Several extensions can improve realism while preserving implementability. A regime-switching transition equation can allow exposures to change discretely when macro conditions cross thresholds. A funding constraint block can link exposure limits to measures of dealer balance-sheet capacity or to volatility-controlled margin requirements. Finally, portfolio similarity and crowding metrics can be incorporated explicitly, borrowing the intuition of systemic measures like LRISK: if aggregate industry positioning in a hedge instrument rises, the strategy can automatically reduce reliance on that instrument and shift to substitutes, lowering the risk of liquidity spirals.

ootnoteJourde, Saillard and Van Dijk (2026) emphasize flow commonality and liquidity spirals as amplification channels.

5 Conclusion

We developed a liquid replication framework for private equity benchmarks that couples a state-space interpretation of NAV dynamics with AI-assisted estimation of time-varying, liquidity-constrained exposures to hedgeable risk premia. The intended output is a strategy that can be implemented as an overlay using liquid instruments, while remaining auditable for governance purposes.

A core takeaway is that "liquid replication" re-allocates, rather than eliminates, liquidity risk. The relevant fragility shifts from lockups and capital calls to the market depth and crowding of the hedging set. This mirrors the broader liquidity-transformation lesson: normal-times convenience can coexist with stress-time amplification when capacity is impaired. Practically, this motivates explicit liquidity budgets, stress tests of turnover and margin, and monitoring for flow commonality across strategies.

In operational terms, the replica is best viewed as a *control layer*: it provides a daily handle on exposures that are otherwise slow to observe and slow to trade. Used appropriately, it can improve decision quality around rebalancing, liquidity budgeting, and stress preparation. Used inappropriately—for example, as a levered substitute for a PE program—it can create new risks. We therefore emphasize governance, conservative constraints, and explicit communication of what the model is (and is not) designed to replicate.

Extensions include richer nonlinear dynamics (e.g., regime-switching transitions) and explicit funding constraints to tie replication more tightly to macro-financial conditions. These additions would sharpen the link to systemic liquidity-risk mechanisms while preserving the central objective: improving trust in illiquid exposures via transparent, liquid, hedgeable replication.

References

Viral V Acharya, Conor Kehoe, and Michael Reyner. Private equity vs. plc boards in the uk: A comparison of practices and effectiveness. *Journal of Applied Corporate Finance*, 21(1):45–56, 2009.

Andrew Ang, Bingxu Chen, William Goetzmann, and Ludovic Phalippou. Estimating private equity returns from limited partner cash flows. Technical Report 13-83, Columbia Business School Research Paper, 2014.

Ulf Axelson, Tim Jenkinson, Per Strömberg, and Michael Weisbach. Borrow cheap, buy high? the determinants of leverage and pricing in buyouts. *The Journal of Finance*, pages 2223–2269, 2013.

BCG. *Global Asset Management 2020*, May 2020. URL https://image-src.bcg.com/Images/BCG-Global-Asset-Management-2020-May-2020-r_tcm9-247209.pdf. Retrieved from Boston Consulting Group.

Eric Benhamou, Jean-Jacques Ohana, and Beatrice Guez. GRIP: Graphical models revealing insights for portfolio replication - a learning approach. *SSRN Electronic Journal*, April 2024. URL https://ssrn.com/abstract=4780148.

Trond Döskeland and Per Strömberg. Evaluating investments in unlisted equity for the norwegian government pension fund global (gpfg). *Swedish House of Finance Research Paper*, 1(23-03), 2018.

Arpit Gupta and Stijn Van Nieuwerburgh. Valuing private equity investments strip by strip. *The Journal of Finance*, 76(6):3255–3307, 2021.

Michael Jensen. Agency costs of free cash flow, corporate finance, and takeovers. *The American Economic Review*, pages 323–329, 1986.

Steven Kaplan and Per Strömberg. Leveraged buyouts and private equity. *Journal of Economic Perspective*, 23:121–146, 2009.

Steven N. Kaplan and Antoinette Schoar. Private equity performance: Returns, persistence, and capital flows. *The Journal of Finance*, 60(4):1791–1823, 2005. ISSN 00221082, 15406261. URL http://www.jstor.org/stable/3694854.

McKinsey. *McKinsey's Private Markets Annual Review*, February 19 2020. URL https://www.mckinsey.com/industries/private-equity-and-principal-investors/our-insights/mckinseys-private-markets-annual-review. Retrieved from McKinsey.

McKinsey. Mckinsey's private markets annual review 2023. *McKinsey Insights*, 2023. URL https://www.mckinsey.com/industries/private-capital/our-insights/mckinseys-private-markets-annual-review. Accessed online at https://www.mckinsey.com/industries/private-capital/our-insights/mckinseys-private-markets-annual-review.

Kevin P Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.

Taylor D. Nadauld, Berk A. Sensoy, Keith Vorkink, and Michael S. Weisbach. The liquidity cost of private equity investments: Evidence from secondary market transactions. *Journal of Financial Economics*, 132(3):158–181, 2019. ISSN 0304-405X. URL https://www.sciencedirect.com/science/article/pii/S0304405X18303062.

Jean-Jacques Ohana, Eric Benhamou, David Saltiel, and Beatrice Guez. Deep decoding of strategies. *SSRN Electronic Journal*, June 2022. URL https://ssrn.com/abstract=4128693.

L. Pastor and R. Stambaugh. Liquidity risk and expected stock returns. *The Journal of Political Economy*, pages 642–685, 2003.

Daniel Rasmussen and Brian Chinono. Leveraged small value equities. *Available at SSRN 2639647*, 2015. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2639647. Retrieved from SSRN.

David T. Robinson and Berk A. Sensoy. Cyclicality, performance measurement, and cash flow liquidity in private equity. *Journal of Financial Economics*, 122(3):521–543, 2016. ISSN 0304-405X. doi: https://doi.org/10.1016/j.jfineco.2016.09.008. URL https://www.sciencedirect.com/science/article/pii/S0304405X16301593.

Thierry Roncalli and Jerome Teiletche. An alternative approach to alternative beta. *Available at SSRN 1035521*, 2007.

B. Sensoy and D. Robinson. Cyclicality, performance measurement, and cash flow liquidity in private equity. *Journal of Financial Economics*, 2016.

Erik Stafford. Replicating private equity with value investing, homemade leverage, and hold-to-maturity accounting. *The Review of Financial Studies*, 35(1):299–342, 02 2021. ISSN 0893-9454. doi: 10.1093/rfs/hhab020. URL https://doi.org/10.1093/rfs/hhab020.

Thomson Reuters. *Thomson Reuters Private Equity Buyout Index Methodology*, 2014. URL https://www.thomsonreuters.com/content/dam/openweb/documents/pdf/tr-com-financial/methodology/private-equity-buyout-index-methodology.pdf. Available at https://www.thomsonreuters.com/content/dam/openweb/documents/pdf/tr-com-financial/methodology/private-equity-buyout-index-methodology.pdf.