Robust Asset Pricing Tests under Unconditional Heteroskedasticity and Autocorrelation: Revisiting the Factor Zoo

Xinyu Wang, Shaoping Wang, Zhenxin Wang

School of Economics
Huazhong University of Science and Technology

Stylized Facts in Financial Time Series

- **Stylized Facts:** Financial returns often display:
    - **Heteroskedasticity** — time-varying volatility (Engle, 1982; Bollerslev, 1986)
    - **Autocorrelation** — serial dependence in returns
- **Caveat:** These features violate i.i.d. assumptions in classical inference, which may lead to unreliable conclusion
- **Solution: Newey-West (1987) adjustment** is commonly applied in empirical asset pricing (Bali et al, 2016).

Classical HAC Estimation and Its Pitfalls

- Newey-West uses a kernel (i.e., Bartlett) to estimate long-run variance (LRV) - HAC estimator
- However, regardless of kernel or bandwidth choices, kernel-based HAC estimators often (Andrews, 1991):
  - **Over-reject** nulls in finite samples
  - Perform poorly when autocorrelation is strong
- Root problem: reliance on vanishing bandwidth assumptions (Kiefer and Vogelsang, 2005). ($k/T \to 0$ with $T \to \infty$, small-b asymptotics)
  - A clever technical assumption that substantially simplifies asymptotic calculations
  - In practice, there is a given sample size, and some fraction of sample autocovariances is used to estimate the asymptotic variance. Nothing could change the fact that a positive fraction is being used a particular data set.

Fixed-b and Series-Based Alternatives

- **Fixed-b asymptotics** (Kiefer and Vogelsang, 2005):
  - Set bandwidth as fixed fraction of sample size ($k/T \to b$ with $T \to \infty$)
  - Improves finite-sample behavior
  - Test statistics converge to nonstandard limiting distribution
- **Orthonormal series-based HAR inference** (Sun, 2011, 2013):
  - Project residuals onto orthonormal basis (e.g., Fourier)
  - Converge to standard t and F distributions under fixed-b

The Remaining Gap: Unconditional Volatility

- Both HAC and HAR estimators only account for **conditionally heteroskedasticity**.
- In empirical asset pricing applications, volatility is often driven by unconditional shifts:
  - Financial crises
  - Regime changes
  - Exogenous shocks
- Such shifts are not captured by standard assumptions and undermine inference.
- **Motivation**: Build a procedure robust to both serial dependence and unconditional heteroskedasticity.

Our Proposal: Robust t-Test with Dependent Wild Bootstrap

- We develop a **robust t-test** under the fixed-b framework that:
  - Accommodates **unconditional heteroskedasticity**
  - Handles **serial correlation**
  - Yields a **standard t distribution** under general conditions
- **Finite-sample Performance Enhancement**: We introduce the **Dependent Wild Bootstrap** (DWB) (Shao, 2010)
  - Preserves both volatility shifts and dependence
  - Outperforms traditional bootstrap methods in financial time series

1. Introduction
2. Model and Preliminaries
3. Series LRV Estimator and Its Asymptotic Properties
4. Bootstrap HAR t-Test
5. Monte Carlo simulations
6. Application
7. Conclusion

Asset Pricing Regression Framework

We begin with a standard linear factor model:

$$y_t = \alpha + \boldsymbol{\beta}'\mathbf{f}_t + u_t, \quad t = 1, \ldots, T$$

- $y_t = r_t - r_t^f$: excess return over the risk-free rate
- $\mathbf{f}_t = (f_{1t}, \ldots, f_{mt})'$: vector of m pricing factors
- $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_m)'$: factor loadings
- u_t: idiosyncratic error term

OLS Estimation and Scaled Representation

Compact regression form:

$$y_t = X_t'\theta + u_t$$

where $X_t = (1, \mathbf{f}_t')'$ and $\theta = (\alpha, \boldsymbol{\beta}')'$.
OLS estimator:

$$\sqrt{T}(\hat{\theta} - \theta) = \left(\frac{1}{T}\sum_{t=1}^{T} X_t X_t'\right)^{-1} \left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} X_t u_t\right)$$

Goal: Explore the distribution of $\hat{\theta}$ under unconditional heteroskedastic and dependent errors.

Assumption 1: Design Matrix Regularity

Assumption 1 (Factor Process and Regressors)

Let $\mathbf{X}_t = (1, \mathbf{f}_t')'$.

- $\mathbf{f}_t$ is independent of u_t for all t
- $\sup_{t \leq T} \mathbb{E}[||\mathbf{X}_t||^{\nu}] < \infty$ for some $\nu \geq 4$
- $T^{-1}\mathbf{X}'\mathbf{X}$ is positive definite
- $Q = \mathbb{E}[X_t X_t']$ is finite and non-singular
- $\{X_t\}$ is stationary and ergodic
- $\theta \in \text{int}(\Theta) \subset \mathbb{R}^{m+1}$

Assumption 2: Error Process Structure

Assumption 2 (Time-Varying Linear Process)

$$u_t = g(t/T)\, v_t = g(t/T)\, C(L)\varepsilon_t$$

where:

- $\varepsilon_t \overset{i.i.d.}{\sim} (0,1)$ with $\mathbb{E}[\varepsilon_t^{\nu}] < \infty$
- $C(L) = \sum_{\ell=0}^{\infty} c_\ell L^\ell$, with $c_0 = 1$
- $C(1) \in (0,\infty)$, and $\sum_{\ell=0}^{\infty} \ell |c_\ell| < \infty$
- $g(\cdot)$ is deterministic, measurable, and uniformly bounded, with a finite number of discontinuities and satisfying a Lipschitz condition except at those discontinuities

Consistency and Asymptotic Normality

Proposition (Consistency)

Under Assumptions 1–2:

$$\hat{\theta} \xrightarrow{p} \theta \quad \text{as } T \to \infty$$

Proposition (Asymptotic Normality)

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}(0, \Omega)$$

where

$$\Omega = \sum_{j=-\infty}^{\infty} \mathbb{E}[X_t u_t u_{t-j} X'_{t-j}] = \left(\int_0^1 g^2(x)\, dx \right) \cdot Q^{-1} \cdot \left(\sum_{\ell=0}^{\infty} c_\ell^2 \right)$$

1. Introduction
2. Model and Preliminaries
3. **Series LRV Estimator and Its Asymptotic Properties**
4. Bootstrap HAR t-Test
5. Monte Carlo simulations
6. Application
7. Conclusion

From OLS to Robust t-Tests

Goal: Test the linear hypothesis:

$$H_0 : R\theta = r$$

where R is $J \times (m+1)$, and r is a $J \times 1$ vector.

Standard t-statistic:

$$t = \frac{\sqrt{T}(R\hat{\theta}_T - r)}{\sqrt{R\hat{\Omega}R'}}$$

Challenge: Construct a valid estimator $\hat{\Omega}$ for the long-run variance (LRV) matrix Ω, accounting for serial dependence and time-varying volatility.

Baseline: Classical Kernel-Based HAC Estimator

Classical HAC Estimator (Newey-West):

$$\hat{\Omega} = \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} Q(t,s) \hat{u}_t \hat{u}_s$$

- $\hat{u}_t = y_t - X_t' \hat{\theta}$: OLS residuals
- $Q(t,s)$: kernel weighting function (e.g., Bartlett)

Limitations:

- Sensitive to bandwidth and kernel choice
- Finite-sample over-rejection

Series-Based Estimation: Orthonormal Basis

Following Sun (2011, 2013, 2014), we project the error process onto orthonormal basis functions:

Assumption (Series Basis)

$\{\phi_k(x)\}_{k=1}^K$ is a sequence of continuously differentiable and orthonormal basis functions on $L_2[0,1]$ satisfying $\int_0^1 \phi_k(x)\,dx = 0$.

Define the weighting function:

$$Q_K(t,s) = \frac{1}{K}\sum_{k=1}^{K}\phi_k\left(\frac{t}{T}\right)\phi_k\left(\frac{s}{T}\right)$$

This provides a smooth approximation of low-frequency dependence in $\hat{u}_t$.

Series-Based LRV Estimator: Matrix Form

Using the series weighting function $Q_K(t,s)$, we define the LRV estimator as:

$$\hat{\Omega}_K = \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T} Q_K(t,s)\,\hat{u}_t\hat{u}_s$$

Expanded form:

$$\hat{\Omega}_K = \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[\frac{1}{K}\sum_{k=1}^{K}\phi_k\left(\frac{t}{T}\right)\phi_k\left(\frac{s}{T}\right)\right]\hat{u}_t\hat{u}_s$$

This version mirrors the structure of kernel HAC but uses orthonormal series weighting.

Series-Based Estimator: Projection Form

Key simplification:

$$\hat{\Omega}_K = \frac{1}{K}\sum_{k=1}^{K}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\phi_k\left(\frac{t}{T}\right)\hat{u}_t\right)^2$$

Interpretation:

- Each term is the squared projection of the error process onto a basis function.
- Summing over K low-frequency components captures the smoothed LRV.

Advantage: Robust to both autocorrelation and nonstationary volatility.

Asymptotic Distribution of the Test Statistic

Proposition

Let Assumptions 1–3 hold. As $T \to \infty$ with fixed K:

(a) ***Variance estimator:***

$$\frac{\hat{\Omega}_K}{\Omega} \xrightarrow{p} \frac{\chi^2_K}{K}$$

(b) ***Test statistic:***

$$t \xrightarrow{d} t_K$$

Conclusion: Series-based LRV leads to standard t-distribution under fixed-b, enabling valid and convenient inference.

Why Bootstrap the HAR t-Test?

- The fixed-K HAR t-test is asymptotically valid under:
  - General forms of heteroskedasticity
  - Serial correlation
- **Problem:** The limiting t_K distribution may poorly approximate the **finite-sample distribution**, especially when T is small or moderate; does not fully exploit the sample information.
- **Solution:** Apply a bootstrap method that preserves both:
  - Serial dependence
  - Time-varying volatility

The Dependent Wild Bootstrap (DWB)

- We adopt the **Dependent Wild Bootstrap (DWB)** developed by Shao (2010) and further analyzed by Leucht and Neumann (2013)
- **Key Idea:** Use a sequence of random multipliers $\{e_t^*\}$ that mimic dependent and heteroskedastic patterns in original residual sequence.
- **We use:** Ornstein–Uhlenbeck (OU) process discretization:

$$e_t^* = \exp(-1/b_T)\, e_{t-1}^* + v_t, \quad v_t \sim N(0, 1 - \exp(-2/b_T))$$

  - Doukhan et al. (2015) show that this DWB process:
    - Is asymptotically valid
    - Outperforms other bootstraps in practice

Bootstrap HAR t-Test Procedure

1. **Estimate OLS:**

$$y_t = X_t'\hat{\theta}, \quad \hat{u}_t = y_t - X_t'\hat{\theta}$$

   Compute t using HAR estimator.

2. **Simulate multipliers:**

$$e_t^* = \exp(-1/b_T)e_{t-1}^* + v_t, \quad v_t \sim \mathcal{N}(0, 1 - \exp(-2/b_T))$$

3. **Generate bootstrap pseudo-errors and pseudo-samples:**

$$u_t^* = \hat{u}_t e_t^*, \quad y_t^* = X_t'\hat{\theta} + u_t^*$$

   Re-estimate $\hat{\theta}^*$ and compute t_b^* using the same HAR t-test structure.

4. **Repeat:** Run steps 2–3 for B times to form $\{t_b^*\}_{b=1}^B$.

Bootstrap p-Value and Properties

Compute empirical bootstrap p-value:

$$p^* = \frac{1}{B}\sum_{b=1}^{B} \mathbb{I}\{t_b^* > t\}$$

Key Benefits:

- Captures finite-sample features of the test statistic
- Particularly in settings with complex serial dependence and nonstationary volatility, which are prevalent in financial return data
- Provides more reliable inference in small samples

Asymptotic Validity of the Bootstrapped Test Statistic

Proposition

Let all Assumptions hold. As $T \to \infty$ with fixed K, then we have

$$\sup_{x \in \mathbb{R}^p} \left| P^* \left(\sqrt{T} \left(\hat{\beta}^* - \hat{\beta} \right) \leq x \right) - P \left(\sqrt{T} \left(\hat{\beta} - \beta \right) \leq x \right) \right| = o_p(1)$$

1. Introduction
2. Model and Preliminaries
3. Series LRV Estimator and Its Asymptotic Properties
4. Bootstrap HAR t-Test
5. **Monte Carlo simulations**
6. Application
7. Conclusion

Simulation Setup

We consider the famous Fama and French (1993) three-factor model:

$$y_t = \alpha + \sum_{\ell=1}^{3} \beta f_{\ell t} + u_t$$

- $f_{\ell t}$: Simulated factors mimicking Fama and French (1993) (Market, SMB, HML)
- u_t: Innovations drawn from 4 different specifications
- Simulation horizon: $t = -49, \ldots, T$ with burn-in of 50 obs

Factor DGP: FF3 Calibration with AR-GARCH process

Each factor $f_{\ell t}$ is calibrated on factor's monthly time series over the period Jan 2010 to Apr 2024, using the data from Kenneth French website with AR(1)-GARCH(1,1) process:

$$
\begin{aligned}
f_{\ell t} &= 0.84 - 0.04 f_{\ell,t-1} + \sqrt{\eta_{\ell t}}\zeta_{\ell t}, \text{ for } \ell = 1 \ \text{ (Market factor)},\\
f_{\ell t} &= 0.06 - 0.07 f_{\ell,t-1} + \sqrt{\eta_{\ell t}}\zeta_{\ell t}, \text{ for } \ell = 2 \ (SMB),\\
f_{\ell t} &= 0.16 + 0.17 f_{\ell,t-1} + \sqrt{\eta_{\ell t}}\zeta_{\ell t}, \text{ for } \ell = 3 \ (HML),
\end{aligned}
$$

$$
\begin{aligned}
\eta_{\ell t} &= 1.43 + 0.71\eta_{\ell,t-1} + 0.24\zeta^2_{\ell,t-1}, \text{ for } \ell = 1 \ \text{ (Market factor)},\\
\eta_{\ell t} &= 2.24 + 0.44\eta_{\ell,t-1} + 0.33\zeta^2_{\ell,t-1}, \text{ for } \ell = 2 \ (SMB),\\
\eta_{\ell t} &= 0.28 + 0.79\eta_{\ell,t-1} + 0.20\zeta^2_{\ell,t-1},, \text{ for } \ell = 3 \ (HML).
\end{aligned}
$$

Error Scenarios: i.i.d. and Heteroskedastic Cases

(i) i.i.d. case:

$$u_t \sim i.i.d. N(0,1)$$

(ii) Heteroskedasticity case:

$$u_t = g(t/T)\,\varepsilon_t, \quad \varepsilon_t \sim N(0,1)$$

- **Abrupt break:**

  $$g^2(r) = \sigma_0^2 + (\sigma_1^2 - \sigma_0^2) \cdot \mathbf{1}_{\{r \geq \tau\}}$$

  Break location $\tau \in \{0.1, 0.5, 0.9\}$, steepness $\delta = \sigma_1/\sigma_0 \in \{0.2, 5\}$

- **Smooth shift:**

  $$g^2(r) = \sigma_0^2 + (\sigma_1^2 - \sigma_0^2) \cdot r^m$$

  $m \in \{1, 2\}$

Error Scenarios: Serial and Combined Cases

(iii) Serial correlation:

$$u_t = \phi u_{t-1} + e_t, \quad e_t \sim N(0,1)$$

- Persistence levels: $\phi \in \{0.2, 0.7\}$

(iv) Combined heteroskedasticity + serial correlation:

$$u_t = g\left(\frac{t}{T}\right) \cdot v_t, \quad v_t = \phi v_{t-1} + e_t$$

- Combines settings from (ii) and (iii)
- Captures structural volatility shifts with persistence

Summary of Simulation Design

Factor DGP:

- Calibrated AR(1)-GARCH(1,1) from real data
- Burn-in of 50 periods

Error Processes:

- (i) i.i.d.
- (ii) Heteroskedasticity (step/smooth shifts)
- (iii) Serial correlation (AR)
- (iv) Combined (nonstationary + autocorrelated)

Varying:

- Break locations τ, steepness δ
- AR persistence ϕ
- Sample size T, replications R

Finite sample performance

Table 1: Size of considered tests under serial correlation and break shift

			$T = 100$				$T = 200$				$T = 400$			
ϕ	τ	δ	t	t_{NW}	t_{HAR}	t_{DWB}	t	t_{NW}	t_{HAR}	t_{DWB}	t	t_{NW}	t_{HAR}	t_{DWB}
0	0	1	0.060	0.070	0.046	0.048	0.055	0.060	0.063	0.061	0.052	0.055	0.052	0.049
	0.1	0.2	0.039	0.054	0.035	0.051	0.058	0.048	0.034	0.044	0.057	0.062	0.033	0.045
		5	0.052	0.070	0.050	0.053	0.049	0.061	0.059	0.062	0.048	0.052	0.055	0.053
	0.5	0.2	0.059	0.069	0.015	0.035	0.056	0.062	0.018	0.042	0.055	0.057	0.024	0.044
		5	0.054	0.061	0.024	0.040	0.045	0.054	0.028	0.052	0.041	0.053	0.027	0.049
	0.9	0.2	0.057	0.070	0.048	0.051	0.055	0.059	0.060	0.053	0.052	0.056	0.054	0.053
		5	0.045	0.050	0.034	0.051	0.043	0.048	0.037	0.058	0.043	0.045	0.028	0.047
0.2	0	1	0.095	0.079	0.052	0.051	0.093	0.071	0.066	0.066	0.084	0.060	0.052	0.051
	0.1	0.2	0.078	0.074	0.036	0.048	0.094	0.058	0.033	0.046	0.095	0.071	0.035	0.049
		5	0.093	0.076	0.050	0.051	0.090	0.068	0.061	0.061	0.087	0.059	0.052	0.050
	0.5	0.2	0.098	0.082	0.016	0.032	0.090	0.069	0.020	0.041	0.095	0.065	0.023	0.050
		5	0.100	0.076	0.023	0.040	0.082	0.061	0.025	0.049	0.077	0.058	0.027	0.043
	0.9	0.2	0.096	0.077	0.050	0.052	0.090	0.069	0.057	0.059	0.087	0.063	0.055	0.053
		5	0.080	0.063	0.033	0.051	0.084	0.061	0.039	0.062	0.083	0.053	0.028	0.043
0.7	0	1	0.243	0.138	0.055	0.063	0.241	0.126	0.065	0.070	0.234	0.113	0.050	0.056
	0.1	0.2	0.257	0.147	0.040	0.057	0.265	0.123	0.036	0.054	0.242	0.115	0.034	0.049
		5	0.250	0.136	0.052	0.059	0.226	0.126	0.057	0.061	0.231	0.113	0.053	0.054
	0.5	0.2	0.262	0.150	0.017	0.032	0.254	0.126	0.019	0.034	0.242	0.116	0.019	0.041
		5	0.258	0.138	0.020	0.034	0.248	0.125	0.026	0.041	0.230	0.095	0.022	0.038
	0.9	0.2	0.254	0.145	0.054	0.064	0.244	0.128	0.059	0.065	0.239	0.115	0.053	0.055
		5	0.230	0.114	0.041	0.060	0.226	0.116	0.044	0.059	0.236	0.101	0.027	0.043

Finite sample performance

1. Serial correlation induces substantial size distortion in the traditional t test, and the Newey-West adjustment only partially mitigates the distortion when persistence is moderate. In contrast, the t_{HAR} and its DWB version exhibit strong size control
2. The fixed-K HAR t-test tends to under-reject when breaks occur in the middle of the sample. In contrast, the DWB version achieves better size accuracy.

Finite sample performance

Table 2: Size of considered tests under serial correlation and smooth shift

			$T=100$				$T=200$				$T=400$			
ϕ	τ	δ	t	t_{NW}	t_{HAR}	t_{DWB}	t	t_{NW}	t_{HAR}	t_{DWB}	t	t_{NW}	t_{HAR}	t_{DWB}
0	1	0.2	0.059	0.072	0.040	0.046	0.056	0.065	0.053	0.055	0.056	0.057	0.046	0.053
		5	0.052	0.063	0.047	0.051	0.045	0.060	0.055	0.062	0.046	0.049	0.044	0.048
	2	0.2	0.057	0.071	0.040	0.044	0.058	0.064	0.056	0.061	0.054	0.057	0.050	0.051
		5	0.045	0.062	0.041	0.053	0.050	0.060	0.047	0.061	0.046	0.050	0.040	0.052
0.2	1	0.2	0.102	0.081	0.043	0.051	0.091	0.070	0.056	0.058	0.091	0.065	0.048	0.053
		5	0.091	0.076	0.048	0.052	0.089	0.067	0.057	0.064	0.083	0.057	0.043	0.050
	2	0.2	0.104	0.082	0.048	0.050	0.093	0.070	0.058	0.059	0.090	0.066	0.050	0.054
		5	0.092	0.074	0.043	0.051	0.087	0.066	0.047	0.054	0.082	0.059	0.040	0.051
0.7	1	0.2	0.253	0.142	0.047	0.062	0.242	0.126	0.051	0.059	0.238	0.115	0.048	0.057
		5	0.142	0.082	0.046	0.053	0.135	0.080	0.053	0.063	0.129	0.101	0.041	0.051
	2	0.2	0.257	0.146	0.050	0.061	0.236	0.130	0.055	0.055	0.232	0.115	0.052	0.057
		5	0.244	0.138	0.043	0.054	0.234	0.125	0.048	0.051	0.230	0.100	0.038	0.051

1. Both the traditional t and t_{NW} tests exhibit persistent over-rejection regardless of the presence of heteroskedasticity, and the distortion worsens with stronger serial correlation
2. Both the t_{HAR} and t_{DWB} tests maintain accurate size control across all setting.

Finite sample performance

Table 3: Power of considered tests under serial correlation and break

			$T=100$				$T=200$				$T=400$			
ϕ	τ	δ	t	t_{NW}	t_{HAR}	t_{DWB}	t	t_{NW}	t_{HAR}	t_{DWB}	t	t_{NW}	t_{HAR}	t_{DWB}
0	0	1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.1	0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		5	0.483	0.467	0.371	0.369	0.759	0.738	0.521	0.559	0.949	0.945	0.781	0.798
	0.5	0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		5	0.705	0.706	0.678	0.666	0.921	0.912	0.845	0.851	0.996	0.995	0.962	0.960
	0.9	0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		5	0.997	0.989	0.944	0.933	1.000	1.000	0.983	0.986	1.000	1.000	0.999	0.999
0.2	0	1	1.000	1.000	0.998	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.1	0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		5	0.350	0.358	0.288	0.294	0.597	0.587	0.391	0.432	0.829	0.823	0.638	0.647
	0.5	0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		5	0.532	0.533	0.569	0.557	0.793	0.789	0.734	0.739	0.973	0.965	0.892	0.891
	0.9	0.2	1.000	1.000	0.999	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		5	0.967	0.951	0.894	0.883	0.998	0.994	0.950	0.951	1.000	1.000	0.993	0.994
0.7	0	1	0.467	0.455	0.572	0.628	0.730	0.718	0.746	0.833	0.938	0.934	0.968	0.971
	0.1	0.2	0.989	0.983	0.995	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		5	0.098	0.096	0.112	0.128	0.125	0.121	0.132	0.165	0.176	0.174	0.188	0.207
	0.5	0.2	0.740	0.733	0.901	0.865	0.921	0.913	0.978	0.960	0.997	0.996	1.000	0.997
		5	0.129	0.134	0.238	0.184	0.185	0.190	0.302	0.278	0.232	0.228	0.448	0.415
	0.9	0.2	0.497	0.470	0.614	0.655	0.751	0.756	0.815	0.855	0.954	0.953	0.977	0.980
		5	0.373	0.374	0.423	0.440	0.442	0.455	0.570	0.593	0.646	0.647	0.807	0.781

Finite sample performance

Table 4: Power of considered tests under serial correlation and model smooth function

			$T=100$				$T=200$				$T=400$			
ϕ	τ	δ	t	t_{NW}	t_{HAR}	t_{DWB}	t	t_{NW}	t_{HAR}	t_{DWB}	t	t_{NW}	t_{HAR}	t_{DWB}
0	1	0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		5	0.696	0.684	0.550	0.557	0.915	0.908	0.737	0.765	0.997	0.996	0.947	0.937
	2	0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		5	0.831	0.843	0.713	0.708	0.978	0.973	0.876	0.894	1.000	1.000	0.982	0.982
0.2	1	0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		5	0.537	0.528	0.438	0.434	0.785	0.775	0.566	0.646	0.970	0.966	0.854	0.847
	2	0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		5	0.676	0.691	0.594	0.586	0.895	0.889	0.762	0.788	0.996	0.995	0.934	0.939
0.7	1	0.2	0.681	0.671	0.824	0.834	0.918	0.911	0.958	0.963	0.999	0.997	0.998	0.998
		5	0.294	0.305	0.263	0.276	0.463	0.464	0.340	0.403	0.720	0.711	0.602	0.596
	2	0.2	0.588	0.574	0.724	0.751	0.853	0.845	0.897	0.931	0.981	0.980	0.992	0.994
		5	0.160	0.164	0.215	0.207	0.199	0.201	0.262	0.295	0.305	0.300	0.446	0.439

Finite sample performance

1. The traditional t test suffers from severe size distortion under serial correlation, while the Newey-West adjustment t_{NW} only slightly alleviates this issue.
2. Both t_{HAR} and t_{DWB} demonstrate robust size control even in the presence of substantial serial dependence
3. t_{HAR} tends to under-reject under certain heteroskedasticity settings, whereas t_{DWB} consistently maintains accurate size across all scenarios
4. t_{DWB} achieves this without sacrificing power.

Revisiting the "Factor Zoo" in China

- Our method is applicable to prominent asset pricing tests:
  - α-tests, factor redundancy, spanning, and univariate significance tests
- We re-evaluate **31 pricing anomalies** in the Chinese stock market, previously found significant under standard methods (Hsu et al., 2018' Liu et al., 2019; Jansen et al., 2021; Li et al., 2024; Wang and Zhu, 2024).
- **Key finding:** Only a small fraction remain significant using our robust procedure.

Anomaly returns with 1-week holding period

Anomaly	Return	t	p	t_{HAC}	p_{HAC}	t_{HAR}	p_{HAR}	p_{DWB}
Trading-based								
1. Liquidity								
ILLIQ	0.65	4.33	**0.00**	3.65	**0.00**	2.25	0.11	0.11
YT	1.10	6.35	**0.00**	5.86	**0.00**	7.06	**0.01**	**0.00**
ABT	0.59	6.18	**0.00**	5.30	**0.00**	2.93	0.06	0.06
VDTV	-0.88	-5.93	**0.00**	-4.49	**0.00**	-3.23	**0.05**	**0.03**
VTURN	0.07	0.42	0.68	0.40	0.69	0.36	0.74	0.72
Size	0.39	2.52	**0.01**	2.22	**0.03**	1.60	0.21	0.18
2. Risk								
VOL	3.07	13.17	**0.00**	11.93	**0.00**	10.67	**0.00**	**0.00**
MAX	3.81	21.24	**0.00**	20.15	**0.00**	16.08	**0.00**	**0.00**
3. Past Returns								
REV	7.05	39.36	**0.00**	32.48	**0.00**	33.83	**0.00**	**0.00**
ABR	1.03	7.85	**0.00**	7.65	**0.00**	16.53	**0.00**	**0.00**
RR	0.11	1.28	0.20	1.30	0.19	1.24	0.30	0.31
RS	0.30	3.61	**0.00**	3.62	**0.00**	6.93	**0.01**	**0.00**
SUE	0.58	6.34	**0.00**	5.89	**0.00**	22.35	**0.00**	**0.00**
TES	-0.15	-1.81	0.07	-1.80	0.07	-1.43	0.25	0.28
Accounting-based								
4. Profitability								
ROE	0.12	1.31	0.19	1.28	0.20	0.92	0.43	0.46
ROA	0.20	2.27	**0.02**	2.08	**0.04**	2.02	0.14	0.13
GPLAQ	-0.19	-2.05	**0.04**	-2.09	**0.04**	-2.80	0.07	0.11
OPLA	0.24	2.65	**0.01**	2.72	**0.01**	1.79	0.17	0.17
OPLE	0.09	1.01	0.31	1.04	0.30	0.65	0.56	0.60
SGQ	0.53	6.68	**0.00**	6.55	**0.00**	6.69	**0.01**	**0.01**
5. Value								
EP	0.11	1.95	**0.05**	1.96	**0.05**	2.36	0.10	0.14
BM	-0.51	-2.28	**0.02**	-2.16	**0.03**	-1.98	0.14	0.13
CP	0.07	1.76	0.08	1.69	0.09	0.90	0.43	0.46
6. Others								
AAG	0.10	0.95	0.34	0.88	0.38	0.80	0.48	0.53
ACC	-0.04	-0.58	0.56	-0.62	0.54	-0.70	0.53	0.51
NOA	0.00	-0.01	1.00	-0.01	1.00	-0.01	0.99	0.99
7. Intangible								
ALA	-0.38	-3.60	**0.00**	-3.54	**0.00**	-3.25	0.05	0.05
CTA	-0.27	-2.74	**0.01**	-2.35	**0.02**	-1.56	0.22	0.25
RA1	0.20	2.04	**0.04**	2.28	**0.02**	1.32	0.28	0.30
TAN	-0.31	-2.49	**0.01**	-2.29	**0.02**	-1.87	0.16	0.17

t & t_{NW} tests: 21 significant anomalies V.S. t_{HAR} & t_{DWB}: 9 significant

Anomaly returns with 4-week holding period

Anomaly	Return	t	p	t_{HAC}	p_{HAC}	t_{HAR}	p_{HAR}	p_{DWB}
Trading-based								
1. Liquidity								
ILLIQ	0.56	6.64	**0.00**	3.78	**0.00**	2.61	0.08	0.10
YT	1.08	11.63	**0.00**	6.54	**0.00**	6.30	**0.01**	**0.02**
ABT	0.57	10.53	**0.00**	6.30	**0.00**	3.68	**0.03**	**0.01**
VDTV	-0.05	-0.40	0.69	-0.21	0.83	-0.11	0.92	0.96
VTURN	0.02	0.25	0.80	0.15	0.88	0.12	0.91	0.93
Size	0.35	4.10	**0.00**	2.26	**0.02**	1.52	0.23	0.23
2. Risk								
VOL	1.87	17.10	**0.00**	10.33	**0.00**	7.79	**0.00**	**0.00**
MAX	2.53	27.81	**0.00**	18.37	**0.00**	14.09	**0.00**	**0.01**
3. Past return								
REV	1.62	20.94	**0.00**	13.04	**0.00**	10.77	**0.00**	**0.00**
ABR	0.50	10.06	**0.00**	6.44	**0.00**	6.46	**0.01**	**0.00**
RR	0.06	1.43	0.15	0.84	0.40	0.67	0.55	0.58
RS	0.26	6.16	**0.00**	3.67	**0.00**	6.15	**0.01**	**0.00**
SUE	0.48	9.88	**0.00**	5.59	**0.00**	21.15	**0.00**	**0.00**
TES	-0.12	-2.97	**0.00**	-1.66	0.10	-1.08	0.36	0.42
Accounting-based								
4. Profitability								
ROE	0.10	2.12	**0.03**	1.23	0.22	0.78	0.49	0.54
ROA	0.18	3.73	**0.00**	2.09	**0.04**	1.94	0.15	0.15
GPLAQ	-0.17	-4.03	**0.00**	-2.25	**0.02**	-2.38	0.10	0.10
OPLA	0.22	5.22	**0.00**	2.95	**0.00**	1.73	0.18	0.19
OPLE	0.10	2.50	**0.01**	1.44	0.15	0.82	0.47	0.51
SGQ	0.49	12.58	**0.00**	7.37	**0.00**	7.30	**0.01**	**0.00**
5. Value								
EP	0.09	3.29	**0.00**	1.90	0.06	1.83	0.16	0.18
BM	-0.43	-4.37	**0.00**	-2.50	**0.01**	-1.99	0.14	0.16
CP	0.07	3.22	**0.00**	1.76	0.08	0.85	0.46	0.48
6. Others								
AAG	0.09	1.54	0.12	0.87	0.38	0.75	0.51	0.54
ACC	-0.04	-1.35	0.18	-0.82	0.41	-0.78	0.49	0.51
NOA	-0.01	-0.11	0.91	-0.06	0.95	-0.08	0.94	0.94
7. Intangible								
ALA	-0.37	-7.32	**0.00**	-4.10	**0.00**	-3.46	**0.04**	**0.05**
CTA	-0.29	-5.25	**0.00**	-2.84	**0.00**	-1.81	0.17	0.16
RA1	0.10	2.06	**0.04**	1.40	0.16	1.23	0.31	0.25
TAN	-0.32	-4.74	**0.00**	-2.64	**0.01**	-2.03	0.14	0.13

t tests: 24 significant anomalies V.S. t_{NW}: 18 significant V.S. t_{HAR} & t_{DWB}: 9

Revisiting the "Factor Zoo" in China

1. These results highlight the extent to which conventional inference methods may overstate the statistical significance of anomalies, thus resulting in the so-called "factor zoo" problem.
2. By accounting for heteroskedasticity and serial dependence, our proposed framework offers more credible inference and helps taming such problems, thereby separate genuine pricing patterns from statistical artifacts.

1. Introduction
2. Model and Preliminaries
3. Series LRV Estimator and Its Asymptotic Properties
4. Bootstrap HAR t-Test
5. Monte Carlo simulations
6. Application
7. **Conclusion**

Main Contributions

- **Methodological**: We develop a new t-test that extends HAR inference to settings with **unconditional volatility shifts and serial correlation**.
- **Bootstrapping Innovation**: We introduce a **dependent wild bootstrap version** of the test and prove its validity under fixed-b asymptotics.
- **Empirical Insight**: We show that many asset pricing anomalies in the Chinese market may be **statistical artifacts** when more robust inference is applied.
- Together, these contributions offer a new benchmark for empirical asset pricing in the presence of realistic error structures.

Conclusion

- We propose a robust inference framework for asset pricing tests using **fixed-b asymptotic theory**.
- Our method accommodates both **unconditional heteroskedasticity** and **serial dependence**.
- To improve finite-sample accuracy, we enhance the test with the **dependent wild bootstrap (DWB)**.
- We derive the asymptotic distribution and establish the **validity of the bootstrap-enhanced t-test**.
- **Monte Carlo simulations** confirm superior size control and power relative to standard t, t_{NW}, and t_{HAR} tests.
- Applied to the Chinese **"factor zoo"**, our method shows that most anomalies are not robust—highlighting the need for credible inference.

Takeaway: Credible inference under realistic error structures is essential for reliable asset pricing conclusions.

Thank you for listening !