Beyond First-Order Bias in Predictive Regressions: Estimation, Inference, and Return Predictability *

Liang Jiang[a], Ron Kaniel[b], Yayi Yan[c], and Jun Yu[d]

[a]International School of Finance, Fudan University.

[b]Simon School of Business, University of Rochester.

[c]School of Statistics and Data Science, Shanghai University of Finance and Economics.

[d]Faculty of Business Administration, University of Macau

March 15, 2026

Abstract

This paper reveals and addresses a large and overlooked source of distortion in predictive regressions: higher-order bias. While existing corrections only address first-order bias, we show that higher-order terms in both direct regression (DR) and implication (IM) based estimators can seriously mislead both statistical and economic conclusions, especially with persistent predictors and long horizons. We then develop iBoot, a unified framework that integrates indirect inference with bootstrap methods. iBoot (i) jointly corrects first- and higher-order biases in both estimation and inference for DR and IM estimators, and (ii) applies seamlessly to both short- and long-horizon regressions. Simulations demonstrate that iBoot achieves minimal bias, accurate confidence intervals, correct test size, and high power. When applied to classic return predictors, iBoot materially changes the economic significance of return predictability. This paper thus establishes a new foundation for credible and robust estimation and inference in predictive regressions.

Keywords: Bias correction, predictive regression, indirect inference, bootstrap, return predictability inference

JEL codes: G11, G12, G17

*We would like to thank Mikhail Golosov, Dashan Huang, Raymond Kan, Kai Li, Nadya Malenko, Jun Tu, Keli Xu, and seminar/conference participants at Hong Kong Polytechnic University, Chinese University of Hong Kong, Shanghai Jiao Tong University for helpful discussions and comments.

1 Introduction

For decades, economists and practitioners have sought to uncover time-varying patterns in expected market returns. In this pursuit, researchers have identified numerous predictors for future returns – such as dividend-price ratios, earnings-price ratios, and other financial metrics – typically analyzed through predictive regressions. A central challenge in this framework is that the ordinary least squares (OLS) estimator is biased in finite samples, distorting both coefficient estimates and statistical inference. To address this issue, the literature has developed first-order (order $1/T$) bias corrections for short-horizon (Stambaugh, 1999) and long-horizon (Boudoukh et al., 2022) predictive regressions, which have become standard tools in empirical practice.

This paper shows that such corrections are inadequate in many empirically relevant settings. As illustrated in Figure 1, when predictor persistence – measured by the autoregressive parameter ρ – is high, the true bias of the OLS estimator can greatly exceed its first-order approximation, accounting for nearly one-quarter of total distortion even at short horizon and rising sharply at longer horizons. Standard first-order bias corrections thus systematically underestimate actual bias and can lead to misleading conclusions about return predictability. This problem is pervasive in practice, as many financial predictors are highly persistent.[1] Yet it has been largely overlooked in the literature, calling into question conventional approaches to estimation and inference in predictive regressions.

Why does the gap between true and first-order bias widen as persistence increases? To answer this question, we derive exact bias expressions for both the direct regression (DR) and implication-based (IM) estimators and then develop analytical second-order (order $1/T^2$) approximations to clarify the underlying mechanism.[2] The key insight is that second-order bias decays very slowly as predictor persistence rises. In particular, the approximation includes terms proportional to $1/(1-\rho)^k$, where $k = 2, 3$ denotes the power on $(1-\rho)$. As ρ approaches one, these terms grow rapidly, causing the bias beyond the first-order to become large in finite samples despite being asymptotically negligible. This explains the substantial divergence between true and first-order bias – hereafter referred to as higher-order bias – documented in Figure 1.

To address this issue, we develop *iBoot*, a unified estimation and inference framework that integrates indirect inference with bootstrap methods. Indirect inference has been widely used for bias reduction in economics and finance (Gouriéroux et al., 2000; Phillips and Yu, 2009; Gouriéroux et al., 2010; Jiang et al., 2018). In particular, Phillips (2012) shows that indirect inference can

[1] In the Goyal and Welch (2008) dataset, roughly one-third of annual predictors have $\widehat{\rho} > 0.85$, and more than two-thirds of quarterly and monthly predictors exceed $\widehat{\rho} > 0.9$. See Figure 15 in the Appendix for persistence estimates of all predictors in Goyal and Welch (2008).

[2] We focus on second-order terms for exposition: they are the leading components beyond first order, while higher-order terms quickly become analytically intractable. Importantly, the estimation and inference framework introduced later in the paper can accommodate bias correction beyond second order.

eliminate nearly all bias in estimating ρ and is robust to its stationary and nonstationary values.

Correcting the bias in estimating ρ alone, however, is not sufficient in predictive regressions. As our exact bias expressions demonstrate, the biases of the DR and IM estimators are considerably more intricate, especially at long horizons. They depend not only on the bias in the autoregressive coefficient estimator $\widehat{\rho}$, but also on interaction terms involving $\widehat{\rho}$ and its higher-order powers, such as $\widehat{\rho}^2$, $\widehat{\rho}^3$. Even if each component is individually unbiased, their interactions may not be, because these terms are correlated. Consequently, indirect inference by itself cannot fully address the bias in predictive regressions.

Our iBoot framework overcomes this challenge by combining indirect inference with bootstrap methods. Indirect inference first delivers nearly unbiased short-horizon parameter estimates, which are then used to generate bootstrap samples that closely mimic the true data-generating process. These realistic bootstrap samples are subsequently employed to eliminate remaining bias arising from interactions among $\widehat{\rho}$ and its powers at long horizons and to build precise confidence intervals and hypothesis tests.

The resulting iBoot estimators, named as *iBoot DR* and *iBoot IM*, are the first to jointly correct first- and higher-order bias across short and long horizons. Extensive Monte Carlo simulations show that iBoot consistently delivers minimal bias across all sample sizes, persistence levels, and horizons. Gains are particularly pronounced for persistent predictors and long horizons, where conventional methods fail.

The accompanying inference procedure likewise provides the first valid hypothesis tests and confidence intervals that account for higher-order bias. Compared with OLS with the Newey-West or Hodrick standard errors, the first-order bias-corrected estimator of Boudoukh et al. (2022) (BIR), and the IM method of Xu (2020) (IM), iBoot achieves three distinct improvements: (i) accurate confidence interval coverage across all horizons, (ii) exact size control even in small samples, and (iii) higher statistical power when sizes are matched. This dual strength – exact inference with strong power – makes iBoot a robust and general-purpose tool for predictive regression analysis.

In the empirical analysis, we first apply iBoot to several well-known return predictors in Goyal et al. (2024). In the direct regression (DR) setting, this includes highly persistent predictors such as dividend yield, forward implied variance, and government investment; in the implication-based (IM) setting, we examine the treasury bill rate, cross-section based tail risk, and equity issuing activity. Across monthly, quarterly, and annual frequencies, we find that iBoot leads to significant revisions relative to conventional methods. For instance, iBoot DR uncovers significant return predictability of the dividend yield at the monthly frequency across horizons, overturning the insignificance reported by OLS and BIR; iBoot IM eliminates the spurious long-horizon significance found by IM for the monthly treasury bill rate. Similar revisions occur at quarterly and annual frequencies for forward implied variance and government investment at long horizons in DR, as well as for cross-sectional tail risk and equity issuing activity in IM, highlighting the pervasive higher-order bias across predictors, frequencies, and horizons.

We then extend the analysis to the full universe of predictors in Goyal et al. (2024). Across these predictors and horizons, iBoot frequently overturns conclusions from conventional methods. At the monthly frequency, for example, iBoot changes statistical significance relative to BIR for 8 of the 35 predictors at the short horizon – five shifting from insignificance to significance and three from significance to insignificance – and for 19 predictors at longer horizons – seven shifting from insignificance to significance and twelve in the opposite direction. It also reverses inference relative to both OLS and BIR: five predictors that are insignificant under both methods become significant under iBoot, while six previously significant predictors become insignificant. Sometimes, iBoot restores significance for predictors whose predictive power is removed by BIR and even occasionally reverses the sign of estimated coefficients.

These patterns indicate that correcting only first-order bias, as in BIR, can leave substantial distortions in predictive regressions – distortions that iBoot helps mitigate. Importantly, the resulting changes in predictability conclusions stem not only from more accurate point estimates but also from more accurate confidence intervals. By jointly correcting higher-order bias and precisely accounting for finite-sample uncertainty, iBoot delivers sharper and more credible inference, reshaping empirical conclusions on return predictability of these predictors.

Our contributions are threefold. First, we demonstrate that both DR and IM estimators exhibit sizable higher-order bias and that standard first-order corrections may substantially underestimate true bias, often leading to misleading conclusions about return predictability – an issue largely overlooked in existing work. Second, we develop iBoot, a unified estimation and inference framework that jointly corrects first- and higher-order biases in both DR and IM estimators across horizons. Compared to state-of-the-art methods, iBoot delivers markedly lower bias, more accurate confidence intervals, correct test sizes, and higher power, offering a comprehensive and reliable solution for predictive regression analysis. Third, we provide new empirical evidence showing that correcting higher-order bias materially changes conclusions about return predictability. Together, these contributions establish a new foundation for credible and robust estimation and inference in predictive regressions.

Our paper builds on several strands of literature. The most closely related works are Stambaugh (1999), Boudoukh et al. (2022), and Kan and Pan (2022).The first two papers develop first-order bias corrections for short- and long-horizon DR estimators, respectively. The latter further derive second-order asymptotic approximations for the mean and variance of the long-horizon DR estimator under normality of the error terms, but their proposed test statistic relies only on a first-order bias correction.[3] Amihud and Hurvich (2004) propose an augmented regression to mitigate first-order bias in short-horizon DR settings, and Boudoukh et al. (2022) also derive the first-order bias of the IM estimator under $\beta_1 = 0$. We complement and advance this literature along three dimensions. First, we identify and quantify large higher-order bias in both DR and IM estimators, without imposing normality. Second, we reveal that the first-order corrections can be inadequate in empirically relevant settings. Third, we introduce a unified framework that corrects both first- and higher-order distortions in estimation and inference across horizons.

The second strand concerns inference for return predictability, including Hodrick (1992), Valkanov (2003), Campbell and Yogo (2006), Hjalmarsson (2011), Kostakis et al. (2015), and Xu (2020). These works propose valid testing procedures for return predictability but neglect higher-order bias. iBoot complements and extends this literature by offering an inference that remains reliable even when such bias is large.

Finally, our study connects to the broader econometric literature on indirect inference and bootstrapping (Smith, 1993; Gourieroux et al., 1993; Efron, 1979; Phillips, 2015). While indirect inference has been used for bias reduction and the bootstrap for resampling, we are the first to integrate the two into a unified framework that delivers robust estimation and inference for predictive regressions.

The remainder of the paper is organized as follows. Section 2 analyzes the higher-order bias in conventional DR and IM estimators. Section 3 introduces the iBoot estimators and the corresponding inference methods. Section 4 presents simulation results. Section 5 reports empirical findings. Section 6 concludes.

2 High-order bias in predictive regression

The predictive model is often formulated as

$$r_t = \alpha_1 + \beta_1 x_{t-1} + u_t, \tag{2.1}$$

$$x_t = c + \rho x_{t-1} + v_t, \tag{2.2}$$

[3] Kendall (1954) and Marriott and Pope (1954) derive the first-order bias of $\widehat{\rho}_i$, the i-th order sample autocorrelation of the predictor. These results are used by Stambaugh (1999) and Boudoukh et al. (2022) to obtain first-order bias expressions for $\widehat{\beta}_h$ at short and long horizons, respectively. Kiviet and Phillips (2012) derive the second-order bias of $\widehat{\rho}$ in autoregressive models, but their analysis is restricted to the autoregressive equation and does not extend to predictive regressions, nor does it provide results for $\widehat{\rho}_i$, which is central to long-horizon bias. Likewise, Yu (2012) study second-order bias in an Ornstein-Uhlenbeck process with a known mean, a setting that differs from the predictive regression framework considered here.

where r_t is the logarithm of excess market return, x_t is the predictor, $|\rho| < 1$, $[u_t, v_t]' \sim i.i.d.(0, \Sigma)$ with $\Sigma = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix}$, and $E[u_t|v_t] = \frac{\sigma_{uv}}{\sigma_v^2} v_t$.[4]

We can add up the returns in Eq.(2.1) from t to $t+h-1$ ($h \geq 1$) and keep x_{t-1} as the predictor by backward substitution to obtain the h-horizon predictive regression model as:

$$\sum_{i=0}^{h-1} r_{t+i} = \alpha_h + \beta_h x_{t-1} + u_{h,t}, \tag{2.3}$$

where α_h is the aggregated intercept, $u_{h,t} = \sum_{i=0}^{h-1} u_{t+i} + 1_{[h\geq 2]} \left(\sum_{i=0}^{h-2} \sum_{j=0}^{h-2-i} \rho^j v_{t+i} \right) \beta$ is the aggregated error, and $1_{[\cdot]}$ is the indicator function.

Eq.(2.3) accommodates short- and long-horizon predictive regression. When $h = 1$, it reduces to the short-horizon predictive regression in Eq.(2.1). When $h \geq 2$, it becomes the long-horizon predictive regression discussed in Boudoukh et al. (2022). β_h is the parameter of interest in predictive regression, which quantifies how much a predictor x_{t-1} can predict future returns.

The literature offers two main approaches to estimate β_h. The first is the DR approach, which estimates β_h by applying OLS directly to Eq. (2.3). This method is widely used in empirical asset pricing studies. The second is the IM approach, which first estimates β_1 and the predictor's persistence ρ from Eqs. (2.1) and (2.2), respectively, and then computes

$$\beta_h = \beta_1 \sum_{i=0}^{h-1} \rho^i. \tag{2.4}$$

This approach, dating back to Campbell and Shiller (1988), has been adopted in numerous subsequent studies, including Hodrick (1992), Boudouk and Richardson (1994), Cochrane (2008), Lettau and Van Nieuwerburgh (2008), and Xu (2020). When the prediction horizon is one period ($h = 1$), these two approaches are equivalent and the slope coefficient β_1 is estimated by the short-horizon regression in Eq. (2.1).

In the remainder of this section, we derive the bias formulas for both DR and IM estimators and show that each exhibits substantial higher-order bias. We further explain why this bias becomes particularly pronounced for highly persistent predictors and long forecasting horizons.

[4] We do not require normal distribution for u_t and v_t as in Stambaugh (1999), Boudoukh et al. (2022), and Kan and Pan (2022), but only require $E[u_t|v_t]$ is linear in v_t. So, our setting is more general. In the Appendix, we consider non-normal and GARCH(1,1) errors in simulations, and the results are consistent with our theoretical findings.

2.1 Bias in the DR estimator

By the OLS formula, the DR estimator of β_h is given by

$$\widehat{\beta}_h = \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \sum_{i=0}^{h-1} r_{t+i}}{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1}^2}, \tag{2.5}$$

where $\tilde{x}_{t-1} = x_{t-1} - \sum_{t=1}^{T-h+1} x_{t-1}/(T-h+1)$. $\widehat{\beta}_h$ is the conventional DR estimator for β_h and is well known to be biased in finite sample.

The following proposition gives the exact bias formula for $\widehat{\beta}_h$.

Proposition 2.1. *The bias of $\widehat{\beta}_h$ in Eq.(2.5) is*

$$E[\widehat{\beta}_h] - \beta_h = \left[\beta_1 + \frac{\sigma_{uv}}{\sigma_v^2}(1-\rho)\right] \sum_{i=0}^{h-1} E[\widehat{\rho}_i - \rho_i] + \frac{\sigma_{uv}}{\sigma_v^2} E\left[\widehat{\rho}_h - \rho_h\right] \tag{2.6}$$

where $\rho_i := \rho^i$ is the i-th order autocorrelation of x_t, $\widehat{\rho}_i$ is its OLS estimator for $i \geq 1$, and $\widehat{\rho}_0 = 1$.

When $h = 1$, the exact bias formula reduces to the classic result in Stambaugh (1999),

$$E[\widehat{\beta}_1] - \beta_1 = \frac{\sigma_{uv}}{\sigma_v^2} E[\widehat{\rho} - \rho],$$

showing that the bias in $\widehat{\beta}_1$ is driven entirely by the bias in $\widehat{\rho}$ and the covariance-to-variance ratio σ_{uv}/σ_v^2.

For $h \geq 2$, the bias structure becomes more involved. The bias of $\widehat{\beta}_h$ arises from the cumulative bias in the higher-order autocorrelations $\widehat{\rho}_i$, as well as the persistence parameter ρ, the error covariance ratio σ_{uv}/σ_v^2, and the forecast horizon h under the null $\beta_1 = 0$. When $\beta_1 \neq 0$, the bias of $\widehat{\beta}_h$ additionally depends on β_1 itself, consistent with Boudoukh et al. (2022).

Figure 2 further demonstrates the magnitude of higher-order bias by comparing the true bias of $\widehat{\beta}_h$ (solid line) against its first-order one (dashed line). Like Figure 1, these biases are simulated under identical parameters except for $\widehat{\Sigma}$, the sample covariance of u_t and v_t, which we compute using the predictor scaled risk-neutral vix (rsvix). The different sign of the bias in Figure 2 from that in Figure 1 is due to the negative correlation between u_t and v_t, which leads to a positive bias in $\widehat{\rho}$ for the rsvix predictor.

A key insight from Figure 2 is that as $\rho \to 1$, the true bias of $\widehat{\beta}_h$ increasingly diverges from its first-order approximation. At the short horizon (top panel), when $\rho = 0.99$, the true bias of $\widehat{\beta}_1$ is 0.325, compared with a first-order value of 0.25, so higher-order bias contributes 0.075 – about 30% of the first-order magnitude. The gap widens further at longer horizons. For the DR estimator at $h = 3$ (left side of the bottom panel), the true bias (about 0.95) exceeds its first-order approximation (about 0.73) by roughly 0.22, reflecting the compounding effects of high persistence and longer forecast horizons, even though the proportional contribution of higher-order bias remains similar.

Why is the higher-order bias so large? This seems counterintuitive because higher-order terms like $1/T^k$ ($k \geq 2$) are typically negligible in asymptotic expansions. To explain this, we expand the bias formula for $\widehat{\beta}_h$ (Eq. 2.6) to the second order, which leads to Proposition 2.2.

Proposition 2.2 demonstrates that the bias of $\widehat{\beta}_h$ depends on multiple elements: persistence (ρ), the covariance-to-variance ratio (σ_{uv}/σ_v^2), prediction horizon (h), sample size (T), and the higher moments of v_t (skewness γ_1 and excess kurtosis γ_2). The first-order component, given by the sum of the leading terms in (2.7) and (2.8), coincides with Proposition 2 of Boudoukh et al. (2022), while the remaining terms introduce additional second-order terms that capture nonlinear interactions among these quantities.

Proposition 2.2. *For any $h \geq 1$, the approximate bias of the OLS estimator $\widehat{\beta}_h$ is*

$$
\begin{aligned}
& E[\widehat{\beta}_h - \beta_h] \\
= \; & -\frac{1}{T}\frac{\sigma_{uv}}{\sigma_v^2}\left[\frac{2\rho(1-\rho^h)}{1-\rho} + h(1+\rho)\right] - \frac{1}{T^2}\frac{\sigma_{uv}}{\sigma_v^2}\left[\frac{A_1^{DR}(\rho,h)}{(1-\rho)^2(1+\rho)} + \frac{(h^2-3h)\rho + h^2 - h + 2}{2}\right] && (2.7) \\
& -\frac{\beta_1}{T}\left[\frac{(2h-1)(\rho-\rho^h)}{1-\rho} + h - 1\right] - \frac{\beta_1}{T^2}\left[\frac{A_2^{DR}(\rho,h)}{(1-\rho)^3(1+\rho)} + \frac{(h-1)^2\rho}{1-\rho} + \frac{h^2-h+2}{2}\right] && (2.8) \\
& -\frac{1}{T^2}\frac{\sigma_{uv}}{\sigma_v^2}\left[\frac{2\gamma_2 A_3^{DR}(\rho,h)}{1-\rho} + \frac{4\gamma_1^2 A_4^{DR}(\rho,h)}{(1-\rho)(1+\rho+\rho^2)}\right] - \frac{\beta_1}{T^2}\left[\frac{2\gamma_2 A_5^{DR}(\rho,h)}{(1-\rho)^2} + \frac{4\gamma_1^2 A_6^{DR}(\rho,h)}{(1-\rho)^2(1+\rho+\rho^2)}\right] && (2.9) \\
& + o(T^{-2}),
\end{aligned}
$$

where $\gamma_1 = E(v_t^3)/\sigma_v^3$ and $\gamma_2 = E(v_t^4)/\sigma_v^4 - 3$ denote the skewness and excess kurtosis of v_t, respectively, and $\{A_i^{DR}(\rho,h)\}_i$ are nonlinear functions of ρ and h that are provided in the proof. In addition, we have $-(\rho^3+\rho^2+3\rho+1) \leq A_1^{DR}(\rho,h) \leq \frac{16}{e}$, $-(3\rho^3+\rho^2+\rho+1) \leq A_2^{DR}(\rho,h) \leq \frac{6}{e} + \frac{24}{e^2}$, where e is Euler's number, and for $i = 3,4,5,6$, $|A_i^{DR}(\rho,h)| \leq 1$ uniformly over $\rho \in [0,1]$ and $h \geq 1$.

The key economic implication is that these second-order terms can be strongly amplified by persistence. Several components scale with $1/(1-\rho)^k$ for integers $k \geq 1$; in particular, the leading terms associated with σ_{uv}/σ_v^2 and β_1 grow at rates $1/(1-\rho)^2$ and $1/(1-\rho)^3$. As ρ approaches one, these terms increase rapidly, so the second-order bias may remain sizable even when asymptotically negligible. The coefficient functions $A_1^{DR}(\rho,h)$ and $A_2^{DR}(\rho,h)$ reinforce this amplification: A_1^{DR} approaches its maximum (about $16/e$) as $\rho \to 1$, while A_2^{DR} converges to $24/e^2 + 6/e$ when both ρ and h are large. Together, these features explain why higher-order bias can become substantial for highly persistent predictors.

The left panel of Figure 3 illustrates this amplification. Under $\beta_1 = 0$, $\rho = 0.97$, and $\sigma_{uv}/\sigma_v^2 = -5$, the second-order component remains economically meaningful relative to the first-order bias. At $h = 3$, the first-order bias is about 0.6, while the second-order bias is roughly 0.2 – around one-third as large. Both increase with the horizon h, reflecting the joint effect of persistence and forecast length and mirroring the simulation evidence in Figures 1 and 2.[5]

Our expression reduces to Lemma 6 of Kan and Pan (2022) under normality ($\gamma_1 = \gamma_2 = 0$). When v_t is non-normal, the terms in (2.9) enter that depend explicitly on skewness (γ_1) and excess kurtosis (γ_2). These terms introduce further interactions between persistence and higher moments, implying that non-Gaussian predictors may exacerbate higher-order distortions in finite samples.

2.2 Bias in the IM estimator

The conventional IM estimator of β_h (Hodrick, 1992; Lettau and Van Nieuwerburgh, 2008; Xu, 2020) exploits the structural relationship $\beta_h = \beta_1 \sum_{i=0}^{h-1} \rho^i$,

$$\widehat{\beta}_h^{IM} = \widehat{\beta}_1 \sum_{i=0}^{h-1} (\widehat{\rho})^i, \tag{2.10}$$

where $\widehat{\beta}_1$ and $\widehat{\rho}$ are estimated from Eq.(2.1) and Eq.(2.2), respectively.

The following proposition gives the exact bias formula for $\widehat{\beta}_h^{IM}$.

Proposition 2.3. *The bias of $\widehat{\beta}_h^{IM}$ in Eq.(2.10) is*

$$E[\widehat{\beta}_h^{IM}] - \beta_h = \left[\beta_1 + \frac{\sigma_{uv}}{\sigma_v^2}(1-\rho)\right] \sum_{i=0}^{h-1} E\left[\widehat{\rho}^i - \rho^i\right] + \frac{\sigma_{uv}}{\sigma_v^2} E\left[\widehat{\rho}^h - \rho^h\right], \tag{2.11}$$

where ρ^i for $i \geq 1$ is the i-th power of ρ and $\widehat{\rho}$ is the OLS estimator for ρ.

Proposition 2.3 shows that the bias of the IM estimator $\widehat{\beta}_h^{IM}$ is similar to that of the DR estimator $\widehat{\beta}_h$ in Eq.(2.6), with the key difference being that the IM bias depends on powers of a single estimator $\widehat{\rho}$, whereas the DR bias depends on multiple estimators $\widehat{\rho}_i$, $i = 1, \ldots, h$. This distinction arises because the IM estimator constructs β_h from estimates of β_1 and ρ alone, while the DR estimator directly estimates β_h from the long-horizon regression.

Proposition 2.3 yields insights parallel to those in Proposition 2.1. When $h = 1$, the IM bias naturally reduces to the bias of $\widehat{\beta}_1$. For $h \geq 2$, the bias accumulates through the terms $(\widehat{\rho})^i$ across horizons, and is shaped by β_1, ρ, σ_{uv}/σ_v^2, and h.

To our knowledge, this paper is the first to derive an exact analytical bias formula for the IM estimator, $\widehat{\beta}_h^{IM}$, in predictive regressions. Boudoukh et al. (2022) obtained its first-order bias for the IM estimator under $\beta_1 = 0$ based on a bivariate Taylor expansion, without deriving the exact form, the first-order bias when $\beta \neq 0$ or accounting for higher-order bias.

[5] In the Appendix, we also plot the first- and second-order biases computed from the exact bias formula in Proposition 2.1 under $\beta_1 \neq 0$, and obtain qualitatively similar patterns.

Proposition 2.4. *For any $h \geq 1$, the approximate bias of the implied estimator $\widehat{\beta}_h^{IM}$ is*

$$
\begin{aligned}
& E(\widehat{\beta}_h^{IM} - \beta_h) \\
= \quad & -\frac{1}{T}\frac{\sigma_{uv}}{\sigma_v^2}\left[\frac{2(\rho-\rho^{h+1})}{1-\rho} + h\rho^{h-1}(1+\rho)\right] - \frac{1}{T^2}\frac{\sigma_{uv}}{\sigma_v^2}\frac{A_1^{IM}(\rho,h)}{(1-\rho)^2(1+\rho)} && (2.12)\\
& -\frac{\beta_1}{T}\left[\frac{2\rho(1-\rho^h)}{(1-\rho)^2} - \frac{2h\rho^h}{1-\rho} + \frac{h(h-1)(1+\rho)\rho^{h-2}}{2}\right] - \frac{\beta_1}{T^2}\frac{A_2^{IM}(\rho,h)}{(1-\rho)^3(1+\rho)} && (2.13)\\
& -\frac{1}{T^2}\frac{\sigma_{uv}}{\sigma_v^2}\left[\frac{\gamma_2 A_3^{IM}(\rho,h)}{1-\rho} + \frac{\gamma_1^2 A_4^{IM}(\rho,h)}{(1-\rho)(1+\rho+\rho^2)}\right] - \frac{\beta_1}{T^2}\left[\frac{\gamma_2 A_5^{IM}(\rho,h)}{(1-\rho)^2} + \frac{\gamma_1^2 A_6^{IM}(\rho,h)}{(1-\rho)^2(1+\rho+\rho^2)}\right] && (2.14)\\
& +o(T^{-2}),
\end{aligned}
$$

in which $\{A_i^{IM}(\rho,h)\}_i$ are nonlinear functions of ρ and h that are provided in the proof. In addition, for $1 \leq i \leq 6$, $|A_i^{IM}(\rho,h)|$ are bounded uniformly over $\rho \in [0,1]$ and $h \geq 1$.

Analogous to the DR estimator, the bias of the IM estimator can also be expanded to the second order, yielding Propostion 2.4. The leading terms in Eq. (7.14) and (7.14) give the first-order bias, while the remaining terms capture the second-order bias.

As with the DR estimator, the bias of the IM estimator depends on persistence ρ, the covariance-to-variance ratio σ_{uv}/σ_v^2, the prediction horizon h, and sample size T. When v_t is non-normal, additional components enter that depend on skewness γ_1 and excess kurtosis γ_2. Similarly, the second-order terms include factors that scale as $1/(1-\rho)^k$ for $k \geq 1$. In particular, the leading components associated with σ_{uv}/σ_v^2 and β_1 grow at rates $1/(1-\rho)^2$ and $1/(1-\rho)^3$, so the bias can become large as $\rho \to 1$. The coefficient functions $A_i^{IM}(\rho,h)$, although bounded, further amplify these effects when persistence is high. Consequently, higher-order bias can be economically meaningful for persistent predictors.

The right panel of Figure 3 shows that IM estimator's second-order bias is slightly smaller than that of the DR estimator but remains sizable. At $h = 4$, for example, the first-order bias is about 0.8, and the second-order bias is roughly 0.2 – one-fourth as large. Both biases rise with the forecast horizon, consistent with the DR estimator's pattern.

2.3 Discussion

In sum, Propositions 2.2 and 2.4 explain why higher-order bias in both DR and IM estimators can be large, especially with highly persistent predictors and long forecast horizons. Such distortions undermine both estimation accuracy and statistical inference in predictive regressions, highlighting the need for more reliable methods.

A natural way is to correct estimators using explicit second-order bias formulas. In practice, however, this approach performs poorly. Our unreported simulations show that second-order corrections are unstable because of the bias terms scaling with $1/(1-\rho)^k$ ($k = 2, 3$). When $\widehat{\rho}$ is close to one, even small estimation errors can cause these terms to explode, sharply inflating variance and degrading finite-sample performance.[6]

To resolve high-order bias and the associated inference challenges, we introduce in the next section *iBoot*: a unified framework that combines indirect inference and bootstrapping. The name reflects its hybrid nature – combining indirect inference ("i") with bootstrapping ("Boot"). iBoot addresses the higher-order bias issues identified in Section 2 and consists of two key components:

- The iBoot estimator, which corrects both first- and higher-order biases in short- and long-horizon regressions for the DR and IM estimators. (detailed in Section 3.1).

- The iBoot inference method, which enables valid statistical inference following bias correction of the corresponding estimators (presented in Section 3.2).

3 The iBoot framework

3.1 Reducing bias in predictive regressions: the iBoot estimator

The iBoot estimator is obtained through two stages that integrate indirect inference and bootstrapping:

1. **Indirect inference stage**: Compute an initial indirect inference estimator, $\widehat{\beta}_h^{II}$, by matching dynamics between simulated and observed data.

2. **Bootstrap stage**: Leverage $\widehat{\beta}_h^{II}$ to design a bootstrap procedure to further correct multi-horizon estimation biases.

3.1.1 Indirect inference stage

Indirect inference, proposed by Smith (1993) and Gourieroux et al. (1993), is a simulation-based method designed to correct finite-sample bias – an issue that is particularly severe in autoregressive models with persistent predictors. The approach exploits the fact that the OLS estimator in an AR(1) model is systematically biased downward. By simulating data across a range of true ρ values and estimating them via OLS, one can construct a binding function: a mapping between the true parameter ρ and the mean of its OLS estimates across simulations.

Importantly, Andrews (1993) show that the distribution of the OLS estimator $\widehat{\rho}$ in AR(1) models does not depend on other parameters such as constant c or σ_v^2 in Equation 2.2. Hence, the binding function can be generated using simple AR(1) simulations with $c = 0$ and $\sigma_v^2 = 1$. This function captures the systematic bias pattern of $\widehat{\rho}$ across different true ρ values. When applied to real

[6]These simulation results are available upon request.

data, the function allows one to invert this mapping and "recover" the true ρ, yielding the indirect inference estimator.

The left panel of Figure 4 helps illustrate this process. The dash-dot line represents the binding function, while the dotted line marks the true $\rho = 0.95$. The dashed line shows the OLS estimator $\widehat{\rho}$ from real data, whose value (below 0.9) reflects a clear downward bias. By inverting the binding function, we obtain the indirect inference estimator $\widehat{\rho}^{II}$, whose value (above 0.9) lies closer to the true ρ. As shown by Phillips (2012), indirect inference effectively mitigates finite-sample bias in AR(1) settings and performs robustly across different persistence levels. The right panel of Figure 4 confirms this, comparing the scaled bias densities of $\widehat{\rho}$ and $\widehat{\rho}^{II}$, with the latter exhibiting far smaller bias.

To construct the iBoot estimator, we first apply indirect inference to correct the bias in $\widehat{\beta}_h$ for predictive regressions. For the single-horizon case ($h = 1$), Proposition 2.1 quantifies the bias of the OLS estimator $\widehat{\beta}_1$ as:

$$E[\widehat{\beta}_1] - \beta_1 = \phi E\left[\widehat{\rho} - \rho\right], \tag{3.1}$$

where $\widehat{\rho}$ is the OLS estimator of ρ and $\phi = \sigma_{uv}/\sigma_v^2$.

This bias term can be eliminated by the indirect inference estimator $\overline{\overline{\beta}}_1^{II}$, defined as

$$\overline{\overline{\beta}}_1^{II} = \widehat{\beta}_1 - \phi\left(\widehat{\rho} - \widehat{\rho}^{II}\right). \tag{3.2}$$

The following proposition states that $\overline{\overline{\beta}}_1^{II}$ is unbiased for β_1, providing the foundation for iBoot's ability to reduce both first- and higher-order biases in predictive regressions.

Proposition 3.1. *Suppose the binding function for $\widehat{\rho}$ is affine, then the indirect inference estimator $\overline{\overline{\beta}}_1^{II}$ in Eq.(3.2) is unbiased for β_1.*

Since ϕ is unobserved, $\overline{\overline{\beta}}_1^{II}$ is infeasible. We therefore replace ϕ with its unbiased estimator $\widehat{\phi}$ from Amihud and Hurvich (2004), yielding the feasible indirect inference estimator:

$$\widehat{\beta}_1^{II} = \widehat{\beta}_1 - \widehat{\phi}\left(\widehat{\rho} - \widehat{\rho}^{II}\right). \tag{3.3}$$

For the long-horizon case ($h > 1$), we generalize this approach to construct an indirect inference DR estimator for β_h:

$$\widehat{\beta}_h^{II} = \widehat{\beta}_h - \left[\widehat{\beta}_1^{II} + \widehat{\phi}(1 - \widehat{\rho}^{II})\right] \sum_{i=0}^{h-1} \left(\widehat{\rho}_i - \widehat{\rho}_i^{II}\right) - \widehat{\phi}\left(\widehat{\rho}_h - \widehat{\rho}_h^{II}\right), \tag{3.4}$$

where $\widehat{\rho}_i^{II}$ is the indirect inference estimator for ρ_i.

Correcting individual components of $\widehat{\beta}_h$ with unbiased estimators substantially reduces bias, yet $\widehat{\beta}_h^{II}$ retains remaining bias due to inclusion of $\widehat{\phi}$, interactions between $\widehat{\beta}_1^{II}$ or $\widehat{\phi}$ and $\widehat{\rho}_i - \widehat{\rho}_i^{II}$, and between $\widehat{\rho}^{II}$ and $\widehat{\rho}_i - \widehat{\rho}_i^{II}$. Typically, the multiplicative terms are correlated, introducing a non-zero covariance that perpetuates bias.

For the IM estimator, we apply the same principle using Eq. (2.4), yielding:

$$\widehat{\beta}_h^{IM,II} = \widehat{\beta}_1^{II} \sum_{i=0}^{h-1} \left(\widehat{\rho}^{II}\right)^i . \tag{3.5}$$

Similar to $\widehat{\beta}_h^{II}$, the indirect inference corrections in $\widehat{\beta}_1^{II}$ and $\widehat{\rho}^{II}$ alleviate much of the finite-sample bias individually. However, $\widehat{\beta}_h^{IM,II}$ still inherits remaining bias from nonlinear interactions between $\widehat{\beta}_1^{II}$ and $(\widehat{\rho}^{II})^i$.

To eliminate this remaining bias in both the DR and IM estimators, we develop a bootstrap correction stage, detailed next.

3.1.2 Bootstrap stage

To remove the remaining bias in $\widehat{\beta}_h^{II}$ and $\widehat{\beta}_h^{IM,II}$, we introduce a wild bootstrap procedure. Originally proposed by Liu (1988), the wild bootstrap is particularly suited for predictive regressions because it preserves the contemporaneous correlation between return and predictor innovations.

Using the indirect inference estimates $\widehat{\rho}^{II}$ and $\widehat{\beta}_1^{II}$, we generate B bootstrap replications: $\widehat{\beta}_h^{II,1}, \ldots, \widehat{\beta}_h^{II,B}$ for the DR estimator and $\widehat{\beta}_h^{IM,II,1}, \ldots, \widehat{\beta}_h^{IM,II,B}$ for the IM estimator. The remaining bias is then approximated as

$$\frac{1}{B}\sum_{b=1}^{B}\left(\widehat{\beta}_h^{II,b} - \beta_h^*\right) \quad \text{and} \quad \frac{1}{B}\sum_{b=1}^{B}\left(\widehat{\beta}_h^{IM,II,b} - \beta_h^*\right),$$

where $\beta_h^* = \sum_{i=0}^{h-1}\left(\widehat{\rho}^{II}\right)^i \widehat{\beta}_1^{II}$ denotes the "true" parameter value in the bootstrap world. Subtracting these estimated biases from $\widehat{\beta}_h^{II}$ and $\widehat{\beta}_h^{IM,II}$ yields the iBoot estimator for β_h. A step-by-step implementation of this estimator is presented in Appendix 7.9.

A natural question is whether first-order bias-corrected estimators (e.g., Stambaugh, 1999) could be used in place of indirect inference when generating bootstrap samples. They cannot. As the analytical example in Appendix 7.10 shows, first-order corrections leave higher-order bias unaddressed, and this residual distortion is inherited by the bootstrap, leading to contaminated bias estimates.

This observation clarifies the dual role of indirect inference in our framework. First, it delivers bias-reduced estimates of the short-horizon parameters. Second, by anchoring the bootstrap to these improved estimates, it produces samples that more closely approximate the true data-generating process. These more accurate bootstrap samples are then used to remove the remaining long-horizon distortions and to construct reliable confidence intervals and hypothesis tests.

3.2 Inferring return predictability: the iBoot inference method

How can statistical inference for return predictability be conducted free from both first- and higher-order bias? The standard approach typically relies on confidence intervals or tests assuming normality for β_h. However, the OLS estimator for β_h often deviates strongly from normality in small samples, particularly with persistent predictors (Campbell and Yogo, 2006). Even with bias correction, using normal-distribution critical values can produce sizable distortions in test size, undermining the reliability of statistical inference.

To overcome these limitations, our iBoot framework employs indirect inference to estimate ρ and utilizes bootstrap replications, as described in Section 3.1.2, to construct confidence intervals and hypothesis tests for β_h. This approach avoids large-sample asymptotic approximations, accounts for both first- and higher-order biases, and delivers more accurate inference for return predictability.

The iBoot inference procedure consists of two variants: iBoot DR and iBoot IM, tailored respectively to the DR and IM estimators. When $h = 1$, the two methods coincide because the DR and IM estimators are identical. In this case, the bootstrapped indirect inference estimator suffices to approximate the sampling distribution, as bias is already corrected through indirect inference.

When $h > 1$, the procedures diverge due to the distinct bias structures of the DR and IM estimators. For the iBoot DR estimator, inference is based on the iBoot DR point estimate, adjusted for bias using a centered bootstrap. The centered bootstrap is particularly useful when the bootstrap distribution is asymmetric or biased: by recentering around the mean of bootstrap estimates, we approximate the distribution of $\widehat{\beta}_h^{iBootDR}$ and invert it to obtain the confidence interval. Hypothesis testing is then straightforward – reject $\mathbb{H}_0 : \beta_h = 0$ if zero lies outside the corresponding confidence interval.

For the iBoot IM estimator, inference is based directly on the bootstrapped indirect-inference IM estimates. Although the procedure is related to Xu (2020), it differs in three important respects. First, we replace their first-order bias-corrected estimator of β_1 with the indirect-inference estimator $\widehat{\beta}1^{II}$. As shown in Section 3.1.2, first-order corrections tend to understate the bias in $\widehat{\beta}1$, whereas indirect inference more effectively accounts for higher-order distortions. Second, Xu (2020) correct only the bias in β_1, leaving the estimators of α_1, c, and ρ unadjusted when generating bootstrap samples. In contrast, iBoot IM bias-corrects all key parameters prior to bootstrapping, leading to bootstrap samples that better approximate the underlying data-generating process. As a result, iBoot IM achieves improved size control and higher power in our simulations (Section 4). Third, their bootstrap is based on the t-ratio, which limits its use for interval estimation. Our approach instead bootstraps the estimator itself, allowing for the construction of both confidence intervals and hypothesis tests within a unified framework.

Taken together, these differences make iBoot IM a more comprehensive and practically useful inference procedure for IM-based return predictability analysis. A detailed step-by-step algorithm is provided in Appendix 7.11.

3.3 Discussion

Overall, the iBoot framework provides a unified solution to longstanding challenges in predictive regressions. By combining indirect inference with bootstrapping, iBoot corrects both first- and higher-order biases that undermine conventional estimation and inference. The framework applies seamlessly to both DR and IM estimators, allowing researchers to work with their preferred specification while benefiting from robust bias correction and reliable inference.

Although the second-order bias expansions in Propositions 2.2 and 2.4 are derived under stationarity ($\rho < 1$) to illustrate the source and magnitude of higher-order bias, the iBoot framework itself remains valid well beyond this setting. In particular, iBoot continues to perform reliably when predictors are highly persistent, local to unity, or even exhibit a unit root. This robustness stems from two features. First, indirect inference delivers accurate bias correction for $\widehat{\rho}$ even in near-unit-root or unit-root environments, as shown by Phillips (2012). Second, the exact bias expressions in Propositions 2.1 and 2.3 hold for all values of ρ. Together, these properties ensure that iBoot can effectively mitigate bias in both $\widehat{\beta}_h$ and $\widehat{\beta}_h^{IM}$ and deliver credible inference across the full range of empirically relevant persistence levels.

4 Monte Carlo Simulations

We conduct Monte Carlo simulations to evaluate the finite-sample performance of our iBoot estimation and inference methods relative to alternatives. The data-generating process (DGP) follows the predictive regression model in (2.1) with $\alpha_1 = 0$ and $\beta_1 = 0$. The predictor x_t is generated from (2.2) with $c = 0$ and persistence levels $\rho \in \{0.5, 0.7, 0.8, 0.82, 0.87, 0.9, 0.92, 0.95, 0.97, 0.99\}$. The error terms (u_t, v_t) are drawn from a bivariate normal distribution with covariance matrix $\Sigma = \begin{bmatrix} 0.04 & -0.018 \\ -0.018 & 0.01 \end{bmatrix}$, implying $\sigma_{uv}/\sigma_v^2 = -1.8$. We consider sample sizes $T \in \{30, 60, 100, 300, 600\}$, where $T = 30, 60, 100$ reflect typical annual samples and $T = 300, 600$ correspond to quarterly and monthly samples. Each design uses $B = 10{,}000$ bootstrap replications and $S = 10{,}000$ Monte Carlo simulations.

4.1 Bias comparison against alternatives

Here we evaluate the finite-sample bias of our iBoot estimators relative to existing bias-reduction methods. The central finding is that iBoot delivers economically and statistically meaningful improvements over first-order bias-corrected estimators, especially when predictors are highly persistent and forecast horizons are long.

We begin by comparing the bias of the iBoot DR estimator ($\widehat{\beta}_h^{iBootDR}$) with two benchmark DR estimators: the OLS estimator ($\widehat{\beta}_h$) and the bias-corrected estimator ($\widehat{\beta}_h^{BIR}$) proposed by Stambaugh (1999) for $h = 1$ and Boudoukh et al. (2022) for $h > 1$.

Figures 5 and 6 report biases across different values of ρ, h, and T. Consistent with existing evidence, OLS exhibits substantial bias that intensifies as persistence rises. More importantly, although the BIR estimator reduces this bias, it remains noticeably biased when ρ is close to one or when h is large. The iBoot DR (red solid line) estimator, by contrast, consistently produces the smallest bias, with particularly pronounced gains relative to BIR under high persistence and long horizons. For example, when $T = 30$ and $\rho = 0.9$, the BIR bias at $h = 1$ is about 0.095, whereas the iBoot bias is around 0.02 and close to zero. Even in larger samples, improvements remain economically meaningful: at $T = 600$ and $\rho = 0.99$, the BIR bias reaches roughly 0.15 at $h = 48$, while the iBoot bias is negligible. These results indicate that correcting only first-order bias is insufficient in persistent environments, whereas iBoot effectively addresses the higher-order components documented earlier.

We next examine the IM estimator. For $h > 1$, we compare the iBoot IM estimator, $\widehat{\beta}_h^{iBootIM}$, with the conventional IM estimator, $\widehat{\beta}_h^{IM}$, and its first-order bias-corrected version, $\widehat{\beta}_h^{FOIM}$. Figure 7 shows a similar pattern. Although FO IM reduces bias relative to the conventional IM estimator, it leaves non-negligible distortions under strong persistence and long horizons. The advantage of iBoot IM (red solid line) becomes particularly clear in such cases. For instance, at $T = 60$ and $\rho = 0.95$, the FO IM bias at $h = 3$ is about 0.1, whereas the iBoot IM bias is nearly zero. The gap persists in larger samples: at $T = 600$ and $\rho = 0.99$, the FO IM bias is about 0.1 at $h = 48$, while the iBoot IM bias remains close to zero. Thus, as with the DR estimator, higher-order bias materially affects IM estimation under high persistence, and iBoot substantially mitigates these distortions relative to first-order corrections.

Tables 7–12 in Appendix summarize the numerical bias values for all estimators across different settings, corroborating the graphical findings. Overall, both iBoot DR and iBoot IM estimators substantially mitigate finite-sample bias across all configurations, outperforming existing methods – especially under high persistence, long horizons, and limited sample sizes.

4.2 Size and power comparison against alternatives

We evaluate the size and power of the proposed *iBoot* inference methods against four widely used alternatives for testing return predictability under $\mathbb{H}_0 : \beta_h = 0$ versus $\mathbb{H}_1 : \beta_h \neq 0$. Specifically, we compare:

1. NW: OLS estimator with Newey and West (1987) standard errors;

2. HOD: OLS estimator with Hodrick (1992) standard errors;

3. BIR: Bias-corrected estimator with analytical AR(1) standard errors from Boudoukh et al. (2022);

4. IM: Xu (2020)'s implied method using Stambaugh (1999)'s bias-corrected estimator and bootstrapped t-ratios;

5. iBoot DR: iBoot inference method for the DR estimator proposed in this paper;

6. iBoot IM: iBoot inference method for the IM estimator proposed in this paper.

NW and HOD are popular methods in predictive regressions for confidence intervals and hypothesis testing. NW accounts for autocorrelation and heteroskedasticity in standard errors, while HOD adjusts standard errors for overlapping data in long-horizon regressions. BIR corrects first-order bias and applies analytical AR(1) standard errors. IM uses a first-order bias-corrected estimator and bootstraps t-ratios for inference. Our proposed iBoot DR and iBoot IM methods combine indirect inference with bootstrapping to correct both first- and higher-order biases. Among these, NW, HOD, BIR, and iBoot DR apply to direct regressions, while IM and iBoot IM apply to implied regressions.

Figure 8 compares empirical sizes at $T = 30$ ($h = 1, 2$), $T = 60$ ($h = 1, 3$), $T = 100$ ($h = 1, 5$), and $T = 300$ ($h = 1, 16$). iBoot IM (solid line) consistently maintains the nominal 5% size across all scenarios – even with small samples ($T = 30$), high persistence ($\rho = 0.99$), and long horizons ($h = 48$). iBoot DR (dash-dotted line) also performs well, with minor deviations only when ρ approaches 1 at extreme horizons (e.g., $h = 5$ when $T = 100$). These deviations arise because larger horizons amplify noise in multi-period returns, increasing estimator bias, leading to challenges for correct bootstrap approximation in small samples. When the sample size grows (e.g., $T = 600$, Figure 9), iBoot DR also achieves near-perfect size control across all horizons.

While also affected by the bias amplification at long horizons, iBoot IM is less sensitive to the horizon even in small sample because it exploits the structural relationship between β_h, β_1, and ρ. As a result, its bootstrap distributions mimic those of short horizons, enabling iBoot IM to retain exact size control even for large h.

All alternative methods exhibit significantly larger size distortions. At $T = 30$ and $h = 1$, NW, HOD, BIR, and IM all exceed 10% size across all ρ, with NW and HOD reaching nearly 30% as $\rho \to 1$. Even at $T = 100$, distortions persist: all methods remain above 10% for $\rho > 0.9$, and NW and HOD exceed 20% near $\rho = 1$. The problem worsens at longer horizons – at $T = 30$, $h = 2$, and $\rho = 0.99$, NW exceeds 40%, HOD about 30%, BIR around 20%.

These distortions stem from two factors: (i) sizable estimator bias, especially under high persistence and long horizons, and (ii) poor normal approximations in small samples with persistent predictors (Campbell and Yogo, 2006). To disentangle their relative contributions, our unreportd simulations replace the original estimators with the iBoot DR estimator while retaining normal critical values. Although this reduces size distortions, rejection rates remain around 10% at $T = 30$ and $\rho = 0.99$. When we further replace normal critical values with iBoot bootstrap values, size falls to the nominal 5%. This comparison underscores that both bias correction and improved distribution approximations are necessary for reliable inference. Although IM employs bootstrapping, its reliance on a first-order bias-corrected estimator leaves higher-order bias unaddressed, resulting in persistent size distortions.

Larger samples improve all methods, especially at short horizons. As shown in Figure 9, at $T = 600$ and $h = 1$, NW, HOD, BIR, and IM approach the 5% nominal size for most ρ, but all exceed 10% as $\rho \to 1$. At $h = 48$, distortions reappear: NW rises above 30%, HOD around 15%, and BIR and IM near 10%. In contrast, iBoot DR and iBoot IM maintain 5% size across all ρ and horizons.

Figure 10 highlights the power advantages of iBoot relative to existing methods. Power is defined as the rejection rate of $H_0 : \beta_h = 0$ when false (with $\beta_h = \beta_1(1 - \rho^h)/(1 - \rho)$ and $\beta_1 = 0.16$). To ensure fair comparisons, we focus on the large-sample cases ($T = 300, 600$) and report size-adjusted power.[7] This guarantees that differences in power reflect genuine sensitivity to alternatives rather than size distortions.

Across all settings, iBoot IM achieves the highest power while maintaining exact size control. IM ranks second, consistent with Xu (2020), who show that the standard t-statistic for β_h weakens as $h(1 - \rho) \to \infty$ because multi-period return aggregation amplifies noise relative to the predictor's signal. The IM-based methods overcome this by exploiting the structural link between β_h, β_1, and ρ, preserving power even at long horizons. This explains why iBoot IM and IM outperform other approaches for large h.

Among direct regression methods, iBoot DR is the most powerful, followed by NW and BIR. At $h = 1$, iBoot DR's advantage is modest but widens sharply with horizon length. For example, at $T = 600$ and $h = 48$, iBoot DR attains around 60% power for $\rho < 0.88$ and nearly 100% for $\rho > 0.92$, whereas NW and BIR reach only 40% and 60%, respectively.

Overall, the results highlight the superior performance of the iBoot framework. It uniquely

[7] For each method, we first obtain an empirical critical value under H_0 that controls the size at 5%, and then compute the corresponding size-adjusted power under H_1.

combines exact size control with strong power across almost all persistence levels, horizons, and sample sizes – setting a new benchmark for reliable inference in predictive regressions.

4.3 Summary

Our Monte Carlo simulations demonstrate that the iBoot estimators consistently outperform alternative DR and IM estimators in bias reduction across various sample sizes, persistence levels, and prediction horizons. The iBoot inference methods deliver more precise exact size control while maintaining high power across nearly all scenarios. In contrast, NW, HOD, BIR, and IM exhibit substantial size distortions, particularly in small samples, with persistent predictors, or at long horizons.

Appendix 7.12.3 reports confidence interval coverage across methods. The results parallel the size findings: iBoot IM achieves coverage rates nearly identical to the nominal 95% level across all designs, while iBoot DR performs closely behind, with only slight undercoverage at extreme horizons. In contrast, NW, HOD, and BIR display marked undercoverage under high persistence and long horizons.

For robustness, Appendix 7.13 examines the case $\beta_1 \neq 0$ and reach an even stronger conclusion: iBoot maintains minimal bias and accurate inference, with performance advantages that become more pronounced when $\beta_1 = -0.16$. Appendices 7.14 and 7.15 further consider non-normal and GARCH(1,1) errors respectively and show that iBoot's superior bias correction and inference accuracy remain intact under these alternative data-generating processes.

Overall, these findings underscore two key insights: (1) higher-order bias is sizable and materially distorts inference in predictive regressions; and (2) by jointly addressing both first- and higher-order bias, the iBoot DR and iBoot IM methods provide a reliable and robust solution for credible estimation and inference in this setting.

5 Empirical Results

Having demonstrated the magnitude of higher-order bias and the effectiveness of the iBoot framework in simulations, we now assess its practical value using real-world data.

We present DR and IM results separately in Sections 5.1 and 5.2 because they target different estimands and may lead to different conclusions about return predictability. For each approach, we begin with widely studied predictors and then extend the analysis to the full predictor set in Goyal et al. (2024).

The predictive regressions are estimated at monthly, quarterly, and annual frequencies with horizons ranging one to forty-eight months, one to sixteen quarters, and from one to five years, respectively. Following Boudoukh et al. (2022), the sample begins in 1968. Stock returns and predictors are obtained from Goyal et al. (2024). For DR, we compare OLS with the Newey-West (NW) standard errors, the bias-corrected estimator from Boudoukh et al. (2022) (BIR), and iBoot DR. For IM, we compare the estimator of Xu (2020) and iBoot IM.

5.1 Empirical results for DR estimators

5.1.1 Results for three well-known predictors

For the DR estimators, we start by focusing on three widely studied predictors – dividend yield (dy), forward implied variance ($impvar$), and government investment ($govik$). These variables are either highly persistent or exhibit large covariance-to-variance ratios, making them especially susceptible to higher-order bias. They also exhibit some of the largest shifts in significance under iBoot DR relative to OLS and BIR. As such, they provide an ideal setting to illustrate both the statistical and economic implications of correcting higher-order bias.

Tables 1-3 summarize the results for each predictor across frequencies. Each table lists the sample period, horizon (h), sample size (T), estimated persistence ($\hat{\rho}_h$), covariance-to-variance ratio ($\hat{\sigma}_{uv}/\hat{\sigma}_v^2$), beta estimates for the three DR estimators, and the percentage p-value changes between iBoot DR and the other two methods. Asterisks denote significance at the 10%, 5%, and 1% levels.

Dividend yield (dy)

The dividend yield (dy) is a common predictor of stock returns, with early evidence documented by Shiller (1984), who finds significant positive short-horizon predictability at the annual frequency over 1872-1983. In our sample, all estimated coefficients on dy remain positive, but statistical significance varies sharply across methods, as shown in Table 1.

At the annual frequency, OLS and BIR yield insignificant estimates at all horizons, whereas iBoot DR detects significant positive predictability at short horizons, aligning with the original findings of Shiller (1984). At the monthly frequency, OLS and BIR remain insignificant throughout, while iBoot DR is significant at every horizon, providing evidence of both short- and long-horizon return predictability. At the quarterly frequency, none of the methods achieves conventional significance; however, iBoot DR consistently delivers substantially smaller p-values than OLS and BIR, with reductions of roughly 10%-50%.

Taken together, these results show that iBoot DR uncovers stronger and more consistent evidence of return predictability from dy than conventional approaches.

Forward implied variance ($impvar$)

The forward implied variance ($impvar$) is another well-known return predictor introduced by Bakshi et al. (2011), who document significant positive predictability at the one-month horizon using data from 1998M09-2008M09. Using our longer sample, we find that the three methods lead to sharply different conclusions about *impvar*'s predictive power, as shown in Table 2.

At monthly and quarterly frequencies, all methods yield positive but insignificant estimates at short horizons. As the horizon increases, estimates turn negative, but inference diverges: OLS remains insignificant throughout, BIR exhibits scattered significance, while iBoot DR becomes consistently significant at longer horizons. Even when both iBoot DR and BIR reject the null, their magnitudes differ substantially. For example, at $h = 48$ (monthly), both are significant at the 5% level, yet the iBoot DR estimate (-36.44) is about 20% larger in absolute value than BIR's.

Table 31 in the Appendix reports results for the shorter sample used by Bakshi et al. (2011). A similar pattern of method-dependent significance emerges. At short horizons, all methods produce positive estimates, consistent with the results in Table 2.[8] At longer horizons ($h = 16, 20$), iBoot DR begins to detect significant negative predictability, while BIR and OLS remain insignificant. This pattern suggests that higher-order bias may have masked the emergence of negative predictability at longer horizons in earlier studies. Overall, once higher-order bias is corrected, iBoot provides more consistent and robust evidence of *impvar*'s long-run return predictability than existing methods.

Government investment (*govik*)

Government investment (*govik*) has been widely studied as a return predictor since Belo and Yu (2013), who document significant positive short-horizon predictability at the quarterly frequency over 1947Q2-2010Q4. Using an extended sample from 1968 to 2023, we find that iBoot frequently revises the conclusions about *govik*'s predictive power made by OLS and BIR, as shown in Table 3.

At the quarterly frequency, BIR yield insignificant estimates at all horizons, eliminating the significant negative predictability detected by OLS at long horizons (h = 8-16). In contrast, iBoot DR uncovers significant negative predictability at most horizons (h = 2-16), overturning the conclusions of OLS at short horizons and BIR throughout.

Monthly data for *govik* are unavailable, but the annual results closely mirror the quarterly findings. BIR again produces insignificant estimates, removing the significant negative predictability identified by OLS at long horizons. By contrast, iBoot DR detects significant negative predictability at most horizons, reversing the conclusions of both OLS and BIR.

Overall, iBoot DR reveals a more pervasive and robust negative relationship between *govik* and stock returns than existing methods. This contrasts with the positive predictability reported by Belo and Yu (2013). However, when the sample is restricted to their study period, all three methods yield significant positive estimates at short horizons, consistent with their findings (see Table 32 in

[8] Bakshi et al. (2011) report significant positive predictability for *impvar* by including control variables in the short-horizon regression. Without controls, however, Goyal et al. (2024) find insignificant estimates, consistent with our results.

the Appendix). This pattern suggests that the predictive relationship between *govik* and returns may evolve over time and that higher-order bias can obscure this evolution. Consequently, iBoot DR provides a more nuanced and reliable assessment of *govik*'s return predictability.

Discussions

Our empirical analysis shows that the iBoot DR frequently leads to different conclusions about return predictability than OLS and BIR. These differences stem from two tightly linked components of the iBoot framework – higher-order bias correction and robust inference – since bootstrap samples used in inference are generated from bias-corrected estimates.

To illustrate their respective roles, we plot the empirical densities of the OLS, BIR, and iBoot DR estimators at selected horizons and frequencies where differences in magnitude and significance are most pronounced. Figure 11 displays fitted normal densities for OLS and BIR using their respective point estimates and standard errors (Newey-West and BIR) and the bootstrap density for iBoot DR.

For monthly dy at $h = 1$, all three densities are centered at nearly the same value (consistent with Table 1), indicating mild bias differences. iBoot DR's distribution, however, is much concentrated, reflecting smaller standard errors. Here, the gain in significance under iBoot DR is driven primarily by improved inference rather than bias correction.

For quarterly *impvar* at $h = 12$, the BIR's density shifts slightly to the left relative to OLS, indicating some bias correction, but the iBoot DR density shifts much further left, indicating substantial higher-order bias correction. This shift alone explains the change in significance: replacing the OLS or BIR point estimate with the iBoot DR estimate – while holding their standard errors fixed – yields t-statistics around -1.75, significant at the 10% level. In this case, bias correction is the primary factor driving the altered conclusion.

For annual *govik* at $h = 1$, the iBoot DR density exhibits only a modest shift relative to OLS and BIR, suggesting limited bias correction. Instead, the key difference is dispersion: the iBoot DR distribution is far tighter, implying substantially smaller standard errors. This improvement in precision primarily drives the emergence of statistical significance.

Overall, the iBoot framework delivers materially different empirical conclusions from traditional methods. Both higher-order bias correction and robust inference are essential, with their relative importance varying across predictors and horizons. Together, they help explain why earlier studies may have overstated – or understated – the economic significance of return predictability.

5.1.2 DR Results for all predictors in Goyal et al. (2024)

We extend our analysis to the full set of predictors in Goyal et al. (2024) – 47 annual, 40 quarterly, and 35 monthly – evaluating iBoot's performance across a broader universe. Using the same horizons (annual: $h = 1 - 5$; quarterly: $h = 1 - 16$; monthly: $h = 1 - 48$), we record for each predictor whether iBoot DR alters its statistical significance relative to OLS and BIR.

We classify significance changes into six categories: (1) OLS significant $\rightarrow$ iBoot DR insignificant (green); (2) OLS insignificant $\rightarrow$ iBoot DR significant (yellow); (3) BIR significant $\rightarrow$ iBoot DR insignificant (purple); (4) BIR insignificant $\rightarrow$ iBoot DR significant (red); (5) both OLS and BIR significant $\rightarrow$ iBoot DR insignificant (blue); (6) both OLS and BIR insignificant $\rightarrow$ iBoot DR significant (orange).

Figure 12(a) summarizes significance shifts in the short horizon ($h = 1$). Relative to OLS, six monthly and five annual predictors gain significance under iBoot, while two (one monthly, one annual) lose significance. Relative to BIR, five monthly, one quarterly, and seven annual predictors gain significance, while three monthly, two quarterly, and one annual predictor lose significance. When measured against both OLS and BIR, iBoot identifies five monthly and four annual significant predictors and eliminates significance for one monthly predictor.

Figure 12(b) reports corresponding results at long horizons ($h > 1$). Relative to OLS, iBoot increases significance for eleven monthly, five quarterly, and three annual predictors, while reducing significance for thirteen monthly, fourteen quarterly, and eight annual predictors. Compared with BIR, iBoot increases significance for seven monthly, five quarterly, and seven annual predictors but reduces significance for twelve monthly and ten quarterly predictors. Against both OLS and BIR, six monthly, three quarterly, and three annual predictors gain significance, whereas six monthly and six quarterly predictors lose significance.

We also examine two special cases. First, among predictors for which BIR removes significance relative to OLS, iBoot DR restores significance for one quarterly and three annual predictors at the short horizon, and for three monthly, four quarterly, and five annual predictors at longer horizons. Second, we identify predictors for which iBoot DR reverses the sign of the estimated coefficient relative to BIR. Such reversals occur for eight monthly, ten quarterly, and two annual predictors, although most remain statistically insignificant under both methods.

Overall, iBoot frequently removes overstated significance in both OLS and BIR – consistent with its well-controlled size in simulations – while also uncovering predictive relationships that conventional methods fail to detect. These results highlight the importance of correcting higher-order bias and using robust inference when assessing return predictability in the direct regression framework.

5.2 Empirical results for IM estimators

5.2.1 Results for three well-known predictors

We next turn to implication-based (IM) estimators, again first focusing on three persistent predictors: the Treasury bill rate (*tbl*), cross-section-based tail risk (*tail*), and equity issuing activity (*eqis*). Because the IM approach of Xu (2020) already relies on bootstrap inference, differences between IM and iBoot IM arise mainly from improved bias correction – and the more accurate sampling distributions it induces – rather than from changes in the inference procedure itself. These predictors, therefore, offer a clean setting for isolating the role of higher-order bias correction within the IM framework.

As in Section 5.1, Tables 4-6 report the sample period, horizon (h), sample size (T), estimated persistence ($\hat{\rho}_h$), covariance-to-variance ratio ($\hat{\sigma}_{uv}/\hat{\sigma}_v^2$), and estimated coefficients for IM and iBoot IM, with significance denoted by conventional asterisks.

Treasury bill rate (*tbl*)

The treasury bill rate (tbl) is another commonly used predictor of stock returns proposed by Fama and Schwert (1977), who used direct regressions to document significant predictability at short horizons over 1953-1975. Here, using the IM methods over our extended sample, we find that iBoot IM substantially revises both the magnitude and significance of IM estimates for *tbl*, as shown in Table 4.

At the monthly frequency, IM reports 10% significance at all horizons, whereas all iBoot IM estimates are insignificant. The two methods also differ markedly in magnitude. For example, at $h = 48$, the IM estimate (-3.506) is about 35% smaller in absolute value than the iBoot IM estimate (-5.365). The combination of smaller magnitudes and higher significance under IM suggests that its bootstrap distribution is overly concentrated, likely reflecting uncorrected higher-order bias that distorts the sampling distribution.

At quarterly and annual frequencies, both IM and iBoot IM yield insignificant estimates throughout. Thus, iBoot IM uncovers a more pervasive and robust insignificant relationship between *tbl* and stock returns than IM, calling into question of significant predictability of the *tbl* predictor.

Cross-section based tail risk (*tail*)

The cross-section based tail risk (*tail*) is a more recent predictor proposed by Kelly and Jiang (2014), who document significant positive return predictability at the monthly frequency over 1963-2010 across one-month, one-year, and three-year horizons. Our sample extends through 2023 and allows us to examine predictability at monthly, quarterly, and annual frequencies, which we report in Table 5.

At the monthly and annual frequencies, both IM and iBoot IM yield positive and statistically significant estimates at all horizons. The two methods produce similar magnitudes, though iBoot IM consistently delivers smaller p-values, indicating stronger inference.

At the quarterly frequency, however, the results diverge sharply. While IM finds no significant predictability at any horizon, iBoot IM detects significant positive effects throughout. The point estimates are similar – for example, at $h = 16$, IM yields 1.212 versus 1.231 for iBoot IM – so the difference in inference cannot be attributed to magnitude alone. Instead, it reflects differences in bootstrap accuracy. The IM bootstrap samples are generated using a first-order bias-corrected β but uncorrected estimates of ρ, and the intercepts c and α. In contrast, iBoot IM applies higher-order bias correction to β, ρ, and the associated c and α before bootstrapping, producing a more accurate sampling distribution. This improvement enhances statistical power and allows iBoot IM to uncover quarterly predictability that IM fails to detect.

Equity issuing activity (*eqis*)

Equity issuing activity (*eqis*) is another well-known predictor introduced by Baker and Wurgler (2000), who document significant negative return predictability at the annual frequency over 1928-1997. Using a more recent sample from 1968-2023, we reexamine this relation at the annual frequency as well, as quarterly and monthly data for *eqis* are unavailable.

Table 6 reports the results. Both IM and iBoot IM produce negative estimates across all horizons, consistent with prior evidence. However, their inference differs substantially. IM detects significance only at short horizons ($h = 1, 2$), whereas iBoot IM finds statistically significant predictability at all horizons. Differences in magnitude also emerge at longer horizons: for instance, at $h = 5$, the IM estimate (-1.267) is about 10% smaller in absolute value than the iBoot IM estimate (-1.414). Together, these results suggest that uncorrected higher-order bias in IM attenuates both magnitudes and statistical significance, while iBoot IM restores stronger and more robust evidence of return predictability.

Discussions

As in Section 5.1, we examine empirical densities for IM and iBoot IM at horizons and frequencies where differences in magnitude and significance are most pronounced for these predictors.

Figure 13 plots the distributions for monthly *tbl* at $h = 1$, quarterly *tail* at $h = 16$, and annual *eqis* at $h = 5$. In each case, the iBoot IM distribution shifts noticeably relative to IM, providing clear evidence of higher-order bias correction. Because both approaches rely on bootstrap inference, the resulting changes in statistical significance primarily reflect improved bias correction and the more accurate bootstrap sampling generated by iBoot IM's bias-corrected parameter estimates.

5.2.2 IM Results for all predictors in Goyal et al. (2024)

As with the DR estimators, we extend our IM analysis to the full set of predictors in Goyal et al. (2024) – 47 annual, 40 quarterly, and 35 monthly – to assess iBoot IM across a broad universe. Using the same horizons (annual: $h = 1$-5; quarterly: $h = 1$-16; monthly: $h = 1$-48), we document for each predictor whether iBoot IM changes its statistical significance relative to IM.

Here, significance shifts fall into two categories: (1) IM insignificant $\rightarrow$ iBoot IM significant (blue); and (2) IM significant $\rightarrow$ iBoot IM insignificant (orange);

Figure 14(a) displays results for the short horizon ($h = 1$). Five monthly, one quarterly, and five annual predictors gain significance under iBoot, while two monthly and one quarterly predictors lose significance.

Figure 14(b) summarizes the long-horizon results ($h > 1$). Here, five monthly, two quarterly, and six annual predictors gain significance under iBoot, whereas two monthly, one quarterly, and one annual predictor lose significance.

Overall, the full-sample IM results echo the patterns seen for DR: iBoot IM frequently corrects overstated significance (reflecting its superior size control) while also uncovering understated significance (reflecting its higher power). These adjustments further underscore the role of higher-order bias and robust inference in shaping empirical conclusions about return predictability in the implication based regression framework.

6 Conclusion

This paper demonstrates that conventional predictive regression methods systematically underestimate bias and misstate inference due to neglected higher-order effects. We analytically derive the magnitude and structure of these biases for both direct regression (DR) and implication (IM) estimators, showing that they can be large for persistent predictors and long horizons.

To address this problem, we introduce iBoot, a unified estimation and inference framework that integrates indirect inference with bootstrap methods. iBoot is the first approach to jointly correct both first- and higher-order biases across horizons. Extensive simulations show that iBoot achieves minimal bias, precise coverage, and exact size control while maintaining high power, outperforming OLS, BIR, and IM. Empirically, applying iBoot to the comprehensive predictor set of Goyal et al. (2024) reveals that higher-order bias corrections materially alter the evidence on return predictability and iBoot frequently overturns the conclusions made by existing methods.

By quantifying higher-order bias and providing a complete toolkit to correct it and make sound inference, this paper establishes a new empirical foundation of predictive regression analysis, offering researchers and practitioners a framework for credible estimation and inference.

References

Amihud, Y. and C. M. Hurvich (2004). Predictive regressions: A reduced-bias estimation method. *Journal of Financial and Quantitative Analysis 39*(4), 813–841.

Andrews, D. W. (1993). Exactly median-unbiased estimation of first order autoregressive/unit root models. *Econometrica: Journal of the Econometric Society*, 139–165.

Baker, M. and J. Wurgler (2000). The equity share in new issues and aggregate stock returns. *the Journal of Finance 55*(5), 2219–2257.

Bakshi, G., G. Panayotov, and G. Skoulakis (2011). Improving the predictability of real economic activity and asset returns with forward variances inferred from option portfolios. *Journal of Financial Economics 100*(3), 475–495.

Bao, Y. and A. Ullah (2010). Expectation of quadratic forms in normal and nonnormal variables with applications. *Journal of Statistical Planning and Inference 140*(5), 1193–1205.

Belo, F. and J. Yu (2013). Government investment and the stock market. *Journal of Monetary Economics 60*(3), 325–339.

Boudouk, J. and M. Richardson (1994). The statistics of long-horizon regressions revisited 1. *Mathematical Finance 4*(2), 103–119.

Boudoukh, J., R. Israel, and M. Richardson (2022). Biases in long-horizon predictive regressions. *Journal of Financial Economics 145*(3), 937–969.

Campbell, J. Y. and R. J. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies 1*(3), 195–228.

Campbell, J. Y. and M. Yogo (2006). Efficient tests of stock return predictability. *Journal of Financial Economics 81*(1), 27–60.

Cochrane, J. H. (2008). The dog that did not bark: A defense of return predictability. *The Review of Financial Studies 21*(4), 1533–1575.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics 7*(1), 1–26.

Fama, E. F. and G. W. Schwert (1977). Asset returns and inflation. *Journal of financial economics 5*(2), 115–146.

Gourieroux, C., A. Monfort, and E. Renault (1993). Indirect inference. *Journal of Applied Econometrics 8*(S1), S85–S118.

Gouriéroux, C., P. C. Phillips, and J. Yu (2010). Indirect inference for dynamic panel models. *Journal of Econometrics 157*(1), 68–77.

Gouriéroux, C., E. Renault, and N. Touzi (2000). 13 calibration by simulation for small sample bias correction. *Simulation-based inference in econometrics: Methods and applications*, 328.

Goyal, A. and I. Welch (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies 21*(4), 1455–1508.

Goyal, A., I. Welch, and A. Zafirov (2024). A comprehensive 2022 look at the empirical performance of equity premium prediction. *The Review of Financial Studies 37*(11), 3490–3557.

Hjalmarsson, E. (2011). New methods for inference in long-horizon regressions. *Journal of Financial and Quantitative Analysis 46*(3), 815–839.

Hodrick, R. J. (1992). Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *The Review of Financial Studies 5*(3), 357–386.

Jiang, L., X. Wang, and J. Yu (2018). New distribution theory for the estimation of structural break point in mean. *Journal of Econometrics 205*(1), 156–176.

Kan, R. and J. Pan (2022). Finite sample analysis of predictive regressions with long-horizon returns. *Rotman School of Management Working Paper* (3790052).

Kelly, B. and H. Jiang (2014). Tail risk and asset prices. *The Review of Financial Studies 27*(10), 2841–2871.

Kendall, M. G. (1954). Note on bias in the estimation of autocorrelation. *Biometrika 41*(3-4), 403–404.

Kiviet, J. F. and G. D. Phillips (2012). Higher-order asymptotic expansions of the least-squares estimation bias in first-order dynamic regression models. *Computational Statistics & Data Analysis 56*(11), 3705–3729.

Kostakis, A., T. Magdalinos, and M. P. Stamatogiannis (2015). Robust econometric inference for stock return predictability. *The Review of Financial Studies 28*(5), 1506–1553.

Lettau, M. and S. Van Nieuwerburgh (2008). Reconciling the return predictability evidence: The review of financial studies: Reconciling the return predictability evidence. *The Review of Financial Studies 21*(4), 1607–1652.

Liu, R. Y. (1988). Bootstrap procedures under some non-iid models. *The annals of statistics*, 1696–1708.

Marriott, F. and J. Pope (1954). Bias in the estimation of autocorrelations. *Biometrika 41*(3/4), 390–402.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix. *Econometrica 55*, 703–708.

Phillips, P. C. (2012). Folklore theorems, implicit maps, and indirect inference. *Econometrica 80*(1), 425–454.

Phillips, P. C. (2015). Halbert white jr. memorial jfec lecture: Pitfalls and possibilities in predictive regression. *Journal of Financial Econometrics 13*(3), 521–555.

Phillips, P. C. and J. Yu (2009). Simulation-based estimation of contingent-claims prices. *The Review of Financial Studies 22*(9), 3669–3705.

Shiller, R. J. (1984). Stock prices and social dynamics. *Brookings papers on economic activity 1984*(2), 457–510.

Smith, A. A. (1993). Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics 8*(S1), S63–S84.

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics 54*(3), 375–421.

Ullah, A. (2004). *Finite sample econometrics*. OUP Oxford.

Valkanov, R. (2003). Long-horizon regressions: theoretical results and applications. *Journal of Financial Economics 68*(2), 201–232.

Xu, K.-L. (2020). Testing for multiple-horizon predictability: Direct regression based versus implication based. *The Review of Financial Studies 33*(9), 4403–4443.

Yu, J. (2012). Bias in the estimation of the mean reversion parameter in continuous time models. *Journal of Econometrics 169*(1), 114–122.

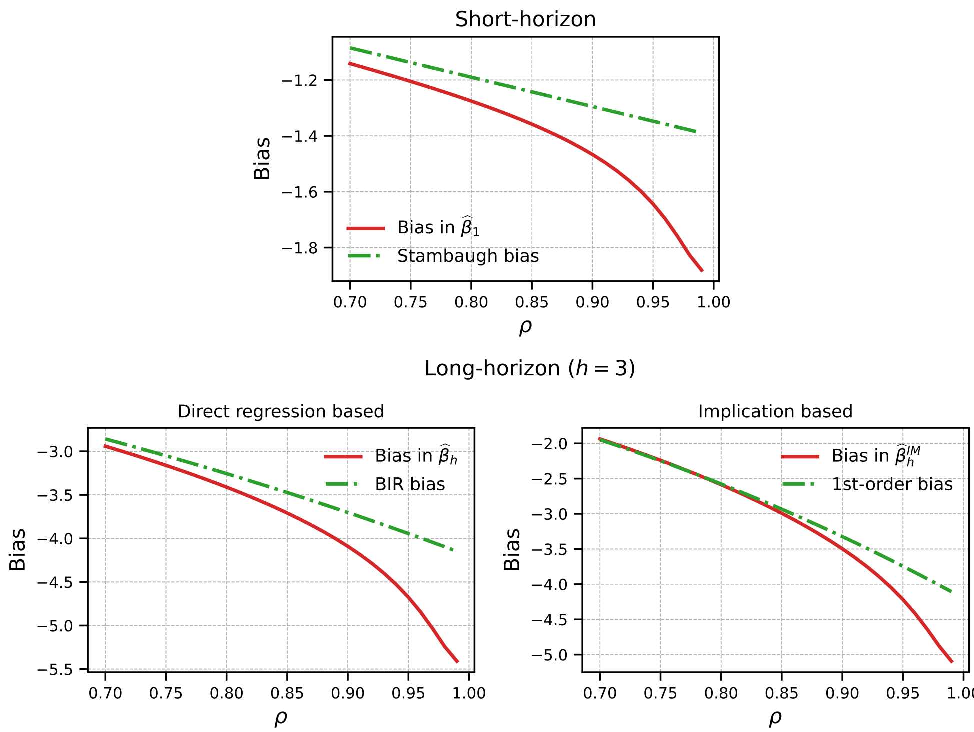


Figure 1: The solid line is the true bias in predictive regressions, and the dashed line is its first-order approximation: Stambaugh bias (Stambaugh, 1999) for $h = 1$; BIR bias (Boudoukh et al., 2022) for the direct regression (DR) estimator and the first-order bias in implication-based (IM) estimator for $h = 3$. Simulations use $r_t = \alpha_1 + \beta_1 x_{t-1} + u_t$ and $x_t = c + \rho x_{t-1} + v_t$, with $\alpha_1 = \beta_1 = c = 0$, and $[u_t, v_t]' \sim i.i.d.N(0, \Sigma)$ with the covariance matrix $\Sigma = \begin{bmatrix} 2.50 \times 10^{-2} & 7.00 \times 10^{-5} \\ 7.00 \times 10^{-5} & 2.00 \times 10^{-6} \end{bmatrix}$, estimated from the cross-sectional premium (csp) predictor.

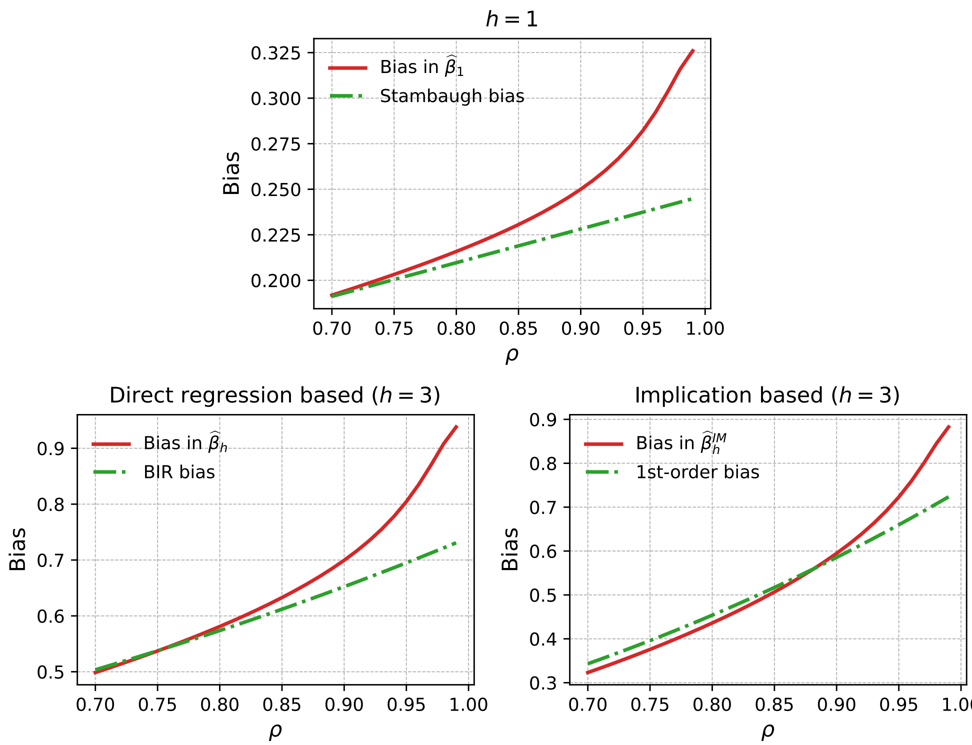


Figure 2: The solid line is the true bias in predictive regressions, and the dashed line is its first-order approximation: Stambaugh bias (Stambaugh, 1999) for $h = 1$; BIR bias (Boudoukh et al., 2022) for the direct regression (DR) estimator and the first-order bias in implication-based (IM) estimator for $h = 3$. Simulations use $r_t = \alpha_1 + \beta_1 x_{t-1} + u_t$ and $x_t = c + \rho x_{t-1} + v_t$, with $\alpha_1 = \beta_1 = c = 0$, and $[u_t, v_t]' \sim i.i.d.N(0, \Sigma)$ with the covariance matrix, $\Sigma = \begin{bmatrix} 3.60 \times 10^{-2} & -3.70 \times 10^{-3} \\ -3.70 \times 10^{-3} & 6.00 \times 10^{-4} \end{bmatrix}$, estimated from the scaled risk-neutral vix (rsvix) predictor

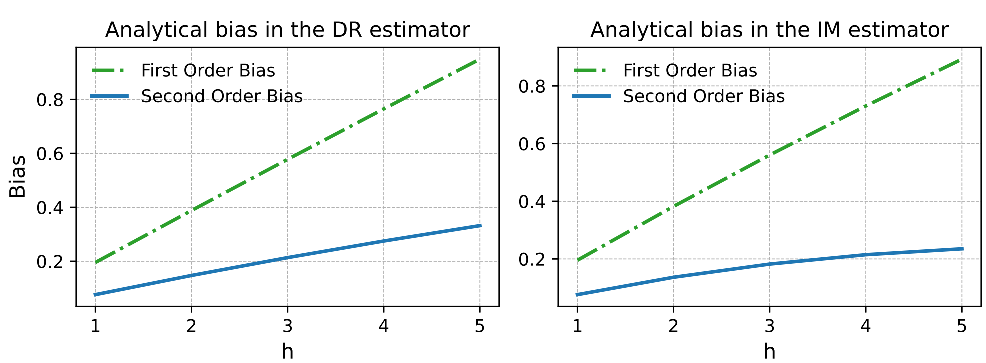


Figure 3: First- (dashed) and second-order (solid) for the DR and IM estimators, calculated from the analytic formulae in Proposition 2.2 for DR and Proposition 2.4 for IM. Parameters are set to $\rho = 0.97$, $\beta_1 = 0$, $\sigma_{uv}/\sigma_v^2 = -5$, $\gamma_1 = \gamma_2 = 0$, and $T = 100$.

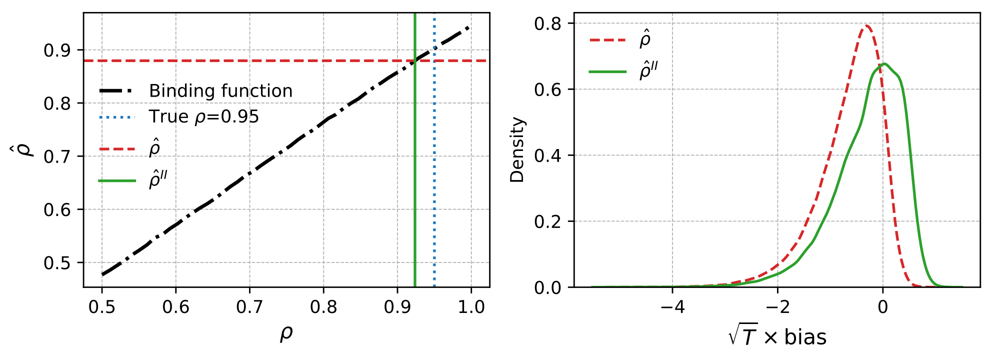


Figure 4: Bias reduction via indirect inference in an AR(1) model with $\rho = 0.95$ and $T = 60$. Left: comparison of OLS (broken red line) and indirect inference (solid green line) estimators. Right: sampling distributions of $\sqrt{T}(\widehat{\rho} - \rho)$ (broken) and $\sqrt{T}(\widehat{\rho}^{II} - \rho)$ (solid) based on 50,000 simulations.

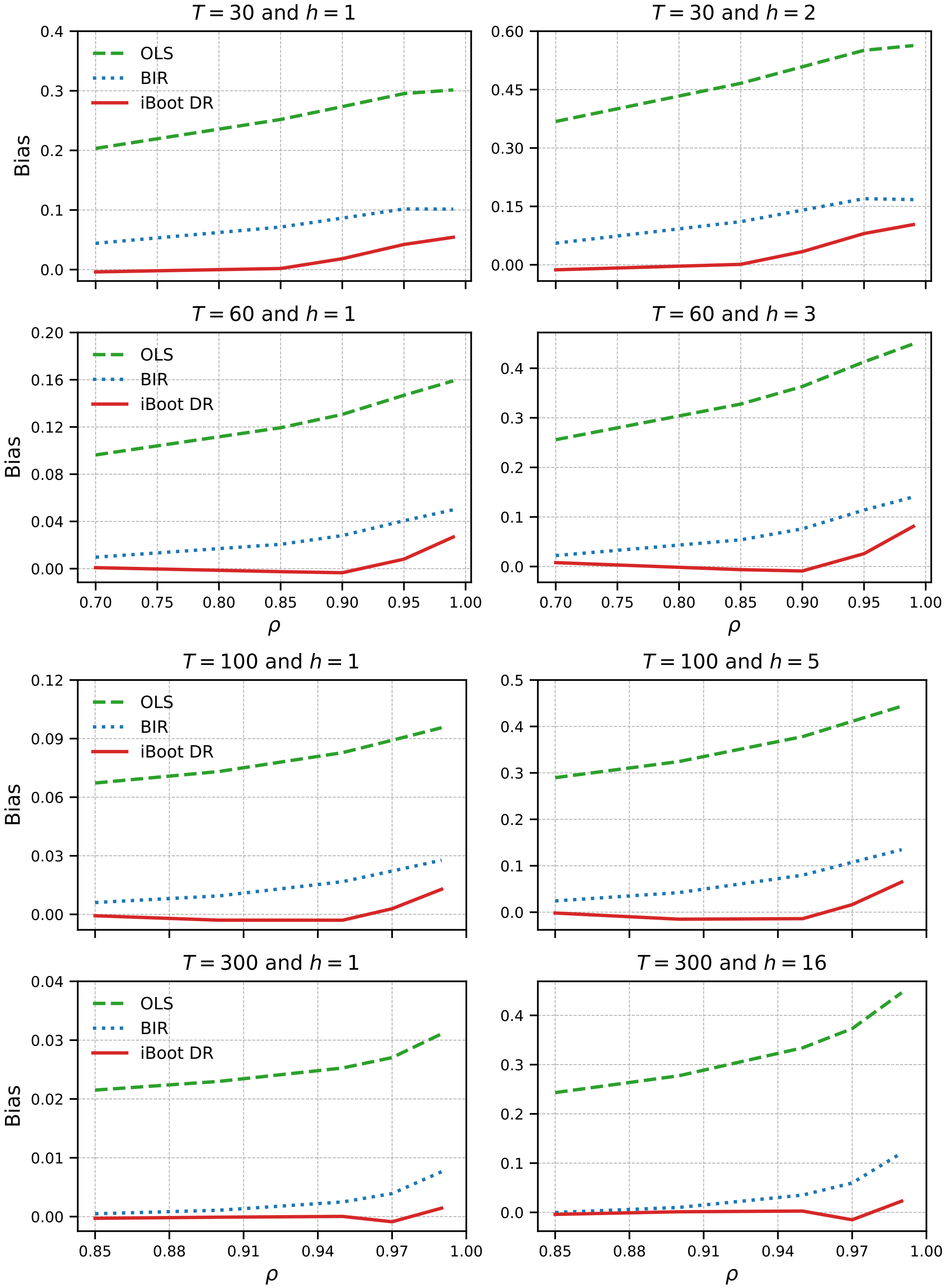


Figure 5: Bias plots for various DR estimators. OLS means the OLS estimator, BIR the BIR estimator proposed in Boudoukh et al. (2022), and iBoot DR the iBoot estimator proposed in this paper.

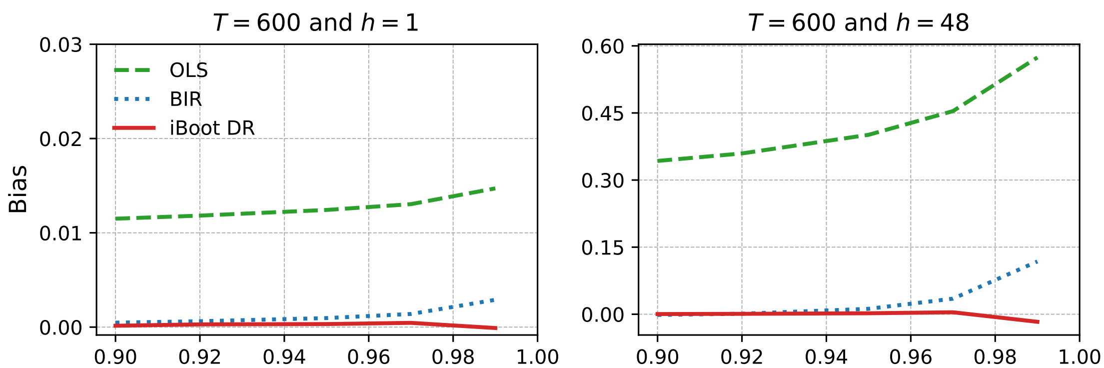


Figure 6: Bias plots for various DR estimators across different horizons when $T = 600$. OLS means the OLS estimator, BIR the BIR estimator proposed in Boudoukh et al. (2022), and iBoot DR the iBoot estimator proposed in this paper.

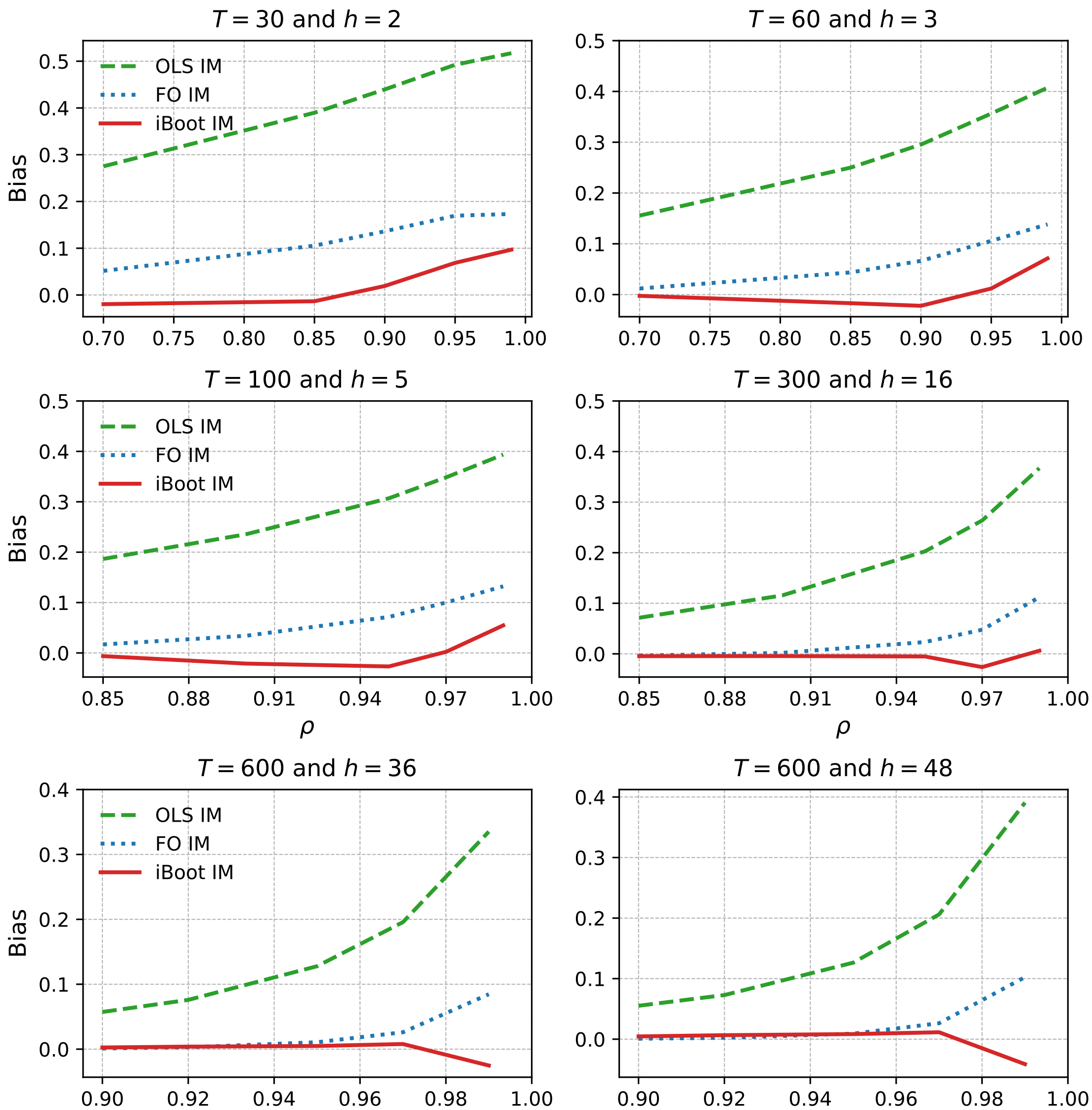


Figure 7: Bias of various Implication-Based (IM) estimators at long horizons. Compares the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator proposed in this paper.

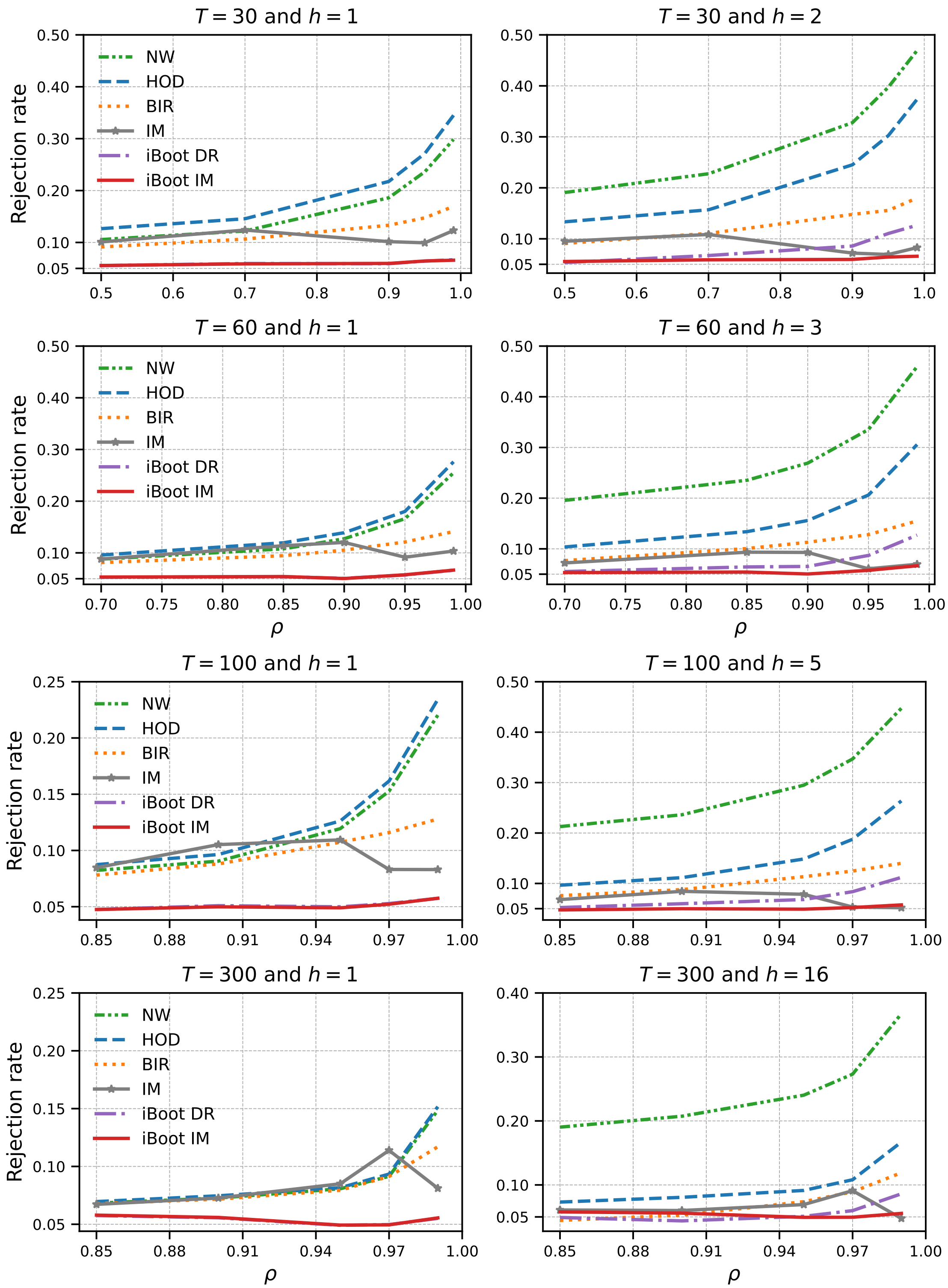


Figure 8: Size curves under $H_0 : \beta_1 = 0$ for various tests: Newey-West (NW), Hodrick (HOD), Boudoukh et al. (2022) (BIR), Xu (2020) (IM), and the proposed iBoot tests for Direct Regression (iBoot DR) and Implication-Based (iBoot IM) estimators.

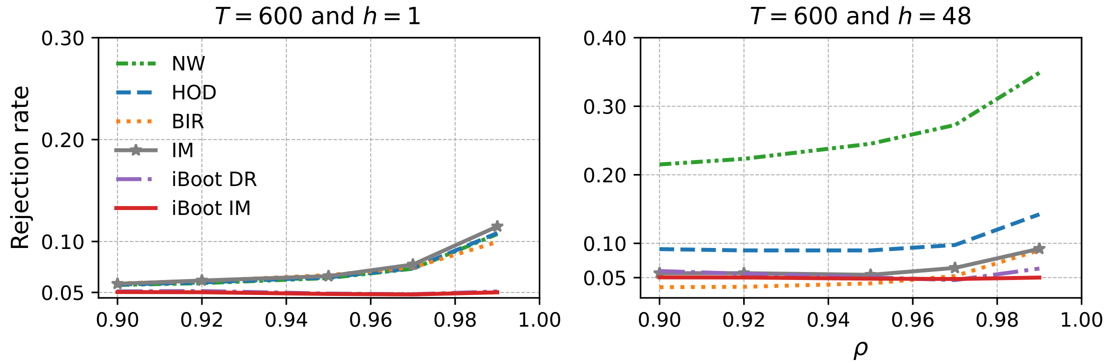


Figure 9: Size curves under $H_0 : \beta_1 = 0$ for various tests across horizons when $T = 600$: Newey-West (NW), Hodrick (HOD), Boudoukh et al. (2022) (BIR), Xu (2020) (IM), and the proposed iBoot tests for Direct Regression (iBoot DR) and Implication-Based (iBoot IM) estimators.

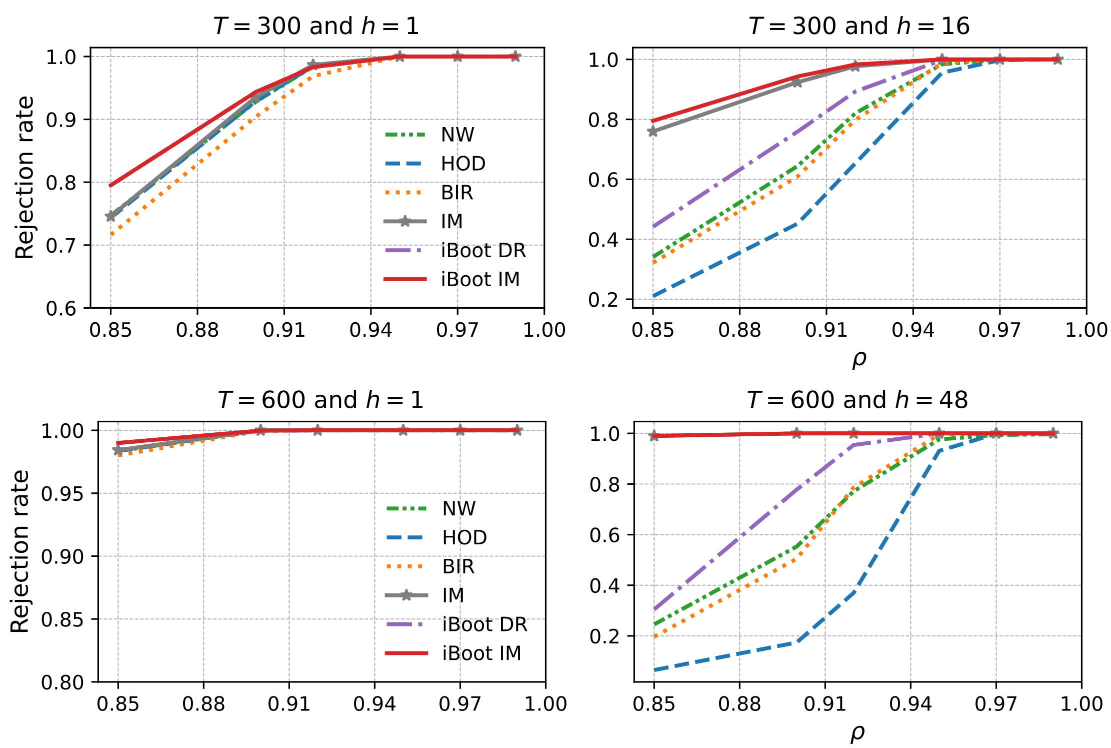


Figure 10: Size-adjusted power curves for testing $H_0 : \beta_1 = 0$ when the true $\beta_1 = 0.16$, across horizons. The figure compares tests from Newey-West (NW), Hodrick (HOD), Boudoukh et al. (2022) (BIR), Xu (2020) (IM), and our proposed iBoot DR and iBoot IM, for $T = 300$ and $T = 600$.

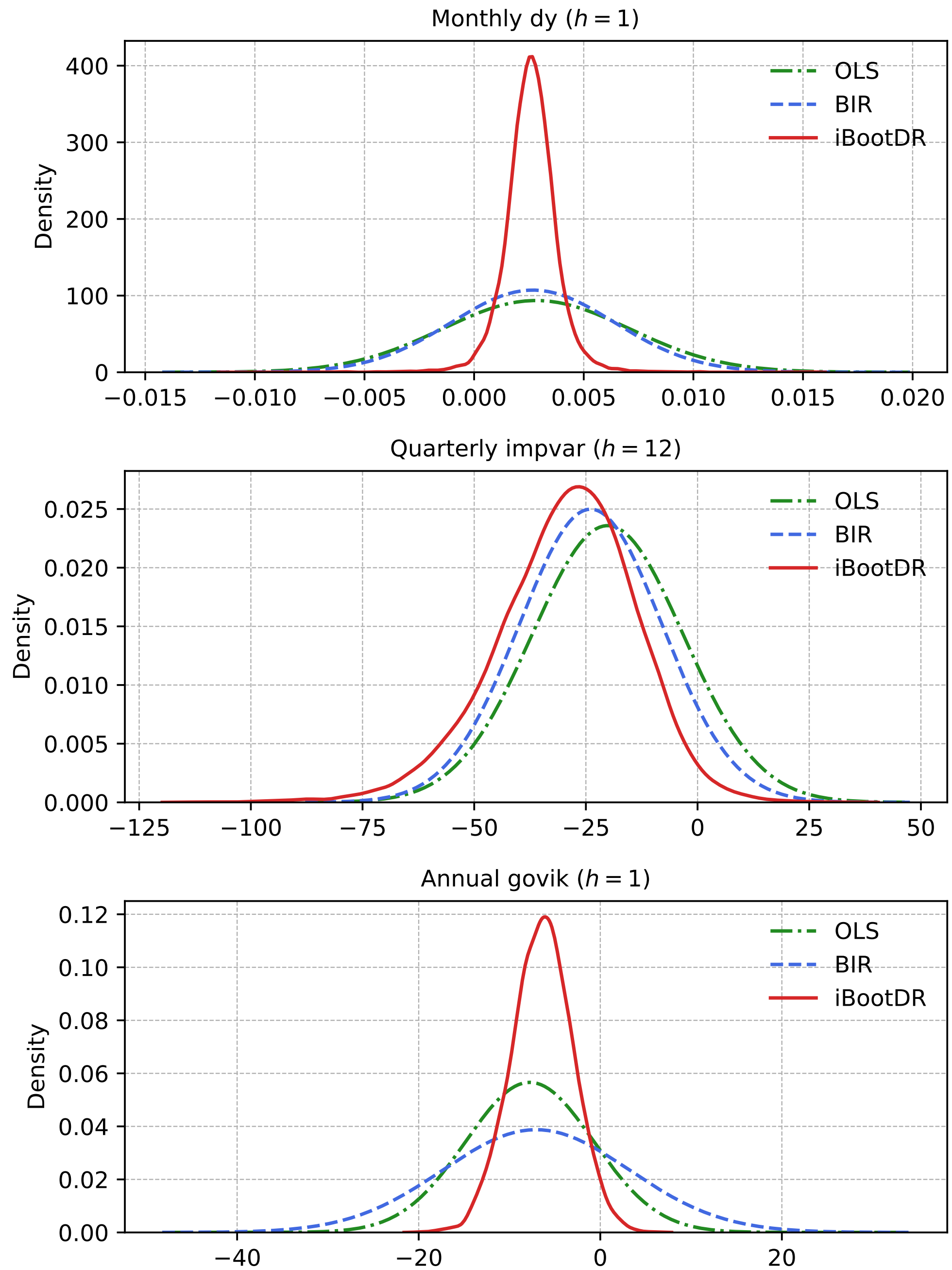
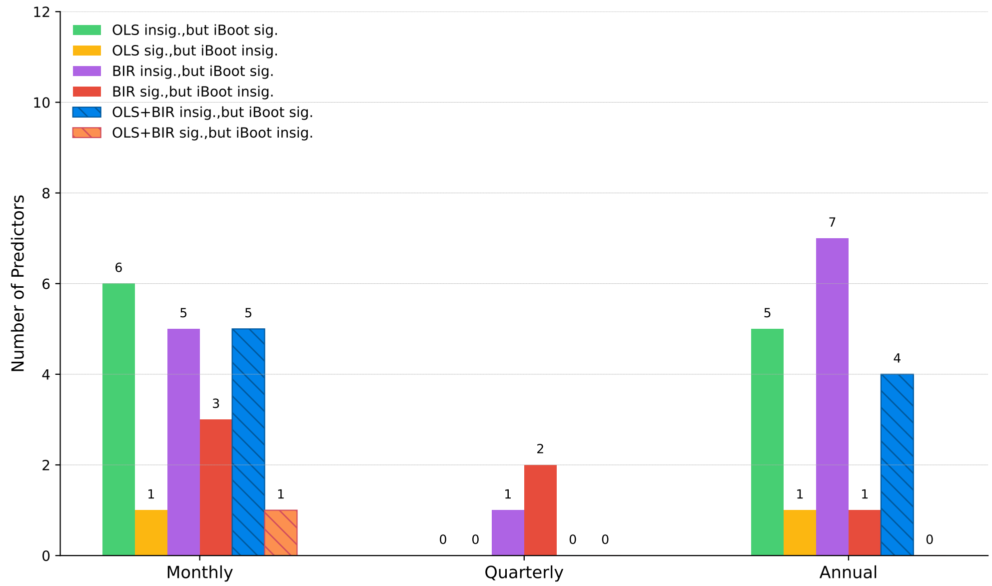
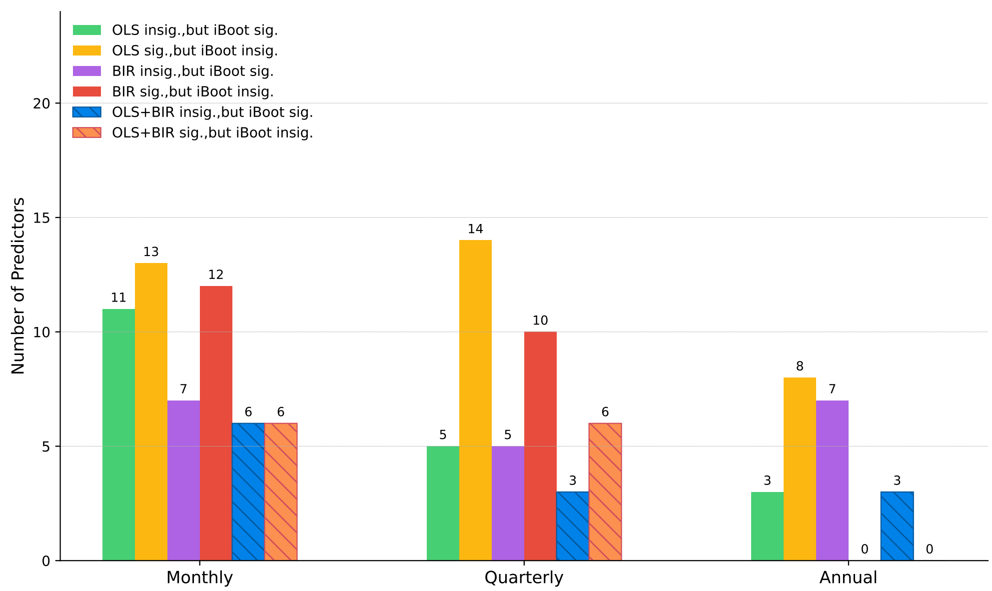


Figure 11: Empirical densities of OLS, BIR, and iBoot DR estimators for three predictors: monthly dy (h=1), quarterly impvar (h=12), and annual govik (h=1) at horzions where iBoot DR changes the significance of both OLS and BIR. OLS and BIR densities are fitted normal distributions using point estimates with Newey-West and analytical standard errors given by Boudoukh et al. (2022), respectively. iBoot uses the bootstrap replications proposed in Section 3.2. Data are from Goyal et al. (2024).



(a) Short horizon ($h = 1$)



(b) Long horizon ($h > 1$)

Figure 12: Changes in statistical significance across all predictors in Goyal et al. (2024) – 35 monthly, 40 quarterly, and 47 annual – when using iBoot estimators compared to OLS and BIR for DR estimators in short (Panel a $h = 1$) and long horizons (Panel b $h > 1$). Each panel classifies changes into six categories: (1) OLS insignificant but iBoot DR significant (green); (2) OLS significant but iBoot DR insignificant (yellow); (3) BIR insignificant but iBoot DR significant (purple); (4) BIR significant but iBoot DR insignificant (red); (5) Both OLS and BIR insignificant but iBoot DR significant (blue); and (6) Both OLS and BIR significant but iBoot DR insignificant (orange). OLS and BIR use Newey-West and analytical standard errors given by Boudoukh et al. (2022), respectively. iBoot uses the inference method proposed in Section 3.2. Data are from Goyal et al. (2024).

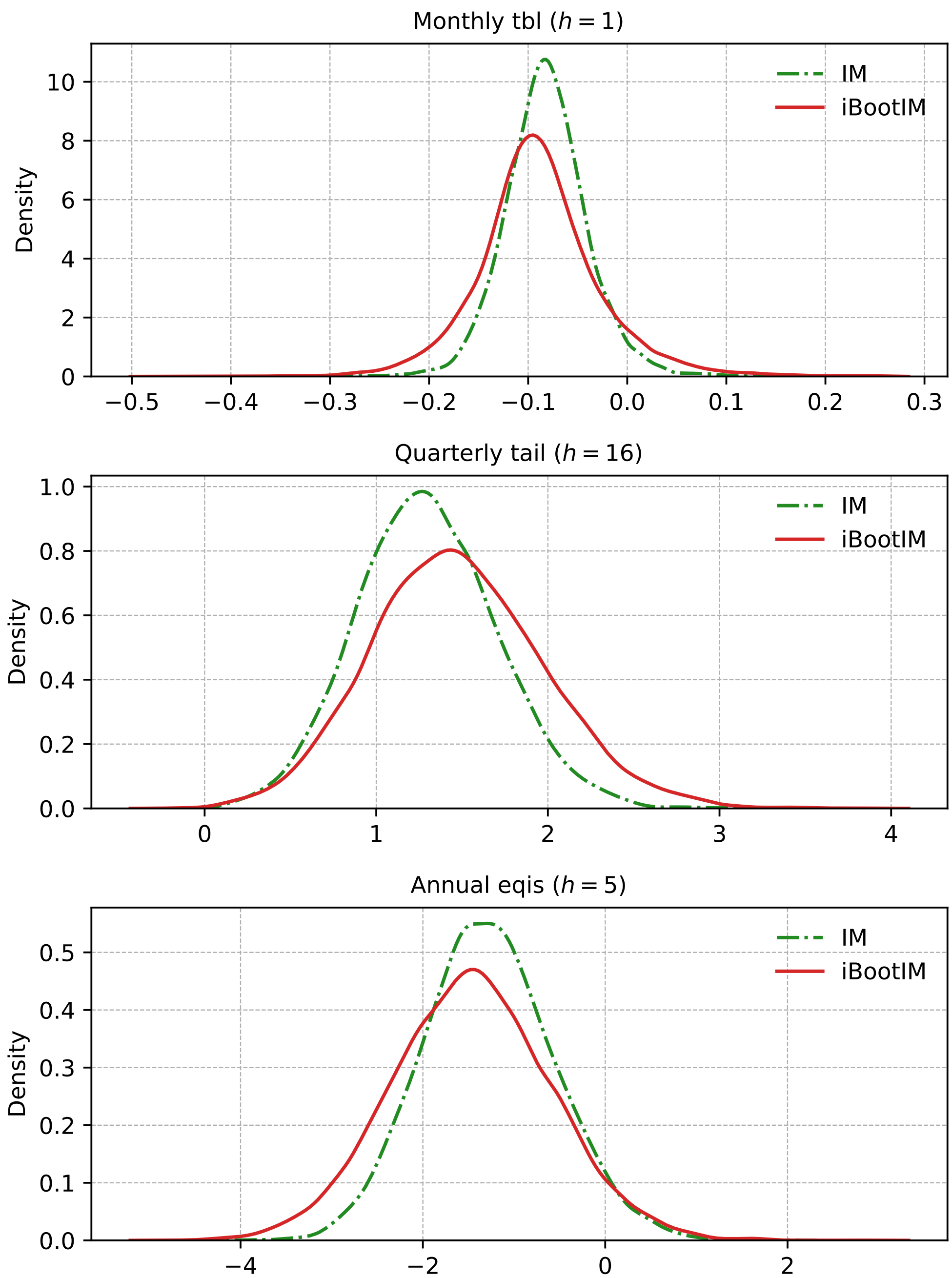
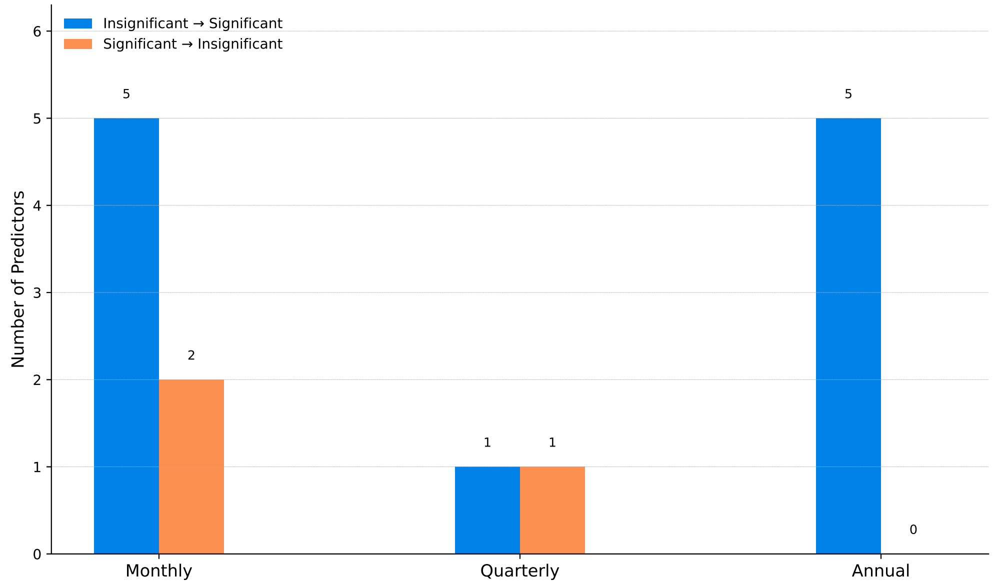
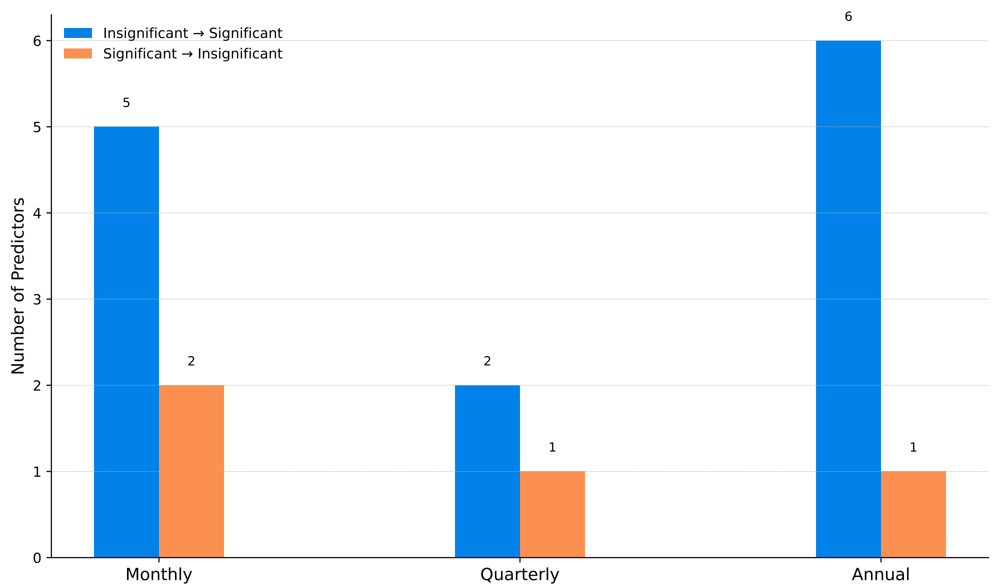


Figure 13: Empirical densities of IM and iBoot IM estimators for three predictors: monthly tbl (h=1), quarterly tail (h=16), and annual eqis (h=5) at horzions where iBoot IM changes the significance of IM. IM densities uses the bootstrapped sample given by Xu (2020) and iBoot uses the bootstrap replications proposed in Section 3.2. Data are from Goyal et al. (2024).



(a) Short horizon ($h = 1$)



(b) Long horizon ($h > 1$)

Figure 14: Changes in statistical significance across all predictors in Goyal et al. (2024) – 35 monthly, 40 quarterly, and 47 annual – when using iBoot estimators compared to IM for IM estimators in short (Panel a $h = 1$) and long horizons (Panel b $h > 1$). Each panel classifies changes into two categories: (1) IM insignificant but iBoot IM significant (blue) and (2) IM significant but iBoot IM insignificant (orange). Inference for IM uses bootstrapped t-ratios (Xu, 2020) and iBoot uses the inference method proposed in Section 3.2. Data are from Goyal et al. (2024).

Table 1: DR estimators – Dividend yield (dy)

This table presents Direct Regression (DR) estimators of the dividend yield (dy) across monthly, quarterly, and annual frequencies. For each, we report the sample period, number of observations (T), autoregressive coefficient ($\hat{\rho}_h$), covariance-to-variance ratio ($\hat{\sigma}_{uv}/\hat{\sigma}_v^2$), and the coefficient estimates from OLS ($\widehat{\beta}_h$), BIR (Boudoukh et al., 2022), and iBoot DR. OLS inference uses Newey-West standard errors; BIR, analytical; and iBoot, our proposed method from Section 3.2. $\%\Delta p^{OLS}$ and $\%\Delta p^{BIR}$ are the percentage p-value change of the iBoot relative to OLS and BIR. Data source: Goyal et al. (2024). Asterisks indicate significance at the 10%, 5%, and 1% levels.

Sample Period	T	$\hat{\rho}_h$	$\hat{\sigma}_{uv}/\hat{\sigma}_v^2$	h	$\hat{\beta}_h$	$\hat{\beta}_h^{BIR}$	$\hat{\beta}_h^{iBootDR}$	$\%\Delta p^{OLS}$	$\%\Delta p^{BIR}$
				Panel A: Monthly					
1968M02-2023M12	671	0.995	-0.02	1	0.003	0.003	0.003**	-93.3	-92.7
				2	0.005	0.005	0.005**	-92.4	-93.0
				12	0.040	0.039	0.039**	-93.7	-94.4
				24	0.069	0.070	0.069**	-92.1	-92.8
				30	0.071	0.074	0.073*	-87.6	-88.3
				36	0.077	0.082	0.080*	-84.1	-84.9
				48	0.097	0.107	0.102*	-80.1	-82.2
				Panel B: Quarterly					
1968Q1-2023Q4	224	0.983	-0.04	1	0.011	0.011	0.011	-47.0	-43.5
				2	0.023	0.021	0.021	-36.4	-42.7
				4	0.041	0.039	0.039	-24.5	-38.0
				8	0.066	0.065	0.065	-21.7	-32.3
				12	0.078	0.080	0.079	-17.4	-26.6
				16	0.094	0.103	0.099	-8.2	-22.8
				Panel C: Annual					
1968-2023	56	0.946	0.09	1	0.032	0.038	0.036**	-91.2	-89.0
				2	0.052	0.067	0.062	-80.8	-78.2
				3	0.051	0.077	0.066	-66.3	-62.5
				4	0.081	0.122	0.105	-68.1	-68.3
				5	0.080	0.137	0.114	-57.7	-58.7

Table 2: DR estimators – Forward implied variance (*impvar*)

This table presents Direct Regression (DR) estimators of the forward implied variance (*impvar*) across monthly, quarterly, and annual frequencies. For each, we report the sample period, number of observations (T), autoregressive coefficient ($\hat{\rho}_h$), covariance-to-variance ratio ($\hat{\sigma}_{uv}/\hat{\sigma}_v^2$), and the coefficient estimates from OLS ($\widehat{\beta}_h$), BIR (Boudoukh et al., 2022), and iBoot DR. OLS inference uses Newey-West standard errors; BIR, analytical; and iBoot, our proposed method from Section 3.2. $\%\Delta p^{OLS}$ and $\%\Delta p^{BIR}$ are the percentage p-value change of the iBoot relative to OLS and BIR. Data source: Goyal et al. (2024). Asterisks indicate significance at the 10%, 5%, and 1% levels.

Sample Period	T	$\hat{\rho}_h$	$\hat{\sigma}_{uv}/\hat{\sigma}_v^2$	h	$\hat{\beta}_h$	$\hat{\beta}_h^{BIR}$	$\hat{\beta}_h^{iBootDR}$	$\%\Delta p^{OLS}$	$\%\Delta p^{BIR}$
				Panel A: Monthly					
1996M01-2023M02	326	0.790	-13.58	1	0.456	0.315	0.265	15.0	11.7
				2	0.769	0.504	0.435	22.2	13.6
				12	1.39	0.247	-1.02	7.6	-9.1
				24	-7.30	-9.36	-14.01	-74.4	-60.5
				30	-13.56	-16.10	-22.79**	-90.5	-80.6
				36	-20.39	-23.43*	-32.02***	-96.3	-88.1
				48	-26.84	-30.95**	-36.44**	-94.3	-68.3
				Panel B: Quarterly					
1996Q1-2022Q4	108	0.607	-22.24	1	3.41	2.82	2.75	10.3	30.3
				2	5.95	4.94	4.83	31.0	4.4
				4	6.20	4.57	3.72	80.9	6.9
				8	-2.93	-5.63	-8.91	-50.5	-39.9
				12	-20.13	-23.94	-29.61**	-84.3	-72.5
				16	-28.30	-33.34*	-33.29*	-77.5	-32.4
				Panel C: Annual					
1996-2022	27	0.431	-41.33	1	3.40	-0.249	0.355	16.4	-17.8
				2	-6.08	-12.88	-13.53	-45.3	-25.1
				3	-17.92	-27.77	-31.60	-47.2	-40.6
				4	-29.88	-42.88	-42.80	-46.6	-24.1
				5	-26.17	-42.54	-33.98	-14.8	42.8

Table 3: DR estimators – Government investment ($govik$)

This table presents Direct Regression (DR) estimators of the government investment ($govik$) across quarterly and annual frequencies. For each, we report the sample period, number of observations (T), autoregressive coefficient ($\hat{\rho}_h$), covariance-to-variance ratio ($\hat{\sigma}_{uv}/\hat{\sigma}_v^2$), and the coefficient estimates from OLS ($\widehat{\beta}_h$), BIR (Boudoukh et al., 2022), and iBoot DR. OLS inference uses Newey-West standard errors; BIR, analytical; and iBoot, our proposed method from Section 3.2. $\%\Delta p^{OLS}$ and $\%\Delta p^{BIR}$ are the percentage p-value change of the iBoot relative to OLS and BIR. Data source: Goyal et al. (2024). Asterisks indicate significance at the 10%, 5%, and 1% levels.

Sample Period	T	$\hat{\rho}_h$	$\hat{\sigma}_{uv}/\hat{\sigma}_v^2$	h	$\hat{\beta}_h$	$\hat{\beta}_h^{BIR}$	$\hat{\beta}_h^{iBootDR}$	$\%\Delta p^{OLS}$	$\%\Delta p^{BIR}$
					Panel A: Monthly				
					Not available				
					Panel B: Quarterly				
1968Q1-2023Q4	224	0.947	13.95	1	-1.44	-1.20	-1.42	-28.5	-56.4
				2	-3.56	-3.10	-3.55**	-84.9	-93.8
				4	-8.36	-7.54	-8.38**	-86.2	-96.6
				8	-18.16**	-16.85	-18.45**	-41.3	-96.3
				12	-24.31**	-22.76	-24.95**	74.9	-94.1
				16	-31.14***	-29.53	-31.94**	256.9	-92.2
					Panel C: Annual				
1968-2023	56	0.767	11.64	1	-7.77	-7.07	-7.51*	-80.2	-89.1
				2	-15.37	-14.46	-15.18**	-79.1	-94.1
				3	-25.04**	-24.25	-24.88**	-39.0	-93.0
				4	-30.23**	-29.81	-30.34*	97.1	-85.6
				5	-41.44**	-41.56	-41.38*	138.7	-82.1

Table 4: IM estimators – Treasury bill rate (*tbl*)

This table presents implication based (IM) estimators of the treasury bill rate (*tbl*) across monthly, quarterly, and annual frequencies. For each, we report the sample period, number of observations (T), autoregressive coefficient ($\hat{\rho}_h$), covariance-to-variance ratio ($\hat{\sigma}_{uv}/\hat{\sigma}_v^2$), and the coefficient estimates from IM and iBoot IM. Inference for IM uses bootstrapped t-ratios (Xu, 2020) and iBoot uses the inference method proposed in Section 3.2. $\%\Delta p^{IM}$ is the percentage p-value change of the iBoot relative to IM. Data source: Goyal et al. (2024). Asterisks indicate significance at the 10%, 5%, and 1% levels.

Sample Period	T	$\hat{\rho}_h$	$\hat{\sigma}_{uv}/\hat{\sigma}_v^2$	h	$\hat{\beta}_h^{IM}$	$\hat{\beta}_h^{iBootIM}$	$\%\Delta p^{IM}$
			Panel A: Monthly				
1968M02-2023M12	671	0.991	-1.16	1	-0.089*	-0.106	67.0
				2	-0.177*	-0.213	66.3
				12	-1.016*	-1.296	64.9
				24	-1.932*	-2.630	60.6
				30	-2.356*	-3.307	56.2
				36	-2.759*	-3.990	54.0
				48	-3.506*	-5.365	48.9
			Panel B: Quarterly				
1968Q1-2023Q4	224	0.954	-0.67	1	-0.221	-0.240	7.1
				2	-0.432	-0.475	6.3
				4	-0.825	-0.933	5.0
				8	-1.509	-1.790	2.3
				12	-2.075	-2.564	-1.0
				16	-2.545	-3.253	-6.1
			Panel C: Annual				
1968-2023	56	0.863	0.43	1	-0.648	-0.543	-13.2
				2	-1.207	-1.074	-13.3
				3	-1.689	-1.586	-13.5
				4	-2.106	-2.075	-13.6
				5	-2.465	-2.536	-13.4

Table 5: IM estimators – Cross-section based tail-risk (*tail*)

This table presents implication based (IM) estimators of the cross-section based tail-risk (*tail*) across monthly, quarterly, and annual frequencies. For each, we report the sample period, number of observations (T), autoregressive coefficient ($\hat{\rho}_h$), covariance-to-variance ratio ($\hat{\sigma}_{uv}/\hat{\sigma}_v^2$), and the coefficient estimates from IM and iBoot IM. Inference for IM uses bootstrapped t-ratios (Xu, 2020) and iBoot uses the inference method proposed in Section 3.2. $\%\Delta p^{IM}$ is the percentage p-value change of the iBoot relative to IM. Data source: Goyal et al. (2024). Asterisks indicate significance at the 10%, 5%, and 1% levels.

Sample Period	T	$\hat{\rho}_h$	$\hat{\sigma}_{uv}/\hat{\sigma}_v^2$	h	$\hat{\beta}_h^{IM}$	$\hat{\beta}_h^{iBootIM}$	$\%\Delta p^{IM}$
			Panel A: Monthly				
1968M02-2023M12	671	0.910	-0.05	1	0.077**	0.076***	-93.3
				2	0.148**	0.146***	-93.1
				12	0.582**	0.585***	-90.7
				24	0.770**	0.780***	-89.7
				30	0.809**	0.821***	-89.3
				36	0.831**	0.843***	-89.1
				48	0.851**	0.863***	-88.6
			Panel B: Quarterly				
1968Q1-2023Q4	224	0.856	0.20	1	0.190	0.188***	-99.3
				2	0.353	0.350***	-99.3
				4	0.611	0.612***	-99.3
				8	0.940	0.951***	-99.2
				12	1.117	1.134***	-99.2
				16	1.212	1.231***	-99.2
			Panel C: Annual				
1968-2023	56	0.733	-1.0	1	1.252***	1.251***	-100.0
				2	2.169***	2.179***	-100.0
				3	2.843***	2.858***	-100.0
				4	3.336***	3.344***	-100.0
				5	3.699**	3.683***	-100.0

Table 6: IM estimators – Equity issuing activity ($eqis$)

This table presents implication based (IM) estimators of the equity issuing activity ($eqis$) across monthly, quarterly, and annual frequencies. For each, we report the sample period, number of observations (T), autoregressive coefficient ($\hat{\rho}_h$), covariance-to-variance ratio ($\hat{\sigma}_{uv}/\hat{\sigma}_v^2$), and the coefficient estimates from IM and iBoot IM. Inference for IM uses bootstrapped t-ratios (Xu, 2020) and iBoot uses the inference method proposed in Section 3.2. $\%\Delta p^{IM}$ is the percentage p-value change of the iBoot relative to IM. Data source: Goyal et al. (2024). Asterisks indicate significance at the 10%, 5%, and 1% levels.

Sample Period	T	$\hat{\rho}_h$	$\hat{\sigma}_{uv}/\hat{\sigma}_v^2$	h	$\hat{\beta}_h^{IM}$	$\hat{\beta}_h^{iBootIM}$	$\%\Delta p^{IM}$
Panel A: Monthly							
Not available							
Panel B: Quarterly							
Not available							
Panel C: Annual							
1968-2023	56	0.769	0.05	1	-0.401*	-0.407*	2.3
				2	-0.709*	-0.744*	-4.3
				3	-0.946	-1.019*	-12.8
				4	-1.128	-1.240*	-20.0
				5	-1.267	-1.414*	-25.5

7 Appendix

7.1 Estimates of the predictor persistence ρ in Goyal and Welch (2008)

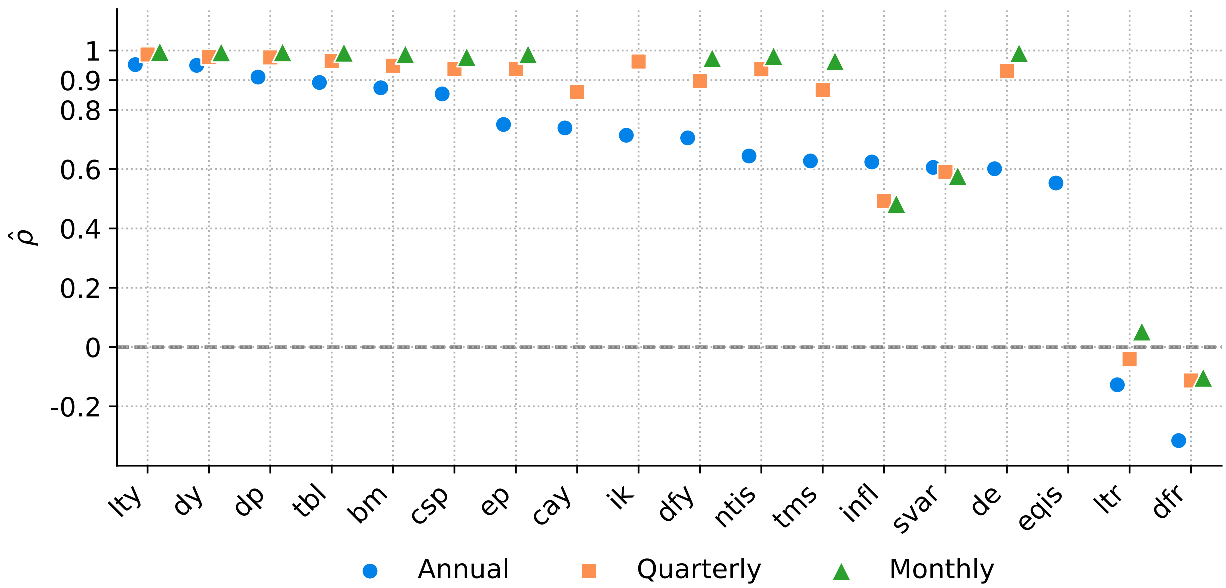


Figure 15: The values of $\widehat{\rho}$. The list of predictors is from Goyal and Welch (2008). The sample period is from 1926 to 2023. Data sources: Goyal and Welch (2008).

7.2 Proof of Proposition 2.1

By (2.3), we have

$$\begin{aligned}
\widehat{\beta}_h &= \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \sum_{i=0}^{h-1} r_{t+i}}{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1}^2} \\
&= \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \sum_{i=0}^{h-1} (\alpha + \beta x_{t+i-1} + u_{t+i})}{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1}^2} \\
&= \beta_1 \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \sum_{i=0}^{h-1} x_{t+i-1}}{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1}^2} + \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \sum_{i=0}^{h-1} u_{t+i}}{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1}^2} \\
&= \beta_1 \sum_{i=0}^{h-1} \widehat{\rho}_i + \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \sum_{i=0}^{h-1} u_{t+i}}{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1}^2}, \qquad (7.1)
\end{aligned}$$

where $\tilde{x}_{t-1} = x_{t-1} - \sum_{t=1}^{T-h+1} x_{t-1}/(T-h+1)$, which implies

$$\widehat{\beta}_h - \beta_h = \beta_1 \sum_{i=0}^{h-1} (\widehat{\rho}_i - \rho_i) + \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \sum_{i=0}^{h-1} u_{t+i}}{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1}^2}. \qquad (7.2)$$

Now assume that $E[u_t|v_t]$ is linear in v_t and set $e_{t+i} = u_{t+i} - \frac{\sigma_{uv}}{\sigma_v^2} v_{t+i}$. Then from (7.2), we have

$$\begin{aligned}
\widehat{\beta}_h - \beta_h &= \beta_1 \sum_{i=0}^{h-1} (\widehat{\rho}_i - \rho_i) + \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \sum_{i=0}^{h-1} (e_{t+i} + \frac{\sigma_{uv}}{\sigma_v^2} v_{t+i})}{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1}^2} \\
&= \beta_1 \sum_{i=1}^{h-1} (\widehat{\rho}_i - \rho_i) + \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \sum_{i=0}^{h-1} (\frac{\sigma_{uv}}{\sigma_v^2} v_{t+i})}{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1}^2} + \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \sum_{i=0}^{h-1} e_{t+i}}{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1}^2}. \qquad (7.3)
\end{aligned}$$

By backward substitutions,

$$\begin{aligned}
x_t &= c + \rho x_{t-1} + v_t \\
x_{t+1} &= c(1+\rho) + \rho^2 x_{t-1} + v_{t+1} + \rho v_t \\
&\dots \\
x_{t+i-1} &= c(1 + \rho + \cdots + \rho^{i-1}) + \rho^i x_{t-1} + v_{t+i-1} + \rho v_{t+i-2} + \cdots + \rho^{i-1} v_t \qquad (7.4) \\
x_{t+i} &= c(1 + \rho + \cdots + \rho^i) + \rho^{i+1} x_{t-1} + v_{t+i} + \rho v_{t+i-1} + \cdots + \rho^i v_t. \qquad (7.5)
\end{aligned}$$

From (7.5) we have

$$\begin{aligned}
\widehat{\rho}_{i+1} &= \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} x_{t+i}}{\tilde{x}_{t-1}^2} \\
&= \rho_{i+1} + \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \left(v_{t+i} + \rho v_{t+i-1} + \cdots + \rho^i v_t\right)}{\tilde{x}_{t-1}^2}. \qquad (7.6)
\end{aligned}$$

From (7.4) we have

$$
\begin{aligned}
\widehat{\rho}_i &= \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} x_{t+i-1}}{\tilde{x}_{t-1}^2} \\
&= \rho_i + \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \left(v_{t+i-1} + \rho v_{t+i-2} + \cdots + \rho^{i-1} v_t \right)}{\tilde{x}_{t-1}^2},
\end{aligned}
$$

which implies that

$$
\rho \widehat{\rho}_i = \rho \rho_i + \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \left(\rho v_{t+i-1} + \rho^2 v_{t+i-2} + \cdots + \rho^i v_t \right)}{\tilde{x}_{t-1}^2}. \tag{7.7}
$$

Subtracting (7.6) by (7.7), we have

$$
\frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} v_{t+i}}{\tilde{x}_{t-1}^2} = (\widehat{\rho}_{i+1} - \rho_{i+1}) - (\rho \widehat{\rho}_i - \rho \rho_i), \tag{7.8}
$$

which can be inserted into (7.3) to obtain

$$
\widehat{\beta}_h - \beta_h = \beta_1 \sum_{i=0}^{h-1} (\widehat{\rho}_i - \rho_i) + \frac{\sigma_{uv}}{\sigma_v^2} \sum_{i=0}^{h-1} [(\widehat{\rho}_{i+1} - \rho_{i+1}) - (\rho \widehat{\rho}_i - \rho \rho_i)] + \frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \sum_{i=0}^{h-1} e_{t+i}}{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1}^2}. \tag{7.9}
$$

Taking the expectations of both sides, we have

$$
\begin{aligned}
\widehat{\beta}_h - \beta_h &= \beta_1 \sum_{i=0}^{h-1} E[\widehat{\rho}_i - \rho_i] + \frac{\sigma_{uv}}{\sigma_v^2} \left\{ \sum_{i=0}^{h-1} E\left[\widehat{\rho}_{i+1} - \rho_{i+1}\right] - \rho \sum_{i=0}^{h-1} E\left[\widehat{\rho}_i - \rho_i\right] \right\} \\
&= \left[\beta_1 + \frac{\sigma_{uv}}{\sigma_v^2} (1 - \rho) \right] \sum_{i=0}^{h-1} E[\widehat{\rho}_i - \rho_i] + \frac{\sigma_{uv}}{\sigma_v^2} E\left[\widehat{\rho}_h - \rho_h\right],
\end{aligned} \tag{7.10}
$$

which is obtained by

$$
\begin{aligned}
E\left[\frac{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1} \sum_{i=0}^{h-1} e_{t+i}}{\sum_{t=1}^{T-h+1} \tilde{x}_{t-1}^2} \right] &= E\left\{ \sum_{i=0}^{h-1} \sum_{t=1}^{T-h+1} \frac{\tilde{x}_{t-1}}{\sum_{t=1}^{T} \tilde{x}_{t-1}^2} E\left[e_{t+i} | x_0, v_0, v_1, \ldots, v_T \right] \right\} \\
&= 0,
\end{aligned}
$$

because $E\left[e_{t+i} | x_0, v_0, v_1, \ldots, v_T\right] = E\left[e_{t+i} | v_{t+i}\right] = 0$. Q.E.D.

7.3 Proof of Proposition 2.2

We first derive the bias formula for $\widehat{\rho}_i$ up to the order $O(T^{-2})$. We then show that for $1 \leq i \leq h$, the approximate bias of the OLS estimator

$$\widehat{\rho}_i = \sum_{t=1}^{T-h+1} \widetilde{x}_{t-1} x_{t+i-1} \Big/ \left(\sum_{t=1}^{T-h+1} \widetilde{x}_{t-1}^2 \right),$$

which is based on a sample of size $T-h+1$, is given by

$$\begin{aligned}
E(\widehat{\rho}_i - \rho_i) &= -\frac{1}{T-h+1}\left[(1+\rho)\frac{1-\rho^i}{1-\rho} + 2i\rho^i\right] \\
&\quad -\frac{1}{(T-h+1)^2}\left(\frac{i\rho^i(3\rho^2+8\rho+3)}{1-\rho^2} - \frac{i(\rho+1)}{1-\rho} - (5i^2+1)\rho^i + 1\right) \\
&\quad +\frac{2\gamma_2 i\rho^i}{(T-h+1)^2} + \frac{4\gamma_1^2(1+\rho)}{(T-h+1)^2(1+\rho+\rho^2)}\frac{\rho^{i+1}(1-\rho^i)}{1-\rho} + o(T^{-2}). \qquad (7.11)
\end{aligned}$$

Equation (7.11) implies that

$$\begin{aligned}
\sum_{i=1}^{h-1} E(\widehat{\rho}_i - \rho_i) &= -\frac{1}{n}\left[h-1+\frac{(2h-1)(\rho-\rho^h)}{1-\rho}\right] \\
&\quad -\frac{1}{n^2}\left[\frac{(\rho^h-1)(1+\rho+\rho^2+3\rho^3)}{(1-\rho)^3(1+\rho)} + \frac{h\rho^h(7\rho^2+2\rho-3)}{(1-\rho)^2(1+\rho)}\right. \\
&\quad \left. +\frac{5h^2\rho^h+(h-h^2)\rho}{1-\rho} + \frac{3h-h^2}{2}\right] - \frac{1}{n^2}\left[\frac{2\gamma_2(h\rho^h(1-\rho)-\rho(1-\rho^h))}{(1-\rho)^2}\right] \\
&\quad -\frac{1}{n^2}\left[\frac{4\gamma_1^2(-\rho^2+\rho^{h+1}+\rho^{h+2}-\rho^{2h+1})}{(1-\rho)^2(1+\rho+\rho^2)}\right] + o(T^{-2}) \\
&= -\frac{1}{T}\left[h-1+\frac{(2h-1)(\rho-\rho^h)}{1-\rho}\right] - \frac{h-1}{T^2}\left[h-1+\frac{(2h-1)(\rho-\rho^h)}{1-\rho}\right] \\
&\quad -\frac{1}{T^2}\left[\frac{(\rho^h-1)(1+\rho+\rho^2+3\rho^3)}{(1-\rho)^3(1+\rho)} + \frac{h\rho^h(7\rho^2+2\rho-3)}{(1-\rho)^2(1+\rho)}\right. \\
&\quad \left. +\frac{h^2(5\rho^h-\rho)+h}{1-\rho} + \frac{h-h^2}{2}\right] - \frac{1}{T^2}\left[\frac{2\gamma_2(h\rho^h(1-\rho)-\rho(1-\rho^h))}{(1-\rho)^2}\right] \\
&\quad -\frac{1}{T^2}\left[\frac{4\gamma_1^2(-\rho^2+\rho^{h+1}+\rho^{h+2}-\rho^{2h+1})}{(1-\rho)^2(1+\rho+\rho^2)}\right] + o(T^{-2}) \\
&= -\frac{1}{T}\left[h-1+\frac{(2h-1)(\rho-\rho^h)}{1-\rho}\right] \\
&\quad -\frac{1}{T^2}\left[\frac{(\rho^h-1)(1+\rho+\rho^2+3\rho^3)}{(1-\rho)^3(1+\rho)} + \frac{h\rho^h(7\rho^2+2\rho-3)}{(1-\rho)^2(1+\rho)}\right. \\
&\quad \left. +\frac{(3h^2+3h-1)\rho^h+(h-1)^2\rho}{1-\rho} + \frac{h^2-h+2}{2}\right] \\
&\quad -\frac{1}{T^2}\left[\frac{2\gamma_2(h\rho^h(1-\rho)-\rho(1-\rho^h))}{(1-\rho)^2}\right]
\end{aligned}$$

$$-\frac{1}{T^2}\left[\frac{4\gamma_1^2(-\rho^2+\rho^{h+1}+\rho^{h+2}-\rho^{2h+1})}{(1-\rho)^2(1+\rho+\rho^2)}\right]+o(T^{-2})$$

in which $n = T - h + 1$. In addition, by using Proposition 2.1, we have

$$\begin{aligned}
& E[\widehat{\beta}_h-\beta_h] \\
= \; & -\frac{1}{T}\frac{\sigma_{uv}}{\sigma_v^2}\left[h(1+\rho)+\frac{2\rho(1-\rho^h)}{1-\rho}\right]-\frac{\beta_1}{T}\left[h-1+\frac{(2h-1)(\rho-\rho^h)}{1-\rho}\right] \\
& -\frac{1}{T^2}\frac{\sigma_{uv}}{\sigma_v^2}\left[\frac{A_1^{DR}(\rho,h)}{(1-\rho)^2(1+\rho)}+\frac{(h^2-3h)\rho}{2}+\frac{h^2-h+2}{2}+\frac{2\gamma_2A_3^{DR}(\rho,h)}{1-\rho}+\frac{4\gamma_1^2A_4^{DR}(\rho,h)}{(1-\rho)(1+\rho+\rho^2)}\right] \\
& -\frac{\beta_1}{T^2}\left[\frac{A_2^{DR}(\rho,h)}{(1-\rho)^3(1+\rho)}+\frac{(h-1)^2\rho}{1-\rho}+\frac{h^2-h+2}{2}+\frac{2\gamma_2A_5^{DR}(\rho,h)}{(1-\rho)^2}+\frac{4\gamma_1^2A_6^{DR}(\rho,h)}{(1-\rho)^2(1+\rho+\rho^2)}\right]+o(T^{-2})
\end{aligned}$$

with

$$\begin{aligned}
A_1^{DR}(\rho,h) &= h\rho^h(1-\rho)(8\rho^2+8\rho)+(\rho^3+\rho^2+3\rho+1)(\rho^h-1) \\
&\quad -\rho^h(1-\rho)(1-\rho^2) \\
A_2^{DR}(\rho,h) &= (3h^2-1)\rho^h(1-\rho)(1-\rho^2)+h\rho^h(1-\rho)(4\rho^2+2\rho) \\
&\quad +(3\rho^3+\rho^2+\rho+1)(\rho^h-1), \\
A_3^{DR}(\rho,h) &= -\rho+\rho^{h+1}, \quad A_4^{DR}(\rho,h)=-\rho^2+\rho^{2h+2}, \\
A_5^{DR}(\rho,h) &= -\rho(1-\rho^h)+h\rho^h(1-\rho), \quad A_6^{DR}(\rho,h)=-\rho^2+\rho^{h+1}+\rho^{h+2}-\rho^{2h+1}.
\end{aligned}$$

For the subsequent derivations, we assume for notational simplicity that the sample size is $T-i+1$ when estimating ρ_i. Note that $x_t = \sum_{j=0}^{i-1} c\rho^j + \rho^i x_{t-i} + \sum_{j=0}^{i-1}\rho^j v_{t-j}$, then we can decompose the OLS estimator as

$$\widehat{\rho}_i=\rho_i+\sum_{k=1}^{i}\rho^{i-k}(\mathbf{x}_{-i}'\mathbf{M}_i\mathbf{x}_{-i})^{-1}\mathbf{x}_{-i}'\mathbf{M}_i\mathbf{v}_{k,T-i+k}=\rho_i+\sum_{k=1}^{i}\rho^{i-k}\frac{N_{i,k}}{D_i},$$

where $\mathbf{v}_{k,T-i+k}=[v_k,v_{k+1},...,v_{T-i+k}]^\top$, $N_{i,k}=\mathbf{x}_{-i}'\mathbf{M}_i\mathbf{v}_{k,T-i+k}$, $D_i=\mathbf{x}_{-i}'\mathbf{M}_i\mathbf{x}_{-i}$, $\mathbf{M}_i=\mathbf{I}_{T-i+1}-\frac{1}{T-i+1}\boldsymbol{\ell}_{T-i+1}\boldsymbol{\ell}_{T-i+1}'$ is the projection matrix, $\boldsymbol{\ell}_{T-i+1}=[1,...,1]'$, $\mathbf{x}_{-i}=[x_0,...,x_{T-i}]'$ and $\mathbf{x}_i=[x_i,...,x_T]'$.

Define the selection matrices $\mathbf{J}_{i,k}=\left[\mathbf{I}_{T-i+1},\mathbf{0}_{(T-i+1)\times(k-1)}\right]$ and $\mathbf{S}_{i,k}=\left[\mathbf{0}_{(T-i+1)\times(k-1)},\mathbf{I}_{T-i+1}\right]$, and thus we have $\mathbf{x}_{-i}=\mathbf{J}_{i,k}\mathbf{x}_{-(i-k+1)}$, $\mathbf{v}_{k,T-k+i}=\mathbf{S}_{i,k}\mathbf{v}_{1,T-k+i}$, $N_{i,k}=\mathbf{x}_{-(i-k+1)}'\mathbf{J}_{i,k}'\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-k+i}$

and $D_i = \mathbf{x}'_{-(i-k+1)}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{-(i-k+1)}$. In addition, define $\mathbf{F}_{T-i+k} = [1, \rho, ..., \rho^{T-i+k-1}]'$ and

$$\mathbf{C}_{T-i+k} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ \rho & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho^{T-i+k-2} & \cdots & \rho & 1 & 0 \end{bmatrix}.$$

Then, we have

$$\begin{aligned}
\mathbf{x}_{-(i-k+1)} &= \mathbf{F}_{T-i+k}x_0 + \mathbf{C}_{T-i+k}\boldsymbol{\ell}_{T-i+k}c + \mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k} = \mathbf{x}_{T-i+k,D} + \mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k}, \\
N_{i,k} &= \mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k} + \mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}, \\
D_i &= \mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D} + 2\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k} \\
&\quad +\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k},
\end{aligned}$$

where $\mathbf{x}_{T-i+k,D} = \mathbf{F}_{T-i+k}x_0 + \mathbf{C}_{T-i+k}\boldsymbol{\ell}_{T-i+k}c$.

Then, we can decompose $\widehat{\rho}_i - \rho_i$ as follows

$$\begin{aligned}
\widehat{\rho}_i - \rho_i &= \sum_{k=1}^{i}\rho^{i-k}\frac{N_{i,k}}{D_i - E(D_i) + E(D_i)} = \sum_{k=1}^{i}\rho^{i-k}\frac{N_{i,k}}{E(D_i)} \times \frac{1}{1 + (D_i - E(D_i))(E(D_i))^{-1}} \\
&= \sum_{k=1}^{i}\rho^{i-k}\frac{N_{i,k}}{E(D_i)}\big(1 - (D_i - E(D_i))(E(D_i))^{-1} + (D_i - E(D_i))^2(E(D_i))^{-2} \\
&\quad -(D_i - E(D_i))^3(E(D_i))^{-3}\big) + o_P(T^{-2}) \\
&= \sum_{k=1}^{i}\rho^{i-k}\left(a_{k,-1/2} + a_{k,-1} + a_{k,-3/2} + a_{k,-2}\right) + o_P(T^{-2}),
\end{aligned}$$

where the second equality follows from the Taylor series expansion for $1/(1+x)$ and the facts that $N_{i,k} = O_P(T^{1/2})$, $D_i = O_P(T)$, $D_i - E(D_i) = O_P(T^{1/2})$ and

$$a_{k,-l/2} = (-1)^{l-1}N_{i,k}\left(D_i - E(D_i)\right)^{l-1}\left[E(D_i)\right]^{-l} = O_P(T^{-l/2})$$

for $l = 1, ..., 4$.

Given the stochastic expansion of $\widehat{\rho}_i - \rho_i$, the approximate bias, up to order $O(T^{-2})$, can be written as

$$E(\widehat{\rho}_i - \rho_i) = \sum_{k=1}^{i}\rho^{i-k}E\left(a_{k,-1/2} + a_{k,-1} + a_{k,-3/2} + a_{k,-2}\right) + o(T^{-2})$$

$$= \sum_{k=1}^{i} \rho^{i-k} \left(4\frac{E(N_{i,k})}{E(D_i)} - 6\frac{E(N_{i,k}D_i)}{[E(D_i)]^2} + 4\frac{E(N_{i,k}D_i^2)}{[E(D_i)]^3} - \frac{E(N_{i,k}D_i^3)}{[E(D_i)]^4} \right) + o(T^{-2}).$$

Then, to complete the proof, it is sufficient to compute the terms $E(N_{i,k}D_i^l)$ for $l = 0, 1, 2, 3$ and $E(D_i)$.

Consider $E(N_{i,k})$ first. Since x_0 and $\mathbf{v}_{1,T-i+k}$ are mutually independent and by using Lemma 7.2 (1), we have

$$\begin{aligned} E(N_{i,k}) &= E(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}) \\ &= \sigma_v^2 \text{tr}\left\{\mathbf{S}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\right\} \\ &= -\frac{\sigma_v^2}{1-\rho} + \frac{\sigma_v^2(k-1)}{(T-i+1)(1-\rho)} + \frac{\sigma_v^2}{(T-i+1)(1-\rho)^2} + o(T^{-1}). \end{aligned}$$

Consider $E(D_i)$, write

$$\begin{aligned} E(D_i) &= E(\mathbf{x}'_{-(i-k+1)}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{-(i-k+1)}) \\ &= E(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k}) + E(\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}). \end{aligned}$$

For the first term in the right hand side (r.h.s.), by using Lemma 7.2 (2), we have

$$\begin{aligned} & E(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k}) \\ = {} & \sigma_v^2 \text{tr}\{\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\} \\ = {} & \sigma_v^2 \left(\frac{T-i+1}{1-\rho^2} - \frac{1}{(1-\rho^2)^2} - \frac{1}{(1-\rho)^2} \right) + O(T^{-1}). \end{aligned}$$

For the second term on the r.h.s., by using the identity $\mathbf{C}_{T-i+k}\boldsymbol{\ell}_{T-i+k} = (\boldsymbol{\ell}_{T-i+k} - \mathbf{F}_{T-i+k})/(1-\rho)$ and Lemma 7.2 (3), we have

$$\begin{aligned} & E(\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}) \\ = {} & E(x_0^2\mathbf{F}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{F}_{T-i+k}) + c^2E(\boldsymbol{\ell}'_{T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\boldsymbol{\ell}_{T-i+k}) \\ & +E(2x_0c\mathbf{F}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\boldsymbol{\ell}_{T-i+k}) \\ = {} & \left(\frac{c^2}{(1-\rho)^2} + \frac{\sigma_v^2}{1-\rho^2} \right) \left(\mathbf{F}'_{T-i+1}\mathbf{M}_i\mathbf{F}_{T-i+1} \right) \\ & + \frac{c^2}{(1-\rho)^2} \times \left((\boldsymbol{\ell}_{T-i+1} - \mathbf{F}_{T-i+1})'\mathbf{M}_i(\boldsymbol{\ell}_{T-i+1} - \mathbf{F}_{T-i+1}) \right) \\ & +2\frac{c^2}{(1-\rho)^2} \left((\boldsymbol{\ell}_{T-i+1} - \mathbf{F}_{T-i+1})'\mathbf{M}_i\mathbf{F}_{T-i+1} \right) \\ = {} & \frac{\sigma_v^2}{1-\rho^2} \left(\mathbf{F}'_{T-i+1}\mathbf{M}_i\mathbf{F}_{T-i+1} \right) = \frac{\sigma_v^2}{(1-\rho^2)^2} + O(T^{-1}), \end{aligned}$$

which implies that $E(D_i) = \frac{(T-i+1)\sigma_v^2}{1-\rho^2} - \frac{\sigma_v^2}{(1-\rho)^2} + O(T^{-1})$.

Consider $E(N_{i,k}D_i)$, write

$$\begin{aligned}
E(N_{i,k}D_i) &= E(\mathbf{x}'_{-(i-k+1)}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{-(i-k+1)}\mathbf{x}'_{-(i-k+1)}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}) \\
&= 2E(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}) \\
&\quad +2E(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}) \\
&\quad +E(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k}\cdot\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}) \\
&\quad +E(\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}\cdot\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}) \\
&\quad +E(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k}\cdot\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}).
\end{aligned}$$

For the first term on the r.h.s., by the identity $\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D} = \mathbf{M}_i\mathbf{F}_{T-i+1}(x_0 - c/(1-\rho))$ and using Lemma 7.2 (4), we have

$$\begin{aligned}
&E(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}) \\
= \;& \sigma_v^2 \mathrm{tr}\left\{E(\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k})\right\} \\
= \;& \frac{\sigma_v^4}{1-\rho^2}\mathrm{tr}\left\{\mathbf{F}'_{T-i+1}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{F}_{T-i+1}\right\} \\
= \;& \frac{\rho^k\sigma_v^4}{(1-\rho^2)^3} + O(T^{-1}).
\end{aligned}$$

For the second term on the r.h.s., note that $\mathbf{C}_{T-i+k}\boldsymbol{\ell}_{T-i+k} = (\boldsymbol{\ell}_{T-i+k} - \mathbf{F}_{T-i+k})/(1-\rho)$ and thus $E(\mathbf{x}_{T-i+k,D}) = \boldsymbol{\ell}_{T-i+k}\frac{c}{1-\rho}$ and $\mathbf{M}_i\mathbf{J}_{i,k}E(\mathbf{x}_{T-i+k,D}) = 0$, we have

$$E(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}) = 0.$$

Similarly, the third term on the r.h.s is also equal to zero. For the fourth term, we have

$$\begin{aligned}
&E(\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}\cdot\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{i,k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}) \\
= \;& E(\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D})\cdot E(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{i,k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}) \\
= \;& \left(\frac{\sigma_v^2}{(1-\rho^2)^2} + O(T^{-1})\right)\left(-\frac{\sigma_v^2}{1-\rho} + \frac{\sigma_v^2(k-1)}{(T-i+1)(1-\rho)} + \frac{\sigma_v^2}{(T-i+1)(1-\rho)^2} + o(T^{-1})\right) \\
= \;& -\frac{\sigma_v^4}{(1-\rho^2)^2(1-\rho)} + O(T^{-1}).
\end{aligned}$$

Finally, for the fifth term, by using Lemma 7.1 (2) and using Lemmas 7.2 (1), (2), (5), and (6), we have

$$E(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k}\cdot\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k})/\sigma_v^4$$

$$
\begin{aligned}
&= -\frac{(T-i+1)(1+\rho-2\rho^k)}{(1-\rho^2)^2} + \left(\frac{\rho^2+3\rho+2-4\rho^k}{(1-\rho^2)^3} + \frac{(k-1)(1+\rho-2\rho^k)}{(1-\rho^2)^2}\right. \\
&\quad \left. -\frac{1}{(1-\rho)^3} - \frac{\gamma_2}{(1-\rho)(1-\rho^2)}\right) + O(T^{-1}).
\end{aligned}
$$

Combining the above analysis, we have

$$
\begin{aligned}
E(N_{i,k}D_i) &= -\frac{(T-i+1)(1+\rho-2\rho^k)\sigma_v^4}{(1-\rho^2)^2} - \sigma_v^4\left(\frac{\rho^3+2\rho^2+\rho+2\rho^k}{(1-\rho^2)^3}\right. \\
&\quad \left. -\frac{(k-1)(1+\rho-2\rho^k)}{(1-\rho^2)^2} + \frac{\gamma_2}{(1-\rho)(1-\rho^2)}\right) + O(T^{-1}).
\end{aligned}
$$

Next, consider $E(N_{i,k}D_i^2)$. Since $(E(D_i))^3$ is of order $O(T^3)$, it is sufficient to consider the terms of order at least $O(T)$. Then, we have

$$
\begin{aligned}
E(N_{i,k}D_i^2) &= E\left[(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k})^2\mathbf{v}'_{i,k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}\right] \\
&\quad +2E(\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}) \\
&\quad \times E(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k}\cdot\mathbf{v}'_{i,k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}) \\
&\quad +4E\left[\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k}\right. \\
&\quad \left.\times\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}\right] + O(1).
\end{aligned}
$$

For the first term on the r.h.s., by using Lemma 7.1 (3), Lemmas 7.2 (1), (2), (6), (9) and (10) and Lemma 7.3, we have

$$
\begin{aligned}
&E\left[(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k})^2\mathbf{v}'_{i,k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}\right] \\
&= \frac{(T-i+1)^2(4\rho^k-1-\rho)\sigma_v^6}{(1-\rho^2)^3} + (T-i+1)\sigma_v^6\left[\frac{(4k+12)\rho^k+(k-1)(1+\rho)}{(1-\rho^2)^3}\right. \\
&\quad \left.+\frac{20\rho^{k+2}-8\rho^{k+1}-16\rho^k-4\rho^3-7\rho^2-4\rho-1}{(1-\rho^2)^4}\right] + \frac{4\gamma_1^2(T-i+1)\rho^{2k}}{(1-\rho^3)(1-\rho^2)^2} \\
&\quad +\frac{(8\rho^k-3-3\rho)\gamma_2\sigma_v^6(T-i+1)}{(1-\rho^2)^3}.
\end{aligned}
$$

For the second term on the r.h.s., by using the above analyses, we have

$$
\begin{aligned}
&E(\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}) \\
&\times E(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k}\cdot\mathbf{v}'_{i,k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}) \\
&= \left(\frac{\sigma_v^2}{(1-\rho^2)^2} + O(T^{-1})\right) \times \left(-\frac{(T-i+1)(1+\rho-2\rho^k)\sigma_v^4}{(1-\rho^2)^2} + O(1)\right) \\
&= -\frac{(T-i+1)(1+\rho-2\rho^k)\sigma_v^6}{(1-\rho^2)^4} + O(1).
\end{aligned}
$$

For the third term on the r.h.s., by using the identity $\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D} = \mathbf{M}_i\mathbf{F}_{T-i+1}(x_0 - c/(1-\rho))$, Lemma 7.1 (2) and Lemmas 7.2 (2), (4), (7) and (8), we have

$$\begin{aligned}
& E\left[\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k}\right. \\
& \left.\times\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}\right] \\
= \; & \frac{\sigma_v^2}{1-\rho^2}E\left[\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k}\right. \\
& \left.\times\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{F}_{T-i+1}\mathbf{F}'_{T-i+1}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}\right] \\
= \; & \frac{(T-i+1)\sigma_v^6\rho^k}{(1-\rho^2)^4} + O(1).
\end{aligned}$$

Combining the above results, we have

$$\begin{aligned}
E(N_{i,k}D_i^2) \; = \; & \frac{(T-i+1)^2(4\rho^k-1-\rho)\sigma_v^6}{(1-\rho^2)^3} + (T-i+1)\sigma_v^6\left[\frac{(4k+12)\rho^k+(k-1)(1+\rho)}{(1-\rho^2)^3}\right. \\
& \left.+\frac{20\rho^{k+2}-8\rho^{k+1}-8\rho^k-4\rho^3-7\rho^2-6\rho-3}{(1-\rho^2)^4}\right] + \frac{4\gamma_1^2(T-i+1)\rho^{2k}}{(1-\rho^3)(1-\rho^2)^2} \\
& +\frac{(8\rho^k-3-3\rho)\gamma_2\sigma_v^6(T-i+1)}{(1-\rho^2)^3}.
\end{aligned}$$

Next, consider $E(N_{i,k}D_i^3)$. Since $(E(D_i))^4$ is of order $O(T^4)$, it is sufficient to consider the terms of order at least $O(T^2)$. Then, we have

$$\begin{aligned}
E(N_{i,k}D_i^3) \; = \; & E\left[(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k})^3\mathbf{v}'_{i,k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}\right] \\
& +3E(\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}) \\
& \times E((\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k})^2\cdot\mathbf{v}'_{i,k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}) \\
& +6E\left[(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k})^2\right. \\
& \left.\times\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}\right] + O(T).
\end{aligned}$$

For the first term on the r.h.s., by using Lemma 7.1 (4), Lemmas 7.2 (1), (2), (6) and (9)–(11) and Lemmas 7.3–7.4, we have

$$\begin{aligned}
& E\left[(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k})^3\mathbf{v}'_{i,k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}\right] \\
= \; & \frac{(T-i+1)^3\sigma_v^8(6\rho^k-\rho-1)}{(1-\rho^2)^4} + (T-i+1)^2\sigma_v^8\left[\frac{(54-18k)\rho^{k+2}+(18k+18)\rho^k}{(1-\rho^2)^5}\right. \\
& \left.+\frac{(-3-6\rho^2)}{(1-\rho^2)^4(1-\rho)} + \frac{1-12\rho^k+(k-1)(1-\rho)}{(1-\rho^2)^3(1-\rho)^2} - \frac{3}{(1-\rho^2)^2(1-\rho)^3}\right] \\
& +\frac{\gamma_2(T-i+1)^2\sigma_v^8(30\rho^k-6-6\rho)}{(1-\rho^2)^4} + \frac{12\gamma_1^2\rho^{2k}(T-i+1)^2\sigma_v^8}{(1-\rho^2)^3(1-\rho^3)} + O(T).
\end{aligned}$$

For the second term on the r.h.s., by using the above analyses, we have

$$\begin{aligned}
& E(\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}) \\
& \times E((\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k})^2 \cdot \mathbf{v}'_{i,k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}) \\
= & \left(\frac{\sigma_v^2}{(1-\rho^2)^2}+O(T^{-1})\right)\times\left(\frac{(T-i+1)^2(4\rho^k-1-\rho)\sigma_v^6}{(1-\rho^2)^3}+O(T)\right) \\
= & \frac{(T-i+1)^2(4\rho^k-1-\rho)\sigma_v^8}{(1-\rho^2)^5}+O(T).
\end{aligned}$$

For the third term on the r.h.s., by using the identity $\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D} = \mathbf{M}_i\mathbf{F}_{T-i+1}(x_0 - c/(1-\rho))$, Lemma 7.1 (3) and Lemmas 7.2 (2), (4) and (8), we have

$$\begin{aligned}
& E\left[(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k})^2\right. \\
& \left.\times\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{x}_{T-i+k,D}\mathbf{x}'_{T-i+k,D}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}\right] \\
= & \frac{\sigma_v^2}{1-\rho^2}E\left[(\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\mathbf{v}_{1,T-i+k})^2\right. \\
& \left.\times\mathbf{v}'_{1,T-i+k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{F}_{T-i+1}\mathbf{F}'_{T-i+1}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{v}_{1,T-i+k}\right] \\
= & \frac{(T-i+1)^2\sigma_v^8\rho^k}{(1-\rho^2)^5}+O(T).
\end{aligned}$$

Combining the above analyses, we have

$$\begin{aligned}
E(N_{i,k}D_i^3) = & \frac{(T-i+1)^3\sigma_v^8(6\rho^k-\rho-1)}{(1-\rho^2)^4}+(T-i+1)^2\sigma_v^8\left[\frac{(54-18k)\rho^{k+2}+(18k+36)\rho^k}{(1-\rho^2)^5}\right. \\
& \left.+\frac{(-6-6\rho^2)}{(1-\rho^2)^4(1-\rho)}+\frac{1-12\rho^k+(k-1)(1-\rho)}{(1-\rho^2)^3(1-\rho)^2}-\frac{3}{(1-\rho^2)^2(1-\rho)^3}\right] \\
& +\frac{\gamma_2(T-i+1)^2\sigma_v^8(30\rho^k-6-6\rho)}{(1-\rho^2)^4}+\frac{12\gamma_1^2\rho^{2k}(T-i+1)^2\sigma_v^8}{(1-\rho^2)^3(1-\rho^3)}+O(T).
\end{aligned}$$

By using the explicit formula of the terms $E(N_{i,k}D_i^l)$ for $l = 0,1,2,3$ and $E(D_i)$ and using the identity $(E(D_i))^q = \left(\frac{(T-i+1)\sigma_v^2}{1-\rho^2}\right)^q\left(1+qx+O(x^2)\right)$ with $x = -\frac{1+\rho}{(T-i+1)(1-\rho)}$ for $q \in \{-2,-3,-4\}$, we are able to obtain (here the sample size is $T-i+1$)

$$\begin{aligned}
E(\widehat{\rho}_i-\rho_i) = & -\frac{1}{T-i+1}\left[(1+\rho)\frac{1-\rho^i}{1-\rho}+2i\rho^i\right] \\
& -\frac{1}{(T-i+1)^2}\left(\frac{i\rho^i(3\rho^2+8\rho+3)}{1-\rho^2}-\frac{i(\rho+1)}{1-\rho}-(5i^2+1)\rho^i+1\right) \\
& +\frac{2\gamma_2 i\rho^i}{(T-i+1)^2}+\frac{4\gamma_1^2(1+\rho)}{(T-i+1)^2(1+\rho+\rho^2)}\frac{\rho^{i+1}(1-\rho^i)}{1-\rho}+o(T^{-2}).
\end{aligned}$$

To complete the proof, we now state the properties of $A_i^{DR}(\rho,h)$ for $1 \le i \le 6$. Consider the upper bound of $A_1^{DR}(\rho,h)$ first, for any $\rho\in[0,1]$ and $h\geq 1$, we have

$$\begin{aligned} A_1^{DR}(\rho,h) &= h\rho^h(1-\rho)(8\rho^2+8\rho)+(\rho^3+\rho^2+3\rho+1)(\rho^h-1)-\rho^h(1-\rho)(1-\rho^2) \\ &\leq 16h\rho^h(1-\rho)\leq \frac{16}{e} \end{aligned}$$

provided that

$$\max_{h\geq 1}\sup_{\rho\in[0,1]}(1-\rho)h\rho^h=\max_{h\geq 1}\left(1-\frac{1}{h+1}\right)^{h+1}\leq \frac{1}{e}. \tag{7.12}$$

Note that in (7.12) the function $\sup_{\rho\in[0,1]}(1-\rho)h\rho^h$ attains its maximum at $\rho=\frac{h}{h+1}$, that is, at a value of ρ close to 1. For the lower bound, note that when ρ is far away from 1 the term $h\rho^h(1-\rho)$ is small, we have

$$A_1^{DR}(\rho,h)\geq -(\rho^3+\rho^2+3\rho+1)\geq -6$$

for any $\rho\in[0,1]$. Similarly, for $A_2^{DR}(\rho,h)$, we have

$$A_2^{DR}(\rho,h) \quad\leq\quad 6h^2\rho^h(1-\rho)^2+6h\rho^h(1-\rho)\leq\frac{24}{e^2}+\frac{6}{e}$$

and

$$A_2^{DR}(\rho,h) \quad\geq\quad -(3\rho^3+\rho^2+\rho+1)\geq -6.$$

When ρ is close to zero, the negative term $-(3\rho^3+\rho^2+\rho+1)$ becomes dominant, causing $A_2^{DR}(\rho,h)$ to move toward their lower bounds. In addition, when ρ is close to 1 and h is large, $A_2^{DR}(\rho,h)$ converges $\frac{24}{e^2}$ to since $h^2\rho^h(1-\rho)^2\leq\frac{4}{e^2}$.

In addition, by using $\max_{h\geq 1}\sup_{\rho\in[0,1]}(1-\rho)h\rho^h\leq\frac{1}{e}$, it is easy to show $|A_i^{DR}(\rho,h)|\leq 1$ for $i=3,4,5,6$. Q.E.D.

7.4 Proof of Proposition 2.3

By (2.4) and (2.10), we have

$$
\begin{aligned}
\widehat{\beta}_h^{IM} - \beta_h &= \widehat{\beta}_1 \sum_{i=0}^{h-1} \widehat{\rho}^i - \beta_1 \sum_{i=0}^{h-1} \rho^i \\
&= \widehat{\beta}_1 \sum_{i=0}^{h-1} \widehat{\rho}^i - \beta_1 \sum_{i=0}^{h-1} \rho^i + \widehat{\beta}_1 \sum_{i=0}^{h-1} \rho^i - \widehat{\beta}_1 \sum_{i=0}^{h-1} \rho^i \\
&= \widehat{\beta}_1 \sum_{i=0}^{h-1} (\widehat{\rho}^i - \rho^i) + (\widehat{\beta}_1 - \beta_1) \sum_{i=0}^{h-1} \rho^i.
\end{aligned}
$$

Using $\widehat{\beta}_1 - \beta_1 = \frac{\sigma_{uv}}{\sigma_v^2}(\widehat{\rho} - \rho) + \frac{\sum_{t=1}^T \tilde{x}_{t-1} e_t}{\sum_{t-1}^T \tilde{x}_{t-1}^2}$, we have

$$
\begin{aligned}
\widehat{\beta}_h^{IM} - \beta_h &= \left[\beta_1 + \frac{\sigma_{uv}}{\sigma_v^2}(\widehat{\rho} - \rho) + \frac{\sum_{t=1}^T \tilde{x}_{t-1} e_t}{\sum_{t=1}^T \tilde{x}^2}\right] \sum_{i=0}^{h-1} (\widehat{\rho}^i - \rho^i) + \left[\frac{\sigma_{uv}}{\sigma_v^2}(\widehat{\rho} - \rho) + \frac{\tilde{x}_{t-1} e_t}{\tilde{x}^2}\right] \sum_{i=0}^{h-1} \rho^i \\
&= \beta_1 \sum_{i=0}^{h-1} (\widehat{\rho}^i - \rho^i) + \frac{\sigma_{uv}}{\sigma_v^2}(\widehat{\rho} - \rho) \sum_{i=0}^{h-1} (\widehat{\rho}^i - \rho^i) + \frac{\sum_{t=1}^T \tilde{x}_{t-1} e_t}{\sum_{t=1}^T \tilde{x}^2} \sum_{i=0}^{h-1} (\widehat{\rho}^i - \rho^i) \\
&+ \frac{\sigma_{uv}}{\sigma_v^2}(\widehat{\rho} - \rho) \sum_{i=0}^{h-1} \rho^i + \frac{\sum_{t=1}^T \tilde{x}_{t-1} e_t}{\sum_{t=1}^T \tilde{x}^2} \sum_{i=0}^{h-1} \rho^i \\
&= \beta_1 \sum_{i=0}^{h-1} (\widehat{\rho}^i - \rho^i) + \frac{\sigma_{uv}}{\sigma_v^2}(\widehat{\rho} - \rho) \sum_{i=0}^{h-1} \widehat{\rho}^i + \frac{\sum_{t=1}^T \tilde{x}_{t-1} e_t}{\sum_{t=1}^T \tilde{x}^2} \sum_{i=0}^{h-1} \widehat{\rho}^i \\
&= \beta_1 \sum_{i=0}^{h-1} (\widehat{\rho}^i - \rho^i) + \frac{\sigma_{uv}}{\sigma_v^2} \left\{ \sum_{i=0}^{h-1} \left[\widehat{\rho}^{i+1} - \rho^{i+1}\right] - \rho \sum_{i=0}^{h-1} \left[\widehat{\rho}^i - \rho^i\right] \right\} + \frac{\sum_{t=1}^T \tilde{x}_{t-1} e_t}{\sum_{t=1}^T \tilde{x}^2} \sum_{i=0}^{h-1} \widehat{\rho}^i
\end{aligned}
$$

Taking expectations on both sides, we have

$$
\begin{aligned}
E[\widehat{\beta}_h^{IM}] - \beta_h &= \beta_1 \sum_{i=0}^{h-1} E[\widehat{\rho}^i - \rho^i] + \frac{\sigma_{uv}}{\sigma_v^2} \left\{ \sum_{i=0}^{h-1} E\left[\widehat{\rho}^{i+1} - \rho^{i+1}\right] - \rho \sum_{i=0}^{h-1} E\left[\widehat{\rho}^i - \rho^i\right] \right\} \\
&= \left[\beta_1 + \frac{\sigma_{uv}}{\sigma_v^2}(1 - \rho)\right] \sum_{i=0}^{h-1} E\left[\widehat{\rho}^i - \rho^i\right] + \frac{\sigma_{uv}}{\sigma_v^2} E\left[\widehat{\rho}^h - \rho^h\right] \qquad (7.13)
\end{aligned}
$$

which is obtained by noting that

$$
\begin{aligned}
E\left(\frac{\sum_{t=1}^T \tilde{x}_{t-1} e_t}{\sum_{t=1}^T \tilde{x}^2} \widehat{\rho}^i\right) &= E\left\{ E\left(\frac{\sum_{t=1}^T \tilde{x}_{t-1} e_t}{\sum_{t=1}^T \tilde{x}^2} \widehat{\rho}^i | x_0, v_0, v_1, \ldots, v_T\right) \right\} \\
&= E\left\{ \sum_{t=1}^T \frac{\tilde{x}_{t-1} \widehat{\rho}^i}{\sum_{t=1}^T \tilde{x}^2} E(e_t | x_0, v_0, v_1, \ldots, v_T) \right\} \\
&= 0,
\end{aligned}
$$

because $E\left[e_{t+i} | x_0, v_0, v_1, \ldots, v_T\right] = E\left[e_{t+i} | v_{t+i}\right] = 0$. Q.E.D.

7.5 Proof of Proposition 2.4

Recall that

$$E[\widehat{\beta}_h^{IM,II}-\beta_h]=\beta_1\sum_{i=0}^{h-1}E[(\widehat{\rho})^i-\rho^i]+\frac{\sigma_{uv}}{\sigma_v^2}\left\{\sum_{i=0}^{h-1}E\left[(\widehat{\rho})^{i+1}-\rho^{i+1}\right]-\rho\sum_{i=0}^{h-1}E\left[(\widehat{\rho})^i-\rho^i\right]\right\}$$

and define $f(\rho)=\sum_{i=0}^{h-1}\rho^i=\frac{1-\rho^h}{1-\rho}$. By using a fourth-order Taylor expansion of $f(\rho)=\frac{1-\rho^h}{1-\rho}$ and $\rho f(\rho)$, we have

$$\begin{aligned} f(\widehat{\rho})-f(\rho) &= f_\rho(\rho)(\widehat{\rho}-\rho)+\tfrac{1}{2}f_{\rho\rho}(\rho)(\widehat{\rho}-\rho)^2+\tfrac{1}{6}f_{\rho\rho\rho}(\rho)(\widehat{\rho}-\rho)^3 \\ &\quad +\tfrac{1}{24}f_{\rho\rho\rho\rho}(\rho)(\widehat{\rho}-\rho)^4+o\big(|\widehat{\rho}-\rho|^4\big), \end{aligned}$$

and

$$\begin{aligned} \widehat{\rho}f(\widehat{\rho})-\rho f(\rho) &= [f(\rho)+\rho f_\rho(\rho)](\widehat{\rho}-\rho)+\tfrac{1}{2}[2f_\rho(\rho)+\rho f_{\rho\rho}(\rho)](\widehat{\rho}-\rho)^2 \\ &\quad +\tfrac{1}{6}[3f_{\rho\rho}(\rho)+\rho f_{\rho\rho\rho}(\rho)](\widehat{\rho}-\rho)^3+\tfrac{1}{24}[4f_{\rho\rho\rho}(\rho)+\rho f_{\rho\rho\rho\rho}(\rho)](\widehat{\rho}-\rho)^4+o\big(|\widehat{\rho}-\rho|^4\big), \end{aligned}$$

with

$$\begin{aligned} f_\rho(\rho) &= \frac{1-\rho^h}{(1-\rho)^2}-\frac{h\rho^{h-1}}{1-\rho}, \quad f_{\rho\rho}(\rho)=c\frac{2(1-\rho^h)}{(1-\rho)^3}-\frac{2h\rho^{h-1}}{(1-\rho)^2}-\frac{h(h-1)\rho^{h-2}}{1-\rho}, \\ f_{\rho\rho\rho}(\rho) &= \frac{6(1-\rho^h)}{(1-\rho)^4}-\frac{6h\rho^{h-1}}{(1-\rho)^3}-\frac{3h(h-1)\rho^{h-2}}{(1-\rho)^2}-\frac{h(h-1)(h-2)\rho^{h-3}}{1-\rho}, \\ f_{\rho\rho\rho\rho}(\rho) &= \frac{24(1-\rho^h)}{(1-\rho)^5}-\frac{24h\rho^{h-1}}{(1-\rho)^4}-\frac{12h(h-1)}{(1-\rho)^3}-\frac{4h(h-1)(h-2)\rho^{h-3}}{(1-\rho)^2} \\ &\quad -\frac{h(h-1)(h-2)(h-3)\rho^{h-4}}{1-\rho}. \end{aligned}$$

It follows that

$$\begin{aligned} & E\left(\widehat{\beta}_h^{IM,II}-\beta_h\right) \\ = \; & \beta_1\left[f_\rho(\rho)E(\widehat{\rho}-\rho)+\tfrac{1}{2}f_{\rho\rho}(\rho)E(\widehat{\rho}-\rho)^2+\tfrac{1}{6}f_{\rho\rho\rho}(\rho)E(\widehat{\rho}-\rho)^3+\tfrac{1}{24}f_{\rho\rho\rho\rho}(\rho)E(\widehat{\rho}-\rho)^4\right] \\ & +\frac{\sigma_{uv}}{\sigma_v^2}\left[f(\rho)E(\widehat{\rho}-\rho)+f_\rho(\rho)E(\widehat{\rho}-\rho)^2+\tfrac{1}{2}f_{\rho\rho}(\rho)E(\widehat{\rho}-\rho)^3+\tfrac{1}{6}f_{\rho\rho\rho}(\rho)E(\widehat{\rho}-\rho)^4\right]+o(T^{-2}). \end{aligned}$$

To complete the proof, it suffices to compute the term $E\left[(\widehat{\rho}-\rho)^k\right]$ for $k=1,2,3,4$.

Consider $E\left[(\widehat{\rho}-\rho)^2\right]$ first. According to the proof of Proposition 2.2, we have

$$\widehat{\rho}-\rho=\frac{N_{1,1}}{D_1}=a_{1,-1/2}+a_{1,-1}+a_{1,-3/2}+o_P(T^{-3/2}),$$

in which $a_{1,-l/2} = (-1)^{l-1} N_{1,1} \left(D_1 - E(D_1)\right)^{l-1} \left[E(D_1)\right]^{-l} = O_P(T^{-l/2})$, which follows that

$$
\begin{aligned}
E\left[(\widehat{\rho}-\rho)^2\right] &= E(a_{1,-1/2}^2 + a_{1,-1}^2 + 2a_{1,-1/2}a_{1,-1} + 2a_{1,-1/2}a_{1,-3/2}) + o(T^{-2}) \\
&= 6\frac{E(N_{1,1}^2)}{[E(D_1)]^2} - 8\frac{E(N_{1,1}^2 D_1)}{[E(D_1)]^3} + 3\frac{E(N_{1,1}^2 D_1^2)}{[E(D_1)]^4} + o(T^{-2}).
\end{aligned}
$$

For $E(N_{1,1}^2)$, write

$$
\begin{aligned}
& E(N_{1,1}^2) = E\left[\left(\mathbf{x}_{-1}'\mathbf{M}_1\mathbf{v}_{1,T}\right)^2\right] \\
= \; & E\left(\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{v}_{1,T}\cdot\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{v}_{1,T}\right) + E\left(\mathbf{v}_{1,T}'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\cdot\mathbf{v}_{1,T}'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right).
\end{aligned}
$$

For the first term, we have

$$
E\left(\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{v}_{1,T}\cdot\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{v}_{1,T}\right) = \sigma_v^2 E\left(\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{x}_{T,D}\right) = \frac{\sigma_v^4}{(1-\rho^2)^2} + O(T^{-1}).
$$

For the second term, replacing $\mathbf{K}_{1,1}$ with $\left(\mathbf{K}_{1,1} + \mathbf{C}_T'\mathbf{M}_1\right)/2$ and using Lemma 7.1(2) and 7.5(1), we have

$$
\begin{aligned}
& E\left(\mathbf{v}_{1,T}'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\cdot\mathbf{v}_{1,T}'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right)/\sigma_v^4 \\
= \; & \frac{T}{1-\rho^2} - \frac{1}{(1-\rho^2)^2} - \frac{1}{(1-\rho)^2} + O(T^{-1}),
\end{aligned}
$$

which implies that

$$
E(N_{1,1}^2) = \frac{T\sigma_v^4}{1-\rho^2} - \frac{\sigma_v^4}{(1-\rho)^2} + O(T^{-1}).
$$

For $E(N_{1,1}^2 D_1)$, since $(E(D_1))^3$ is of order $O(T^3)$, it is sufficient to consider the terms of order at least $O(T)$. Then, write

$$
\begin{aligned}
& E(N_{1,1}^2 D_1) = E\left[\mathbf{x}_{-1}'\mathbf{M}_1\mathbf{x}_{-1}\left(\mathbf{x}_{-1}'\mathbf{M}_1\mathbf{v}_{1,T}\right)^2\right] \\
= \; & E\left(\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{x}_{T,D}\right)E\left[\left(\mathbf{v}_{1,T}'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right)^2\right] + E\left(\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\left(\mathbf{v}_{1,T}'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right)^2\right) \\
& + E\left(\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\cdot\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{v}_{1,T}\cdot\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{v}_{1,T}\right) + O(1).
\end{aligned}
$$

For the first term, using Lemma 7.1(2) and Lemma 7.5(1), we have

$$
E\left(\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{x}_{T,D}\right)E\left[\left(\mathbf{v}_{1,T}'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right)^2\right] = \frac{T\sigma_v^6}{(1-\rho^2)^3} + O(1).
$$

For the second term, by using Lemma 7.1(3) and Lemma 7.5(2)–(4), we have

$$
\frac{E\left(\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\left(\mathbf{v}_{1,T}'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right)^2\right)}{\sigma_v^6} = \frac{T^2}{(1-\rho^2)^2} + \frac{T(2\rho^2-8\rho)}{(1-\rho^2)^3} + \frac{2\gamma_2 T}{(1-\rho^2)^2} + \frac{2\gamma_1^2 T\rho}{(1-\rho^2)(1-\rho^3)} + O(1).
$$

For the third term, by using Lemma 7.1(2) and Lemma 7.5(5), we have

$$\begin{aligned}
& E\left(\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\cdot\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{v}_{1,T}\cdot\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{v}_{1,T}\right)\\
= \;& E\left(\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\cdot\mathbf{v}_{1,T}'\mathbf{M}_1\mathbf{F}_T\mathbf{F}_T'\mathbf{M}_1\mathbf{v}_{1,T}\right)E\left[(x_0-c/(1-\rho))^2\right]\\
= \;& \frac{T\sigma_v^6}{(1-\rho^2)^3}+O(1),
\end{aligned}$$

which implied that

$$E(N_{1,1}^2D_1)=\frac{T^2\sigma_v^6}{(1-\rho^2)^2}+\frac{T(2\rho^2-8\rho+2)\sigma_v^6}{(1-\rho^2)^3}+\frac{2\gamma_2T\sigma_v^6}{(1-\rho^2)^2}+\frac{2\gamma_1^2T\rho\sigma_v^6}{(1-\rho^2)(1-\rho^3)}+O(1).$$

For $E(N_{1,1}^2D_1^2)$, since $(E(D_1))^4$ is of order $O(T^4)$, it is sufficient to consider the terms of order at least $O(T^2)$. Then, write

$$\begin{aligned}
& E(N_{1,1}^2D_1^2)=E\left[\left(\mathbf{x}_{-1}'\mathbf{M}_1\mathbf{x}_{-1}\right)^2\left(\mathbf{x}_{-1}'\mathbf{M}_1\mathbf{v}_{1,T}\right)^2\right]\\
= \;& 2E\left(\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{x}_{T,D}\right)E\left[\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\left(\mathbf{v}_{1,T}'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right)^2\right]\\
& +E\left[\left(\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right)^2\left(\mathbf{v}_{1,T}'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right)^2\right]\\
& +E\left[\left(\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right)^2\cdot\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{v}_{1,T}\cdot\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{v}_{1,T}\right]+O(T).
\end{aligned}$$

For the first term, by using Lemma 7.1(3) and Lemma 7.5, we have

$$E\left(\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{x}_{T,D}\right)E\left[\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\left(\mathbf{v}_{1,T}'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right)^2\right]=\frac{T^2\sigma_v^8}{(1-\rho^2)^4}+O(T).$$

For the second term, by using Lemma 7.1(4) and Lemmas 7.5–7.6, we have

$$\begin{aligned}
& E\left(\left(\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right)^2\left(\mathbf{v}_{1,T}'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right)^2\right)/\sigma_v^8\\
= \;& \frac{T^3}{(1-\rho^2)^3}+\frac{T^2(15\rho^2-14\rho+4)}{(1-\rho^2)^4}+\frac{5\gamma_2T^2}{(1-\rho^2)^3}+\frac{4\gamma_1^2T^2\rho}{(1-\rho^2)^2(1-\rho^3)}+O(T).
\end{aligned}$$

For the third term, by using Lemma 7.1(3) and Lemma 7.5, we have

$$\begin{aligned}
& E\left(\left(\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right)^2\cdot\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{v}_{1,T}\cdot\mathbf{x}_{T,D}'\mathbf{M}_1\mathbf{v}_{1,T}\right)\\
= \;& E\left(\left(\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\right)^2\cdot\mathbf{v}_{1,T}'\mathbf{M}_1\mathbf{F}_T\mathbf{F}_T'\mathbf{M}_1\mathbf{v}_{1,T}\right)E\left[(x_0-c/(1-\rho))^2\right]\\
= \;& \frac{T^2\sigma_v^8}{(1-\rho^2)^4}+O(T),
\end{aligned}$$

which follows that

$$E(N_{1,1}^2D_1^2)=\frac{T^3\sigma_v^8}{(1-\rho^2)^3}+\frac{T^2(15\rho^2-14\rho+6)\sigma_v^8}{(1-\rho^2)^4}+\frac{5\gamma_2T^2\sigma_v^8}{(1-\rho^2)^3}+\frac{4\gamma_1^2T^2\rho\sigma_v^8}{(1-\rho^2)^2(1-\rho^3)}+O(T).$$

Combing the above results and using $[E(D_1)]^q = \left(\frac{T\sigma_v^2}{1-\rho^2}\right)^q (1 + qx + O(x^2))$ with $x = -\frac{1+\rho}{T(1-\rho)}$ for $q \in \{-2, -3, -4\}$, we are able to obtain

$$E\left[(\widehat{\rho} - \rho)^2\right] = \frac{1-\rho^2}{T} + \frac{23\rho^2 + 10\rho - 4}{T^2} - \frac{\gamma_2(1-\rho^2)}{T^2} - \frac{4\gamma_1^2\rho(1-\rho)(1+\rho)^2}{T^2(1+\rho+\rho^2)}.$$

Similarly, we decompose $E\left[(\widehat{\rho} - \rho)^3\right]$ as

$$\begin{aligned} E\left[(\widehat{\rho} - \rho)^3\right] &= E(a_{1,-1/2}^3 + 3a_{1,-1/2}^2 a_{1,-1}) + o(T^{-2}) \\ &= 4\frac{E(N_{1,1}^3)}{[E(D_1)]^3} - 3\frac{E(N_{1,1}^3 D_1)}{[E(D_1)]^4} + o(T^{-2}). \end{aligned}$$

For $E(N_{1,1}^3)$, since $(E(D_1))^3$ is of order $O(T^3)$, it is sufficient to consider the terms of order at least $O(T)$. Then, by using Lemma 7.1(3) and Lemmas 7.5–7.6, we have

$$E(N_{1,1}^3) = E\left[\left(\mathbf{x}_{-1}'\mathbf{M}_1\mathbf{v}_{1,T}\right)^3\right] = E\left[\left(\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{M}_1\mathbf{v}_{1,T}\right)^3\right] + O(1) = -\frac{3T}{(1-\rho^2)(1+\rho)} + O(1).$$

For $E(N_{1,1}^3 D_1)$, since $(E(D_1))^4$ is of order $O(T^4)$, it is sufficient to consider the terms of order at least $O(T^2)$. Then, by using Lemma 7.1(4) and Lemmas 7.5–7.7, we have

$$\begin{aligned} E(N_{1,1}^3 D_1) &= E\left[\mathbf{x}_{-1}'\mathbf{M}_1\mathbf{x}_{-1}\left(\mathbf{x}_{-1}'\mathbf{M}_1\mathbf{v}_{1,T}\right)^3\right] = E\left[\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{K}_{1,1}\mathbf{v}_{1,T}\left(\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{M}_1\mathbf{v}_{1,T}\right)^3\right] + O(T) \\ &= \frac{T^2(9\rho - 3)}{(1-\rho^2)^3} + \frac{\gamma_1^2 T^2}{(1-\rho^2)(1-\rho)} + O(T). \end{aligned}$$

Combining the above analysis, we can obtain

$$E\left[(\widehat{\rho} - \rho)^3\right] = -\frac{(1-\rho^2)(15\rho + 3)}{T^2} - \frac{3\gamma_1^2(1-\rho^2)^2(1+\rho)}{T^2} + o(T^{-2}).$$

Finally, we consider $E\left[(\widehat{\rho} - \rho)^4\right]$, write

$$E\left[(\widehat{\rho} - \rho)^4\right] = E(a_{1,-1/2}^4) + o(T^{-2}) = \frac{E(N_{1,1}^4)}{[E(D_1)]^4} + o(T^{-2}).$$

For $E(N_{1,1}^4)$, since $(E(D_1))^4$ is of order $O(T^4)$, it is sufficient to consider the terms of order at least $O(T^2)$. Then, by using Lemma 7.1(4) and Lemmas 7.5–7.7, we have

$$\begin{aligned} E(N_{1,1}^4) &= E\left[\left(\mathbf{x}_{-1}'\mathbf{M}_1\mathbf{v}_{1,T}\right)^4\right] = E\left[\left(\mathbf{v}_{1,T}'\mathbf{C}_T'\mathbf{M}_1\mathbf{v}_{1,T}\right)^4\right] + O(T) \\ &= \frac{3T^2}{(1-\rho^2)^2} + O(T), \end{aligned}$$

which implies that

$$E\left[(\widehat{\rho} - \rho)^4\right] = \frac{3(1-\rho^2)^2}{T^2} + o(T^{-2}).$$

Combing the above results, we are able to compute the second-order bias of $\widehat{\beta}_h^{IM}$, that is,

$$\begin{aligned}
& E(\widehat{\beta}_h^{IM} - \beta_h) \\
= \; & -\frac{1}{T}\frac{\sigma_{uv}}{\sigma_v^2}\left[\frac{2(\rho-\rho^{h+1})}{1-\rho} + h\rho^{h-1}(1+\rho)\right] - \frac{1}{T^2}\frac{\sigma_{uv}}{\sigma_v^2}\frac{A_1^{IM}(\rho,h)}{(1-\rho)^2(1+\rho)} \\
& -\frac{\beta_1}{T}\left[\frac{2\rho(1-\rho^h)}{(1-\rho)^2} - \frac{2h\rho^h}{1-\rho} + \frac{h(h-1)(1+\rho)\rho^{h-2}}{2}\right] - \frac{\beta_1}{T^2}\frac{A_2^{IM}(\rho,h)}{(1-\rho)^3(1+\rho)} \\
& -\frac{1}{T^2}\frac{\sigma_{uv}}{\sigma_v^2}\left[\frac{\gamma_2 A_3^{IM}(\rho,h)}{1-\rho} + \frac{\gamma_1^2 A_4^{IM}(\rho,h)}{(1-\rho)(1+\rho+\rho^2)}\right] - \frac{\beta_1}{T^2}\left[\frac{\gamma_2 A_5^{IM}(\rho,h)}{(1-\rho)^2} + \frac{\gamma_1^2 A_6^{IM}(\rho,h)}{(1-\rho)^2(1+\rho+\rho^2)}\right] \\
& +o(T^{-2}),
\end{aligned}$$

where

$$\begin{aligned}
A_1^{IM}(\rho,h) &= (1-\rho^h)(-2\rho^3-3\rho^2+\rho+4) + h\rho^{h-1}(1-\rho^2)(11\rho^2-2\rho-4) \\
&\quad +\frac{1}{2}h(h-1)(h-2)\rho^{h-3}(1-\rho^2)^3 - 6h(h-1)\rho^{h-1}(1-\rho^2)^2, \\
A_2^{IM}(\rho,h) &= (\rho^h-1)(2\rho^3+3\rho^2-\rho-4) + h\rho^{h-1}(1-\rho)(2\rho^3+3\rho^2-\rho-4) \\
&\quad +\frac{1}{2}h(h-1)\rho^{h-2}(1-\rho)(1-\rho^2)(11\rho^2-2\rho-4) - 2h(h-1)(h-2)\rho^{h-2}(1-\rho^2)^2(1-\rho) \\
&\quad +\frac{1}{8}h(h-1)(h-2)(h-3)\rho^{h-4}(1-\rho^2)^3(1-\rho), \\
A_3^{IM}(\rho,h) &= (1-\rho^h)(1-\rho) - h\rho^{h-1}(1-\rho^2), \\
A_4^{IM}(\rho,h) &= (1-\rho^h)\left[3(1+\rho)^3(1-\rho^3) + 4(\rho^2+\rho)\right] - (1-\rho^2)h\rho^{h-1}\left[3(1+\rho)^2(1-\rho^3) + 4(\rho^2+\rho)\right] \\
&\quad -\frac{3}{2}h(h-1)\rho^{h-2}(1-\rho^2)^2(1+\rho)(1+\rho+\rho^2), \\
A_5^{IM}(\rho,h) &= (1-\rho^h)(1-\rho) - h\rho^{h-1}(1-\rho)^2 - \frac{1}{2}h(h-1)\rho^{h-2}(1-\rho^2)(1-\rho), \\
A_6^{IM}(\rho,h) &= (1-\rho^h)[4\rho^2+4\rho+3(1+\rho)^3(1+\rho+\rho^2)] - (1-\rho^2)h\rho^{h-1}[4\rho+3(1+\rho)^2(1+\rho+\rho^2)] \\
&\quad -\frac{1}{2}h(h-1)\rho^{h-2}(1-\rho^2)^2[4\rho+3(1+\rho)(1+\rho+\rho^2)] \\
&\quad -\frac{1}{2}h(h-1)(h-2)\rho^{h-3}(1-\rho^2)^2(1-\rho^3).
\end{aligned}$$

To complete the proof, we state the properties of $A_i^{IM}(\rho,h)$ for $1 \le i \le 6$. Consider $A_1^{IM}(\rho,h)$ first. The upper bound for terms like $h\rho^{h-1}(1-\rho^2)$ is found via differentiation. Setting the derivative to zero gives a unique interior maximizer at $\rho^* = \sqrt{(h-1)/(h+1)}$. Thus, for any $\rho \in [0,1]$ and $h \ge 1$, we have

$$\sup_{\rho\in[0,1]} h\rho^{h-1}(1-\rho^2) \le \frac{2h}{h+1}\left(\frac{h-1}{h+1}\right)^{(h-1)/2} \le 1.$$

Similarly, for any $\rho \in [0,1]$ and $h \ge 1$, one can show that the three terms

$$h(h-1)\rho^{h-2}(1-\rho^2)^2, \quad h(h-1)(h-2)\rho^{h-3}(1-\rho^2)^3, \quad h(h-1)(h-2)(h-3)\rho^{h-4}(1-\rho^2)^4$$

are all uniformly bounded in (ρ, h), which completes the proof. Q.E.D.

7.6 Technical Lemmas

Suppose that v_t has finite moments $m_j = E(v_t^j)$ up to the eighth order:

$$
\begin{aligned}
&m_1 = 0, \quad m_2 = \sigma_v^2, \quad m_3 = \gamma_1\sigma_v^3, \quad m_4 = (\gamma_2+3)\sigma_v^4, \quad m_5 = (\gamma_3+10\gamma_1)\sigma_v^5 \\
&m_6 = (\gamma_4+10\gamma_1^2+15\gamma_2+15)\sigma_v^6, \quad m_7 = (\gamma_5+21\gamma_3+35\gamma_2\gamma_1+105\gamma_1)\sigma_v^7 \\
&m_8 = (\gamma_6+28\gamma_4+56\gamma_3\gamma_1+35\gamma_2^2+210\gamma_2+280\gamma_1^2+105)\sigma_v^8,
\end{aligned} \tag{7.14}
$$

in which γ_1 and γ_2 are the Pearson's measures of skewness and kurtosis, $\gamma_3, ..., \gamma_6$ can be regarded as measures for deviation from normality since $\gamma_3, ..., \gamma_6$ are all zeros for a normal distribution.

Lemma 7.1. *Suppose that $\mathbf{v} = [v_1, ..., v_T]^\top$ has mean zero (with ts elements i.i.d.) and has moments as specified in (7.14). Then, for any symmetric matrix $\mathbf{N}_i$, we have*

1. $E(\mathbf{v}\cdot\mathbf{v}'\mathbf{N}_1\mathbf{v})/\sigma_v^3 = \gamma_1(\mathbf{I}_T\odot\mathbf{N}_1)\ell_T$*, where $\odot$ denotes Hadamard product, ℓ_T is a T-dimension vector of ones and $\mathbf{I}_T$ is a $T\times T$-dimension identity matrix;*
2. $E(\mathbf{v}'\mathbf{N}_1\mathbf{v}\cdot\mathbf{v}'\mathbf{N}_2\mathbf{v})/\sigma_v^4 = \gamma_2\mathrm{tr}(\mathbf{N}_1\odot\mathbf{N}_2)+\mathrm{tr}(\mathbf{N}_1)\mathrm{tr}(\mathbf{N}_2)+2\mathrm{tr}(\mathbf{N}_1\mathbf{N}_2)$*;*
3.

$$
\begin{aligned}
& E(\mathbf{v}'\mathbf{N}_1\mathbf{v}\cdot\mathbf{v}'\mathbf{N}_2\mathbf{v}\cdot\mathbf{v}'\mathbf{N}_3\mathbf{v})/\sigma_v^6 \\
= \;& \mathrm{tr}\{\mathbf{N}_1\}\,\mathrm{tr}\{\mathbf{N}_2\}\,\mathrm{tr}\{\mathbf{N}_3\}+2\mathrm{tr}\{\mathbf{N}_1\}\,\mathrm{tr}\{\mathbf{N}_2\mathbf{N}_3\}+2\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_2\}\,\mathrm{tr}\{\mathbf{N}_3\} \\
& +2\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_3\}\,\mathrm{tr}\{\mathbf{N}_2\}+8\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_2\mathbf{N}_3\}+\gamma_4\mathrm{tr}\{\mathbf{N}_1\odot\mathbf{N}_2\odot\mathbf{N}_3\}+\gamma_2\mathrm{tr}\{\mathbf{N}_3\}\mathrm{tr}\{\mathbf{N}_1\odot\mathbf{N}_2\} \\
& +\gamma_2\mathrm{tr}\mathbf{N}_1\mathrm{tr}\{\mathbf{N}_3\odot\mathbf{N}_2\}+\gamma_2\mathrm{tr}\mathbf{N}_2\mathrm{tr}\{\mathbf{N}_3\odot\mathbf{N}_1\}+2\gamma_2\mathrm{tr}\{(\mathbf{N}_3\odot(\mathbf{N}_1\mathbf{N}_2+\mathbf{N}_2\mathbf{N}_1))\} \\
& +2\gamma_2\mathrm{tr}\{\mathbf{N}_2\odot(\mathbf{N}_3\mathbf{N}_1+\mathbf{N}_1\mathbf{N}_3)\}+2\gamma_2\mathrm{tr}\{\mathbf{N}_1\odot(\mathbf{N}_3\mathbf{N}_2+\mathbf{N}_2\mathbf{N}_3)\} \\
& +4\gamma_1^2\mathrm{tr}\{(\mathbf{N}_1\odot\mathbf{N}_2)\mathbf{N}_3\}+2\gamma_1^2\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{N}_2)\mathbf{N}_3(\mathbf{I}_T\odot\mathbf{N}_1)\boldsymbol{\ell}_T \\
& +2\gamma_1^2\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{N}_3)\mathbf{N}_2(\mathbf{I}_T\odot\mathbf{N}_1)\boldsymbol{\ell}_T+2\gamma_1^2\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{N}_2)\mathbf{N}_1(\mathbf{I}_T\odot\mathbf{N}_3)\boldsymbol{\ell}_T;
\end{aligned}
$$

4.

$$
\begin{aligned}
& E(\mathbf{v}'\mathbf{N}_1\mathbf{v}\cdot\mathbf{v}'\mathbf{N}_2\mathbf{v}\cdot\mathbf{v}'\mathbf{N}_3\mathbf{v}\cdot\mathbf{v}'\mathbf{N}_4\mathbf{v})/\sigma_v^8 \\
= \;& \mathrm{tr}\{\mathbf{N}_1\}\,\mathrm{tr}\{\mathbf{N}_2\}\,\mathrm{tr}\{\mathbf{N}_3\}\,\mathrm{tr}\{\mathbf{N}_4\}+2\,[\mathrm{tr}\{\mathbf{N}_1\}\,\mathrm{tr}\{\mathbf{N}_2\}\,\mathrm{tr}\{\mathbf{N}_3\mathbf{N}_4\}+\mathrm{tr}\{\mathbf{N}_1\}\,\mathrm{tr}\{\mathbf{N}_3\}\,\mathrm{tr}\{\mathbf{N}_2\mathbf{N}_4\} \\
& +\mathrm{tr}\{\mathbf{N}_1\}\,\mathrm{tr}\{\mathbf{N}_4\}\,\mathrm{tr}\{\mathbf{N}_2\mathbf{N}_3\}+\mathrm{tr}\{\mathbf{N}_2\}\,\mathrm{tr}\{\mathbf{N}_3\}\,\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_4\} \\
& +\mathrm{tr}\{\mathbf{N}_2\}\,\mathrm{tr}\{\mathbf{N}_4\}\,\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_3\}+\mathrm{tr}\{\mathbf{N}_3\}\,\mathrm{tr}\{\mathbf{N}_4\}\,\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_2\}] \\
& +4\,[\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_2\}\,\mathrm{tr}\{\mathbf{N}_3\mathbf{N}_4\}+\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_3\}\,\mathrm{tr}\{\mathbf{N}_2\mathbf{N}_4\}+\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_4\}\,\mathrm{tr}\{\mathbf{N}_2\mathbf{N}_3\}] \\
& +8\,[\mathrm{tr}\{\mathbf{N}_1\}\,\mathrm{tr}\{\mathbf{N}_2\mathbf{N}_3\mathbf{N}_4\}+\mathrm{tr}\{\mathbf{N}_2\}\,\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_3\mathbf{N}_4\}+ \\
& +\mathrm{tr}\{\mathbf{N}_3\}\,\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_2\mathbf{N}_4\}+\mathrm{tr}\{\mathbf{N}_4\}\,\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_2\mathbf{N}_3\}] \\
& +16\,[\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_3\mathbf{N}_4\mathbf{N}_2\}+\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_4\mathbf{N}_2\mathbf{N}_3\}+\mathrm{tr}\{\mathbf{N}_1\mathbf{N}_4\mathbf{N}_3\mathbf{N}_2\}] \\
& +\gamma_2 f_{\gamma_2}+\gamma_4 f_{\gamma_4}+\gamma_6 f_{\gamma_6}+\gamma_1^2 f_{\gamma_1^2}+\gamma_2^2 f_{\gamma_2^2}+\gamma_1\gamma_3 f_{\gamma_1\gamma_3},
\end{aligned}
$$

in which

$$
f_{\gamma_2} \;=\; \mathrm{tr}(\mathbf{N}_1)\mathrm{tr}(\mathbf{N}_2)\mathrm{tr}(\mathbf{N}_3\odot\mathbf{N}_4)+\mathrm{tr}(\mathbf{N}_1)\mathrm{tr}(\mathbf{N}_3)\mathrm{tr}(\mathbf{N}_2\odot\mathbf{N}_4)+\mathrm{tr}(\mathbf{N}_1)\mathrm{tr}(\mathbf{N}_4)\mathrm{tr}(\mathbf{N}_2\odot\mathbf{N}_3)+
$$

$$
\begin{aligned}
& +\mathrm{tr}(\mathbf{N}_2)\mathrm{tr}(\mathbf{N}_3)\mathrm{tr}(\mathbf{N}_1 \odot \mathbf{N}_4) + \mathrm{tr}(\mathbf{N}_2)\mathrm{tr}(\mathbf{N}_4)\mathrm{tr}(\mathbf{N}_1 \odot \mathbf{N}_3) + \mathrm{tr}(\mathbf{N}_3)\mathrm{tr}(\mathbf{N}_4)\mathrm{tr}(\mathbf{N}_1 \odot \mathbf{N}_2) \\
& +2\big[\ell_T'(\mathbf{N}_1 \odot \mathbf{N}_2)\ell_T\mathrm{tr}(\mathbf{N}_3 \odot \mathbf{N}_4) + \ell_T'(\mathbf{N}_1 \odot \mathbf{N}_3)\ell_T\mathrm{tr}(\mathbf{N}_2 \odot \mathbf{N}_4) + \ell_T'(\mathbf{N}_1 \odot \mathbf{N}_4)\ell_T\mathrm{tr}(\mathbf{N}_2 \odot \mathbf{N}_3) \\
& +\ell_T'(\mathbf{N}_2 \odot \mathbf{N}_3)\ell_T\mathrm{tr}(\mathbf{N}_1 \odot \mathbf{N}_4) + \ell_T'(\mathbf{N}_2 \odot \mathbf{N}_4)\ell_T\mathrm{tr}(\mathbf{N}_1 \odot \mathbf{N}_3) + \ell_T'(\mathbf{N}_3 \odot \mathbf{N}_4)\ell_T\mathrm{tr}(\mathbf{N}_1 \odot \mathbf{N}_2)\big] \\
& +4\big[\mathrm{tr}(\mathbf{N}_1)\mathrm{tr}(\mathbf{N}_2 \odot (\mathbf{N}_3\mathbf{N}_4)) + \mathrm{tr}(\mathbf{N}_1)\mathrm{tr}(\mathbf{N}_3 \odot (\mathbf{N}_2\mathbf{N}_4)) + \mathrm{tr}(\mathbf{N}_1)\mathrm{tr}(\mathbf{N}_4 \odot (\mathbf{N}_2\mathbf{N}_3)) \\
& +\mathrm{tr}(\mathbf{N}_2)\mathrm{tr}(\mathbf{N}_1 \odot (\mathbf{N}_3\mathbf{N}_4)) + \mathrm{tr}(\mathbf{N}_2)\mathrm{tr}(\mathbf{N}_3 \odot (\mathbf{N}_1\mathbf{N}_4)) + \mathrm{tr}(\mathbf{N}_2)\mathrm{tr}(\mathbf{N}_4 \odot (\mathbf{N}_1\mathbf{N}_3)) \\
& +\mathrm{tr}(\mathbf{N}_3)\mathrm{tr}(\mathbf{N}_1 \odot (\mathbf{N}_2\mathbf{N}_4)) + \mathrm{tr}(\mathbf{N}_3)\mathrm{tr}(\mathbf{N}_2 \odot (\mathbf{N}_1\mathbf{N}_4)) + \mathrm{tr}(\mathbf{N}_3)\mathrm{tr}(\mathbf{N}_4 \odot (\mathbf{N}_1\mathbf{N}_2)) \\
& +\mathrm{tr}(\mathbf{N}_4)\mathrm{tr}(\mathbf{N}_1 \odot (\mathbf{N}_2\mathbf{N}_3)) + \mathrm{tr}(\mathbf{N}_4)\mathrm{tr}(\mathbf{N}_2 \odot (\mathbf{N}_1\mathbf{N}_3)) + \mathrm{tr}(\mathbf{N}_4)\mathrm{tr}(\mathbf{N}_3 \odot (\mathbf{N}_1\mathbf{N}_2))\big] \\
& +8\big[\mathrm{tr}((\mathbf{I}_T \odot \mathbf{N}_1)\mathbf{N}_2\mathbf{N}_3\mathbf{N}_4) + \mathrm{tr}((\mathbf{I}_T \odot \mathbf{N}_1)\mathbf{N}_4\mathbf{N}_2\mathbf{N}_3) + \mathrm{tr}((\mathbf{I}_T \odot \mathbf{N}_1)\mathbf{N}_3\mathbf{N}_2\mathbf{N}_4) \\
& +\mathrm{tr}((\mathbf{I}_T \odot \mathbf{N}_2)\mathbf{N}_1\mathbf{N}_3\mathbf{N}_4) + \mathrm{tr}((\mathbf{I}_T \odot \mathbf{N}_2)\mathbf{N}_1\mathbf{N}_4\mathbf{N}_3) + \mathrm{tr}((\mathbf{I}_T \odot \mathbf{N}_2)\mathbf{N}_3\mathbf{N}_1\mathbf{N}_4) \\
& +\mathrm{tr}((\mathbf{I}_T \odot \mathbf{N}_3)\mathbf{N}_1\mathbf{N}_2\mathbf{N}_4) + \mathrm{tr}((\mathbf{I}_T \odot \mathbf{N}_3)\mathbf{N}_1\mathbf{N}_4\mathbf{N}_2) + \mathrm{tr}((\mathbf{I}_T \odot \mathbf{N}_3)\mathbf{N}_2\mathbf{N}_1\mathbf{N}_4) \\
& +\mathrm{tr}((\mathbf{I}_T \odot \mathbf{N}_4)\mathbf{N}_1\mathbf{N}_2\mathbf{N}_3) + \mathrm{tr}((\mathbf{I}_T \odot \mathbf{N}_4)\mathbf{N}_1\mathbf{N}_3\mathbf{N}_2) + \mathrm{tr}((\mathbf{I}_T \odot \mathbf{N}_4)\mathbf{N}_2\mathbf{N}_1\mathbf{N}_3)\big] \\
& +16\big[\ell_T'(\mathbf{I}_T \odot (\mathbf{N}_1\mathbf{N}_2))(\mathbf{I}_T \odot (\mathbf{N}_3\mathbf{N}_4))\ell_T + \ell_T'(\mathbf{I}_T \odot (\mathbf{N}_1\mathbf{N}_3))(\mathbf{I}_T \odot (\mathbf{N}_2\mathbf{N}_4))\ell_T \\
& +\ell_T'(\mathbf{I}_T \odot (\mathbf{N}_1\mathbf{N}_4))(\mathbf{I}_T \odot (\mathbf{N}_2\mathbf{N}_3))\ell_T\big], \\
f_{\gamma_4} \quad = \quad & \mathrm{tr}(\mathbf{N}_1)\mathrm{tr}(\mathbf{N}_2 \odot \mathbf{N}_3 \odot \mathbf{N}_4) + \mathrm{tr}(\mathbf{N}_2)\mathrm{tr}(\mathbf{N}_1 \odot \mathbf{N}_3 \odot \mathbf{N}_4) + \mathrm{tr}(\mathbf{N}_3)\mathrm{tr}(\mathbf{N}_1 \odot \mathbf{N}_2 \odot \mathbf{N}_4) \\
& +\mathrm{tr}(\mathbf{N}_4)\mathrm{tr}(\mathbf{N}_1 \odot \mathbf{N}_2 \odot \mathbf{N}_3) + 4\big[\mathrm{tr}(\mathbf{N}_1 \odot \mathbf{N}_2 \odot (\mathbf{N}_3\mathbf{N}_4)) + \mathrm{tr}(\mathbf{N}_1 \odot \mathbf{N}_3 \odot (\mathbf{N}_2\mathbf{N}_4)) \\
& +\mathrm{tr}(\mathbf{N}_1 \odot \mathbf{N}_4 \odot (\mathbf{N}_2\mathbf{N}_3)) + \mathrm{tr}(\mathbf{N}_2 \odot \mathbf{N}_3 \odot (\mathbf{N}_1\mathbf{N}_4)) \\
& +\mathrm{tr}(\mathbf{N}_2 \odot \mathbf{N}_4 \odot (\mathbf{N}_1\mathbf{N}_3)) + \mathrm{tr}(\mathbf{N}_3 \odot \mathbf{N}_4 \odot (\mathbf{N}_1\mathbf{N}_2))\big], \\
f_{\gamma_6} \quad = \quad & \mathrm{tr}(\mathbf{N}_1 \odot \mathbf{N}_2 \odot \mathbf{N}_3 \odot \mathbf{N}_4), \\
f_{\gamma_1^2} \quad = \quad & 2\Big[\ell_T'((\mathbf{I}_T \odot \mathbf{N}_2)\mathbf{N}_3(\mathbf{I}_T \odot \mathbf{N}_4))\ell_T\mathrm{tr}(\mathbf{N}_1) + \ell_T'((\mathbf{I}_T \odot \mathbf{N}_2)\mathbf{N}_4(\mathbf{I}_T \odot \mathbf{N}_3))\ell_T\mathrm{tr}(\mathbf{N}_1) \\
& +\ell_T'((\mathbf{I}_T \odot \mathbf{N}_1)\mathbf{N}_3(\mathbf{I}_T \odot \mathbf{N}_4))\ell_T\mathrm{tr}(\mathbf{N}_2) + \ell_T'((\mathbf{I}_T \odot \mathbf{N}_1)\mathbf{N}_4(\mathbf{I}_T \odot \mathbf{N}_3))\ell_T\mathrm{tr}(\mathbf{N}_2) \\
& +\ell_T'((\mathbf{I}_T \odot \mathbf{N}_1)\mathbf{N}_2(\mathbf{I}_T \odot \mathbf{N}_4))\ell_T\mathrm{tr}(\mathbf{N}_3) + \ell_T'((\mathbf{I}_T \odot \mathbf{N}_1)\mathbf{N}_4(\mathbf{I}_T \odot \mathbf{N}_2))\ell_T\mathrm{tr}(\mathbf{N}_3) \\
& +\ell_T'((\mathbf{I}_T \odot \mathbf{N}_1)\mathbf{N}_2(\mathbf{I}_T \odot \mathbf{N}_3))\ell_T\mathrm{tr}(\mathbf{N}_4) + \ell_T'((\mathbf{I}_T \odot \mathbf{N}_1)\mathbf{N}_3(\mathbf{I}_T \odot \mathbf{N}_2))\ell_T\mathrm{tr}(\mathbf{N}_4) \\
& +\ell_T'((\mathbf{I}_T \odot \mathbf{N}_3)\mathbf{N}_2(\mathbf{I}_T \odot \mathbf{N}_4))\ell_T\mathrm{tr}(\mathbf{N}_1) + \ell_T'((\mathbf{I}_T \odot \mathbf{N}_3)\mathbf{N}_1(\mathbf{I}_T \odot \mathbf{N}_4))\ell_T\mathrm{tr}(\mathbf{N}_2) \\
& +\ell_T'((\mathbf{I}_T \odot \mathbf{N}_2)\mathbf{N}_1(\mathbf{I}_T \odot \mathbf{N}_4))\ell_T\mathrm{tr}(\mathbf{N}_3) + \ell_T'((\mathbf{I}_T \odot \mathbf{N}_2)\mathbf{N}_1(\mathbf{I}_T \odot \mathbf{N}_3))\ell_T\mathrm{tr}(\mathbf{N}_4)\Big] \\
& +4\Big[\ell_T'((\mathbf{I}_T \odot \mathbf{N}_1)\mathbf{N}_2\mathbf{N}_3(\mathbf{I}_T \odot \mathbf{N}_4))\ell_T + \ell_T'((\mathbf{I}_T \odot \mathbf{N}_1)\mathbf{N}_2\mathbf{N}_4(\mathbf{I}_T \odot \mathbf{N}_3))\ell_T \\
& +\ell_T'((\mathbf{I}_T \odot \mathbf{N}_2)\mathbf{N}_1\mathbf{N}_3(\mathbf{I}_T \odot \mathbf{N}_4))\ell_T + \ell_T'((\mathbf{I}_T \odot \mathbf{N}_2)\mathbf{N}_1\mathbf{N}_4(\mathbf{I}_T \odot \mathbf{N}_3))\ell_T \\
& +\ell_T'((\mathbf{I}_T \odot \mathbf{N}_3)\mathbf{N}_1\mathbf{N}_2(\mathbf{I}_T \odot \mathbf{N}_4))\ell_T + \ell_T'((\mathbf{I}_T \odot \mathbf{N}_3)\mathbf{N}_1\mathbf{N}_4(\mathbf{I}_T \odot \mathbf{N}_2))\ell_T \\
& +\ell_T'((\mathbf{I}_T \odot \mathbf{N}_4)\mathbf{N}_1\mathbf{N}_2(\mathbf{I}_T \odot \mathbf{N}_3))\ell_T + \ell_T'((\mathbf{I}_T \odot \mathbf{N}_4)\mathbf{N}_1\mathbf{N}_3(\mathbf{I}_T \odot \mathbf{N}_2))\ell_T \\
& +\ell_T'((\mathbf{I}_T \odot \mathbf{N}_1)\mathbf{N}_3\mathbf{N}_4(\mathbf{I}_T \odot \mathbf{N}_2))\ell_T + \ell_T'((\mathbf{I}_T \odot \mathbf{N}_2)\mathbf{N}_3\mathbf{N}_4(\mathbf{I}_T \odot \mathbf{N}_1))\ell_T \\
& +\ell_T'((\mathbf{I}_T \odot \mathbf{N}_3)\mathbf{N}_2\mathbf{N}_4(\mathbf{I}_T \odot \mathbf{N}_1))\ell_T + \ell_T'((\mathbf{I}_T \odot \mathbf{N}_4)\mathbf{N}_2\mathbf{N}_3(\mathbf{I}_T \odot \mathbf{N}_1))\ell_T\Big] \\
& +4\Big[\ell_T'(\mathbf{N}_2 \odot \mathbf{N}_3 \odot \mathbf{N}_4)\ell_T\mathrm{tr}(\mathbf{N}_1) + \ell_T'(\mathbf{N}_1 \odot \mathbf{N}_3 \odot \mathbf{N}_4)\ell_T\mathrm{tr}(\mathbf{N}_2) \\
& +\ell_T'(\mathbf{N}_1 \odot \mathbf{N}_2 \odot \mathbf{N}_4)\ell_T\mathrm{tr}(\mathbf{N}_3) + \ell_T'(\mathbf{N}_1 \odot \mathbf{N}_2 \odot \mathbf{N}_3)\ell_T\mathrm{tr}(\mathbf{N}_4)\Big] \\
& +8\Big[\ell_T'(\mathbf{N}_1 \odot \mathbf{N}_2)\mathbf{N}_3(\mathbf{I}_T \odot \mathbf{N}_4)\ell_T + \ell_T'(\mathbf{N}_1 \odot \mathbf{N}_2)\mathbf{N}_4(\mathbf{I}_T \odot \mathbf{N}_3)\ell_T \\
& +\ell_T'(\mathbf{N}_1 \odot \mathbf{N}_3)\mathbf{N}_2(\mathbf{I}_T \odot \mathbf{N}_4)\ell_T + \ell_T'(\mathbf{N}_1 \odot \mathbf{N}_3)\mathbf{N}_4(\mathbf{I}_T \odot \mathbf{N}_2)\ell_T \\
& +\ell_T'(\mathbf{N}_1 \odot \mathbf{N}_4)\mathbf{N}_2(\mathbf{I}_T \odot \mathbf{N}_3)\ell_T + \ell_T'(\mathbf{N}_1 \odot \mathbf{N}_4)\mathbf{N}_3(\mathbf{I}_T \odot \mathbf{N}_2)\ell_T
\end{aligned}
$$

$$
\begin{aligned}
&+\ell_T'(\mathbf{N}_2\odot\mathbf{N}_3)\mathbf{N}_1(\mathbf{I}_T\odot\mathbf{N}_4)\ell_T+\ell_T'(\mathbf{N}_2\odot\mathbf{N}_3)\mathbf{N}_4(\mathbf{I}_T\odot\mathbf{N}_1)\ell_T\\
&+\ell_T'(\mathbf{N}_2\odot\mathbf{N}_4)\mathbf{N}_1(\mathbf{I}_T\odot\mathbf{N}_3)\ell_T+\ell_T'(\mathbf{N}_2\odot\mathbf{N}_4)\mathbf{N}_3(\mathbf{I}_T\odot\mathbf{N}_1)\ell_T\\
&+\ell_T'(\mathbf{N}_3\odot\mathbf{N}_4)\mathbf{N}_1(\mathbf{I}_T\odot\mathbf{N}_2)\ell_T+\ell_T'(\mathbf{N}_3\odot\mathbf{N}_4)\mathbf{N}_2(\mathbf{I}_T\odot\mathbf{N}_1)\ell_T\Big]\\
&+16\Big[\mathrm{tr}(\mathbf{N}_1(\mathbf{N}_2\odot\mathbf{N}_3)\mathbf{N}_4)+\mathrm{tr}(\mathbf{N}_1(\mathbf{N}_2\odot\mathbf{N}_4)\mathbf{N}_3)+\mathrm{tr}(\mathbf{N}_1(\mathbf{N}_3\odot\mathbf{N}_4)\mathbf{N}_2)\\
&+\mathrm{tr}(\mathbf{N}_2(\mathbf{N}_1\odot\mathbf{N}_3)\mathbf{N}_4)+\mathrm{tr}(\mathbf{N}_2(\mathbf{N}_1\odot\mathbf{N}_4)\mathbf{N}_3)+\mathrm{tr}(\mathbf{N}_3(\mathbf{N}_1\odot\mathbf{N}_2)\mathbf{N}_4)\Big],
\end{aligned}
$$

$$
\begin{aligned}
f_{\gamma_2^2} &= \mathrm{tr}(\mathbf{N}_1\odot\mathbf{N}_2)\,\mathrm{tr}(\mathbf{N}_3\odot\mathbf{N}_4)+\mathrm{tr}(\mathbf{N}_1\odot\mathbf{N}_3)\,\mathrm{tr}(\mathbf{N}_2\odot\mathbf{N}_4)+\mathrm{tr}(\mathbf{N}_1\odot\mathbf{N}_4)\,\mathrm{tr}(\mathbf{N}_2\odot\mathbf{N}_3)\\
&+4[\ell_T'(\mathbf{I}_T\odot\mathbf{N}_1)(\mathbf{N}_2\odot\mathbf{N}_3)(\mathbf{I}_T\odot\mathbf{N}_4)\ell_T+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_1)(\mathbf{N}_2\odot\mathbf{N}_4)(\mathbf{I}_T\odot\mathbf{N}_3)\ell_T\\
&+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_1)(\mathbf{N}_3\odot\mathbf{N}_4)(\mathbf{I}_T\odot\mathbf{N}_2)\ell_T+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_2)(\mathbf{N}_1\odot\mathbf{N}_3)(\mathbf{I}_T\odot\mathbf{N}_4)\ell_T\\
&+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_2)(\mathbf{N}_1\odot\mathbf{N}_4)(\mathbf{I}_T\odot\mathbf{N}_3)\ell_T+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_3)(\mathbf{N}_1\odot\mathbf{N}_2)(\mathbf{I}_T\odot\mathbf{N}_4)\ell_T]\\
&+8\ell_T'(\mathbf{N}_1\odot\mathbf{N}_2\odot\mathbf{N}_3\odot\mathbf{N}_4)\ell_T,\\
f_{\gamma_1\gamma_3} &= 2[\ell_T'(\mathbf{I}_T\odot\mathbf{N}_1)\mathbf{N}_2(\mathbf{I}_T\odot\mathbf{N}_3\odot\mathbf{N}_4)\ell_T+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_1)\mathbf{N}_3(\mathbf{I}_T\odot\mathbf{N}_2\odot\mathbf{N}_4)\ell_T\\
&+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_1)\mathbf{N}_4(\mathbf{I}_T\odot\mathbf{N}_2\odot\mathbf{N}_3)\ell_T+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_2)\mathbf{N}_1(\mathbf{I}_T\odot\mathbf{N}_3\odot\mathbf{N}_4)\ell_T\\
&+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_2)\mathbf{N}_3(\mathbf{I}_T\odot\mathbf{N}_1\odot\mathbf{N}_4)\ell_T+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_2)\mathbf{N}_4(\mathbf{I}_T\odot\mathbf{N}_1\odot\mathbf{N}_3)\ell_T\\
&+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_3)\mathbf{N}_1(\mathbf{I}_T\odot\mathbf{N}_2\odot\mathbf{N}_4)\ell_T+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_3)\mathbf{N}_2(\mathbf{I}_T\odot\mathbf{N}_1\odot\mathbf{N}_4)\ell_T\\
&+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_3)\mathbf{N}_4(\mathbf{I}_T\odot\mathbf{N}_1\odot\mathbf{N}_2)\ell_T+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_4)\mathbf{N}_1(\mathbf{I}_T\odot\mathbf{N}_2\odot\mathbf{N}_3)\ell_T\\
&+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_4)\mathbf{N}_2(\mathbf{I}_T\odot\mathbf{N}_1\odot\mathbf{N}_3)\ell_T+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_4)\mathbf{N}_3(\mathbf{I}_T\odot\mathbf{N}_1\odot\mathbf{N}_2)\ell_T]\\
&+8[\ell_T'(\mathbf{I}_T\odot\mathbf{N}_1)(\mathbf{N}_2\odot\mathbf{N}_3\odot\mathbf{N}_4)\ell_T+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_2)(\mathbf{N}_1\odot\mathbf{N}_3\odot\mathbf{N}_4)\ell_T\\
&+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_3)(\mathbf{N}_1\odot\mathbf{N}_2\odot\mathbf{N}_4)\ell_T+\ell_T'(\mathbf{I}_T\odot\mathbf{N}_4)(\mathbf{N}_1\odot\mathbf{N}_2\odot\mathbf{N}_3)\ell_T].
\end{aligned}
$$

Parts (1)–(3) of Lemma 7.1 can be found in Appendix A.7 of Ullah (2004), while part (4) is Theorem 2 of Bao and Ullah (2010).

Let $\mathbf{J}_{i,k} = \left[\mathbf{I}_{T-i+1}, \mathbf{0}_{(T-i+1)\times(k-1)}\right]$, $\mathbf{S}_{i,k} = \left[\mathbf{0}_{(T-i+1)\times(k-1)}, \mathbf{I}_{T-i+1}\right]$, $\mathbf{F}_{T-i+k} = [1, \rho, ..., \rho^{T-i+k-1}]'$ and

$$
\mathbf{C}_{T-i+k} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ \rho & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho^{T-i+k-2} & \cdots & \rho & 1 & 0 \end{bmatrix}.
$$

Define $\mathbf{W}_{i,k} = \mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}$ and $\mathbf{K}_{i,k} = \mathbf{S}'_{i,k}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}$.

Lemma 7.2. *Given the structure of $\mathbf{J}_{i,k}$, $\mathbf{S}_{i,k}$, $\mathbf{C}_{T-i+k}$, $\mathbf{F}_{T-i+k}$ and $\mathbf{M}_i$, we have*

1. $\mathrm{tr}\left\{\mathbf{K}_{i,k}\right\} = -\frac{1}{1-\rho} + \frac{k-1}{(T-i+1)(1-\rho)} + \frac{1}{(T-i+1)(1-\rho)^2} + o(T^{-1})$;
2. $\mathrm{tr}\{\mathbf{W}_{i,k}\} = \frac{T-i+1}{1-\rho^2} - \frac{1}{(1-\rho^2)^2} - \frac{1}{(1-\rho)^2} + O(T^{-1})$;
3. $\mathbf{F}'_{T-i+1}\mathbf{M}_i\mathbf{F}_{T-i+1} = \frac{1}{1-\rho^2} - \frac{1}{(T-i+1)(1-\rho)^2} + o(T^{-1})$;

4. $\operatorname{tr}\left\{\mathbf{F}'_{T-i+1}\mathbf{M}_i\mathbf{S}_{i,k}\mathbf{C}'_{T-i+k}\mathbf{J}'_{i,k}\mathbf{M}_i\mathbf{F}_{T-i+1}\right\} = \frac{\rho^k}{(1-\rho^2)^2} + O(T^{-1})$;

5. $\operatorname{tr}\{\mathbf{W}_{i,k} \odot \mathbf{K}_{i,k}\} = -\frac{1}{(1-\rho^2)(1-\rho)} + O(T^{-1})$;

6. $\operatorname{tr}\{\mathbf{W}_{i,k}\mathbf{K}_{i,k}\} = \frac{(T-i+1)\rho^k}{(1-\rho^2)^2} - \frac{(k-1)\rho^k}{(1-\rho^2)^2} - \frac{1}{(1-\rho)^3} - \frac{2\rho^k}{(1-\rho^2)^3} + O(T^{-1})$;

7. $\operatorname{tr}\{\mathbf{W}_{i,k} \odot \mathbf{S}'_{i,k}\mathbf{M}_i\mathbf{F}_{T-i+1}\mathbf{F}'_{T-i+1}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\} = O(1)$;

8. $\operatorname{tr}\{\mathbf{W}_{i,k}\mathbf{S}'_{i,k}\mathbf{M}_i\mathbf{F}_{T-i+1}\mathbf{F}'_{T-i+1}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\} = O(1)$,
$\operatorname{tr}\{\mathbf{W}^2_{i,k}\mathbf{S}'_{i,k}\mathbf{M}_i\mathbf{F}_{T-i+1}\mathbf{F}'_{T-i+1}\mathbf{M}_i\mathbf{J}_{i,k}\mathbf{C}_{T-i+k}\} = O(1)$;

9. $\operatorname{tr}\{\mathbf{W}^2_{i,k}\} = \frac{(T-i+1)(1+\rho^2)}{(1-\rho^2)^3} + O(1)$;

10. $\operatorname{tr}\{\mathbf{W}^2_{i,k}\mathbf{K}_{i,k}\} = \frac{(k+1)(T-i+1)\rho^k}{(1-\rho^2)^3} + \frac{3(T-i+1)\rho^{(k+2)}}{(1-\rho^2)^4} + O(1)$;

11. $\operatorname{tr}\{\mathbf{W}^3_{i,k}\mathbf{K}_{i,k}\} = O(T)$, $\operatorname{tr}\{\mathbf{W}^3_{i,k}\} = O(T)$.

Lemma 7.3. *Given the structure of* $\mathbf{J}_{i,k}$, $\mathbf{S}_{i,k}$, $\mathbf{C}_{T-i+k}$, $\mathbf{F}_{T-i+k}$ *and* $\mathbf{M}_i$, *we have*

1. $\operatorname{tr}\{\mathbf{W}_{i,k} \odot \mathbf{W}_{i,k} \odot \mathbf{K}_{i,k}\} = O(1)$;

2. $\operatorname{tr}\{\mathbf{W}_{i,k} \odot \mathbf{W}_{i,k}\} = \frac{T-i+1}{(1-\rho^2)^2} + O(1)$;

3. $\operatorname{tr}\left\{\mathbf{W}^2_{i,k} \odot \mathbf{K}_{i,k}\right\} = O(1)$, $\operatorname{tr}\left\{\mathbf{W}^2_{i,k} \odot \mathbf{W}_{i,k}\right\} = O(T)$;

4. $\operatorname{tr}\{(\mathbf{W}_{i,k}\mathbf{K}_{i,k}) \odot \mathbf{W}_{i,k}\} = \frac{(T-i+1)\rho^k}{(1-\rho^2)^3} + O(1)$, $\operatorname{tr}\{(\mathbf{K}_{i,k}\mathbf{W}_{i,k}) \odot \mathbf{W}_{i,k}\} = \frac{(T-i+1)\rho^k}{(1-\rho^2)^3} + O(1)$;

5. $\operatorname{tr}\{(\mathbf{W}_{i,k} \odot \mathbf{W}_{i,k})\mathbf{K}_{i,k}\} = \frac{T\rho^{2k}}{(1-\rho^3)(1-\rho^2)^2} + O(1)$;

6. $\boldsymbol{\ell}'_{T-i+k}(\mathbf{I}_{T-i+k}\odot\mathbf{W}_{i,k})\mathbf{K}_{i,k}(\mathbf{I}_{T-i+k}\odot\mathbf{W}_{i,k})\boldsymbol{\ell}_{T-i+k} = O(1)$, $\boldsymbol{\ell}'_{T-i+k}(\mathbf{I}_{T-i+k}\odot\mathbf{W}_{i,k})\mathbf{W}_{i,k}(\mathbf{I}_{T-i+k}\odot \mathbf{K}_{i,k})\boldsymbol{\ell}_{T-i+k} = O(1)$;

Lemma 7.4. *Given the structure of* $\mathbf{J}_{i,k}$, $\mathbf{S}_{i,k}$, $\mathbf{C}_{T-i+k}$, $\mathbf{F}_{T-i+k}$ *and* $\mathbf{M}_i$, *we have*

1. $\operatorname{tr}\{\mathbf{W}_{i,k} \odot \mathbf{W}_{i,k} \odot \mathbf{W}_{i,k} \odot \mathbf{K}_{i,k}\} = O(1)$, $\operatorname{tr}\{\mathbf{W}_{i,k} \odot \mathbf{W}_{i,k} \odot \mathbf{W}_{i,k}\} = O(1)$;

2. $\operatorname{tr}\left\{\mathbf{W}^2_{i,k} \odot \mathbf{W}_{i,k} \odot \mathbf{K}_{i,k}\right\} = O(1)$;

3. $\operatorname{tr}\{(\mathbf{W}_{i,k}\mathbf{K}_{i,k}) \odot \mathbf{W}_{i,k} \odot \mathbf{W}_{i,k}\} = O(T)$;

4. $\boldsymbol{\ell}'_{T-i+k}(\mathbf{I}_{T-i+k} \odot \mathbf{W}_{i,k})(\mathbf{W}_{i,k} \odot \mathbf{K}_{i,k})(\mathbf{I}_{T-i+k} \odot \mathbf{W}_{i,k})\boldsymbol{\ell}_{T-i+k} = O(T)$;

5. $\boldsymbol{\ell}'_{T-i+k}(\mathbf{W}_{i,k} \odot \mathbf{W}_{i,k} \odot \mathbf{W}_{i,k} \odot \mathbf{K}_{i,k})\boldsymbol{\ell}_{T-i+k} = O(T)$;

6. $\boldsymbol{\ell}'_{T-i+k}(\mathbf{I}_{T-i+k} \odot \mathbf{W}_{i,k})\mathbf{W}_{i,k}(\mathbf{I}_{T-i+k} \odot \mathbf{W}_{i,k} \odot \mathbf{K}_{i,k})\boldsymbol{\ell}_{T-i+k} = O(1)$;

7. $\boldsymbol{\ell}'_{T-i+k}(\mathbf{I}_{T-i+k} \odot \mathbf{W}_{i,k})\mathbf{K}_{i,k}(\mathbf{I}_{T-i+k} \odot \mathbf{W}_{i,k} \odot \mathbf{W}_{i,k})\boldsymbol{\ell}_{T-i+k} = O(1)$;

8. $\boldsymbol{\ell}'_{T-i+k}(\mathbf{I}_{T-i+k} \odot \mathbf{K}_{i,k})\mathbf{W}_{i,k}(\mathbf{I}_{T-i+k} \odot \mathbf{W}_{i,k} \odot \mathbf{W}_{i,k})\boldsymbol{\ell}_{T-i+k} = O(1)$;

9. $\boldsymbol{\ell}'_{T-i+k}(\mathbf{I}_{T-i+k} \odot \mathbf{K}_{i,k})(\mathbf{W}_{i,k} \odot \mathbf{W}_{i,k} \odot \mathbf{W}_{i,k})\boldsymbol{\ell}_{T-i+k} = O(1)$;

10. $\boldsymbol{\ell}'_{T-i+k}(\mathbf{I}_{T-i+k}\odot\mathbf{K}_{i,k})(\mathbf{W}_{i,k}\odot\mathbf{W}_{i,k}\odot\mathbf{W}_{i,k})\boldsymbol{\ell}_{T-i+k}=O(1)$;

11. $\boldsymbol{\ell}'_{T-i+k}(\mathbf{W}_{i,k}\odot\mathbf{W}_{i,k})\boldsymbol{\ell}_{T-i+k}=O(T)$;

12. $\boldsymbol{\ell}'_{T-i+k}(\mathbf{W}_{i,k}\odot\mathbf{K}_{i,k})\boldsymbol{\ell}_{T-i+k}=\frac{(T-i+1)\rho^k}{(1-\rho^2)^2}+O(1)$;

13. $\operatorname{tr}\left\{(\mathbf{I}_{T-i+k}\odot\mathbf{W}_{i,k})\mathbf{W}_{i,k}^2\mathbf{K}_{i,k}\right\}=O(T)$, $\operatorname{tr}\{(\mathbf{I}_{T-i+k}\odot\mathbf{W}_{i,k})\mathbf{W}_{i,k}\mathbf{K}_{i,k}\mathbf{W}_{i,k}\}=O(T)$;

14. $\operatorname{tr}\left\{(\mathbf{I}_{T-i+k}\odot\mathbf{K}_{i,k})\mathbf{W}_{i,k}^3\right\}=O(1)$, $\boldsymbol{\ell}'_{T-i+k}(\mathbf{I}_{T-i+k}\odot\mathbf{W}_{i,k}\mathbf{K}_{i,k})(\mathbf{I}_{T-i+k}\odot\mathbf{W}_{i,k}^2)\boldsymbol{\ell}_{T-i+k}=O(T)$;

15. $\boldsymbol{\ell}'_{T-i+k}(\mathbf{I}_{T-i+k}\odot\mathbf{W}_{i,k})\mathbf{W}_{i,k}\mathbf{K}_{i,k}(\mathbf{I}_{T-i+k}\odot\mathbf{W}_{i,k})\boldsymbol{\ell}_{T-i+k}=O(1)$,
$\boldsymbol{\ell}'_{T-i+k}(\mathbf{I}_{T-i+k}\odot\mathbf{W}_{i,k})\mathbf{W}_{i,k}^2(\mathbf{I}_{T-i+k}\odot\mathbf{K}_{i,k})\boldsymbol{\ell}_{T-i+k}=O(1)$;

16. $\boldsymbol{\ell}'_{T-i+k}(\mathbf{W}_{i,k}\odot\mathbf{W}_{i,k}\odot\mathbf{W}_{i,k})\boldsymbol{\ell}_{T-i+k}=O(T)$,
$\boldsymbol{\ell}'_{T-i+k}(\mathbf{W}_{i,k}\odot\mathbf{W}_{i,k}\odot\mathbf{K}_{i,k})\boldsymbol{\ell}_{T-i+k}=\frac{(T-i+1)\rho^{2k}}{(1-\rho^2)^2(1-\rho^3)}+O(1)$;

17. $\boldsymbol{\ell}'_{T-i+k}(\mathbf{W}_{i,k}\odot\mathbf{W}_{i,k})\mathbf{K}_{i,k}(\mathbf{I}_{T-i+k}\odot\mathbf{W}_{i,k})\boldsymbol{\ell}_{T-i+k}=O(1)$, $\boldsymbol{\ell}'_{T-i+k}(\mathbf{W}_{i,k}\odot\mathbf{W}_{i,k})\mathbf{W}_{i,k}(\mathbf{I}_{T-i+k}\odot\mathbf{K}_{i,k})\boldsymbol{\ell}_{T-i+k}=O(1)$, $\boldsymbol{\ell}'_{T-i+k}(\mathbf{W}_{i,k}\odot\mathbf{K}_{i,k})\mathbf{W}_{i,k}(\mathbf{I}_{T-i+k}\odot\mathbf{W}_{i,k})\boldsymbol{\ell}_{T-i+k}=O(1)$;

18. $\operatorname{tr}\{\mathbf{W}_{i,k}(\mathbf{W}_{i,k}\odot\mathbf{W}_{i,k})\mathbf{K}_{i,k}\}=O(T)$, $\operatorname{tr}\{\mathbf{W}_{i,k}(\mathbf{W}_{i,k}\odot\mathbf{K}_{i,k})\mathbf{W}_{i,k}\}=O(T)$.

Lemma 7.5. *Given the structure of* $\mathbf{C}_T$, $\mathbf{F}_T$ *and* $\mathbf{M}_1$, *we have*

1. $\operatorname{tr}\{\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}\}=O(T^{-1})$, $\operatorname{tr}\left\{\mathbf{K}_{1,1}^2\right\}=-\frac{1}{(1-\rho)^2}+O(T^{-1})$;

2 $\operatorname{tr}\left\{\mathbf{W}_{1,1}(\mathbf{K}_{1,1})^2\right\}=\frac{T\rho^2}{(1-\rho^2)^3}+O(1)$, $\operatorname{tr}\left\{\left(\mathbf{K}'_{1,1}\right)^3\mathbf{C}_T\right\}=\frac{T\rho^2}{(1-\rho^2)^3}+O(1)$,
$\operatorname{tr}\left\{\mathbf{W}_{1,1}\mathbf{K}_{1,1}\mathbf{K}'_{1,1}\right\}=\frac{T}{(1-\rho^2)^2}+\frac{2T\rho^2}{(1-\rho^2)^3}+O(1)$;

3. $\operatorname{tr}\{\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}\}=O(T^{-1})$, $\operatorname{tr}\left\{\mathbf{W}_{1,1}\odot(\mathbf{K}_{1,1})^2\right\}=O(1)$, $\operatorname{tr}\left\{\mathbf{W}_{1,1}\odot\left(\mathbf{K}_{1,1}\mathbf{K}'_{1,1}\right)\right\}=\frac{T}{(1-\rho^2)^2}+O(1)$, $\operatorname{tr}\{(\mathbf{W}_{1,1}\mathbf{K}_{1,1})\odot\mathbf{K}_{1,1}\}=O(1)$, $\operatorname{tr}\left\{\left(\mathbf{W}_{1,1}\mathbf{K}'_{1,1}\right)\odot\mathbf{K}_{1,1}\right\}=O(1)$,
$\operatorname{tr}\{(\mathbf{K}_{1,1}\mathbf{W}_{1,1})\odot\mathbf{K}_{1,1}\}=O(1)$, $\operatorname{tr}\{\mathbf{W}_{1,1}(\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1})\}=\frac{T\rho}{(1-\rho^2)(1-\rho^3)}+O(1)$,
$\operatorname{tr}\left\{\mathbf{W}_{1,1}\left(\mathbf{K}_{1,1}\odot\mathbf{K}'_{1,1}\right)\right\}=O(1)$, $\operatorname{tr}\{\mathbf{W}_{1,1}\odot\mathbf{W}_{1,1}\}=\frac{T}{(1-\rho^2)^2}+O(1)$;

4. $\boldsymbol{\ell}'_T(\mathbf{I}_T\odot\mathbf{K}_{1,1})\mathbf{W}_{1,1}(\mathbf{I}_T\odot\mathbf{K}_{1,1})\boldsymbol{\ell}_T=o(1)$, $\boldsymbol{\ell}'_T(\mathbf{I}_T\odot\mathbf{K}_{1,1})\mathbf{K}_{1,1}(\mathbf{I}_T\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T=o(1)$, $\boldsymbol{\ell}'_T(\mathbf{I}_T\odot\mathbf{K}_{1,1})\mathbf{W}_{1,1}(\mathbf{I}_T\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T=o(1)$;

5. $\operatorname{tr}\{\mathbf{W}_{1,1}\odot\mathbf{M}_1\mathbf{F}_T\mathbf{F}'_T\mathbf{M}_1\}=O(1)$, $\operatorname{tr}\{\mathbf{W}_{1,1}\mathbf{M}_1\mathbf{F}_T\mathbf{F}'_T\mathbf{M}_1\}=O(1)$,
$\operatorname{tr}\left\{\mathbf{W}_{1,1}^2\mathbf{M}_1\mathbf{F}_T\mathbf{F}'_T\mathbf{M}_1\right\}=O(1)$, $\operatorname{tr}\{\mathbf{W}_{1,1}\odot\mathbf{W}_{1,1}\odot\mathbf{M}_1\mathbf{F}_T\mathbf{F}'_T\mathbf{M}_1\}=O(1)$,
$\operatorname{tr}\{(\mathbf{W}_{1,1}\odot\mathbf{W}_{1,1})\mathbf{M}_1\mathbf{F}_T\mathbf{F}'_T\mathbf{M}_1\}=O(1)$, $\operatorname{tr}\left\{(\mathbf{W}_{1,1})^2\odot\mathbf{M}_1\mathbf{F}_T\mathbf{F}'_T\mathbf{M}_1\right\}=O(1)$,
$\operatorname{tr}\{\mathbf{W}_{1,1}\odot(\mathbf{W}_{1,1}\mathbf{M}_1\mathbf{F}_T\mathbf{F}'_T\mathbf{M}_1)\}=O(1)$;

6. $\boldsymbol{\ell}'_T(\mathbf{I}_T\odot\mathbf{W}_{1,1})\mathbf{W}_{1,1}(\mathbf{I}_T\odot\mathbf{M}_1\mathbf{F}_T\mathbf{F}'_T\mathbf{M}_1)\boldsymbol{\ell}_T=O(1)$, $\boldsymbol{\ell}'_T(\mathbf{I}_T\odot\mathbf{W}_{1,1})\mathbf{M}_1\mathbf{F}_T\mathbf{F}'_T\mathbf{M}_1(\mathbf{I}_T\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T=o(1)$.

Lemma 7.6. *Given the structure of* $\mathbf{C}_T$, $\mathbf{F}_T$ *and* $\mathbf{M}_1$, *we have*

1. $\operatorname{tr}\left\{\mathbf{W}_{1,1}^2\mathbf{K}_{1,1}\right\} = O(T)$, $\operatorname{tr}\left\{\mathbf{W}_{1,1}^2\mathbf{K}_{1,1}^2\right\} = O(T)$, $\operatorname{tr}\left\{\mathbf{W}_{1,1}^3\right\} = O(T)$, $\ell_T'\left(\mathbf{W}_{1,1}\odot\mathbf{W}_{1,1}\right)\ell_T = O(T)$, $\ell_T'\left(\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}\right)\ell_T = \frac{T}{1-\rho^2}+O(1)$, $\ell_T'\left(\mathbf{K}_{1,1}'\odot\mathbf{K}_{1,1}\right)\ell_T = O(1)$,
   $\operatorname{tr}\left\{(\mathbf{W}_{1,1})^2\odot\mathbf{K}_{1,1}\right\} = O(1)$, $\operatorname{tr}\left\{\mathbf{W}_{1,1}\mathbf{K}_{1,1}\odot\mathbf{C}_T'\mathbf{K}_{1,1}\right\} = O(T)$, $\operatorname{tr}\left\{\mathbf{W}_{1,1}\mathbf{K}_{1,1}'\odot\mathbf{W}_{1,1}\right\} = O(T)$,
   $\operatorname{tr}\left\{\left(\mathbf{I}_T\odot\mathbf{W}_{1,1}\right)\mathbf{W}_{1,1}(\mathbf{K}_{1,1}+\mathbf{K}_{1,1}')^2\right\} = O(T)$, $\operatorname{tr}\left\{\left(\mathbf{I}_T\odot\mathbf{W}_{1,1}\right)(\mathbf{K}_{1,1}+\mathbf{K}_{1,1}')\mathbf{W}_{1,1}(\mathbf{K}_{1,1}+\mathbf{K}_{1,1}')\right\} = O(T)$, $\operatorname{tr}\left\{\left(\mathbf{I}_T\odot\mathbf{K}_{1,1}\right)\mathbf{W}_{1,1}^2\mathbf{K}_{1,1}\right\} = O(1)$, $\operatorname{tr}\left\{\left(\mathbf{I}_T\odot\mathbf{K}_{1,1}\right)\mathbf{W}_{1,1}\mathbf{K}_{1,1}\mathbf{W}_{1,1}\right\} = O(1)$,
   $\ell_T'\left(\mathbf{I}_T\odot\mathbf{K}_{1,1}^2\right)\left(\mathbf{I}_T\odot\mathbf{W}_{1,1}^2\right)\ell_T = O(1)$, $\ell_T'\left(\mathbf{I}_T\odot(\mathbf{W}_{1,1}\mathbf{K}_{1,1}')\right)^2\ell_T = O(T)$,
   $\ell_T'\left(\mathbf{I}_T\odot\mathbf{W}_{1,1}\right)\left(\mathbf{I}_T\odot\mathbf{W}_{1,1}^2\right)\ell_T = O(T)$, $\ell_T'\left(\mathbf{I}_T\odot(\mathbf{W}_{1,1}\mathbf{K}_{1,1})\right)^2\ell_T = O(T)$;

2. $\operatorname{tr}\left\{\mathbf{W}_{1,1}\odot\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1}^2\right\} = O(1)$, $\operatorname{tr}\left\{\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1}\odot\mathbf{W}_{1,1}\mathbf{W}_{1,1}'\right\} = O(1)$,
   $\operatorname{tr}\left\{\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1}\odot\mathbf{W}_{1,1}\mathbf{K}_{1,1}\right\} = O(1)$, $\operatorname{tr}\left\{\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1}\odot\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1}\right\} = O(T^{-1})$;

3. $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{W}_{1,1})\mathbf{W}_{1,1}(\mathbf{I}_T\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = O(1)$, $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{W}_{1,1})\mathbf{K}_{1,1}^2(\mathbf{I}_T\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = O(1)$,
   $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{K}_{1,1})\mathbf{W}_{1,1}^2(\mathbf{I}_T\odot\mathbf{K}_{1,1})\boldsymbol{\ell}_T = o(1)$, $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{W}_{1,1})\mathbf{W}_{1,1}\mathbf{K}_{1,1}(\mathbf{I}_T\odot\mathbf{K}_{1,1})\boldsymbol{\ell}_T = o(1)$,
   $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{W}_{1,1})\mathbf{W}_{1,1}\mathbf{K}_{1,1}'(\mathbf{I}_T\odot\mathbf{K}_{1,1})\boldsymbol{\ell}_T = o(1)$, $\boldsymbol{\ell}_T'(\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = \frac{T\rho}{(1-\rho^2)(1-\rho^3)}+O(1)$,
   $\boldsymbol{\ell}_T'(\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}'\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = O(1)$, $\boldsymbol{\ell}_T'(\mathbf{K}_{1,1}\odot\mathbf{W}_{1,1}\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = O(T)$,
   $\boldsymbol{\ell}_T'(\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}')\mathbf{W}_{1,1}(\mathbf{I}_T\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = o(1)$, $\boldsymbol{\ell}_T'(\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1})\mathbf{W}_{1,1}(\mathbf{I}_T\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = O(1)$,
   $\boldsymbol{\ell}_T'(\mathbf{W}_{1,1}\odot\mathbf{W}_{1,1})\mathbf{K}_{1,1}(\mathbf{I}_T\odot\mathbf{K}_{1,1})\boldsymbol{\ell}_T = o(1)$, $\boldsymbol{\ell}_T'(\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1})\mathbf{K}_{1,1}(\mathbf{I}_T\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = O(1)$,
   $\boldsymbol{\ell}_T'(\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1})\mathbf{K}_{1,1}'(\mathbf{I}_T\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = O(1)$, $\boldsymbol{\ell}_T'(\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1})\mathbf{W}_{1,1}(\mathbf{I}_T\odot\mathbf{K}_{1,1})\boldsymbol{\ell}_T = o(1)$;

4. $\operatorname{tr}\left\{\left(\mathbf{W}_{1,1}\odot\mathbf{W}_{1,1}\right)\mathbf{W}_{1,1}\right\} = O(T)$, $\operatorname{tr}\left\{\left(\mathbf{W}_{1,1}\odot\mathbf{W}_{1,1}\right)\mathbf{K}_{1,1}^2\right\} = O(T)$,
   $\operatorname{tr}\left\{\left(\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1}\right)\mathbf{W}_{1,1}\mathbf{K}_{1,1}\right\} = O(T)$, $\operatorname{tr}\left\{\left(\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1}\right)\mathbf{W}_{1,1}\mathbf{K}_{1,1}'\right\} = O(T)$;

5. $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{W}_{1,1})\left(\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}\right)(\mathbf{I}_T\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = O(T)$,
   $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{W}_{1,1})\left(\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}'\right)(\mathbf{I}_T\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = O(1)$,
   $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{K}_{1,1})\left(\mathbf{K}_{1,1}\odot\mathbf{C}_T'\mathbf{K}_{1,1}\right)(\mathbf{I}_T\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = O(1)$,
   $\boldsymbol{\ell}_T'(\mathbf{K}_{1,1}'\odot\mathbf{K}_{1,1}\odot\mathbf{W}_{1,1}\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = O(1)$,
   $\boldsymbol{\ell}_T'(\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}\odot\mathbf{W}_{1,1}\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = O(T)$;

6. $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{W}_{1,1})\mathbf{K}_{1,1}(\mathbf{I}_T\odot\mathbf{K}_{1,1}\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = o(1)$, $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{W}_{1,1})\mathbf{K}_{1,1}'(\mathbf{I}_T\odot\mathbf{K}_{1,1}\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = o(1)$,
   $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{K}_{1,1})\mathbf{W}_{1,1}(\mathbf{I}_T\odot\mathbf{K}_{1,1}\odot\mathbf{W}_{1,1})\boldsymbol{\ell}_T = o(1)$, $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{W}_{1,1})\mathbf{W}_{1,1}(\mathbf{I}_T\odot\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1})\boldsymbol{\ell}_T = o(1)$,
   $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{W}_{1,1})(\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1})\boldsymbol{\ell}_T = O(T)$, $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{W}_{1,1})(\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}')\boldsymbol{\ell}_T = O(1)$,
   $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{K}_{1,1})(\mathbf{W}_{1,1}\odot\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1})\boldsymbol{\ell}_T = O(1)$;

7. $\operatorname{tr}\left\{\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}\right\} = o(1)$, $\operatorname{tr}\left\{\mathbf{K}_{1,1}\odot(\mathbf{K}_{1,1})^2\right\} = o(1)$,
   $\operatorname{tr}\left\{(\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1})\mathbf{K}_{1,1}\right\} = O(1)$, $\operatorname{tr}\left\{(\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}')\mathbf{K}_{1,1}\right\} = O(1)$,
   $\boldsymbol{\ell}_T'(\mathbf{I}_T\odot\mathbf{K}_{1,1})\mathbf{K}_{1,1}(\mathbf{I}_T\odot\mathbf{K}_{1,1})\boldsymbol{\ell}_T = o(1)$.

Lemma 7.7. *Given the structure of $\mathbf{C}_T$, $\mathbf{F}_T$ and $\mathbf{M}_1$, we have*

1. $\operatorname{tr}\left\{\mathbf{W}_{1,1}\mathbf{K}_{1,1}^3\right\} = O(T)$, $\operatorname{tr}\left\{\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}\right\} = O(T^{-2})$,
   $\operatorname{tr}\left\{\left(\mathbf{I}_T\odot\mathbf{W}_{1,1}\right)\left(\mathbf{K}_{1,1}'+\mathbf{K}_{1,1}\right)^3\right\} = O(T)$, $\operatorname{tr}\left\{\left(\mathbf{I}_T\odot\mathbf{K}_{1,1}\right)\mathbf{W}_{1,1}\left(\mathbf{K}_{1,1}'+\mathbf{K}_{1,1}\right)^2\right\} = O(1)$,
   $\ell_T'\left\{\mathbf{I}_T\odot(\mathbf{W}_{1,1}\mathbf{K}_{1,1})\right\}\left\{\mathbf{I}_T\odot\left(\mathbf{K}_{1,1}'+\mathbf{K}_{1,1}\right)^2\right\}\ell_T = O(T)$,
   $\operatorname{tr}\left\{\mathbf{W}_{1,1}\odot\mathbf{K}_{1,1}\odot\left(\mathbf{K}_{1,1}'+\mathbf{K}_{1,1}\right)^2\right\} = O(1)$, $\operatorname{tr}\left\{\mathbf{K}_{1,1}\odot\mathbf{K}_{1,1}\odot(\mathbf{W}_{1,1}\mathbf{K}_{1,1})\right\} = O(T^{-1})$;

2. $\ell_T' \left(\mathbf{I}_T \odot \mathbf{W}_{1,1}\right) \left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right)^2 \left(\mathbf{I}_T \odot \mathbf{K}_{1,1}\right) \ell_T = O(T^{-1})$,
$\ell_T' \left(\mathbf{I}_T \odot \mathbf{K}_{1,1}\right) \mathbf{W}_{1,1}\mathbf{K}_{1,1} \left(\mathbf{I}_T \odot \mathbf{K}_{1,1}\right) \ell_T = O(T^{-2})$,
$\ell_T' \left(\mathbf{W}_{1,1} \odot \mathbf{K}_{1,1}\right) \left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right) \left(\mathbf{I}_T \odot \mathbf{K}_{1,1}\right) \ell_T = O(T^{-1})$,
$\ell_T' \left[\left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right) \odot \left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right)\right] \mathbf{W}_{1,1} \left(\mathbf{I}_T \odot \mathbf{K}_{1,1}\right) \ell_T = O(T^{-1})$,
$\ell_T' \left[\left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right) \odot \left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right)\right] \mathbf{K}_{1,1} \left(\mathbf{I}_T \odot \mathbf{W}_{1,1}\right) \ell_T = O(1)$,
$\mathrm{tr}\left\{\mathbf{W}_{1,1} \left[\left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right) \odot \left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right)\right] \mathbf{K}_{1,1}\right\} = O(T)$,
$\mathrm{tr}\left\{\left(\mathbf{W}_{1,1} \odot \mathbf{K}_{1,1}\right) \left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right)^2\right\} = O(T)$;

3. $\ell_T' \left(\mathbf{K}_{1,1} \odot \mathbf{K}_{1,1} \odot \mathbf{K}_{1,1}\right) \ell_T = \frac{T}{1-\rho^3} + O(1)$,
$\ell_T' \left(\mathbf{I}_T \odot \mathbf{W}_{1,1}\right) \left[\left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right) \odot \left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right)\right] \left(\mathbf{I}_T \odot \mathbf{K}_{1,1}\right) \ell_T = O(1)$,
$\ell_T' \left(\mathbf{I}_T \odot \mathbf{K}_{1,1}\right) \left(\mathbf{W}_{1,1} \odot \mathbf{K}_{1,1}\right) \left(\mathbf{I}_T \odot \mathbf{K}_{1,1}\right) \ell_T = O(T^{-1})$,
$\ell_T' \left\{\left(\mathbf{W}_{1,1}\right) \odot \left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right) \odot \left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right) \odot \left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right)\right\} \ell_T = O(T)$,
$\ell_T' \left(\mathbf{I}_T \odot \mathbf{W}_{1,1}\right) \mathbf{K}_{1,1} \left(\mathbf{I}_T \odot \mathbf{K}_{1,1} \odot \mathbf{K}_{1,1}\right) \ell_T = O(T^{-2})$,
$\ell_T' \left(\mathbf{I}_T \odot \mathbf{K}_{1,1}\right) \mathbf{W}_{1,1} \left(\mathbf{I}_T \odot \mathbf{K}_{1,1} \odot \mathbf{K}_{1,1}\right) \ell_T = O(T^{-3})$,
$\ell_T' \left(\mathbf{I}_T \odot \mathbf{K}_{1,1}\right) \mathbf{K}_{1,1} \left(\mathbf{I}_T \odot \mathbf{W}_{1,1} \odot \mathbf{K}_{1,1}\right) \ell_T = O(T^{-2})$,
$\ell_T' \left(\mathbf{I}_T \odot \mathbf{W}_{1,1}\right) \left[\left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right) \odot \left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right) \odot \left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right)\right] \ell_T = O(T)$,
$\ell_T' \left(\mathbf{I}_T \odot \mathbf{K}_{1,1}\right) \left[\mathbf{W}_{1,1} \odot \left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right) \odot \left(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}\right)\right] \ell_T = O(1)$;

4. $\mathrm{tr}\left\{\mathbf{K}_{1,1}^4\right\} = O(1)$, $\mathrm{tr}\left\{(\mathbf{I}_T \odot \mathbf{K}_{1,1})(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1})^3\right\} = O(T)$,
$\ell_T'(\mathbf{I}_T \odot (\mathbf{K}_{1,1}' + \mathbf{K}_{1,1})^2)(\mathbf{I}_T \odot (\mathbf{K}_{1,1}' + \mathbf{K}_{1,1})^2)\ell_T = O(1)$,
$\mathrm{tr}\left\{\mathbf{K}_{1,1} \odot \mathbf{K}_{1,1} \odot (\mathbf{K}_{1,1}' + \mathbf{K}_{1,1})^2\right\} = O(T^{-1})$,
$\mathrm{tr}\left\{\mathbf{K}_{1,1} \odot \mathbf{K}_{1,1} \odot \mathbf{K}_{1,1} \odot \mathbf{K}_{1,1}\right\} = O(T^{-3})$;

5. $\ell_T'(\mathbf{I}_T \odot \mathbf{K}_{1,1})\mathbf{K}_{1,1}(\mathbf{I}_T \odot \mathbf{K}_{1,1})\ell_T = O(1)$,
$\ell_T'(\mathbf{I}_T \odot \mathbf{K}_{1,1})(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1})^2(\mathbf{I}_T \odot \mathbf{K}_{1,1})\ell_T = O(1)$,
$\ell_T'(\mathbf{K}_{1,1} \odot \mathbf{K}_{1,1} \odot \mathbf{K}_{1,1}')\ell_T = O(1)$,
$\ell_T'((\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}) \odot (\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}))\mathbf{K}_{1,1}(\mathbf{I}_T \odot \mathbf{K}_{1,1})\ell_T = O(1)$,
$\mathrm{tr}\left\{(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1})((\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}) \odot (\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}))(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1})\right\} = O(T)$;

6. $\ell_T'(\mathbf{I}_T \odot \mathbf{K}_{1,1})((\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}) \odot (\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}))(\mathbf{I}_T \odot \mathbf{K}_{1,1})\ell_T = O(1)$,
$\ell_T' \left\{(\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}) \odot (\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}) \odot (\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}) \odot (\mathbf{K}_{1,1}' + \mathbf{K}_{1,1})\right\} \ell_T = O(T)$;

7. $\ell_T'(\mathbf{I}_T \odot \mathbf{K}_{1,1})\mathbf{K}_{1,1}(\mathbf{I}_T \odot \mathbf{K}_{1,1} \odot \mathbf{K}_{1,1})\ell_T = O(1)$,
$\ell_T'(\mathbf{I}_T \odot \mathbf{K}_{1,1})((\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}) \odot (\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}) \odot (\mathbf{K}_{1,1}' + \mathbf{K}_{1,1}))\ell_T = O(1)$.

Proof of Lemmas 7.2–7.7. Given the structure of $\mathbf{K}_{i,k}$ and $\mathbf{W}_{i,k}$, we can obtain the results in Lemmas 7.2–7.7 by straightforward albeit tedious calculations. Numerically, one can easily verify these formulas by using a computer program. □

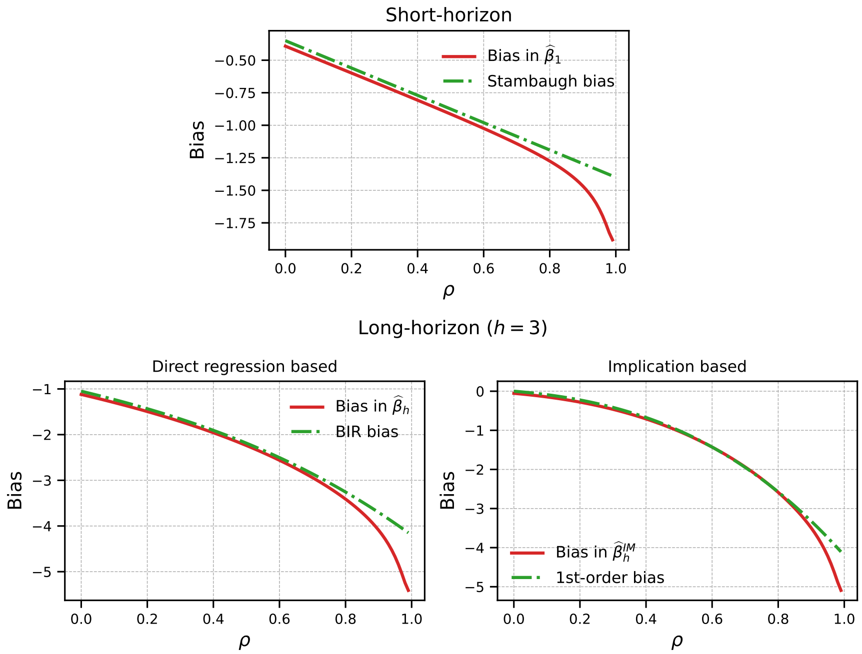
7.7 Proof of Proposition 3.1

Taking expectations in Eq. (3.2) leads to:

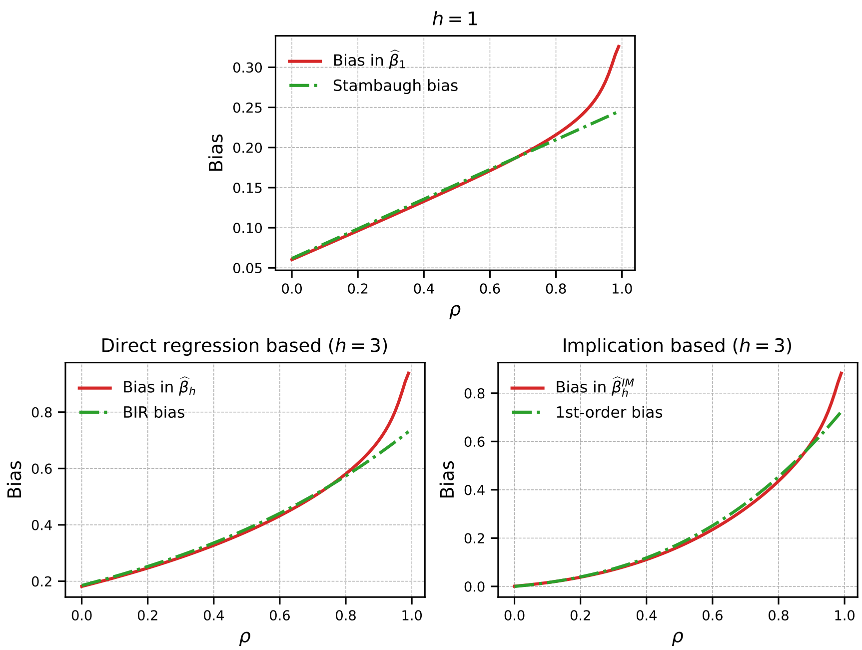
$$E[\overline{\overline{\widehat{\beta}}}_1^{II}] \quad = \quad E[\widehat{\beta}_1] - \phi E[\widehat{\rho} - \widehat{\rho}^{II}] = E[\widehat{\beta}_1] - \phi E[\widehat{\rho} - \rho] = E[\widehat{\beta}_1] - E[\widehat{\beta}_1] + \beta_1 = \beta_1,$$

where the second equation holds because $E[\widehat{\rho}^{II}] = \rho$ when the binding function is affine, shown in Gouriéroux et al. (2000), and the third equation follows from (3.1).

7.8 Additional Bias Figures when ρ varies from 0 to 1



(a) csp



(b) rsvix

Figure 16: Bias in the predictive regression for $T = 100$ and horizons $h = 1$ (short) and $h = 3$ (long). The solid line is the total bias in predictive regressions, and the dashed line is its first-order approximation: Stambaugh bias (Stambaugh, 1999) for $h = 1$; BIR bias (Boudoukh et al., 2022) for the direct regression (DR) estimator and the first-order bias in implication-based (IM) estimator for $h = 3$. Simulations use Eqs. (2.1) and (2.2) with $\alpha_1 = \beta_1 = c = 0$ and Σ are from cross-sectional premium (csp) and scaled risk-neutral vix (rsvix) residual covariance matrix, respectively.

7.9 Algorithm to implement the iBoot estimator

The iBoot estimator can be implemented in the following steps:

Step 1. For a given value ρ in $(-1, 1)$, simulate the AR(1) process x_t with $c = 0$ and $\sigma^2 = 1$ using the same sample size T as the original data. Compute the OLS estimate $\widetilde{\rho}$ from each simulated data. Repeat this procedure S times to obtain $\{\widetilde{\rho}_1, ..., \widetilde{\rho}_S\}$ and calculate the binding function value at ρ as $b_T(\rho) = \frac{1}{S}\sum_{i=1}^{S}\widetilde{\rho}_i$. Choose a large S and repeat the simulation over a fine grid of ρ in $(-1, 1)$ to obtain the full binding function $b_T(\rho)$.

Step 2. Using the original data, estimate $\widehat{\rho}$ via OLS. Obtain the indirect inference estimate $\widehat{\rho}^{II}$ by numerically inverting the binding function $b_T(\widehat{\rho}^{II}) = \widehat{\rho}$.

Step 3. Use $\widehat{\rho}^{II}$ to calculate $\widehat{c}^{II} = \frac{1}{T}\sum_{t=1}^{T} x_t - \widehat{\rho}^{II}\frac{1}{T}\sum_{t=1}^{T} x_{t-1}$, $\widehat{\beta}_1^{II} = \widehat{\beta}_1 - \frac{\widehat{\sigma}_{uv}}{\widehat{\sigma}_v^2}(\widehat{\rho} - \widehat{\rho}^{II})$, $\widehat{\alpha}_1^{II} = \frac{1}{T}\sum_{t=1}^{T} r_t - \widehat{\beta}_1^{II}\frac{1}{T}\sum_{t=1}^{T} x_{t-1}$, and calculate the residuals as $\widehat{u}_t^{II} = r_t - \widehat{\alpha}_1^{II} - \widehat{\beta}_1^{II} x_{t-1}$ and $\widehat{v}_t^{II} = x_t - \widehat{c}^{II} - \widehat{\rho}^{II} x_{t-1}$.

Step 4. Generate wild bootstrap samples by scaling residuals $(u_t^b, v_t^b) = \xi_t \cdot (\widehat{u}_t^{II}, \widehat{v}_t^{II})$ for $1 \leq t \leq T$, in which $\{\xi_t\}_{t=1}^{T}$ is a sequence of i.i.d. random variables, and calculate the bootstrap sample (r_t^b, x_t^b) as follows:

$$
\begin{aligned}
r_t^b &= \widehat{\alpha}_1^{II} + \widehat{\beta}_1^{II} x_{t-1}^b + u_t^b, \\
x_t^b &= \widehat{c}^{II} + \widehat{\rho}^{II} x_{t-1}^b + v_t^b,
\end{aligned}
$$

with initialization $x_0^b = x_0$.

Step 5. From each bootstrap sample, compute:

- the DR estimator $\widehat{\beta}_h^{II,b}$ following Eq. (3.4), and
- the IM estimator $\widehat{\beta}_h^{IM,II,b}$ following Eq. (3.5).

Step 6. Repeat Steps 2-5 B times to obtain bootstrap replicates $\widehat{\beta}_h^{II,1}, \ldots, \widehat{\beta}_h^{II,B}$ for the DR estimator and $\widehat{\beta}_h^{IM,II,1}, \ldots, \widehat{\beta}_h^{IM,II,B}$ for the IM estimator. Compute the average bootstrap biases:

$$
\text{bias}_h^{DR} = \frac{1}{B}\sum_{b=1}^{B}\left(\widehat{\beta}_h^{II,b} - \beta_h^*\right), \quad \text{bias}_h^{IM} = \frac{1}{B}\sum_{b=1}^{B}\left(\widehat{\beta}_h^{IM,II,b} - \beta_h^*\right),
$$

where $\beta_h^* = \sum_{i=0}^{h-1}(\widehat{\rho}^{II})^i\widehat{\beta}_1^{II}$ is the "true" value in the bootstrap world.

Step 7. Define the final iBoot estimators as:

$$
\widehat{\beta}_h^{iBootDR} = \widehat{\beta}_h^{II} - \text{bias}_h^{DR}, \quad \widehat{\beta}_h^{iBootIM} = \widehat{\beta}_h^{IM,II} - \text{bias}_h^{IM}.
$$

7.10 Why do bootstrap methods using first-order bias correction fail to remove bias effectively?

To illustrate, consider the single-horizon regression with true parameters (α_1, β_1) and generate bootstrap samples as:

$$r_t^* = \alpha_1 + \beta_1 x_{t-1} + u_t^*,$$
$$x_t^* = \widehat{c}_{bc} + \widehat{\rho}_{bc} x_{t-1} + v_t^*,$$

where $\widehat{c}_{bc}$ and $\widehat{\rho}_{bc}$ are first-order bias-corrected estimates from the original data, and u_t^*, v_t^* are i.i.d. resampled residuals preserving their empirical covariance.

We then generate B bootstrap estimates $\widetilde{\beta}_1^1, \ldots, \widetilde{\beta}_1^B$ and estimate the bias as $\frac{1}{B}\sum_{b=1}^B(\widetilde{\beta}_1^b - \beta_1)$. By Proposition 2.2, this bias is approximately

$$-\frac{\sigma_{uv}}{\sigma_v^2}\left(\frac{3\widehat{\rho}_{bc}+1}{T} + \frac{9\widehat{\rho}_{bc}^3 + 6\widehat{\rho}_{bc}^2 - 5\widehat{\rho}_{bc}}{T^2(1-\widehat{\rho}_{bc}^2)}\right)$$

up to $O(T^{-2})$ for large B (assuming $\gamma_1 = \gamma_2 = 0$).

However, this bootstrap bias substantially underestimates the true bias. The reason is that $\widehat{\rho}_{bc}$ itself remains downward biased due to uncorrected higher-order terms (see Eq. (7.11)). This bias is further magnified by the denominator term $1 - \widehat{\rho}_{bc}^2$, especially when ρ is near unity. Consequently, the bootstrap methods – even when combined with first-order corrections – fail to remove the bias effectively.

7.11 Algorithm to implement the iBoot inference method

The iBoot inference procedure builds on the iBoot estimation steps described in Appendix 7.9, with the following additional steps:

Step 1. When $h = 1$, define $\widehat{\beta}_1^{iBoot\ DR}(\alpha)$ and $\widehat{\beta}_1^{iBoot\ IM}(\alpha)$ as the α-quantile of the B bootstrap estimates $\left\{\widehat{\beta}_1^{II,1}, \ldots, \widehat{\beta}_1^{II,B}\right\}$, since the DR and IM estimators coincide at horizon $h = 1$.

When $h > 1$, construct the quantiles separately for the two estimators. For the iBoot DR estimator, let $\widehat{\beta}_h^{iBoot\ DR}(\alpha)$ denote the α-quantile of the B bootstrap estimates centered around the iBoot DR point estimate, i.e.,

$$\left\{\widehat{\beta}_h^{iBootDR} + \left(\widehat{\beta}_h^{II,1} - \frac{1}{B}\sum_{b=1}^{B}\widehat{\beta}_h^{II,b}\right), \ldots, \widehat{\beta}_h^{iBootDR} + \left(\widehat{\beta}_h^{II,B} - \frac{1}{B}\sum_{b=1}^{B}\widehat{\beta}_h^{II,b}\right)\right\}.$$

For the iBoot IM estimator, let $\widehat{\beta}_h^{iBoot\ IM}(\alpha)$ be the α-quantile of the B bootstrap estimates $\left\{\widehat{\beta}_h^{II,IM,1}, \ldots, \widehat{\beta}_h^{II,IM,B}\right\}$.

Step 2. Construct the $(1-\alpha)$ confidence intervals for β_h as

$$[\widehat{\beta}_h^{iBoot\ DR}(\alpha/2),\ \widehat{\beta}_h^{iBoot\ DR}(1-\alpha/2)] \quad \text{and} \quad [\widehat{\beta}_h^{iBoot\ IM}(\alpha/2),\ \widehat{\beta}_h^{iBoot\ IM}(1-\alpha/2)]$$

for the iBoot DR and iBoot IM estimators, respectively.

Step 3. To test the null hypothesis $\mathbb{H}_0 : \beta_h = 0$ against the alternative $\mathbb{H}_1 : \beta_h \neq 0$, reject $\mathbb{H}_0$ at significance level α if zero is not contained in the corresponding confidence interval, i.e., if

$$0 \notin [\widehat{\beta}_h^{iBoot\ DR}(\alpha/2),\ \widehat{\beta}_h^{iBoot\ DR}(1-\alpha/2)] \quad \text{or} \quad 0 \notin [\widehat{\beta}_h^{iBoot\ IM}(\alpha/2),\ \widehat{\beta}_h^{iBoot\ IM}(1-\alpha/2)].$$

7.12 Additional tables and figures in Section 4

7.12.1 Bias tables of the DR estimators

Tables 7-9 presents the bias, variance, and RMSE for the estimators from direct regression: $\hat{\beta}_h$, $\hat{\beta}_h^{BIR}$, and $\hat{\beta}_h^{iBoot\ DR}$ at various values of ρ, h, and T, corresponding to Figures 5-6 in the main text.

7.12.2 Bias tables of the IM estimators

Tables 10-12 presents the bias, variance, and RMSE for the implied estimators: $\hat{\beta}_h^{IM}$, $\hat{\beta}_h^{FO\ IM}$, and $\hat{\beta}_h^{iBoot\ IM}$ at various values of ρ, h, and T, corresponding to Figure 7 in the main text.

Table 7: Bias, Variance, and RMSE of the DR Estimators when $T = 30, 60$ in Section 4

This table reports the bias, variance, and RMSE of the direct regression (DR) estimators: the OLS estimator $\hat{\beta}_h$, the bias-corrected estimator $\hat{\beta}_h^{BIR}$ proposed by Boudoukh et al. (2022), and iBoot estimator $\hat{\beta}_h^{iBootDR}$ proposed in this paper. The simulation design follows Section 4.

ρ	h	$\hat{\beta}_h$			$\hat{\beta}_h^{BIR}$			$\hat{\beta}_h^{iBootDR}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
T = 30										
0.70	1	0.203	0.106	0.383	0.044	0.123	0.354	-0.004	0.143	0.379
	2	0.368	0.294	0.655	0.056	0.380	0.619	-0.013	0.494	0.703
0.85	1	0.252	0.089	0.391	0.071	0.104	0.331	0.002	0.126	0.355
	2	0.466	0.258	0.690	0.111	0.343	0.596	0.001	0.446	0.668
0.90	1	0.273	0.083	0.397	0.086	0.097	0.324	0.018	0.114	0.338
	2	0.509	0.242	0.707	0.140	0.324	0.586	0.034	0.401	0.634
0.95	1	0.295	0.075	0.403	0.102	0.088	0.314	0.042	0.099	0.317
	2	0.551	0.220	0.724	0.170	0.299	0.572	0.080	0.345	0.593
0.99	1	0.301	0.068	0.398	0.101	0.080	0.300	0.054	0.083	0.293
	2	0.563	0.202	0.721	0.167	0.277	0.553	0.103	0.291	0.549
T = 60										
0.70	1	0.096	0.044	0.232	0.010	0.048	0.219	0.001	0.049	0.222
	2	0.179	0.138	0.412	0.014	0.158	0.397	0.002	0.169	0.411
	3	0.256	0.253	0.564	0.022	0.299	0.547	0.008	0.337	0.581
0.85	1	0.119	0.033	0.217	0.021	0.036	0.190	-0.003	0.041	0.203
	2	0.227	0.107	0.398	0.035	0.125	0.355	-0.006	0.151	0.388
	3	0.328	0.204	0.558	0.054	0.250	0.503	-0.006	0.321	0.566
0.90	1	0.131	0.028	0.213	0.028	0.031	0.178	-0.003	0.037	0.193
	2	0.251	0.094	0.396	0.050	0.111	0.336	-0.007	0.138	0.371
	3	0.363	0.182	0.560	0.076	0.225	0.481	-0.009	0.295	0.543
0.95	1	0.147	0.024	0.213	0.041	0.026	0.166	0.008	0.030	0.173
	2	0.284	0.080	0.401	0.075	0.095	0.317	0.016	0.112	0.335
	3	0.413	0.155	0.571	0.114	0.195	0.456	0.026	0.239	0.489
0.99	1	0.159	0.020	0.212	0.050	0.021	0.154	0.027	0.022	0.152
	2	0.308	0.067	0.402	0.093	0.079	0.296	0.053	0.084	0.294
	3	0.449	0.130	0.576	0.140	0.165	0.430	0.081	0.178	0.430

Table 8: Bias, Variance, and RMSE of the DR Estimators when $T = 100, 300$ in Section 4

This table reports the bias, variance, and RMSE of the direct regression (DR) estimators: the OLS estimator $\hat{\beta}_h$, the bias-corrected estimator $\hat{\beta}_h^{BIR}$ proposed by Boudoukh et al. (2022), and iBoot estimator $\hat{\beta}_h^{iBootDR}$ proposed in this paper. The simulation design follows Section 4.

ρ	h	$\hat{\beta}_h$			$\hat{\beta}_h^{BIR}$			$\hat{\beta}_h^{iBootDR}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 100					
0.70	1	0.055	0.024	0.165	0.002	0.025	0.159	-0.001	0.026	0.160
	3	0.145	0.145	0.408	0.002	0.161	0.402	-0.003	0.172	0.415
	5	0.222	0.304	0.594	0.002	0.343	0.586	-0.005	0.384	0.619
0.85	1	0.067	0.016	0.143	0.006	0.017	0.130	-0.001	0.018	0.133
	3	0.186	0.108	0.378	0.016	0.123	0.351	-0.002	0.136	0.369
	5	0.290	0.243	0.572	0.024	0.288	0.537	-0.002	0.343	0.585
0.90	1	0.073	0.013	0.135	0.009	0.014	0.117	-0.003	0.016	0.125
	3	0.206	0.092	0.367	0.026	0.105	0.325	-0.009	0.126	0.355
	5	0.324	0.210	0.562	0.042	0.254	0.506	-0.015	0.327	0.572
0.95	1	0.083	0.010	0.130	0.017	0.011	0.104	-0.003	0.013	0.112
	3	0.237	0.074	0.361	0.048	0.085	0.296	-0.009	0.105	0.324
	5	0.378	0.172	0.561	0.080	0.212	0.468	-0.014	0.276	0.525
0.99	1	0.096	0.007	0.128	0.028	0.008	0.092	0.013	0.008	0.092
	3	0.277	0.056	0.364	0.081	0.065	0.268	0.039	0.070	0.267
	5	0.444	0.132	0.573	0.135	0.166	0.429	0.065	0.183	0.433
					T = 300					
0.70	1	0.018	0.007	0.088	-0.000	0.008	0.087	-0.001	0.008	0.087
	8	0.106	0.209	0.469	-0.003	0.218	0.467	-0.004	0.233	0.482
	16	0.188	0.478	0.717	-0.010	0.497	0.705	-0.009	0.586	0.766
0.85	1	0.022	0.004	0.070	0.000	0.004	0.067	-0.000	0.005	0.067
	8	0.139	0.167	0.432	0.002	0.180	0.424	-0.001	0.194	0.440
	16	0.243	0.444	0.710	-0.000	0.482	0.695	-0.004	0.553	0.744
0.90	1	0.023	0.003	0.062	0.001	0.003	0.058	-0.000	0.003	0.058
	8	0.156	0.137	0.402	0.008	0.150	0.387	0.000	0.160	0.401
	16	0.277	0.390	0.683	0.010	0.437	0.661	0.001	0.503	0.709
0.95	1	0.025	0.002	0.052	0.002	0.002	0.046	0.000	0.002	0.047
	8	0.183	0.095	0.358	0.020	0.106	0.326	0.001	0.118	0.343
	16	0.334	0.291	0.635	0.035	0.344	0.588	0.003	0.422	0.650
0.99	1	0.031	0.001	0.045	0.008	0.001	0.034	0.001	0.001	0.035
	8	0.236	0.054	0.331	0.061	0.062	0.256	0.012	0.071	0.267
	16	0.446	0.174	0.610	0.121	0.218	0.482	0.023	0.266	0.517

Table 9: Bias, Variance, and RMSE of the DR Estimators when $T = 600$ in Section 4

This table reports the bias, variance, and RMSE of the direct regression (DR) estimators: the OLS estimator $\hat{\beta}_h$, the bias-corrected estimator $\hat{\beta}_h^{BIR}$ proposed by Boudoukh et al. (2022), and iBoot estimator $\hat{\beta}_h^{iBootDR}$ proposed in this paper. The simulation design follows Section 4.

ρ	h	$\hat{\beta}_h$			$\hat{\beta}_h^{BIR}$			$\hat{\beta}_h^{iBootDR}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 600					
0.70	1	0.009	0.004	0.060	-0.000	0.004	0.060	-0.000	0.004	0.060
	12	0.077	0.176	0.427	0.001	0.179	0.424	0.001	0.190	0.436
	24	0.147	0.400	0.649	0.007	0.406	0.637	0.008	0.469	0.685
	36	0.217	0.611	0.811	0.009	0.620	0.788	0.013	0.795	0.892
	48	0.275	0.806	0.939	-0.004	0.817	0.904	0.001	1.178	1.085
0.85	1	0.011	0.002	0.046	0.000	0.002	0.045	0.000	0.002	0.045
	12	0.100	0.154	0.405	0.005	0.160	0.400	0.003	0.169	0.411
	24	0.181	0.399	0.657	0.011	0.414	0.643	0.011	0.461	0.679
	36	0.254	0.638	0.838	0.010	0.660	0.812	0.013	0.795	0.892
	48	0.316	0.859	0.979	-0.004	0.887	0.942	0.000	1.174	1.084
0.90	1	0.011	0.001	0.039	0.000	0.001	0.038	0.000	0.001	0.038
	12	0.113	0.129	0.377	0.007	0.136	0.369	0.005	0.144	0.379
	24	0.202	0.361	0.634	0.014	0.383	0.619	0.011	0.423	0.650
	36	0.277	0.599	0.822	0.012	0.633	0.796	0.011	0.746	0.864
	48	0.343	0.824	0.970	-0.001	0.869	0.932	0.000	1.112	1.054
0.95	1	0.012	0.001	0.031	0.001	0.001	0.029	0.000	0.001	0.029
	12	0.132	0.088	0.325	0.013	0.095	0.309	0.006	0.100	0.316
	24	0.237	0.272	0.573	0.022	0.301	0.549	0.012	0.337	0.580
	36	0.324	0.478	0.764	0.020	0.535	0.732	0.010	0.628	0.793
	48	0.401	0.687	0.921	0.012	0.771	0.878	0.002	0.967	0.983
0.97	1	0.013	0.001	0.027	0.001	0.001	0.024	0.000	0.001	0.024
	12	0.143	0.066	0.295	0.017	0.072	0.269	0.006	0.077	0.278
	24	0.263	0.212	0.530	0.031	0.241	0.492	0.011	0.278	0.527
	36	0.364	0.384	0.719	0.036	0.450	0.672	0.011	0.552	0.743
	48	0.454	0.567	0.879	0.035	0.674	0.822	0.004	0.874	0.935
0.99	1	0.015	0.000	0.024	0.003	0.000	0.019	-0.000	0.000	0.020
	12	0.168	0.041	0.263	0.035	0.045	0.215	-0.001	0.054	0.232
	24	0.318	0.136	0.487	0.068	0.161	0.407	-0.005	0.202	0.450
	36	0.452	0.254	0.677	0.096	0.318	0.572	-0.010	0.426	0.653
	48	0.574	0.384	0.845	0.118	0.502	0.718	-0.017	0.705	0.840

Table 10: Bias, Variance, and RMSE of the IM Estimators when $T = 30, 60$ in Section 4

This table reports the bias, variance, and RMSE of the implied (IM) estimators: the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator $\hat{\beta}_h^{iBootIM}$ proposed in this paper. The simulation design follows Section 4.

ρ	h	$\hat{\beta}_h^{IM}$			$\hat{\beta}_h^{FO\ IM}$			$\hat{\beta}_h^{iBootIM}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 30					
0.70	1	0.203	0.106	0.383	0.032	0.125	0.354	-0.004	0.143	0.379
	2	0.275	0.212	0.536	0.052	0.315	0.563	-0.020	0.398	0.631
0.85	1	0.252	0.089	0.391	0.057	0.106	0.330	0.002	0.126	0.355
	2	0.390	0.199	0.592	0.105	0.302	0.560	-0.013	0.388	0.623
0.90	1	0.273	0.083	0.397	0.072	0.098	0.322	0.018	0.114	0.338
	2	0.440	0.190	0.619	0.136	0.291	0.556	0.019	0.353	0.594
0.95	1	0.295	0.075	0.403	0.087	0.089	0.311	0.042	0.098	0.317
	2	0.492	0.177	0.648	0.170	0.273	0.549	0.069	0.306	0.558
0.99	1	0.301	0.068	0.398	0.086	0.081	0.297	0.054	0.083	0.293
	2	0.517	0.164	0.657	0.173	0.256	0.535	0.097	0.259	0.518
					T = 60					
0.70	1	0.096	0.044	0.232	0.007	0.048	0.219	0.001	0.049	0.222
	2	0.139	0.107	0.355	0.011	0.131	0.362	-0.000	0.135	0.368
	3	0.156	0.159	0.428	0.012	0.209	0.458	-0.003	0.219	0.468
0.85	1	0.119	0.033	0.217	0.017	0.036	0.190	-0.003	0.041	0.203
	2	0.198	0.090	0.360	0.032	0.112	0.337	-0.008	0.135	0.367
	3	0.250	0.148	0.459	0.044	0.203	0.453	-0.017	0.255	0.505
0.90	1	0.131	0.028	0.213	0.024	0.031	0.178	-0.003	0.037	0.193
	2	0.226	0.082	0.364	0.047	0.102	0.323	-0.011	0.127	0.356
	3	0.295	0.138	0.474	0.066	0.191	0.442	-0.022	0.249	0.499
0.95	1	0.147	0.024	0.213	0.037	0.026	0.166	0.008	0.030	0.173
	2	0.263	0.071	0.374	0.073	0.089	0.307	0.012	0.104	0.323
	3	0.356	0.122	0.499	0.106	0.172	0.428	0.012	0.207	0.456
0.99	1	0.159	0.020	0.212	0.046	0.021	0.153	0.027	0.022	0.152
	2	0.293	0.059	0.381	0.093	0.075	0.289	0.051	0.079	0.285
	3	0.407	0.104	0.520	0.138	0.149	0.410	0.071	0.157	0.402

Table 11: Bias, Variance, and RMSE of the IM Estimators when $T = 100, 300$ in Section 4

This table reports the bias, variance, and RMSE of the implied (IM) estimators: the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator $\hat{\beta}_h^{iBootIM}$ proposed in this paper. The simulation design follows Section 4.

ρ	h	$\hat{\beta}_h^{IM}$			$\hat{\beta}_h^{FO\ IM}$			$\hat{\beta}_h^{iBootIM}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 100					
0.70	1	0.055	0.024	0.165	0.001	0.025	0.159	-0.001	0.026	0.160
	3	0.092	0.097	0.325	0.001	0.115	0.339	-0.005	0.117	0.342
	5	0.095	0.145	0.393	-0.002	0.181	0.426	-0.008	0.185	0.431
0.85	1	0.067	0.016	0.143	0.005	0.017	0.130	-0.001	0.018	0.133
	3	0.148	0.084	0.325	0.013	0.102	0.320	-0.003	0.110	0.331
	5	0.187	0.151	0.431	0.017	0.202	0.450	-0.006	0.222	0.471
0.90	1	0.073	0.013	0.135	0.008	0.014	0.117	-0.003	0.016	0.125
	3	0.174	0.074	0.324	0.023	0.091	0.303	-0.011	0.108	0.329
	5	0.236	0.142	0.444	0.034	0.194	0.442	-0.021	0.243	0.493
0.95	1	0.083	0.010	0.130	0.015	0.011	0.104	-0.003	0.013	0.112
	3	0.213	0.062	0.328	0.045	0.077	0.281	-0.012	0.095	0.309
	5	0.307	0.125	0.468	0.071	0.176	0.425	-0.027	0.228	0.478
0.99	1	0.096	0.007	0.128	0.026	0.008	0.092	0.013	0.008	0.092
	3	0.259	0.048	0.339	0.080	0.060	0.257	0.036	0.064	0.255
	5	0.394	0.100	0.505	0.132	0.144	0.402	0.055	0.155	0.397
					T = 300					
0.70	1	0.018	0.007	0.088	-0.000	0.008	0.087	-0.001	0.008	0.087
	8	0.030	0.067	0.260	-0.003	0.073	0.270	-0.003	0.073	0.271
	16	0.025	0.076	0.277	-0.004	0.082	0.286	-0.002	0.082	0.286
0.85	1	0.022	0.004	0.070	0.000	0.004	0.067	-0.000	0.005	0.067
	8	0.073	0.087	0.304	-0.000	0.100	0.316	-0.004	0.101	0.318
	16	0.071	0.136	0.375	-0.003	0.161	0.401	-0.005	0.163	0.404
0.90	1	0.023	0.003	0.062	0.001	0.003	0.058	-0.000	0.003	0.058
	8	0.101	0.085	0.309	0.004	0.100	0.316	-0.002	0.102	0.319
	16	0.115	0.158	0.414	0.002	0.199	0.446	-0.004	0.204	0.452
0.95	1	0.025	0.002	0.052	0.002	0.002	0.046	0.000	0.002	0.047
	8	0.143	0.071	0.302	0.016	0.085	0.293	-0.001	0.092	0.303
	16	0.203	0.163	0.452	0.023	0.221	0.470	-0.005	0.247	0.497
0.99	1	0.031	0.001	0.045	0.007	0.001	0.034	0.001	0.001	0.035
	8	0.214	0.046	0.303	0.059	0.056	0.244	0.008	0.065	0.255
	16	0.367	0.123	0.508	0.112	0.179	0.438	0.006	0.217	0.466

Table 12: Bias, Variance, and RMSE of the IM Estimators when $T = 600$ in Section 4

This table reports the bias, variance, and RMSE of the implied (IM) estimators: the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator $\hat{\beta}_h^{iBootIM}$ proposed in this paper. The simulation design follows Section 4.

ρ	h	$\hat{\beta}_h^{IM}$			$\hat{\beta}_h^{FO\ IM}$			$\hat{\beta}_h^{iBootIM}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 600					
0.70	1	0.009	0.004	0.060	-0.000	0.004	0.060	-0.000	0.004	0.060
	12	0.014	0.037	0.192	-0.001	0.038	0.195	-0.001	0.038	0.195
	24	0.013	0.038	0.195	-0.001	0.039	0.198	-0.000	0.039	0.198
	36	0.013	0.038	0.195	-0.001	0.039	0.198	-0.000	0.039	0.198
	48	0.013	0.038	0.195	-0.001	0.039	0.198	-0.000	0.039	0.198
0.85	1	0.011	0.002	0.046	0.000	0.002	0.045	0.000	0.002	0.045
	12	0.039	0.059	0.246	0.000	0.064	0.253	-0.000	0.064	0.253
	24	0.035	0.077	0.279	-0.000	0.084	0.289	0.000	0.084	0.289
	36	0.034	0.080	0.285	-0.001	0.087	0.294	0.001	0.086	0.294
	48	0.033	0.081	0.286	-0.001	0.087	0.295	0.002	0.086	0.294
0.90	1	0.011	0.001	0.039	0.000	0.001	0.038	0.000	0.001	0.038
	12	0.060	0.064	0.259	0.003	0.071	0.266	0.000	0.071	0.266
	24	0.061	0.101	0.324	0.002	0.116	0.340	0.001	0.116	0.340
	36	0.057	0.115	0.344	0.001	0.131	0.362	0.003	0.130	0.361
	48	0.055	0.120	0.350	0.001	0.135	0.368	0.005	0.134	0.366
0.95	1	0.012	0.001	0.031	0.001	0.001	0.029	0.000	0.001	0.029
	12	0.094	0.058	0.258	0.008	0.066	0.257	0.002	0.067	0.258
	24	0.122	0.123	0.371	0.011	0.151	0.388	0.003	0.154	0.392
	36	0.128	0.166	0.427	0.011	0.211	0.459	0.005	0.215	0.464
	48	0.126	0.193	0.457	0.009	0.248	0.498	0.008	0.252	0.502
0.97	1	0.013	0.001	0.027	0.001	0.001	0.024	0.000	0.001	0.024
	12	0.115	0.049	0.249	0.014	0.056	0.237	0.004	0.058	0.242
	24	0.170	0.119	0.385	0.023	0.152	0.390	0.006	0.161	0.401
	36	0.196	0.178	0.465	0.026	0.242	0.492	0.008	0.261	0.511
	48	0.206	0.223	0.515	0.026	0.316	0.562	0.011	0.344	0.587
0.99	1	0.015	0.000	0.024	0.003	0.000	0.019	-0.000	0.000	0.020
	12	0.151	0.034	0.238	0.033	0.040	0.202	-0.004	0.048	0.218
	24	0.258	0.095	0.402	0.062	0.127	0.361	-0.012	0.160	0.400
	36	0.335	0.158	0.520	0.085	0.233	0.490	-0.025	0.310	0.557
	48	0.391	0.215	0.607	0.102	0.345	0.596	-0.041	0.484	0.697

7.12.3 Coverage rates comparison against alternatives

We evaluate the empirical coverage of 95% confidence intervals from our iBoot estimators against three alternatives: (1) Newey-West (NW) intervals using the OLS estimator and Newey-West standard errors; (2) Hodrick (HOD) intervals using the OLS estimator and Hodrick standard errors; and (3) BIR intervals using the first-order bias-reduced estimator of Boudoukh et al. (2022) with their analytical AR(1) standard errors.

The top panel of Figure 17 plots coverage rates of β_h across different persistence levels (ρ) for small samples ($T = 30$, $h = 1, 2$; $T = 60$, $h = 1, 3$). iBoot IM (solid line) tracks the 95% nominal level almost perfectly across all scenarios – even with small samples, high persistence ($\rho = 0.99$), and long horizons. iBoot DR (dash-dotted line) performs the second best, with mild undercoverage at extreme ρ and h.

By contrast, NW, HOD, and BIR intervals exhibit severe undercoverage. At $T = 30$ and $h = 1$, all three fall below 90% for every ρ, with NW and HOD dropping under 80% (HOD below 70%) as ρ approaches one. At $h = 2$, coverage deteriorates further: near $\rho = 1$, NW falls below 60%, HOD around 60%, and BIR near 80%. In all these cases, iBoot IM consistently maintains accurate 95% coverage, while iBoot DR remains close behind.

Coverage improves as sample size increases, as expected, yet iBoot IM remains uniquely precise. For instance, at $T = 60$ and $h = 1$, NW and HOD stay below 80% and BIR dips under 90% when ρ nears one, whereas both iBoot IM and iBoot DR match the 95% nominal rate. At $h = 3$, discrepancies widen again – NW plunges below 60%, HOD near 70%, and BIR around 85% – while iBoot IM holds steady and iBoot DR shows only slight undercoverage.

The bottom panel of Figure 17 (larger samples, $T = 100, 300$) reinforces this pattern. At $h = 1$, alternatives improve but never achieve iBoot IM's exact coverage, and the gap expands with longer horizons. Even at $T = 300$, as $\rho \to 1$, NW and HOD barely reach 85%, BIR barely reaches 90%, while iBoot IM stays at 95% and iBoot DR remains close. At $h = 16$, the divergence becomes striking: NW coverage collapses below 70%, HOD to 80%, and BIR to 90%, whereas iBoot IM continues to align precisely with the nominal level.

Finally, Figure 18 (large sample, $T = 600$) confirms these results. Even with large sample, alternatives still under-cover when ρ is high, particularly at long horizons ($h = 12, 24, 36, 48$): NW coverage falls near or below 70%, HOD around 85%, and BIR below 90%. In contrast, iBoot IM maintains exact 95% coverage throughout, and iBoot DR performs nearly as well.

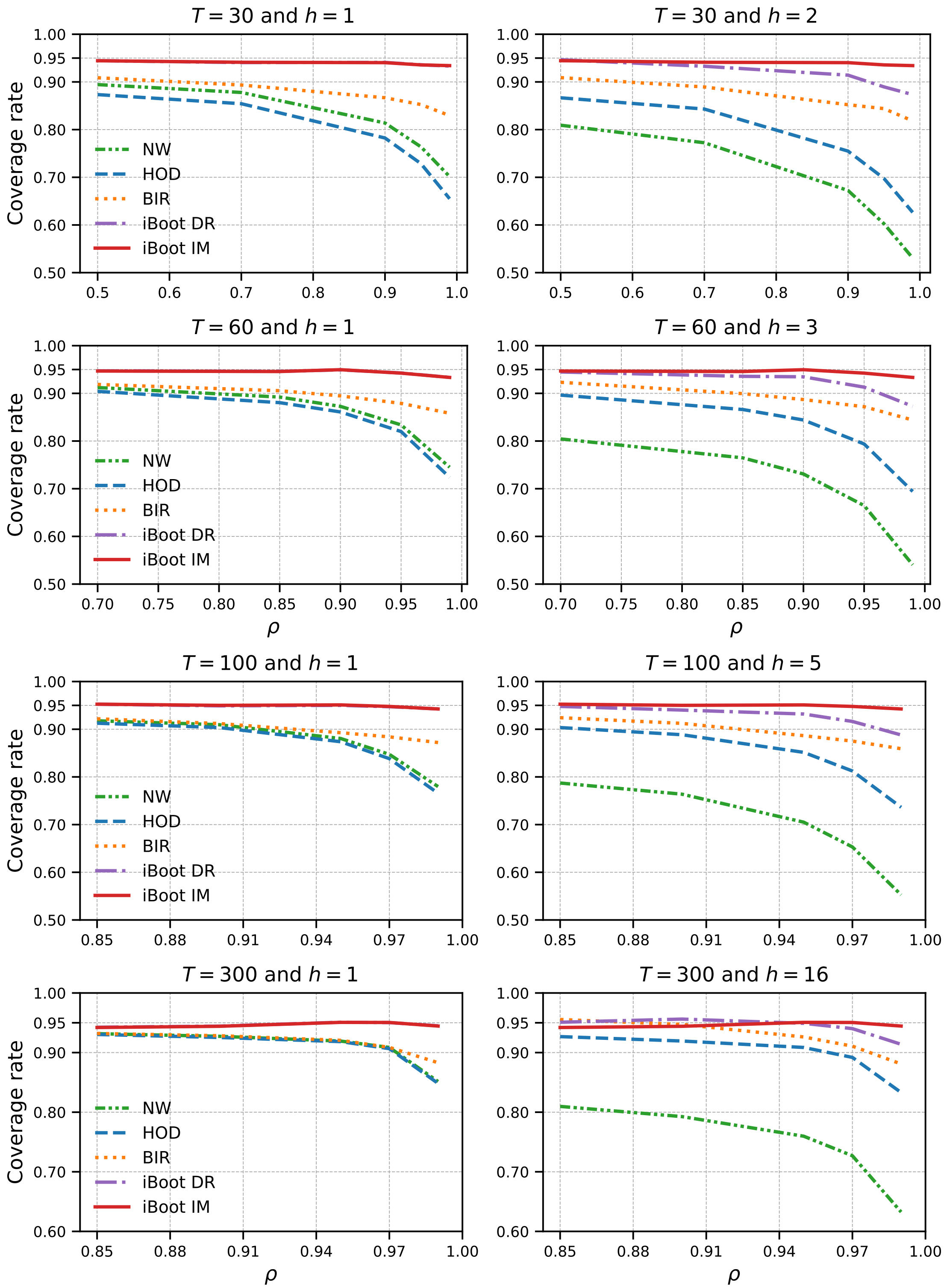


Figure 17: Empirical coverage rates of 95% Confidence Intervals (CIs) for various methods. NW means the Newey-West CIs, HOD the Hodrick CIs, BIR the CIs using the BIR estimator and the standard errors proposed in Boudoukh et al. (2022), iBoot DR and iBoot IM the CIs proposed in this paper, for the DR and IM estimators respectively.

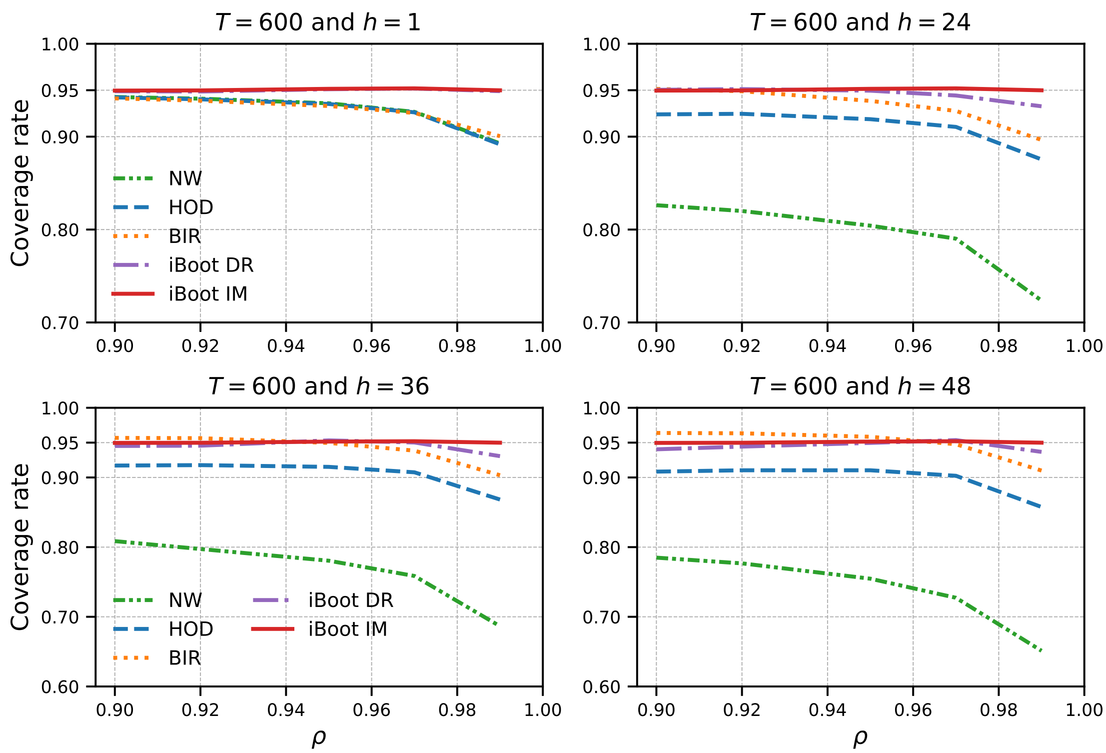


Figure 18: Empirical coverage rates of 95% Confidence Intervals (CIs) for various methods. NW means the Newey-West CIs, HOD the Hodrick CIs, BIR the CIs using the BIR estimator and the standard errors proposed in Boudoukh et al. (2022), iBoot DR and iBoot IM the CIs proposed in this paper, for the DR and IM estimators respectively.

7.13 Simulation Results When the true parameter $\beta_1 = -0.16$

This section reports additional simulation results when the true parameter is $\beta_1 = -0.16$. The simulation design follows Section 4, except for the value of β_1.

Figures 19-20 display the bias of various DR estimators across sample sizes and horizons, and Figure 21 shows results for the IM estimators. The OLS bias becomes even larger when $\beta_1 = -0.16$ than when $\beta_1 = 0$. The BIR and FO IM estimators reduce this bias partially, but the iBoot estimators consistently yield the smallest bias across all designs. Detailed values appear in Tables 13-18.

Figures 22-23 present the empirical coverage rates of 95% confidence intervals, and Figures 24-25 show size curves under $H_0 : \beta_1 = -0.16$. The iBoot methods maintain coverage close to the nominal level and exhibit accurate size control, whereas OLS, HOD, and BIR display even greater size distortions than in the $\beta_1 = 0$ case. Even at $T = 600$, their distortions persist across horizons, while iBoot remains close to the nominal size. Because of these distortions, we omit power comparisons for this setting.

Overall, the results for $\beta_1 = -0.16$ reinforce our main findings: iBoot effectively removes bias and maintains valid inference for predictive regressions with persistent regressors and long horizons. The performance gap between iBoot and conventional methods becomes even more pronounced when the true coefficient is -0.16, confirming the robustness of iBoot across parameter values.

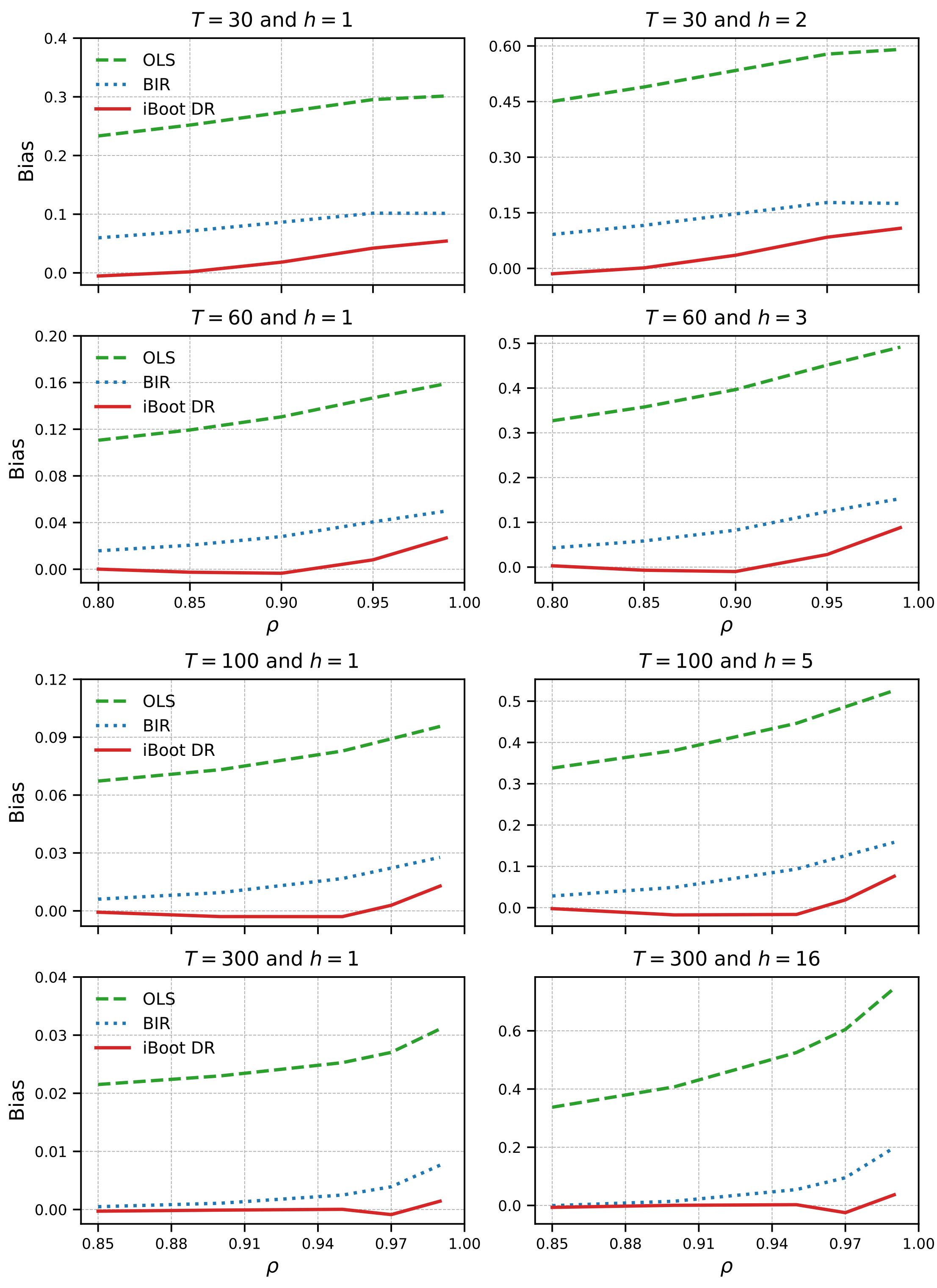


Figure 19: Bias plots for various DR estimators when $\beta_1 = -0.16$. OLS means the OLS estimator, BIR the BIR estimator proposed in Boudoukh et al. (2022), and iBoot DR the iBoot estimator proposed in this paper.

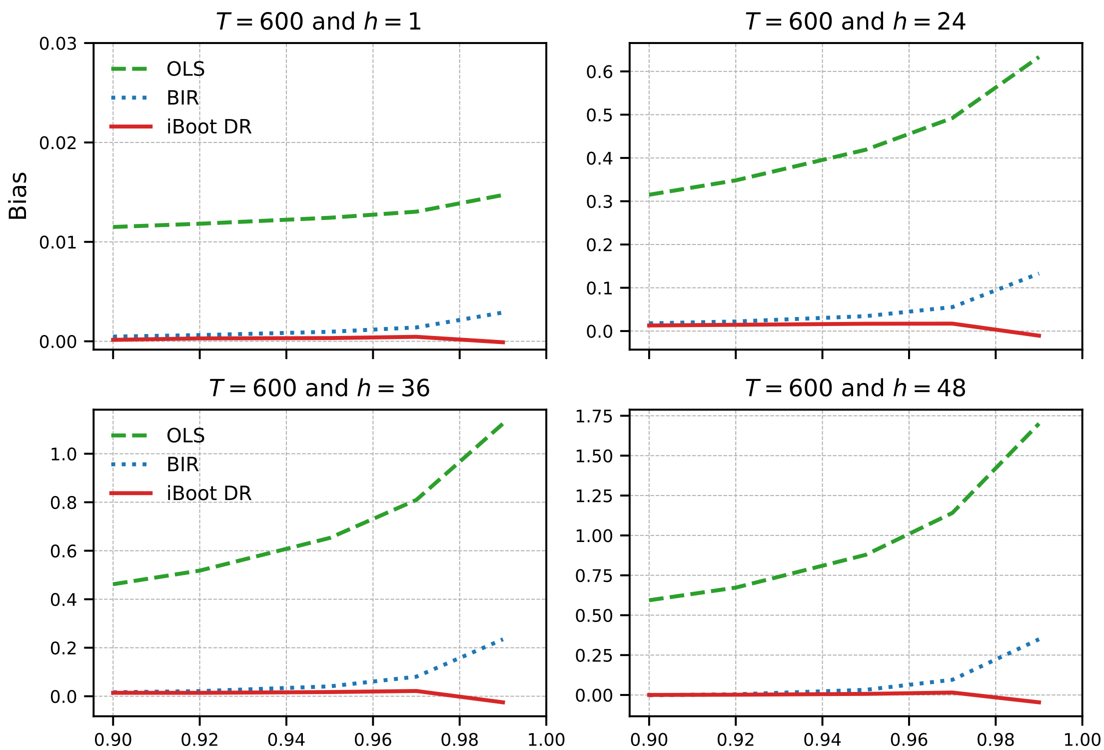


Figure 20: Bias plots for various DR estimators when $\beta_1 = -0.16$ across different horizons when $T = 600$. OLS means the OLS estimator, BIR the BIR estimator proposed in Boudoukh et al. (2022), and iBoot DR the iBoot estimator proposed in this paper.

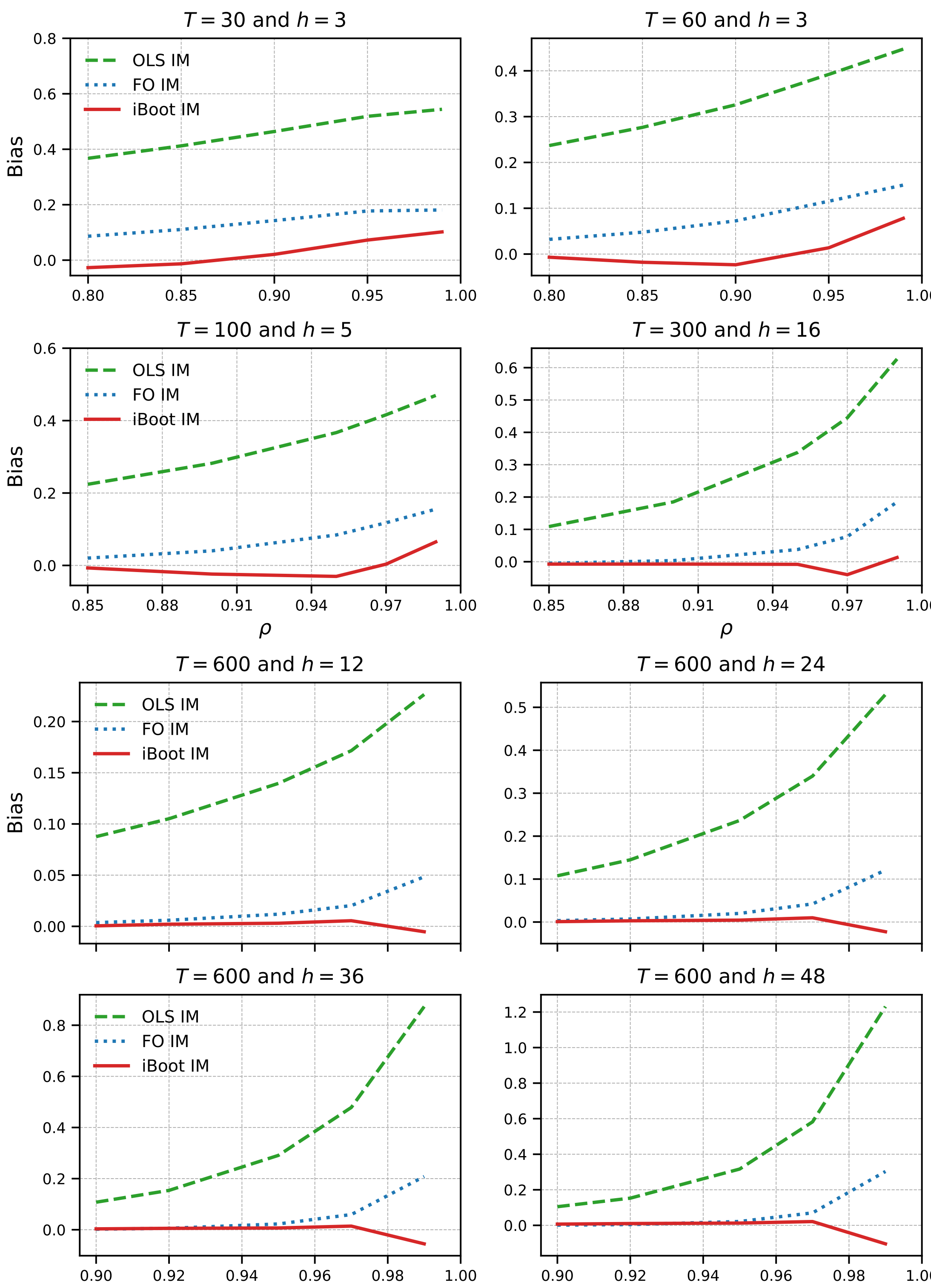


Figure 21: Bias of various Implication-Based (IM) estimators at long horizons when $\beta_1 = -0.16$. Compares the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator proposed in this paper.

Table 13: Bias, Variance, and RMSE of the DR Estimators when $\beta_1 = -0.16$ and $T = 30, 60$

This table reports the bias, variance, and RMSE of the direct regression (DR) estimators: the OLS estimator $\hat{\beta}_h$, the bias-corrected estimator $\hat{\beta}_h^{BIR}$ proposed by Boudoukh et al. (2022), and iBoot estimator $\hat{\beta}_h^{iBootDR}$ proposed in this paper. The simulation design follows Section 4 except for $\beta_1 = -0.16$.

ρ	h	$\hat{\beta}_h$			$\hat{\beta}_h^{BIR}$			$\hat{\beta}_h^{iBootDR}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 30					
0.70	1	0.203	0.106	0.383	0.044	0.123	0.354	-0.004	0.143	0.379
	2	0.387	0.317	0.683	0.059	0.411	0.644	-0.013	0.531	0.729
0.85	1	0.252	0.089	0.391	0.071	0.104	0.331	0.002	0.126	0.355
	2	0.489	0.279	0.720	0.116	0.371	0.620	0.001	0.481	0.694
0.90	1	0.273	0.083	0.397	0.086	0.097	0.324	0.018	0.114	0.338
	2	0.534	0.261	0.739	0.147	0.350	0.610	0.036	0.433	0.659
0.95	1	0.295	0.075	0.403	0.102	0.088	0.314	0.042	0.099	0.317
	2	0.578	0.239	0.757	0.178	0.323	0.596	0.084	0.372	0.616
0.99	1	0.301	0.068	0.398	0.101	0.080	0.300	0.054	0.083	0.293
	2	0.591	0.219	0.754	0.175	0.300	0.575	0.108	0.314	0.571
					T = 60					
0.70	1	0.096	0.044	0.232	0.010	0.048	0.219	0.001	0.049	0.222
	2	0.188	0.148	0.428	0.015	0.169	0.412	0.002	0.181	0.426
	3	0.278	0.288	0.604	0.024	0.341	0.584	0.008	0.383	0.619
0.85	1	0.119	0.033	0.217	0.021	0.036	0.190	-0.003	0.041	0.203
	2	0.238	0.116	0.415	0.037	0.134	0.369	-0.006	0.162	0.403
	3	0.358	0.235	0.603	0.058	0.288	0.540	-0.007	0.369	0.607
0.90	1	0.131	0.028	0.213	0.028	0.031	0.178	-0.003	0.037	0.193
	2	0.262	0.102	0.413	0.052	0.119	0.349	-0.008	0.148	0.385
	3	0.397	0.210	0.606	0.083	0.260	0.517	-0.010	0.340	0.583
0.95	1	0.147	0.024	0.213	0.041	0.026	0.166	0.008	0.030	0.173
	2	0.297	0.087	0.418	0.078	0.102	0.329	0.017	0.121	0.348
	3	0.451	0.181	0.620	0.124	0.227	0.492	0.028	0.277	0.527
0.99	1	0.159	0.020	0.212	0.050	0.021	0.154	0.027	0.022	0.152
	2	0.323	0.072	0.420	0.097	0.085	0.308	0.056	0.091	0.306
	3	0.491	0.152	0.627	0.153	0.192	0.465	0.088	0.207	0.464

Table 14: Bias, Variance, and RMSE of the DR Estimators when $\beta_1 = -0.16$ and $T = 100, 300$

This table reports the bias, variance, and RMSE of the direct regression (DR) estimators: the OLS estimator $\hat{\beta}_h$, the bias-corrected estimator $\hat{\beta}_h^{BIR}$ proposed by Boudoukh et al. (2022), and iBoot estimator $\hat{\beta}_h^{iBootDR}$ proposed in this paper. The simulation design follows Section 4 except for $\beta_1 = -0.16$.

ρ	h	$\hat{\beta}_h$			$\hat{\beta}_h^{BIR}$			$\hat{\beta}_h^{iBootDR}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
0.70	1	0.055	0.024	0.165	0.002	0.025	0.159	-0.001	0.026	0.160
	3	0.158	0.166	0.437	0.003	0.184	0.429	-0.004	0.195	0.442
	5	0.254	0.375	0.663	0.002	0.425	0.652	-0.006	0.474	0.689
0.85	1	0.067	0.016	0.143	0.006	0.017	0.130	-0.001	0.018	0.133
	3	0.203	0.125	0.407	0.017	0.142	0.377	-0.002	0.156	0.395
	5	0.338	0.311	0.652	0.028	0.369	0.608	-0.002	0.437	0.661
0.90	1	0.073	0.013	0.135	0.009	0.014	0.117	-0.003	0.016	0.125
	3	0.225	0.106	0.396	0.028	0.121	0.349	-0.010	0.145	0.381
	5	0.381	0.273	0.647	0.049	0.330	0.576	-0.018	0.425	0.652
0.95	1	0.083	0.010	0.130	0.017	0.011	0.104	-0.003	0.013	0.112
	3	0.259	0.086	0.391	0.052	0.099	0.318	-0.009	0.122	0.349
	5	0.446	0.227	0.653	0.093	0.280	0.537	-0.016	0.364	0.603
0.99	1	0.096	0.007	0.128	0.028	0.008	0.092	0.013	0.008	0.092
	3	0.302	0.065	0.395	0.088	0.076	0.289	0.042	0.081	0.288
	5	0.526	0.178	0.674	0.158	0.222	0.497	0.076	0.245	0.501
					T = 300					
0.70	1	0.018	0.007	0.088	-0.000	0.008	0.087	-0.001	0.008	0.087
	8	0.126	0.277	0.541	-0.004	0.289	0.538	-0.005	0.309	0.556
	16	0.234	0.688	0.862	-0.013	0.716	0.847	-0.012	0.846	0.920
0.85	1	0.022	0.004	0.070	0.000	0.004	0.067	-0.000	0.005	0.067
	8	0.173	0.240	0.520	0.002	0.259	0.509	-0.002	0.277	0.527
	16	0.337	0.764	0.937	-0.001	0.834	0.913	-0.007	0.956	0.978
0.90	1	0.023	0.003	0.062	0.001	0.003	0.058	-0.000	0.003	0.058
	8	0.199	0.203	0.493	0.009	0.222	0.472	-0.000	0.237	0.487
	16	0.408	0.737	0.950	0.014	0.831	0.912	0.001	0.955	0.977
0.95	1	0.025	0.002	0.052	0.002	0.002	0.046	0.000	0.002	0.047
	8	0.237	0.147	0.451	0.025	0.164	0.406	0.001	0.182	0.426
	16	0.525	0.628	0.951	0.055	0.742	0.863	0.003	0.904	0.951
0.99	1	0.031	0.001	0.045	0.008	0.001	0.034	0.001	0.001	0.035
	8	0.310	0.088	0.429	0.080	0.099	0.325	0.015	0.115	0.339
	16	0.746	0.436	0.996	0.199	0.539	0.760	0.037	0.658	0.812

Table 15: Bias, Variance, and RMSE of the DR Estimators when $\beta_1 = -0.16$ and $T = 600$

This table reports the bias, variance, and RMSE of the direct regression (DR) estimators: the OLS estimator $\hat{\beta}_h$, the bias-corrected estimator $\hat{\beta}_h^{BIR}$ proposed by Boudoukh et al. (2022), and iBoot estimator $\hat{\beta}_h^{iBootDR}$ proposed in this paper. The simulation design follows Section 4 except for $\beta_1 = -0.16$.

ρ	h	$\hat{\beta}_h$			$\hat{\beta}_h^{BIR}$			$\hat{\beta}_h^{iBootDR}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
T = 600										
0.70	1	0.009	0.004	0.060	-0.000	0.004	0.060	-0.000	0.004	0.060
	12	0.094	0.246	0.505	0.001	0.251	0.501	0.001	0.266	0.516
	24	0.184	0.591	0.790	0.006	0.601	0.775	0.007	0.696	0.835
	36	0.273	0.917	0.996	0.009	0.931	0.965	0.014	1.204	1.098
	48	0.350	1.216	1.157	-0.006	1.235	1.111	0.000	1.805	1.344
0.85	1	0.011	0.002	0.046	0.000	0.002	0.045	0.000	0.002	0.045
	12	0.132	0.246	0.513	0.005	0.256	0.506	0.004	0.271	0.520
	24	0.259	0.750	0.904	0.012	0.780	0.883	0.011	0.869	0.932
	36	0.377	1.276	1.191	0.012	1.323	1.150	0.015	1.599	1.265
	48	0.480	1.766	1.413	-0.006	1.829	1.352	-0.000	2.439	1.562
0.90	1	0.011	0.001	0.039	0.000	0.001	0.038	0.000	0.001	0.038
	12	0.155	0.220	0.494	0.009	0.232	0.482	0.005	0.244	0.494
	24	0.315	0.785	0.940	0.018	0.835	0.914	0.013	0.923	0.961
	36	0.462	1.461	1.294	0.016	1.553	1.246	0.014	1.831	1.353
	48	0.594	2.131	1.576	-0.002	2.262	1.504	0.000	2.899	1.703
0.95	1	0.012	0.001	0.031	0.001	0.001	0.029	0.000	0.001	0.029
	12	0.188	0.163	0.446	0.016	0.175	0.419	0.007	0.183	0.428
	24	0.419	0.730	0.951	0.034	0.809	0.900	0.017	0.901	0.949
	36	0.653	1.614	1.428	0.041	1.818	1.349	0.017	2.128	1.459
	48	0.878	2.679	1.858	0.032	3.045	1.745	0.007	3.792	1.947
0.97	1	0.013	0.001	0.027	0.001	0.001	0.024	0.000	0.001	0.024
	12	0.208	0.127	0.413	0.024	0.138	0.372	0.008	0.147	0.384
	24	0.492	0.638	0.938	0.055	0.724	0.853	0.017	0.827	0.909
	36	0.810	1.572	1.492	0.081	1.842	1.360	0.022	2.235	1.495
	48	1.141	2.869	2.042	0.095	3.440	1.857	0.015	4.414	2.101
0.99	1	0.015	0.000	0.024	0.003	0.000	0.019	-0.000	0.000	0.020
	12	0.249	0.083	0.380	0.051	0.091	0.306	-0.003	0.109	0.331
	24	0.633	0.474	0.936	0.133	0.554	0.756	-0.011	0.701	0.837
	36	1.124	1.344	1.615	0.235	1.652	1.307	-0.025	2.205	1.485
	48	1.701	2.799	2.386	0.349	3.598	1.929	-0.046	5.018	2.241

Table 16: Bias, Variance, and RMSE of the IM Estimators when $\beta_1 = -0.16$ and $T = 30, 60$

This table reports the bias, variance, and RMSE of the implied (IM) estimators: the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator $\hat{\beta}_h^{iBootIM}$ proposed in this paper. The simulation design follows Section 4 except for $\beta_1 = -0.16$.

ρ	h	$\hat{\beta}_h^{IM}$			$\hat{\beta}_h^{FO\ IM}$			$\hat{\beta}_h^{iBootIM}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
0.70	1	0.203	0.106	0.383	0.032	0.125	0.354	-0.004	0.143	0.379
	2	0.293	0.233	0.565	0.054	0.344	0.589	-0.020	0.435	0.660
0.85	1	0.252	0.089	0.391	0.057	0.106	0.330	0.002	0.126	0.355
	2	0.412	0.218	0.623	0.111	0.329	0.584	-0.013	0.422	0.650
0.90	1	0.273	0.083	0.397	0.072	0.098	0.322	0.018	0.114	0.338
	2	0.464	0.208	0.651	0.143	0.317	0.580	0.021	0.384	0.620
0.95	1	0.295	0.075	0.403	0.087	0.089	0.311	0.042	0.098	0.317
	2	0.519	0.194	0.680	0.177	0.297	0.573	0.073	0.333	0.582
0.99	1	0.301	0.068	0.398	0.086	0.081	0.297	0.054	0.083	0.293
	2	0.544	0.180	0.690	0.181	0.279	0.558	0.102	0.282	0.540
					T = 60					
0.70	1	0.096	0.044	0.232	0.007	0.048	0.219	0.001	0.049	0.222
	2	0.147	0.117	0.372	0.012	0.143	0.378	0.000	0.148	0.384
	3	0.174	0.186	0.466	0.013	0.245	0.495	-0.002	0.256	0.506
0.85	1	0.119	0.033	0.217	0.017	0.036	0.190	-0.003	0.041	0.203
	2	0.209	0.099	0.377	0.034	0.122	0.351	-0.008	0.147	0.383
	3	0.277	0.174	0.501	0.048	0.237	0.490	-0.018	0.298	0.547
0.90	1	0.131	0.028	0.213	0.024	0.031	0.178	-0.003	0.037	0.193
	2	0.237	0.089	0.381	0.049	0.110	0.336	-0.011	0.138	0.371
	3	0.326	0.162	0.518	0.072	0.223	0.478	-0.023	0.291	0.540
0.95	1	0.147	0.024	0.213	0.037	0.026	0.166	0.008	0.030	0.173
	2	0.276	0.077	0.391	0.076	0.096	0.319	0.013	0.113	0.337
	3	0.392	0.144	0.546	0.115	0.201	0.463	0.014	0.243	0.493
0.99	1	0.159	0.020	0.212	0.046	0.021	0.153	0.027	0.022	0.152
	2	0.307	0.065	0.399	0.097	0.081	0.301	0.053	0.085	0.297
	3	0.447	0.124	0.569	0.151	0.175	0.444	0.078	0.184	0.436

Table 17: Bias, Variance, and RMSE of the IM Estimators when $\beta_1 = -0.16$ and $T = 100, 300$

This table reports the bias, variance, and RMSE of the implied (IM) estimators: the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator $\hat{\beta}_h^{iBootIM}$ proposed in this paper. The simulation design follows Section 4 except for $\beta_1 = -0.16$.

ρ	h	$\hat{\beta}_h^{IM}$			$\hat{\beta}_h^{FO\ IM}$			$\hat{\beta}_h^{iBootIM}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
T = 100										
0.70	1	0.055	0.024	0.165	0.001	0.025	0.159	-0.001	0.026	0.160
	3	0.103	0.114	0.352	0.001	0.135	0.367	-0.005	0.137	0.370
	5	0.115	0.189	0.449	-0.002	0.236	0.486	-0.009	0.241	0.491
0.85	1	0.067	0.016	0.143	0.005	0.017	0.130	-0.001	0.018	0.133
	3	0.163	0.098	0.354	0.014	0.119	0.346	-0.003	0.128	0.358
	5	0.224	0.201	0.501	0.020	0.267	0.517	-0.007	0.294	0.542
0.90	1	0.073	0.013	0.135	0.008	0.014	0.117	-0.003	0.016	0.125
	3	0.192	0.087	0.352	0.025	0.107	0.327	-0.012	0.127	0.356
	5	0.282	0.190	0.520	0.040	0.259	0.510	-0.024	0.324	0.570
0.95	1	0.083	0.010	0.130	0.015	0.011	0.104	-0.003	0.013	0.112
	3	0.233	0.073	0.357	0.049	0.090	0.304	-0.013	0.111	0.334
	5	0.367	0.170	0.552	0.084	0.236	0.493	-0.030	0.307	0.555
0.99	1	0.096	0.007	0.128	0.026	0.008	0.092	0.013	0.008	0.092
	3	0.284	0.056	0.370	0.087	0.070	0.278	0.039	0.075	0.276
	5	0.470	0.138	0.599	0.155	0.195	0.468	0.065	0.210	0.463
T = 300										
0.70	1	0.018	0.007	0.088	-0.000	0.008	0.087	-0.001	0.008	0.087
	8	0.038	0.095	0.311	-0.004	0.104	0.322	-0.004	0.104	0.323
	16	0.033	0.117	0.343	-0.005	0.126	0.355	-0.003	0.126	0.355
0.85	1	0.022	0.004	0.070	0.000	0.004	0.067	-0.000	0.005	0.067
	8	0.096	0.133	0.377	-0.000	0.152	0.390	-0.005	0.154	0.393
	16	0.109	0.259	0.521	-0.005	0.309	0.556	-0.007	0.314	0.561
0.90	1	0.023	0.003	0.062	0.001	0.003	0.058	-0.000	0.003	0.058
	8	0.133	0.133	0.388	0.005	0.156	0.394	-0.003	0.158	0.398
	16	0.185	0.333	0.606	0.003	0.419	0.647	-0.007	0.429	0.655
0.95	1	0.025	0.002	0.052	0.002	0.002	0.046	0.000	0.002	0.047
	8	0.189	0.113	0.386	0.021	0.136	0.369	-0.002	0.146	0.383
	16	0.338	0.382	0.704	0.038	0.512	0.716	-0.008	0.573	0.757
0.99	1	0.031	0.001	0.045	0.007	0.001	0.034	0.001	0.001	0.035
	8	0.284	0.075	0.395	0.077	0.092	0.312	0.010	0.107	0.327
	16	0.627	0.325	0.847	0.185	0.460	0.703	0.013	0.557	0.747

Table 18: Bias, Variance, and RMSE of the IM Estimators when $\beta_1 = -0.16$ and $T = 600$

This table reports the bias, variance, and RMSE of the implied (IM) estimators: the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator $\hat{\beta}_h^{iBootIM}$ proposed in this paper. The simulation design follows Section 4 except for $\beta_1 = -0.16$.

ρ	h	$\hat{\beta}_h^{IM}$			$\hat{\beta}_h^{FO\ IM}$			$\hat{\beta}_h^{iBootIM}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 600					
0.70	1	0.009	0.004	0.060	-0.000	0.004	0.060	-0.000	0.004	0.060
	12	0.019	0.055	0.236	-0.001	0.058	0.240	-0.001	0.058	0.240
	24	0.017	0.059	0.243	-0.001	0.061	0.247	-0.000	0.061	0.247
	36	0.017	0.059	0.243	-0.001	0.061	0.247	-0.000	0.061	0.247
	48	0.017	0.059	0.243	-0.001	0.061	0.247	-0.000	0.061	0.247
0.85	1	0.011	0.002	0.046	0.000	0.002	0.045	0.000	0.002	0.045
	12	0.057	0.103	0.326	0.001	0.112	0.335	-0.000	0.112	0.335
	24	0.057	0.162	0.406	-0.000	0.177	0.421	0.001	0.177	0.421
	36	0.054	0.177	0.424	-0.001	0.193	0.439	0.002	0.192	0.439
	48	0.053	0.181	0.428	-0.001	0.196	0.443	0.003	0.195	0.441
0.90	1	0.011	0.001	0.039	0.000	0.001	0.038	0.000	0.001	0.038
	12	0.088	0.118	0.354	0.004	0.131	0.361	0.000	0.131	0.362
	24	0.108	0.253	0.514	0.003	0.289	0.538	0.001	0.290	0.539
	36	0.108	0.324	0.580	0.002	0.373	0.611	0.003	0.372	0.610
	48	0.105	0.356	0.606	0.001	0.408	0.638	0.007	0.404	0.636
0.95	1	0.012	0.001	0.031	0.001	0.001	0.029	0.000	0.001	0.029
	12	0.140	0.112	0.363	0.012	0.128	0.358	0.003	0.130	0.360
	24	0.237	0.374	0.656	0.020	0.456	0.676	0.004	0.466	0.683
	36	0.291	0.663	0.865	0.023	0.842	0.918	0.007	0.861	0.928
	48	0.317	0.912	1.006	0.022	1.184	1.088	0.013	1.208	1.099
0.97	1	0.013	0.001	0.027	0.001	0.001	0.024	0.000	0.001	0.024
	12	0.172	0.097	0.356	0.020	0.111	0.334	0.006	0.116	0.341
	24	0.340	0.398	0.716	0.042	0.499	0.708	0.010	0.531	0.728
	36	0.478	0.844	1.036	0.059	1.129	1.064	0.014	1.218	1.104
	48	0.582	1.356	1.302	0.070	1.900	1.380	0.021	2.074	1.440
0.99	1	0.015	0.000	0.024	0.003	0.000	0.019	-0.000	0.000	0.020
	12	0.226	0.070	0.348	0.048	0.081	0.289	-0.005	0.098	0.314
	24	0.530	0.355	0.797	0.121	0.460	0.689	-0.023	0.584	0.764
	36	0.873	0.922	1.298	0.208	1.307	1.162	-0.055	1.733	1.318
	48	1.231	1.785	1.817	0.303	2.729	1.679	-0.104	3.782	1.947

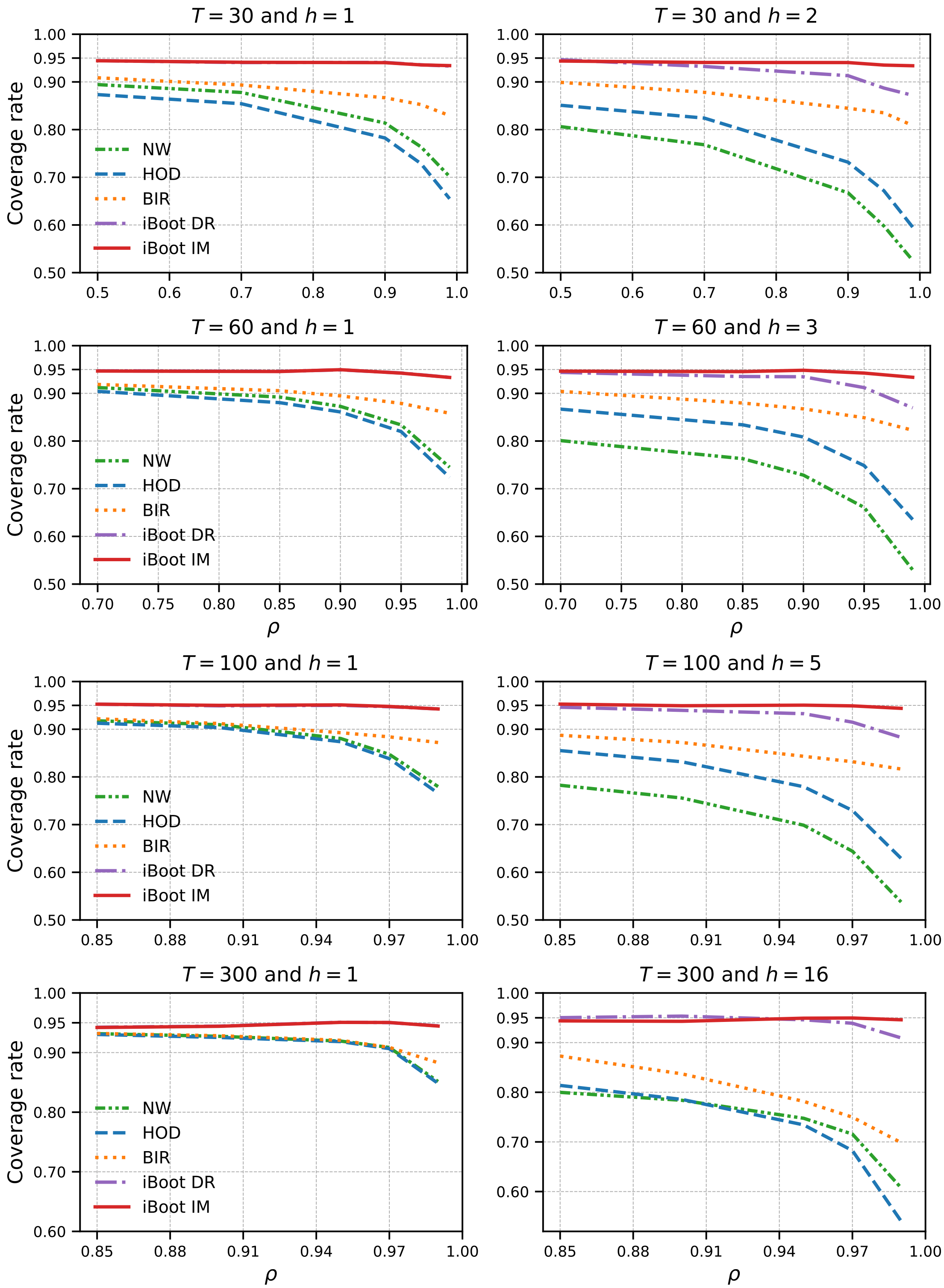


Figure 22: Empirical coverage rates of 95% Confidence Intervals (CIs) for various methods when $\beta_1 = -0.16$. NW means the Newey-West CIs, HOD the Hodrick CIs, BIR the CIs using the BIR estimator and the standard errors proposed in Boudoukh et al. (2022), iBoot DR and iBoot IM the CIs proposed in this paper, for the DR and IM estimators respectively.

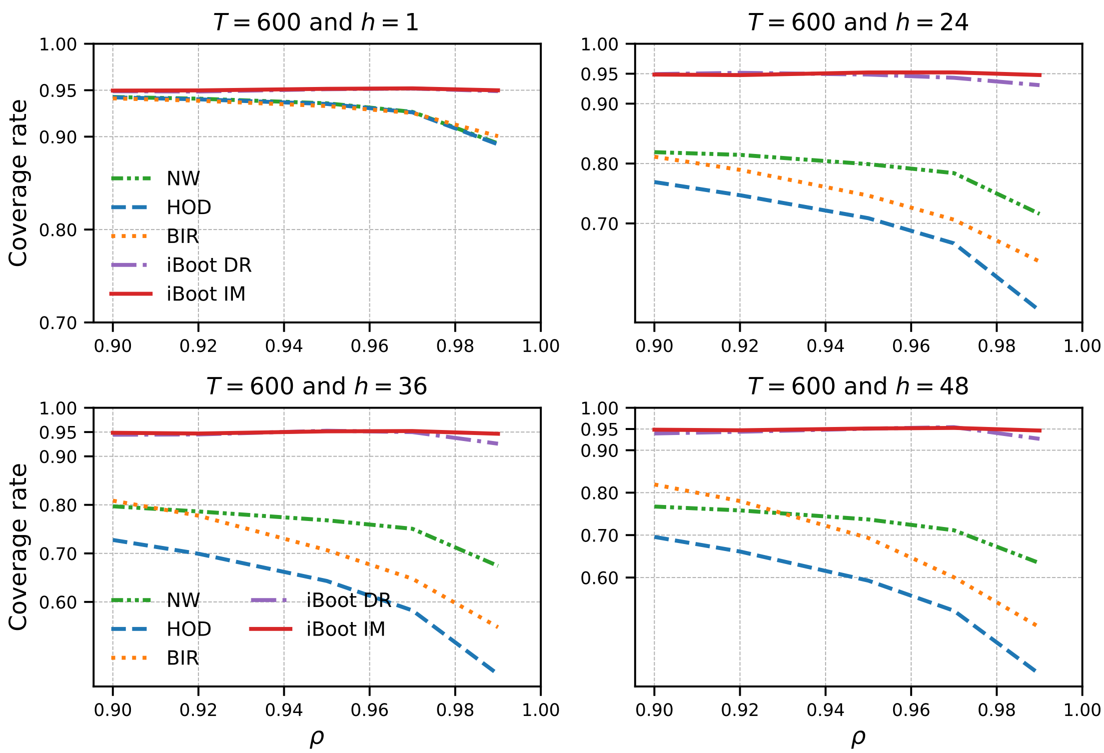


Figure 23: Empirical coverage rates of 95% Confidence Intervals (CIs) for various methods when $\beta_1 = -0.16$. NW means the Newey-West CIs, HOD the Hodrick CIs, BIR the CIs using the BIR estimator and the standard errors proposed in Boudoukh et al. (2022), iBoot DR and iBoot IM the CIs proposed in this paper, for the DR and IM estimators respectively.

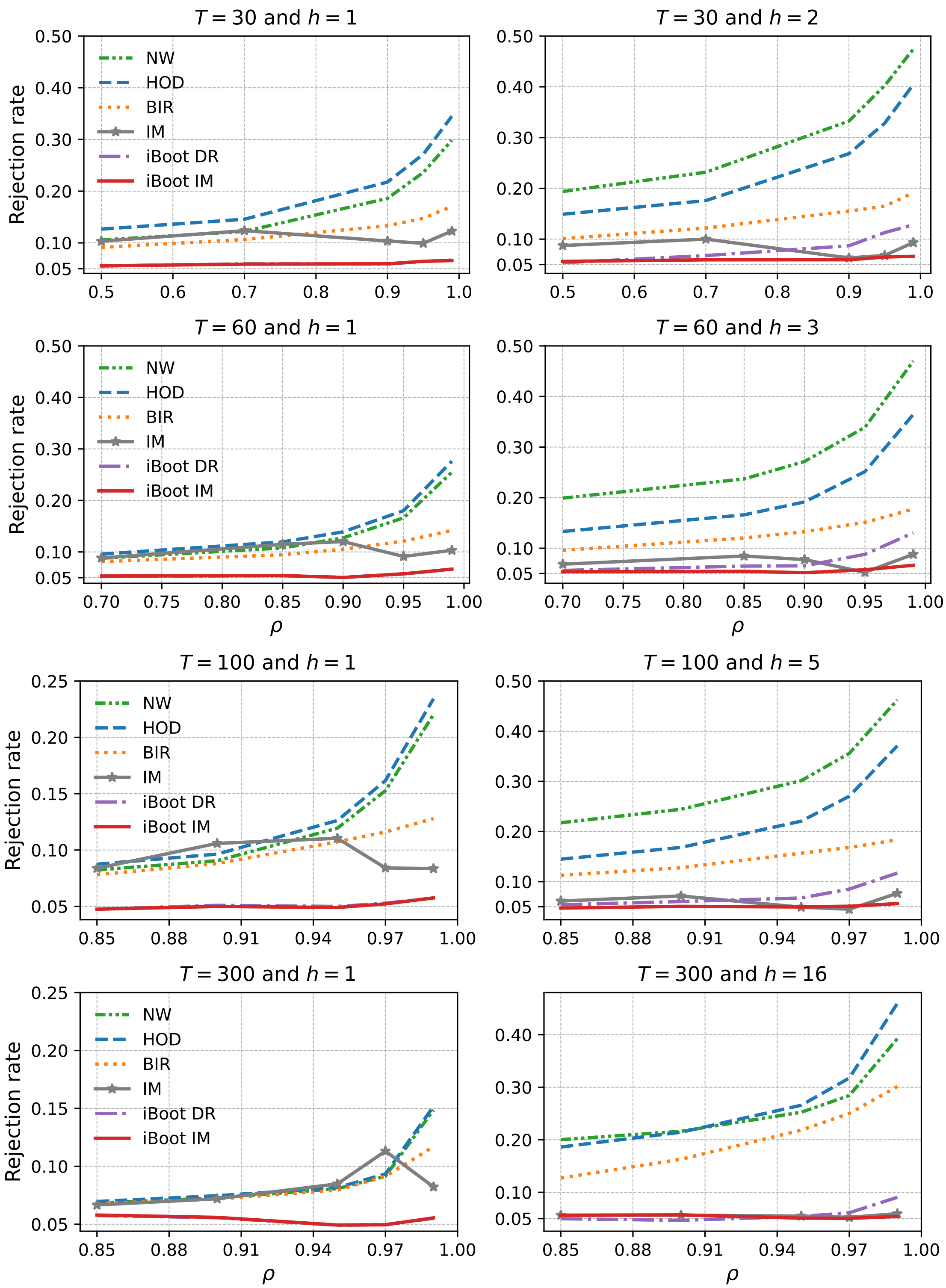


Figure 24: Size curves under $H_0 : \beta_1 = -0.16$ for various tests: Newey-West (NW), Hodrick (HOD), Boudoukh et al. (2022) (BIR), Xu (2020) (IM), and the proposed iBoot tests for Direct Regression (iBoot DR) and Implication-Based (iBoot IM) estimators.

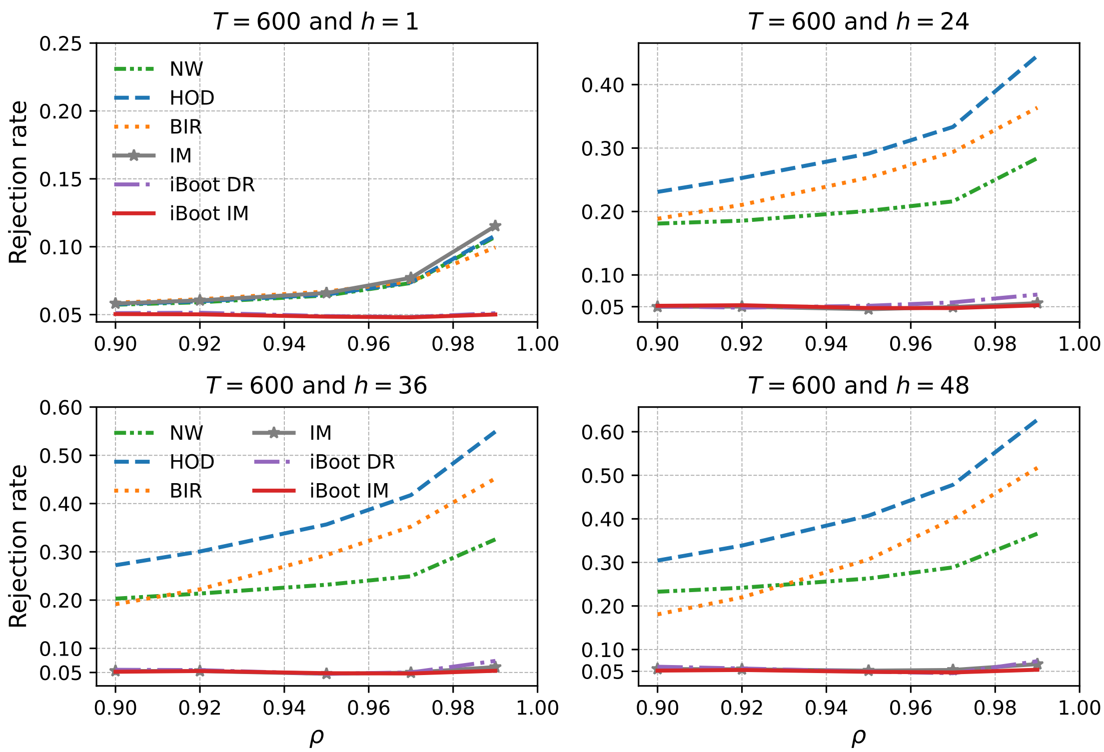


Figure 25: Size curves under $H_0 : \beta_1 = -0.16$ for various tests across horizons when $T = 600$: Newey-West (NW), Hodrick (HOD), Boudoukh et al. (2022) (BIR), Xu (2020) (IM), and the proposed iBoot tests for Direct Regression (iBoot DR) and Implication-Based (iBoot IM) estimators.

7.14 Simulation Results When Errors are Non-normal

This subsection examines the robustness of our results when the regression errors (u_t and v_t) in Eqs. (2.1) and (2.2) are drawn from non-normal distributions. Specifically, we generate v_t from a Pearson distribution with mean zero, standard deviation σ_v, skewness 1, and excess kurtosis 5. We then construct $u_t = \frac{\sigma_{uv}}{\sigma_v^2} v_t + e_t$, where e_t is independently drawn from a normal distribution with mean zero and variance $\sigma_e^2 = \sigma_u^2 - \frac{\sigma_{uv}^2}{\sigma_v^2}$. This ensures that u_t and v_t share the same mean and covariance structure as in the main simulations. All other parameter settings follow Section 4.

Figures 26-27 display the biases of alternative DR estimators under non-normal errors, while Figure 28 shows the corresponding results for IM estimators. The associated bias tables appear in Tables 19-21 for DR estimators and Tables 22-24 for IM estimators. Figures 29-33 report the coverage rates, test sizes, and powers.

Across all metrics – bias, coverage, size, and power – the results closely mirror those in Section 4. The iBoot estimators maintain minimal bias and accurate inference, confirming that the proposed framework remains robust even when the error terms deviate substantially from normality.

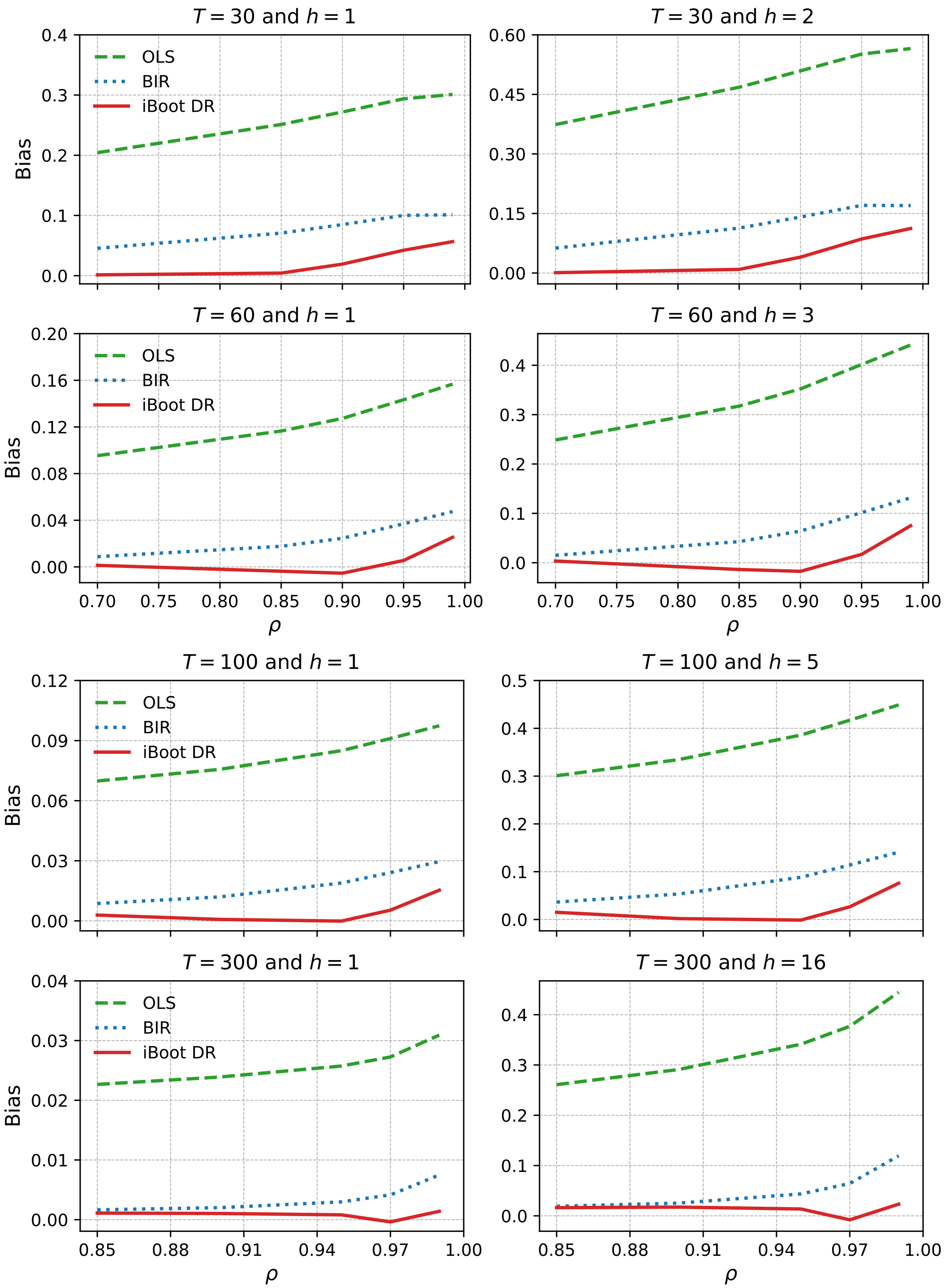


Figure 26: Bias plots for various DR estimators with non-normal errors. OLS means the OLS estimator, BIR the BIR estimator proposed in Boudoukh et al. (2022), and iBoot DR the iBoot estimator proposed in this paper.

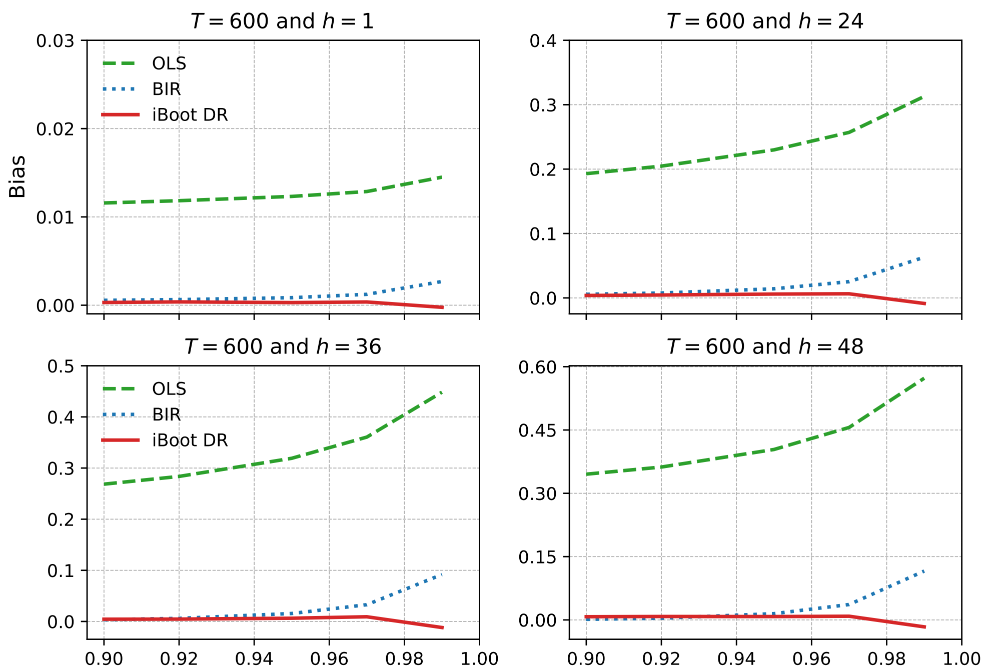


Figure 27: Bias plots for various DR estimators with non-normal errors when $T = 600$. OLS means the OLS estimator, BIR the BIR estimator proposed in Boudoukh et al. (2022), and iBoot DR the iBoot estimator proposed in this paper.

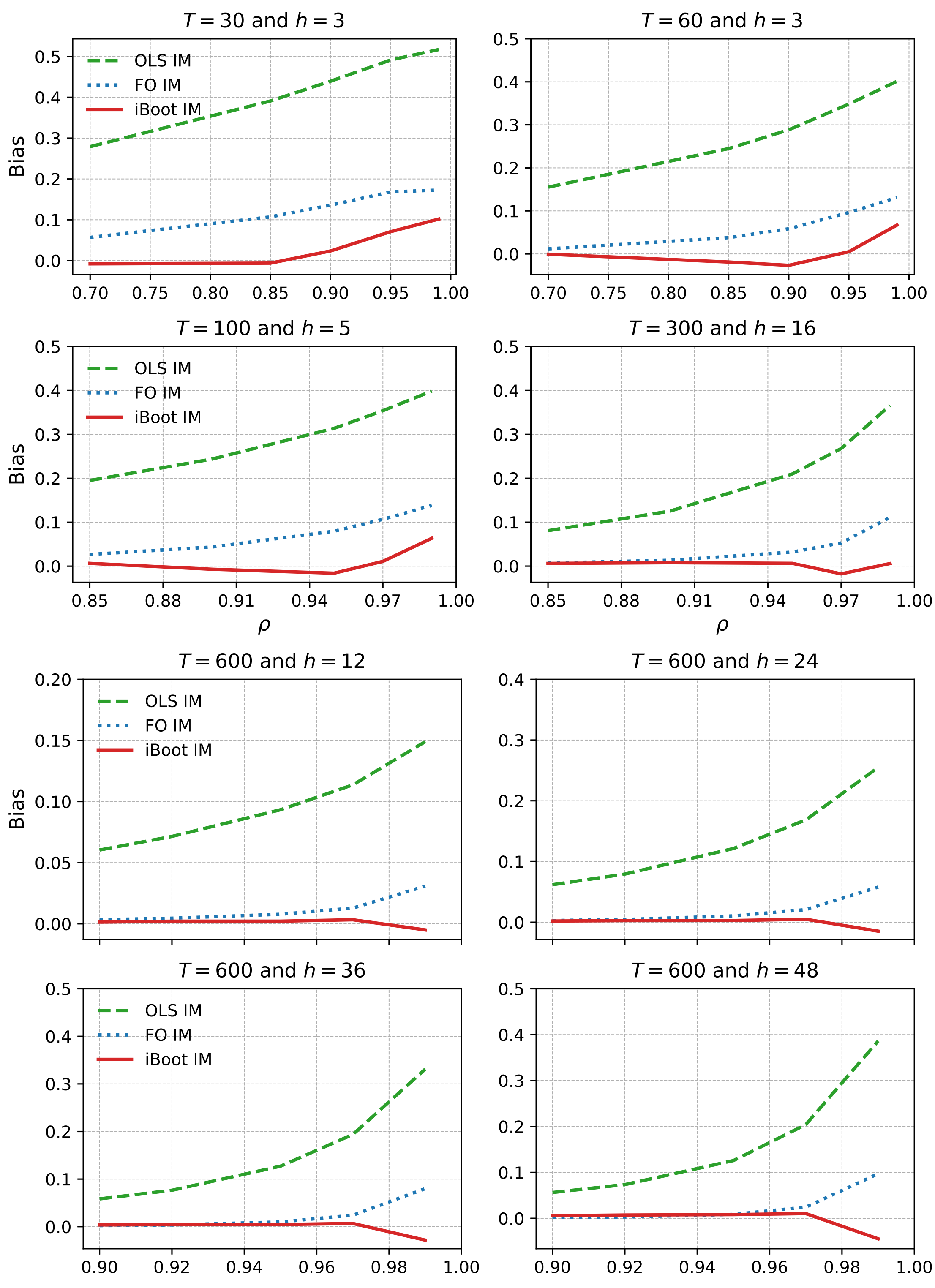


Figure 28: Bias of various Implication-Based (IM) estimators with non-normal errors at long horizons. Compares the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator proposed in this paper.

Table 19: Bias, Variance, and RMSE of the DR Estimators with Non-normal Errors

This table reports the bias, variance, and RMSE of the direct regression (DR) estimators when the errors in the predictive regression are non-normal and $T = 30, 60$: the OLS estimator $\hat{\beta}_h$, the bias-corrected estimator $\hat{\beta}_h^{BIR}$ proposed by Boudoukh et al. (2022), and iBoot estimator $\hat{\beta}_h^{iBootDR}$ proposed in this paper. The simulation design follows Section 7.14.

ρ	h	$\hat{\beta}_h$			$\hat{\beta}_h^{BIR}$			$\hat{\beta}_h^{iBootDR}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 30					
0.70	1	0.204	0.104	0.382	0.045	0.121	0.351	0.001	0.140	0.375
	2	0.374	0.295	0.660	0.063	0.380	0.619	0.001	0.490	0.700
0.85	1	0.251	0.088	0.388	0.071	0.102	0.328	0.004	0.123	0.351
	2	0.468	0.259	0.691	0.113	0.342	0.596	0.009	0.442	0.665
0.90	1	0.272	0.081	0.394	0.085	0.095	0.320	0.019	0.111	0.334
	2	0.509	0.241	0.708	0.141	0.323	0.585	0.040	0.397	0.631
0.95	1	0.294	0.074	0.401	0.100	0.087	0.312	0.042	0.097	0.314
	2	0.552	0.223	0.726	0.170	0.301	0.575	0.086	0.344	0.593
0.99	1	0.301	0.069	0.399	0.101	0.081	0.302	0.057	0.084	0.295
	2	0.565	0.210	0.728	0.170	0.287	0.562	0.112	0.299	0.558
					T = 60					
0.70	1	0.095	0.044	0.229	0.009	0.047	0.217	0.001	0.048	0.220
	2	0.176	0.132	0.404	0.011	0.152	0.390	0.001	0.162	0.403
	3	0.249	0.243	0.552	0.015	0.287	0.536	0.003	0.323	0.569
0.85	1	0.116	0.032	0.213	0.018	0.035	0.187	-0.004	0.040	0.200
	2	0.221	0.103	0.390	0.028	0.120	0.348	-0.009	0.145	0.381
	3	0.317	0.197	0.546	0.043	0.241	0.493	-0.014	0.310	0.557
0.90	1	0.127	0.028	0.210	0.025	0.030	0.176	-0.005	0.036	0.191
	2	0.243	0.091	0.388	0.042	0.107	0.330	-0.012	0.133	0.365
	3	0.352	0.176	0.548	0.064	0.218	0.471	-0.017	0.286	0.535
0.95	1	0.143	0.024	0.210	0.037	0.026	0.165	0.006	0.030	0.172
	2	0.276	0.078	0.393	0.067	0.092	0.310	0.010	0.108	0.329
	3	0.402	0.152	0.560	0.102	0.191	0.448	0.017	0.231	0.481
0.99	1	0.157	0.020	0.210	0.047	0.021	0.154	0.025	0.022	0.152
	2	0.303	0.066	0.398	0.088	0.079	0.294	0.050	0.082	0.291
	3	0.441	0.130	0.570	0.132	0.166	0.428	0.075	0.176	0.426

Table 20: Bias, Variance, and RMSE of the DR Estimators with Non-normal Errors

This table reports the bias, variance, and RMSE of the direct regression (DR) estimators when the errors in the predictive regression are non-normal and $T = 100, 300$: the OLS estimator $\hat{\beta}_h$, the bias-corrected estimator $\hat{\beta}_h^{BIR}$ proposed by Boudoukh et al. (2022), and iBoot estimator $\hat{\beta}_h^{iBootDR}$ proposed in this paper. The simulation design follows Section 7.14.

ρ	h	$\hat{\beta}_h$			$\hat{\beta}_h^{BIR}$			$\hat{\beta}_h^{iBootDR}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 100					
0.70	1	0.058	0.024	0.164	0.004	0.025	0.157	0.002	0.025	0.158
	3	0.154	0.150	0.417	0.011	0.166	0.408	0.008	0.177	0.421
	5	0.235	0.313	0.607	0.016	0.353	0.594	0.011	0.393	0.627
0.85	1	0.070	0.016	0.145	0.009	0.017	0.130	0.003	0.018	0.133
	3	0.194	0.114	0.389	0.024	0.129	0.360	0.009	0.142	0.377
	5	0.301	0.253	0.586	0.036	0.300	0.549	0.015	0.354	0.595
0.90	1	0.076	0.013	0.138	0.012	0.014	0.119	0.001	0.016	0.126
	3	0.213	0.097	0.377	0.034	0.110	0.334	0.002	0.131	0.362
	5	0.335	0.219	0.575	0.053	0.265	0.517	0.002	0.338	0.582
0.95	1	0.085	0.011	0.133	0.019	0.011	0.107	-0.000	0.013	0.115
	3	0.243	0.078	0.370	0.054	0.090	0.304	-0.000	0.111	0.333
	5	0.386	0.179	0.572	0.088	0.221	0.478	-0.001	0.287	0.536
0.99	1	0.097	0.008	0.132	0.030	0.008	0.096	0.015	0.009	0.096
	3	0.281	0.060	0.372	0.086	0.069	0.277	0.046	0.075	0.277
	5	0.449	0.138	0.583	0.141	0.174	0.440	0.076	0.191	0.444
					T = 300					
0.70	1	0.020	0.007	0.087	0.001	0.007	0.085	0.001	0.007	0.086
	8	0.113	0.200	0.462	0.005	0.209	0.457	0.005	0.223	0.472
	16	0.211	0.458	0.709	0.013	0.475	0.689	0.016	0.558	0.747
0.85	1	0.023	0.004	0.069	0.002	0.004	0.065	0.001	0.004	0.066
	8	0.146	0.162	0.428	0.010	0.173	0.417	0.007	0.187	0.432
	16	0.261	0.434	0.709	0.019	0.468	0.685	0.016	0.536	0.732
0.90	1	0.024	0.003	0.061	0.002	0.003	0.056	0.001	0.003	0.056
	8	0.162	0.133	0.399	0.014	0.145	0.381	0.008	0.156	0.395
	16	0.291	0.385	0.685	0.025	0.429	0.655	0.017	0.493	0.702
0.95	1	0.026	0.002	0.051	0.003	0.002	0.045	0.001	0.002	0.046
	8	0.186	0.094	0.359	0.023	0.105	0.325	0.006	0.116	0.341
	16	0.341	0.293	0.640	0.044	0.344	0.588	0.014	0.421	0.649
0.99	1	0.031	0.001	0.045	0.007	0.001	0.034	0.001	0.001	0.035
	8	0.235	0.055	0.332	0.060	0.062	0.257	0.011	0.072	0.269
	16	0.445	0.174	0.610	0.119	0.218	0.482	0.023	0.268	0.518

Table 21: Bias, Variance, and RMSE of the DR Estimators with Non-normal Errors

This table reports the bias, variance, and RMSE of the direct regression (DR) estimators when the errors in the predictive regression are non-normal and $T = 600$: the OLS estimator $\hat{\beta}_h$, the bias-corrected estimator $\hat{\beta}_h^{BIR}$ proposed by Boudoukh et al. (2022), and iBoot estimator $\hat{\beta}_h^{iBootDR}$ proposed in this paper. The simulation design follows Section 7.14.

ρ	h	$\hat{\beta}_h$			$\hat{\beta}_h^{BIR}$			$\hat{\beta}_h^{iBootDR}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 600					
0.70	1	0.009	0.004	0.060	0.000	0.004	0.060	0.000	0.004	0.060
	12	0.078	0.175	0.425	0.002	0.178	0.422	0.002	0.189	0.434
	24	0.142	0.384	0.636	0.001	0.391	0.625	0.003	0.450	0.671
	36	0.208	0.589	0.795	-0.001	0.598	0.774	0.004	0.766	0.875
	48	0.275	0.782	0.926	-0.004	0.794	0.891	0.005	1.146	1.070
0.85	1	0.011	0.002	0.046	0.000	0.002	0.045	0.000	0.002	0.045
	12	0.099	0.152	0.402	0.003	0.158	0.397	0.003	0.167	0.409
	24	0.173	0.382	0.642	0.003	0.397	0.630	0.004	0.442	0.665
	36	0.244	0.618	0.823	0.001	0.641	0.800	0.005	0.772	0.878
	48	0.318	0.839	0.970	-0.001	0.868	0.932	0.007	1.147	1.071
0.90	1	0.012	0.001	0.039	0.001	0.001	0.038	0.000	0.001	0.038
	12	0.111	0.126	0.372	0.005	0.133	0.365	0.003	0.140	0.375
	24	0.193	0.346	0.619	0.006	0.367	0.606	0.004	0.406	0.637
	36	0.269	0.583	0.810	0.003	0.617	0.786	0.004	0.727	0.853
	48	0.345	0.808	0.963	0.001	0.853	0.924	0.008	1.089	1.044
0.95	1	0.012	0.001	0.031	0.001	0.001	0.029	0.000	0.001	0.029
	12	0.128	0.085	0.319	0.009	0.092	0.303	0.003	0.096	0.310
	24	0.230	0.261	0.560	0.014	0.289	0.538	0.006	0.324	0.569
	36	0.319	0.469	0.756	0.016	0.525	0.725	0.006	0.616	0.785
	48	0.404	0.678	0.917	0.015	0.761	0.873	0.008	0.951	0.975
0.97	1	0.013	0.001	0.027	0.001	0.001	0.024	0.000	0.001	0.024
	12	0.140	0.064	0.289	0.014	0.069	0.264	0.004	0.074	0.272
	24	0.257	0.204	0.520	0.025	0.232	0.483	0.006	0.268	0.517
	36	0.361	0.379	0.714	0.033	0.443	0.667	0.009	0.543	0.737
	48	0.456	0.562	0.878	0.037	0.668	0.818	0.009	0.864	0.930
0.99	1	0.015	0.000	0.023	0.003	0.000	0.019	-0.000	0.000	0.020
	12	0.165	0.039	0.258	0.032	0.043	0.211	-0.004	0.052	0.228
	24	0.313	0.131	0.479	0.063	0.155	0.399	-0.009	0.196	0.443
	36	0.448	0.250	0.672	0.092	0.313	0.567	-0.012	0.421	0.649
	48	0.573	0.380	0.841	0.116	0.497	0.715	-0.016	0.699	0.836

Table 22: Bias, Variance, and RMSE of the IM Estimators with Non-normal Errors

This table reports the bias, variance, and RMSE of the implied (IM) estimators when the errors in the predictive regression are non-normal and $T = 30, 60$: the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator $\hat{\beta}_h^{iBootIM}$ proposed in this paper. The simulation design follows Section 7.14.

ρ	h	$\hat{\beta}_h^{IM}$			$\hat{\beta}_h^{FO\ IM}$			$\hat{\beta}_h^{iBootIM}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 30					
0.70	1	0.204	0.104	0.382	0.033	0.122	0.351	0.001	0.140	0.375
	2	0.279	0.211	0.538	0.057	0.311	0.560	-0.008	0.389	0.624
0.85	1	0.251	0.088	0.388	0.057	0.103	0.327	0.004	0.123	0.351
	2	0.391	0.199	0.593	0.107	0.300	0.558	-0.006	0.382	0.618
0.90	1	0.272	0.081	0.394	0.070	0.096	0.318	0.019	0.111	0.334
	2	0.439	0.190	0.619	0.136	0.288	0.554	0.024	0.347	0.590
0.95	1	0.294	0.074	0.401	0.085	0.088	0.309	0.042	0.097	0.314
	2	0.491	0.178	0.648	0.168	0.273	0.549	0.071	0.303	0.555
0.99	1	0.301	0.069	0.399	0.086	0.082	0.299	0.057	0.084	0.295
	2	0.517	0.170	0.661	0.173	0.264	0.542	0.102	0.265	0.525
					T = 60					
0.70	1	0.095	0.044	0.229	0.006	0.047	0.217	0.001	0.048	0.220
	2	0.138	0.106	0.353	0.010	0.130	0.360	0.001	0.134	0.366
	3	0.155	0.159	0.427	0.012	0.209	0.457	-0.001	0.217	0.466
0.85	1	0.116	0.032	0.213	0.014	0.035	0.187	-0.004	0.040	0.200
	2	0.194	0.089	0.356	0.027	0.110	0.333	-0.010	0.132	0.364
	3	0.245	0.148	0.456	0.038	0.202	0.451	-0.019	0.252	0.502
0.90	1	0.127	0.028	0.210	0.021	0.030	0.175	-0.005	0.036	0.191
	2	0.220	0.081	0.359	0.041	0.101	0.320	-0.014	0.125	0.354
	3	0.289	0.137	0.470	0.058	0.190	0.440	-0.027	0.247	0.497
0.95	1	0.143	0.024	0.210	0.033	0.026	0.164	0.005	0.030	0.172
	2	0.257	0.071	0.370	0.066	0.088	0.305	0.007	0.104	0.322
	3	0.348	0.123	0.494	0.096	0.173	0.426	0.005	0.207	0.455
0.99	1	0.157	0.020	0.210	0.044	0.022	0.153	0.025	0.022	0.152
	2	0.289	0.060	0.379	0.088	0.076	0.290	0.048	0.079	0.285
	3	0.401	0.107	0.518	0.131	0.153	0.412	0.067	0.158	0.403

Table 23: Bias, Variance, and RMSE of the IM Estimators with Non-normal Errors

This table reports the bias, variance, and RMSE of the implied (IM) estimators when the errors in the predictive regression are non-normal and $T = 100, 300$: the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator $\hat{\beta}_h^{iBootIM}$ proposed in this paper. The simulation design follows Section 7.14.

ρ	h	$\hat{\beta}_h^{IM}$			$\hat{\beta}_h^{FO\ IM}$			$\hat{\beta}_h^{iBootIM}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 100					
0.70	1	0.058	0.024	0.164	0.003	0.025	0.157	0.002	0.025	0.158
	3	0.099	0.095	0.323	0.007	0.113	0.336	0.003	0.114	0.338
	5	0.104	0.142	0.391	0.007	0.178	0.421	0.002	0.181	0.425
0.85	1	0.070	0.016	0.145	0.007	0.017	0.130	0.003	0.018	0.133
	3	0.154	0.085	0.329	0.020	0.103	0.321	0.006	0.110	0.331
	5	0.195	0.152	0.436	0.027	0.202	0.451	0.006	0.221	0.470
0.90	1	0.076	0.013	0.138	0.011	0.014	0.119	0.001	0.016	0.126
	3	0.180	0.076	0.329	0.030	0.093	0.306	-0.001	0.110	0.331
	5	0.244	0.144	0.451	0.044	0.197	0.446	-0.007	0.244	0.494
0.95	1	0.085	0.011	0.133	0.018	0.011	0.107	-0.000	0.013	0.115
	3	0.218	0.064	0.334	0.051	0.079	0.286	-0.005	0.099	0.314
	5	0.314	0.128	0.476	0.080	0.180	0.431	-0.016	0.234	0.484
0.99	1	0.097	0.008	0.132	0.028	0.008	0.096	0.015	0.009	0.096
	3	0.263	0.051	0.347	0.084	0.064	0.266	0.042	0.068	0.265
	5	0.399	0.105	0.514	0.138	0.151	0.413	0.063	0.164	0.410
					T = 300					
0.70	1	0.020	0.007	0.087	0.001	0.007	0.085	0.001	0.007	0.086
	8	0.035	0.065	0.257	0.003	0.070	0.265	0.002	0.071	0.266
	16	0.031	0.073	0.272	0.002	0.079	0.281	0.004	0.079	0.281
0.85	1	0.023	0.004	0.069	0.002	0.004	0.065	0.001	0.004	0.066
	8	0.080	0.083	0.298	0.007	0.095	0.308	0.004	0.096	0.310
	16	0.081	0.128	0.367	0.007	0.152	0.390	0.006	0.154	0.392
0.90	1	0.024	0.003	0.061	0.002	0.003	0.056	0.001	0.003	0.056
	8	0.107	0.081	0.303	0.011	0.095	0.308	0.005	0.096	0.310
	16	0.125	0.149	0.406	0.013	0.187	0.433	0.008	0.191	0.437
0.95	1	0.026	0.002	0.051	0.003	0.002	0.045	0.001	0.002	0.046
	8	0.147	0.068	0.299	0.020	0.081	0.286	0.005	0.087	0.296
	16	0.210	0.155	0.446	0.032	0.209	0.458	0.006	0.232	0.481
0.99	1	0.031	0.001	0.045	0.007	0.001	0.034	0.001	0.001	0.035
	8	0.213	0.045	0.301	0.058	0.055	0.242	0.007	0.064	0.253
	16	0.366	0.120	0.504	0.111	0.175	0.432	0.006	0.213	0.461

Table 24: Bias, Variance, and RMSE of the IM Estimators with Non-normal Errors

This table reports the bias, variance, and RMSE of the implied (IM) estimators when the errors in the predictive regression are non-normal and $T = 600$: the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator $\hat{\beta}_h^{iBootIM}$ proposed in this paper. The simulation design follows Section 7.14.

ρ	h	$\hat{\beta}_h^{IM}$			$\hat{\beta}_h^{FO\ IM}$			$\hat{\beta}_h^{iBootIM}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 600					
0.70	1	0.009	0.004	0.060	0.000	0.004	0.060	0.000	0.004	0.060
	12	0.015	0.036	0.192	0.000	0.038	0.195	0.000	0.038	0.195
	24	0.014	0.038	0.195	0.000	0.039	0.198	0.000	0.039	0.198
	36	0.014	0.038	0.195	0.000	0.039	0.198	0.000	0.039	0.198
	48	0.014	0.038	0.195	0.000	0.039	0.198	0.000	0.039	0.198
0.85	1	0.011	0.002	0.046	0.000	0.002	0.045	0.000	0.002	0.045
	12	0.041	0.059	0.246	0.002	0.064	0.253	0.001	0.064	0.253
	24	0.037	0.077	0.279	0.001	0.083	0.289	0.002	0.083	0.289
	36	0.035	0.080	0.285	0.001	0.087	0.294	0.003	0.086	0.294
	48	0.035	0.081	0.286	0.001	0.087	0.295	0.004	0.086	0.294
0.90	1	0.012	0.001	0.039	0.001	0.001	0.038	0.000	0.001	0.038
	12	0.060	0.063	0.259	0.003	0.071	0.266	0.001	0.071	0.266
	24	0.062	0.101	0.324	0.003	0.115	0.340	0.002	0.116	0.340
	36	0.058	0.115	0.343	0.002	0.131	0.361	0.004	0.130	0.361
	48	0.056	0.119	0.350	0.002	0.135	0.368	0.006	0.134	0.366
0.95	1	0.012	0.001	0.031	0.001	0.001	0.029	0.000	0.001	0.029
	12	0.093	0.057	0.256	0.008	0.065	0.255	0.002	0.066	0.257
	24	0.121	0.122	0.370	0.010	0.149	0.387	0.003	0.152	0.390
	36	0.127	0.165	0.425	0.010	0.209	0.457	0.005	0.213	0.462
	48	0.126	0.191	0.455	0.009	0.246	0.496	0.008	0.250	0.500
0.97	1	0.013	0.001	0.027	0.001	0.001	0.024	0.000	0.001	0.024
	12	0.114	0.048	0.247	0.013	0.055	0.236	0.003	0.058	0.240
	24	0.169	0.118	0.382	0.021	0.150	0.388	0.005	0.159	0.399
	36	0.194	0.176	0.462	0.024	0.239	0.489	0.007	0.257	0.507
	48	0.204	0.221	0.512	0.024	0.312	0.559	0.010	0.340	0.583
0.99	1	0.015	0.000	0.023	0.003	0.000	0.019	-0.000	0.000	0.020
	12	0.149	0.033	0.235	0.031	0.039	0.200	-0.005	0.047	0.216
	24	0.255	0.094	0.398	0.058	0.125	0.358	-0.015	0.158	0.397
	36	0.331	0.156	0.515	0.080	0.230	0.486	-0.028	0.306	0.554
	48	0.386	0.213	0.602	0.097	0.341	0.592	-0.045	0.480	0.694

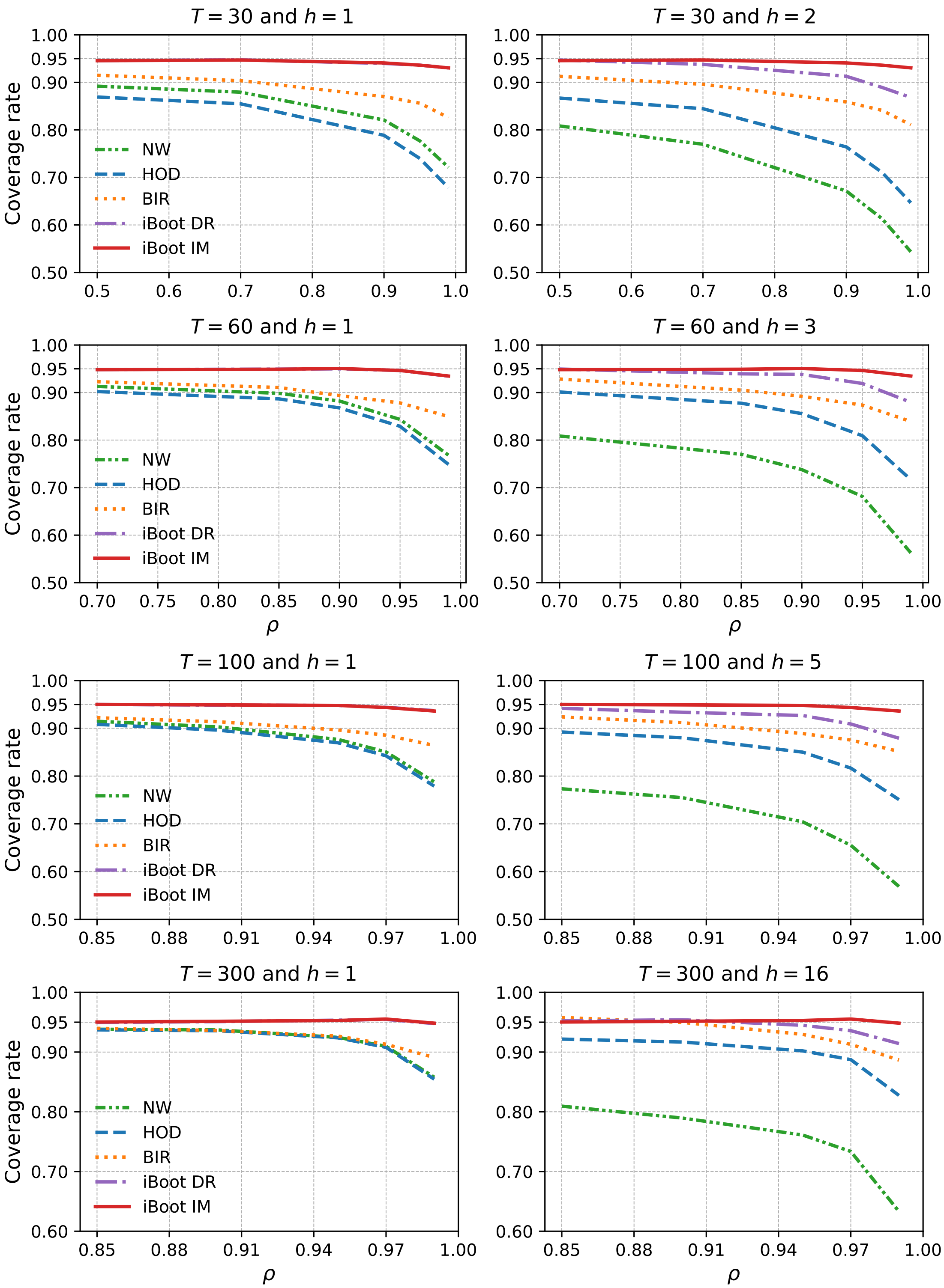


Figure 29: Empirical coverage rates of 95% Confidence Intervals (CIs) for various methods when the errors are non-normal. NW means the Newey-West CIs, HOD the Hodrick CIs, BIR the CIs using the BIR estimator and the standard errors proposed in Boudoukh et al. (2022), iBoot DR and iBoot IM the CIs proposed in this paper, for the DR and IM estimators respectively.

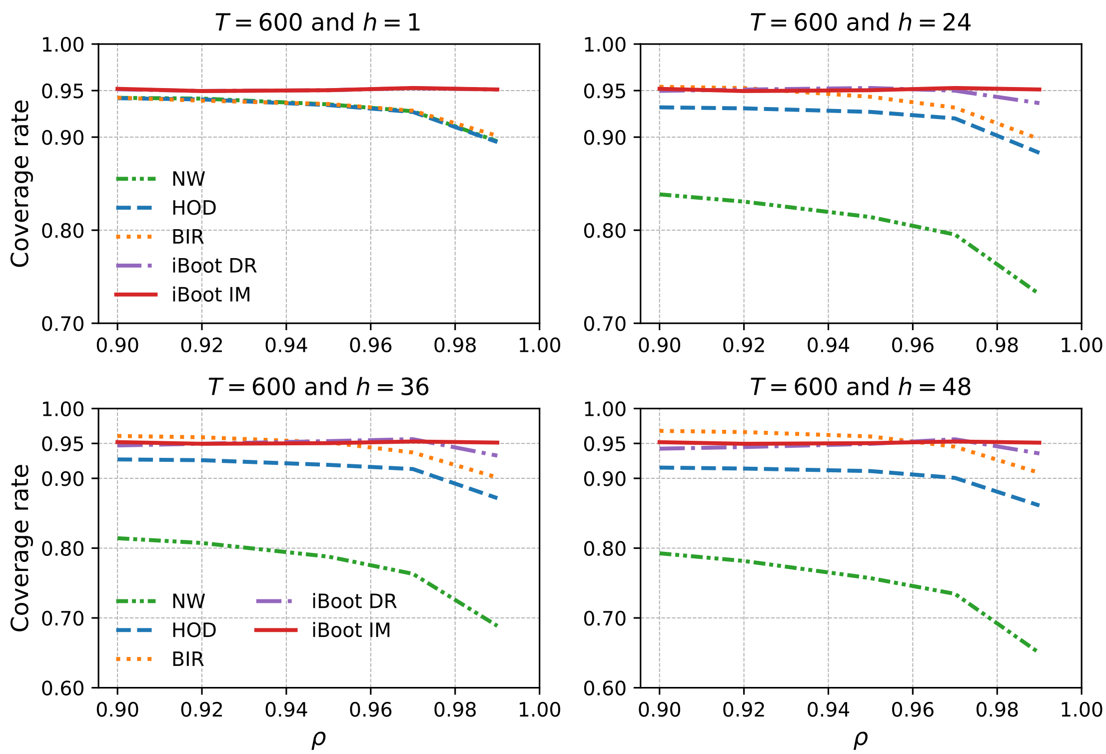


Figure 30: Empirical coverage rates of 95% Confidence Intervals (CIs) for various methods when the errors are non-normal. NW means the Newey-West CIs, HOD the Hodrick CIs, BIR the CIs using the BIR estimator and the standard errors proposed in Boudoukh et al. (2022), iBoot DR and iBoot IM the CIs proposed in this paper, for the DR and IM estimators respectively.

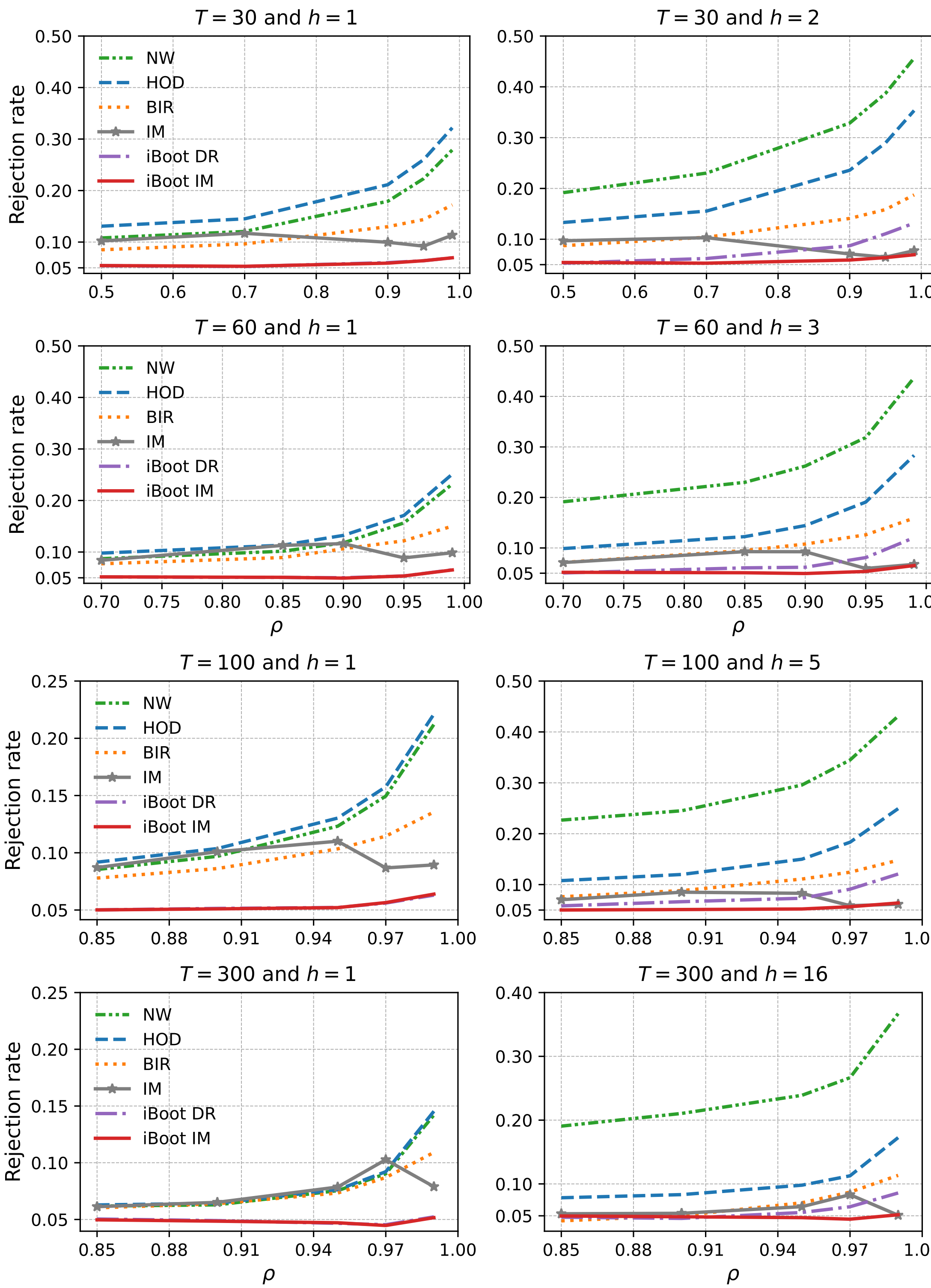


Figure 31: Size curves under $H_0 : \beta_1 = 0$ for various tests with non-normal errors: Newey-West (NW), Hodrick (HOD), Boudoukh et al. (2022) (BIR), Xu (2020) (IM), and the proposed iBoot tests for Direct Regression (iBoot DR) and Implication-Based (iBoot IM) estimators.

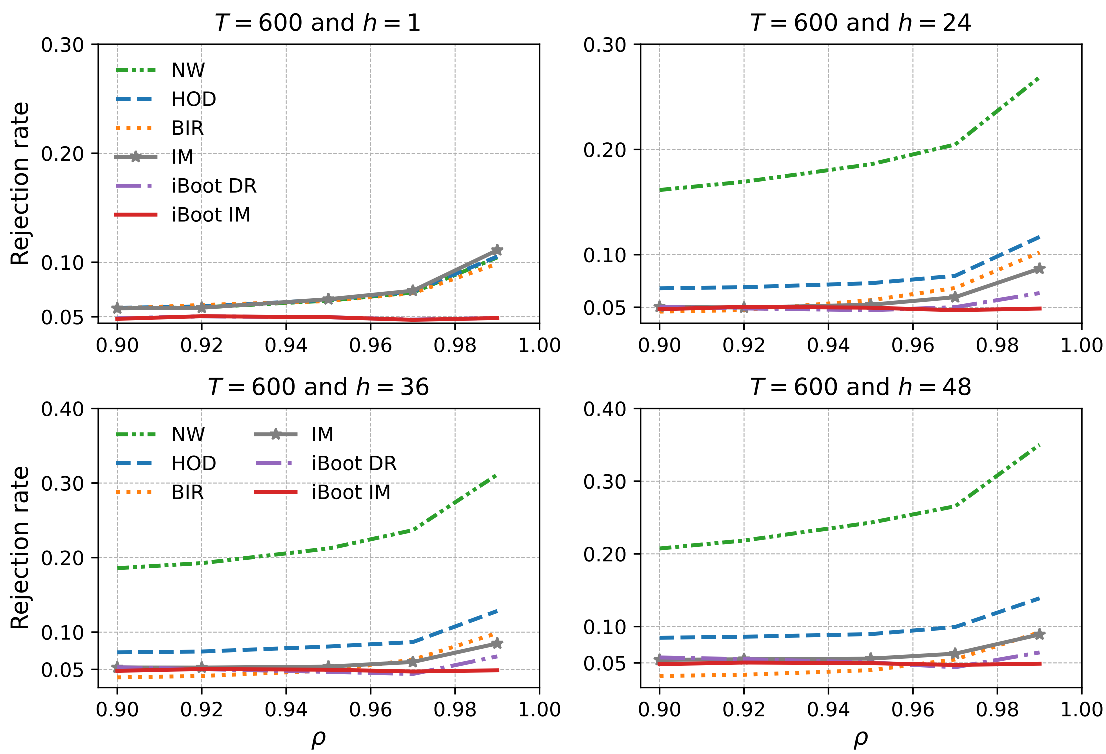


Figure 32: Size curves under $H_0 : \beta_1 = 0$ for various tests across horizons with non-normal errors and $T = 600$: Newey-West (NW), Hodrick (HOD), Boudoukh et al. (2022) (BIR), Xu (2020) (IM), and the proposed iBoot tests for Direct Regression (iBoot DR) and Implication-Based (iBoot IM) estimators.

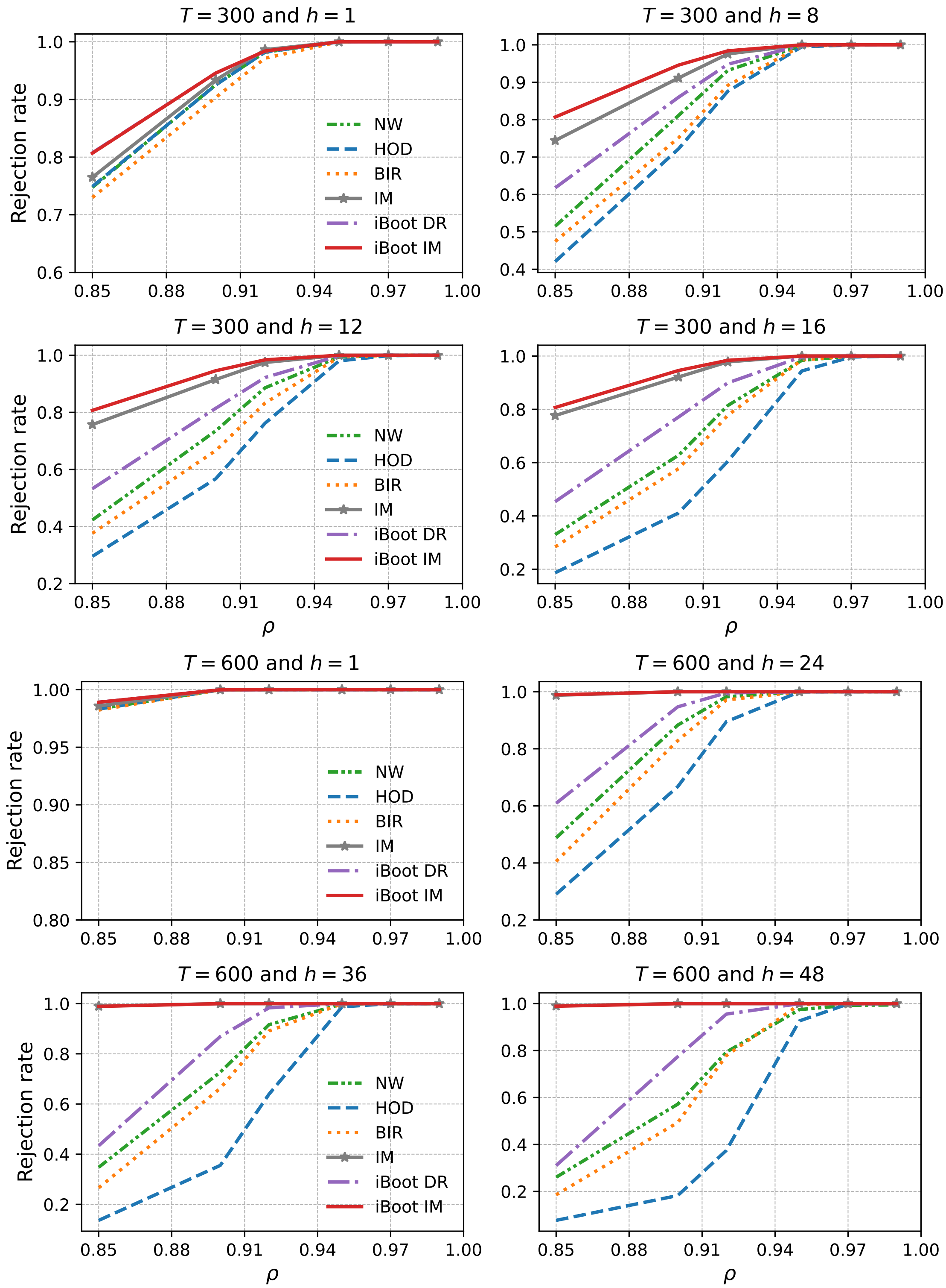


Figure 33: Size-adjusted power curves for testing $H_0 : \beta_1 = 0$ when the true $\beta_1 = 0.16$, across horizons with non normal errors. The figure compares tests from Newey-West (NW), Hodrick (HOD), Boudoukh et al. (2022) (BIR), Xu (2020) (IM), and our proposed iBoot DR and iBoot IM, for $T = 300$ and $T = 600$.

7.15 Simulation Results When Errors are GARCH(1,1)

This subsection tests the robustness of our findings when the regression errors (u_t and v_t) in Eqs. (2.1) and (2.2) follow GARCH(1,1) dynamics. In this setting, the errors are serially dependent and non-normally distributed. We generate them using a two-step procedure.

First, we simulate two independent GARCH(1,1) processes:

$$e_{u,t} = \sigma_{u,t}\varepsilon_{u,t}, \quad \sigma_{u,t}^2 = \omega_u + \alpha_u e_{u,t-1}^2 + \beta_u \sigma_{u,t-1}^2,$$

$$e_{v,t} = \sigma_{v,t}\varepsilon_{v,t}, \quad \sigma_{v,t}^2 = \omega_v + \alpha_v e_{v,t-1}^2 + \beta_v \sigma_{v,t-1}^2,$$

where $\varepsilon_{u,t}$ and $\varepsilon_{v,t}$ are i.i.d. standard normal innovations. The parameters are set to $\alpha_u = \alpha_v = 0.05$, $\beta_u = \beta_v = 0.8$, and $\omega_u = \sigma_u^2(1 - \alpha_u - \beta_u)$, $\omega_v = \sigma_v^2(1 - \alpha_v - \beta_v)$.

Second, we introduce contemporaneous correlation by premultiplying $[e_{u,t}, e_{v,t}]'$ with the Cholesky (square-root) matrix of the target covariance matrix Σ, producing $[u_t, v_t]'$. This procedure ensures that u_t and v_t share the same mean and covariance structure as in the baseline simulations while exhibiting GARCH(1,1) volatility. All other design parameters remain identical to Section 4.

Figures 34-35 display the biases of alternative DR estimators, and Figure 36 presents the corresponding IM results. The detailed bias values appear in Tables 25-27 and Tables 28-30. Coverage rates, test sizes, and power results are shown in Figures 37-41.

Across all metrics – bias, coverage, size, and power – the results mirror those from the baseline case with homoskedastic normal errors. The iBoot estimators continue to deliver minimal bias and accurate inference, demonstrating that our framework remains robust to GARCH(1,1) volatility and non-normality.

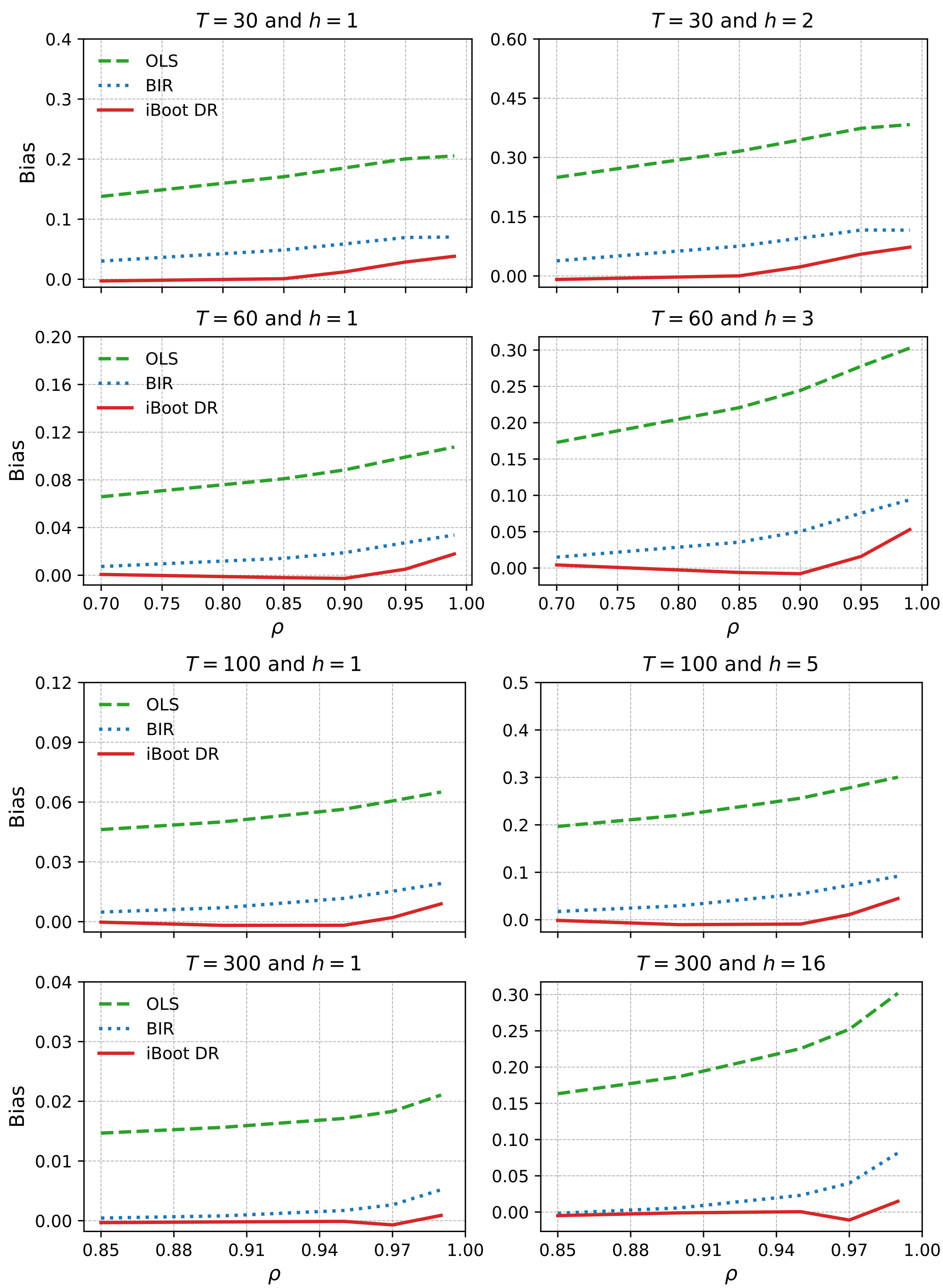


Figure 34: Bias plots for various DR estimators with GARCH(1,1) errors. OLS means the OLS estimator, BIR the BIR estimator proposed in Boudoukh et al. (2022), and iBoot DR the iBoot estimator proposed in this paper.

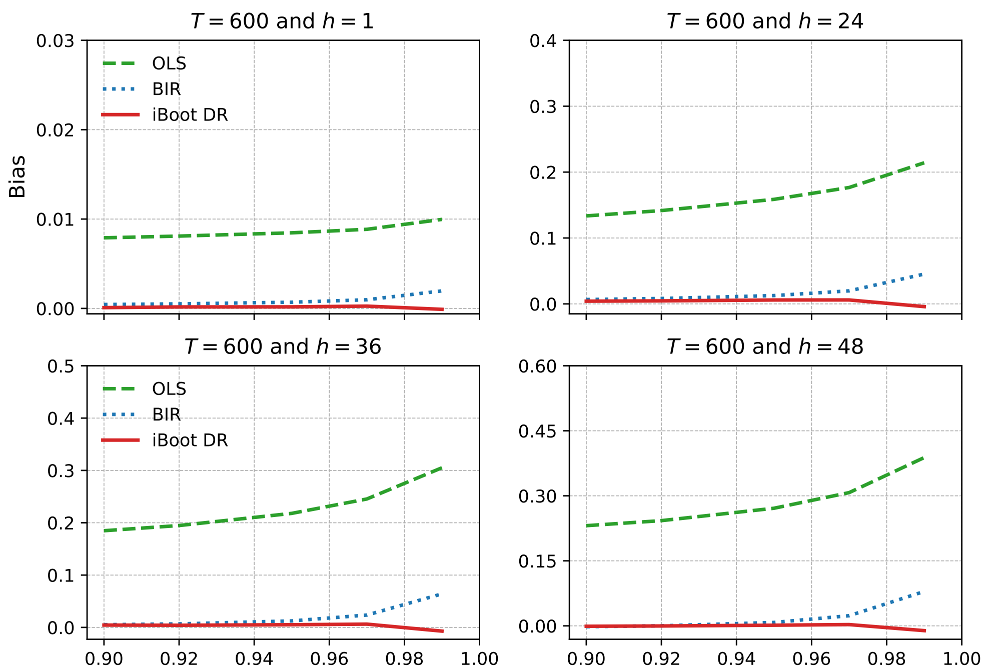


Figure 35: Bias plots for various DR estimators with GARCH(1,1) errors and $T = 600$. OLS means the OLS estimator, BIR the BIR estimator proposed in Boudoukh et al. (2022), and iBoot DR the iBoot estimator proposed in this paper.

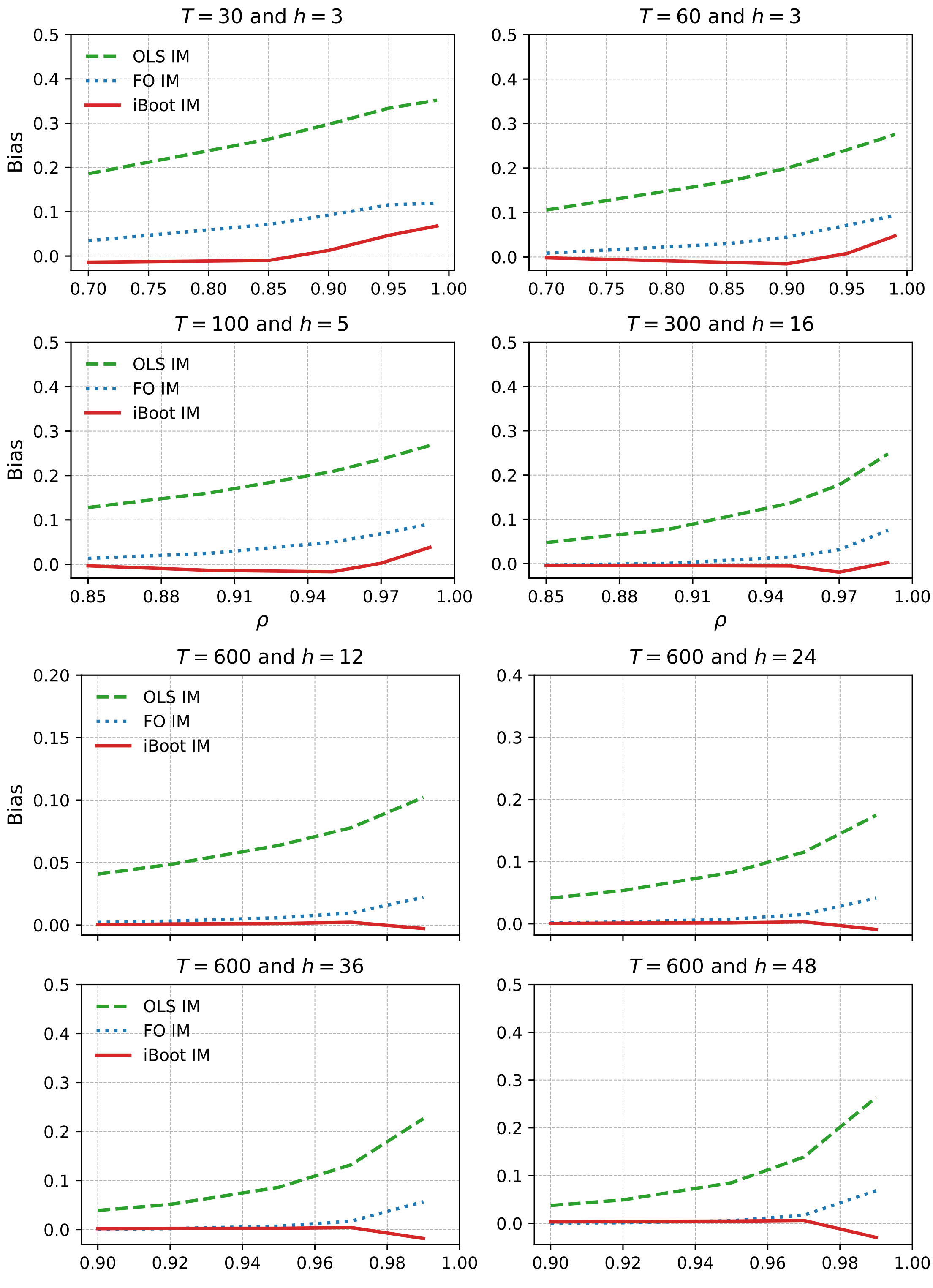


Figure 36: Bias of various Implication-Based (IM) estimators with GARCH(1,1) errors at long horizons. Compares the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator proposed in this paper.

Table 25: Bias, Variance, and RMSE of the DR Estimators with GARCH(1,1) Errors

This table reports the bias, variance, and RMSE of the direct regression (DR) estimators when the errors in the predictive regression are GARCH(1,1) and $T = 30, 60$: the OLS estimator $\hat{\beta}_h$, the bias-corrected estimator $\hat{\beta}_h^{BIR}$ proposed by Boudoukh et al. (2022), and iBoot estimator $\hat{\beta}_h^{iBootDR}$ proposed in this paper. The simulation design follows Section 7.15.

ρ	h	$\hat{\beta}_h$			$\hat{\beta}_h^{BIR}$			$\hat{\beta}_h^{iBootDR}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
T = 30										
0.70	1	0.138	0.046	0.255	0.030	0.054	0.235	-0.003	0.064	0.252
	2	0.250	0.127	0.435	0.038	0.166	0.409	-0.009	0.218	0.467
0.85	1	0.171	0.039	0.261	0.049	0.046	0.220	0.001	0.056	0.237
	2	0.316	0.111	0.459	0.076	0.149	0.394	0.000	0.196	0.443
0.90	1	0.185	0.036	0.266	0.059	0.043	0.215	0.012	0.050	0.224
	2	0.345	0.104	0.472	0.096	0.141	0.387	0.023	0.176	0.420
0.95	1	0.200	0.033	0.270	0.070	0.039	0.209	0.029	0.043	0.210
	2	0.374	0.095	0.485	0.116	0.130	0.379	0.055	0.151	0.392
0.99	1	0.205	0.030	0.269	0.070	0.036	0.202	0.038	0.037	0.197
	2	0.383	0.089	0.485	0.116	0.123	0.369	0.073	0.129	0.367
T = 60										
0.70	1	0.066	0.019	0.154	0.007	0.021	0.146	0.001	0.022	0.148
	2	0.122	0.059	0.272	0.010	0.068	0.261	0.002	0.074	0.271
	3	0.173	0.107	0.371	0.015	0.128	0.358	0.004	0.146	0.382
0.85	1	0.081	0.014	0.144	0.014	0.015	0.125	-0.002	0.018	0.134
	2	0.154	0.046	0.263	0.024	0.054	0.233	-0.004	0.065	0.256
	3	0.221	0.086	0.367	0.036	0.106	0.328	-0.006	0.139	0.372
0.90	1	0.088	0.012	0.142	0.019	0.013	0.117	-0.003	0.016	0.127
	2	0.169	0.040	0.262	0.033	0.047	0.220	-0.006	0.059	0.244
	3	0.244	0.076	0.369	0.050	0.095	0.313	-0.008	0.127	0.356
0.95	1	0.099	0.010	0.142	0.027	0.011	0.109	0.005	0.013	0.114
	2	0.191	0.034	0.266	0.050	0.040	0.207	0.010	0.048	0.219
	3	0.278	0.065	0.377	0.076	0.083	0.297	0.016	0.101	0.319
0.99	1	0.108	0.008	0.142	0.034	0.009	0.102	0.018	0.010	0.101
	2	0.208	0.029	0.269	0.063	0.034	0.195	0.035	0.036	0.194
	3	0.303	0.056	0.384	0.094	0.071	0.283	0.053	0.077	0.282

Table 26: Bias, Variance, and RMSE of the DR Estimators with GARCH(1,1) Errors

This table reports the bias, variance, and RMSE of the direct regression (DR) estimators when the errors in the predictive regression are GARCH(1,1) and $T = 100, 300$: the OLS estimator $\hat{\beta}_h$, the bias-corrected estimator $\hat{\beta}_h^{BIR}$ proposed by Boudoukh et al. (2022), and iBoot estimator $\hat{\beta}_h^{iBootDR}$ proposed in this paper. The simulation design follows Section 7.15.

ρ	h	$\hat{\beta}_h$			$\hat{\beta}_h^{BIR}$			$\hat{\beta}_h^{iBootDR}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 100					
0.70	1	0.038	0.011	0.110	0.002	0.011	0.106	-0.000	0.011	0.106
	3	0.099	0.063	0.271	0.003	0.071	0.266	-0.002	0.076	0.275
	5	0.151	0.131	0.392	0.002	0.148	0.385	-0.004	0.168	0.409
0.85	1	0.046	0.007	0.095	0.005	0.007	0.086	-0.000	0.008	0.088
	3	0.127	0.047	0.251	0.012	0.054	0.232	-0.001	0.060	0.244
	5	0.197	0.104	0.377	0.017	0.123	0.352	-0.001	0.149	0.386
0.90	1	0.050	0.006	0.090	0.007	0.006	0.077	-0.002	0.007	0.083
	3	0.140	0.040	0.243	0.019	0.046	0.214	-0.006	0.055	0.235
	5	0.220	0.089	0.371	0.029	0.109	0.331	-0.010	0.142	0.377
0.95	1	0.056	0.004	0.087	0.012	0.005	0.069	-0.002	0.005	0.074
	3	0.161	0.032	0.240	0.033	0.037	0.195	-0.006	0.046	0.214
	5	0.256	0.073	0.372	0.054	0.091	0.306	-0.009	0.119	0.345
0.99	1	0.065	0.003	0.086	0.019	0.003	0.062	0.009	0.004	0.061
	3	0.188	0.024	0.244	0.055	0.028	0.177	0.027	0.030	0.176
	5	0.301	0.056	0.383	0.092	0.071	0.283	0.045	0.079	0.284
					T = 300					
0.70	1	0.012	0.003	0.059	0.000	0.003	0.058	-0.001	0.003	0.059
	8	0.071	0.090	0.309	-0.003	0.094	0.307	-0.004	0.101	0.318
	16	0.126	0.206	0.471	-0.009	0.214	0.463	-0.009	0.256	0.506
0.85	1	0.015	0.002	0.047	0.000	0.002	0.045	-0.000	0.002	0.045
	8	0.093	0.072	0.284	0.001	0.078	0.279	-0.002	0.084	0.290
	16	0.163	0.190	0.465	-0.002	0.206	0.454	-0.005	0.239	0.489
0.90	1	0.016	0.001	0.041	0.001	0.001	0.038	-0.000	0.001	0.038
	8	0.105	0.059	0.264	0.005	0.065	0.254	-0.001	0.069	0.264
	16	0.187	0.166	0.448	0.006	0.187	0.432	-0.001	0.217	0.466
0.95	1	0.017	0.001	0.034	0.002	0.001	0.030	-0.000	0.001	0.031
	8	0.124	0.041	0.237	0.013	0.046	0.215	-0.000	0.051	0.227
	16	0.226	0.124	0.419	0.023	0.147	0.385	0.001	0.183	0.428
0.99	1	0.021	0.000	0.030	0.005	0.000	0.023	0.001	0.001	0.023
	8	0.160	0.024	0.222	0.042	0.027	0.170	0.007	0.031	0.178
	16	0.302	0.075	0.407	0.082	0.094	0.318	0.015	0.117	0.342

Table 27: Bias, Variance, and RMSE of the DR Estimators with GARCH(1,1) Errors

This table reports the bias, variance, and RMSE of the direct regression (DR) estimators when the errors in the predictive regression are GARCH(1,1) and $T = 600$: the OLS estimator $\hat{\beta}_h$, the bias-corrected estimator $\hat{\beta}_h^{BIR}$ proposed by Boudoukh et al. (2022), and iBoot estimator $\hat{\beta}_h^{iBootDR}$ proposed in this paper. The simulation design follows Section 7.15.

ρ	h	$\hat{\beta}_h$			$\hat{\beta}_h^{BIR}$			$\hat{\beta}_h^{iBoot}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 600					
0.70	1	0.007	0.002	0.041	0.000	0.002	0.040	0.000	0.002	0.040
	12	0.052	0.077	0.283	0.001	0.079	0.280	0.001	0.084	0.289
	24	0.097	0.173	0.427	0.002	0.176	0.419	0.002	0.205	0.452
	36	0.143	0.262	0.531	0.002	0.266	0.516	0.004	0.346	0.588
	48	0.184	0.343	0.614	-0.004	0.348	0.590	-0.001	0.513	0.716
0.85	1	0.007	0.001	0.031	0.000	0.001	0.030	0.000	0.001	0.030
	12	0.068	0.067	0.268	0.003	0.070	0.265	0.002	0.074	0.273
	24	0.119	0.171	0.431	0.004	0.178	0.422	0.004	0.200	0.447
	36	0.169	0.272	0.548	0.004	0.282	0.531	0.005	0.344	0.587
	48	0.213	0.364	0.640	-0.004	0.377	0.614	-0.002	0.508	0.713
0.90	1	0.008	0.001	0.026	0.000	0.001	0.025	0.000	0.001	0.025
	12	0.076	0.056	0.249	0.005	0.059	0.244	0.003	0.063	0.251
	24	0.134	0.155	0.415	0.007	0.164	0.405	0.004	0.183	0.427
	36	0.185	0.255	0.537	0.005	0.270	0.519	0.004	0.322	0.567
	48	0.231	0.348	0.634	-0.002	0.368	0.607	-0.001	0.480	0.693
0.95	1	0.008	0.000	0.021	0.001	0.000	0.019	0.000	0.000	0.019
	12	0.089	0.038	0.215	0.008	0.041	0.204	0.003	0.044	0.209
	24	0.159	0.116	0.376	0.013	0.129	0.359	0.006	0.145	0.381
	36	0.218	0.203	0.500	0.012	0.228	0.477	0.005	0.270	0.520
	48	0.271	0.289	0.602	0.008	0.326	0.571	0.001	0.416	0.645
0.97	1	0.009	0.000	0.018	0.001	0.000	0.016	0.000	0.000	0.016
	12	0.097	0.029	0.195	0.011	0.031	0.177	0.003	0.034	0.184
	24	0.177	0.090	0.349	0.020	0.103	0.322	0.006	0.120	0.347
	36	0.245	0.163	0.472	0.024	0.192	0.438	0.006	0.238	0.488
	48	0.307	0.239	0.577	0.023	0.286	0.535	0.003	0.378	0.614
0.99	1	0.010	0.000	0.016	0.002	0.000	0.012	-0.000	0.000	0.013
	12	0.113	0.018	0.175	0.023	0.020	0.142	-0.001	0.024	0.153
	24	0.214	0.058	0.322	0.045	0.069	0.266	-0.004	0.087	0.296
	36	0.305	0.107	0.448	0.064	0.135	0.374	-0.007	0.184	0.429
	48	0.388	0.161	0.559	0.079	0.213	0.468	-0.011	0.305	0.552

Table 28: Bias, Variance, and RMSE of the IM Estimators with GARCH(1,1) Errors

This table reports the bias, variance, and RMSE of the implied (IM) estimators when the errors in the predictive regression are GARCH(1,1) and $T = 30, 60$: the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator $\hat{\beta}_h^{iBootIM}$ proposed in this paper. The simulation design follows Section 7.15.

ρ	h	$\hat{\beta}_h^{IM}$			$\hat{\beta}_h^{FO\ IM}$			$\hat{\beta}_h^{iBootIM}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 30					
0.70	1	0.138	0.046	0.255	0.022	0.055	0.236	-0.003	0.064	0.252
	2	0.186	0.091	0.354	0.035	0.138	0.373	-0.014	0.176	0.420
0.85	1	0.171	0.039	0.261	0.039	0.047	0.219	0.001	0.056	0.237
	2	0.264	0.085	0.393	0.071	0.132	0.370	-0.010	0.171	0.414
0.90	1	0.185	0.036	0.266	0.049	0.043	0.214	0.012	0.050	0.224
	2	0.298	0.081	0.412	0.092	0.127	0.368	0.013	0.154	0.393
0.95	1	0.200	0.033	0.270	0.059	0.039	0.207	0.029	0.043	0.210
	2	0.334	0.076	0.433	0.116	0.119	0.364	0.047	0.134	0.368
0.99	1	0.205	0.030	0.269	0.060	0.036	0.200	0.038	0.037	0.197
	2	0.352	0.072	0.442	0.119	0.114	0.358	0.068	0.115	0.346
					T = 60					
0.70	1	0.066	0.019	0.154	0.005	0.021	0.146	0.001	0.022	0.148
	2	0.094	0.046	0.235	0.008	0.058	0.240	-0.000	0.060	0.244
	3	0.106	0.069	0.282	0.009	0.092	0.303	-0.002	0.096	0.311
0.85	1	0.081	0.014	0.144	0.012	0.015	0.125	-0.002	0.018	0.134
	2	0.134	0.039	0.238	0.022	0.049	0.222	-0.006	0.059	0.243
	3	0.169	0.063	0.303	0.030	0.088	0.298	-0.012	0.112	0.335
0.90	1	0.088	0.012	0.142	0.016	0.013	0.117	-0.003	0.016	0.127
	2	0.153	0.035	0.241	0.032	0.044	0.212	-0.008	0.055	0.235
	3	0.200	0.058	0.314	0.045	0.082	0.290	-0.015	0.108	0.329
0.95	1	0.099	0.010	0.142	0.025	0.011	0.109	0.005	0.013	0.114
	2	0.178	0.030	0.248	0.049	0.038	0.201	0.008	0.045	0.212
	3	0.241	0.052	0.331	0.071	0.074	0.281	0.008	0.090	0.299
0.99	1	0.108	0.008	0.142	0.031	0.009	0.101	0.018	0.010	0.101
	2	0.198	0.026	0.255	0.063	0.033	0.191	0.034	0.034	0.188
	3	0.275	0.045	0.347	0.093	0.065	0.271	0.047	0.068	0.265

Table 29: Bias, Variance, and RMSE of the IM Estimators with GARCH(1,1) Errors

This table reports the bias, variance, and RMSE of the implied (IM) estimators when the errors in the predictive regression are GARCH(1,1) and $T = 100, 300$: the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator $\hat{\beta}_h^{iBootIM}$ proposed in this paper. The simulation design follows Section 7.15.

ρ	h	$\hat{\beta}_h^{IM}$			$\hat{\beta}_h^{FO\ IM}$			$\hat{\beta}_h^{iBootIM}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 100					
0.70	1	0.038	0.011	0.110	0.001	0.011	0.106	-0.000	0.011	0.106
	3	0.064	0.042	0.216	0.002	0.051	0.226	-0.002	0.052	0.228
	5	0.066	0.064	0.261	0.000	0.080	0.284	-0.004	0.082	0.287
0.85	1	0.046	0.007	0.095	0.004	0.007	0.086	-0.000	0.008	0.088
	3	0.101	0.037	0.216	0.010	0.045	0.212	-0.001	0.048	0.220
	5	0.128	0.065	0.286	0.013	0.088	0.298	-0.003	0.098	0.313
0.90	1	0.050	0.006	0.090	0.006	0.006	0.077	-0.002	0.007	0.083
	3	0.119	0.032	0.215	0.017	0.040	0.201	-0.007	0.048	0.219
	5	0.161	0.061	0.295	0.025	0.085	0.292	-0.014	0.108	0.328
0.95	1	0.056	0.004	0.087	0.011	0.005	0.069	-0.002	0.005	0.074
	3	0.145	0.027	0.219	0.032	0.033	0.186	-0.008	0.042	0.204
	5	0.209	0.054	0.312	0.050	0.076	0.281	-0.017	0.100	0.317
0.99	1	0.065	0.003	0.086	0.018	0.003	0.061	0.009	0.004	0.061
	3	0.176	0.021	0.228	0.055	0.026	0.172	0.025	0.028	0.169
	5	0.268	0.044	0.340	0.091	0.064	0.268	0.038	0.068	0.264
					T = 300					
0.70	1	0.012	0.003	0.059	-0.000	0.003	0.058	-0.000	0.003	0.059
	8	0.020	0.030	0.175	-0.003	0.033	0.182	-0.003	0.033	0.182
	16	0.016	0.034	0.186	-0.003	0.037	0.193	-0.002	0.037	0.193
0.85	1	0.015	0.002	0.047	0.000	0.002	0.045	-0.000	0.002	0.045
	8	0.049	0.038	0.202	-0.000	0.044	0.211	-0.003	0.045	0.213
	16	0.048	0.060	0.250	-0.003	0.072	0.268	-0.004	0.073	0.270
0.90	1	0.016	0.001	0.041	0.001	0.001	0.038	-0.000	0.001	0.038
	8	0.068	0.037	0.205	0.003	0.044	0.210	-0.002	0.045	0.212
	16	0.078	0.069	0.274	0.001	0.088	0.297	-0.004	0.090	0.301
0.95	1	0.017	0.001	0.034	0.002	0.001	0.030	-0.000	0.001	0.031
	8	0.097	0.031	0.200	0.011	0.037	0.193	-0.002	0.040	0.201
	16	0.137	0.070	0.298	0.015	0.096	0.310	-0.005	0.109	0.329
0.99	1	0.021	0.000	0.030	0.005	0.000	0.023	0.001	0.001	0.023
	8	0.145	0.020	0.203	0.040	0.025	0.163	0.004	0.029	0.170
	16	0.248	0.053	0.339	0.075	0.079	0.290	0.002	0.095	0.309

Table 30: Bias, Variance, and RMSE of the IM Estimators with GARCH(1,1) Errors

This table reports the bias, variance, and RMSE of the implied (IM) estimators when the errors in the predictive regression are GARCH(1,1) and $T = 600$: the OLS IM estimator defined in (2.10), the first-order bias-reduced estimator (FO IM) derived from the $1/T$ terms in Proposition 2.4 under $\beta_1 = 0$, and the iBoot IM estimator $\hat{\beta}_h^{iBootIM}$ proposed in this paper. The simulation design follows Section 7.15.

ρ	h	$\hat{\beta}_h^{IM}$			$\hat{\beta}_h^{FO\ IM}$			$\hat{\beta}_h^{iBootIM}$		
		Bias	Var	RMSE	Bias	Var	RMSE	Bias	Var	RMSE
					T = 600					
0.70	1	0.007	0.002	0.041	0.000	0.002	0.040	0.000	0.002	0.040
	12	0.010	0.017	0.129	-0.000	0.017	0.131	-0.000	0.017	0.131
	24	0.009	0.017	0.131	-0.000	0.018	0.133	0.000	0.018	0.133
	36	0.009	0.017	0.131	-0.000	0.018	0.133	0.000	0.018	0.133
	48	0.009	0.017	0.131	-0.000	0.018	0.133	0.000	0.018	0.133
0.85	1	0.007	0.001	0.031	0.000	0.001	0.030	0.000	0.001	0.030
	12	0.027	0.026	0.164	0.001	0.029	0.169	0.000	0.029	0.169
	24	0.024	0.034	0.186	-0.000	0.037	0.193	0.001	0.037	0.193
	36	0.023	0.036	0.190	-0.000	0.039	0.197	0.001	0.039	0.196
	48	0.023	0.036	0.191	-0.000	0.039	0.197	0.002	0.039	0.197
0.90	1	0.008	0.001	0.026	0.000	0.001	0.025	0.000	0.001	0.025
	12	0.041	0.028	0.173	0.002	0.031	0.177	0.000	0.032	0.177
	24	0.041	0.045	0.216	0.001	0.051	0.226	0.001	0.051	0.227
	36	0.039	0.051	0.229	0.001	0.058	0.241	0.002	0.058	0.241
	48	0.037	0.053	0.233	0.000	0.060	0.245	0.003	0.060	0.244
0.95	1	0.008	0.000	0.021	0.001	0.000	0.019	0.000	0.000	0.019
	12	0.064	0.025	0.171	0.006	0.029	0.170	0.001	0.029	0.172
	24	0.083	0.054	0.246	0.007	0.066	0.258	0.001	0.068	0.260
	36	0.086	0.072	0.283	0.007	0.093	0.305	0.002	0.095	0.308
	48	0.085	0.084	0.302	0.006	0.109	0.331	0.005	0.112	0.334
0.97	1	0.009	0.000	0.018	0.001	0.000	0.016	0.000	0.000	0.016
	12	0.078	0.021	0.165	0.010	0.025	0.157	0.002	0.026	0.161
	24	0.115	0.052	0.255	0.015	0.066	0.258	0.003	0.071	0.266
	36	0.132	0.077	0.307	0.017	0.106	0.326	0.004	0.115	0.339
	48	0.139	0.096	0.340	0.017	0.138	0.372	0.006	0.152	0.390
0.99	1	0.010	0.000	0.016	0.002	0.000	0.012	-0.000	0.000	0.013
	12	0.102	0.015	0.159	0.022	0.017	0.134	-0.003	0.021	0.145
	24	0.175	0.041	0.268	0.041	0.055	0.239	-0.009	0.070	0.266
	36	0.226	0.068	0.345	0.057	0.102	0.324	-0.018	0.137	0.370
	48	0.264	0.092	0.402	0.068	0.150	0.394	-0.029	0.213	0.463

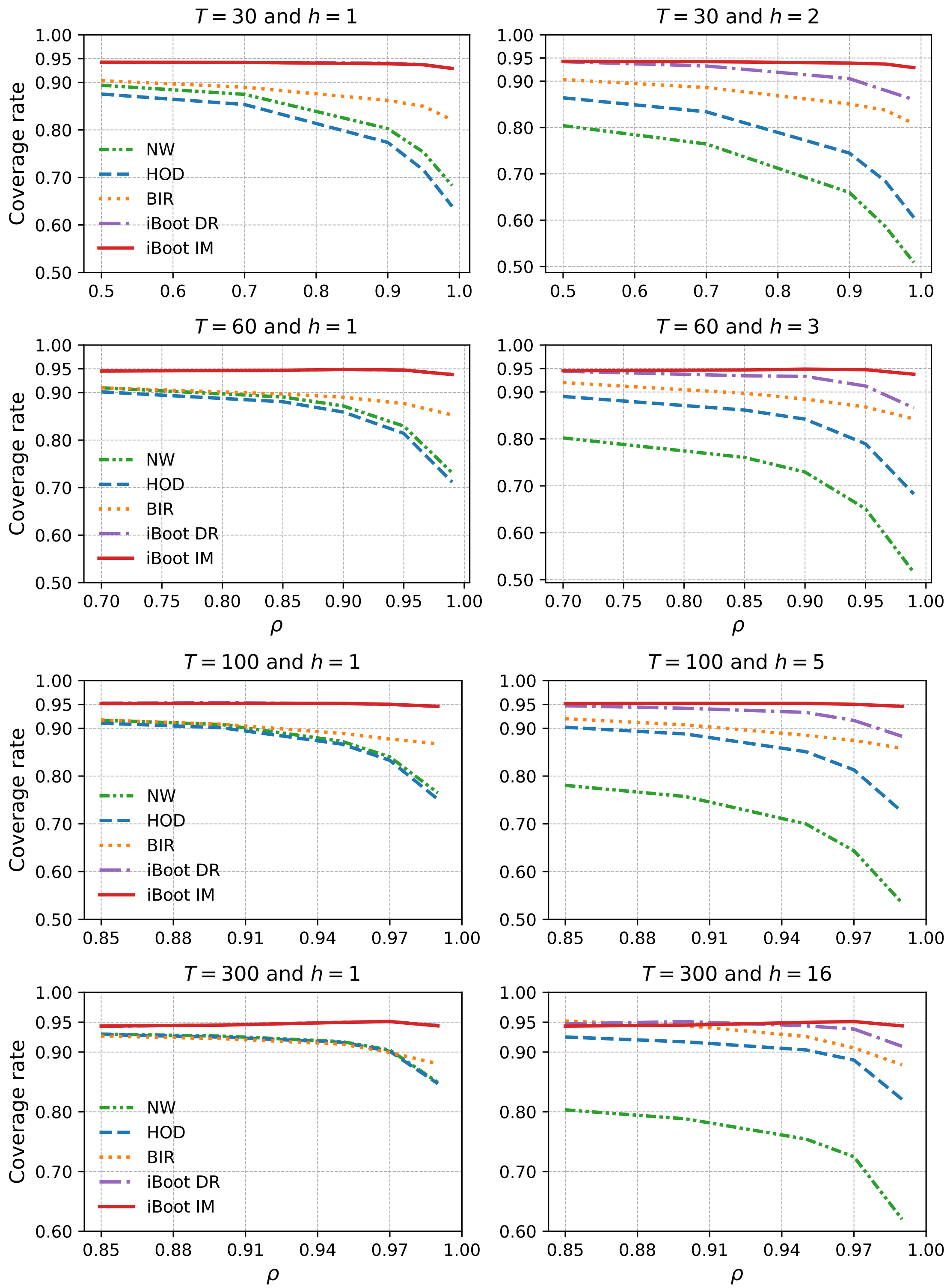


Figure 37: Empirical coverage rates of 95% Confidence Intervals (CIs) for various methods when the errors are GARCH(1,1). NW means the Newey-West CIs, HOD the Hodrick CIs, BIR the CIs using the BIR estimator and the standard errors proposed in Boudoukh et al. (2022), iBoot DR and iBoot IM the CIs proposed in this paper, for the DR and IM estimators respectively.

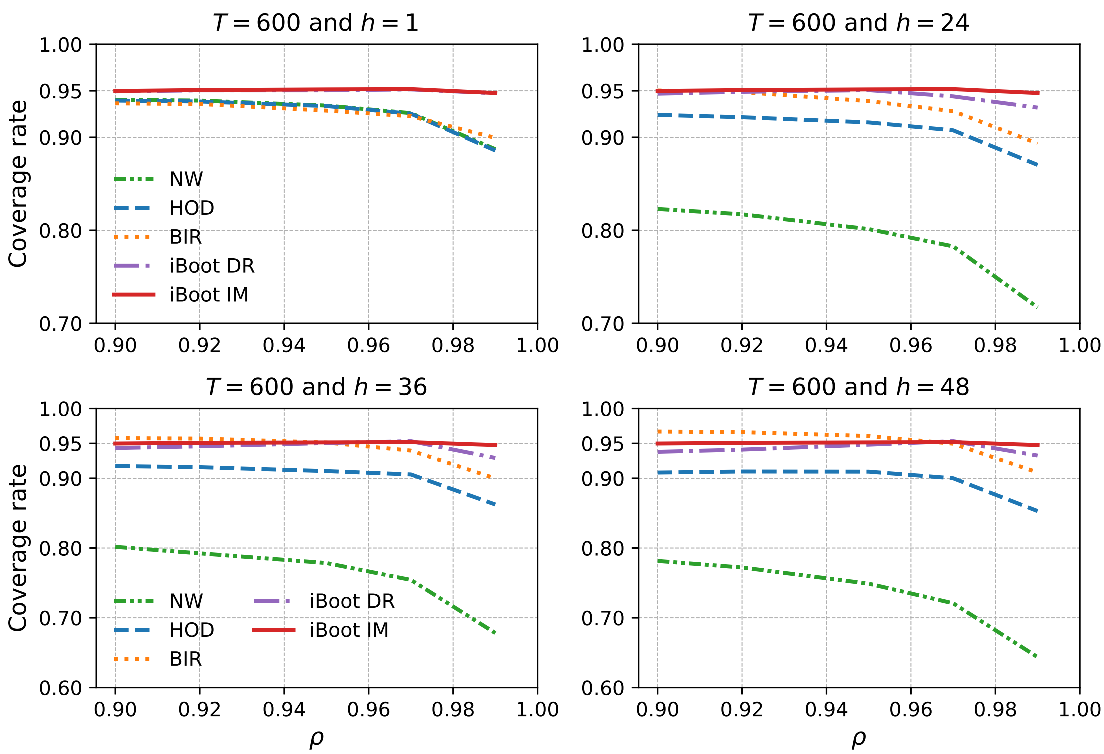


Figure 38: Empirical coverage rates of 95% Confidence Intervals (CIs) for various methods when the errors are GARCH(1,1). NW means the Newey-West CIs, HOD the Hodrick CIs, BIR the CIs using the BIR estimator and the standard errors proposed in Boudoukh et al. (2022), iBoot DR and iBoot IM the CIs proposed in this paper, for the DR and IM estimators respectively.

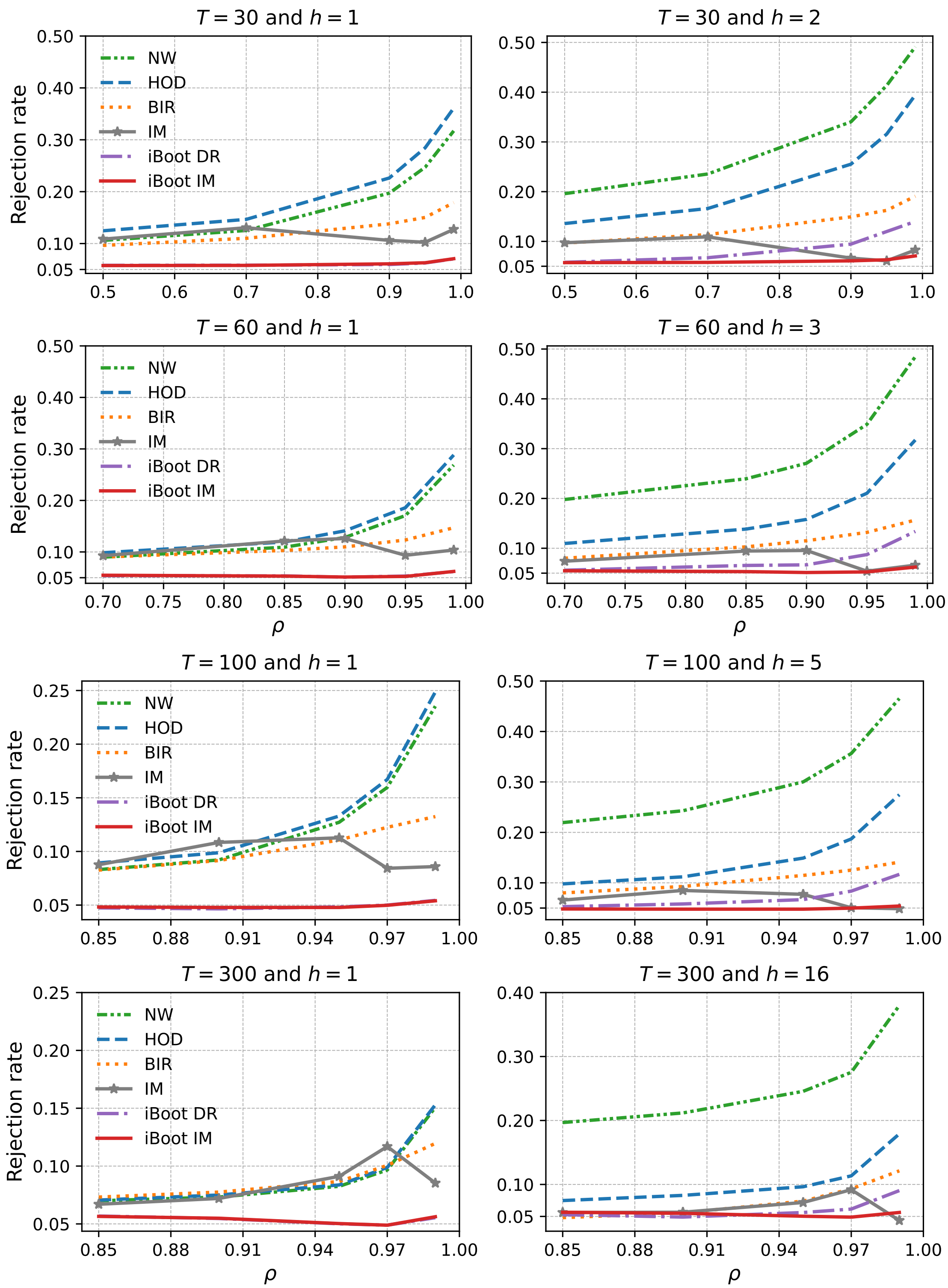


Figure 39: Size curves under $H_0 : \beta_1 = 0$ for various tests with GARCH(1,1) errors: Newey-West (NW), Hodrick (HOD), Boudoukh et al. (2022) (BIR), Xu (2020) (IM), and the proposed iBoot tests for Direct Regression (iBoot DR) and Implication-Based (iBoot IM) estimators.

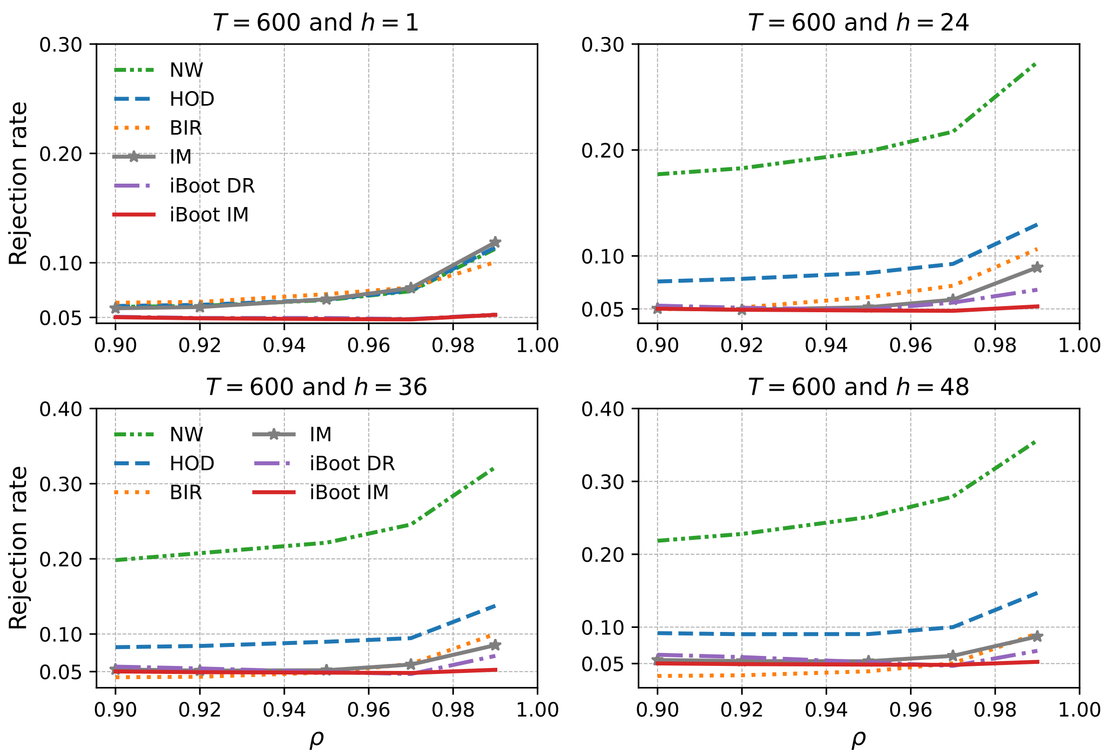


Figure 40: Size curves under $H_0 : \beta_1 = 0$ for various tests across horizons with GARCH(1,1) errors and $T = 600$: Newey-West (NW), Hodrick (HOD), Boudoukh et al. (2022) (BIR), Xu (2020) (IM), and the proposed iBoot tests for Direct Regression (iBoot DR) and Implication-Based (iBoot IM) estimators.

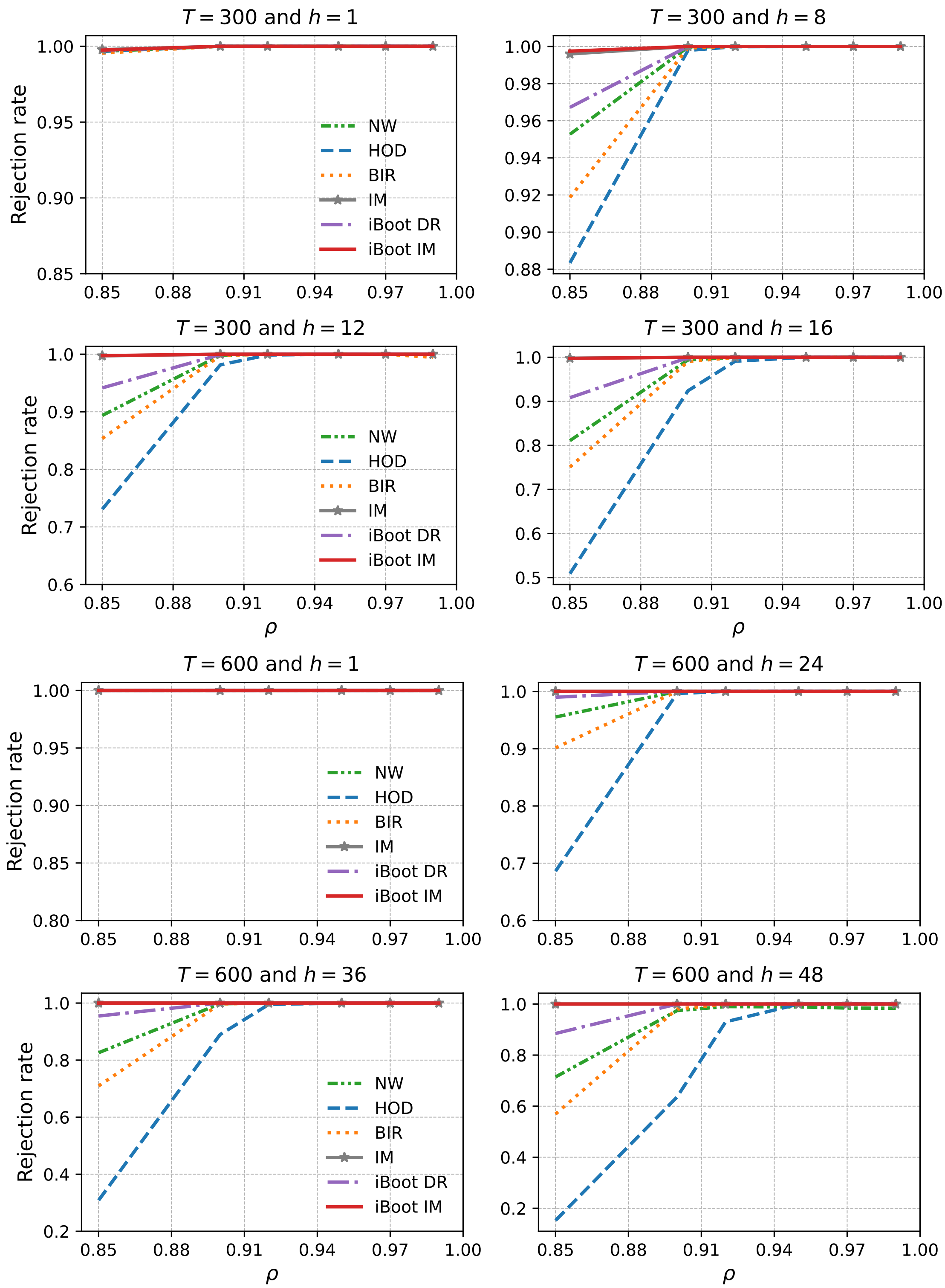


Figure 41: Size-adjusted power curves for testing $H_0 : \beta_1 = 0$ when the true $\beta_1 = 0.16$, across horizons with GARCH(1,1) errors. The figure compares tests from Newey-West (NW), Hodrick (HOD), Boudoukh et al. (2022) (BIR), Xu (2020) (IM), and our proposed iBoot DR and iBoot IM, for $T = 300$ and $T = 600$.

Table 31: DR estimators in the original sample – Forward implied variance (*impvar*)

This table presents Direct Regression (DR) estimators of the forward implied variance (*impvar*) in the original sample (Sep 1998-Sep 2008) in Bakshi et al. (2011). We report the sample period, number of observations (T), autoregressive coefficient ($\hat{\rho}_h$), covariance-to-variance ratio ($\hat{\sigma}_{uv}/\hat{\sigma}_v^2$), and the coefficient estimates from OLS ($\widehat{\beta}_h$), BIR (Boudoukh et al., 2022), and iBoot DR. OLS inference uses Newey-West standard errors; BIR, analytical; and iBoot, our proposed method from Section 3.2. Asterisks indicate significance at the 10%, 5%, and 1% levels.

Sample Period	T	$\hat{\rho}_h$	$\hat{\sigma}_{uv}/\hat{\sigma}_v^2$	h	$\hat{\beta}_h$	$\hat{\beta}_h^{BIR}$	$\hat{\beta}_h^{iBootDR}$
199809-200809	121	0.682	-11.77	1	2.254	1.956	1.940
				2	2.058	1.548	1.448
				12	-8.151	-9.720	-14.766
				16	-13.812	-15.795	-23.090*
				20	-23.662	-26.096	-36.842**
				24	-33.832*	-36.757**	-49.493***
				30	-44.896***	-48.637**	-62.101***
				36	-51.203***	-55.876**	-63.704***
				48	-54.263***	-61.260**	-60.084**

Table 32: DR estimators in the original sample – Government investment (*govik*)

This table presents Direct Regression (DR) estimators of the government investment (*govik*) in the original sample (1947Q2-2010Q4) in Belo and Yu (2013). We report the sample period, number of observations (T), autoregressive coefficient ($\hat{\rho}_h$), covariance-to-variance ratio ($\hat{\sigma}_{uv}/\hat{\sigma}_v^2$), and the coefficient estimates from OLS ($\widehat{\beta}_h$), BIR (Boudoukh et al., 2022), and iBoot DR. OLS inference uses Newey-West standard errors; BIR, analytical; and iBoot, our proposed method from Section 3.2. Asterisks indicate significance at the 10%, 5%, and 1% levels.

Sample Period	T	$\hat{\rho}_h$	$\hat{\sigma}_{uv}/\hat{\sigma}_v^2$	h	$\hat{\beta}_h$	$\hat{\beta}_h^{BIR}$	$\hat{\beta}_h^{iBootDR}$
1947Q2-2010Q4	255	0.993	3.38	1	1.195**	1.248**	1.213***
				2	2.362***	2.488**	2.415***
				4	4.799***	5.130**	4.960***
				8	9.491***	10.473**	10.049***
				12	13.263***	15.218**	14.493***
				16	15.701***	18.952**	17.831***