Dissecting Anomalies in Conditional Asset Pricing[*]

ONLINE APPENDIX

Valentina Raponi[†] Paolo Zaffaroni[‡]

[*]We are grateful to Bryzgalova, Lerner, Lettau, and Pelger (2025) and Freyberger, Hoeppner, Neuhierl, and Weber (2025) for sharing with us their data. We thank S. Bryzgalova, V. DeMiguel, P. Gagliardini (SOFIE discussant), N. Gospodinov (AFA discussant), D. Massacci (York discussant), M. Pelger (AFA chair), G. Skoulakis (WFA discussant), M. Weber, and seminar participants at Berlin, IESE, Lancaster, LBS, Imperial College London, and Rotterdam for comments. This project has received funding from Grant PID2020-115069GB-I00 and by the Horizon 2020 programme under the Marie Sklodowska-Curie grant agreement No. 801370.

[†]IESE Business School, University of Navarra, Barcelona, Spain. Email: vraponi@iese.edu

[‡]Imperial College London, ~~London,~~ UK and Sapienza University of Rome, ~~Rome,~~ Italy.
Email: p.zaffaroni@imperial.ac.uk (corresponding author)

Contents

OA.1 Literature Review

The empirical literature on asset pricing anomalies is very extensive, with more than 400 papers proposing (or *dissecting*) anomalies deemed relevant for explaining the cross-sectional variation of stock returns (see Hou, Chen, and Zhang (2020) for a detailed list). These findings have spurred a growing literature that attempts to summarize (or *digest*) such cross-sectional variation with new risk factors.[1] Nevertheless, many asset-specific characteristics cannot be subsumed by risk factors and remain important determinants of average equity returns (see Daniel and Titman (1997), Lewellen (2015), and Dong, Yan, Rapach, and Zhou (2022), among others).

The apparent significance of such a wide range of anomalies can, in part, be attributed to a lack of robust statistical methodologies. A recent example is Hou, Chen, and Zhang (2020), which casts doubt on the empirical validity of 452 anomalies proposed in the asset pricing and accounting literature, showing that 65% fail to explain the cross-section of average stock returns, with the largest failure rate (96%) found in the trading frictions literature. This concern becomes even more severe once multiple-testing adjustments are applied (see Harvey, Liu, and Zhu (2016)).

The two-pass methodology augmented with anomalies has been studied and extended in many directions. Important results, valid under large-T and assuming constant premia, were derived by Jagannathan and Wang (1998), who established the limiting distribution of the CSR OLS estimator under the null hypothesis of no effect of anomalies. Brennan, Chordia, and Subrahmanyam (1998) propose first netting out average returns from exposures to common risk factors, and then regressing the risk-adjusted average returns on firms' characteristics to test for anomalies. Their approach was further extended by Avramov and Chordia (2006), allowing for time variation in risk exposures through observed state variables. Chordia, Goyal, and Shanken (2015) examine the two-pass estimator when N is much larger than T, and where anomaly variables also vary over time, proposing a bootstrap procedure for standard error estimation.

Beyond the two-pass framework, alternative approaches have been developed to quantify the economic and statistical relevance of anomalies. Notable examples include the semi-parametric

[1] Prominent examples are the Fama and French (1993) and Fama and French (2015) factors, the momentum factors of Carhart (1997) and Jegadeesh and Titman (1993), the liquidity factors of Pástor and Stambaugh (2003) and Acharya and Pedersen (2005), the idiosyncratic risk factor of Ang, Hodrick, Xing, and Zhang (2006), the four q-factors of Hou, Chen, and Zhang (2015), and the lucky factors of Stambaugh and Yuan (2017), among many others.

estimation of Connor and Linton (2007), the Projected Principal Component Analysis of Fan, Liao, and Wang (2016), the Instrumented Principal Component Analysis of Kelly, Pruitt, and Su (2019), and the Bayesian approach of Kozak, Nagel, and Santosh (2020). Other studies examine how firms' characteristics affect investors' portfolio choices (see, e.g., DeMiguel, Martin-Utrera, Nogales, and Uppal (2020) and Kim, Korajczyk, and Neuhierl (2021)). Using non-parametric methods, Freyberger, Neuhierl, and Weber (2020) show that characteristics play a crucial role in model selection and return predictability.

Most of the empirical literature on anomalies relies on portfolio data constructed from a relatively small set of asset-specific predictors. While portfolios reduce sampling variability in estimated risk exposures, they also reduce dispersion in the cross-sectional distribution of risk exposures and residual risk, leading to larger standard errors of premia estimates (see Ang, Liu, and Schwarz (2020)). Moreover, portfolios make it difficult to investigate the joint effect of a high-dimensional set of anomalies. They may also be prone to data-snooping biases, particularly when the same dataset is repeatedly examined (see Lo and MacKinlay (1990), Brennan, Chordia, and Subrahmanyam (1998), Conrad, Cooper, and Kaul (2003), Barras, Scaillet, and Wermers (2010), McLean and Pontiff (2016), and Chen (2021)). These issues can be substantially mitigated—if not avoided—by our approach, which applies to large cross-sections of individual assets, datasets that are typically much less scrutinized than portfolio-level data.

OA.2 Conditional Asset Pricing with Anomalies: The General Case without Risk-Free Asset

This section, which can be read independently of the main manuscript, generalizes our methodology when the risk-free asset is not tradable and $\mathbf{R}_t$ denotes gross returns. We now assume that asset returns are governed by the following conditional asset pricing factor model:

$$\mathbf{R}_t = \boldsymbol{\alpha}_{t-1} + \mathbf{B}_{t-1}\mathbf{f}_t + \boldsymbol{\epsilon}_t, \tag{OA.1}$$

where $\mathbf{R}_t$ denotes the $N \times 1$ vector of gross asset returns at time t, $\boldsymbol{\alpha}_{t-1} = (\alpha_{1,t-1}, \ldots, \alpha_{N,t-1})'$, $\mathbf{B}_{t-1} = (\boldsymbol{\beta}_{1,t-1}, \ldots, \boldsymbol{\beta}_{N,t-1})'$, where $\boldsymbol{\beta}_{i,t-1} = (\beta_{i1,t-1}, \ldots, \beta_{iK_\mathrm{f},t-1})'$ is the vector of time-varying risk exposures on K_f observed risk factors $\mathbf{f}_t = (f_{1,t}, \ldots, f_{K_\mathrm{f},t})'$, and $\boldsymbol{\epsilon}_t = (\epsilon_{1,t}, \ldots, \epsilon_{N,t})'$.

When conditional no-arbitrage and full diversification of the mean-variance frontier hold (see Chamberlain and Rothschild (1983, Corollary 1) and Hansen and Richard (1987) for an extension to a conditional asset pricing setup), exact pricing follows. That is:

$$\mathrm{E}\left[R_{it}|I_{t-1}, \mathbf{\Pi}\right] = \gamma_{0,t-1} + \boldsymbol{\gamma}'_{\mathrm{f,t-1}}\boldsymbol{\beta}_{i,t-1}, \tag{OA.2}$$

where $\mathrm{E}[\cdot]$ denotes the expectation operator, I_t represents the information set available up to time t, and $\mathbf{\Pi}$ defines the complete set of parameters, known to the agent when evaluating expected returns, with $\{\boldsymbol{\gamma}_0, \boldsymbol{\gamma}_{\mathrm{f}}, \mathbf{B}_{1:T}\} \subset \mathbf{\Pi}$, where $\boldsymbol{\gamma}_0 = (\gamma_{0,1}, \cdots, \gamma_{0,T-1})'$ denotes the zero-beta rate vector, $\boldsymbol{\gamma}_{\mathrm{f}} = (\boldsymbol{\gamma}_{\mathrm{f},1}, \cdots, \boldsymbol{\gamma}_{\mathrm{f},T-1})'$ denotes the risk premia matrix associated with the observed risk factors $\mathbf{f}_t$, and $\mathbf{B}_{1:T} = (\mathbf{B}_1, \cdots, \mathbf{B}_{T-1})$ is the matrix of risk exposures.

However, in this paper, we are specifically interested in situations where (OA.2) might not hold, and we replace it by

$$\mathrm{E}\left[R_{it}|I_{t-1}, \mathbf{\Pi}\right] = a_{i,t-1} + \gamma_{0,t-1} + \boldsymbol{\gamma}'_{\mathrm{f,t-1}}\boldsymbol{\beta}_{i,t-1}, \tag{OA.3}$$

for some time-varying and asset-specific *pricing errors* $\mathbf{a} \subset \mathbf{\Pi}$, with $\mathbf{a} = (\mathbf{a}_1, \cdots, \mathbf{a}_{T-1})'$, setting $\mathbf{a}_{t-1} = (a_{1,t-1}, \cdots, a_{N,t-1})'$, where the pricing errors $a_{i,t-1}$ are governed by some *observed* anomaly variates, represented by a $K_{\mathrm{z}} \times 1$ vector of *asset-specific* time-varying variables, $\mathbf{z}_{i,t-1}$, which we refer to as *anomalies*. Formally, we assume that

$$a_{i,t-1} = \boldsymbol{\gamma}'_{\mathrm{z,t-1}}\mathbf{z}_{i,t-1}, \quad \text{for } i = 1, \cdots, N, \text{ and } t = 1, \ldots, T, \tag{OA.4}$$

for some vector of coefficients $\boldsymbol{\gamma}_{\mathrm{z}} \subset \mathbf{\Pi}$, where $\boldsymbol{\gamma}_{\mathrm{z}} = (\boldsymbol{\gamma}_{\mathrm{z},1}, \cdots, \boldsymbol{\gamma}_{\mathrm{z,T-1}})'$ denotes the *anomalies' premia* matrix. Clearly, should all the elements of $\boldsymbol{\gamma}_{\mathrm{z}}$ be zero, then exact pricing (OA.2) holds, and no anomaly affects the cross-section of expected returns.

Using (OA.4), the asset pricing relationship in (OA.5) becomes

$$\mathrm{E}\left[R_{it}|I_{t-1}, \mathbf{\Pi}\right] = \gamma_{0,t-1} + \boldsymbol{\gamma}'_{\mathrm{f,t-1}}\boldsymbol{\beta}_{i,t-1} + \boldsymbol{\gamma}'_{\mathrm{z,t-1}}\mathbf{z}_{i,t-1}. \tag{OA.5}$$

The expression in (OA.5) represents our reference asset pricing restriction.[2] It is worth noticing that, while allowing for anomalies, condition (OA.5) does not necessarily represent a deviation from no-arbitrage, but only from exact pricing.[3]

[2]Condition (OA.5) implies that the agent has full information on the anomaly variables $\mathbf{z}_{i,t-1}$ for every stock. If one suspects that the agent's information is not complete (for example, because firms' balance sheet data is released less frequently or with delays), then the asset pricing restriction (OA.5) generalizes to $\mathrm{E}\left[R_{i,t}|I_{t-1}, \mathbf{\Pi}\right] = \gamma_{0,t-1} + \boldsymbol{\gamma}'_{\mathrm{f,t-1}}\boldsymbol{\beta}_{i,t-1} + \boldsymbol{\gamma}'_{\mathrm{z,t-1}}\mathrm{E}\left[\mathbf{z}_{i,t-1}|I_{t-1}, \mathbf{\Pi}\right]$.

[3]See Proposition OA.3 in Section OA.9 for more details.

Now, under (OA.5), the asset pricing model in (1) generalizes to

$$\mathbf{R}_t = \gamma_{0,t-1}\mathbf{1}_N + \mathbf{Z}_{t-1}\boldsymbol{\gamma}_{\mathrm{z,t-1}} + \mathbf{B}_{t-1}\boldsymbol{\delta}_{\mathrm{f,t-1}} + \boldsymbol{\epsilon}_t, \tag{OA.6}$$

where $\mathbf{1}_N$ denotes a $N \times 1$ vector of ones, $\mathbf{Z}_{t-1} = (\mathbf{z}_{1,t-1}, \cdots, \mathbf{z}_{N,t-1})'$ represents the $N \times K_{\mathrm{z}}$ matrix of anomalies at time $t-1$, and where we set

$$\boldsymbol{\delta}_{\mathrm{f,t-1}} = \boldsymbol{\gamma}_{\mathrm{f,t-1}} + \mathbf{f}_t - \mathrm{E}\left[\mathbf{f}_t | I_{t-1}, \boldsymbol{\Pi}\right], \tag{OA.7}$$

which we denominate as the vector of *ex-post* risk premia.[4]

OA.2.1 Anomalies with Time-Varying Premia: OLS-Based Estimation

The results of the previous section show that the conventional approach is not valid whenever one postulates time variation in the (true) anomalies' premia and unless strict orthogonality conditions are satisfied. We now introduce our new methodology, valid when $N \to \infty$ and T remains fixed, and show how all these challenges related to the conventional approach can be resolved, using a new OLS *bias-adjusted* estimator of the time-varying premia $\boldsymbol{\delta}_{\mathrm{f},t-1}$ and $\boldsymbol{\gamma}_{\mathrm{z},t-1}$.[5]

Consider again the *conditional* asset pricing model with locally-constant risk exposures in (OA.6), and rewrite it as

$$\mathbf{R}_t = \mathbf{Z}_{t-1}\boldsymbol{\gamma}_{\mathrm{z},t-1} + \mathbf{X}\boldsymbol{\Gamma}_{\mathrm{f},t-1} + \boldsymbol{\epsilon}_t \tag{OA.8}$$

comma

where $\mathbf{X} = (\mathbf{1}_N, \mathbf{B})$ and $\boldsymbol{\Gamma}_{\mathrm{f},t-1} = (\gamma_{0,t-1}, \boldsymbol{\delta}'_{\mathrm{f},t-1})'$, with $\boldsymbol{\delta}_{\mathrm{f},t-1}$ defined in (OA.7). Since the matrix $\mathbf{X}$ in (OA.8) is unknown, one needs first to estimate the risk exposures $\mathbf{B}$ to make (OA.8) feasible. The conventional two-pass approach typically advocates a simple OLS regression of $\mathbf{R}_t$ on an intercept and the observed risk factors $\mathbf{f}_t$, that is:

$$\hat{\mathbf{B}} = \mathbf{R}'\mathbb{M}_{\mathbf{1}_{T-1}}\mathbf{F}(\mathbf{F}'\mathbb{M}_{\mathbf{1}_{T-1}}\mathbf{F})^{-1} = \mathbf{R}'\mathbf{P}, \tag{OA.9}$$

where $\hat{\mathbf{B}} = (\hat{\boldsymbol{\beta}}_1, \ldots, \hat{\boldsymbol{\beta}}_N)'$, $\mathbf{F} = (\mathbf{f}_2, \cdots \mathbf{f}_T)'$, $\mathbf{R} = (\mathbf{R}_2, \cdots, \mathbf{R}_T)'$, and $\mathbf{P} = \mathbb{M}_{\mathbf{1}_{T-1}}\mathbf{F}(\mathbf{F}'\mathbb{M}_{\mathbf{1}_{T-1}}\mathbf{F})^{-1}$, where we assume that $\mathbf{P}'\mathbf{P} = (\mathbf{F}'\mathbb{M}_{\mathbf{1}_{T-1}}\mathbf{F})^{-1} > 0$ for every T (see Assumption OA.3 in Section OA.3.1), and where, as said before, the matrix $\mathbb{M}_{1_{T-1}} = \mathbf{I}_{T-1} - \mathbf{1}_{T-1}\mathbf{1}'_{T-1}/(T-1)$ de-means

[4] The notion of *ex-post* risk premia was originally coined by Shanken (1992) to denote a noisy yet unbiased version of the ex-ante risk premia $\boldsymbol{\gamma}_{\mathrm{f,t-1}}$ due to the unexpected factor outcomes $\mathbf{f}_t - \mathrm{E}\left[\mathbf{f}_t | I_{t-1}, \boldsymbol{\Pi}\right]$, arising whenever one considers the fixed-T case. Shanken (1992) considers the time-average of $\boldsymbol{\delta}_{\mathrm{f,t-1}}$ with constant premia and expected value, yielding $\boldsymbol{\delta}_{\mathrm{f}} = \boldsymbol{\gamma}_{\mathrm{f}} + \bar{\mathbf{f}} - \mathrm{E}[\mathbf{f}_t]$, whereas here we consider the time-varying ex-post risk premia (OA.7), with a slight abuse of notation as $\boldsymbol{\delta}_{\mathrm{f,t-1}}$ contains $\mathbf{f}_t$, justified by latency of $\boldsymbol{\delta}_{\mathrm{f,t-1}}$.

[5] All the results are established under several regularity conditions, reported in Section OA.3.1.

the data, yielding $\mathbb{M}_{1_{T-1}}\mathbf{R} = \mathbf{R} - \mathbf{1}_{T-1}\bar{\mathbf{R}}'$ and $\mathbb{M}_{1_{T-1}}\mathbf{F} = \mathbf{F} - \mathbf{1}_{T-1}\bar{\mathbf{f}}'$, setting $\bar{\mathbf{R}} = \sum_{t=2}^{T}\mathbf{R}_t/(T-1)$ and $\bar{\mathbf{f}} = \sum_{t=2}^{T}\mathbf{f}_t/(T-1)$.

It is clear that the OLS estimator in (OA.9) *excludes* (on purpose) the potential effect of the anomalies $\mathbf{Z}_{t-1}$, as well as the time variation of their premia. This could induce sources of bias in the estimates, further exacerbated if $\mathbf{f}_t$ and $\mathbf{Z}_{t-1}$ were potentially correlated across time, making $\hat{\mathbf{B}}$ invalid. This challenge is resolved by introducing the following smoothness Assumption OA.1, which constraints the time variation of the premia parameters by implying their (temporal) orthogonality with the risk factors $\mathbf{f}_t$. As the time-series dimension T gets small (and as long as $T > K_{\mathrm{f}} + 1$), this assumption appears mild, especially in terms of anomalies' premia, where the observed (time-varying) $\mathbf{z}_{i,t-1}$ could account for most of the time variation of their overall contribution to expected returns.

Assumption OA.1 (*Smoothness of the premia parameters*)**.** *Define the* $(T-1) \times K_{\mathrm{f}}$ *matrix*

$$\breve{\boldsymbol{\delta}}_{\mathrm{f}} = (\breve{\boldsymbol{\delta}}_{f,1}, \ldots, \breve{\boldsymbol{\delta}}_{\mathrm{f},T-1})',$$

where

$$\breve{\boldsymbol{\delta}}_{\mathrm{f},t-1} = \boldsymbol{\delta}_{\mathrm{f},t-1} - \mathbf{f}_t = \boldsymbol{\gamma}_{\mathrm{f},t-1} - E(\mathbf{f}_t \mid I_{t-1}, \boldsymbol{\Pi}).$$

Let $\boldsymbol{\Delta}_z$ *denote the* $(T-1) \times N$ *matrix*

$$\boldsymbol{\Delta}_z = \begin{bmatrix} \boldsymbol{\gamma}_{z,1}' - \boldsymbol{\gamma}_{\mathrm{z}}' & \mathbf{0}_{K_{\mathrm{z}}}' & \cdots & \mathbf{0}_{K_{\mathrm{z}}}' \\ \mathbf{0}_{K_{\mathrm{z}}}' & \boldsymbol{\gamma}_{z,2}' - \boldsymbol{\gamma}_{\mathrm{z}}' & \cdots & \mathbf{0}_{K_{\mathrm{z}}}' \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{K_{\mathrm{z}}}' & \mathbf{0}_{K_{\mathrm{z}}}' & \cdots & \boldsymbol{\gamma}_{\mathrm{z},t-1}' - \boldsymbol{\gamma}_{\mathrm{z}}' \end{bmatrix} \begin{bmatrix} \mathbf{Z}_1' \\ \mathbf{Z}_2' \\ \vdots \\ \mathbf{Z}_{T-1}' \end{bmatrix},$$

for some constant $K_{\mathrm{z}} \times 1$ *vector* $\boldsymbol{\gamma}_{\mathrm{z}}$ *satisfying*

$$\frac{1}{N}\sum_{i=1}^{N}(\mathbf{Z}_i'\mathbf{Z}_i)^{-1}\mathbf{Z}_i'\mathbf{R}_i \;\to_p\; \boldsymbol{\gamma}_{\mathrm{z}}.$$

Then, the following conditions are assumed to hold:

$$\mathbf{P}'\boldsymbol{\gamma}_0 = \mathbf{0}_{K_{\mathrm{f}}}, \qquad \mathbf{P}'\breve{\boldsymbol{\delta}}_{\mathrm{f}} = \mathbf{0}_{K_{\mathrm{f}} \times K_{\mathrm{f}}}, \qquad \frac{\mathbf{P}'\boldsymbol{\Delta}_z\boldsymbol{\epsilon}}{N} = o_p\left(\frac{1}{\sqrt{N}}\right).$$

Assumption OA.1 is extremely mild in most cases of interest. When the premia parameters are *locally* constant over the time interval of size T, and the test assets and risk factors are expressed as excess returns under the assumption that a risk-free asset is tradable, Assumption OA.1 is automatically satisfied.[6]

Although Assumption OA.1 helps mitigate the bias arising from the temporal correlation between the factors and the premia parameters, an additional source of bias remains in the estimation of (OA.9) due to the exclusion of $\mathbf{Z}_{t-1}$ in the first pass. This extra bias is addressed by *orthogonalizing* the anomaly variables $\mathbf{Z}_{t-1}$ with respect to the observable factors $\mathbf{f}_t$, *before* running the two-step procedure. Therefore, $\mathbf{Z}_{t-1}$ can be interpreted as representing the *net* component of the anomaly variables that affects expected returns, thus eliminating any *indirect* influence (or confounding effect) arising from the risk factors.[7]

To better understand the implications of the orthogonalization on the model's parameters and their corresponding interpretations, consider the case where the researcher posits a model that involves a set of *initial* anomalies $\mathbf{Z}^{\dagger}_{t-1} = (\mathbf{z}^{\dagger}_{1,t-1}, \cdots, \mathbf{z}^{\dagger}_{N,t-1})'$, such that:

$$\mathbf{R}_t = \boldsymbol{\alpha}^{\dagger}_{t-1} + \mathbf{Z}^{\dagger}_{t-1}\boldsymbol{\gamma}_{\mathrm{z},t-1} + \mathbf{B}^{\dagger}\mathbf{f}_{t-1} + \mathbf{e_t} \qquad \text{(OA.10)}$$

comma

where $\mathbf{B}^{\dagger}$ will be, in general, different from $\mathbf{B}$. Then, starting from (OA.10), one can construct the *orthogonal* anomalies $\mathbf{Z}_{t-1}$ as the residuals from projecting $\mathbf{Z}^{\dagger}_{t-1}$ onto the unit constant and $\mathbf{f}_t$, and re-centering, leading to:

$$\mathbf{z}_{i,t-1} = \mathbf{z}^{\dagger}_{i,t-1} - \hat{\boldsymbol{\Sigma}}_{\mathbf{z}^{\dagger}_i\mathbf{f}}\hat{\boldsymbol{\Sigma}}^{-1}_{\mathbf{f}}(\mathbf{f}_t - \bar{\mathbf{f}}), \qquad \text{(OA.11)}$$

where $\hat{\boldsymbol{\Sigma}}_{\mathbf{z}^{\dagger}_i\mathbf{f}} = \widehat{\mathrm{Cov}}(\mathbf{z}^{\dagger}_{i,t-1}, \mathbf{f}') = \frac{1}{T-1}\mathbf{Z}^{\dagger\prime}_i\mathbb{M}_{1_{T-1}}\mathbf{F}$, and $\hat{\boldsymbol{\Sigma}}_{\mathbf{f}} = \widehat{\mathrm{Var}}(\mathbf{f}) = \frac{1}{T-1}\mathbf{F}'\mathbb{M}_{1_{T-1}}\mathbf{F}$, where $\mathbf{Z}^{\dagger}_i = (\mathbf{z}^{\dagger}_{i,1}, \cdots, \mathbf{z}^{\dagger}_{i,T-1})'$.and where we use $\widehat{\mathrm{Cov}}(\cdot)$ and $\widehat{\mathrm{Var}}(\cdot)$ to denote the sample covariance and sample variance estimators, respectively. Then, replacing (OA.11) in (OA.10), and imposing the asset

[6] One can avoid imposing the smoothness conditions of Assumption OA.1, thereby allowing for time-series dependence between the time-varying premia and the risk factors, albeit at the cost of more complicated expressions. In particular, (OA.6) can be expressed as a panel data model with interactive fixed effects:

$$\mathbf{R}_t = \boldsymbol{\alpha} + \mathbf{Z}_{t-1}\bar{\boldsymbol{\gamma}}_{\mathrm{z}} + \mathbf{B}\mathbf{f}_t + \mathbf{u}_t,$$

where the error term satisfies $\mathbf{u}_t = \boldsymbol{\xi}_t + \boldsymbol{\Delta}\mathbf{g}_t$, for an asset-specific error $\boldsymbol{\xi}_t$ and a vector of zero-mean latent factors $\mathbf{g}_t$ that may be correlated with the observed risk factors $\mathbf{f}_t$, with risk exposures $\boldsymbol{\Delta}$. Moreover, $\bar{\boldsymbol{\gamma}}_{\mathrm{z}} = T^{-1}\sum_{t=1}^{T}\boldsymbol{\gamma}_{t-1,\mathrm{z}}$. Assumption OA.1 implies orthogonality between $\mathbf{f}_t$ and $\mathbf{u}_t$, which restores the consistency of the OLS estimator $\hat{\mathbf{B}}$.

An alternative estimator for $\mathbf{B}$ exists (see Avarucci and Zaffaroni (2022)) that avoids Assumption OA.1, but at the cost of a more involved analysis of the CSR in the second pass. Details are available upon request.

[7] The orthogonalization between anomalies and risk factors implies that $\mathbf{Z}_{t-1}$ are no longer pre-determined. By standard arguments, this introduces a bias asymptotically irrelevant under our large-N–fixed-T sampling scheme.

pricing restriction in (OA.5), we get model (OA.8), where, setting $\bar{\boldsymbol{\gamma}}_{\mathrm{z}} = T^{-1}\sum_{t=1}^{T}\boldsymbol{\gamma}_{t-1,z}$,

$$\boldsymbol{\beta}_i \quad = \quad \boldsymbol{\beta}_i^{\dagger} + \hat{\boldsymbol{\Sigma}}_{\mathbf{f}}^{-1}\hat{\boldsymbol{\Sigma}}'_{\mathbf{z}_i^{\dagger}\mathbf{f}}\bar{\boldsymbol{\gamma}}_{\mathrm{z}}. \tag{OA.12}$$

From (OA.12), it is easy to see that, after the orthogonalization of the anomaly variables, the (transformed) $\mathbf{B}$ takes now into account not only the direct effect of the risk factors on the cross-section of expected returns but also the indirect effect of $\mathbf{f}_t$, trough its possible time-series dependence with $\mathbf{Z}_{t-1}^{\dagger}$.[8]

This setup is extremely convenient and allows us to estimate the matrix $\mathbf{B}$ by simply using (20), without now incurring any source of bias coming from the exclusion of anomalies from the first-pass regression or due to the potential correlation between risk factors and anomalies. Therefore, the feasible version of (OA.8) becomes

$$\mathbf{R}_t = \hat{\mathbf{X}}\boldsymbol{\Gamma}_{\mathrm{f},t-1} + \mathbf{Z}_{t-1}\boldsymbol{\gamma}_{\mathrm{z},t-1} + \boldsymbol{\eta}_t, \tag{OA.13}$$

setting $\hat{\mathbf{X}} = (\mathbf{1}_N, \hat{\mathbf{B}})$, with $\hat{\mathbf{B}}$ defined in (20), $\boldsymbol{\eta}_t = \boldsymbol{\epsilon}_t - (\hat{\mathbf{X}} - \mathbf{X})\boldsymbol{\Gamma}_{\mathrm{f},t-1}$, and where $\mathbf{Z}_{t-1}$ satisfies (OA.11), hence being uncorrelated (even in sample) with the risk factors. Running a single CSR OLS regression on (OA.13) yields the time-varying OLS estimator

$$\begin{bmatrix}\hat{\boldsymbol{\Gamma}}_{\mathrm{f},t-1}\\ \hat{\gamma}_{\mathrm{z},t-1}\end{bmatrix} = \begin{bmatrix}\hat{\mathbf{X}}'\hat{\mathbf{X}} & \hat{\mathbf{X}}'\mathbf{Z}_{t-1}\\ \mathbf{Z}'_{t-1}\hat{\mathbf{X}} & \mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}\end{bmatrix}^{-1}\begin{bmatrix}\hat{\mathbf{X}}'\mathbf{R}_t\\ \mathbf{Z}'_{t-1}\mathbf{R}_t\end{bmatrix}, \tag{OA.14}$$

where $\hat{\boldsymbol{\Gamma}}_{\mathrm{f},t-1} = (\hat{\gamma}_{0,t-1}, \hat{\boldsymbol{\delta}}'_{\mathrm{f},t-1})'$.[9] The estimator in (OA.14) generalizes the conventional estimator $\tilde{\gamma}_{\mathrm{z},t-1}$ in (21) to the case of when both anomalies and (estimated) risk exposures are used as regressors in the feasible model.[10] The two estimators coincide when $\hat{\mathbf{X}}'\mathbf{Z}_{t-1} = \mathbf{0}_{1+K_{\mathrm{f}}\times K_{\mathrm{z}}}$, a condition which is, however, not warranted in general. When such orthogonality condition is violated, then $\hat{\boldsymbol{\gamma}}_{\mathrm{z},t-1}$ in (OA.14) remains valid, but $\tilde{\gamma}_{\mathrm{z},t-1}$ in (21) becomes biased.[11]

[8]Notice that, by taking T small, one mitigates the effect of the likely time-variation of the (conditional) second moments of the $\mathbf{f}_t$ and $\mathbf{Z}_{t-1}^{\dagger}$ which is not taken explicitly into account by the sample estimates $\hat{\boldsymbol{\Sigma}}_{\mathbf{f}}^{-1}$ and $\hat{\boldsymbol{\Sigma}}_{\mathbf{z}_i^{\dagger}\mathbf{f}}$.

[9]To simplify the exposition we are slightly abusing the notation by indexing the OLS estimator (OA.14) to time $t-1$ instead to time t.

[10]The OLS premia estimates (OA.14) can also be interpreted as portfolio returns (see Fama (1976)). In particular, it is immediate to see how $\hat{\gamma}_{0,t-1}$ would represent a unit-long portfolio with a zero spread in both the anomalies and the factor's risk exposures, whereas $\hat{\boldsymbol{\gamma}}_{\mathrm{z},t-1}$ would represent a zero-investment long-short portfolio which is factor-neutral, i.e., an anomalies-based portfolio, with specular properties applying to $\hat{\boldsymbol{\delta}}_{\mathrm{f},t-1}$.

[11]To clarify, notice that the orthogonality condition that we impose between $\mathbf{Z}_{t-1}$ and $\mathbf{f}_t$ represents a *time-series* restriction, which does not imply the *cross-sectional* restriction $\hat{\mathbf{X}}'\mathbf{Z}_{t-1} = \mathbf{0}_{1+K_{\mathrm{f}}\times K_{\mathrm{z}}}$.

Although (OA.14) resolves the bias coming from the potential lack of orthogonality between the risk factors and the anomalies, unfortunately, other sources of bias arise in our large-N–fixed-T set-up. The reason is that $\hat{\mathbf{B}}$ does not converge to $\mathbf{B}$ when T is fixed, making the OLS estimator in (OA.14) biased due the EIV effect.[12]

However, we show that such biases can be consistently estimated, leading to our new *bias-adjusted* CSR OLS estimator:

$$\begin{bmatrix} \hat{\mathbf{\Gamma}}^*_{\mathrm{f},t-1} \\ \hat{\gamma}^*_{\mathrm{z},t-1} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{X}}'\hat{\mathbf{X}} - N\hat{\mathbf{\Lambda}}_1 & \hat{\mathbf{X}}'\mathbf{Z}_{t-1} \\ \mathbf{Z}'_{t-1}\hat{\mathbf{X}} & \mathbf{Z}'_{t-1}\mathbf{Z}_{t-1} \end{bmatrix}^{-1} \begin{bmatrix} \hat{\mathbf{X}}'\mathbf{R}_t - N\hat{\mathbf{\Lambda}}_{2,t-1} \\ \mathbf{Z}'_{t-1}\mathbf{R}_t \end{bmatrix}, \tag{OA.15}$$

where $\hat{\mathbf{\Gamma}}^*_{\mathrm{f},t-1} = (\hat{\gamma}^*_{0,t-1}, \hat{\boldsymbol{\delta}}^{*\prime}_{\mathrm{f},t-1})'$, and where we set

$$\hat{\mathbf{\Lambda}}_1 = \begin{bmatrix} 0 & \mathbf{0}'_{K_\mathrm{f}} \\ \mathbf{0}_{K_\mathrm{f}} & \hat{\sigma}^2\mathbf{P}'\mathbf{P} \end{bmatrix}, \qquad \hat{\mathbf{\Lambda}}_{2,t-1} = \hat{\sigma}^2 \begin{bmatrix} 0 \\ \mathbf{P}'\boldsymbol{\imath}_{t-1,T-1} \end{bmatrix}, \tag{OA.16}$$

where $\boldsymbol{\imath}_{s,T-1}$ denotes the s-th row of the identity matrix $\mathbf{I}_{T-1}$, and where

$$\hat{\sigma}^2 = \frac{\mathrm{tr}(\hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}})}{N(T-K-2)}, \tag{OA.17}$$

with $\mathrm{tr}(\cdot)$ denoting the trace operator, $K = K_\mathrm{f} + K_\mathrm{z}$, and where $\hat{\boldsymbol{\epsilon}}$ represents the OLS residuals, defined as $\hat{\boldsymbol{\epsilon}}_i = \mathbf{M}_{\tilde{\mathbf{D}}_i}\mathbf{R}_i$, with $\mathbb{M}_{\tilde{\mathbf{D}}_i} = \mathbf{I}_{T-1} - \tilde{\mathbf{D}}_i(\tilde{\mathbf{D}}'_i\tilde{\mathbf{D}}_i)^{-1}\tilde{\mathbf{D}}'_i$, and $\tilde{\mathbf{D}}_i = (\mathbf{D}, \tilde{\mathbf{Z}}_i)$, with $\mathbf{D} = (\mathbf{1}_{T-1}, \mathbf{F})$, and $\tilde{\mathbf{Z}}_i = \mathbb{M}_{\mathbf{1}_{T-1}}\mathbf{Z}_i$.[13]

The following theorem establishes the limiting properties of our novel bias-adjusted estimator. Let $\mathbf{Z} = (\mathbf{z}_1, ..., \mathbf{z}_N)'$ define the overall $N \times K_\mathrm{z}(T-1)$ matrix of anomalies, with $\mathbf{z}_i$ being the $K_\mathrm{z}(T-1) \times 1$ vector $\mathbf{z}_i = \left(\mathrm{z}^{(1)}_{i,1}, \cdots, \mathrm{z}^{(1)}_{i,T-1}, \cdots, \mathrm{z}^{(K_\mathrm{z})}_{i,1}, \cdots, \mathrm{z}^{(K_\mathrm{z})}_{i,T-1}\right)'$, with $\mathrm{z}^{(j)}_{i,T-1}$ denoting the value of the jth anomaly for stock i at time t. Let $\mathbf{0}_a$ and $\mathbf{1}_a$ denote an $a \times 1$ vector of zeros and ones, respectively. The following $K_\mathrm{z}(T-1) \times K_\mathrm{z}$ matrices of constants

comma

$$\mathbb{J} = \frac{1}{T-1}\begin{bmatrix} \mathbf{1}_{T-1} & \mathbf{0}_{T-1} & \dots & \mathbf{0}_{T-1} \\ \mathbf{0}_{T-1} & \mathbf{1}_{T-1} & \dots & \mathbf{0}_{T-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{T-1} & \mathbf{0}_{T-1} & \dots & \mathbf{1}_{T-1} \end{bmatrix} = \left(\mathbf{I}_{K_\mathrm{z}} \otimes \frac{\mathbf{1}_{T-1}}{(T-1)}\right) = \frac{1}{T-1}\sum_{s=1}^{T-1}\mathbb{J}_s \tag{OA.18}$$

[12]Moreover, as the estimator (OA.14) is evaluated at each point in time, a second source of bias arises (besides the EIV) because $\mathbf{P}'\boldsymbol{\imath}_{t-1,T-1}$ could be, in general, different from $\mathbf{0}_{K_\mathrm{f}}$.

defined?

[13]Note that, while computation of the OLS estimator $\hat{\boldsymbol{\beta}}_i$ only requires the regressors $\mathbf{D}$, the corresponding residuals must be evaluated for both $\mathbf{D}_i$ and $\tilde{\mathbf{Z}}_i$, as it always happens in regressions with orthogonal independent variables.

with

$$\mathbb{J}_s = \begin{bmatrix} \boldsymbol{\iota}_{s,T-1} & \mathbf{0}_{T-1} & \dots & \mathbf{0}_{T-1} \\ \mathbf{0}_{T-1} & \boldsymbol{\iota}_{s,T-1} & \dots & \mathbf{0}_{T-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{T-1} & \mathbf{0}_{T-1} & \dots & \boldsymbol{\iota}_{s,T-1} \end{bmatrix} = (\mathbf{I}_{K_z} \otimes \boldsymbol{\iota}_{s,T-1}) \quad \text{for } 1 \leq s \leq T-1, \tag{OA.19}$$

are needed to evaluate, respectively, the sample means of the anomaly variables and to select the s-th observation from the $\mathbf{Z}$ matrix, yielding $\mathbf{Z}\mathbb{J} = (T-1)\sum_{s=1}^{T-1}\mathbf{Z}_s$ and $\mathbf{Z}\mathbb{J}_s = \mathbf{Z}_s$. Finally, let $\otimes$, vec$(\cdot)$ and $\odot$ denote the Kronecker product, the vec operator, and the Hadamard product, respectively.

Theorem OA.1 (Large-N consistency and asymptotic normality of the time varying bias-adjusted CSR OLS estimator)**.** *As $N \to \infty$, under Assumptions OA.3 –OA.9 (listed in Section OA.3.1), then*

(i)

$$\hat{\mathbf{\Gamma}}^*_{\mathrm{f},t-1} - \mathbf{\Gamma}_{\mathrm{f},t-1} = O_p\left(\frac{1}{\sqrt{N}}\right) \quad \textit{and} \quad \hat{\boldsymbol{\gamma}}^*_{\mathrm{z},t-1} - \boldsymbol{\gamma}_{\mathrm{z},t-1} = O_p\left(\frac{1}{\sqrt{N}}\right), \tag{OA.20}$$

(ii)

$$\sqrt{N}\begin{bmatrix} \hat{\mathbf{\Gamma}}^*_{\mathrm{f},t-1} - \mathbf{\Gamma}_{\mathrm{f},t-1} \\ \hat{\boldsymbol{\gamma}}^*_{\mathrm{z},t-1} - \boldsymbol{\gamma}_{\mathrm{z},t-1} \end{bmatrix} \to_d \mathcal{N}\left(\mathbf{0}_{K+1}, \mathbf{L}^{-1}_{t-1}\mathbf{O}_{t-1}\mathbf{L}^{-1\prime}_{t-1}\right), \tag{OA.21}$$

for some $\mathbf{L}_{t-1} > 0$ and $\mathbf{O}_{t-1}$.[14]

Proof. See Section OA.5.

To conduct statistical inference, we need a consistent estimator of the asymptotic covariance matrix in (OA.21), which we present in the next theorem.

Theorem OA.2 (Standard errors of the time varying bias-adjusted CSR OLS estimator)**.** *As $N \to \infty$, under Assumptions OA.3–OA.9, and the identification condition $\kappa_4 = 0$,*

$$\hat{\mathbf{L}}^{-1}_{t-1}\,\hat{\mathbf{O}}_{t-1}\,\hat{\mathbf{L}}^{-1\prime}_{t-1} \;\to_p\; \mathbf{L}^{-1}_{t-1}\,\mathbf{O}_{t-1}\,\mathbf{L}^{-1\prime}_{t-1} \quad \text{comma} \tag{OA.22}$$

[14]To ease the exposition, the definition of $\mathbf{L}_{t-1}$ and $\mathbf{O}_{t-1}$ has been relegated to the proof of the theorem (see (OA.84)).

where

$$\hat{\mathbf{L}}_{t-1} = \frac{1}{N}\begin{bmatrix} \hat{\mathbf{X}}'\hat{\mathbf{X}} - N\hat{\mathbf{\Lambda}}_1 & \hat{\mathbf{X}}'\mathbf{Z}_{t-1} \\ \mathbf{Z}'_{t-1}\hat{\mathbf{X}} & \mathbf{Z}'_{t-1}\mathbf{Z}_{t-1} \end{bmatrix}, \quad \text{and} \quad \hat{\mathbf{O}}_{t-1} = \begin{bmatrix} \hat{\mathbf{U}}_{t-1} & \hat{\sigma}^2\hat{\mathbf{G}}_{t-1}\hat{\mathbf{H}}'_{t-1} \\ \hat{\sigma}^2\hat{\mathbf{H}}_{t-1}\hat{\mathbf{G}}'_{t-1} & \hat{\mathbf{H}}_{t-1}\hat{\mathbf{\Sigma}}_{\mathrm{U}}\hat{\mathbf{H}}'_{t-1}, \end{bmatrix} \tag{OA.23}$$

comma

with $\hat{\mathbf{U}}_{t-1} = \hat{\sigma}^2\hat{\mathbf{Q}}'_{t-1}\hat{\mathbf{Q}}_{t-1}\left(\hat{\mathbf{\Sigma}}_{\mathrm{X}} - \hat{\mathbf{\Lambda}}_1\right) + \begin{bmatrix} 0 & \mathbf{0}'_{K_{\mathrm{f}}} \\ \mathbf{0}_{K_{\mathrm{f}}} & \hat{\mathbf{V}}'_{t-1}\hat{\mathbf{U}}_{\epsilon}\hat{\mathbf{V}}_{t-1} \end{bmatrix}$, *setting* $\overline{\mathbb{M}}_{\tilde{\mathrm{D}}} = N^{-1}\sum_{i=1}^{N}\mathbb{M}_{\tilde{\mathrm{D}}_i}$, $\hat{\mathbf{\Sigma}}_{\mathrm{X}} = N^{-1}\hat{\mathbf{X}}'\hat{\mathbf{X}}$, $\hat{\mathbf{\Sigma}}_{\mathrm{ZB}} = N^{-1}\mathbf{Z}'\hat{\mathbf{B}}$, $\hat{\boldsymbol{\mu}}_{\mathrm{z},T-1} = N^{-1}\mathbf{Z}'\mathbf{1}_N$, *and* $\hat{\mathbf{\Sigma}}_{\mathrm{U}} = (\hat{\sigma}^2\boldsymbol{I}_{T-1}\otimes\mathbf{Z}'\mathbf{Z}/N)$, *with* $\hat{\mathbf{\Lambda}}_1$ *and* $\hat{\sigma}^2$ *defined in* (OA.16) *and* (OA.17), *respectively,* κ_4 *defined in* (OA.56), *and we define the following matrices*

$$\begin{aligned}
\hat{\mathbf{Q}}_{t-1} &= \boldsymbol{\imath}_{t-1,T-1} - \mathbf{P}\hat{\boldsymbol{\delta}}^*_{\mathrm{f},t-1}, \quad \hat{\mathbf{H}}_{t-1} = \hat{\mathbf{Q}}'_{t-1}\otimes\mathbb{J}'_{t-1}, \\
\hat{\mathbf{G}}_{t-1} &= \left[\hat{\mathbf{Q}}_{t-1}\otimes\hat{\boldsymbol{\mu}}_{\mathrm{z},T-1}, \hat{\mathbf{Q}}_{t-1}\otimes\hat{\mathbf{\Sigma}}_{\mathrm{ZB}}\right]', \text{ and} \\
\hat{\mathbf{V}}_{t-1} &= (\hat{\mathbf{Q}}_{t-1}\otimes\mathbf{P}) - \left(\frac{\mathrm{vec}(\overline{\mathbb{M}}_{\tilde{\mathrm{D}}})}{(T-K-2)}\right)\hat{\mathbf{Q}}'_{t-1}\mathbf{P},
\end{aligned}$$

where $\hat{\mathbf{U}}_{\epsilon}$ *is obtained plugging* $\kappa_4 = 0$ *and* $\hat{\sigma}^4 = N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T-1}\hat{\epsilon}^4_{it}/3\,\mathrm{tr}\left(\overline{\mathbb{M}}^{(2)}_{\tilde{\mathrm{D}}}\right)$, *with* $\overline{\mathbb{M}}^{(2)}_{\tilde{\mathrm{D}}} = \frac{1}{N}\sum_{i=1}^{N}\left(\mathbb{M}_{\tilde{\mathrm{D}}_i}\odot\mathbb{M}_{\tilde{\mathrm{D}}_i}\right)$, *into* $\mathbf{U}_{\epsilon} = \mathbf{U}_{\epsilon}(\kappa_4, \sigma^4)$.[15]

Proof. See Section OA.5.

The square root of the diagonal elements of $\hat{\mathbf{L}}^{-1}_{t-1}\hat{\mathbf{O}}_{t-1}\hat{\mathbf{L}}^{-1}_{t-1}$ in (OA.23), divided by $\sqrt{N}$, represent the standard errors of the premia estimators $\hat{\mathbf{\Gamma}}^*_{\mathrm{f},t-1}$ and $\hat{\gamma}^*_{\mathrm{z},t-1}$, which can be used to construct asymptotically valid confidence intervals.

OA.2.2 Anomalies with Time-Varying Premia: WLS-Based Estimation

Fama and French (2008) and Hou, Chen, and Zhang (2020), among others, recognize that most of the empirical results on asset pricing anomalies can be seriously affected by the presence of micro-cap stocks. Small-cap equities typically show higher returns than large-cap stocks, but they also tend to have the largest cross-sectional dispersions both in terms of returns and anomaly variables. To mitigate this effect, Hou, Chen, and Zhang (2020) consider a WLS estimator of the premia parameters, with the weights being proportional to the corresponding stock's market capitalization.

[15] See Remark OA.7 to Assumption OA.7.

Formally, let $\left(\hat{\boldsymbol{\Gamma}}^{(\mathrm{w})\prime}_{\mathrm{f},t-1}, \hat{\boldsymbol{\gamma}}^{(\mathrm{w})\prime}_{\mathrm{z},t-1}\right)'$ denote the WLS estimator of the $(K+1)$-vector of premia coefficients, where $\hat{\boldsymbol{\Gamma}}^{(\mathrm{w})}_{\mathrm{f},t-1} = \left(\hat{\gamma}^{(\mathrm{w})\prime}_{0,t-1}, \hat{\boldsymbol{\delta}}^{(\mathrm{w})\prime}_{\mathrm{f},t-1}\right)'$ denotes the WLS premia estimator of the zero-beta rate and the K_{f} risk factors, while $\hat{\boldsymbol{\gamma}}^{(\mathrm{w})}_{\mathrm{z},t-1}$ refers to the WLS premia estimator of the K_{z} anomaly variables. Let $\mathbf{W}_{t-1}$ be an $N \times N$ diagonal matrix containing the asset-specific weights at time $t-1$, i.e. $\mathbf{W}_{t-1} = \mathrm{diag}\left(\mathrm{w}_{1,t-1}, ..., \mathrm{w}_{N,t-1}\right)$, where we assume $\mathrm{w}_{i,t} > 0$ for every asset i and period t without great loss of generality. Then, following Hou, Chen, and Zhang (2020), we have:

$$\begin{bmatrix} \hat{\boldsymbol{\Gamma}}^{(\mathrm{w})}_{\mathrm{f},t-1} \\ \\ \hat{\boldsymbol{\gamma}}^{(\mathrm{w})}_{\mathrm{z},t-1} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{X}}'\mathbf{W}_{t-1}\hat{\mathbf{X}} & \hat{\mathbf{X}}'\mathbf{W}_{t-1}\mathbf{Z}_{t-1} \\ \\ \mathbf{Z}'_{t-1}\mathbf{W}_{t-1}\hat{\mathbf{X}} & \mathbf{Z}'_{t-1}\mathbf{W}_{t-1}\mathbf{Z}_{t-1} \end{bmatrix}^{-1} \begin{bmatrix} \hat{\mathbf{X}}'\mathbf{W}_{t-1}\mathbf{R}_t \\ \\ \mathbf{Z}'_{t-1}\mathbf{W}_{t-1}\mathbf{R}_t \end{bmatrix} \quad \text{comma} \tag{OA.24}$$

where, for each stock i, the weight $\mathrm{w}_{i,t-1}$ in $\mathbf{W}_{t-1}$ is given by the corresponding stock market capitalization at time $t-1$.

Similarly to the (un-weighted) time-varying OLS estimator (OA.15), we can show that analogous conclusions apply to the WLS estimator in (OA.24). Indeed, whenever one wants to estimate time-varying premia under the traditional large-T–fixed-N setting, we show that the estimator in (OA.24) would be invalid, because it is affected by a random (hence, unpredictable) bias.[16]In the large-N–fixed-T set-up, instead, the WLS estimator in (OA.24) is still contaminated by several sources of bias which, however, can be consistently estimated, yielding our novel bias-adjusted CSR WLS estimator:

$$\begin{bmatrix} \hat{\boldsymbol{\Gamma}}^{*(\mathrm{w})}_{\mathrm{f},t-1} \\ \\ \hat{\boldsymbol{\gamma}}^{*(\mathrm{w})}_{\mathrm{z},t-1} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{X}}'\mathbf{W}_{t-1}\hat{\mathbf{X}} - N\hat{\boldsymbol{\Lambda}}^{(\mathrm{w})}_{1,t-1} & \hat{\mathbf{X}}'\mathbf{W}_{t-1}\mathbf{Z}_{t-1} \\ \\ \mathbf{Z}'_{t-1}\mathbf{W}_{t-1}\hat{\mathbf{X}} & \mathbf{Z}'_{t-1}\mathbf{W}_{t-1}\mathbf{Z}_{t-1} \end{bmatrix}^{-1} \begin{bmatrix} \hat{\mathbf{X}}'\mathbf{W}_{t-1}\mathbf{R}_t - N\hat{\boldsymbol{\Lambda}}^{(\mathrm{w})}_{2,t-1} \\ \\ \mathbf{Z}'_{t-1}\mathbf{W}_{t-1}\mathbf{R}_t \end{bmatrix}, \tag{OA.25}$$

where $\hat{\boldsymbol{\Gamma}}^{*(\mathrm{w})}_{t-1} = \left(\hat{\gamma}^{*(\mathrm{w})}_{0,t-1}, \hat{\boldsymbol{\delta}}^{*(\mathrm{w})\prime}_{\mathrm{f},t-1}\right)'$, and where we set

$$\hat{\boldsymbol{\Lambda}}^{(\mathrm{w})}_{1,t-1} = \begin{bmatrix} 0 & \mathbf{0}'_{K_f} \\ \\ \mathbf{0}_{K_f} & \hat{\sigma}^{2(\mathrm{w})}_{t-1}\mathbf{P}'\mathbf{P} \end{bmatrix}, \qquad \hat{\boldsymbol{\Lambda}}^{(\mathrm{w})}_{2,t-1} = \hat{\sigma}^{2(\mathrm{w})}_{t-1} \begin{bmatrix} 0 \\ \\ \mathbf{P}'\boldsymbol{\imath}_{t-1,T-1} \end{bmatrix} \tag{OA.26}$$

with

$$\hat{\sigma}^{2(\mathrm{w})}_{t-1} = \frac{\mathrm{tr}(\hat{\boldsymbol{\epsilon}}\mathbf{W}_{t-1}\hat{\boldsymbol{\epsilon}}')}{N(T-K-2)} \tag{OA.27}$$

[16]Details are available upon request.

Before establishing the asymptotic properties of the WLS estimator in (OA.25), it is important to highlight some necessary remarks. The choice of using the stock's market capitalization in the $\mathbf{W}_{t-1}$ matrix makes the weighting scheme parameter-free. On one hand, this simplifies the WLS analysis, where the weights are typically defined as functions of unknown parameters to be estimated, such as setting them to be inversely proportional to the regression error variance. On the other hand, however, market capitalization will be very likely correlated - both cross-sectionally and over time - with returns and other anomalies, making the asymptotic analysis of the estimator non-trivial. For this reason, we need to impose some conditions on the sample moments of anomalies, weights, and asset-specific errors. Specifically, we assume that each asset-specific error is uncorrelated with past values of both anomaly variables and weights, but could be potentially correlated with their contemporary and future values.[17]

The limiting behaviour of the weights plays a crucial role in determining the statistical properties of the WLS estimator, especially when $N \to \infty$. In particular, a condition that the weights should satisfy is the so-called *granularity* assumption, which guarantees that the weights dissipate to zero sufficiently fast for every asset, as $N \to \infty$. When the granularity assumption fails, then the WLS estimator exhibits a random limit, making both estimation and inference invalid. Therefore, in the following theorems, we establish the limiting properties of the WLS estimator in (OA.25) under the assumption that granularity holds.[18]

Theorem OA.3. *As $N \to \infty$, under Assumptions OA.9, OA.3 –OA.8, and OA.10–OA.13,*

(i)

$$\hat{\mathbf{\Gamma}}^{*(\mathrm{w})}_{\mathrm{f},t-1} - \mathbf{\Gamma}_{\mathrm{f},t-1} = O_p\left(\frac{1}{\sqrt{N}}\right), \qquad \hat{\boldsymbol{\gamma}}^{*(\mathrm{w})}_{\mathrm{z},t-1} - \boldsymbol{\gamma}_{\mathrm{z},t-1} = O_p\left(\frac{1}{\sqrt{N}}\right). \tag{OA.28}$$

(ii)

$$\sqrt{N}\begin{bmatrix} \hat{\mathbf{\Gamma}}^{*(\mathrm{w})}_{\mathrm{f},t-1} - \mathbf{\Gamma}_{\mathrm{f},t-1} \\ \hat{\boldsymbol{\gamma}}^{*(\mathrm{w})}_{\mathrm{z},t-1} - \boldsymbol{\gamma}_{\mathrm{z},t-1} \end{bmatrix} \to_d \mathcal{N}\left(\mathbf{0}_{\mathrm{K}+1}, \mathbf{L}^{-1}_{t-1}\mathbf{O}^{(\mathrm{w})}_{t-1}\mathbf{L}^{-1}_{t-1}\right), \tag{OA.29}$$

where $\mathbf{L}_{t-1} > 0$ is the same as in Theorem 1, and for for some $\mathbf{O}^{(\mathrm{w})}_{t-1}$.[19]

Proof. See Section OA.5.

[17] See Assumption OA.13 in Section OA.3.2.

[18] See Assumption OA.10 and Section OA.8.

[19] The definition of $\mathbf{O}^{(\mathrm{w})}_{t-1}$ has been relegated to the proof of the theorem (see (OA.93)).

The next theorem shows how to construct asymptotically-valid standard errors for the WLS estimator.

Theorem OA.4. *As $N \to \infty$, under Assumptions OA.9, OA.3 –OA.8, OA.10–OA.13, and the identification condition $\kappa_4 = 0$,*

$$\hat{\mathbf{L}}_{t-1}^{(\mathrm{w})-1}\,\hat{\mathbf{O}}_{t-1}^{(\mathrm{w})}\,\hat{\mathbf{L}}_{t-1}^{(\mathrm{w})-1\prime} \quad \to_p \quad \mathbf{L}_{t-1}^{-1}\,\mathbf{O}_{t-1}^{(\mathrm{w})}\,\mathbf{L}_{t-1}^{-1\prime} \qquad \text{comma} \tag{OA.30}$$

where

$$\begin{aligned} \hat{\mathbf{L}}_{t-1}^{(\mathrm{w})} &= \tfrac{1}{N}\begin{bmatrix} \hat{\mathbf{X}}'\mathbf{W}_{t-1}\hat{\mathbf{X}} - N\hat{\boldsymbol{\Lambda}}_{1,t-1}^{(\mathrm{w})} & \hat{\mathbf{X}}'\mathbf{W}_{t-1}\mathbf{Z}_{t-1} \\ \mathbf{Z}_{t-1}'\mathbf{W}_{t-1}\hat{\mathbf{X}} & \mathbf{Z}_{t-1}'\mathbf{W}_{t-1}\mathbf{Z}_{t-1} \end{bmatrix}, \\ \text{and} \quad \hat{\mathbf{O}}_{t-1}^{(\mathrm{w})} &= \hat{\lambda}_{t-1}\begin{bmatrix} \hat{\boldsymbol{\mu}}_{\mathrm{x}}\hat{\boldsymbol{\mu}}_{\mathrm{x}}' & \hat{\boldsymbol{\mu}}_{\mathrm{x}}\hat{\boldsymbol{\mu}}_{\mathrm{z}}' \\ \hat{\boldsymbol{\mu}}_{\mathrm{z}}\hat{\boldsymbol{\mu}}_{\mathrm{x}}' & \hat{\boldsymbol{\mu}}_{\mathrm{z}}\hat{\boldsymbol{\mu}}_{\mathrm{z}}' \end{bmatrix} + \hat{\mathbf{M}}_{t-1}^{(\mathrm{w})} \qquad \text{comma} \end{aligned} \tag{OA.31}$$

with

$$\hat{\mathbf{M}}_{t-1}^{(\mathrm{w})} = \begin{bmatrix} 0 & \mathbf{0}_{K_{\mathrm{f}}}' & \mathbf{0}_{K_{\mathrm{z}}}' \\ \mathbf{0}_{K_{\mathrm{f}}} & \hat{\mu}_{\mathrm{w},t-1}^{2}\,\hat{\mathbf{V}}_{t-1}^{(\mathrm{w})\prime}\hat{\mathbf{U}}_{\epsilon}\hat{\mathbf{V}}_{t-1}^{(\mathrm{w})} & \mathbf{0}_{K_{\mathrm{f}}\times K_{\mathrm{z}}} \\ \mathbf{0}_{K_{\mathrm{z}}} & \mathbf{0}_{K_{\mathrm{z}}\times K_{\mathrm{f}}} & \hat{\mathbf{H}}_{t-1}^{(\mathrm{w})}\hat{\boldsymbol{\Sigma}}_{\mathrm{U}}^{(\mathrm{w})}\hat{\mathbf{H}}_{t-1}^{(\mathrm{w})\prime} + \hat{\mathbf{S}}_{t-1}^{(\mathrm{w})} + \hat{\mathbf{S}}_{t-1}^{(\mathrm{w})} \end{bmatrix} \quad \text{comma} \tag{OA.32}$$

setting $\hat{\boldsymbol{\mu}}_{\mathrm{x}} = (1, \hat{\boldsymbol{\mu}}_{\beta}')'$, $\hat{\boldsymbol{\mu}}_{\beta} = N^{-1}\hat{\mathbf{B}}\mathbf{1}_N$, $\hat{\boldsymbol{\mu}}_{\mathrm{z}} = N^{-1}\mathbb{J}'\mathbf{Z}'\mathbf{1}_N$, $\hat{\mu}_{\mathrm{w},t-1}^2 = N^{-1}\mathbf{1}_N'\mathbf{W}_{t-1}^2\mathbf{1}_N$, $\hat{\boldsymbol{\Sigma}}_{\mathrm{U}}^{(\mathrm{w})} = (\hat{\sigma}_{t-1}^{2(\mathrm{w})}\mathbf{I}_{T-1} \otimes N^{-1}\mathbf{Z}'\mathbf{Z})$, $\hat{\boldsymbol{\Sigma}}_{\mathrm{ZW}} = N^{-1}\mathbf{Z}'\mathbf{W}$, *and* $\hat{\boldsymbol{\Sigma}}_{\mathrm{V}} = \left(\hat{\sigma}_{t-1}^{2(\mathrm{w})}\mathbf{I}_{T-1} \otimes N^{-1}\sum_{i=1}^{N}\mathbf{w}_i\mathbf{w}_i'\right)$, *where* $\mathbf{w}_i = (\mathrm{w}_{i,1}, ..., \mathrm{w}_{i,T-1})'$, *with* $\hat{\boldsymbol{\Lambda}}_{1,t-1}^{(\mathrm{w})}$ *and* $\hat{\sigma}_{t-1}^{2(\mathrm{w})}$ *defined in* (OA.26) *and* (OA.27), *respectively, and we define the following matrices, where* $\hat{\mathbf{U}}_{\epsilon}$ *is defined in Theorem 2,*

$$\begin{aligned} \hat{\mathbf{Q}}_{t-1}^{(\mathrm{w})} &= \boldsymbol{\imath}_{t-1,T-1} - \mathbf{P}\hat{\boldsymbol{\delta}}_{\mathrm{f},t-1}^{*(\mathrm{w})}, \quad \hat{\mathbf{H}}_{t-1}^{(\mathrm{w})} = \hat{\mathbf{Q}}_{t-1}^{(\mathrm{w})\prime} \otimes \mathbb{J}_{t-1}', \\ \hat{\mathbf{Y}}_{t-1} &= \hat{\mathbf{Q}}_{t-1}^{(\mathrm{w})} \otimes \boldsymbol{\imath}_{t-1,T-1}, \quad \hat{\lambda}_{t-1} = \hat{\mathbf{Y}}_{t-1}'\hat{\boldsymbol{\Sigma}}_{\mathrm{V}}\hat{\mathbf{Y}}_{t-1}, \\ \hat{\mathbf{S}}_{t-1} &= \hat{\boldsymbol{\mu}}_{\mathrm{z}}\hat{\mathbf{Y}}_{t-1}'(\hat{\sigma}_{t-1}^{2(\mathrm{w})}\boldsymbol{I}_{T-1} \otimes \hat{\boldsymbol{\Sigma}}_{\mathrm{ZW}}')\hat{\mathbf{H}}_{t-1}^{(\mathrm{w})\prime}, \\ \hat{\mathbf{V}}_{t-1}^{(\mathrm{w})} &= (\hat{\mathbf{Q}}_{t-1}^{(\mathrm{w})} \otimes \mathbf{P}) - \left(\frac{\mathrm{vec}(\overline{\mathbf{M}}_{\tilde{\mathrm{D}}})}{T-K-2}\right)\hat{\mathbf{Q}}_{t-1}^{(\mathrm{w})\prime}\mathbf{P}. \end{aligned}$$

Proof. Omitted, as it closely parallels the proof of Theorem OA.2. See Section OA.5.

OA.2.3 Global Misspecification

All the results so far established assume that the asset pricing model (OA.8) is *correctly specified*, meaning that the true model does not omit any relevant variable (either a risk factor or an anomaly variable) or it does not include any irrelevant one.[20] When this assumption is violated - a very likely scenario in practice - then the issue of *global misspecification* arises, which, if ignored, could seriously compromise our inferential results.[21] Indeed, misspecification affects the standard errors obtained in the previous sections, with the risk of making an anomaly appear significant when instead its premium is null or making it statistically irrelevant when instead its effect is non-zero. Therefore, the objective of this section is to extend our methodology and robustify our inferential results to the case of a generic deviation from exact pricing of unknown form, i.e., global misspecification.

Consider the asset pricing restriction in (OA.5) and assume now that, beyond the presence of anomalies, there is a further deviation from exact pricing, due to potential global misspecification. That is, assume that:

$$\mathrm{E}[R_{i,t}|I_{t-1},\mathbf{\Pi}] = \tilde{\gamma}_{0,t-1} + \tilde{\boldsymbol{\delta}}'_{\mathrm{f},t-1}\boldsymbol{\beta}_i + \tilde{\boldsymbol{\gamma}}'_{z,t-1}\mathbf{z}_{i,t-1} + \mathrm{m}_{i,t-1}, \tag{OA.33}$$

where $\mathrm{m}_{i,t-1}$ represents an additional pricing error, accounting for the fact that the postulated model could potentially specify the wrong set of variables. In other words, one could think that the overall deviation from exact pricing in (OA.2) has now a semiparametric structure, with the parametric part being linear in the $\mathbf{z}_{i,t-1}$, and a non-parametric component coming from the misspecification error $\mathrm{m}_{i,t-1}$, which is completely unspecified. Our objective is to test for the statistical relevance of the anomalies $\mathbf{z}_{i,t-1}$, *regardless* of whether they represent or not the full set of variables describing the true asset pricing model, that is regardless of whether $\mathrm{m}_{i,t-1}$ is zero or not.

[20] A different form of misspecification, not explored in this paper, occurs when one (or more) vector of betas is a linear combination of the other ones, implying that $\mathbf{X}$ is not full-column rank. This happens, for example, when one or more of the candidate risk factors have zero (or almost zero) betas, a situation which is often referred to as the issue of *spurious* or *useless* factors. See, e.g., Jagannathan and Wang (1998), Kan and Zhang (1999a,b), Kleibergen (2009), Gospodinov, Kan, and Robotti (2014), Bryzgalova (2014), Burnside (2016), Ahn, Horenstein, and Wang (2018), Kleibergen and Zhan (2014, 2020), and Anatolyev and Mikusheva (2022), among others. The less restrictive cases of semi-strong, when $\mathbf{B}'\mathbf{B}/N = o(1)$ (see Connor and Korajczyk (2024)), and weak factors, when $\mathbf{B}'\mathbf{B}/N = O(1/N)$ (see Lettau and Pelger (2020), Giglio, Xiu, and Zhang (2025), and Kim, Raponi, and Zaffaroni (2023)), are also ruled out by our assumptions. Kim, Raponi, and Zaffaroni (2023) develop an inferential procedure to test for spurious and weak factors, valid when N is large and T is fixed. update date

[21] Global misspecification has been studied widely in the large-T sampling scheme; see Jagannathan and Wang (1998), Shanken and Zhou (2007), Hou and Kimmel (2006), and Kan, Robotti, and Shanken (2013), among others. Gagliardini, Ossola, and Scaillet (2016) and Raponi, Robotti, and Zaffaroni (2020) show how to robustify their risk premia estimator to global misspecification in the large-N–large-T and in the large-N–fixed-T settings, respectively.

The relationship in (OA.33) implies that the parameters $\tilde{\boldsymbol{\Gamma}}_{\mathrm{f},t-1}$, and $\tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1}$, represent the so-called *pseudo*-true values of the premia coefficients. Formally, let $\mathbf{c}_{\mathrm{z}}$ and $\mathbf{c}_{\mathrm{f}}$ denote two arbitrary vectors of dimension K_{z} and $K_{\mathrm{f}}+1$, respectively. Then, by generalizing Shanken and Zhou (2007) and Raponi, Robotti, and Zaffaroni (2020), we define the pseudo-true premia parameters

$$(\tilde{\boldsymbol{\Gamma}}_{\mathrm{f},t-1}', \tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1}')' = \underset{\mathbf{c}_{\mathrm{f}},\mathbf{c}_{\mathrm{z}}}{\operatorname{argmin}} \frac{1}{N}\Big(\mathrm{E}[\mathbf{R}_t|\mathbf{I}_{t-1},\boldsymbol{\Pi}] - \mathbf{Z}_{t-1}\mathbf{c}_{\mathrm{z}} - \mathbf{X}\mathbf{c}_{\mathrm{f}}\Big)'\Big(\mathrm{E}[\mathbf{R}_t|\mathbf{I}_{t-1},\boldsymbol{\Pi}] - \mathbf{Z}_{t-1}\mathbf{c}_{\mathrm{z}} + \mathbf{X}\mathbf{c}_{\mathrm{f}}\Big), \tag{OA.34}$$

When the model is correctly specified, then $\tilde{\boldsymbol{\Gamma}}_{\mathrm{f},t-1} = \boldsymbol{\Gamma}_{\mathrm{f},t-1}$ and $\tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1} = \boldsymbol{\gamma}_{t-1,z}$, that is, we recover the true risk and anomalies' premia. However, the cross-sectional correlation between $\mathbf{m}_t$ and $\boldsymbol{\epsilon}_s$, arising as a result of global misspecification, induces further biases to the CSR OLS estimator, which nevertheless can be consistently estimated, leading to our novel misspecification-robust premia estimators $\hat{\boldsymbol{\Gamma}}^{*(\mathrm{m})}_{\mathrm{f},t-1} = (\hat{\gamma}^{*(\mathrm{m})}_{0,t-1}, \hat{\boldsymbol{\delta}}^{*(\mathrm{m})\prime}_{\mathrm{f},t-1})'$ and $\hat{\boldsymbol{\gamma}}^{*(\mathrm{m})}_{\mathrm{z},t-1}$, defined as follows.

$$\begin{bmatrix} \hat{\boldsymbol{\Gamma}}^{*(\mathrm{m})}_{\mathrm{f},t-1} \\ \hat{\boldsymbol{\gamma}}^{*(\mathrm{m})}_{\mathrm{z},t-1} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{X}}'\hat{\mathbf{X}} - N(\hat{\boldsymbol{\Lambda}}_1 + \hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{1,t-1}) & \hat{\mathbf{X}}'\mathbf{Z}_{t-1} - N\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{3,t-1} \\ \mathbf{Z}_{t-1}'\hat{\mathbf{X}} & \mathbf{Z}_{t-1}'\mathbf{Z}_{t-1} \end{bmatrix}^{-1} \begin{bmatrix} \hat{\mathbf{X}}'\mathbf{R}_t - N(\hat{\boldsymbol{\Lambda}}_{2,t-1} + \hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{2,t-1}) \\ \mathbf{Z}_{t-1}'\mathbf{R}_t \end{bmatrix}, \tag{OA.35}$$

setting $\hat{\boldsymbol{\Lambda}}_1$ and $\hat{\boldsymbol{\Lambda}}_{2,t-1}$ are defined in (OA.16), and where we define the additional bias-correction terms

$$\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{1,t-1} = \frac{1}{N}\begin{bmatrix} \mathbf{0}_{K_{\mathrm{f}}+1}' \\ \mathbf{P}'\hat{\boldsymbol{\Psi}}_{\mathrm{D}\hat{\mathrm{X}}} \end{bmatrix}, \ \hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{2,t-1} = \frac{1}{N}\begin{bmatrix} 0 \\ \mathbf{P}'\hat{\boldsymbol{\Psi}}_{\mathrm{DR}} - \hat{\sigma}^2\mathbf{P}'\hat{\boldsymbol{\Psi}}_{\mathrm{D}\tilde{\mathrm{D}}} \end{bmatrix}, \text{ and } \hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{3,t-1} = \frac{1}{N}\begin{bmatrix} \mathbf{0}_{K_z}' \\ \mathbf{P}'\hat{\boldsymbol{\Psi}}_{\mathrm{DZ}} \end{bmatrix}, \tag{OA.36}$$

with $\hat{\boldsymbol{\Psi}}_{\mathrm{D}\hat{\mathrm{X}}} = \begin{bmatrix} \mathbf{M}^{(-1)}_{\mathrm{D},t-1}\hat{\boldsymbol{\epsilon}}\hat{\mathbf{X}} \\ \mathbf{0}_{(T-t+1)\times(K_{\mathrm{f}}+1)} \end{bmatrix}$, $\hat{\boldsymbol{\Psi}}_{\mathrm{DZ}} = \begin{bmatrix} \mathbf{M}^{(-1)}_{\mathrm{D},t-1}\hat{\boldsymbol{\epsilon}}\mathbf{Z}_{t-1} \\ \mathbf{0}_{(T-t+1)\times K_{\mathrm{z}}} \end{bmatrix}$, $\hat{\boldsymbol{\Psi}}_{\mathrm{DR}} = \begin{bmatrix} \mathbf{M}^{(-1)}_{\mathrm{D},t-1}\hat{\boldsymbol{\epsilon}}\mathbf{R}_t \\ \mathbf{0}_{T-t+1} \end{bmatrix}$, and

$\hat{\boldsymbol{\Psi}}_{\mathrm{D}\tilde{\mathrm{D}}} = \begin{bmatrix} \mathbf{M}^{(-1)}_{\mathrm{D},t-1}\mathbb{M}_{\mathrm{D}}\boldsymbol{\imath}_{t-1,T-1} \\ \mathbf{0}_{T-t+1} \end{bmatrix}$, setting the $(t-2)\times(T-1)$ matrix $\mathbf{M}^{(-1)}_{\mathrm{D},t-1} = \mathbb{M}_{11}^{-1}[\mathbf{I}_{t-2}, \mathbf{0}_{(t-2)\times(T-t+1)}]$,

where $\mathbb{M}_{11}$ denotes the $(t-2)\times(t-2)$ top-left block of $\mathbb{M}_D = \mathbf{I}_{T-1} - \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'$.[22]

Some additional regularity assumptions are needed, as follows.

[22] We use the partition $\mathbb{M}_D = \begin{bmatrix} \mathbb{M}_{11} & \mathbb{M}_{12} \\ \mathbb{M}_{21} & \mathbb{M}_{22} \end{bmatrix}$.

Assumption OA.2. *(mixed-moments of pricing errors). We assume:*

(i) For every $2 \leq s, t \leq T$, let $\boldsymbol{\theta}_{t-1,m} = (\theta_{t-3,m}, \theta_{t-4,m}, \ldots, \theta_{0,m})'$, defined as $\frac{1}{N}\sum_{i=1}^{N} \epsilon_{i,s} m_{i,t-1} \rightarrow_p \theta_{t-1-s,m}$, such that $\theta_{u,m} = 0$ for $u < 0$. Then,

$$\frac{1}{N}\boldsymbol{\epsilon}\mathbf{m}_{t-1} \rightarrow_p \begin{bmatrix} \boldsymbol{\theta}_{t-1,m} \\ \mathbf{0}_{T-t+1} \end{bmatrix},$$

(ii) For every $2 \leq s, t \leq T$,

$$\frac{1}{N}\mathbf{m}'_{t-1}\mathbf{m}_{t-1} \rightarrow_p \sigma_{t-1mm}.$$

(iii) For every $2 \leq s, t \leq T$,

$$\frac{1}{N}\sum_{i=1}^{N} \mathbf{P}_{\tilde{D}_i} \boldsymbol{\epsilon}_i m_{i,t-1} \rightarrow_p \mathbf{0}_{T-1}.$$

Remark OA.1. *To better understand Assumption OA.2-(i), note that although we do not impose any parameterization on* $\mathrm{m}_{i,t-1}$*, simple considerations suggest that* $\mathrm{m}_{i,t-1}$ *is cross-sectionally correlated with* $\epsilon_{i,s}$ *for every* $s \leq t$*. As an illustrative example, consider the case where some relevant risk factors and anomaly variables are omitted from the true model* (OA.6)*, and assume that no other sources of misspecification are present. In that case, the asset-pricing model can be written as*

$$\begin{aligned} \mathbf{R}_t &= \mathbf{Z}_{t-1}\tilde{\gamma}_{\mathrm{z},t-1} + \mathbf{X}\tilde{\boldsymbol{\Gamma}}_{\mathrm{f},t-1} + \boldsymbol{\epsilon}_t, \quad \textit{with} \\ \boldsymbol{\epsilon}_t &= \breve{\boldsymbol{\epsilon}}_t + \breve{\mathbf{Z}}_{t-1}\breve{\gamma}_{\mathrm{z},t-1} + \breve{\mathbf{B}}\breve{\mathbf{f}}_t, \end{aligned} \tag{OA.37}$$

where $\tilde{\boldsymbol{\Gamma}}_{\mathrm{f},t-1} = (\tilde{\gamma}_{0,t-1}, \tilde{\boldsymbol{\delta}}'_{\mathrm{f},t-1})'$*,* $\breve{\mathbf{Z}}_{t-1}$ *denotes the* $N \times \breve{K}_{\mathrm{z}}$ *matrix of excluded anomalies with corresponding premia* $\breve{\gamma}_{\mathrm{z},t-1}$*, and* $\breve{\mathbf{B}}$ *is the* $N \times \breve{K}_{\mathrm{f}}$ *matrix of exposures associated with the* $\breve{K}_{\mathrm{f}}$ *omitted factors* $\breve{\mathbf{f}}_t$*, with ex-post risk premia* $\breve{\boldsymbol{\delta}}_{\mathrm{f},t-1}$*. Here* $\breve{\boldsymbol{\epsilon}}_t$ *represents the* pure *asset-specific component of asset returns, coinciding with* $\boldsymbol{\epsilon}_t$ *in the case of correct model specification.*

Combining (OA.33) *with* (OA.37)*, one obtains*

$$\mathbf{m}_{t-1} = (m_{1,t-1}, \ldots, m_{N,t-1})' = \breve{\mathbf{Z}}_{t-1}\breve{\gamma}_{\mathrm{z},t-1} + \breve{\mathbf{B}}\breve{\boldsymbol{\delta}}_{\mathrm{f},t-1},$$

which implies that $\mathbf{m}_t$ *and* $\boldsymbol{\epsilon}_s$ *are potentially correlated (both across assets and across time) through* $\breve{\mathbf{Z}}_t$*,* $\breve{\mathbf{B}}$*, or both, whenever* $s \leq t$*. In contrast,* $\mathbf{m}_t$ *and* $\boldsymbol{\epsilon}_s$ *are uncorrelated whenever* $s > t$*, which follows immediately from the temporal i.i.d. assumption on* $\epsilon_{i,t}$*.*

Remark OA.2. *Assumption (OA.2)-(ii) rules out explosive limiting behavior of the average squared pricing errors. Assumption (OA.2)-(iii) simplifies the formulae and reinforces the idea that the risk exposures to the omitted risk factors and anomalies are cross-sectionally unrelated to the anomalies and the risk exposures of the factors in the candidate model.*

In the following, we first establish the additional biases induced by global misspecification for the CSR OLS estimator. We then construct the misspecification-robust, bias-adjusted estimator of the premia in (OA.35) and derive its limiting properties. Finally, we show how to construct asymptotically valid standard errors.

Proposition OA.1 (biases of CSR OLS — time-varying estimator with misspecification)**.** *Under Assumptions OA.9–OA.8 and OA.2, as $N \to \infty$,*

$$\begin{bmatrix} \hat{\mathbf{\Gamma}}_{t-1} \\ \hat{\gamma}_{t-1,z} \end{bmatrix} \to_p \begin{bmatrix} \tilde{\mathbf{\Gamma}}_{t-1} \\ \tilde{\gamma}_{t-1,z} \end{bmatrix} + \mathbf{C}_{t-1}^{-1} \left(- \begin{bmatrix} \mathbf{\Lambda}_1 \\ \mathbf{0}_{K_z \times (K_f+1)} \end{bmatrix} \tilde{\mathbf{\Gamma}}_{t-1} + \begin{bmatrix} \mathbf{\Lambda}_{t-1,2} \\ \mathbf{0}_{K_z} \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{P}' \begin{bmatrix} \boldsymbol{\theta}_{t-1,m} \\ \mathbf{0}_{T-t+1} \end{bmatrix} \\ \mathbf{0}_{K_z} \end{bmatrix} \right), \quad \text{(OA.38)}$$

where $\mathbf{C}_{t-1}$, $\mathbf{\Lambda}_1$, and $\mathbf{\Lambda}_{t-1,2}$ are defined as

$$\mathbf{\Lambda}_1 = \begin{bmatrix} 0 & \mathbf{0}'_{K_f} \\ \mathbf{0}_{K_f} & \sigma^2 \mathbf{P}'\mathbf{P} \end{bmatrix}, \qquad \mathbf{\Lambda}_{2,t-1} = \sigma^2 \begin{bmatrix} 0 \\ \mathbf{P}'\boldsymbol{\imath}_{t-1,T-1} \end{bmatrix}, \quad \text{(OA.39)}$$

and

$$\mathbf{C}_{t-1} = \begin{bmatrix} \mathbf{\Sigma}_X + \mathbf{\Lambda}_1 & \mathbf{\Sigma}'_{ZX,t-1} \\ \mathbf{\Sigma}_{ZX,t-1} & \mathbb{J}'_{t-1}\mathbf{\Sigma}_Z\mathbb{J}_{t-1} \end{bmatrix}. \quad \text{(OA.40)}$$

Here $\boldsymbol{\theta}_{t-1,m}$ is defined in Assumption OA.2.

Proof. For part (i), one needs to consider the additional term

$$\frac{1}{N} \begin{bmatrix} \hat{\mathbf{X}}' \\ \mathbf{Z}'_{t-1} \end{bmatrix} \mathbf{m}_{t-1} = \frac{1}{N} \begin{bmatrix} 0 \\ (\hat{\mathbf{B}} - \mathbf{B})'\mathbf{m}_{t-1} \\ \mathbf{0}_{K_z} \end{bmatrix},$$

where

$$\frac{1}{N}(\hat{\mathbf{B}} - \mathbf{B})'\mathbf{m}_{t-1} = \frac{1}{N}\mathbf{P}'\boldsymbol{\epsilon}\mathbf{m}_{t-1} \to \mathbf{P}' \begin{bmatrix} \theta_{t-3,m} \\ \theta_{t-4,m} \\ \vdots \\ \theta_{0,m} \\ \mathbf{0}_{T-t+1} \end{bmatrix} = \mathbf{P}' \begin{bmatrix} \boldsymbol{\theta}_{t-1,m} \\ \mathbf{0}_{T-t+1} \end{bmatrix},$$

using the property $(\mathbf{X}, \mathbf{Z}_{t-1})'\mathbf{m}_{t-1} = \mathbf{0}_{K_z+K_f+1}$. For part (ii), the result follows since $\hat{\mathbf{X}} \to_p \mathbf{X}$ as $T \to \infty$, under the assumed conditions $\mathbf{P}'\boldsymbol{\epsilon} \to_p \mathbf{0}_{K_f \times N}$ and $(\mathbf{X}, \mathbf{Z}_{t-1})$ having full column rank.

■

To construct the bias-adjusted estimator, we proceed as follows.

A natural estimator for $\boldsymbol{\theta}_{t-1,m}$ is

$$\hat{\boldsymbol{\theta}}_{t-1,m} = \frac{1}{N}\hat{\boldsymbol{\epsilon}}'_{2,t-1}\hat{\mathbf{m}}^{prelim}_{t-1},$$

with

$$\hat{\mathbf{m}}^{prelim}_{t-1} = \mathbf{R}_t - (\hat{\mathbf{X}}, \mathbf{Z}_{t-1})\begin{bmatrix}\hat{\boldsymbol{\Gamma}}^{prelim\prime}_{t-1} \\ \hat{\gamma}^{prelim\prime}_{t-1z}\end{bmatrix},$$

setting

$$\boldsymbol{\epsilon} = \begin{bmatrix}\boldsymbol{\epsilon}_{2,t-1} \\ \boldsymbol{\epsilon}_{t,T-1}\end{bmatrix}, \qquad \boldsymbol{\epsilon}_{2,t-1} = (\boldsymbol{\epsilon}_2, \ldots, \boldsymbol{\epsilon}_{t-1})', \qquad \boldsymbol{\epsilon}_{t,T-1} = (\boldsymbol{\epsilon}_t, \ldots, \boldsymbol{\epsilon}_{T-1})',$$

and *assuming* the availability of preliminary consistent estimators for the premia parameters $(\hat{\boldsymbol{\Gamma}}^{prelim\prime}_{t-1}, \hat{\gamma}^{prelim\prime}_{t-1z})$.

Notice that $\boldsymbol{\epsilon}_{t,T-1}$ and $\mathbf{m}_{t-1}$ are mutually independent, and hence their covariance is zero. Lemma 8 shows that $N^{-1}\hat{\boldsymbol{\epsilon}}'_{2,t-1}\hat{\mathbf{m}}^{prelim}_{t-1}$ is biased, but that by adding a suitable bias-correction term one can construct a valid estimator for $\boldsymbol{\theta}_{t-1,m}$. This, in turn, yields the overall bias-adjustment term for the CSR OLS estimator, as follows:

$$\begin{bmatrix} 0 \\ \mathbf{P}'\begin{bmatrix}\hat{\boldsymbol{\theta}}^{prelim}_{t-1,m} \\ \mathbf{0}_{T-t+1}\end{bmatrix} \\ \mathbf{0}_{K_z}\end{bmatrix} = \begin{bmatrix}\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{2,t-1} \\ \mathbf{0}_{K_z}\end{bmatrix} - \begin{bmatrix}\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{1,t-1} \\ \mathbf{0}_{K_z \times K_f+1}\end{bmatrix}\hat{\boldsymbol{\Gamma}}^{prelim}_{t-1} - \begin{bmatrix}\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{3,t-1} \\ \mathbf{0}_{K_z \times K_z}\end{bmatrix}\hat{\gamma}^{prelim}_{t-1z},$$

setting

$$\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{1,t-1} = \frac{1}{N}\begin{bmatrix}\mathbf{0}'_{K_\mathrm{f}+1} \\ \mathbf{P}'\hat{\boldsymbol{\Psi}}_{\mathrm{D}\hat{\mathrm{X}}}\end{bmatrix}, \;\; \hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{2,t-1} = \frac{1}{N}\begin{bmatrix}0 \\ \mathbf{P}'\hat{\boldsymbol{\Psi}}_{\mathrm{DR}} - \hat{\sigma}^2\mathbf{P}'\hat{\boldsymbol{\Psi}}_{\mathrm{D}\tilde{\mathrm{D}}}\end{bmatrix}, \text{ and } \hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{3,t-1} = \frac{1}{N}\begin{bmatrix}\mathbf{0}'_{K_z} \\ \mathbf{P}'\hat{\boldsymbol{\Psi}}_{\mathrm{DZ}}\end{bmatrix},$$

with $\hat{\boldsymbol{\Psi}}_{\mathrm{D}\hat{\mathrm{X}}} \equiv \begin{bmatrix}\mathbf{M}^{(-1)}_{\mathrm{D},t-1}\hat{\boldsymbol{\epsilon}}\hat{\mathbf{X}} \\ \mathbf{0}_{(T-t+1)\times(K_\mathrm{f}+1)}\end{bmatrix}$, $\hat{\boldsymbol{\Psi}}_{\mathrm{DZ}} \equiv \begin{bmatrix}\mathbf{M}^{(-1)}_{\mathrm{D},t-1}\hat{\boldsymbol{\epsilon}}\mathbf{Z}_{t-1} \\ \mathbf{0}_{(T-t+1)\times K_z}\end{bmatrix}$, $\hat{\boldsymbol{\Psi}}_{\mathrm{DR}} \equiv \begin{bmatrix}\mathbf{M}^{(-1)}_{\mathrm{D},t-1}\hat{\boldsymbol{\epsilon}}\mathbf{R}_t \\ \mathbf{0}_{T-t+1}\end{bmatrix}$, and

$\hat{\boldsymbol{\Psi}}_{\mathrm{D}\tilde{\mathrm{D}}} \equiv \begin{bmatrix}\mathbf{M}^{(-1)}_{\mathrm{D},t-1}\mathbb{M}_{\mathrm{D}}\boldsymbol{\imath}_{t-1,T-1} \\ \mathbf{0}_{T-t+1}\end{bmatrix}$, setting the $(t-2)\times(T-1)$ matrix $\mathbf{M}^{(-1)}_{\mathrm{D},t-1} = \mathbb{M}^{-1}_{11}[\mathbf{I}_{t-2}, \mathbf{0}_{(t-2)\times(T-t+1)}]$, and where $\mathbb{M}_{11}$ denotes the $(t-2)\times(t-2)$ top-left block of $\mathbb{M}_D = \mathbf{I}_{T-1} - \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'$, where we use the partition $\mathbb{M}_D = \begin{bmatrix} \mathbb{M}_{11} & \mathbb{M}_{12} \\ \mathbb{M}_{21} & \mathbb{M}_{22} \end{bmatrix}$.

Estimator (OA.35) is then obtained as the 'fixed-point' solution to the following system of equations:

$$\begin{bmatrix} \hat{\boldsymbol{\Gamma}}^{*(\mathrm{m})}_{\mathrm{f},t-1} \\ \hat{\gamma}^{*(\mathrm{m})}_{\mathrm{z},t-1} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\Gamma}}_{t-1} \\ \hat{\gamma}_{t-1,z} \end{bmatrix} - \hat{\mathbf{C}}^{-1}_{t-1} \left(- \begin{bmatrix} \hat{\boldsymbol{\Lambda}}_1 \\ \mathbf{0}_{K_z \times K_f+1} \end{bmatrix} \hat{\boldsymbol{\Gamma}}^{*(\mathrm{m})}_{\mathrm{f},t-1} + \begin{bmatrix} \hat{\boldsymbol{\Lambda}}_{t-1,2} \\ \mathbf{0}_{K_z} \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{P}' \begin{bmatrix} \hat{\boldsymbol{\theta}}^{*}_{t-1,m} \\ \mathbf{0}_{T-t+1} \end{bmatrix} \\ \mathbf{0}_{K_z} \end{bmatrix} \right),$$

that is, setting

$$\begin{bmatrix} 0 \\ \mathbf{P}' \begin{bmatrix} \hat{\boldsymbol{\theta}}^{*}_{t-1,m} \\ \mathbf{0}_{T-t+1} \end{bmatrix} \\ \mathbf{0}_{K_z} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{2,t-1} \\ \mathbf{0}_{K_z} \end{bmatrix} - \begin{bmatrix} \hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{1,t-1} \\ \mathbf{0}_{K_z \times K_f+1} \end{bmatrix} \hat{\boldsymbol{\Gamma}}^{*(\mathrm{m})}_{\mathrm{f},t-1} - \begin{bmatrix} \hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{3,t-1} \\ \mathbf{0}_{K_z \times K_z} \end{bmatrix} \hat{\gamma}^{*(\mathrm{m})}_{\mathrm{z},t-1},$$

obtained replacing $\hat{\boldsymbol{\Gamma}}^{prelim}_{\mathrm{f},t-1}, \hat{\gamma}^{prelim}_{\mathrm{z},t-1}$ with $\hat{\boldsymbol{\Gamma}}^{*(\mathrm{m})}_{\mathrm{f},t-1}, \hat{\gamma}^{*(\mathrm{m})}_{\mathrm{z},t-1}$ into $\hat{\boldsymbol{\theta}}^{prelim}_{t-1,m}$ above.

The following theorem derives the asymptotic properties of our robust estimator, extending the results of Theorem OA.1.

Theorem OA.5. *As $N \to \infty$, under Assumptions OA.9, OA.3 –OA.8, and OA.2, then*

(i)

$$\hat{\boldsymbol{\Gamma}}^{*(\mathrm{m})}_{\mathrm{f},t-1} - \tilde{\boldsymbol{\Gamma}}_{\mathrm{f},t-1} = O_p\left(\frac{1}{\sqrt{N}}\right), \quad \hat{\gamma}^{*(\mathrm{m})}_{\mathrm{z},t-1} - \tilde{\gamma}_{\mathrm{z},t-1} = O_p\left(\frac{1}{\sqrt{N}}\right). \tag{OA.41}$$

(ii)

$$\sqrt{N} \begin{bmatrix} \hat{\boldsymbol{\Gamma}}^{*(\mathrm{m})}_{\mathrm{f},t-1} - \tilde{\boldsymbol{\Gamma}}_{\mathrm{f},t-1} \\ \hat{\gamma}^{*(\mathrm{m})}_{\mathrm{z},t-1} - \tilde{\gamma}_{\mathrm{z},t-1} \end{bmatrix} \to_d \mathcal{N}\left(\mathbf{0}_{\mathrm{K}+1}, \mathbf{L}^{-1}_{t-1}\, \mathbf{O}^{(\mathrm{m})}_{t-1}\, \mathbf{L}_{t-1}\right) \tag{OA.42}$$

with $\mathbf{O}^{(\mathrm{m})}_{t-1} = \mathbf{O}_{t-1} + \boldsymbol{\Omega}^{(\mathrm{m})}_{t-1}$, for some $\boldsymbol{\Omega}^{(\mathrm{m})}_{t-1}$, and with $\mathbf{L}_{t-1} > 0$ and $\mathbf{O}_{t-1}$ being the same as in Theorem OA.1.[23]

Proof. See Section OA.5

In the following theorem we show that $\hat{\mathbf{L}}^{(\mathrm{m})}_{t-1} \to_p \mathbf{L}_{t-1}$ and $\hat{\boldsymbol{\Omega}}^{(\mathrm{m})}_{t-1} \to_p \boldsymbol{\Omega}^{(\mathrm{m})}_{t-1}$ as $N \to \infty$, under the same assumptions as Theorem OA.5 and with $\kappa_4 = 0$. This result guarantees that the estimators $\hat{\mathbf{L}}^{(\mathrm{m})}_{t-1}$ and $\hat{\boldsymbol{\Omega}}^{(\mathrm{m})}_{t-1}$ yield asymptotically valid standard errors.[24]

[23] To ease the exposition, the definition of $\boldsymbol{\Omega}^{(\mathrm{m})}_{t-1}$ has been relegated to the proof of the theorem (see (OA.95)).

[24] We omit the proof of Theorem OA.6, as it closely parallels the proof of Theorem 2.

Theorem OA.6 (standard errors of CSR OLS — time-varying estimator robust to misspecification)**.** *Under Assumptions OA.9–OA.8 and OA.2, and the identification condition $\kappa_4 = 0$, as $N \to \infty$, it holds that*

$$\hat{\mathbf{L}}^{(\mathrm{m})}_{t-1} = \frac{1}{N}\begin{bmatrix} \hat{\mathbf{X}}'\hat{\mathbf{X}} - N(\hat{\boldsymbol{\Lambda}}_1 + \hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{1,t-1}) & \hat{\mathbf{X}}'\mathbf{Z}_{t-1} - N\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{3,t-1} \\ \mathbf{Z}'_{t-1}\hat{\mathbf{X}} & \mathbf{Z}'_{t-1}\mathbf{Z}_{t-1} \end{bmatrix} \to_p \mathbf{L}^{(\mathrm{m})}_{t-1}$$

and

$$\hat{\boldsymbol{\Omega}}^{(\mathrm{m})}_{t-1} = \hat{\boldsymbol{\Omega}}_{1,t-1} + \hat{\boldsymbol{\Omega}}_{2,t-1} + \hat{\boldsymbol{\Omega}}_{3,t-1} + \hat{\boldsymbol{\Omega}}'_{3,t-1} + \hat{\boldsymbol{\Omega}}_{4,t-1} + \hat{\boldsymbol{\Omega}}'_{4,t-1} \to_p \boldsymbol{\Omega}^{(\mathrm{m})}_{t-1},$$

setting

$$\hat{\boldsymbol{\Omega}}_{1,t-1} = \hat{\sigma}^2\hat{\sigma}_{t-1mm}\begin{bmatrix} 0 & \mathbf{0}'_{K_f} & \mathbf{0}'_{K_z} \\ \mathbf{0}_{K_f} & \mathbf{P}'\begin{bmatrix}\mathbb{M}_{11}^{-1}\mathbb{M}_{12} \\ \mathbf{I}_{T-t+1}\end{bmatrix}\begin{bmatrix}\mathbb{M}'_{12}\mathbb{M}_{11}^{-1} & \mathbf{I}_{T-t+1}\end{bmatrix}\mathbf{P} & \mathbf{0}_{K_f\times K_z} \\ \mathbf{0}_{K_z} & \mathbf{0}_{K_z\times K_f} & \mathbf{0}_{K_z\times K_z}\end{bmatrix},$$

$$\hat{\boldsymbol{\Omega}}_{2,t-1} = \begin{bmatrix} 0 & \mathbf{0}'_{K_f} & \mathbf{0}'_{K_z} \\ \mathbf{0}_{K_f} & \hat{\mathbf{A}}_{t-1}\hat{\mathbf{U}}_{\epsilon}\hat{\mathbf{A}}'_{t-1} & \mathbf{0}_{K_f\times K_z} \\ \mathbf{0}_{K_z} & \mathbf{0}_{K_z\times K_f} & \mathbf{0}_{K_z\times K_z}\end{bmatrix},$$

$$\hat{\boldsymbol{\Omega}}_{3,t-1} = \begin{bmatrix} 0 & \mathbf{0}'_{K_f} & \mathbf{0}'_{K_z} \\ \mathbf{0}_{K_f} & -\mathbf{P}'\begin{bmatrix}\mathbb{M}_{11}^{-1}\mathbb{M}_{12} \\ \mathbf{I}_{T-t+1}\end{bmatrix}\begin{bmatrix}\mathbf{0}_{T-t+1\times(t-2)^2} & (\hat{\boldsymbol{\theta}}'_{t-1,m}\otimes\hat{\sigma}^2\mathbf{I}_{T-1}) & (\hat{\sigma}^2\mathbf{I}_{T-1}\otimes\hat{\boldsymbol{\theta}}'_{t-1,m}) & \mathbf{0}_{T-t+1\times(T-t+1)^2}\end{bmatrix}\hat{\mathbf{V}}_{t-1} & \mathbf{0}_{K_f\times K_z} \\ \mathbf{0}_{K_z} & \mathbf{0}_{K_z\times K_f} & \mathbf{0}_{K_z\times K_z}\end{bmatrix},$$

$$\hat{\boldsymbol{\Omega}}_{4,t-1} = \begin{bmatrix} 0 & \mathbf{0}'_{K_f} & \mathbf{0}'_{K_z} \\ \mathbf{0}_{K_f} & -\mathbf{P}'\begin{bmatrix}\mathbb{M}_{11}^{-1}\mathbb{M}_{12} \\ \mathbf{I}_{T-t+1}\end{bmatrix}\begin{bmatrix}\mathbf{0}_{T-t+1\times(t-2)^2} & (\hat{\boldsymbol{\theta}}'_{t-1,m}\otimes\hat{\sigma}^2\mathbf{I}_{T-1}) & (\hat{\sigma}^2\mathbf{I}_{T-1}\otimes\hat{\boldsymbol{\theta}}'_{t-1,m}) & \mathbf{0}_{T-t+1\times(T-t+1)^2}\end{bmatrix}\hat{\mathbf{A}}'_{t-1} & \mathbf{0}_{K_f\times K_z} \\ \mathbf{0}_{K_z} & \mathbf{0}_{K_z\times K_f} & \mathbf{0}_{K_z\times K_z}\end{bmatrix},$$

where

$$\hat{\boldsymbol{\theta}}_{t-1,m} = \mathbf{M}^{(-1)}_{\mathrm{D},t-1}\left(\frac{1}{N}\hat{\boldsymbol{\epsilon}}'\hat{\mathbf{m}}_{t-1} - \hat{\sigma}^2\mathbb{M}_D\boldsymbol{\iota}_{t-1,T-1}\right)$$ comma

with

$$\hat{\mathbf{m}}_{t-1} = \mathbf{R}_t - (\hat{\mathbf{X}}, \mathbf{Z}_{t-1})(\hat{\boldsymbol{\Gamma}}^{*(\mathrm{m})\prime}_{\mathrm{f},t-1}, \hat{\boldsymbol{\gamma}}^{*(\mathrm{m})\prime}_{\mathrm{z},t-1})', \quad \boldsymbol{M}^{(-1)}_{D,t-1} = \mathbb{M}_{11}^{-1}[\mathbf{I}_{t-2}, \mathbf{0}_{(t-2)\times(T-t+1)}]$$ comma

and

$$\hat{\sigma}_{t-1mm} = \frac{\hat{\mathbf{m}}'_{t-1}\hat{\mathbf{m}}_{t-1}}{N} - \hat{\sigma}^2\hat{\mathbf{Q}}'_{t-1}\hat{\mathbf{Q}}_{t-1} + 2\hat{\boldsymbol{\delta}}^{*(\mathrm{m})\prime}_{\mathrm{f},t-1}\mathbf{P}'\begin{bmatrix}\hat{\boldsymbol{\theta}}_{t-1,m} \\ \mathbf{0}_{T-t+1}\end{bmatrix},$$

and where

$$\hat{\mathbf{A}}_{t-1,T-1} \qquad = -\mathbf{P}' \begin{bmatrix} \left(\hat{\mathbf{Q}}'_{t-1} \otimes \mathbb{M}_{11}^{-1}(\mathbf{I}_{t-2}, \mathbf{0}_{t-2\times T-t+1})\overline{\mathbb{M}}_{\tilde{D}}\right)\left(\mathbf{I}^2_{T-1} - vec(\mathbf{I}_{T-1})\frac{vec'(\overline{\mathbb{M}}_{\tilde{D}})}{T-K_f-K_z-2}\right) \\ \mathbf{0}_{T-t+1\times(T-1)^2} \end{bmatrix},$$

while all the other quantities are defined in Theorem OA.2.

Proof. Omitted, as it closely parallels the proof of Theorem OA.2. See Section OA.5.

OA.2.4 Measuring Anomalies' Contribution Jointly: The R-Squared Test

Despite the considerable literature on asset pricing anomalies, how much of the cross-sectional variation in expected returns is accounted for by betas and how much by anomalies is still unclear, still representing a challenging question.

Offering a simple criterion that can answer this question and allow for conducting formal inference on (joint) anomalies' contribution is the objective of this section. One could consider the ratios of the (cross-sectional) variance of the betas' component and the anomalies' component to the overall (cross-sectional) variance of average returns, to measure their relative contribution. Specifically, suppose one has estimated the model (OA.13) using our bias-adjusted CSR OLS estimator, hence obtaining:

$$\mathbf{R}_t = \hat{\mathbf{X}}\hat{\mathbf{\Gamma}}^*_{\mathrm{f},t-1} + \mathbf{Z}_{t-1}\hat{\gamma}^*_{\mathrm{z},t-1} + \hat{\boldsymbol{\eta}}_t, \tag{OA.43}$$

where $\hat{\boldsymbol{\eta}}_t$ indicates the $N \times 1$ vector of residuals. Then, the fraction of the overall variance explained by the anomaly variables $\mathbf{Z}_{t-1}$ (at any point in time) would simply be

$$\hat{R}^{2(\mathrm{bench})}_{\mathrm{z},t-1} = \frac{\hat{\gamma}^{*\prime}_{\mathrm{z},t-1}\mathbf{Z}'_{t-1}\mathbb{M}_{1_N}\mathbf{Z}_{t-1}\hat{\gamma}^*_{\mathrm{z},t-1}}{\mathbf{R}'_t\mathbb{M}_{1_N}\mathbf{R}_t}. \tag{OA.44}$$

However, despite being a very simple and intuitive measure, the R-squared in (OA.44) could lead to several problems. First, since beta and anomaly components are not necessarily orthogonal cross-sectionally, this can lead to a fraction of the cross-sectional variance explained by the betas and by the anomaly variables - expressed by the sum of the corresponding R-squared - that is jointly greater than 100%. In addition, while (by construction) orthogonality between CSR residuals $\hat{\boldsymbol{\eta}}_t$ and the regressors (both $\hat{\mathbf{X}}$ and $\mathbf{Z}_{t-1}$) is warranted by the conventional CSR OLS estimator in (OA.14), this does not hold when considering our bias-adjusted estimator $\left(\hat{\mathbf{\Gamma}}^{*\prime}_{\mathrm{f},t-1}, \hat{\boldsymbol{\gamma}}^{*\prime}_{\mathrm{z},t-1}\right)'$ of (OA.15), implying that $\hat{R}^{2(\mathrm{bench})}_{\mathrm{z},t-1}$ is even wrongly centred.[25]

To overcome such (lack of) orthogonality issues, let us rearrange the estimated asset pricing model (OA.43) as follows:

$$
\begin{aligned}
\mathbf{R}_t &= \hat{\mathbf{X}}\hat{\mathbf{\Gamma}}^{*}_{\mathrm{f},t-1} + \mathbf{Z}_{t-1}\hat{\boldsymbol{\gamma}}^{*}_{\mathrm{z},t-1} + \mathbb{P}_{[\hat{\mathbf{X}},\mathbf{Z}_{t-1}]}\hat{\boldsymbol{\eta}}_t + \mathbb{M}_{[\hat{\mathbf{X}},\mathbf{Z}_{t-1}]}\hat{\boldsymbol{\eta}}_t \\
&= \mathbb{P}_{\hat{\mathbf{X}}}\left(\hat{\mathbf{X}}\hat{\mathbf{\Gamma}}^{*}_{\mathrm{f},t-1} + \mathbf{Z}_{t-1}\hat{\boldsymbol{\gamma}}^{*}_{\mathrm{z},t-1} + \mathbb{P}_{[\hat{\mathbf{X}},\mathbf{Z}_{t-1}]}\hat{\boldsymbol{\eta}}_t\right) + \mathbb{M}_{\hat{\mathbf{X}}}\left(\mathbf{Z}_{t-1}\hat{\boldsymbol{\gamma}}^{*}_{\mathrm{z},t-1} + \mathbb{P}_{[\hat{\mathbf{X}},\mathbf{Z}_{t-1}]}\hat{\boldsymbol{\eta}}_t\right) + \mathbb{M}_{[\hat{\mathbf{X}},\mathbf{Z}_{t-1}]}\hat{\boldsymbol{\eta}}_t.
\end{aligned}
\tag{OA.45}
$$

where we use the notation $\mathbb{M}_A = \mathbf{I}_a - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' = \mathbf{I}_a - \mathbb{P}_A$, with $\mathbb{P}_A = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$, for a generic matrix $\mathbf{A}$ of dimension $a \times b$ and rank $b < a$. Notice that, by construction, the three terms on the right-hand side of (OA.45) are now mutually orthogonal, with the second term reflecting the joint contribution of $\mathbf{Z}_{t-1}$. This yields our proposed R-squared test statistic:

$$
\hat{R}^2_{\mathrm{z},t-1} = \frac{\left(\hat{\boldsymbol{\gamma}}^{*\prime}_{\mathrm{z},t-1}\mathbf{Z}'_{t-1} + \hat{\boldsymbol{\eta}}'_t\mathbb{P}_{[\hat{\mathbf{X}},\mathbf{Z}_{t-1}]}\right)\mathbb{M}_{\hat{X}}\left(\mathbf{Z}_{t-1}\hat{\boldsymbol{\gamma}}^{*}_{\mathrm{z},t-1} + \mathbb{P}_{[\hat{\mathbf{X}},\mathbf{Z}_{t-1}]}\hat{\boldsymbol{\eta}}_t\right)}{\mathbf{R}'_t\mathbb{M}_{1_N}\mathbf{R}_t}, \tag{OA.46}
$$

which satisfies $0 \leq \hat{R}^2_{\mathrm{z},t-1} \leq 1$. Notice that (OA.46) represents a meaningful quantity, which allows us to disentangle the contribution of that portion of anomalies that is *unexplained* by - i.e., orthogonal to - the risk exposures (through the term $\mathbb{M}_{\hat{\mathbf{X}}}\mathbf{Z}_{t-1}$), as well as the contribution that might arise from the term $\mathbb{P}_{[\hat{\mathbf{X}},\mathbf{Z}_{t-1}]}\hat{\boldsymbol{\eta}}_t$, which is not guaranteed to be null in general.

Armed with $\hat{R}^2_{\mathrm{z},t-1}$, one could test for the null hypothesis of zero anomalies' contribution and, in case of rejection, construct an asymptotically valid confidence interval for it. This requires us to establish the limiting statistical properties of $\hat{R}^2_{\mathrm{z},t-1}$, in particular its non-standard limiting distribution, occurring when $\boldsymbol{\gamma}_{\mathrm{z},t-1} = \mathbf{0}_{K_{\mathrm{z}}}$. Therefore, in the following, we derive the asymptotic distribution of $\hat{R}^2_{\mathrm{z},t-1}$, distinguishing between the two complementary cases of zero and non-zero anomalies' premia.[26]

[25]Finally, notice that when the (true) anomalies' premia $\boldsymbol{\gamma}_{\mathrm{z},t-1}$ are zero, then $\hat{R}^{2(\mathrm{bench})}_{\mathrm{z},t-1}$ will converge to zero in probability. This implies that we face a boundary problem - as necessarily $\hat{R}^{2(\mathrm{bench})}_{\mathrm{z},t-1} \geq 0$ - leading to a non-standard limiting distribution of the test statistic, under the null hypothesis of null anomalies' premia.

[26]The limiting behaviour of $\hat{R}^2_{\mathrm{z},t-1}$ resembles the one of the Hansen and Jagannathan (2007) (HJ) distance: both statistics are distributed like linear combinations of iid chi-squares (with one degree of freedom) under the corresponding null hypothesis, and normally distributed under their alternatives.

Theorem OA.7 (R^2 test of anomalies' contribution)**.** *Set the R-squared test statistic equal to*

$$\mathcal{T}^2_{\mathrm{z},t-1} = N\left(\hat{R}^2_{\mathrm{z},t-1} - \frac{\hat{\boldsymbol{\eta}}_t'\mathbb{P}_{[\hat{\mathbf{X}},\mathbf{Z}_{t-1}]}\mathbb{M}_{\hat{\mathbf{X}}}\mathbb{P}_{[\hat{\mathbf{X}},\mathbf{Z}_{t-1}]}\hat{\boldsymbol{\eta}}_t + 2\,\hat{\boldsymbol{\eta}}_t'\mathbb{P}_{[\hat{\mathbf{X}},\mathbf{Z}_{t-1}]}\mathbb{M}_{\hat{\mathbf{X}}}\mathbf{Z}_{t-1}\hat{\gamma}^*_{\mathrm{z},t-1}}{\mathbf{R}_t'\mathbb{M}_{1_N}\mathbf{R}_t}\right). \tag{OA.47}$$

Under Assumptions OA.9, OA.3 –OA.8, and OA.14 as $N \to \infty$, then:

(i) *When $\boldsymbol{\gamma}_{\mathrm{z},t-1} = \mathbf{0}_{K_{\mathrm{z}}}$,*

$$\mathcal{T}_{\mathrm{z},t-1} \to_d \sum_{j=1}^{K_{\mathrm{z}}} d_{j,t-1}\chi^2_{1,j},$$

where $(\chi^2_{1,1},\cdots,\chi^2_{1,K_{\mathrm{z}}})$ are i.i.d χ^2_1-distributed random variables , and $(d_{1,t-1},\cdots,d_{K_{\mathrm{z}},t-1})$ are the K_{z} eigenvalues of the matrix

$$\left(\mathbf{L}^{-1}_{\mathrm{z},t-1}\mathbf{O}_{t-1}\mathbf{L}^{-1\prime}_{\mathrm{z},t-1}\right)^{\frac{1}{2}}\frac{\boldsymbol{\Sigma}_{\mathrm{Z}\hat{\mathrm{X}}\mathrm{Z},t-1}}{\sigma_{\tilde{\mathrm{R}},t}}\left(\mathbf{L}^{-1}_{\mathrm{z},t-1}\mathbf{O}_{t-1}\mathbf{L}^{-1\prime}_{\mathrm{z},t-1}\right)^{\frac{1}{2}},$$

where $\mathbf{L}_{\mathrm{z},t-1} = \left[\mathbf{0}_{K_{\mathrm{z}}\times(K_{\mathrm{f}}+1)},\mathbf{I}_{K_{\mathrm{z}}}\right]\mathbf{L}_{t-1}$, with $\mathbf{L}_{t-1}$ and $\mathbf{O}_{t-1}$ defined in (OA.84)*, and where $N^{-1}\mathbf{R}_t'\mathbb{M}_{1_N}\mathbf{R}_t \to_p \sigma_{\tilde{\mathrm{R}},t} > 0$, while $N^{-1}\mathbf{Z}_{t-1}'\mathbb{M}_{\hat{\mathbf{X}}}\mathbf{Z}_{t-1} \to_p \boldsymbol{\Sigma}_{\mathrm{Z}\hat{\mathrm{X}}\mathrm{Z},t-1}$.*

(ii) *When $\boldsymbol{\gamma}_{t-1,z} \neq \mathbf{0}_{K_z}$,*

$$\mathcal{T}_{\mathrm{z},t-1} \to_p \infty.$$

Moreover, under the additional Assumption OA.14, together with $\kappa_4 = 0$, for any $0 < \alpha < 1$,

$$\Pr\left(\hat{R}^2_{\mathrm{z},t-1} - z_{\alpha/2}\left(\frac{\hat{\omega}_{\mathrm{z},t-1}}{N}\right)^{\frac{1}{2}} \leq R^2_{\mathrm{z},t-1} \leq \hat{R}^2_{\mathrm{z},t-1} + z_{\alpha/2}\left(\frac{\hat{\omega}_{\mathrm{z},t-1}}{N}\right)^{\frac{1}{2}}\right) \to (1-\alpha).$$

where $\hat{\omega}_{\mathrm{z},t-1}$ represents a consistent estimator of the asymptotic variance of $\hat{R}^2_{\mathrm{z},t-1}$, $z_{\alpha/2}$ denotes the $\alpha/2$-th quantile of the standard normal distribution, and $R^2_{t-1,z}$ denotes the (probability) limit of $\hat{R}^2_{t-1,z}$.[27]

Proof. See Section OA.5.

OA.3 Main Assumptions

In this section, we recall all the assumptions required for the validity of our large-N asymptotic theory, together with some explanatory remarks. Assumptions and definitions refer to the more

[27]For the definition of $\hat{\omega}_{\mathrm{z},t-1}$ see (OA.100).

no ( )

general case of Section OA.2; the corresponding quantities that apply when $\gamma_{0,t-1}$ is replaced by the risk-free rate can be readily obtained. All the moments below are assumed to hold conditionally on the factors $\mathbf{F}$, even if not written explicitly, and all the limits are taken as $N \to \infty$.

It is useful to recall the $N \times K_{\mathrm{z}}(T-1)$ matrix of anomalies $\mathbf{Z} = (\mathbf{z}_1, \ldots, \mathbf{z}_N)'$, where $\mathbf{z}_i$ denotes the $K_{\mathrm{z}}(T-1) \times 1$ vector

$$\mathbf{z}_i = \left(z_{i,1}^{(1)}, \ldots, z_{i,T-1}^{(1)}, \ldots, z_{i,1}^{(K_{\mathrm{z}})}, \ldots, z_{i,T-1}^{(K_{\mathrm{z}})}\right)'.$$

The $N \times K_{\mathrm{z}}$ matrix of anomalies at time $t-1$ is defined as $\mathbf{Z}_{t-1} = (\mathbf{z}_{1,t-1}, \ldots, \mathbf{z}_{N,t-1})'$, while the $(T-1) \times K_{\mathrm{z}}$ matrix of anomalies specific to the i-th asset is $\mathbf{Z}_i = (\mathbf{z}_{i,1}, \ldots, \mathbf{z}_{i,T-1})'$, with $\mathbf{z}_{i,t-1} = \left(z_{i,t-1}^{(1)}, \ldots, z_{i,t-1}^{(K_{\mathrm{z}})}\right)'$.

OA.3.1 Assumptions required for the bias-adjusted OLS estimation

Assumption OA.3 (*risk factors and anomalies*)**.** Set $\tilde{\mathbf{Z}}_i = \mathbb{M}_{1_{T-1}} \mathbf{Z}_i$, and $\mathbf{D} = (\mathbf{1}_{T-1}, \mathbf{F})$. Then, for every T, the $(T-1) \times (K+1)$ matrix $\tilde{\mathbf{D}}_i = (\mathbf{D}, \tilde{\mathbf{Z}}_i)$ satisfies

$$\tilde{\mathbf{D}}_i' \tilde{\mathbf{D}}_i > 0 \quad \text{for every } i = 1, ..., N.$$

Assumption OA.4 (*loadings*)**.**

$$\frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\beta}_i \to \boldsymbol{\mu}_\beta \qquad \text{and} \qquad \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\beta}_i \boldsymbol{\beta}_i' \to \boldsymbol{\Sigma}_\beta,$$

such that the matrix

$$\boldsymbol{\Sigma}_{\mathrm{X}} = \begin{bmatrix} 1 & \boldsymbol{\mu}_\beta' \\ \boldsymbol{\mu}_\beta & \boldsymbol{\Sigma}_\beta \end{bmatrix} > 0.$$

Remark OA.3. *Assumption OA.4 states that the limiting cross-sectional averages of the betas and of the squared betas exist. Its second part rules out the possibility of spurious or weak factors, as well as situations in which at least one element of $\boldsymbol{\beta}_i$ is cross-sectionally constant. This, in turn, implies that $\mathbf{X} = (\mathbf{1}_N, \mathbf{B})$ has full column rank for sufficiently large N. For simplicity of exposition, we assume that the $\boldsymbol{\beta}_i$ are non-random.*[28]

[28] See Gagliardini, Ossola, and Scaillet (2016) and Raponi, Robotti, and Zaffaroni (2020) for analyses of asset pricing models with random betas.

Assumption OA.5 (*asset-specific components*)**.** The $N \times 1$ vector of error terms $\boldsymbol{\epsilon}_t$ is independently and identically distributed (i.i.d.) over time with

$$\mathrm{E}[\boldsymbol{\epsilon}_t] = \mathbf{0}_N \quad \text{comma} \tag{OA.48}$$

and with the $N \times N$ variance-covariance matrix satisfying

$$\mathrm{Var}[\boldsymbol{\epsilon}_t] = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2N} \\ \vdots & \vdots & \cdots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_N^2 \end{pmatrix} = \boldsymbol{\Sigma} > 0, \tag{OA.49}$$

where σ_{ij} denotes the (i,j)-th element of $\boldsymbol{\Sigma}$, for every $i,j = 1,\ldots,N$, and with $\sigma_i^2 = \sigma_{ii}$.

Remark OA.4. *The i.i.d. assumption over time is common in many studies, including Shanken (1992) and Raponi, Robotti, and Zaffaroni (2020). Nonetheless, in principle, our large-N asymptotic theory allows the ϵ_{it} to be arbitrarily correlated over time, albeit at the cost of more involved expressions and derivations for the limiting distributions of the estimators. Condition (OA.49) does not impose any specific structure on the elements of $\boldsymbol{\Sigma}$ beyond non-singularity. In particular, we do not require returns' innovations to be uncorrelated across assets or to exhibit homoskedasticity. However, our large-N asymptotic theory must discipline the degree of cross-sectional correlation among the $\boldsymbol{\epsilon}_t$ (as N increases), as specified in Assumption OA.6 below.*

Assumption OA.6. *(cross-sectional moments of asset-specific components)*

(i)
$$\frac{1}{N}\sum_{i=1}^{N}\left(\sigma_i^2 - \sigma^2\right) = o\left(\frac{1}{\sqrt{N}}\right), \tag{OA.50}$$
for some $0 < \sigma^2 < \infty$.

(ii)
$$\sum_{i,j=1}^{N} \mid \sigma_{ij} \mid \mathbb{1}_{\{i \neq j\}} = o\left(N\right). \tag{OA.51}$$

(iii)
$$\frac{1}{N}\sum_{i=1}^{N} \mu_{4i} \to \mu_4, \tag{OA.52}$$
for some $0 < \mu_4 < \infty$, where $\mu_{4i} = \mathrm{E}[\epsilon_{i,t}^4]$.

(iv)
$$\frac{1}{N}\sum_{i=1}^{N}\sigma_i^4 \to \sigma_4, \tag{OA.53}$$
for some $0 < \sigma_4 < \infty$.

(v)
$$\sup_i \mu_{4i} \leq C < \infty, \tag{OA.54}$$
for a generic constant C.

(vi)
$$\mathrm{E}[\epsilon_{i,t}^3] = 0. \tag{OA.55}$$

(vii)
$$\frac{1}{N}\sum_{i=1}^{N}\kappa_{4,iiii} \to \kappa_4, \tag{OA.56}$$
for some $0 \leq |\kappa_4| < \infty$, where $\kappa_{4,iiii} = \kappa_4[\epsilon_{it}, \epsilon_{it}, \epsilon_{it}, \epsilon_{it}]$ denotes the fourth-order cumulant of the asset-specific component $\{\epsilon_{i,t}, \epsilon_{i,t}, \epsilon_{i,t}, \epsilon_{i,t}\}$.

(viii) For every $3 \leq h \leq 8$, all the following mixed cumulants of order h satisfy
$$\sup_{i_1}\sum_{i_2,\ldots,i_h=1}^{N}|\kappa_{h,i_1 i_2 \ldots i_h}| = o\left(N\right), \tag{OA.57}$$
for at least one i_j $(2 \leq j \leq h)$ different from i_1, where $\kappa_{h,i_1,i_2\ldots i_h}$ is the mixed cumulant in the $\{\epsilon_{i_1,s}, \epsilon_{i_2,s}, \cdots, \epsilon_{i_h,s}\}$ of order h.

Remark OA.5. *Assumption OA.6 describes the cross-sectional behavior of the asset-specific components. Specifically, Assumption OA.6(i) limits the cross-sectional heterogeneity of the returns' conditional variance, while Assumption OA.6(ii) sets the maximum degree of (conditional) cross-correlation among asset returns allowed by our theory. These assumptions are not very restrictive and accommodate several forms of strong cross-sectional dependence among the* ϵ_{it}*'s. One example is a factor structure of the following form:*
$$\epsilon_{it} = \lambda_i u_t + \eta_{i,t}, \tag{OA.58}$$
where $u_t \sim i.i.d.(0,1)$ *and* $\eta_{i,t} \sim i.i.d.(0,\sigma_\eta^2)$ *over time and across units, and where* u_t *and* $\eta_{i,s}$ *are mutually independent for every* $i, s,$ *and* t. *The coefficient* λ_i *is such that* $\sum_{i=1}^{N}|\lambda_i| = O(N^\delta)$, *with* $0 \leq \delta < 1/4$, *and* $\lambda_1 + \cdots + \lambda_q \sim CN^\delta$, *for some fixed* $q < N$ *and constant* C.

Although Assumptions OA.6(i) and OA.6(ii) are easily satisfied in the special case $\sigma^2 = \sigma_\eta^2$, notice that the maximum eigenvalue of $\boldsymbol{\Sigma}$ is unbounded as $N \to \infty$.[29] *This contrasts with the standard Asset Pricing Theory (APT), where boundedness of the maximum eigenvalue is the most common assumption (see, e.g., the generalization of the APT by Chamberlain and Rothschild (1983)). Our assumptions are thus milder than those postulated by the APT and therefore more likely to be verified in the data.*[30]

Other special cases encompassed by Assumption OA.6 (with non-zero cross-covariances σ_{ij}) include network and spatial measures of cross-dependence, and suitably modified versions of the block-dependence structure of Gagliardini, Ossola, and Scaillet (2016).[31]

We demonstrate the substantial degree of cross-correlation between the ϵ_{it} allowed by our setup by means of a simple numerical example, illustrated in Table OA.I. Using the data generating process

$$R_{it} = \gamma_0 + Z_{1i,t-1}\gamma_{z_1} + Z_{2i,t-1}\gamma_{z_2} + \beta_i\big(\gamma_1 + f_t - E(f_t)\big) + \epsilon_{it}, \tag{OA.59}$$

we consider balanced panels with a time-series dimension of $T = 36$ and $T = 72$ observations. Specifically, f_t in (OA.59) *is the excess market return (from Kenneth French's website) over January 2013–December 2015 for $T = 36$, and over January 2011–December 2015 for $T = 72$. In addition, $E(f_t)$ in* (OA.59) *is set equal to the time-series mean of f_t over the corresponding sample periods, i.e., 2013–2015 and 2011–2015 for $T = 36$ and $T = 72$, respectively.*

We calibrate the true values for γ_0, γ_1, and β_i in (OA.59) *using a cross-section of 1,000 stocks from the CRSP database, setting them equal to OLS estimates of β_i from the first-pass regressions and to CSR OLS estimates of γ_0 and γ_1. For the anomalies $Z_{1i,t-1}$ and $Z_{2i,t-1}$ in* (OA.59)*, we use data on two firms' characteristics, namely the book-to-market ratio ($Z^\dagger_{1i,t-1}$) and asset growth ($Z^\dagger_{2i,t-1}$), over the two sample periods: December 2012–November 2015 for $T = 36$, and December 2010–November 2015 for $T = 72$. The anomaly variables are lagged with respect to the risk factors*

[29] The maximum eigenvalue of $\boldsymbol{\Sigma}$ is given by $\sup_{\mathbf{c}:\|\mathbf{c}\|=1} \mathbf{c}'\boldsymbol{\Sigma}\mathbf{c}$.

[30] Specifically, Assumption OA.6 allows the maximum eigenvalue of $\boldsymbol{\Sigma}$ to diverge at rate $o\left(\sqrt{N}\right)$. This implies that the row-column norm of $\boldsymbol{\Sigma}$, namely $\sup_{1\le i\le N}\sum_{j=1}^N |\sigma_{ij}|$, can diverge as $N \to \infty$. Gagliardini, Ossola, and Scaillet (2016) allow for an even faster rate of divergence, $o(N)$, but in their setting both T and N grow jointly.

[31] Assumption BD.2 of Gagliardini, Ossola, and Scaillet (2016) on block sizes and block numbers requires that the largest block size shrinks with N and that there are not too many large blocks; in other words, the partition into independent blocks must be sufficiently fine-grained asymptotically. They formally show that such a block-dependence structure is compatible with the unboundedness of the maximum eigenvalue of $\boldsymbol{\Sigma}$.

and the test assets.

We then orthogonalize both $Z^{\dagger}_{1i,t-1}$ *and* $Z^{\dagger}_{2i,t-1}$ *with respect to the market factor* f_t*, that is, we consider*

$$\boldsymbol{Z}_{1i} = \boldsymbol{M}_D \boldsymbol{Z}^{\dagger}_{1i} + \mathbf{1}_{T-1} \bar{Z}^{\dagger}_{1i}, \tag{OA.60}$$

$$\boldsymbol{Z}_{2i} = \boldsymbol{M}_D \boldsymbol{Z}^{\dagger}_{2i} + \mathbf{1}_{T-1} \bar{Z}^{\dagger}_{2i}, \tag{OA.61}$$

where $\boldsymbol{Z}_{1i} = [Z_{1i,1}, \ldots, Z_{1i,T-1}]'$*,* $\boldsymbol{Z}_{2i} = [Z_{2i,1}, \ldots, Z_{2i,T-1}]'$*,* $\boldsymbol{M}_D = \boldsymbol{I}_{T-1} - \boldsymbol{D}(\boldsymbol{D}'\boldsymbol{D})^{-1}\boldsymbol{D}'$*, with* $\boldsymbol{D} = (\mathbf{1}_{T-1}, \boldsymbol{f})$*, and* $\bar{Z}^{\dagger}_{ik} = (T-1)^{-1}\sum_{t=1}^{T-1} Z^{\dagger}_{ki,t}$*, for* $k = 1, 2$*. In this way we ensure that the anomalies and the risk factor are orthogonal, that is,* $\widehat{\mathrm{Cov}}(Z_{1i,t-1}, f_t) = \widehat{\mathrm{Cov}}(Z_{2i,t-1}, f_t) = 0$*, and that the sample means of both* $\boldsymbol{Z}_{1i}$ *and* $\boldsymbol{Z}_{2i}$ *match those of* $\boldsymbol{Z}^{\dagger}_{1i}$ *and* $\boldsymbol{Z}^{\dagger}_{2i}$*.*

We then use (OA.118) *and* (OA.119) *in the data generating process* (OA.59)*, where, in our simulation design, both the factor and the anomalies' data are taken as given and kept fixed throughout. We calibrate* γ_{z_1} *using our bias-adjusted CSR OLS estimator by regressing stock returns on the excess market factor and the book-to-market anomaly, while* γ_{z_2} *is calibrated by estimating the regression of stock returns on the excess market factor and the asset growth anomaly. Finally, the calibration of the error term* ϵ_{it} *in* (OA.59) *is based on (OA.58).*

We compute the following quantity:

$$A(\delta, N) = \frac{1}{N}\sum_{j=1}^{N}\left(\frac{\sum_{i=1,\, i\neq j}^{N} |\sigma_{ij}|}{\sigma_j^2 + \sum_{i=1,\, i\neq j}^{N} |\sigma_{ij}|}\right). \tag{OA.62}$$

For the case of zero cross-correlation, that is, $\sigma_{ij} = 0$ *when* $i \neq j$*, we obtain* $A(\delta, N) = 0$*. In contrast, when the cross-correlations become dominant,* $A(\delta, N)$ *approaches 1. Varying* δ *and* N *(averaged over 1,000 Monte Carlo iterations), we obtain the results reported in Table OA.I.*

This simple numerical example shows that for every value of δ *(which is* proportionally *related to the degree of dependence), the sum of the cross-covariances tends to increase with* N *in relative terms and dominates the behavior of the overall covariance matrix.*

Assumption OA.6(iii) simply requires the existence of the limit of the conditional fourth moment, averaged across assets. In Assumption OA.6(iv), the magnitude of σ_4 *reflects the degree of cross-sectional heterogeneity in the conditional variance of asset returns. Assumption OA.6(v) imposes a bounded fourth-moment condition, uniform across assets, which implies that* $\sup_i \sigma_i^2 \leq C < \infty$*.*

Table OA.I: Conditional Cross-Sectional Dependence: Returns

$A(\delta, N)$	Value
A(0.25, 100)	0.8585
A(0.25, 500)	0.9268
A(0.25, 3000)	0.9628
A(0.245, 100)	0.8560
A(0.245, 500)	0.9235
A(0.245, 3000)	0.9606
A(0.01, 100)	0.5418
A(0.01, 500)	0.5680
A(0.01, 3000)	0.5612

Assumption OA.6(vi) is a convenient symmetry condition, but it is not strictly necessary for our results. This assumption could be relaxed, although the derivation of the asymptotic distribution would then become more cumbersome due to the appearance of additional terms involving the third moment of the disturbance. Assumptions OA.6(vii)–(viii) allow for non-Gaussianity of asset returns whenever $\kappa_4 > 0$*. For example, they are satisfied when the marginal distribution of asset returns follows a Student-t distribution with more than four degrees of freedom. However, when estimating the asymptotic covariance matrix of our bias-adjusted estimators, it is necessary to set* $\kappa_4 = 0$ *purely for identification purposes. That said, higher-order cumulants are not constrained to be zero, implying that even when* $\kappa_4 = 0$*, the distribution is not necessarily Gaussian.*

Assumption OA.7 (*CLT of asset-specific components*)**.** *As* $N \to \infty$*:*

(i)

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \boldsymbol{\epsilon}_i \ \to_d \ \mathcal{N}\big(\mathbf{0}_{T-1}, \sigma^2 \mathbf{I}_{T-1}\big) \,. \tag{OA.63}$$

(ii) Let

$$\mathbf{U}_\epsilon = \lim_{N\to\infty} \frac{1}{N} \sum_{i,j=1}^{N} \mathrm{E}\Big[\mathrm{vec}\big(\boldsymbol{\epsilon}_i \boldsymbol{\epsilon}_i' - \sigma_i^2 \mathbf{I}_{T-1}\big) \mathrm{vec}\big(\boldsymbol{\epsilon}_j \boldsymbol{\epsilon}_j' - \sigma_j^2 \mathbf{I}_{T-1}\big)'\Big].$$

Then,

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathrm{vec}\big(\boldsymbol{\epsilon}_i \boldsymbol{\epsilon}_i' - \sigma_i^2 \mathbf{I}_{T-1}\big) \ \to_d \ \mathcal{N}\big(\mathbf{0}_{(T-1)^2}, \mathbf{U}_\epsilon\big) \,. \tag{OA.64}$$

(iii) For any $T \times 1$ vector $\mathbf{c}$,

defined?

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\mathbf{c}' \otimes \begin{bmatrix} 1 \\ \boldsymbol{\beta}_i \end{bmatrix}\right) \boldsymbol{\epsilon}_i \ \to_d \ \mathcal{N}\big(\mathbf{0}_{K_\mathrm{f}+1}, \, (\mathbf{c}'\mathbf{c})\, \sigma^2 \boldsymbol{\Sigma}_\mathrm{X}\big) \,. \tag{OA.65}$$

Remark OA.6. *From* (OA.65)*, it follows that*

$$N^{-1/2} \sum_{i=1}^{N} \left(\mathbf{c}' \otimes \boldsymbol{\beta}_i\right) \boldsymbol{\epsilon}_i \quad \to_d \quad \mathcal{N}\left(\mathbf{0}_{K_{\mathrm{f}}},\, (\mathbf{c}'\mathbf{c})\, \sigma^2 \boldsymbol{\Sigma}_\beta\right).$$

Primitive conditions for Assumption OA.7 can be derived, but only at the cost of increasing the complexity of our proofs.[32]

Remark OA.7. *The expression for* $\mathbf{U}_\epsilon$ *in* (OA.64) *can be derived in closed form. In particular, the* $T^2 \times T^2$ *matrix* $\mathbf{U}_\epsilon$ *has the following structure:*

$$\mathbf{U}_\epsilon = \left[\begin{array}{c:c:c:c:c} U_{11} & \cdots & U_{1t} & \cdots & U_{1T} \\ \vdots & \ddots & \vdots & & \vdots \\ U_{t1} & \cdots & U_{tt} & \cdots & U_{tT} \\ \vdots & & \vdots & \ddots & \vdots \\ U_{T1} & \cdots & U_{Tt} & \cdots & U_{TT} \end{array}\right].$$

Each block of $\mathbf{U}_\epsilon$ *is a* $T \times T$ *matrix. The blocks along the main diagonal, denoted* U_{tt} *for* $t = 1, 2, \ldots, T$*, are themselves diagonal matrices with* $(\kappa_4 + 2\sigma_4)$ *in the* (t,t)*-th position and* σ_4 *in the* (s,s) *position for every* $s \neq t$*. The blocks outside the main diagonal, denoted* U_{ts} *for* $s, t = 1, 2, \ldots, T$ *with* $s \neq t$*, are all zero matrices except for the* (s,t)*-th element, which equals* σ_4*.*

$$U_{tt} = \begin{array}{cc} & \begin{array}{c} \downarrow \\ t\text{-th column} \end{array} \\ \begin{array}{c} \\ \\ \\ \underset{\substack{t\text{-th} \\ row}}{\rightarrow} \\ \\ \\ \\ \end{array} & \begin{bmatrix} \sigma_4 & \cdots & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \sigma_4 & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & (\kappa_4 + 2\sigma_4) & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & \sigma_4 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & \sigma_4 \end{bmatrix} \end{array}, \quad U_{ts} = \begin{array}{cc} & \begin{array}{c} \downarrow \\ t\text{-th column} \end{array} \\ \begin{array}{c} \\ \\ \\ \underset{\substack{s\text{-th} \\ row}}{\rightarrow} \\ \\ \\ \\ \end{array} & \begin{bmatrix} 0 & \cdots & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & \sigma_4 & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{bmatrix} \end{array}.$$

Assumption OA.8 (*Moments and Central Limit Theorem for anomalies*)**.** *Define the* $K_{\mathrm{z}}(T-1)^2 \times 1$ *vector* $\mathbf{u}_i = \boldsymbol{\epsilon}_i \otimes \mathbf{z}_i$*. The following conditions hold:*

[32] For instance, when (OA.58) holds and all the above assumptions are satisfied, then (OA.63) follows from Theorem 2 of Kuersteiner and Prucha (2013), provided that η_{it} satisfies their martingale difference assumptions (see Assumptions 1 and 2 therein). This result extends directly to (OA.64)–(OA.65) under suitable additional assumptions. Details are available upon request.

(i) *Let* $\boldsymbol{\mu}_{\mathrm{z}} = (\mu_{\mathrm{z}}^{(1)}, \ldots, \mu_{\mathrm{z}}^{(K_{\mathrm{z}})})' = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\mu}_{\mathrm{z}_i}$*, with* $\boldsymbol{\mu}_{\mathrm{z}_i} = \mathrm{E}[\mathbf{z}_{i,s}]$*. Then*

$$\frac{\mathbf{Z}'\mathbf{1}_N}{N} \to_p (\boldsymbol{\mu}_{\mathrm{z}} \otimes \mathbf{1}_{T-1}) = \boldsymbol{\mu}_{\mathrm{z},T-1}.$$

(ii) *Let* $\boldsymbol{\Sigma}_{\mathrm{Z}}$ *be a finite* $K_{\mathrm{z}}(T-1) \times K_{\mathrm{z}}(T-1)$ *matrix such that* $\mathbb{J}'\boldsymbol{\Sigma}_{\mathrm{Z}}\mathbb{J} > 0$ *and* $\mathbb{J}_{t-1}'\boldsymbol{\Sigma}_{\mathrm{Z}}\mathbb{J}_{t-1} > 0$ *for every* $2 \le t \le T$*. Then*

$$\frac{\mathbf{Z}'\mathbf{Z}}{N} \to_p \boldsymbol{\Sigma}_{\mathrm{Z}}.$$

(iii) *Let* $\boldsymbol{\Sigma}_{\mathrm{ZB}}$ *be a finite* $K_{\mathrm{z}}(T-1) \times K_{\mathrm{f}}$ *matrix. Then*

$$\frac{\mathbf{Z}'\mathbf{B}}{N} \to_p \boldsymbol{\Sigma}_{\mathrm{ZB}}.$$

(iv) *Setting* $\boldsymbol{\mu}_{\mathrm{u}_i} = \mathrm{E}[\mathbf{u}_i]$*, we require*

$$\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\mu}_{\mathrm{u}_i} = o\left(\frac{1}{\sqrt{N}}\right).$$

(v) *Setting* $\boldsymbol{\Sigma}_{\mathrm{u},ij} = \mathrm{Cov}[\mathbf{u}_i, \mathbf{u}_j]$*, for* $i, j = 1, \ldots, N$*, we require*

$$\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\Sigma}_{\mathrm{u},ii} \to \boldsymbol{\Sigma}_{\mathrm{U}} = (\sigma^2 \mathbf{I}_{T-1} \otimes \boldsymbol{\Sigma}_{\mathrm{Z}}), \quad \textit{and} \quad \sum_{i,j=1}^{N} \boldsymbol{\Sigma}_{\mathrm{u},ij}\, \mathbb{1}_{i\neq j} = o(N).$$

(vi) *For any* $i, j = 1, \ldots, N$*,*

$$\mathrm{Cov}\big[\mathbf{z}_{i,t}, \boldsymbol{\epsilon}_j' \otimes \boldsymbol{\epsilon}_j'\big] = \mathbf{0}_{K_{\mathrm{z}}\times(T-1)^2}, \qquad \mathrm{Cov}\big[\boldsymbol{\epsilon}_i, \boldsymbol{\epsilon}_j' \otimes (\boldsymbol{u}_j - \mathrm{E}[\boldsymbol{u}_j])'\big] = \mathbf{0}_{(T-1)\times K_{\mathrm{z}}(T-1)^3}.$$

(vii) *Let* $\boldsymbol{\Sigma}_{\mathrm{U}}$ *be a finite* $K_{\mathrm{z}}(T-1)^2 \times K_{\mathrm{z}}(T-1)^2$ *matrix. Then*

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\mathbf{u}_i - \boldsymbol{\mu}_{\mathrm{u}_i}) \to_d \mathcal{N}\big(\mathbf{0}_{K_{\mathrm{z}}(T-1)^2}, \boldsymbol{\Sigma}_{\mathrm{U}}\big).$$

(viii) *Setting* $\boldsymbol{\Sigma}_{\mathrm{u}\epsilon,ij} = \mathrm{Cov}[\boldsymbol{\epsilon}_i \otimes \boldsymbol{\epsilon}_i, \mathbf{u}_j']$*, we require*

$$\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\Sigma}_{\mathrm{u}\epsilon,ii} \to \mathbf{0}_{(T-1)^2\times K_{\mathrm{z}}(T-1)^2}, \quad \textit{and} \quad \frac{1}{N}\sum_{i,j=1}^{N} \boldsymbol{\Sigma}_{\mathrm{u}\epsilon,ij} \to \mathbf{0}_{(T-1)^2\times K_{\mathrm{z}}(T-1)^2}.$$

(ix) *For* $i, j = 1, \ldots, N$*,*

$$\frac{1}{N^2}\sum_{i,j=1}^{N} \mathrm{Cov}\big[\mathbf{u}_i \otimes \mathbf{u}_i,\, \mathbf{u}_j' \otimes \mathbf{u}_j'\big] \to \mathbf{0}_{K_{\mathrm{z}}^2(T-1)^4\times K_{\mathrm{z}}^2(T-1)^4}.$$

(x) *Let* $\mathbb{P}_{\tilde{Z}_i} = \tilde{\mathbf{Z}}_i(\tilde{\mathbf{Z}}_i'\tilde{\mathbf{Z}}_i)^{-1}\tilde{\mathbf{Z}}_i'$, *with generic element* $\mathbb{p}_{i,ts}$ *for* $t, s = 1, \ldots, T-1$, *where* $\tilde{\mathbf{Z}}_i = \mathbb{M}_{1_{T-1}}\mathbf{Z}_i$. *Then for every* $1 \leq t+1, s+1, v_a, u_a \leq (T-1)$ *with* $a = 1, \ldots, 4$, *the following conditions hold:*

$$(x.1)\ \frac{1}{N}\sum_{i=1}^{N}\mathbb{P}_{\tilde{Z}_i} \to_p \mathbb{P}_{\tilde{Z}},$$

$$(x.2)\ \frac{1}{N}\sum_{i=1}^{N}(\mathbb{P}_{\tilde{Z}_i} \odot \mathbb{P}_{\tilde{Z}_i}) \to_p \mathbb{P}_{\tilde{Z}}^{(2)},$$

$$(x.3)\ \frac{1}{N}\sum_{i=1}^{N}\mathbb{P}_{\tilde{Z}_i}\left(\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i' - \sigma_i^2\mathbf{I}_{T-1}\right) = \mathbb{P}_{\tilde{Z}}\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i' - \sigma_i^2\mathbf{I}_{T-1}\right) + o_p\left(\frac{1}{\sqrt{N}}\right),$$

$$(x.4)\ \frac{1}{N^2}\sum_{i,j=1}^{N}\kappa_4\left[\prod_{a=1}^{4}\mathbb{p}_{i,t-1u_a}, \prod_{a=1}^{4}\mathbb{p}_{j,s-1v_a}, \prod_{a=1}^{4}\epsilon_{i,u_a+1}, \prod_{a=1}^{4}\epsilon_{j,v_a+1}\right] = o(1),$$

$$(x.5)\ \frac{1}{N^2}\sum_{i,j=1}^{N}\kappa_3\left[\prod_{a=1}^{4}\mathbb{p}_{i,t-1u_a}, \prod_{a=1}^{4}\mathbb{p}_{j,s-1v_a}, \prod_{a=1}^{4}\epsilon_{i,u_a+1}\right] = o(1),$$

$$(x.6)\ \frac{1}{N^2}\sum_{i,j=1}^{N}\kappa_3\left[\prod_{a=1}^{4}\mathbb{p}_{i,t-1u_a}, \prod_{a=1}^{4}\epsilon_{i,u_a+1}, \prod_{a=1}^{4}\epsilon_{j,v_a+1}\right] = o(1),$$

$$(x.7)\ \frac{1}{N^2}\sum_{i,j=1}^{N}\mathrm{Cov}\left[\prod_{a=1}^{4}\mathbb{p}_{i,t-1u_a}, \prod_{a=1}^{4}\epsilon_{j,v_a+1}\right] = o(1),$$

$$(x.8)\ \frac{1}{N^2}\sum_{i,j=1}^{N}\mathrm{Cov}[\mathbb{p}_{j,su_1}\mathbb{p}_{i,tv_1}, \epsilon_{i,t+1}\epsilon_{j,s+1}\epsilon_{iu_1+1}\epsilon_{jv_1+1}] = o(1),$$

$$(x.9)\ \frac{1}{N}\sum_{i=1}^{N}\mathrm{Cov}\left[\prod_{a=1}^{4}\mathbb{p}_{i,t-1u_a}, \prod_{a=1}^{4}\epsilon_{i,v_a+1}\right] = o(1).$$

where $\kappa_3[\cdot,\cdot,\cdot]$ *and* $\kappa_4[\cdot,\cdot,\cdot,\cdot]$ *denote mixed cumulants of order 3 and 4, respectively.*

(xi) *For every* $3 \leq h \leq 8$, *the following condition holds:*

$$\sup_{i_1}\sum_{i_2,\ldots,i_h=1}^{N}\left|\kappa_{h,i_1i_2\ldots i_h}^{\mathbb{p}}\right| = o(N),$$

for at least one index i_j $(2 \leq j \leq h)$ *different from* i_1, *where* $\kappa_{h,i_1i_2\ldots i_h}^{\mathbb{p}}$ *is the mixed cumulant of* $\{\mathbb{p}_{i_1,t_1-1u_1}, \mathbb{p}_{i_2,t_2-1u_2}, \ldots, \mathbb{p}_{i_h,t_h-1u_h}\}$ *of order* h, *for every* $2 \leq t_1, \ldots, t_h, u_1, \ldots, u_h \leq T$.

Remark OA.8. *By Assumption OA.8(iv), we have*

$$\frac{\mathbf{Z}'\boldsymbol{\epsilon}'}{N} \to_p \boldsymbol{\Sigma}_{Z\epsilon} = \mathbf{0}_{K_z(T-1)\times(T-1)}.$$

Moreover, Assumption OA.8(vii) implies that

$$\sqrt{N}\,\mathrm{vec}\left(\frac{\mathbf{Z}'\boldsymbol{\epsilon}'}{N}\right) = \sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{\epsilon}_i \otimes \mathbf{z}_i)\right) \;\; \to_d \;\; \mathcal{N}\big(\mathbf{0}_{K_{\mathrm{z}}(T-1)^2},\, \boldsymbol{\Sigma}_{\mathrm{U}}\big)\,.$$

Remark OA.9. *Notice that, under our assumptions,*

$$\frac{1}{N}\bar{\mathbf{Z}}'\mathbf{1}_N = \frac{1}{N}\mathbb{J}'\mathbf{Z}'\mathbf{1}_N \;\; \to_p \;\; \boldsymbol{\mu}_{\mathrm{z}}, \qquad \frac{1}{N}\mathbb{J}_{t-1}'\mathbf{Z}'\mathbf{1}_N \;\; \to_p \;\; \boldsymbol{\mu}_{\mathrm{z}}.$$

Both estimators therefore share the same limiting behavior, although the former is a more efficient estimator of $\boldsymbol{\mu}_{\mathrm{z}}$.

Remark OA.10. *Primitive conditions for Assumption OA.8 can be readily obtained. For instance, one may assume that the* $K_{\mathrm{z}} \times 1$ *vector of anomalies* $\mathbf{z}_{it}$ *follows a linear process of the form*

$$\mathbf{z}_{it} = \boldsymbol{\mu}_{iz} + \sum_{k=0}^{\infty}\boldsymbol{\Delta}_{ik}\boldsymbol{\eta}_{i,t-k}, \tag{OA.66}$$

with innovations $\boldsymbol{\eta}_{i,t} \sim (\mathbf{0}_{K_{\mathrm{z}}}, \boldsymbol{\Sigma}_{\eta})$ *i.i.d. across time, and such that* $N^{-1}\sum_{i=1}^{N}\big(\sum_{k=0}^{\infty}\|\boldsymbol{\Delta}_{ik}\|\big) \le C < \infty$, $\boldsymbol{\Delta}_{i0} = \mathbf{I}_{K_{\mathrm{z}}}$, $N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\mu}_{iz}\| \le C < \infty$, *and*

$$\mathrm{E}[\boldsymbol{\eta}_{i,t}\boldsymbol{\epsilon}_t'] \neq \mathbf{0}_{K_{\mathrm{z}}\times N} \qquad \text{for every } 1 \le i \le N.$$

The specification in (OA.66) *represents a standard assumption in time-series econometrics. In this case, the matrices* $\boldsymbol{\Sigma}_{\mathrm{ZZ}}$ *and* $\boldsymbol{\Sigma}_{\mathrm{ZB}}$ *are constant, with their* (a, b) *entries being general functions of* $(a - b)$.[33]

Furthermore, one can easily allow for cyclical and trending behaviors (either deterministic or stochastic) in $\mathbf{z}_{it}$ *as in* (OA.66). *This is particularly relevant for firms' characteristics such as value and size, which typically display significant time variation. In this case, the limit of the cross-sectional averages of the anomaly variables may no longer be constant across time, a situation our methodology can still handle, albeit at the cost of heavier notation and formalism.*

[33] One can further generalize (OA.66) to allow for a (dynamic) factor defined? structure of the form $\mathbf{z}_{it} = \boldsymbol{\mu}_{iz} + \sum_{k=0}^{\infty}\boldsymbol{\Delta}_{ik}\boldsymbol{\eta}_{i,t-k} + \sum_{k=0}^{\infty}\Delta_{ik}^{\dagger}\boldsymbol{\eta}_{t-k}$, where $\boldsymbol{\eta}_t = (\eta_t^1, \ldots, \eta_t^{K_{\mathrm{z}}})' \sim (\mathbf{0}_{K_{\mathrm{z}}}, \boldsymbol{\Sigma}_{\eta^{\dagger}})$ is an i.i.d. sequence of *common* shocks. In this case, the matrices $\boldsymbol{\Sigma}_{\mathrm{ZZ}}$ and $\boldsymbol{\Sigma}_{\mathrm{ZB}}$ may contain random components, since they depend on terms such as $\eta_t^k\eta_s^{k'}$ for $k, k' = 1, \ldots, K_{\mathrm{z}}$ and $t, s = 1, \ldots, T-1$, with generic (t, s) elements not depending only on $(t - s)$. Despite this lack of stationarity, our results continue to hold (conditionally on $\boldsymbol{\eta}_t$) thanks to the fixed-T assumption. The asymptotic distribution of the estimators is mixed-normal, implying that the test statistics retain the conventional chi-square distribution (by standardization), and all testing procedures remain feasible. Details are available upon request.

Remark OA.11. *Assumption OA.8 imposes basic regularity conditions on the behavior of the matrix* $\mathbf{Z}$ *and controls the degree of cross-sectional dependence between anomalies and excess returns, which gradually dissipates as* N *increases.*

To illustrate the degree of cross-sectional dependence permitted by our theory, we run a Monte Carlo experiment using the data-generating process described above, with the error term following

$$\epsilon_{it} = \frac{\sigma}{\tau}\left(u_{it} + \frac{\mathbf{1}'_{K_z}\boldsymbol{\eta}_{it}}{N^{\delta}}\right), \tag{OA.67}$$

where $\tau = 1 + \frac{K_z}{N^{2\delta}}$, u_{it} *is i.i.d. standard normal,* σ^2 *is calibrated as in the uncorrelated case, and the* $K_z \times 1$ *vector* $\boldsymbol{\eta}_{it}$ *is calibrated using the standardized residuals from a VAR(1) fitted to the two anomalies. The parameter* δ *controls the strength of cross-sectional correlation between the shocks and the anomalies: higher values of* δ *imply weaker correlation. For our theoretical results, we require* $\delta \geq 0.5$.

We consider the case of one anomaly ($K_z = 1$) *and report the sample cross-sectional correlation between return innovations* (ϵ_{it}) *and anomaly innovations* (η_{it})*, averaged over the time dimension* ($T = 36$ *or* $T = 72$)*, for different values of* δ*. Specifically, we compute*

move to right

$$\rho(\delta, N) = \frac{1}{T}\sum_{t=1}^{T}\frac{\sum_{i=1}^{N}\epsilon_{it}\eta_{it}}{\left(\sum_{i=1}^{N}\epsilon_{it}^2\right)^{1/2}\left(\sum_{i=1}^{N}\eta_{it}^2\right)^{1/2}}, \quad \tau(\delta, N) = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{\sum_{i=1}^{N}\epsilon_{it}\eta_{it}}{\left(\sum_{i=1}^{N}\epsilon_{it}^2\right)^{1/2}\left(\sum_{i=1}^{N}\eta_{it}^2\right)^{1/2}}\right| \tag{OA.68}$$

Table OA.II reports results for $\delta = \{0, 0.10, 0.25, 0.50, 0.75, 1.00\}$ *and* $T = 36$ *(Panel A) and* $T = 72$ *(Panel B). By* (OA.120)*, cross-correlation is inversely related to* δ*: our theory requires* $\delta > 0$ *for consistency and* $\delta \geq 0.5$ *for asymptotic normality. Table OA.II shows that the setup admits sizable dependence between anomalies and returns. For example, when* $\delta = 0.50$*, both* ρ *and* τ *equal 0.67 when* $N = 100$ *and 0.37 when* $N = 500$*, a substantial correlation even for large* N*. As expected, the correlation weakens as both* N *and* δ *increase.*

Table OA.II: Conditional average cross-correlations between anomalies and asset returns.

	$\rho(\delta, N)$						$\tau(\delta, N)$					
δ	0.00	0.10	0.25	0.50	0.75	1.00	0.00	0.10	0.25	0.50	0.75	1.00
Panel A: $T = 36$												
$N = 10$	0.993	0.990	0.980	0.941	0.835	0.631	0.993	0.990	0.980	0.941	0.835	0.642
$N = 100$	0.993	0.984	0.942	0.669	0.279	0.097	0.993	0.984	0.942	0.669	0.279	0.121
$N = 500$	0.993	0.978	0.883	0.369	0.083	0.017	0.993	0.978	0.883	0.369	0.085	0.041
$N = 1000$	0.993	0.976	0.848	0.282	0.060	0.019	0.993	0.976	0.848	0.282	0.060	0.028
Panel B: $T = 72$												
$N = 10$	0.993	0.989	0.979	0.939	0.838	0.656	0.993	0.989	0.979	0.939	0.838	0.656
$N = 100$	0.993	0.984	0.942	0.663	0.266	0.082	0.993	0.984	0.942	0.663	0.266	0.102
$N = 500$	0.993	0.979	0.885	0.372	0.083	0.016	0.993	0.979	0.885	0.372	0.084	0.036
$N = 1000$	0.994	0.976	0.849	0.273	0.046	0.007	0.994	0.976	0.849	0.273	0.052	0.029

Assumption OA.9 (*Smoothness of the premia parameters*)**.** *Define the $(T-1) \times K_{\mathrm{f}}$ matrix*

$$\breve{\boldsymbol{\delta}}_{\mathrm{f}} = \left(\breve{\boldsymbol{\delta}}_{\mathrm{f},1}, \ldots, \breve{\boldsymbol{\delta}}_{\mathrm{f},T-1}\right)', \quad \text{with} \quad \breve{\boldsymbol{\delta}}_{\mathrm{f},t-1} = \boldsymbol{\delta}_{\mathrm{f},t-1} - \mathbf{f}_t = \boldsymbol{\gamma}_{\mathrm{f},t-1} - E\big(\mathbf{f}_t \mid I_{t-1}, \boldsymbol{\Pi}\big).$$

Also denote the $(T-1) \times N$ matrix

$$\boldsymbol{\Delta}_z = \begin{bmatrix} \boldsymbol{\gamma}'_{z,1} - \boldsymbol{\gamma}'_{\mathrm{z}} & \mathbf{0}'_{K_{\mathrm{z}}} & \cdots & \mathbf{0}'_{K_{\mathrm{z}}} \\ \mathbf{0}'_{K_{\mathrm{z}}} & \boldsymbol{\gamma}'_{z,2} - \boldsymbol{\gamma}'_{\mathrm{z}} & \cdots & \mathbf{0}'_{K_{\mathrm{z}}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}'_{K_{\mathrm{z}}} & \mathbf{0}'_{K_{\mathrm{z}}} & \cdots & \boldsymbol{\gamma}'_{\mathrm{z},T-1} - \boldsymbol{\gamma}'_{\mathrm{z}} \end{bmatrix} \begin{bmatrix} \mathbf{Z}'_1 \\ \mathbf{Z}'_2 \\ \vdots \\ \mathbf{Z}'_{T-1} \end{bmatrix},$$

for some constant $K_{\mathrm{z}} \times 1$ vector $\boldsymbol{\gamma}_{\mathrm{z}}$ satisfying

$$\frac{1}{N} \sum_{i=1}^{N} (\mathbf{Z}'_i \mathbf{Z}_i)^{-1} \mathbf{Z}'_i \mathbf{R}_i \ \to_p \ \boldsymbol{\gamma}_{\mathrm{z}}.$$

Then the following hold:

$$\mathbf{P}'\boldsymbol{\gamma}_0 = \mathbf{0}_{K_{\mathrm{f}}}, \qquad \mathbf{P}'\,\breve{\boldsymbol{\delta}}_{\mathrm{f}} = \mathbf{0}_{K_{\mathrm{f}} \times K_{\mathrm{f}}}, \qquad \frac{\mathbf{P}'\boldsymbol{\Delta}_z\,\boldsymbol{\epsilon}}{N} = o_p\left(\frac{1}{\sqrt{N}}\right).$$

Remark OA.12. *When a risk-free asset is tradable, Assumption OA.9 reduces to the single condition*

$$\frac{\mathbf{P}'\boldsymbol{\Delta}_z\boldsymbol{\epsilon}}{N} = o_p\left(\tfrac{1}{\sqrt{N}}\right),$$

since the first two conditions (on the zero-beta rate and the factor risk premia) are no longer required.

OA.3.2 Additional assumptions required for the WLS estimation

In this section, we introduce the additional assumptions needed for the validity of the WLS estimation described in Section OA.2.2. Before stating the main assumptions, it is useful to introduce some preliminary notation.

For each asset i, let $\mathbf{w}_{i\cdot} = (w_{i,1}, \ldots, w_{i,T-1})'$ denote the $(T-1) \times 1$ vector of weights, and let $\mathbf{w}_{\cdot,t-1} = (w_{1,t-1}, \ldots, w_{N,t-1})'$ denote the $N \times 1$ vector of weights at time $(t-1)$, for every $2 \leq t \leq T$. Collecting all weights across assets and time yields the $N \times (T-1)$ matrix

$$\mathbf{W} = (\mathbf{w}_{\cdot,1}, \ldots, \mathbf{w}_{\cdot,T-1}) = (\mathbf{w}_{1\cdot}, \ldots, \mathbf{w}'_{N\cdot})'.$$

Assumption OA.10. *(CSR WLS weights)*

(i) *As $N \to \infty$,*

$$\frac{\mathbf{1}'_N \mathbf{W}_{t-1} \mathbf{1}_N}{N} \quad \to_p \quad 1.$$

(ii) *For any real number $h > 1$, as $N \to \infty$,*

$$\frac{\mathbf{1}'_N \mathbf{W}^h_{t-1} \mathbf{1}_N}{N} \quad \to_p \quad \mu^h_{w,t-1}.$$

(iii) *As $N \to \infty$,*

$$\frac{1}{N} \sum_{i=1}^{N} \mathbf{w}_{i\cdot} \mathbf{w}'_{i\cdot} \quad \to_p \quad \mathbf{\Sigma}_{\mathrm{W}}.$$

Remark OA.13. *By Assumption OA.10, granularity follows because $N^{-1} \mathbf{1}'_N \mathbf{W}_{t-1} \mathbf{1}_N \to_p 1$ by Assumption OA.10(i), while $N^{-2} \mathbf{1}'_N \mathbf{W}^2_{t-1} \mathbf{1}_N = O(N^{-1}) = o(1)$ by Assumption OA.10(ii).*

Remark OA.14. *An important example of commonly used weights satisfying Assumption OA.10 is given by*

$$w_{i,t} = \frac{N\, w^{\$}_{i,t}}{\sum_{j=1}^{N} w^{\$}_{j,t}}, \tag{OA.69}$$

where $w^{\$}_{i,t}$ denotes the dollar value of the market capitalization of stock i at time t. The multiplication by N in the numerator of (OA.69) *is simply a normalization factor, ensuring that Assumption OA.10(i) holds, i.e., that the sample average of the weights converges to 1 as $N \to \infty$.*

Assumption OA.11 (*Weighted risk exposures*)**.** *Let* $\mathbf{W}_{t-1}$ *satisfy Assumption OA.10 and let the risk exposures* $\boldsymbol{\beta}_i$ *be a non-random sequence. As* $N \to \infty$*,*

$$\frac{1}{N}\mathbf{B}'\mathbf{W}_{t-1}\mathbf{1}_N \ \ \to_p \ \ \boldsymbol{\mu}_\beta, \qquad \frac{1}{N}\mathbf{B}'\mathbf{W}_{t-1}\mathbf{B} \ \ \to_p \ \ \boldsymbol{\Sigma}_\beta, \tag{OA.70}$$

such that

$$\boldsymbol{\Sigma}_{\mathrm{X}} = \begin{bmatrix} 1 & \boldsymbol{\mu}_\beta' \\ \boldsymbol{\mu}_\beta & \boldsymbol{\Sigma}_\beta \end{bmatrix} > 0. \tag{OA.71}$$

Remark OA.15. *Assumption OA.11 generalizes Assumption OA.4 to the case of weighted averages. Notice that, owing to the granularity condition, the weighted averages in* (OA.70) *converge to the same limits as their unweighted counterparts in Assumption OA.4.*

Assumption OA.12. *(Weighted moments and cumulants)*

(i) *Let* $0 < \sigma^2 < \infty$*. Then, for every* $2 \leq t \leq T$*,*

$$\frac{1}{N}\sum_{i=1}^{N} w_{i,t-1}\big(\sigma_i^2 - \sigma^2\big) = o_p\!\left(\frac{1}{\sqrt{N}}\right). \tag{OA.72}$$

(ii) *For every* $2 \leq t \leq T$*,*

$$\sum_{i,j=1}^{N} w_{i,t-1}\,|\sigma_{ij}|\,\mathbb{1}_{\{i\neq j\}} = o_p(N). \tag{OA.73}$$

(iii) *Let* $0 < \mu_4 < \infty$*, and let* $\mu_{4i} = \mathrm{E}[\epsilon_{it}^4]$*. Then, for every* $2 \leq t \leq T$*,*

$$\frac{1}{N}\sum_{i=1}^{N} w_{i,t-1}\mu_{4i} \ \to_p \ \mu_4. \tag{OA.74}$$

(iv) *Let* $0 < \sigma_4 < \infty$*. Then, for every* $2 \leq t \leq T$*,*

$$\frac{1}{N}\sum_{i=1}^{N} w_{i,t-1}\sigma_i^4 \ \to_p \ \sigma_4. \tag{OA.75}$$

(v) *Let* $\kappa_3(a,b,c)$ *denote the third-order cumulant of random variables* a,b,c*. Then*

$$\kappa_3[\epsilon_{i,t},\epsilon_{j,s},w_{j,h}] = 0, \qquad \kappa_3[\epsilon_{i,t},\epsilon_{j,s},\mathbf{z}_{j,h}] = \mathbf{0}_{K_{\mathrm{z}}}. \tag{OA.76}$$

(vi) *Let* $\kappa_{4,iiii} = \kappa_4[\epsilon_{i,t},\epsilon_{i,t},\epsilon_{i,t},\epsilon_{i,t}]$ *denote the fourth-order cumulant of* $\epsilon_{i,t}$*. Then, for some* $0 \leq |\kappa_4| < \infty$ *and for every* $2 \leq t \leq T$*,*

$$\frac{1}{N}\sum_{i=1}^{N} w_{i,t-1}\kappa_{4,iiii} \ \to_p \ \kappa_4. \tag{OA.77}$$

(vii) *For every $3 \leq h \leq 8$, the following mixed cumulants of order h satisfy*

$$\sup_{i_1} \sum_{i_2,\ldots,i_h=1}^{N} |\kappa_{h,w_{i_1,t-1},i_2,\ldots,i_h}| = o(N), \tag{OA.78}$$

and

$$\sup_{i_1} \sum_{i_2,\ldots,i_h=1}^{N} |\kappa_{h,w_{i_1,t-1},\mathbf{z}_{i_2,r},i_3,\ldots,i_h}| = o(N), \tag{OA.79}$$

for at least one index i_j ($2 \leq j \leq h$) different from i_1. Here $\kappa_{h,w_{i_1,t-1},i_2,\ldots,i_h}$ denotes the mixed cumulant of order h in $\{w_{i_1,t-1}, \epsilon_{i_2,s}, \ldots, \epsilon_{i_h,s}\}$, while $\kappa_{h,w_{i_1,t-1},\mathbf{z}_{i_2,r},i_3,\ldots,i_h}$ denotes the mixed cumulant of order h in $\{w_{i_1,t-1}, \mathbf{z}_{i_2,r}, \epsilon_{i_3,s}, \ldots, \epsilon_{i_h,s}\}$.

Remark OA.16. *Assumption OA.12 extends Assumption OA.6 to the case of weighted averages. Notice that, due to the granularity assumption in Assumption OA.10, all the weighted averages in* (OA.72)–(OA.79) *the same limits of the corresponding un-weighted averages in Assumption OA.6.*

Assumption OA.13 (*Weighted moments and CLT of anomalies*)**.** *Define the $(T-1)^2 \times 1$ vector $\mathbf{v}_i = \boldsymbol{\epsilon}_i \otimes \mathbf{w}_i$ and the corresponding $N \times (T-1)^2$ matrix $\mathbf{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_N)'$, such that $\mathrm{E}[\mathbf{v}_i] = \boldsymbol{\mu}_{\mathrm{v}_i} < \infty$, and $\boldsymbol{\Sigma}_{\mathrm{v},ij} = \mathrm{Cov}[\mathbf{v}_i, \mathbf{v}_j]$. Then:*

(i) *As $N \to \infty$,*

$$\frac{\boldsymbol{\epsilon}\left(\mathbf{W}_{t-1} - \mathrm{E}[\mathbf{W}_{t-1}]\right)\boldsymbol{\epsilon}'}{N} \quad \to_p \quad \mathbf{0}_{(T-1)\times(T-1)}.$$

(ii) *As $N \to \infty$,*

$$\frac{\mathbf{Z}_{t-1}'\mathbf{W}_{t-1}\mathbf{1}_N}{N} \quad \to_p \quad \boldsymbol{\mu}_{\mathrm{z},t-1}, \qquad \frac{\mathbf{Z}_{t-1}'\mathbf{W}_{t-1}\mathbf{Z}_{t-1}}{N} \quad \to_p \quad \boldsymbol{\Sigma}_{\mathrm{Z},t-1}.$$

(iii) *Let $\boldsymbol{\Sigma}_{\mathrm{ZW}}$ be a finite $K_{\mathrm{z}}(T-1) \times (T-1)$ matrix. Then, as $N \to \infty$,*

$$\frac{\mathbf{Z}'\mathbf{W}}{N} \quad \to_p \quad \boldsymbol{\Sigma}_{\mathrm{ZW}}.$$

(iv) *As $N \to \infty$,*

$$\frac{1}{N}\left(\mathbf{Z}_{t-1} - \mathrm{E}[\mathbf{Z}_{t-1}]\right)'\left(\mathbf{W}_{t-1} - \mathrm{E}[\mathbf{W}_{t-1}]\right)\boldsymbol{\epsilon}' \quad \to_p \quad \mathbf{0}_{K_{\mathrm{z}}\times(T-1)}.$$

(v) *As $N \to \infty$,*

$$\frac{1}{N}\left(\mathbf{Z}_{t-1} - \mathrm{E}[\mathbf{Z}_{t-1}]\right)'\mathbf{W}_{t-1}\boldsymbol{\epsilon}' - \frac{1}{N}\left(\mathbf{Z}_{t-1} - \mathrm{E}[\mathbf{Z}_{t-1}]\right)'\boldsymbol{\epsilon}' = o_p\left(N^{-1/2}\right).$$

(vi) *As $N \to \infty$,*

$$\frac{\mathbf{X}'\mathbb{M}_{1_N}\mathbf{V}}{N} = o_p\Big(N^{-1/2}\Big), \qquad \frac{\mathbf{Z}'\mathbb{M}_{1_N}\mathbf{V}}{N} = o_p\Big(N^{-1/2}\Big),$$

where $\mathbb{M}_{1_N}$ denotes the projection matrix $\mathbb{I}_N - N^{-1}\mathbf{1}_N\mathbf{1}_N'$.

(vii) *As $N \to \infty$,*

$$\frac{1}{N}\sum_{i=1}^{N}\mathbf{\Sigma}_{\mathrm{v},ii} \quad\to\quad \mathbf{\Sigma}_{\mathrm{V}} = (\sigma^2\mathbf{I}_{T-1}\otimes\mathbf{\Sigma}_{\mathrm{W}}), \qquad \sum_{i=1}^{N}\mathbf{\Sigma}_{\mathrm{v},ij}\,\mathbb{1}_{i\neq j} = o(N).$$

(viii) *As $N \to \infty$,*

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\mathbf{v}_i - \boldsymbol{\mu}_{\mathrm{v}_i}) \quad\to_d\quad \mathcal{N}\big(\mathbf{0}_{(T-1)^2},\mathbf{\Sigma}_{\mathrm{V}}\big), \qquad \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\mu}_{\mathrm{v}_i} = o(N^{-1/2}).$$

(ix) *As $N \to \infty$,*

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\mathbf{z}_i - \boldsymbol{\mu}_{\mathrm{z}_i}) \quad\to_d\quad \mathcal{N}(\mathbf{0}_{K_z(T-1)},\mathbf{\Sigma}_{ZZ}), \qquad \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{\mu}_{\mathrm{z}_i} - \boldsymbol{\mu}_z) = o(N^{-1/2}).$$

Remark OA.17. *Assumption OA.13 establishes the limits of several sample moments involving anomalies, weights, and asset-specific errors, as well as their mutual correlation across time. Specifically, we assume that each asset-specific error is uncorrelated with past values of both anomaly variables and weights, but each $\boldsymbol{\epsilon}_s$ may be correlated with contemporaneous or future values of $\mathbf{Z}_t$ and $\mathbf{W}_v$ whenever $s \le t$ or $s \le v$. This implies that, despite the granularity condition, the limits of certain quantities will depend on the weights, since these may be correlated with both anomalies and asset-specific error components.*

Remark OA.18. *To simplify the analysis, we also impose zero mixed third-moment conditions. This sometimes requires de-meaning the anomaly variables, which in general need not have zero mean. The main moment conditions are summarized in the following assumption.*

Remark OA.19. *Assumption (ix) represents a strengthening of our earlier assumptions, necessitated by the additional challenges introduced by the weighted estimator.*

OA.3.3 Additional assumptions required for the cross-sectional R-squared test

In this Section, we introduce additional assumptions that are required to derive the R-squared test described in Section OA.2.4.

Assumption OA.14.

(i) *As* $N \to \infty$,

$$\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta = o\Big(N^{-1/2}\Big), \qquad \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\beta}_i\boldsymbol{\beta}_i' - \boldsymbol{\Sigma}_\beta = o\Big(N^{-1/2}\Big).$$

(ii) *As* $N \to \infty$,

$$\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left[(\mathbf{z}_i \otimes \mathbf{z}_i) - \mathrm{vec}(\boldsymbol{\Sigma}_{\mathrm{Z}})\right] \;\to_d\; \mathcal{N}(\mathbf{0}_{K_{\mathrm{z}}^2}, \mathbf{U}_{\mathrm{Z}}), \quad \textit{with} \\
&\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}[(\mathbf{z}_i \otimes \mathbf{z}_i) - \mathrm{vec}(\boldsymbol{\Sigma}_{\mathrm{Z}})] = o\Big(N^{-1/2}\Big), \\
&\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}[(\mathbf{z}_i \otimes \mathbf{z}_i) - \mathrm{vec}(\boldsymbol{\Sigma}_{\mathrm{Z}})]\,[(\mathbf{z}_i \otimes \mathbf{z}_i) - \mathrm{vec}(\boldsymbol{\Sigma}_{\mathrm{Z}})]' \;\to\; \mathbf{U}_{\mathrm{Z}}, \\
&\sum_{i,j=1,\, i\neq j}^{N}\mathrm{E}[(\mathbf{z}_i \otimes \mathbf{z}_i) - \mathrm{vec}(\boldsymbol{\Sigma}_{\mathrm{Z}})]\,[(\mathbf{z}_j \otimes \mathbf{z}_j) - \mathrm{vec}(\boldsymbol{\Sigma}_{\mathrm{Z}})]' = o(N), \quad \textit{and} \\
&\frac{1}{N}\sum_{i,j=1}^{N}\mathrm{Cov}\big[(\mathbf{z}_i \otimes \mathbf{z}_i), \mathbf{z}_j'\big] \;\to\; \boldsymbol{\Sigma}_{\mathrm{z}\otimes\mathrm{z}}.
\end{aligned}$$

(iii) *As* $N \to \infty$,

$$\sqrt{N}\left(\frac{\mathbf{Z}'\mathbf{1}_N}{N} - \boldsymbol{\mu}_{\mathrm{z},T-1}\right) \;\to_d\; \mathcal{N}\big(\mathbf{0}_{K_{\mathrm{z}}(T-1)}, \boldsymbol{\Sigma}_{\mathrm{Z}} - \boldsymbol{\mu}_{\mathrm{z},T-1}\boldsymbol{\mu}_{\mathrm{z},T-1}'\big).$$

(iv) *As* $N \to \infty$,

$$\frac{1}{N}\sum_{i,j=1}^{N}\mathrm{Cov}\big((\mathbf{z}_i \otimes \mathbf{z}_i), (\boldsymbol{\epsilon}_j' \otimes \boldsymbol{\epsilon}_j')\big) \;\to\; \boldsymbol{\Sigma}_{\mathrm{Z}\otimes\epsilon} = \mathbf{0}_{\left((T-1)K_z\right)^2 \times (T-1)^2}.$$

(v) *As* $N \to \infty$,

$$\frac{1}{N}\sum_{i,j=1}^{N}\mathrm{Cov}\big((\mathbf{z}_i \otimes \mathbf{z}_i), \mathbf{u}_j'\big) \;\to\; \boldsymbol{\Sigma}_{\mathrm{ZU}} = \mathbf{0}_{\left((T-1)K_z\right)^2 \times (T-1)^2 K_z}.$$

OA.4 Preliminary Lemmas

In this section, we establish several results needed to derive the asymptotic properties of the new CSR OLS estimators. All results below hold as $N \to \infty$. Proofs are omitted to streamline the exposition, but are available upon request.

Lemma 1. *Under Assumptions OA.9 and OA.3–OA.8, we have:*

(i)
$$\hat{\sigma}^2 - \sigma^2 = O_p\left(\tfrac{1}{\sqrt{N}}\right).$$

(ii)
$$\hat{\mathbf{X}}'\hat{\mathbf{X}} = O_p(N).$$

(iii)
$$\hat{\boldsymbol{\Sigma}}_{\mathrm{X}} \quad \to_p \quad \boldsymbol{\Sigma}_{\mathrm{X}} + \boldsymbol{\Lambda}_1.$$

(iv)
$$\frac{(\hat{\mathbf{X}} - \mathbf{X})'(\hat{\mathbf{X}} - \mathbf{X})}{N} \quad \to_p \quad \boldsymbol{\Lambda}_1.$$

(v)
$$\mathbf{X}'\boldsymbol{\epsilon}_t = O_p\left(\sqrt{N}\right), \qquad \mathbf{X}'\bar{\boldsymbol{\epsilon}} = O_p\left(\sqrt{N}\right).$$

(vi)
$$(\hat{\mathbf{X}} - \mathbf{X})'\mathbf{X}\boldsymbol{\Gamma}_{\mathrm{f},t} = O_p\left(\sqrt{N}\right), \qquad (\hat{\mathbf{X}} - \mathbf{X})'\mathbf{X}\bar{\boldsymbol{\Gamma}}_{\mathrm{f}} = O_p\left(\sqrt{N}\right).$$

(vii)
$$(\hat{\mathbf{X}} - \mathbf{X})'\bar{\boldsymbol{\epsilon}} = O_p\left(\sqrt{N}\right).$$

(viii)
$$(\hat{\mathbf{X}} - \mathbf{X})'\boldsymbol{\epsilon}_t = \begin{bmatrix} 0 \\ -\sigma^2\mathbf{P}'\boldsymbol{\imath}_{t-1,T-1} \end{bmatrix} + O_p\left(\sqrt{N}\right).$$

(ix) *When the identification assumption $\kappa_4 = 0$ holds,*

$$\hat{\sigma}_4 \ \to_p \ \sigma_4,$$

where $\hat{\sigma}_4$ is defined in Theorem OA.2.

Lemma 2. *Under Assumptions OA.9 and OA.3–OA.8, we have*

(i)

$$\frac{\bar{\mathbf{Z}}'\bar{\mathbf{Z}}}{N} = \mathbb{J}'\frac{\mathbf{Z}'\mathbf{Z}}{N}\mathbb{J} \to_p \mathbb{J}'\mathbf{\Sigma}_{\mathrm{Z}}\mathbb{J}.$$

(ii)

$$\frac{\bar{\mathbf{Z}}'\hat{\mathbf{X}}}{N} \to_p \left[\boldsymbol{\mu}_{\mathrm{z}}, \mathbb{J}'\mathbf{\Sigma}_{\mathrm{ZB}}\right] = \mathbf{\Sigma}_{\mathrm{ZX}}.$$

(iii)

$$\frac{\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}}{N} = \mathbb{J}_{t-1}'\frac{\mathbf{Z}'\mathbf{Z}}{N}\mathbb{J}_{t-1} \to_p \mathbb{J}_{t-1}'\mathbf{\Sigma}_{\mathrm{Z}}\mathbb{J}_{t-1}.$$

(iv)

$$\frac{\mathbf{Z}_{t-1}'\hat{\mathbf{X}}}{N} \to_p \left[\boldsymbol{\mu}_{\mathrm{z}}, \mathbb{J}_{t-1}'\mathbf{\Sigma}_{\mathrm{ZB}}\right] = \mathbf{\Sigma}_{\mathrm{ZX,t-1}}.$$

Lemma 3. *Under Assumptions OA.9 and OA.3–OA.8, we have*

(i)

$$\frac{1}{N}\hat{\mathbf{X}}'\left(\bar{\boldsymbol{\epsilon}} + (\mathbf{X} - \hat{\mathbf{X}})\bar{\mathbf{\Gamma}}_{\mathrm{f}}\right) \to_p -\mathbf{\Lambda}_1\bar{\mathbf{\Gamma}}_{\mathrm{f}}. \tag{OA.80}$$

(ii)

$$\frac{1}{N}\bar{\mathbf{Z}}'\left(\bar{\boldsymbol{\epsilon}} + (\mathbf{X} - \hat{\mathbf{X}})\bar{\mathbf{\Gamma}}_{\mathrm{f}}\right) \to_p \mathbf{0}_{K_{\mathrm{z}}}. \tag{OA.81}$$

(iii)

$$\frac{1}{N}\hat{\mathbf{X}}'\left(\boldsymbol{\epsilon}_t + (\mathbf{X} - \hat{\mathbf{X}})\mathbf{\Gamma}_{\mathrm{f},t-1}\right) \to_p -\mathbf{\Lambda}_1\mathbf{\Gamma}_{\mathrm{f},t-1} + \mathbf{\Lambda}_{2,t-1}. \tag{OA.82}$$

(iv)

$$\frac{1}{N}\mathbf{Z}_{t-1}'\left(\boldsymbol{\epsilon}_t + (\mathbf{X} - \hat{\mathbf{X}})\mathbf{\Gamma}_{\mathrm{f},t-1}\right) \to_p \mathbf{0}_{K_{\mathrm{z}}}. \tag{OA.83}$$

OA.4.1 Additional Lemmas required for WLS Estimation

This section establishes several preliminary results needed to derive the asymptotic properties of the WLS estimator described in Section OA.2.2. All results below hold as $N \to \infty$.

Lemma 4. *Under Assumptions OA.9, OA.3–OA.8, and OA.10–OA.13,*

(i)

$$\frac{\boldsymbol{\epsilon}\,\mathbf{W}_{t-1}\mathbf{1}_N}{N} \;\to_p\; \mathbf{0}_{T-1}.$$

(ii)

$$\frac{\boldsymbol{\epsilon}\,\mathbf{W}_{t-1}\mathbf{B}}{N} \;\to_p\; \mathbf{0}_{(T-1)\times K_{\mathrm{f}}}.$$

(iii)

$$\frac{\hat{\mathbf{B}}'\mathbf{W}_{t-1}\hat{\mathbf{B}}}{N} \;\to_p\; \boldsymbol{\Sigma}_\beta + \sigma^2\mathbf{P}'\mathbf{P}.$$

(iv)

$$\frac{\boldsymbol{\epsilon}\,\mathbf{W}_{t-1}\boldsymbol{\epsilon}'}{N} \;\to_p\; \sigma^2\mathbf{I}_{T-1}.$$

Lemma 5. *Under Assumptions OA.9, OA.3–OA.8, and OA.10–OA.13,*

(i)

$$\frac{\mathbf{Z}_{t-1}'\mathbf{W}_{t-1}\mathbf{B}}{N} \;\;\to_p\;\; \boldsymbol{\mu}_{\mathrm{z},t-1}\,\boldsymbol{\mu}_\beta'.$$

(ii)

$$\frac{\mathbf{Z}_{t-1}'\mathbf{W}_{t-1}\boldsymbol{\epsilon}_t}{N} \;\;\to_p\;\; \mathbf{0}_{K_{\mathrm{z}}}.$$

(iii)

$$\frac{\mathbf{Z}_{t-1}'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'}{N} \;\;\to_p\;\; \mathbf{0}_{K_{\mathrm{z}}\times(T-1)}.$$

Lemma 6. *Under Assumptions OA.9, OA.3–OA.8, and OA.10–OA.13,*

$$\frac{\hat{\mathbf{X}}'\mathbf{W}_{t-1}\boldsymbol{\epsilon}_t}{N} \;\;\to_p\;\; \sigma^2\begin{bmatrix} 0 \\ \mathbf{P}'\imath_{t,T-1}\end{bmatrix}.$$

Lemma 7. *Under Assumptions OA.9, OA.3–OA.8, and OA.10–OA.13, when $\kappa_4 = 0$, for every $2 \le t \le T$,*

$$\hat{\sigma}_{t-1}^{2(\mathrm{w})} \;\;\to_p\;\; \sigma^2, \qquad \hat{\sigma}_{4,t-1}^{(\mathrm{w})} \;\;\to_p\;\; \sigma_4,$$

with

$$\hat{\sigma}_{4,t-1}^{(\mathrm{w})} = \frac{1}{N}\sum_{s=2}^{T}\sum_{i=1}^{N}\frac{w_{i,t-1}\,\hat{\epsilon}_{i,s}^4}{3\,\mathrm{tr}\Big(\overline{\mathbb{M}}_{\check{\mathrm{D}}}^{(2)}\Big)}.$$

OA.4.2 Additional lemmas required for OLS estimation under model misspecification

Lemma 8. *Under Assumptions OA.9, OA.3–OA.8, and OA.2,*

$$\hat{\boldsymbol{\theta}}_{t-1,m} = \mathbf{S}_{t-1}\left(\frac{1}{N}\hat{\boldsymbol{\epsilon}}'\hat{\mathbf{m}}_{t-1} - \hat{\sigma}^2\mathbf{M}_{\tilde{D}}\boldsymbol{\imath}_{t-1,T-1}\right) \quad \to_p \quad \boldsymbol{\theta}_{t-1,m},$$

where

$$\hat{\mathbf{m}}_{t-1} = \mathbf{R}_t - (\hat{\mathbf{X}}, \mathbf{Z}_{t-1})\begin{bmatrix}\hat{\boldsymbol{\Gamma}}^{*(\mathrm{m})\prime}_{\mathrm{f},t-1} \\ \hat{\boldsymbol{\gamma}}^{*(\mathrm{m})\prime}_{\mathrm{z},t-1}\end{bmatrix},$$

and

$$\mathbf{M}^{(-1)}_{D,t-1} = \mathbf{M}_{11}^{-1}\,(\mathbf{I}_{t-2},\, \mathbf{0}_{t-2\times(T-t+1)}),$$

with $\mathbf{M}_{11}$ *denoting the top-left block of size* $(t-2)\times(t-2)$ *of the matrix* $\mathbf{M}_D$.

Lemma 9. *Under Assumptions OA.9, OA.3–OA.8, and OA.2,*

$$\frac{\mathbf{1}_N'\hat{\mathbf{m}}_{t-1}}{N} \quad \to_p \quad 0,$$

and

$$\hat{\sigma}_{t-1,mm} = \frac{\hat{\mathbf{m}}_{t-1}'\hat{\mathbf{m}}_{t-1}}{N} - \hat{\sigma}^2\hat{\mathbf{Q}}_{t-1}'\hat{\mathbf{Q}}_{t-1} + 2\,\hat{\boldsymbol{\delta}}^{*(\mathrm{m})\prime}_{\mathrm{f},t-1}\mathbf{P}'\begin{bmatrix}\hat{\boldsymbol{\theta}}_{t-1,m} \\ \mathbf{0}_{T-t+1}\end{bmatrix} \quad \to_p \quad \sigma_{t-1,mm}.$$

OA.5 Proofs of Theorems

Proof of Theorem OA.1. First, let us define the quantities that make the asymptotic covariance matrix of the estimator:

$$\mathbf{L}_{t-1} = \begin{bmatrix}\boldsymbol{\Sigma}_{\mathrm{X}} & \boldsymbol{\Sigma}_{\mathrm{ZX,t-1}}' \\ \boldsymbol{\Sigma}_{\mathrm{ZX,t-1}} & \mathbb{J}_{t-1}'\boldsymbol{\Sigma}_{\mathrm{Z}}\mathbb{J}_{t-1}\end{bmatrix}, \quad \text{and} \quad \mathbf{O}_{t-1} = \begin{bmatrix}\mathbf{U}_{t-1} & \sigma^2\mathbf{G}_{t-1}\mathbf{H}_{t-1}' \\ \sigma^2\mathbf{H}_{t-1}\mathbf{G}_{t-1}' & \mathbf{H}_{t-1}\boldsymbol{\Sigma}_{\mathrm{U}}\mathbf{H}_{t-1}'\end{bmatrix}, \tag{OA.84}$$

with $\mathbf{L}_{t-1}$ is assumed to be positive definite, $\mathbf{U}_{t-1} = \sigma^2\mathbf{Q}_{t-1}'\mathbf{Q}_{t-1}\boldsymbol{\Sigma}_{\mathrm{X}} + \begin{bmatrix}0 & \mathbf{0}_{K_{\mathrm{f}}}' \\ \mathbf{0}_{K_{\mathrm{f}}} & \mathbf{V}_{t-1}'\mathbf{U}_{\epsilon}\mathbf{V}_{t-1}\end{bmatrix}$, $N^{-1}\sum_{i=1}^{N}\mathbb{M}_{\tilde{D}_i} \to_p \mathbb{M}_{\tilde{D}}$, $\mathbf{Z}_{t-1}'\hat{\mathbf{X}}/N \to_p \boldsymbol{\Sigma}_{\mathrm{ZX,t-1}}$, $\mathbf{V}_{t-1} \equiv (\mathbf{Q}_{t-1}\otimes\mathbf{P}) - \big(\mathrm{vec}(\mathbb{M}_{\tilde{D}})/(T-2-K)\big)\,\mathbf{Q}_{t-1}'\mathbf{P}$, $\mathbf{G}_{t-1} = \big[\mathbf{Q}_{t-1}\otimes\boldsymbol{\mu}_{\mathrm{z},T-1}, \quad \mathbf{Q}_{t-1}\otimes\boldsymbol{\Sigma}_{\mathrm{ZB}}\big]'$, and $\mathbf{H}_{t-1} = \mathbf{Q}_{t-1}'\otimes\mathbb{J}_{t-1}'$, with $\mathbf{Q}_{t-1} = \boldsymbol{\imath}_{t-1,T-1} - \mathbf{P}\boldsymbol{\delta}_{\mathrm{f},t-1}$, and where $\boldsymbol{\Sigma}_{\mathrm{X}}$, $\mathbf{U}_{\epsilon}$, $\boldsymbol{\Sigma}_{\mathrm{ZB}}$, $\boldsymbol{\Sigma}_{\mathrm{Z}}$, $\boldsymbol{\Sigma}_{\mathrm{U}}$ and $\boldsymbol{\mu}_{\mathrm{z},T-1}$ are defined in Assumptions OA.4, OA.7, and OA.8.

Starting from (35), we rewrite

$$\begin{bmatrix}\hat{\boldsymbol{\Gamma}}^*_{\mathrm{f},t-1}\\ \hat{\gamma}^*_{\mathrm{z},t-1}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Gamma}_{\mathrm{f},t-1}\\ \gamma_{\mathrm{z,t-1}}\end{bmatrix} + \begin{bmatrix}\hat{\mathbf{X}}'\hat{\mathbf{X}} - N\hat{\boldsymbol{\Lambda}}_1 & \hat{\mathbf{X}}'\mathbf{Z}_{t-1}\\ \mathbf{Z}'_{t-1}\hat{\mathbf{X}} & \mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}\end{bmatrix}^{-1} \times$$
$$\times \left(\begin{bmatrix}N\hat{\boldsymbol{\Lambda}}_1\\ \mathbf{0}_{K_{\mathrm{z}}\times(K_{\mathrm{f}}+1)}\end{bmatrix}\boldsymbol{\Gamma}_{\mathrm{f},t-1} + \begin{bmatrix}\hat{\mathbf{X}}'\\ \mathbf{Z}'_{t-1}\end{bmatrix}\left(\boldsymbol{\epsilon}_t + (\mathbf{X}-\hat{\mathbf{X}})\boldsymbol{\Gamma}_{\mathrm{f},t-1}\right) - \begin{bmatrix}N\hat{\boldsymbol{\Lambda}}_{2,t-1}\\ \mathbf{0}_{K_{\mathrm{z}}}\end{bmatrix}\right) \tag{OA.85}$$

By Lemmas 1-3

$$\hat{\boldsymbol{\Lambda}}_1\boldsymbol{\Gamma}_{\mathrm{f},t-1} + \frac{1}{N}\hat{\mathbf{X}}'\left(\boldsymbol{\epsilon}_t + (\mathbf{X}-\hat{\mathbf{X}})\boldsymbol{\Gamma}_{t-1}\right) - \hat{\boldsymbol{\Lambda}}_{2,t-1} = O_p\left(\frac{1}{\sqrt{N}}\right), \quad \text{and that}$$

$$\frac{1}{N}\mathbf{Z}'_{t-1}\left(\boldsymbol{\epsilon}_t + (\mathbf{X}-\hat{\mathbf{X}})\boldsymbol{\Gamma}_{\mathrm{f},t-1}\right) \to_p \mathbf{0}_{K_{\mathrm{z}}}.$$

Moreover, by Lemmas 1 and 2, $N^{-1}\begin{bmatrix}\hat{\mathbf{X}}'\hat{\mathbf{X}} - N\hat{\boldsymbol{\Lambda}}_1 & \hat{\mathbf{X}}'\mathbf{Z}_{t-1}\\ \mathbf{Z}'_{t-1}\hat{\mathbf{X}} & \mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}\end{bmatrix}^{-1} = O_p(1).$

To prove part (ii), noticing that $\boldsymbol{\epsilon}_t = \boldsymbol{\epsilon}'\boldsymbol{\imath}_{t-1,T-1}$, we get that

$$\frac{1}{\sqrt{N}}\left(N\hat{\boldsymbol{\Lambda}}_1\boldsymbol{\Gamma}_{\mathrm{f},t-1} + \hat{\mathbf{X}}'\left(\boldsymbol{\epsilon}_t + (\mathbf{X}-\hat{\mathbf{X}})\boldsymbol{\Gamma}_{\mathrm{f},t-1}\right) - N\hat{\boldsymbol{\Lambda}}_{2,t-1}\right) =$$
$$= \begin{bmatrix}\frac{\mathbf{1}'_N\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}\\ \frac{\mathbf{B}'\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}\end{bmatrix} + \begin{bmatrix}0\\ \mathcal{P}'\frac{\boldsymbol{\epsilon}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1} - \sqrt{N}\hat{\sigma}^2\mathbf{P}'\mathbf{Q}_{t-1}\end{bmatrix}.$$

Moreover, we have that

$$\frac{1}{\sqrt{N}}\mathbf{Z}'_{t-1}\left(\boldsymbol{\epsilon}_t + (\mathbf{X}-\hat{\mathbf{X}})\boldsymbol{\Gamma}_{\mathrm{f},t-1}\right) = \frac{\mathbf{Z}'_{t-1}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}. \tag{OA.86}$$

Therefore, using (OA.86) and (OA.86) in (OA.85), it follows that

$$\sqrt{N}\begin{bmatrix}\hat{\boldsymbol{\Gamma}}^{*}_{\mathrm{f},t-1}-\boldsymbol{\Gamma}_{\mathrm{f},t-1}\\ \hat{\boldsymbol{\gamma}}^{*}_{\mathrm{z},t-1}-\boldsymbol{\gamma}_{\mathrm{z,t-1}}\end{bmatrix} = \begin{bmatrix}\frac{\hat{\mathbf{X}}'\hat{\mathbf{X}}}{N}-\hat{\boldsymbol{\Lambda}}_1 & \frac{\hat{\mathbf{X}}'\mathbf{Z}_{t-1}}{N}\\ \frac{\mathbf{Z}'_{t-1}\hat{\mathbf{X}}}{N} & \frac{\mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}}{N}\end{bmatrix}^{-1}\times$$

$$\left(\begin{bmatrix}\frac{\mathbf{1}'_N\epsilon'}{\sqrt{N}}\mathbf{Q}_{t-1}\\ \frac{\mathbf{B}'\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}\\ \mathbf{0}_{K_{\mathrm{z}}}\end{bmatrix}+\begin{bmatrix}0\\ \mathbf{P}'\frac{\boldsymbol{\epsilon}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}-\sqrt{N}\hat{\sigma}^2\mathbf{P}'\mathbf{Q}_{t-1}\\ \mathbf{0}_{K_{\mathrm{z}}}\end{bmatrix}+\begin{bmatrix}0\\ \mathbf{0}_{K_{\mathrm{z}}}\\ \frac{\mathbf{Z}'_{t-1}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}\end{bmatrix}\right)$$

$$= \begin{bmatrix}\frac{\hat{\mathbf{X}}'\hat{\mathbf{X}}}{N}-\hat{\boldsymbol{\Lambda}}_1 & \frac{\hat{\mathbf{X}}'\mathbf{Z}_{t-1}}{N}\\ \frac{\mathbf{Z}'_{t-1}\hat{\mathbf{X}}}{N} & \frac{\mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}}{N}\end{bmatrix}^{-1}(I_1+I_2+I_3)\,. \tag{OA.87}$$

Now, by Lemmas 1(iv) and 2, we have that

$$\frac{1}{N}\begin{bmatrix}\hat{\mathbf{X}}'\hat{\mathbf{X}}-N\hat{\boldsymbol{\Lambda}}_1 & \hat{\mathbf{X}}'\mathbf{Z}_{t-1}\\ \mathbf{Z}'_{t-1}\hat{\mathbf{X}} & \mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}\end{bmatrix}\to_p \mathbf{L}_{t-1}. \tag{OA.88}$$

Regarding the variances of the terms I_1 and I_2, under Assumption OA.7, we get

$$\mathrm{Var}\left[\begin{bmatrix}\frac{\mathbf{1}'_N\epsilon'}{\sqrt{N}}\mathbf{Q}_{t-1}\\ \frac{\mathbf{B}'\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}\end{bmatrix}\right]\to_p \sigma^2\mathbf{Q}'_{t-1}\mathbf{Q}_{t-1}\boldsymbol{\Sigma}_{\mathrm{X}},\quad\text{and} \tag{OA.89}$$

$$\mathrm{Var}\left[\mathbf{P}'\frac{\boldsymbol{\epsilon}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}-\sqrt{N}\hat{\sigma}^2\mathbf{P}'\mathbf{Q}_{t-1}\right]\to_p \mathbf{V}'_{t-1}\mathbf{U}_{\epsilon}\mathbf{V}_{t-1}. \tag{OA.90}$$

Notice also that, under Assumption OA.6, the two terms I_1 and I_2 are uncorrelated. Consider now the term I_3, and notice that $\mathbf{z}_{i,t}\equiv\mathbb{J}'_t\mathbf{Z}'\boldsymbol{\imath}_{i,N}$. Then, using the properties of the vec$(\cdot)$ operator, recalling that $\mathbf{u}_i\equiv\boldsymbol{\epsilon}_i\otimes\mathbf{z}_i$ and $\mathbf{H}_{t-1}=\mathbf{Q}'_{t-1}\otimes\mathbb{J}'_{t-1}$, it follows that

$$\frac{\mathbf{Z}'_{t-1}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{H}_{t-1}\mathbf{u}_i. \tag{OA.91}$$

Therefore, using (OA.91), and under Assumption OA.8(v), one obtains

$$\mathrm{Var}\left[\frac{\mathbf{Z}'_{t-1}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}\right]=\mathbf{H}_{t-1}\frac{1}{N}\sum_{i,j=1}^{N}\boldsymbol{\Sigma}_{\mathrm{u,ij}}\mathbf{H}'_{t-1}\to\mathbf{H}_{t-1}\boldsymbol{\Sigma}_{\mathrm{U}}\mathbf{H}'_{t-1} \tag{OA.92}$$

Finally, let us consider the covariance terms of I_3 with both I_1 and I_2. To derive the covariance between I_3 and I_2, we establish

$$
\begin{aligned}
&\operatorname{Cov}\left[\mathbf{P}'\frac{\boldsymbol{\epsilon}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1} - \sqrt{N}\hat{\sigma}^2\mathbf{P}'\mathbf{Q}_{t-1}, \mathbf{Q}'_{t-1}\frac{\boldsymbol{\epsilon}\mathbf{Z}_{t-1}}{\sqrt{N}}\right] = \\
&= \mathbf{V}'_{t-1}\frac{1}{N}\sum_{i,j=1}^{N}\boldsymbol{\Sigma}_{u\epsilon,ij}\mathbf{H}'_{t-1} \to \mathbf{0}_{K_{\mathrm{f}}\times K_{\mathrm{z}}}
\end{aligned}
$$

comma

by Assumption OA.8(viii). Finally, to derive the $\operatorname{Cov}(I_1, I'_3)$, we need to calculate

$$
\begin{aligned}
\operatorname{Cov}\left[\begin{bmatrix}\frac{\mathbf{1}'_N\epsilon'}{\sqrt{N}}\mathbf{Q}_{t-1}\\ \frac{\mathbf{B}'\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}\end{bmatrix}, \mathbf{Q}'_{t-1}\frac{\boldsymbol{\epsilon}\mathbf{Z}_{t-1}}{\sqrt{N}}\right] &= \frac{1}{N}\sum_{i,j=1}^{N}\sigma_{ij}\left(\mathbf{Q}'_{t-1}\otimes \mathrm{E}\left[\begin{bmatrix}1\\ \boldsymbol{\beta}_i\end{bmatrix}\mathbf{z}'_j\right]\right)\mathbf{H}'_{t-1}\\
&\to \sigma^2\mathbf{G}_{t-1}\mathbf{H}'_{t-1}.
\end{aligned}
$$

by Assumption OA.8, and where we set

$$
\mathbf{G}_{t-1} = \Big(\left[\mathbf{Q}_{t-1}\otimes\boldsymbol{\mu}_{\mathrm{z},T-1}\,,\ \mathbf{Q}_{t-1}\otimes\boldsymbol{\Sigma}_{\mathrm{ZB}}\right]'\Big)\Big(\left[\mathbf{Q}_{t-1}\otimes\boldsymbol{\mu}_{\mathrm{z},T-1}\,,\ \mathbf{Q}_{t-1}\otimes\boldsymbol{\Sigma}_{\mathrm{ZB}}\right]'\Big)'.
$$

Therefore, putting all the above results together, and recalling that

$$
\mathbf{U}_{t-1} = \sigma^2\mathbf{Q}'_{t-1}\mathbf{Q}_{t-1}\,\boldsymbol{\Sigma}_{\mathrm{X}} + \begin{bmatrix}0 & \mathbf{0}'_{K_{\mathrm{f}}}\\ \mathbf{0}_{K_{\mathrm{f}}} & \mathbf{V}'_{t-1}\mathbf{U}_{\epsilon}\mathbf{V}_{t-1}\end{bmatrix},
$$

we obtain $\mathbf{O}_{t-1}$, concluding the proof.

■

Remark OA.20. *Theorem 1 in the main paper immediately follows as a special case of Theorem OA.1 when all the corresponding quantities that apply to $\gamma_{0,t-1}$ are replaced with the corresponding quantities obtained by setting $\gamma_{0,t-1}$ equal to the risk-free rate.*

Proof of Theorem OA.2. By Lemma 1 and Lemma 2(iii)–(iv), it follows that $\hat{\mathbf{L}}_{t-1} \to_p \mathbf{L}_{t-1}$. By part (i) of Theorem OA.1, we have $\hat{\boldsymbol{\delta}}^*_{\mathrm{f},t-1} \to_p \boldsymbol{\delta}_{\mathrm{f},t-1}$, implying that $\hat{\mathbf{Q}}_{t-1}$ is a consistent estimator of $\mathbf{Q}_{t-1}$. Moreover, as $N\to\infty$, we have $\overline{\mathbb{M}}_{\tilde{\mathrm{D}}} \to_p \mathbb{M}_{\tilde{D}}$, $\hat{\boldsymbol{\mu}}_{T-1,z} \to_p \boldsymbol{\mu}_{T-1,z}$, $\hat{\boldsymbol{\Sigma}}_{\mathrm{ZB}} \to_p \boldsymbol{\Sigma}_{\mathrm{ZB}}$, and $\mathbf{Z}'\mathbf{Z}/N \to_p \boldsymbol{\Sigma}_{\mathrm{Z}}$. It follows that $\hat{\mathbf{V}}_{t-1} \to_p \mathbf{V}_{t-1}$, $\hat{\mathbf{G}}_{t-1} \to_p \mathbf{G}_{t-1}$, and $\hat{\mathbf{H}}_{t-1} \to_p \mathbf{H}_{t-1}$. Finally, a consistent estimator of $\hat{\mathbf{U}}_{t-1}$ requires a consistent estimate of $\mathbf{U}_{\epsilon}$, which can be obtained using Lemma 1(ix). This concludes the proof of Theorem OA.2. ■

Remark OA.21. *Theorem 2 in the main paper immediately follows as a special case of Theorem OA.2 when all the corresponding quantities that apply to $\gamma_{0,t-1}$ are replaced with the corresponding quantities obtained by setting $\gamma_{0,t-1}$ equal to the risk-free rate.*

Proof of Theorem OA.3. First, let us define the quantities that make up the asymptotic covariance matrix of the estimator. Let $\boldsymbol{\mu}_{\mathrm{z}} = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}[\mathbf{z}_{i,t}]$, and $\boldsymbol{\mu}_{\mathrm{x}} = (1, \boldsymbol{\mu}_{\beta}')'$. Set $\lambda_{t-1} = \mathbf{Y}_{t-1}'\boldsymbol{\Sigma}_{\mathrm{V}}\mathbf{Y}_{t-1}$, with $\mathbf{Y}_{t-1} = \mathbf{Q}_{t-1} \otimes \boldsymbol{\imath}_{t-1,T-1}$ and define $\mathbf{S}_{t-1} \equiv \boldsymbol{\mu}_{\mathrm{z}}\mathbf{Y}_{t-1}'(\sigma^2 \boldsymbol{I}_{T-1} \otimes \boldsymbol{\Sigma}_{\mathrm{ZW}}')\mathbf{H}_{t-1}'$.Then, set

$$\mathbf{O}_{t-1}^{(\mathrm{w})} = \lambda_{t-1}\begin{bmatrix} \boldsymbol{\mu}_{\mathrm{x}}\boldsymbol{\mu}_{\mathrm{x}}' & \boldsymbol{\mu}_{\mathrm{x}}\boldsymbol{\mu}_{\mathrm{z}}' \\ \boldsymbol{\mu}_{\mathrm{z}}\boldsymbol{\mu}_{\mathrm{x}}' & \boldsymbol{\mu}_{\mathrm{z}}\boldsymbol{\mu}_{\mathrm{z}}' \end{bmatrix} + \mathbf{M}_{t-1}^{(\mathrm{w})} \qquad \text{comma} \tag{OA.93}$$

with

$$\mathbf{M}_{t-1}^{(\mathrm{w})} \equiv \begin{bmatrix} 0 & \mathbf{0}_{K_{\mathrm{f}}}' & \mathbf{0}_{K_{\mathrm{z}}}' \\ \mathbf{0}_{K_{\mathrm{f}}} & \mathbf{V}_{t-1}'\mathbf{U}_{\epsilon}\mathbf{V}_{t-1} & \mathbf{0}_{K_{\mathrm{f}}\times K_{\mathrm{z}}} \\ \mathbf{0}_{K_{\mathrm{z}}} & \mathbf{0}_{K_{\mathrm{z}}\times K_{\mathrm{f}}} & \mathbf{H}_{t-1}\boldsymbol{\Sigma}_{\mathrm{U}}\mathbf{H}_{t-1}' + \mathbf{S}_{t-1} + \mathbf{S}_{t-1}' \end{bmatrix}$$

where $\mathbf{Q}_{t-1}$, $\mathbf{V}_{t-1}$, and $\mathbf{H}_{t-1}$ are defined in the proof to Theorem 1, $\boldsymbol{\Sigma}_{\mathrm{ZW}}$, $\boldsymbol{\Sigma}_{\mathrm{Z},t-1}$, $\boldsymbol{\Sigma}_{\mathrm{V}}$, and $\boldsymbol{\mu}_{\mathrm{z},t-1}$ are defined in Assumption OA.13, and $\boldsymbol{\Sigma}_{\mathrm{X}}$, $\mathbf{U}_{\epsilon}$, and $\boldsymbol{\Sigma}_{\mathrm{U}}$ are defined in Assumptions OA.4, OA.7, and OA.8.

Next, let us start the proof from the definition in (OA.25) and rewrite

$$\begin{aligned}\begin{bmatrix} \hat{\boldsymbol{\Gamma}}_{\mathrm{f},t-1}^{*(\mathrm{w})} \\ \hat{\gamma}_{\mathrm{z},t-1}^{*(\mathrm{w})} \end{bmatrix} - \begin{bmatrix} \boldsymbol{\Gamma}_{\mathrm{f},t-1} \\ \gamma_{\mathrm{z},t-1} \end{bmatrix} &= \begin{bmatrix} \hat{\mathbf{X}}'\mathbf{W}_{t-1}\hat{\mathbf{X}} - N\hat{\boldsymbol{\Lambda}}_{1,t-1}^{(\mathrm{w})} & \hat{\mathbf{X}}'\mathbf{W}_{t-1}\mathbf{Z}_{t-1} \\ \mathbf{Z}_{t-1}'\mathbf{W}_{t-1}\hat{\mathbf{X}} & \mathbf{Z}_{t-1}'\mathbf{W}_{t-1}\mathbf{Z}_{t-1} \end{bmatrix}^{-1} \times \\ &\times \left(\begin{bmatrix} N\hat{\boldsymbol{\Lambda}}_{1,t-1}^{(\mathrm{w})} \\ \mathbf{0}_{\mathrm{K}+1} \end{bmatrix} \boldsymbol{\Gamma}_{\mathrm{f},t-1} + \begin{bmatrix} \hat{\mathbf{X}}'\mathbf{W}_{t-1} \\ \mathbf{Z}_{t-1}'\mathbf{W}_{t-1} \end{bmatrix} (\boldsymbol{\epsilon}_t + (\mathbf{X} - \hat{\mathbf{X}})\boldsymbol{\Gamma}_{\mathrm{f},t-1}) - \begin{bmatrix} N\hat{\boldsymbol{\Lambda}}_{2,t-1}^{(\mathrm{w})} \\ \mathbf{0}_{K_{\mathrm{z}}} \end{bmatrix} \right).\end{aligned} \tag{OA.94}$$

To streamline the proof, we begin with part (ii) of the theorem and derive the asymptotic distribution of the above expression, since its $\sqrt{N}$-rate of convergence then follows immediately.

First notice that, by Assumptions OA.10, OA.11, and OA.13,

$$\frac{1}{N}\begin{bmatrix} \hat{\mathbf{X}}'\mathbf{W}_{t-1}\hat{\mathbf{X}} - N\hat{\boldsymbol{\Lambda}}_{1,t-1}^{(\mathrm{w})} & \hat{\mathbf{X}}'\mathbf{W}_{t-1}\mathbf{Z}_{t-1} \\ \mathbf{Z}_{t-1}'\mathbf{W}_{t-1}\hat{\mathbf{X}} & \mathbf{Z}_{t-1}'\mathbf{W}_{t-1}\mathbf{Z}_{t-1} \end{bmatrix} \to_p \mathbf{L}_{t-1},$$

where we set $\boldsymbol{\mu}_{\mathrm{X}} = \begin{bmatrix} 1 \\ \boldsymbol{\mu}_\beta \end{bmatrix}$.

Now consider the next term in (OA.94) and observe that

$$\begin{aligned}
&\hat{\mathbf{X}}'\mathbf{W}_{t-1}\boldsymbol{\epsilon}_t + \hat{\mathbf{X}}'\mathbf{W}_{t-1}(\mathbf{X}-\hat{\mathbf{X}})\boldsymbol{\Gamma}_{\mathrm{f},t-1} - N\hat{\boldsymbol{\Lambda}}^{(\mathrm{w})}_{2,t-1} + N\hat{\boldsymbol{\Lambda}}^{(\mathrm{w})}_{1,t-1}\boldsymbol{\Gamma}_{\mathrm{f},t-1} \\
= \;& \mathbf{a}_{11} + \mathbf{a}_{12},
\end{aligned}$$

with

$$\mathbf{a}_{11} = \mathbf{X}'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'\mathbf{Q}_{t-1}, \qquad \mathbf{a}_{12} = \begin{bmatrix} 0 \\ \mathbf{P}'\Big(\boldsymbol{\epsilon}\mathbf{W}_{t-1}\boldsymbol{\epsilon}' - N\hat{\sigma}^{2(\mathrm{w})}_{t-1}\mathbf{I}_{T-1}\Big)\mathbf{Q}_{t-1} \end{bmatrix}.$$

Moreover,

$$\mathbf{Z}'_{t-1}\mathbf{W}_{t-1}\boldsymbol{\epsilon}_t + \mathbf{Z}'_{t-1}\mathbf{W}_{t-1}(\mathbf{X}-\hat{\mathbf{X}})\boldsymbol{\Gamma}_{\mathrm{f},t-1} \;=\; \mathbf{a}_{21} + \mathbf{a}_{22},$$

where

$$\mathbf{a}_{21} = \mathrm{E}[\mathbf{Z}_{t-1}]'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'\mathbf{Q}_{t-1}, \qquad \mathbf{a}_{22} = \big(\mathbf{Z}_{t-1} - \mathrm{E}[\mathbf{Z}_{t-1}]\big)'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'\mathbf{Q}_{t-1}.$$

Let us start with the term $\mathbf{a}_{11}$. Using Assumption OA.13, and noting that $\mathbf{W}_{t-1}\mathbf{1}_N = \mathbf{W}\boldsymbol{\imath}_{t-1,T-1}$, we can write

$$\begin{aligned}
\frac{1}{\sqrt{N}}\mathbf{X}'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'\mathbf{Q}_{t-1} \;&=\; \frac{\mathbf{X}'\mathbf{1}_N}{N}\left(\mathbf{Q}'_{t-1}\otimes\boldsymbol{\imath}'_{t-1,T-1}\right)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\epsilon}_i\otimes\mathbf{w}_i + o_p(1) \\
&\to_d\; \mathcal{N}\big(\mathbf{0}_{K_\mathrm{f}+1},\, \lambda_{t-1}\boldsymbol{\mu}_{\mathrm{X}}\boldsymbol{\mu}'_{\mathrm{X}}\big),
\end{aligned}$$

where $\boldsymbol{\mu}_{\mathrm{X}} = (1, \boldsymbol{\mu}'_\beta)'$, $\lambda_{t-1} = \mathbf{Y}'_{t-1}\boldsymbol{\Sigma}_{\mathrm{V}}\mathbf{Y}_{t-1}$, with $\mathbf{Y}_{t-1} = \big(\mathbf{Q}_{t-1}\otimes\boldsymbol{\imath}_{t-1,T-1}\big)$ and $\boldsymbol{\Sigma}_{\mathrm{V}}$ defined in Assumption OA.13.

Consider now the term $\mathbf{a}_{12}$. By Assumptions OA.12 and OA.13, we obtain

$$\begin{aligned}
\frac{1}{\sqrt{N}}\mathbf{P}'\Big(\boldsymbol{\epsilon}\mathbf{W}_{t-1}\boldsymbol{\epsilon}' - N\hat{\sigma}^{2(\mathrm{w})}_{t-1}\mathbf{I}_{T-1}\Big)\mathbf{Q}_{t-1} \;&=\; \mathbf{V}'_{t-1}\,\mu_{\mathrm{w},t-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathrm{vec}\big(\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}'_i - \sigma_i^2\mathbf{I}_{T-1}\big) + o_p(1) \\
&\to_d\; \mathcal{N}\big(\mathbf{0}_{K_\mathrm{f}},\, \mathbf{V}'_{t-1}\mathbf{U}_\epsilon\mathbf{V}_{t-1}\big),
\end{aligned}$$

where $\mu_{\mathrm{w},t-1} \to_p 1$ by Assumption OA.10, and $\mathbf{U}_\epsilon$ is defined in Assumption OA.7.

For the term $\mathbf{a}_{21}$, using Assumption OA.7 and following the same steps as for $\mathbf{a}_{11}$, we get

$$\begin{aligned}\frac{1}{\sqrt{N}}\mathrm{E}[\mathbf{Z}_{t-1}]'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'\mathbf{Q}_{t-1} \quad &= \quad \frac{1}{\sqrt{N}}\left(\mathrm{E}[\mathbf{Z}_{t-1}]'\left(\mathbf{I}_N - \frac{\mathbf{1}_N\mathbf{1}_N'}{N}\right)\mathbf{W}_{t-1}\boldsymbol{\epsilon}'\mathbf{Q}_{t-1}\right. \\ &\qquad \left. + \mathrm{E}[\mathbf{Z}_{t-1}]'\frac{\mathbf{1}_N\mathbf{1}_N'}{N}\mathbf{W}_{t-1}\boldsymbol{\epsilon}'\mathbf{Q}_{t-1}\right) \\ &\to_d \quad \mathcal{N}\left(\mathbf{0}_{K_\mathrm{z}},\, \lambda_{t-1}\boldsymbol{\mu}_\mathrm{z}\boldsymbol{\mu}_\mathrm{z}'\right).\end{aligned}$$

Finally, for the term $\mathbf{a}_{22}$ we have

$$\begin{aligned}\sqrt{N}\left(\frac{(\mathbf{Z}_{t-1} - \mathrm{E}[\mathbf{Z}_{t-1}])'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'}{N}\right)\mathbf{Q}_{t-1} \quad &= \quad \sqrt{N}\left(\frac{(\mathbf{Z}_{t-1} - \mathrm{E}[\mathbf{Z}_{t-1}])'\boldsymbol{\epsilon}'}{N}\right)\mathbf{Q}_{t-1} + o_p(1) \\ &\to_d \quad \mathcal{N}\left(\mathbf{0}_{K_\mathrm{z}},\, \mathbf{H}_{t-1}\boldsymbol{\Sigma}_\mathrm{U}\mathbf{H}_{t-1}'\right).\end{aligned}$$

We can now derive the asymptotic distribution of the estimator. Using all the results derived above, we obtain

$$\begin{aligned}&\frac{1}{\sqrt{N}}\left(\begin{bmatrix} N\hat{\boldsymbol{\Lambda}}_{1,t-1}^{(\mathrm{w})} \\ \mathbf{0}_{\mathrm{K}+1}\end{bmatrix}\boldsymbol{\Gamma}_{\mathrm{f},t-1} + \begin{bmatrix}\hat{\mathbf{X}}'\mathbf{W}_{t-1} \\ \mathbf{Z}_{t-1}'\mathbf{W}_{t-1}\end{bmatrix}(\boldsymbol{\epsilon}_t + (\mathbf{X} - \hat{\mathbf{X}})\boldsymbol{\Gamma}_{\mathrm{f},t-1}) - \begin{bmatrix} N\hat{\boldsymbol{\Lambda}}_{2,t-1}^{(\mathrm{w})} \\ \mathbf{0}_{K_\mathrm{z}}\end{bmatrix}\right) \\ &= \begin{bmatrix}\frac{\mathbf{1}_N'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1} \\ \left(\frac{\mathbf{B}'\mathbf{1}_N}{N}\right)\frac{\mathbf{1}_N'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1} \\ \mathbf{0}_{K_\mathrm{z}}\end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{P}'\frac{1}{\sqrt{N}}\left(\boldsymbol{\epsilon}\mathbf{W}_{t-1}\boldsymbol{\epsilon}' - N\hat{\sigma}_{t-1}^{2(\mathrm{w})}\mathbf{I}_{T-1}\right)\mathbf{Q}_{t-1} \\ \mathbf{0}_{K_\mathrm{z}}\end{bmatrix} \\ &+ \begin{bmatrix} 0 \\ \mathbf{0}_{K_\mathrm{f}} \\ \left(\frac{\mathrm{E}[\mathbf{Z}_{t-1}]'\mathbf{1}_N}{N}\right)\left(\frac{\mathbf{1}_N'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'}{\sqrt{N}}\right)\mathbf{Q}_{t-1}\end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{0}_{K_\mathrm{f}} \\ \frac{(\mathbf{Z}_{t-1} - \mathrm{E}[\mathbf{Z}_{t-1}])'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}\end{bmatrix} \\ &= \quad \mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3 + \mathbf{a}_4, \quad \text{setting}\end{aligned}$$

$$\mathbf{a}_1 = \begin{bmatrix}\frac{\mathbf{1}_N'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1} \\ \left(\frac{\mathbf{B}'\mathbf{1}_N}{N}\right)\frac{\mathbf{1}_N'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1} \\ \mathbf{0}_{K_\mathrm{z}}\end{bmatrix},\ \mathbf{a}_2 = \begin{bmatrix} 0 \\ \mathbf{P}'\frac{1}{\sqrt{N}}\left(\boldsymbol{\epsilon}\mathbf{W}_{t-1}\boldsymbol{\epsilon}' - N\hat{\sigma}_{t-1}^{2(\mathrm{w})}\mathbf{I}_{T-1}\right)\mathbf{Q}_{t-1} \\ \mathbf{0}_{K_\mathrm{z}}\end{bmatrix},$$

$$\mathbf{a}_3 = \begin{bmatrix} 0 \\ \mathbf{0}_{K_\mathrm{f}} \\ \left(\frac{\mathrm{E}[\mathbf{Z}_{t-1}]'\mathbf{1}_N}{N}\right)\left(\frac{\mathbf{1}_N'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'}{\sqrt{N}}\right)\mathbf{Q}_{t-1}\end{bmatrix},\ \text{and}\ \mathbf{a}_4 = \begin{bmatrix} 0 \\ \mathbf{0}_{K_\mathrm{f}} \\ \frac{(\mathbf{Z}_{t-1} - \mathrm{E}[\mathbf{Z}_{t-1}])'\mathbf{W}_{t-1}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}\end{bmatrix}.$$

Notice that all the terms have zero mean. Therefore, using the results above:

$$\text{Var}[\mathbf{a}_1] \to \lambda_{t-1} \begin{bmatrix} \boldsymbol{\mu}_{\text{X}}\boldsymbol{\mu}'_{\text{X}} & \mathbf{0}_{(K_{\text{f}}+1)\times K_{\text{z}}} \\ \mathbf{0}_{K_{\text{z}}\times(K_{\text{f}}+1)} & \mathbf{0}_{K_{\text{z}}\times K_{\text{z}}} \end{bmatrix}, \quad \text{Var}[\mathbf{a}_2] \to \begin{bmatrix} 0 & \mathbf{0}'_{K_{\text{f}}} & \mathbf{0}'_{K_{\text{z}}} \\ \mathbf{0}_{K_{\text{f}}} & \mathbf{V}'_{t-1}\mathbf{U}_{\epsilon}\mathbf{V}_{t-1} & \mathbf{0}_{K_{\text{f}}\times K_{\text{z}}} \\ \mathbf{0}_{K_{\text{z}}} & \mathbf{0}_{K_{\text{z}}\times K_{\text{f}}} & \mathbf{0}_{K_{\text{z}}\times K_{\text{z}}} \end{bmatrix},$$

$$\text{Var}[\mathbf{a}_3] \to \lambda_{t-1} \begin{bmatrix} 0 & \mathbf{0}'_{K_{\text{f}}} & \mathbf{0}'_{K_{\text{z}}} \\ \mathbf{0}_{K_{\text{f}}} & \mathbf{0}_{K_{\text{f}}\times K_{\text{f}}} & \mathbf{0}_{K_{\text{f}}\times K_{\text{z}}} \\ \mathbf{0}_{K_z} & \mathbf{0}_{K_z\times K_f} & \boldsymbol{\mu}_{\text{z}}\boldsymbol{\mu}'_{\text{z}} \end{bmatrix}, \quad \text{and} \quad \text{Var}[\mathbf{a}_4] \to \begin{bmatrix} 0 & \mathbf{0}'_{K_{\text{f}}} & \mathbf{0}'_{K_{\text{z}}} \\ \mathbf{0}_{K_{\text{f}}} & \mathbf{0}_{K_{\text{f}}\times K_{\text{f}}} & \mathbf{0}_{K_{\text{f}}\times K_{\text{z}}} \\ \mathbf{0}_{K_{\text{z}}} & \mathbf{0}_{K_{\text{z}}\times K_{\text{f}}} & \mathbf{H}_{t-1}\boldsymbol{\Sigma}_{\text{U}}\mathbf{H}'_{t-1} \end{bmatrix}.$$

It remains to evaluate the covariance terms. Under Assumption OA.6, $\text{Cov}[\mathbf{a}_1, \mathbf{a}'_2] \to \mathbf{0}_{(K+1)\times(K+1)}$ and $\text{Cov}[\mathbf{a}_1, \mathbf{a}'_4] \to \mathbf{0}_{(K+1)\times(K+1)}$, while

$$\text{Cov}[\mathbf{a}_1, \mathbf{a}'_3] \quad \to \quad \lambda_{t-1} \begin{bmatrix} \mathbf{0}_{(K_{\text{f}}+1)\times(K_{\text{f}}+1)} & \boldsymbol{\mu}_{\text{X}}\boldsymbol{\mu}'_{z} \\ \mathbf{0}_{K_{\text{z}}\times(K_{\text{f}}+1)} & \mathbf{0}_{K_{\text{z}}\times K_{\text{z}}} \end{bmatrix}.$$

Regarding $\mathbf{a}_2$, notice that under Assumption OA.6, $\text{Cov}[\mathbf{a}_2, \mathbf{a}'_3] \to \mathbf{0}_{(K+1)\times(K+1)}$ and $\text{Cov}[\mathbf{a}_2, \mathbf{a}'_4] \to \mathbf{0}_{(K+1)\times(K+1)}$. Finally,

$$\text{Cov}[\mathbf{a}_3, \mathbf{a}'_4] \quad \to \quad \begin{bmatrix} 0 & \mathbf{0}'_{K_{\text{f}}} & \mathbf{0}'_{K_{\text{z}}} \\ \mathbf{0}_{K_{\text{f}}} & \mathbf{0}_{K_{\text{f}}\times K_{\text{f}}} & \mathbf{0}_{K_{\text{f}}\times K_{\text{z}}} \\ \mathbf{0}_{K_{\text{z}}} & \mathbf{0}_{K_{\text{z}}\times K_{\text{f}}} & \mathbf{S}_{t-1} \end{bmatrix},$$

where

$$\mathbf{S}_{t-1} = \boldsymbol{\mu}_{\text{z}}\mathbf{Y}'_{t-1}\big(\sigma^2\mathbf{I}_{T-1} \otimes \boldsymbol{\Sigma}'_{\text{ZW}}\big)\mathbf{H}'_{t-1}.$$

In fact,

$$\text{Cov}[\mathbf{a}_3, \mathbf{a}'_4] = \boldsymbol{\mu}_{\text{z}}(\mathbf{Q}'_{t-1} \otimes \boldsymbol{\imath}'_{t-1,T-1})\text{E}\left[\frac{1}{N}\sum_{i=1}^{N} w_{i,t-1}(\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}'_i \otimes \mathbf{w}_{i\cdot}\mathbf{z}'_j)\right]$$

Collecting terms concludes the proof. ■

Proof of Theorem OA.5.

Rewrite the premia estimator as:

$$
\begin{aligned}
\begin{bmatrix}\hat{\boldsymbol{\Gamma}}^{*(\mathrm{m})}_{\mathrm{f},t-1}\\ \hat{\boldsymbol{\gamma}}^{*(\mathrm{m})}_{\mathrm{z},t-1}\end{bmatrix} = \begin{bmatrix}\tilde{\boldsymbol{\Gamma}}_{\mathrm{f},t-1}\\ \tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1}\end{bmatrix} &+ \left(\begin{bmatrix} \hat{\mathbf{X}}'\hat{\mathbf{X}} - N\left(\hat{\boldsymbol{\Lambda}}_1+\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{1,t-1}\right) & \hat{\mathbf{X}}'\mathbf{Z}_{t-1} - N\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{3,t-1}\\ \mathbf{Z}'_{t-1}\hat{\mathbf{X}} & \mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}\end{bmatrix}\right)^{-1}\\
&\times \Bigg[\begin{bmatrix} N\left(\hat{\boldsymbol{\Lambda}}_1+\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{1,t-1}\right)\\ \mathbf{0}_{K_{\mathrm{z}}\times(K_{\mathrm{f}}+1)}\end{bmatrix}\tilde{\boldsymbol{\Gamma}}_{\mathrm{f},t-1} + \begin{bmatrix} N\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{3,t-1}\\ \mathbf{0}_{K_{\mathrm{z}}\times K_{\mathrm{z}}}\end{bmatrix}\tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1}\\
&+ \begin{bmatrix}\hat{\mathbf{X}}'\\ \mathbf{Z}'_{t-1}\end{bmatrix}\left(\mathbf{m}_{t-1}+\boldsymbol{\epsilon}_t+(\mathbf{X}-\hat{\mathbf{X}})\tilde{\boldsymbol{\Gamma}}_{\mathrm{f},t-1}\right) - \begin{bmatrix} N\left(\hat{\boldsymbol{\Lambda}}_{2,t-1}+\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{2,t-1}\right)\\ \mathbf{0}_{K_{\mathrm{z}}}\end{bmatrix}\Bigg].
\end{aligned}
$$

Concerning the bias terms, the proof follows the corresponding part of the proof of Theorem 1, except for the additional terms arising from

$$\frac{1}{N}\hat{\mathbf{X}}'\mathbf{m}_{t-1} - \hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{2,t-1} + \hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{1,t-1}\tilde{\boldsymbol{\Gamma}}_{\mathrm{f},t-1} + \hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{3,t-1}\tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1},$$

which are equal (excluding the zero elements in the first row) to:

$$
\begin{aligned}
&\frac{1}{N}(\hat{\mathbf{B}}-\mathbf{B})'\mathbf{m}_{t-1} - \mathbf{P}'\begin{bmatrix}\mathbf{M}_{11}^{-1}\left((\mathbf{I}_{t-2},\mathbf{0}_{t-2\times T-t+1})\frac{1}{N}\hat{\boldsymbol{\epsilon}}\mathbf{R}_t\right)\\ \mathbf{0}_{T-t+1}\end{bmatrix}\\
&\quad+\mathbf{P}'\begin{bmatrix}\mathbf{M}_{11}^{-1}\left((\mathbf{I}_{t-2},\mathbf{0}_{t-2\times T-t+1})\frac{1}{N}\hat{\boldsymbol{\epsilon}}\hat{\mathbf{X}}\right)\\ \mathbf{0}_{T-t+1\times(K_f+1)}\end{bmatrix}\tilde{\boldsymbol{\Gamma}}_{\mathrm{f},t-1}\\
&\quad+\mathbf{P}'\begin{bmatrix}\mathbf{M}_{11}^{-1}\left((\mathbf{I}_{t-2},\mathbf{0}_{t-2\times T-t+1})\frac{1}{N}\hat{\boldsymbol{\epsilon}}\mathbf{Z}_{t-1}\right)\\ \mathbf{0}_{T-t+1\times K_z}\end{bmatrix}\tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1}\\
&\quad+\mathbf{P}'\begin{bmatrix}\mathbf{M}_{11}^{-1}\left(\hat{\sigma}^2(\mathbf{I}_{t-2},\mathbf{0}_{t-2\times T-t+1})\mathbb{M}_{\mathrm{D}}\boldsymbol{\iota}_{t-1,T-1}\right)\\ \mathbf{0}_{T-t+1}\end{bmatrix}\\
&= -\mathbf{P}'\begin{bmatrix}\mathbf{M}_{11}^{-1}\mathbf{M}_{12}\\ \mathbf{I}_{T-t+1}\end{bmatrix}\frac{1}{N}\boldsymbol{\epsilon}_{t,T-1}\mathbf{m}_{t-1}\\
&\quad-\mathbf{P}'\begin{bmatrix}\mathbf{M}_{11}^{-1}\left((\mathbf{I}_{t-2},\mathbf{0}_{t-2\times T-t+1})\left(\frac{1}{N}\sum_{i=1}^N\mathbf{M}_{\tilde{D}_i}\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i'-\hat{\sigma}^2\mathbb{M}_{\mathrm{D}}\right)\mathbf{Q}_{t-1,T-1}\right)\\ \mathbf{0}_{T-t+1}\end{bmatrix}\\
&= O_p(N^{-1/2}),
\end{aligned}
$$

recalling that $(\hat{\mathbf{B}}-\mathbf{B})'\mathbf{m}_{t-1} = \hat{\mathbf{B}}'\mathbf{m}_{t-1}$, $N^{-1}\sum_{i=1}^N\mathbf{M}_{\tilde{D}_i}\boldsymbol{\epsilon}_i m_{i,t-1} = \mathbb{M}_D(N^{-1}\sum_{i=1}^N\boldsymbol{\epsilon}_i m_{i,t-1}) + o_p(1)$, and $\boldsymbol{\epsilon} = \begin{bmatrix}\boldsymbol{\epsilon}_{2,t-1}\\ \boldsymbol{\epsilon}_{t,T-1}\end{bmatrix}$, with $\boldsymbol{\epsilon}_{2,t-1} = (\boldsymbol{\epsilon}_2,\ldots,\boldsymbol{\epsilon}_{t-1})'$, $\boldsymbol{\epsilon}_{t,T-1} = (\boldsymbol{\epsilon}_t,\ldots,\boldsymbol{\epsilon}_{T-1})'$. Notice that the bias adjustment corresponding to the anomalies $\mathbf{Z}_{t-1}$ cancels out for every finite N.

Regarding part (ii), the limiting distribution of $\hat{\mathbf{\Gamma}}^{*(\mathrm{m})}_{\mathrm{f},t-1}$ and $\hat{\boldsymbol{\gamma}}^{*(\mathrm{m})}_{\mathrm{z},t-1}$, and their (joint) asymptotic covariance matrix follow by

$$\begin{aligned}
&\sqrt{N}\begin{bmatrix}\hat{\mathbf{\Gamma}}^{*(\mathrm{m})}_{\mathrm{f},t-1}-\tilde{\mathbf{\Gamma}}_{\mathrm{f},t-1}\\ \hat{\boldsymbol{\gamma}}^{*(\mathrm{m})}_{\mathrm{z},t-1}-\tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1}\end{bmatrix}\\
&=\begin{bmatrix}\frac{\hat{\mathbf{X}}'\hat{\mathbf{X}}}{N}-(\hat{\mathbf{\Lambda}}_1+\hat{\mathbf{\Lambda}}^{(\mathrm{m})}_{1,t-1}) & \frac{\hat{\mathbf{X}}'\mathbf{Z}_{t-1}}{N}-\hat{\mathbf{\Lambda}}^{(\mathrm{m})}_{3,t-1}\\ \frac{\mathbf{Z}'_{t-1}\hat{\mathbf{X}}}{N} & \frac{\mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}}{N}\end{bmatrix}^{-1}\Big(\mathbf{a}_1+\mathbf{a}_2+\mathbf{a}_3+\mathbf{a}_4+\mathbf{a}_5\Big),
\end{aligned}$$

where

$$\mathbf{a}_1=\begin{bmatrix}\frac{\mathbf{1}'_N\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}\\ \frac{\mathbf{B}'\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}\\ \mathbf{0}_{K_z}\end{bmatrix},\qquad \mathbf{a}_2=\begin{bmatrix}0\\ \mathbf{P}'\frac{\boldsymbol{\epsilon}\boldsymbol{\epsilon}'}{\sqrt{N}}\mathbf{Q}_{t-1}-\sqrt{N}\,\hat{\sigma}^2\mathbf{P}'\mathbf{Q}_{t-1}\\ \mathbf{0}_{K_z}\end{bmatrix},$$

$$\mathbf{a}_3=\begin{bmatrix}0\\ \mathbf{0}_{K_f}\\ \sqrt{N}\left(\frac{\mathbf{Z}'_{t-1}\boldsymbol{\epsilon}'}{N}\right)\mathbf{Q}_{t-1}\end{bmatrix},\qquad \mathbf{a}_4=\begin{bmatrix}0\\ -\mathbf{P}'\begin{bmatrix}\mathbf{M}_{11}^{-1}\mathbf{M}_{12}\\ \mathbf{I}_{T-t+1}\end{bmatrix}\frac{1}{\sqrt{N}}\boldsymbol{\epsilon}_{t,T-1}\mathbf{m}_{t-1}\\ \mathbf{0}_{K_z}\end{bmatrix},\quad\text{and}$$

$$\mathbf{a}_5=\begin{bmatrix}0\\ -\mathbf{P}'\begin{bmatrix}\mathbf{M}_{11}^{-1}\Big((\mathbf{I}_{t-2},\mathbf{0}_{t-2\times T-t+1})\sqrt{N}\Big(\frac{1}{N}\sum_{i=1}^N\mathbf{M}_{\tilde{D}_i}\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}'_i-\hat{\sigma}^2\mathbf{M}_D\Big)\mathbf{Q}_{t-1,T-1}\Big)\\ \mathbf{0}_{T-t+1}\end{bmatrix}\\ \mathbf{0}_{K_z}\end{bmatrix}.$$

We follow the corresponding part of the proof of Theorem OA.1 and rely on their results, adding the derivations of Var[$\mathbf{a}_4$], Var[$\mathbf{a}_5$], as well as the covariances of $\mathbf{a}_4$ and $\mathbf{a}_5$ with $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$, thus obtaining:

$$\mathrm{Var}[\mathbf{a}_4]\;=\;\sigma^2\sigma_{t-1mm}\begin{bmatrix}0 & \mathbf{0}'_{K_f} & \mathbf{0}'_{K_z}\\ \mathbf{0}_{K_f} & \mathbf{P}'\begin{bmatrix}\mathbf{M}_{11}^{-1}\mathbf{M}_{12}\\ \mathbf{I}_{T-t+1}\end{bmatrix}\begin{bmatrix}\mathbf{M}'_{12}\mathbf{M}_{11}^{-1} & \mathbf{I}_{T-t+1}\end{bmatrix}\mathbf{P} & \mathbf{0}_{K_f\times K_z}\\ \mathbf{0}_{K_z} & \mathbf{0}_{K_z\times K_f} & \mathbf{0}_{K_z\times K_z}\end{bmatrix},$$

$$\mathrm{Var}[\mathbf{a}_5]\;=\;\begin{bmatrix}0 & \mathbf{0}'_{K_f} & \mathbf{0}'_{K_z}\\ \mathbf{0}_{K_f} & \mathbf{A}_{t-1}\mathbf{U}_\epsilon\mathbf{A}'_{t-1} & \mathbf{0}_{K_f\times K_z}\\ \mathbf{0}_{K_z} & \mathbf{0}_{K_z\times K_f} & \mathbf{0}_{K_z\times K_z}\end{bmatrix},\quad\text{with covariances}$$

$$
\begin{aligned}
\mathrm{Cov}[\mathbf{a}_4, \mathbf{a}_1'] &= \mathbf{0}_{K_z+K_f+1\times K_z+K_f+1},\\
\mathrm{Cov}[\mathbf{a}_4, \mathbf{a}_3'] &\rightarrow \mathbf{0}_{K_z+K_f+1\times K_z+K_f+1},\\
\mathrm{Cov}[\mathbf{a}_5, \mathbf{a}_1'] &= \mathbf{0}_{K_z+K_f+1\times K_z+K_f+1},\\
\mathrm{Cov}[\mathbf{a}_5, \mathbf{a}_3'] &\rightarrow_p \mathbf{0}_{K_f+K_z+1\times K_f+K_z+1},
\end{aligned}
$$

$$
\mathrm{Cov}[\mathbf{a}_4, \mathbf{a}_2'] = \begin{bmatrix} 0 & \mathbf{0}_{K_f}' & \mathbf{0}_{K_z}' \\ \mathbf{0}_{K_f} & -\mathbf{P}'\begin{bmatrix}\mathbf{M}_{11}^{-1}\mathbf{M}_{12}\\ \mathbf{I}_{T-t+1}\end{bmatrix}\begin{bmatrix}\mathbf{0}_{T-t+1\times(t-2)^2} & (\boldsymbol{\theta}_{t-1,m}' \otimes \sigma^2\mathbf{I}_{T-1}) & (\sigma^2\mathbf{I}_{T-1}\otimes\boldsymbol{\theta}_{t-1,m}') & \mathbf{0}_{T-t+1\times(T-t+1)^2}\end{bmatrix}\mathbf{V}_{t-1} & \mathbf{0}_{K_f\times K_z} \\ \mathbf{0}_{K_z} & \mathbf{0}_{K_z\times K_f} & \mathbf{0}_{K_z\times K_z}\end{bmatrix},
$$

$$
\mathrm{Cov}[\mathbf{a}_4, \mathbf{a}_5'] = \begin{bmatrix} 0 & \mathbf{0}_{K_f}' & \mathbf{0}_{K_z}' \\ \mathbf{0}_{K_f} & -\mathbf{P}'\begin{bmatrix}\mathbf{M}_{11}^{-1}\mathbf{M}_{12}\\ \mathbf{I}_{T-t+1}\end{bmatrix}\begin{bmatrix}\mathbf{0}_{T-t+1\times(t-2)^2} & (\boldsymbol{\theta}_{t-1,m}' \otimes \sigma^2\mathbf{I}_{T-1}) & (\sigma^2\mathbf{I}_{T-1}\otimes\boldsymbol{\theta}_{t-1,m}') & \mathbf{0}_{T-t+1\times(T-t+1)^2}\end{bmatrix}\mathbf{A}_{t-1}' & \mathbf{0}_{K_f\times K_z} \\ \mathbf{0}_{K_z} & \mathbf{0}_{K_z\times K_f} & \mathbf{0}_{K_z\times K_z}\end{bmatrix},
$$

and

$$
\mathrm{Cov}[\mathbf{a}_5, \mathbf{a}_2'] = \begin{bmatrix} 0 & \mathbf{0}_{K_f}' & \mathbf{0}_{K_z}' \\ \mathbf{0}_{K_f} & \mathbf{A}_{t-1}\mathbf{U}_\epsilon\mathbf{V}_{t-1} & \mathbf{0}_{K_f\times K_z} \\ \mathbf{0}_{K_z} & \mathbf{0}_{K_z\times K_f} & \mathbf{0}_{K_z\times K_z}\end{bmatrix},
$$

recalling that $\mathbf{m}_{t-1}'\mathbf{1}_N = 0$ by construction and that $\frac{\mathbf{m}_{t-1}'\mathbf{m}_{t-1}}{N} \rightarrow_p \sigma_{t-1,mm}$ by our assumptions, and where we set

$$
\mathbf{A}_{t-1} \equiv -\mathbf{P}'\begin{bmatrix}(\mathbf{Q}_{t-1,T-1}' \otimes \mathbf{M}_{11}^{-1}(\mathbf{I}_{t-2}, \mathbf{0}_{t-2\times T-t+1})\mathbf{M}_D)(\mathbf{I}_{(T-1)^2} - vec(\mathbf{I}_{T-1})\frac{vec'(\mathbf{M}_{\tilde{D}})}{T-K_f-K_z2}) \\ \mathbf{0}_{T-t+1\times(T-1)^2}\end{bmatrix}.
$$

Notice also that, for $\mathrm{Cov}[\mathbf{a}_4, \mathbf{a}_3]$, we used the result $N^{-1}\sum_{i=1}^N m_{i,t-1}\,\boldsymbol{\epsilon}_{t,T-1}\boldsymbol{\epsilon}'\mathbf{Q}_{t-1}\mathbf{z}_{i,t-1}' \rightarrow_p \mathbf{0}_{T-t+1\times K_z}$. Collecting all terms, one finally obtains

$$
\begin{aligned}
\mathbf{M}_{t-1}^{(\mathrm{m})} &= \mathrm{Var}[\mathbf{a}_1] + \mathrm{Var}[\mathbf{a}_2] + \mathrm{Var}[\mathbf{a}_3] + \mathrm{Var}[\mathbf{a}_4] + \mathrm{Var}[\mathbf{a}_5]\\
&+ \mathrm{Cov}[\mathbf{a}_3, \mathbf{a}_1'] + \mathrm{Cov}[\mathbf{a}_1, \mathbf{a}_3'] + \mathrm{Cov}[\mathbf{a}_4, \mathbf{a}_2'] + \mathrm{Cov}[\mathbf{a}_2, \mathbf{a}_4']\\
&+ \mathrm{Cov}[\mathbf{a}_5, \mathbf{a}_2'] + \mathrm{Cov}[\mathbf{a}_2, \mathbf{a}_5'] + \mathrm{Cov}[\mathbf{a}_4, \mathbf{a}_5'] + \mathrm{Cov}[\mathbf{a}_5, \mathbf{a}_4'].
\end{aligned}
$$

The asymptotic covariance matrix of $\hat{\boldsymbol{\Gamma}}^{*(\mathrm{m})}_{\mathrm{f},t-1}$ and $\hat{\gamma}^{*(\mathrm{m})}_{\mathrm{z},t-1}$ is given by $\mathbf{L}^{-1}_{t-1}\left(\mathbf{O}_{t-1}+\mathbf{M}^{(\mathrm{m})}_{t-1}\right)\mathbf{L}^{-1}_{t-1}$, where $\mathbf{L}_{t-1}$ and $\mathbf{O}_{t-1}$ are defined in (OA.84). Moreover, since $N^{-1}\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{3,t-1}=o_p(1)$, we have

$$\hat{\mathbf{L}}^{(\mathrm{m})}_{t-1} \quad = \quad \begin{bmatrix} \frac{\hat{\mathbf{X}}'\hat{\mathbf{X}}}{N}-(\hat{\boldsymbol{\Lambda}}_1+\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{1,t-1}) & \frac{\hat{\mathbf{X}}'\mathbf{Z}_{t-1}}{N}-\hat{\boldsymbol{\Lambda}}^{(\mathrm{m})}_{3,t-1} \\ \frac{\mathbf{Z}'_{t-1}\hat{\mathbf{X}}}{N} & \frac{\mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}}{N} \end{bmatrix} \quad \to_p \quad \mathbf{L}_{t-1}. \tag{OA.95}$$

■

Proof of Theorem OA.7. We need to establish the limiting distribution of (OA.47) under (i) and (ii). First, notice that, by using the definition in (OA.46), and replacing it into (OA.47), we get

$$\mathcal{T}^2_{\mathrm{z},t-1}=N\left(\frac{\hat{\gamma}^{*\prime}_{\mathrm{z},t-1}\mathbf{Z}'_{t-1}\mathbb{M}_{\hat{X}}\mathbf{Z}_{t-1}\hat{\gamma}^{*}_{\mathrm{z},t-1}}{\mathbf{R}'_t\mathbb{M}_{1_N}\mathbf{R}_t}\right). \tag{OA.96}$$

Let us start from the denominator of (OA.96) and rewrite it as $\mathbf{R}'_t\mathbb{M}_{1_N}\mathbf{R}_t=\mathbf{R}'_t\mathbf{R}_t-\mathbf{R}'_t\mathbf{1}_N\mathbf{1}'_N\mathbf{R}_t/N$. Recalling that $\mathbf{R}_t=\gamma_{0,t-1}\mathbf{1}_N+\mathbf{Z}_{t-1}\boldsymbol{\gamma}_{\mathrm{z},t-1}+\mathbf{B}\boldsymbol{\delta}_{\mathrm{f},t-1}+\boldsymbol{\epsilon}_t$, then:

$$\begin{aligned}\frac{\mathbf{R}'_t\mathbf{1}_N}{N} &= \gamma_{0,t-1}+\boldsymbol{\delta}'_{\mathrm{f},t-1}\frac{\mathbf{B}'\mathbf{1}_N}{N}+\boldsymbol{\gamma}'_{\mathrm{z},t-1}\frac{\mathbf{Z}'_{t-1}\mathbf{1}_N}{N}+\frac{\boldsymbol{\epsilon}_t{}'\mathbf{1}_N}{N}\\ &\to_p \gamma_{0,t-1}+\boldsymbol{\delta}'_{\mathrm{f},t-1}\boldsymbol{\mu}_\beta+\boldsymbol{\gamma}'_{\mathrm{z},t-1}\boldsymbol{\mu}_z.\end{aligned}$$

Moreover, using the same arguments,

$$\begin{aligned}\frac{\mathbf{R}'_t\mathbf{R}_t}{N} &\to_p \gamma^2_{0,t-1}+\boldsymbol{\delta}'_{\mathrm{f},t-1}\boldsymbol{\Sigma}_\beta\boldsymbol{\delta}_{\mathrm{f},t-1}+\boldsymbol{\gamma}'_{\mathrm{z},t-1}\mathbb{J}'_{t-1}\boldsymbol{\Sigma}_{\mathrm{Z}}\mathbb{J}_{t-1}\gamma_{\mathrm{z},t-1}+\sigma^2\\ &+ 2\gamma_{0,t-1}\boldsymbol{\mu}_\beta{}'\boldsymbol{\delta}_{\mathrm{f},t-1}+2\gamma_{0,t-1}\boldsymbol{\mu}'_z\gamma_{\mathrm{z},t-1}+2\boldsymbol{\delta}'_{\mathrm{f},t-1}\boldsymbol{\Sigma}'_{\mathrm{ZB}}\mathbb{J}_{t-1}\gamma_{\mathrm{z},t-1}\end{aligned}$$

Therefore, combining terms

$$\begin{aligned}\frac{\mathbf{R}'_t\mathbb{M}_{1_N}\mathbf{R}_t}{N} &= \frac{\mathbf{R}'_t\mathbf{R}_t}{N}-\frac{\mathbf{R}'_t\mathbf{1}_N\mathbf{1}'_N\mathbf{R}_t}{N}\\ &\to_p \boldsymbol{\delta}'_{\mathrm{f},t-1}\left(\boldsymbol{\Sigma}_\beta-\boldsymbol{\mu}_\beta\boldsymbol{\mu}'_\beta\right)\boldsymbol{\delta}_{\mathrm{f},t-1}+\boldsymbol{\gamma}'_{\mathrm{z},t-1}\left(\mathbb{J}'_{t-1}\boldsymbol{\Sigma}_{\mathrm{Z}}\mathbb{J}_{t-1}-\boldsymbol{\mu}_z\boldsymbol{\mu}'_z\right)\gamma_{\mathrm{z},t-1}\\ &+ \sigma^2+2\boldsymbol{\delta}'_{\mathrm{f},t-1}\left(\boldsymbol{\Sigma}'_{\mathrm{ZB}}\mathbb{J}_{t-1}-\boldsymbol{\mu}_\beta\boldsymbol{\mu}'_z\right)\gamma_{\mathrm{z},t-1}\cancel{\equiv}\sigma_{\tilde{\mathrm{R}},t},\end{aligned}$$

we never used equivalent symbol across paper and OA - check

implying that $\sigma_{\tilde{\mathrm{R}},t}>\sigma^2>0$.

Consider now the numerator of (OA.96) and notice that, under Assumptions OA.4–OA.8 and using Lemma 2,

$$\frac{\mathbf{Z}'_{t-1}\mathbb{M}_{\hat{X}}\mathbf{Z}_{t-1}}{N} \to_p \mathbb{J}'_{t-1}\mathbf{\Sigma}_{\mathrm{Z}}\mathbb{J}_{t-1} - \mathbf{\Sigma}_{\mathrm{ZX,t-1}}\left(\mathbf{\Sigma}_{\mathrm{X}} + \sigma^2\mathbf{P}'\mathbf{P}\right)^{-1}\mathbf{\Sigma}'_{\mathrm{ZX,t-1}} = \mathbf{\Sigma}_{\mathrm{Z\hat{X}Z},t-1}.$$

Moreover, using Theorems OA.1-(ii) and 2, we have that

$$\sqrt{N}\left(\hat{\boldsymbol{\gamma}}^*_{\mathrm{z},t-1} - \boldsymbol{\gamma}_{\mathrm{z},t-1}\right) \to_d \mathcal{N}\left(\mathbf{0}_{K_{\mathrm{z}}}, \mathbf{L}^{-1}_{\mathrm{z},t-1}\,\mathbf{O}_{t-1}\,\mathbf{L}^{-1'}_{\mathrm{z},t-1}\right),$$

with $\mathbf{L}_{\mathrm{z},t-1} = \left[\mathbf{0}_{K_{\mathrm{z}}\times(K_{\mathrm{f}}+1)}, \mathbf{I}_{K_{\mathrm{z}}}\right]\mathbf{L}_{t-1}$. Therefore, under the null hypothesis of $\boldsymbol{\gamma}_{\mathrm{z},t-1} = \mathbf{0}_{K_{\mathrm{z}}}$, and denoting for brevity $\hat{\mathbf{V}}_{\mathrm{LOL}} \equiv \hat{\mathbf{L}}^{-1}_{\mathrm{z},t-1}\,\hat{\mathbf{O}}_{t-1}\,\hat{\mathbf{L}}^{-1'}_{\mathrm{z},t-1}$, and $\mathbf{V}_{\mathrm{LOL}} \equiv \mathbf{L}^{-1}_{\mathrm{z},t-1}\,\mathbf{O}_{t-1}\,\mathbf{L}^{-1'}_{\mathrm{z},t-1}$, with $\hat{\mathbf{L}}_{\mathrm{z},t-1} = \left[\mathbf{0}_{K_{\mathrm{z}}\times(K_{\mathrm{f}}+1)}, \mathbf{I}_{K_{\mathrm{z}}}\right]\hat{\mathbf{L}}_{t-1}$, we have that $\sqrt{N}\left(\hat{\mathbf{V}}_{\mathrm{LOL}}\right)^{-\frac{1}{2}}\hat{\boldsymbol{\gamma}}^*_{\mathrm{z},t-1} \to_d \mathcal{N}(\mathbf{0}_{K_{\mathrm{z}}}, \mathbf{I}_{K_{\mathrm{z}}})$, and

$$\frac{\left(\hat{\mathbf{V}}_{\mathrm{LOL}}\right)^{\frac{1}{2}}\left(\frac{\mathbf{Z}'_{t-1}\mathbb{M}_{\hat{X}}\mathbf{Z}_{t-1}}{N}\right)\left(\hat{\mathbf{V}}_{\mathrm{LOL}}\right)^{\frac{1}{2}}}{\mathbf{R}'_t\mathbb{M}_{1_N}\mathbf{R}_t/N} \to_p \frac{\left(\mathbf{V}_{\mathrm{LOL}}\right)^{\frac{1}{2}}\mathbf{\Sigma}_{\mathrm{Z\hat{X}Z},t-1}\left(\mathbf{V}_{\mathrm{LOL}}\right)^{\frac{1}{2}}}{\sigma_{\tilde{\mathrm{R}},t}} = \mathbf{\Theta}_{t-1},$$

with $\mathbf{\Theta}_{t-1}$ being a symmetric and positive definite matrix admitting the spectral decomposition $\mathbf{\Theta}_{t-1} = \mathbf{\Delta}_{1,t-1}\mathbf{\Delta}_{2,t-1}\mathbf{\Delta}'_{1,t-1}$, for an orthogonal matrix $\mathbf{\Delta}_{1,t-1}$ and a diagonal matrix $\mathbf{\Delta}_{2,t-1} = \mathrm{diag}(d_{1,t-1}, \cdots, d_{K_{\mathrm{z}},t-1})$, whose diagonal elements correspond to the eigenvalues of $\mathbf{\Theta}_{t-1}$. Therefore, combining these results into (OA.96), we have

$$\begin{aligned}
\mathcal{T}^2_{\mathrm{z},t-1} &\equiv \frac{\sqrt{N}\hat{\boldsymbol{\gamma}}^{*\prime}_{\mathrm{z},t-1}\left(\frac{\mathbf{Z}'_{t-1}\mathbb{M}_{\hat{X}}\mathbf{Z}_{t-1}}{N}\right)\sqrt{N}\hat{\boldsymbol{\gamma}}^*_{\mathrm{z},t-1}}{\mathbf{R}'_t\mathbb{M}_{1_N}\mathbf{R}_t/N} \\
&= \frac{\sqrt{N}\hat{\boldsymbol{\gamma}}^{*\prime}_{\mathrm{z},t-1}\left(\hat{\mathbf{V}}_{\mathrm{LOL}}\right)^{-\frac{1}{2}}\left(\hat{\mathbf{V}}_{\mathrm{LOL}}\right)^{\frac{1}{2}}\left(\frac{\mathbf{Z}'_{t-1}\mathbb{M}_{\hat{X}}\mathbf{Z}_{t-1}}{N}\right)\left(\hat{\mathbf{V}}_{\mathrm{LOL}}\right)^{\frac{1}{2}}\left(\hat{\mathbf{V}}_{\mathrm{LOL}}\right)^{-\frac{1}{2}}\sqrt{N}\hat{\boldsymbol{\gamma}}^*_{\mathrm{z},t-1}}{\mathbf{R}'_t\mathbb{M}_{1_N}\mathbf{R}_t/N} \\
&\to_d \quad \xi'\mathbf{\Theta}_{t-1}\xi = \xi'\mathbf{\Delta}_{1,t-1}\mathbf{\Delta}_{2,t-1}\mathbf{\Delta}'_{1,t-1}\xi = \sum_{i=1}^{K_{\mathrm{z}}}\chi^2_{1,j}d_{j,t-1}
\end{aligned}$$

denoting $\xi \sim \mathcal{N}(\mathbf{0}'_{K_{\mathrm{z}}}, \mathbf{I}_{K_{\mathrm{z}}})$, and where the last equality follows by noticing that $\xi\mathbf{\Delta}_{1,t-1} \sim \mathcal{N}(\mathbf{0}_{K_{\mathrm{z}}}, \mathbf{I}_{K_{\mathrm{z}}})$, given $\mathbf{\Delta}'_{1,t-1}\mathbf{I}_{K_{\mathrm{z}}}\mathbf{\Delta}_{1,t-1} = \mathbf{I}_{K_{\mathrm{z}}}$. Finally, notice that, under the null hypothesis of part (i),

$$\sigma_{\tilde{\mathrm{R}},t} = \boldsymbol{\delta}'_{\mathrm{f},t-1}\left(\mathbf{\Sigma}_{\beta} - \boldsymbol{\mu}_{\beta}\boldsymbol{\mu}'_{\beta}\right)\boldsymbol{\delta}_{\mathrm{f},t-1} + \sigma^2.$$

This concludes the proof of part (i).

To prove part (ii), let us first define the following $d \times 1$ random vector

$$\boldsymbol{\theta} = \left(\mathrm{vec}\left(\frac{\mathbf{Z}'\mathbf{Z}}{N}\right), \frac{\mathbf{Z}'\mathbf{1}_N}{N}, \mathrm{vec}\left(\frac{\mathbf{Z}'\mathbf{B}}{N}\right), \mathrm{vec}\left(\frac{\mathbf{Z}'\boldsymbol{\epsilon}'}{N}\right), \mathrm{vec}\left(\frac{\mathbf{B}'\boldsymbol{\epsilon}'}{N}\right), \mathrm{vec}\left(\frac{\boldsymbol{\epsilon}\boldsymbol{\epsilon}'}{N}\right), \frac{\boldsymbol{\epsilon}'\mathbf{1}_N}{N} \right) \quad \text{(OA.97)}$$

with $d = ((T-1)K_{\mathrm{z}})^2 + (T-1)K_{\mathrm{z}} + (T-1)K_{\mathrm{z}}K_{\mathrm{f}} + (T-1)^2 K_{\mathrm{z}} + (T-1)K_{\mathrm{f}} + (T-1)^2 + (T-1)$. Then, consider $\hat{R}^2_{\mathrm{z},t-1}$ in (47) and notice that it can be written as a function of $\boldsymbol{\theta}$, such that

$$\hat{R}^2_{\mathrm{z},t-1} = g_{t-1}(\boldsymbol{\theta}),$$

with $g_{t-1}(\cdot)$ being an elementary and differentiable function, made by simple products and ratios of the arguments in $\boldsymbol{\theta}$.[34] All the random quantities in $\boldsymbol{\theta}$ admit a continuous second-order derivative, implying

$$\sqrt{N} \begin{bmatrix} \mathrm{vec}\left(\frac{\mathbf{Z}'\mathbf{Z}}{N} - \boldsymbol{\Sigma}_{\mathrm{Z}}\right) \\ \frac{\mathbf{Z}'\mathbf{1}_N}{N} - \boldsymbol{\mu}_{\mathrm{z},T-1} \\ \mathrm{vec}\left(\frac{\mathbf{Z}'\mathbf{B}}{N} - \boldsymbol{\Sigma}_{\mathrm{ZB}}\right) \\ \mathrm{vec}\left(\frac{\mathbf{Z}'\boldsymbol{\epsilon}'}{N}\right) \\ \mathrm{vec}\left(\frac{\mathbf{B}'\boldsymbol{\epsilon}'}{N}\right) \\ \mathrm{vec}\left(\frac{\boldsymbol{\epsilon}\boldsymbol{\epsilon}'}{N} - \frac{\sigma^2}{T-1}\right) \\ \frac{\boldsymbol{\epsilon}'\mathbf{1}_N}{N} \end{bmatrix} \to_d \mathcal{N}(\mathbf{0}_d, \mathbf{V}_{\boldsymbol{\theta}}),$$

with the $d \times d$ covariance matrix $\mathbf{V}_{\boldsymbol{\theta}}$ having the following form

$$\mathbf{V}_{\boldsymbol{\theta}} = \begin{bmatrix} \mathrm{Var}[\boldsymbol{\theta}_1] & \mathrm{Cov}[\boldsymbol{\theta}_1, \boldsymbol{\theta}_2'] & \cdots & \mathrm{Cov}[\boldsymbol{\theta}_1, \boldsymbol{\theta}_7'] \\ \mathrm{Cov}[\boldsymbol{\theta}_2, \boldsymbol{\theta}_1'] & \mathrm{Var}[\boldsymbol{\theta}_2] & \cdots & \mathrm{Cov}[\boldsymbol{\theta}_2, \boldsymbol{\theta}_7'] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}[\boldsymbol{\theta}_7, \boldsymbol{\theta}_1'] & \mathrm{Cov}[\boldsymbol{\theta}_7, \boldsymbol{\theta}_2'] & \cdots & \mathrm{Var}[\boldsymbol{\theta}_7] \end{bmatrix}, \quad \text{(OA.98)}$$

setting $\boldsymbol{\theta}_1 = \sqrt{N}\mathrm{vec}\left(\frac{\mathbf{Z}'\mathbf{Z}}{N} - \boldsymbol{\Sigma}_{\mathrm{Z}}\right)$, $\boldsymbol{\theta}_2 = \sqrt{N}\left(\frac{\mathbf{Z}'\mathbf{1}_N}{N} - \boldsymbol{\mu}_{\mathrm{z},T-1}\right)$, $\boldsymbol{\theta}_3 = \sqrt{N}\mathrm{vec}\left(\frac{\mathbf{Z}'\mathbf{B}}{N} - \boldsymbol{\Sigma}_{\mathrm{ZB}}\right)$, $\boldsymbol{\theta}_4 = \mathrm{vec}\left(\frac{\mathbf{Z}'\boldsymbol{\epsilon}'}{\sqrt{N}}\right)$, $\boldsymbol{\theta}_5 = \mathrm{vec}\left(\frac{\mathbf{B}'\boldsymbol{\epsilon}'}{\sqrt{N}}\right)$, $\boldsymbol{\theta}_6 = \sqrt{N}\mathrm{vec}\left(\frac{\boldsymbol{\epsilon}\boldsymbol{\epsilon}'}{N} - \frac{\sigma^2}{T-1}\right)$, and $\boldsymbol{\theta}_7 = \frac{\boldsymbol{\epsilon}'\mathbf{1}_N}{\sqrt{N}}$.

Now, by Assumptions OA.8 and OA.14, $\mathrm{Var}(\boldsymbol{\theta}_1) = \mathrm{Var}\left[\mathrm{vec}\left(\frac{\mathbf{Z}'\mathbf{Z}}{\sqrt{N}}\right)\right] \to \mathbf{U}_{\mathrm{Z}}$, $\mathrm{Var}(\boldsymbol{\theta}_2) = \mathrm{Var}\left[\frac{\mathbf{Z}'\mathbf{1}_N}{\sqrt{N}}\right] \to \boldsymbol{\Sigma}_{\mathrm{Z}} - \boldsymbol{\mu}_{\mathrm{z},T-1}\boldsymbol{\mu}_{\mathrm{z},T-1}'$, $\mathrm{Var}(\boldsymbol{\theta}_3) = \mathrm{Var}\left[\mathrm{vec}\left(\frac{\mathbf{Z}'\mathbf{B}}{\sqrt{N}}\right)\right] \to (\boldsymbol{\Sigma}_{\beta} \otimes \boldsymbol{\Sigma}_{\mathrm{Z}}) - \left(\boldsymbol{\mu}_{\beta}\boldsymbol{\mu}_{\beta}' \otimes \boldsymbol{\mu}_{\mathrm{z},T-1}\boldsymbol{\mu}_{\mathrm{z},T-1}'\right) = \tilde{\boldsymbol{\Sigma}}_{\beta\otimes\mathrm{Z}}$, $\mathrm{Var}(\boldsymbol{\theta}_4) = \mathrm{Var}\left[\mathrm{vec}\left(\frac{\mathbf{Z}'\boldsymbol{\epsilon}'}{\sqrt{N}}\right)\right] \to (\sigma^2\mathbf{I}_{T-1} \otimes \boldsymbol{\Sigma}_{\mathrm{Z}})$, $\mathrm{Var}(\boldsymbol{\theta}_5) = \mathrm{Var}\left[\mathrm{vec}\left(\frac{\mathbf{B}'\boldsymbol{\epsilon}'}{\sqrt{N}}\right]\right) \to (\sigma^2\mathbf{I}_{T-1} \otimes \boldsymbol{\Sigma}_{\beta})$, $\mathrm{Var}(\boldsymbol{\theta}_6) = \mathrm{Var}\left[\mathrm{vec}\left(\frac{\boldsymbol{\epsilon}\boldsymbol{\epsilon}'}{\sqrt{N}}\right)\right] \to \mathbf{U}_{\epsilon}$, and $\mathrm{Var}(\boldsymbol{\theta}_7) = \mathrm{Var}\left[\frac{\boldsymbol{\epsilon}'\mathbf{1}_N}{\sqrt{N}}\right] \to \sigma^2$.

[34] To ease the exposition, we do not repor the $g_{t-1}(.)$ function, because it is elementary. Details are available upon request.

Consider now all the covariance terms. Under Assumptions OA.7, OA.8 and OA.14, we have $\text{Cov}[\boldsymbol{\theta}_1, \boldsymbol{\theta}_2'] \to \boldsymbol{\Sigma}_{\text{z}\otimes\text{z}}$, $\text{Cov}[\boldsymbol{\theta}_1, \boldsymbol{\theta}_3'] \to \boldsymbol{\Sigma}_{\text{z}\otimes\text{z}} \left(\boldsymbol{\mu}_\beta' \otimes \mathbf{I}_{K_\text{z}}\right)$, $\text{Cov}[\boldsymbol{\theta}_2, \boldsymbol{\theta}_3'] \to \boldsymbol{\mu}_\beta' \otimes \left(\boldsymbol{\Sigma}_\text{Z} - \boldsymbol{\mu}_{\text{z},T-1}\boldsymbol{\mu}_{\text{z},T-1}'\right)$, $\text{Cov}[\boldsymbol{\theta}_4, \boldsymbol{\theta}_5'] \to \sigma^2 \mathbf{I}_{T-1} \otimes \boldsymbol{\mu}_{\text{z},T-1}\boldsymbol{\mu}_\beta'$, $\text{Cov}[\boldsymbol{\theta}_4, \boldsymbol{\theta}_7'] \to \sigma^2 \mathbf{I}_{T-1} \otimes \boldsymbol{\mu}_{\text{z},T-1}'$, and $\text{Cov}[\boldsymbol{\theta}_5, \boldsymbol{\theta}_7'] \to \sigma^2 \mathbf{I}_{T-1} \otimes \boldsymbol{\mu}_\beta'$. Moreover, under Assumption OA.6, it follows that all the remaining covariance terms are zero matrices. Putting all together gives $\mathbf{V}_{\boldsymbol{\theta}}$.

Therefore, by the mean-value theorem, it follows that

$$\sqrt{N}(\hat{R}^2_{\text{z},t-1} - R^2_{\text{z},t-1}) \to_d \mathcal{N}(0, \omega_{\text{z},t-1}) \quad \text{with} \quad \omega_{\text{z},t-1} = \mathbf{G}_{t-1}' \mathbf{V}_{\boldsymbol{\theta}} \mathbf{G}_{t-1}, \tag{OA.99}$$

setting $\mathbf{G}_{t-1}(\mathbf{x}) = \partial g_{t-1}(\mathbf{x})/\partial \mathbf{x}$ for a generic d-dimensional vector $\mathbf{x}$, and $\mathbf{G}_{t-1} = \mathbf{G}_{t-1}(\mathbf{x}_0)$ with [35]

$$\mathbf{x}_0 = \left(\text{vec}'\left(\boldsymbol{\Sigma}_\text{Z}\right), \boldsymbol{\mu}_{\text{z},T-1}', \text{vec}'(\boldsymbol{\Sigma}_{\text{ZB}}), \mathbf{0}_{(T-1)^2K_\text{z}}', \mathbf{0}_{(T-1)K_\text{f}}', \text{vec}'(\sigma^2\mathbf{I}_{T-1}), \mathbf{0}_{(T-1)}'\right)'.$$

Finally, a consistent estimator $\hat{\omega}_{\text{z},t-1}$ for $\omega_{\text{z},t-1}$ is obtained by replacing $\mathbf{G}_{t-1}$ with

$$\hat{\mathbf{G}}_{t-1} = \mathbf{G}_{t-1}\left((\text{vec}'\left(\frac{\mathbf{Z}'\mathbf{Z}}{N}\right), \text{vec}'\left(\frac{\mathbf{Z}'\mathbf{1}_N}{N}\right), \text{vec}'\left(\frac{\mathbf{Z}'\hat{\mathbf{B}}}{N}\right), \mathbf{0}_{(T-1)^2K_\text{z}}', \mathbf{0}_{(T-1)K_\text{f}}', \text{vec}'(\hat{\sigma}^2\mathbf{I}_{T-1}))', \mathbf{0}_{(T-1)}'\right),$$

and replacing the terms of $\mathbf{V}_{\boldsymbol{\theta}}$ with their sample counterparts, yielding

$$\hat{\omega}_{\text{z},t-1} = \hat{\mathbf{G}}_{t-1}' \hat{\mathbf{V}}_{\boldsymbol{\theta}} \hat{\mathbf{G}}_{t-1}. \tag{OA.100}$$

$\blacksquare$

Remark OA.22. *Theorem 3 in the main paper immediately follows as a special case of Theorem OA.7 when all the corresponding quantities that apply to $\gamma_{0,t-1}$ are replaced with the corresponding quantities obtained by setting $\gamma_{0,t-1}$ equal to the risk-free rate.*

OA.6 Locally-Constant Betas

This section establishes that, under Assumption 1, namely

$$\frac{(\mathbf{B}_s - \mathbf{B})'(\mathbf{B}_s - \mathbf{B})}{N} = o(N^{-1/2}), \tag{OA.101}$$

[35] The mapping from vectorized matrices to the original matrices is given by the linear function $\text{vec}^{-1}_{m\times p}(\mathbf{x}) = (\text{vec}'(\mathbf{I}_p) \otimes \mathbf{I}_m)(\mathbf{I}_p \otimes \mathbf{x})$ mapping the $mp\times$ vector $\mathbf{x}$ into a $m \times p$ matrix.

the entire asymptotic analysis remains unchanged, as if the risk exposures were constant.

Our smoothing assumption is very general and accommodates a wide variety of time-varying patterns for the risk exposures. An important example is the case where

$$\boldsymbol{\beta}_{i,s} = \boldsymbol{\beta}_i + \mathbf{B}_{1i}\mathbf{g}_s + \mathbf{B}_{2i}\mathbf{z}_{is}, \tag{OA.102}$$

for coefficient matrices $\boldsymbol{\beta}_{0i}$ $(K_{\mathrm{f}} \times 1)$, $\mathbf{B}_{1i}$ $(K_{\mathrm{f}} \times K_{\mathrm{g}})$, and $\mathbf{B}_{2i}$ $(K_{\mathrm{f}} \times K_{\mathrm{z}})$, such that

$$\sum_{i=1}^{N} (\mathbf{B}_{1i} \otimes \mathbf{B}_{1i}) = o(N^{1/2}) \quad \text{and} \quad \sum_{i=1}^{N} (\mathbf{B}_{2i} \otimes \mathbf{B}_{2i})(\mathbf{z}_{is} \otimes \mathbf{z}_{is}) = o(N^{1/2}).$$

These conditions, in turn, are implied when $\mathbf{B}_{1i} = \mathbf{B}^*_{1i}/(N^{1/4}\log(N))$ and $\mathbf{B}_{2i} = \mathbf{B}^*_{2i}/(N^{1/4}\log(N))$, for coefficient matrices $\mathbf{B}^*_{1i}$ and $\mathbf{B}^*_{2i}$ satisfying

$$\frac{1}{N}\sum_{i=1}^{N} (\mathbf{B}^*_{1i} \otimes \mathbf{B}^*_{1i}) = O(1), \qquad \frac{1}{N}\sum_{i=1}^{N} (\mathbf{B}^*_{2i} \otimes \mathbf{B}^*_{2i}) = O(1).$$

We will focus on our main estimator, the CSR OLS-type estimator of Section OA.2. Consider again the *conditional* asset pricing model (see (OA.6))

$$\mathbf{R}_t = \mathbf{Z}_{t-1}\boldsymbol{\gamma}_{\mathrm{z},t-1} + \mathbf{X}_{t-1}\boldsymbol{\Gamma}_{\mathrm{f},t-1} + \boldsymbol{\epsilon}_t \tag{OA.103}$$

setting $\mathbf{X}_{t-1} = (1_N, \mathbf{B}_{t-1})$, where $\boldsymbol{\Gamma}_{\mathrm{f},t-1} = (\gamma_{0,t-1}, \boldsymbol{\delta}'_{\mathrm{f},t-1})'$, with $\boldsymbol{\delta}_{\mathrm{f},t-1}$ defined in (OA.7). Then

$$\mathbf{R}_t = \mathbf{Z}_{t-1}\boldsymbol{\gamma}_{\mathrm{z},t-1} + \mathbf{X}\boldsymbol{\Gamma}_{\mathrm{f},t-1} + \boldsymbol{\epsilon}_t + (\mathbf{X}_{t-1} - \mathbf{X})\boldsymbol{\Gamma}_{\mathrm{f},t-1}, \tag{OA.104}$$

and, in matrix form,

$$\mathbf{R} = \boldsymbol{\gamma}_0 1'_N + \boldsymbol{\Delta}_z + \left(\mathbf{I}_{T-1} \otimes \boldsymbol{\gamma}'_z\right)\begin{bmatrix} \mathbf{Z}'_1 \\ \mathbf{Z}'_2 \\ \vdots \\ \mathbf{Z}'_{T-1} \end{bmatrix} + \boldsymbol{\Delta}_B + \boldsymbol{\delta}_{\mathrm{f}}\mathbf{B}' + \boldsymbol{\epsilon},$$

setting

$$\boldsymbol{\Delta}_B = \begin{bmatrix} \boldsymbol{\delta}'_{\mathrm{f},1} & \mathbf{0}'_{K_{\mathrm{f}}} & \cdots & \mathbf{0}'_{K_{\mathrm{f}}} \\ \mathbf{0}'_{K_{\mathrm{f}}} & \boldsymbol{\delta}'_{\mathrm{f},2} & \cdots & \mathbf{0}'_{K_{\mathrm{f}}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}'_{K_{\mathrm{f}}} & \mathbf{0}'_{K_{\mathrm{f}}} & \cdots & \boldsymbol{\delta}'_{\mathrm{f},T-1} \end{bmatrix} \begin{bmatrix} (\mathbf{B}_1 - \mathbf{B})' \\ (\mathbf{B}_2 - \mathbf{B})' \\ \vdots \\ (\mathbf{B}_{T-1} - \mathbf{B})' \end{bmatrix}, \quad \text{and}$$

$$\boldsymbol{\Delta}_z = \begin{bmatrix} \boldsymbol{\gamma}'_{z,1}-\boldsymbol{\gamma}'_{\rm z} & \mathbf{0}'_{K_{\rm z}} & \cdots & \mathbf{0}'_{K_{\rm z}} \\ \mathbf{0}'_{K_{\rm z}} & \boldsymbol{\gamma}'_{z,2}-\boldsymbol{\gamma}'_{\rm z} & \cdots & \mathbf{0}'_{K_{\rm z}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}'_{K_{\rm z}} & \mathbf{0}'_{K_{\rm z}} & \cdots & \boldsymbol{\gamma}'_{{\rm z},t-1}-\boldsymbol{\gamma}'_{\rm z} \end{bmatrix} \begin{bmatrix} \mathbf{Z}'_1 \\ \mathbf{Z}'_2 \\ \vdots \\ \mathbf{Z}'_{T-1} \end{bmatrix}, \boldsymbol{\delta}_{\rm f} = \begin{bmatrix} \boldsymbol{\delta}'_{\rm f,1} \\ \vdots \\ \boldsymbol{\delta}'_{\rm f,T-1} \end{bmatrix},$$

with $\boldsymbol{\Delta}_z$ defined in Assumption OA.9. Then,

$$\begin{aligned} \hat{\mathbf{B}} &= \mathbf{R}'\mathbb{M}_{\mathbf{1}_{T-1}}\mathbf{F}(\mathbf{F}'\mathbb{M}_{\mathbf{1}_{T-1}}\mathbf{F})^{-1} = \mathbf{R}'\mathbf{P} \\ &= \left(1_N\boldsymbol{\gamma}'_0 + \boldsymbol{\Delta}'_z + (\mathbf{Z}_1, \mathbf{Z}_2, \cdots, \mathbf{Z}_{T-1})(\mathbf{I}_{T-1}\otimes\boldsymbol{\gamma}_z) + \boldsymbol{\Delta}'_B + \mathbf{B}\boldsymbol{\delta}'_{\rm f} + \boldsymbol{\epsilon}'\right)\mathbf{P} \\ &= (\boldsymbol{\epsilon}' + \boldsymbol{\Delta}'_B)\mathbf{P} + \mathbf{B}, \end{aligned}$$

where the last equation follows from Assumption OA.9. The term $\boldsymbol{\Delta}'_B\mathbf{P}$ on the right-hand side of $\hat{\mathbf{B}}$ arises when considering time-varying risk exposures, but its effect will be shown to be asymptotically negligible as $N \to \infty$.

We now show that, asymptotically, for our CSR OLS estimator the results of Theorems OA.1–OA.2 are recovered by replacing $\mathbf{B}_{t-1}$ with $\mathbf{B}$. The same conclusion applies to all our asymptotic analyses of the CSR WLS estimator (Theorems OA.3–OA.4), the R^2 test (Theorem OA.7), and the misspecification-robust case (Theorems OA.5–OA.6).[36]

By inspecting our CSR OLS-type estimator, namely

$$\begin{bmatrix} \hat{\boldsymbol{\Gamma}}^*_{{\rm f},t-1} \\ \hat{\boldsymbol{\gamma}}^*_{{\rm z},t-1} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{X}}'\hat{\mathbf{X}} - N\hat{\boldsymbol{\Lambda}}_1 & \hat{\mathbf{X}}'\mathbf{Z}_{t-1} \\ \mathbf{Z}'_{t-1}\hat{\mathbf{X}} & \mathbf{Z}'_{t-1}\mathbf{Z}_{t-1} \end{bmatrix}^{-1} \begin{bmatrix} \hat{\mathbf{X}}'\mathbf{R}_t - N\hat{\boldsymbol{\Lambda}}_{2,t-1} \\ \mathbf{Z}'_{t-1}\mathbf{R}_t \end{bmatrix},$$

one sees that we need to study the following quantities (and show that they are asymptotically equivalent when the locally-constant case is considered). First, considering the terms in the matrix inverse, namely

$$\begin{aligned} &\frac{1}{N}\hat{\mathbf{B}}'\hat{\mathbf{B}} = \frac{1}{N}[\mathbf{P}'(\boldsymbol{\Delta}_B + \boldsymbol{\epsilon}) + \mathbf{B}'][(\boldsymbol{\Delta}_B + \boldsymbol{\epsilon})'\mathbf{P} + \mathbf{B}] \\ &= \underbrace{\frac{1}{N}[\mathbf{P}'\boldsymbol{\epsilon} + \mathbf{B}'][\boldsymbol{\epsilon}'\mathbf{P} + \mathbf{B}]}_{\text{(locally-constant term)}} + \underbrace{\frac{1}{N}\mathbf{P}'\boldsymbol{\Delta}_B\boldsymbol{\Delta}'_B\mathbf{P} + \frac{1}{N}\mathbf{P}'\boldsymbol{\Delta}_B(\mathbf{B} + \boldsymbol{\epsilon}'\mathbf{P}) + \frac{1}{N}(\mathbf{B}' + \mathbf{P}'\boldsymbol{\epsilon})\boldsymbol{\Delta}'_B\mathbf{P}}_{\text{(time-varying terms)}}, \end{aligned}$$

which requires

$$\frac{1}{N}\boldsymbol{\Delta}_B\boldsymbol{\Delta}'_B = o(1), \;\; \frac{1}{N}\boldsymbol{\Delta}_B\mathbf{B} = o(1), \quad \text{and } \boldsymbol{\Delta}_B\boldsymbol{\epsilon}' = o_p(1),$$

[36] Details are available upon request.

and

$$\frac{1}{N}\mathbf{Z}_{t-1}'\hat{\mathbf{B}} = \frac{1}{N}\mathbf{Z}_{t-1}'[\mathbf{B} + (\boldsymbol{\Delta}_B + \boldsymbol{\epsilon})'\mathbf{P}]$$
$$= \underbrace{\frac{1}{N}\mathbf{Z}_{t-1}'[\mathbf{B} + \boldsymbol{\epsilon}'\mathbf{P}]}_{\text{(locally-constant terms)}} + \underbrace{\frac{1}{N}\mathbf{Z}_{t-1}'\boldsymbol{\Delta}_B'P,}_{\text{(time-varying term)}}$$

which requires

$$\frac{1}{N}\mathbf{Z}_{t-1}'\boldsymbol{\Delta}_B' = o_p(1).$$

Consider now the terms on the right side of the matrix inverse. Then,

$$\frac{1}{N}\mathbf{R}_t'\hat{\mathbf{B}} = \frac{1}{N}(\mathbf{Z}_{t-1}\boldsymbol{\gamma}_{\text{z},t-1} + \mathbf{X}\boldsymbol{\Gamma}_{\text{f},t-1} + \boldsymbol{\epsilon}_t + (\mathbf{X}_{t-1} - \mathbf{X})\boldsymbol{\Gamma}_{\text{f},t-1})'[\mathbf{B} + (\boldsymbol{\Delta}_B + \boldsymbol{\epsilon})'\mathbf{P}]$$
$$= \underbrace{\frac{1}{N}(\mathbf{Z}_{t-1}\boldsymbol{\gamma}_{\text{z},t-1} + \mathbf{X}\boldsymbol{\Gamma}_{\text{f},t-1} + \boldsymbol{\epsilon}_t)'[\mathbf{B} + \boldsymbol{\epsilon}'\mathbf{P}]}_{\text{(locally-constant terms)}}$$
$$+ \underbrace{\frac{1}{N}(\mathbf{Z}_{t-1}\boldsymbol{\gamma}_{\text{z},t-1} + \mathbf{X}\boldsymbol{\Gamma}_{\text{f},t-1} + \boldsymbol{\epsilon}_t + (\mathbf{X}_{t-1} - \mathbf{X})\boldsymbol{\Gamma}_{\text{f},t-1})'\boldsymbol{\Delta}_B'\mathbf{P} + \frac{1}{N}((\mathbf{X}_{t-1} - \mathbf{X})\boldsymbol{\Gamma}_{\text{f},t-1})'[\mathbf{B} + \boldsymbol{\epsilon}'\mathbf{P}]}_{\text{(time-varying term)}}, \text{ and}$$
$$\frac{1}{N}\mathbf{Z}_{t-1}'\hat{\mathbf{B}} = \mathbf{Z}_{t-1}'[\mathbf{B} + (\boldsymbol{\Delta}_B + \boldsymbol{\epsilon})'\mathbf{P}] = \underbrace{\frac{1}{N}\mathbf{Z}_{t-1}'[\mathbf{B} + \boldsymbol{\epsilon}'\mathbf{P}]}_{\text{(locally-constant terms)}} + \underbrace{\frac{1}{N}\mathbf{Z}_{t-1}'\boldsymbol{\Delta}_B'}_{\text{(time-varying term)}},$$

which requires strengthening the above conditions to rate $o_p(N^{-1/2})$, namely

$$\frac{1}{N}\boldsymbol{\Delta}_B\boldsymbol{\Delta}_B' = o(N^{-1/2}), \ \frac{1}{N}\boldsymbol{\Delta}_B\mathbf{B} = o(N^{-1/2}), \quad \text{and } \boldsymbol{\Delta}_B\boldsymbol{\epsilon}' = o_p(N^{-1/2}), \text{ and } \frac{1}{N}\mathbf{Z}_{t-1}'\boldsymbol{\Delta}_B' = o_p(N^{-1/2}). \tag{OA.105}$$

It turns out that Assumption 1 delivers precisely the required sufficient condition.

OA.6.1 Simulation Exercise: Time-Variation of Risk Exposures

To illustrate the time variation allowed by Assumption 1, we run the following simulation design. We first partition the monthly sample into *non-overlapping* windows of length T_{win} months $\mathcal{W}_\tau = \{t_\tau, \ldots, t_\tau + T_{\text{win}} - 1\}$ for $t_1 = 1, t_{\tau+1} = t_\tau + T_{\text{win}}$, $t_{T/T_{\text{win}}} = T - T_{\text{win}} + 1$, and $1 \le \tau \le T/T_{\text{win}}$ (assume it integer). Therefore, we assume that the total number of non-overlapping intervals is T/T_{win}, the latter being assumed integer without loss of generality. Within each window $\mathcal{T}_\tau$, we form a balanced cross-section of assets and estimate the Fama–French five-factor (FF5) model by

OLS using as test assets the dataset of single stocks used in our empirical application:

$$R^e_{i,t} = \hat{\alpha}_{i,\tau} + \hat{\beta}^{\text{Mkt}}_{i,\tau}\,(\text{Mkt–RF})_t + \hat{\beta}^{\text{SMB}}_{i,\tau}\,\text{SMB}_t + \hat{\beta}^{\text{HML}}_{i,\tau}\,\text{HML}_t + \hat{\beta}^{\text{RMW}}_{i,\tau}\,\text{RMW}_t + \hat{\beta}^{\text{CMA}}_{i,\tau}\,\text{CMA}_t + \hat{\varepsilon}_{i,t}, \quad t \in \mathcal{T}_\tau,$$

where $R^e_{i,t}$ denotes stock i's excess returns, and a hat $(\hat{})$ denotes estimated quantities.

To illustrate the time-variation allowed by our setting, we focus on the estimated beta on the market factor, i.e., we adopt the calibration $B_{0,\tau,i} = \hat{\beta}^{\text{Mkt}}_{i,\tau}$, which will change from rolling window to rolling window. Next, for a given cross-section size $N \in \{50, 100, 250, 500, 1000\}$, we generate the overall risk exposure as

$$B_{i,t}(\tau, N) \;=\; B_{0,\tau,i} \;+\; B_{1,i}(N)\,z_t, \;\text{ for every } s \in \mathcal{T}_\tau \text{ and } \tau = 1, \cdots, [T/T_{\text{win}}],$$

where $B_{1,i}(N) = u_{1,i}/N^{\,1/4+\varepsilon}$ is invariant in τ for simplicity, where $u_{1,i} \sim NID(0,1)$ is a fixed draw held constant across all N (to make cross-N comparisons meaningful), and $\varepsilon = 0.05$. The common instrument $\{z_t\}$ follows an AR(1), $z_t = \phi_z z_{t-1} + u_t$, $u_t \sim NID(0,1)$, with $\phi_z \in \{0.7, 0.9, 1.0\}$.

Figures OA.1–OA.3 plot, for $T_{\text{win}} \in \{12, 24\}$ and each ϕ_z, the monthly path of $B_{i,t}(\tau, N)$ for a given series i; colors denote different values of N, while the black dashed horizontal segments show the step function $B_{0,\tau,i}$, constant within each T_{win}-month block.

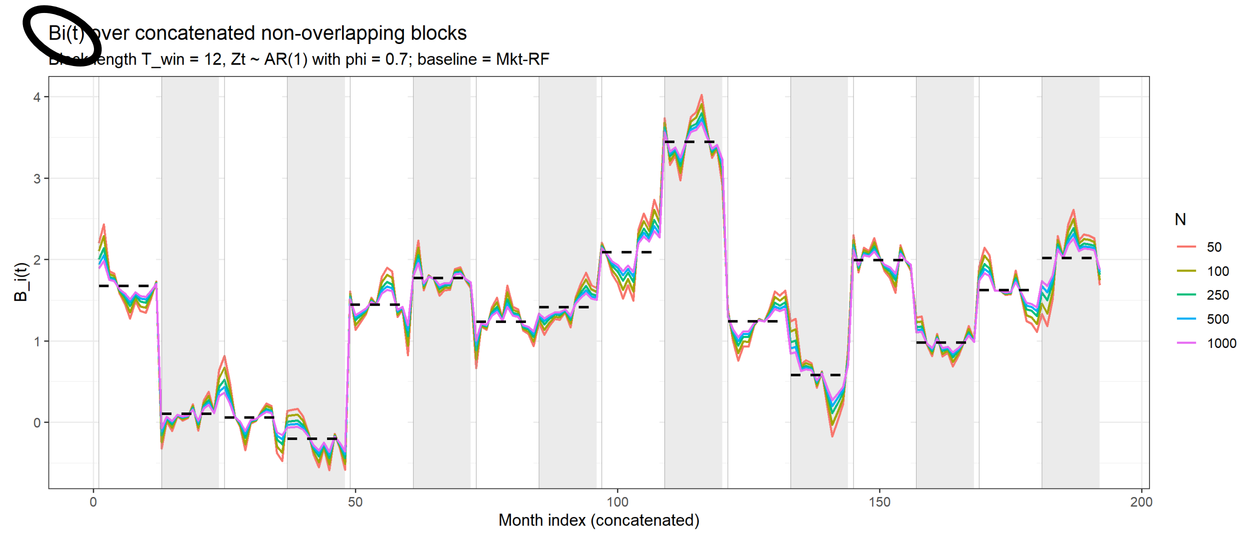
Finally, concatenating all windows for each N, we report, for each series, the sample autocorrelation function ACF(ℓ) of the $B_{i,t}(\tau, N)$, at lags $\ell = 1, \ldots, 12$. Table OA.III presents these ACFs by N (rows) and lag (columns), with separate panels for $\phi_z \in \{0.7, 0.9, 1.0\}$ and side-by-side blocks for $T_{\text{win}} = 12$ and $T_{\text{win}} = 24$.

Overall, the figures and the ACFs show that:

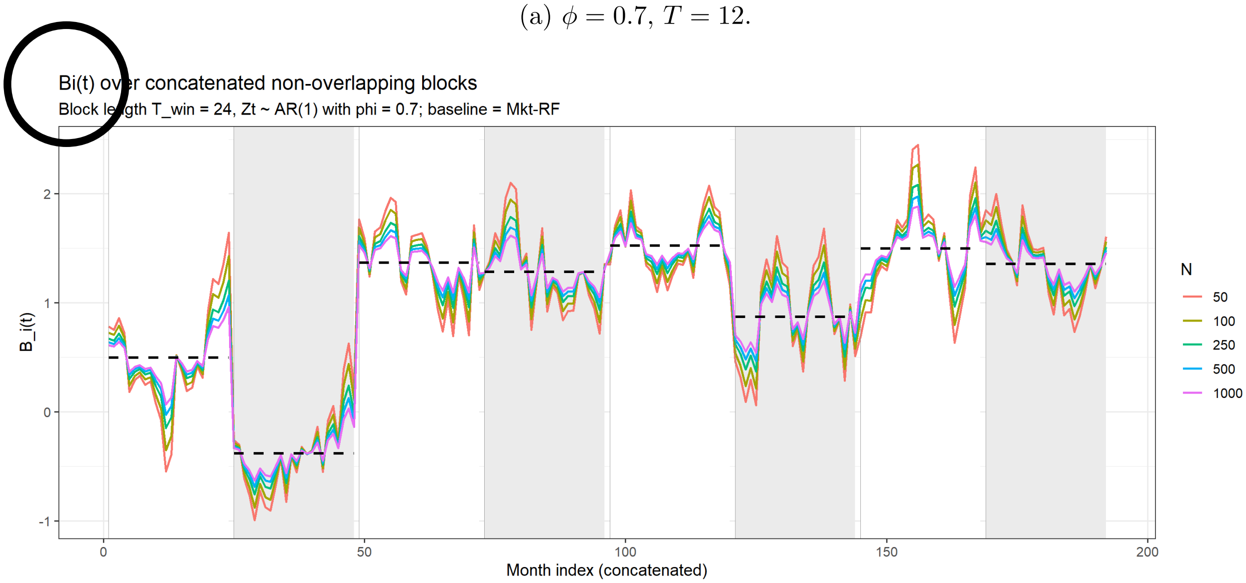
1. The time variation in the $B_{i,t}(\tau, N)$ across the full sample is sizeable and persists for all N, in fact it appears largely independent of N, driven by the empirically-calibrated changes in the $B_{0,\tau,i}$ from one window to the other, across different values of T_{win} and ϕ_z.

2. The dispersion in the $B_{i,t}(\tau, N)$ across each local window, driven by the autocorrelation in the instrument z_t, diminishes as N as increases, satisfying Assumption 1 as the $B_{i,t}(\tau, N)$ converge towards the $B_{0,\tau,i}$. Notice that the autocorrelation (both in population and in sample) within any given window of size T_{win} will not change across different values of N, except at the limit (i.e., when N is de facto infinity), when clearly the betas become locally constant.

Figure OA.1Concatenated monthly series $B_{i,t}(\tau, N)$ for asset i (colored by cross-section size $N \in \{50, 100, 250, 500, 1000\}$). Dashed black horizontal segments plot the blockwise baseline $B_{0,\tau,i} = \hat{\beta}^{\text{Mkt}}_{i,\tau}$ obtained from non-overlapping T_{win}-month FF5 regressions of $(Mkt - RF, SMB, HML, RMW, CMA)$ on excess returns. Within each block, $B_{i,t}(\tau, N) = B_{0,\tau,i} + B_{1,i}(N)\, z_t$ with $B_{1,i}(N) = u_{1,i}/N^{1/4+\varepsilon}$ as $t \in \mathcal{T}_\tau$, for $\varepsilon = 0.05$, $u_{1,i} \sim NID(0,1)$ fixed across N, and z_t follows an AR(1) with $\phi_z = 0.70$. Panels show $T_{\text{win}} = 12$ (top) and $T_{\text{win}} = 24$ (bottom).

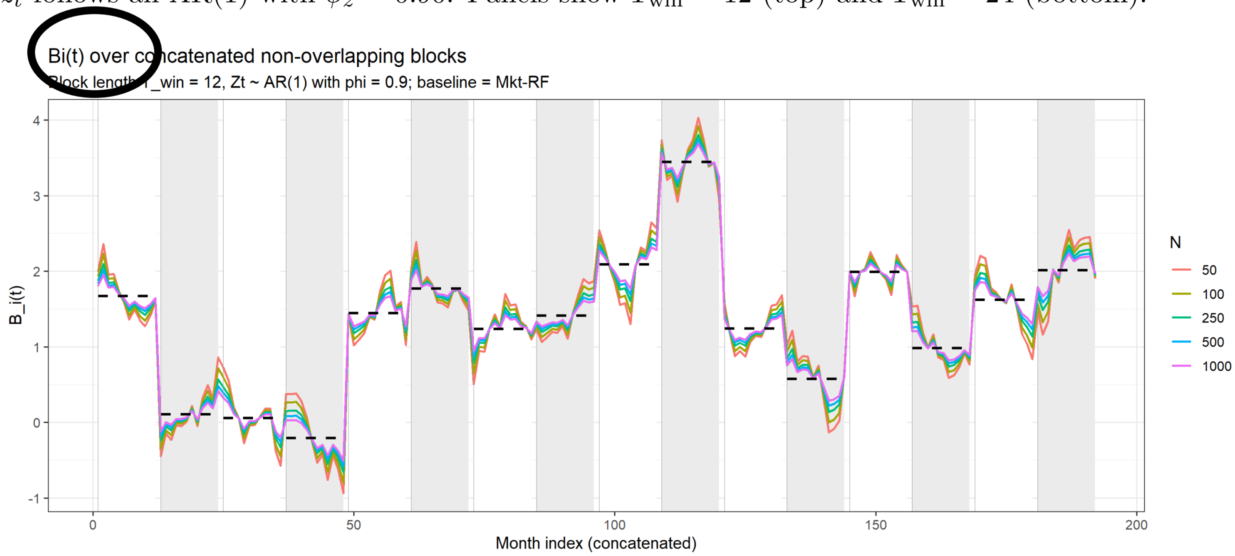


(a) $\phi = 0.7$, $T = 12$.

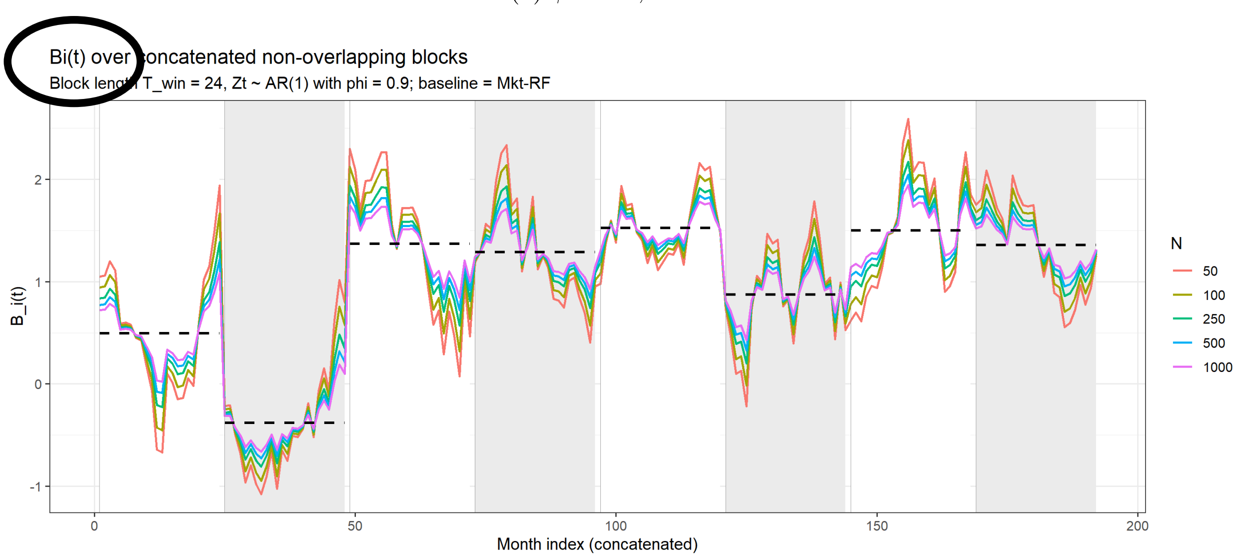


(b) $\phi = 0.7$, $T = 24$.

Figure OA.2Concatenated monthly series $B_{i,t}(\tau, N)$ for asset i (colored by cross-section size $N \in \{50, 100, 250, 500, 1000\}$). Dashed black horizontal segments plot the blockwise baseline $B_{0,\tau,i} = \hat{\beta}^{\text{Mkt}}_{i,\tau}$ obtained from non-overlapping T_{win}-month FF5 regressions of $(Mkt - RF, SMB, HML, RMW, CMA)$ on excess returns. Within each block, $B_{i,t}(\tau, N) = B_{0,\tau,i} + B_{1,i}(N)\, z_t$ with $B_{1,i}(N) = u_{1,i}/N^{1/4+\varepsilon}$ as $t \in \mathcal{T}_\tau$, for $\varepsilon = 0.05$, $u_{1,i} \sim NID(0,1)$ fixed across N, and z_t follows an AR(1) with $\phi_z = 0.90$. Panels show $T_{\text{win}} = 12$ (top) and $T_{\text{win}} = 24$ (bottom).

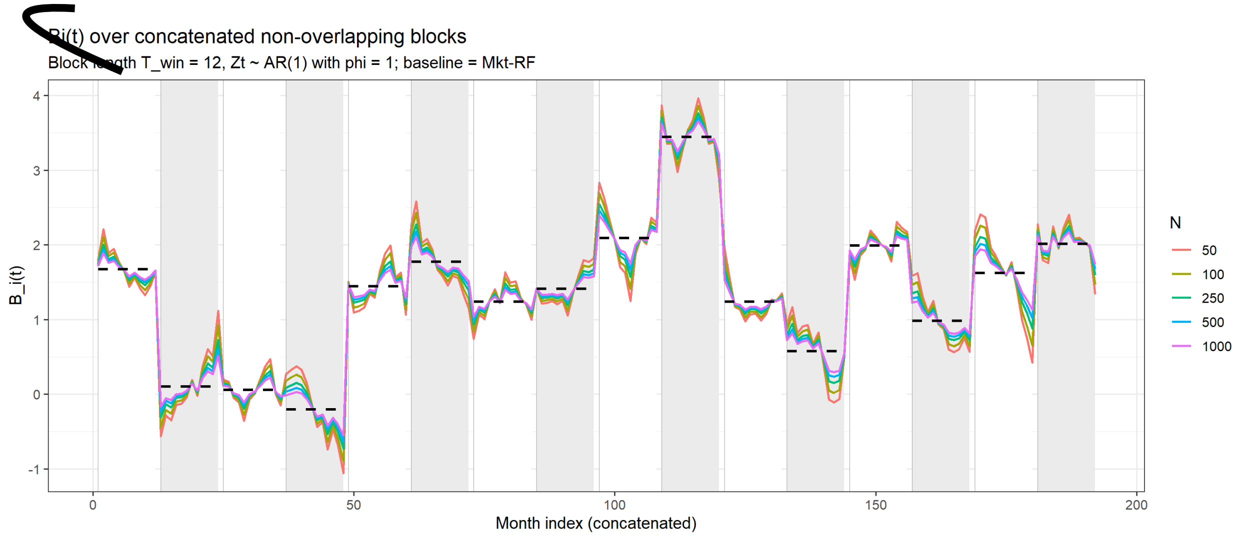


(a) $\phi = 0.9$, $T = 12$.

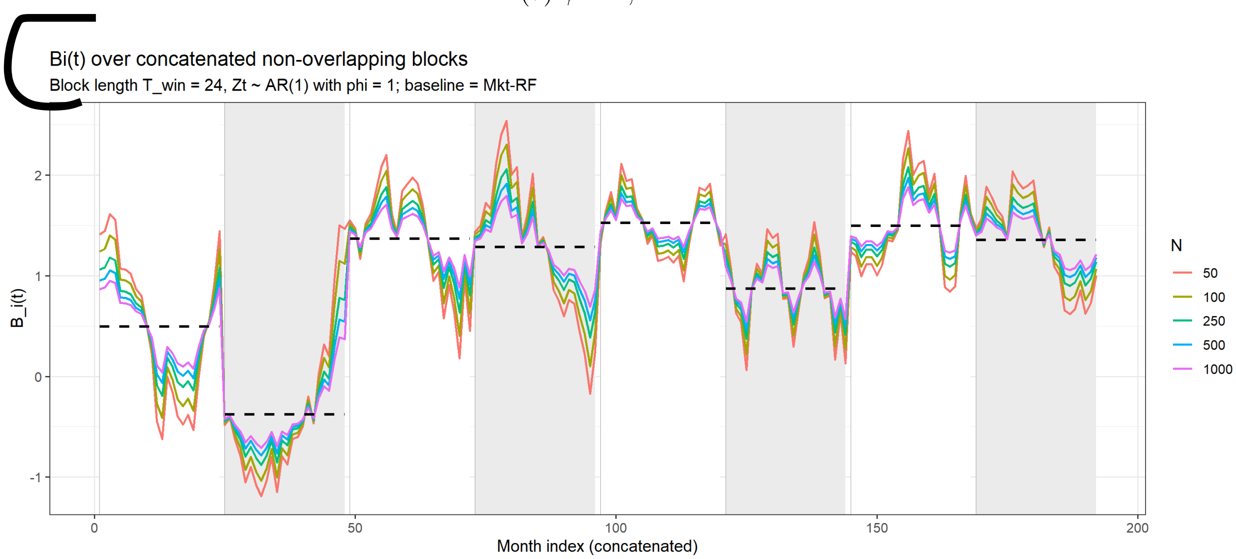


(b) $\phi = 0.9$, $T = 24$.

Figure OA.3Concatenated monthly series $B_{i,t}(\tau, N)$ for asset i (colored by cross-section size $N \in \{50, 100, 250, 500, 1000\}$). Dashed black horizontal segments plot the blockwise baseline $B_{0,\tau,i} = \hat{\beta}^{\text{Mkt}}_{i,\tau}$ obtained from non-overlapping T_{win}-month FF5 regressions of $(Mkt - RF, SMB, HML, RMW, CMA)$ on excess returns. Within each block, $B_{i,t}(\tau, N) = B_{0,\tau,i} + B_{1,i}(N)\, z_t$ with $B_{1,i}(N) = u_{1,i}/N^{1/4+\varepsilon}$ as $t \in \mathcal{T}_\tau$, for $\varepsilon = 0.05$, $u_{1,i} \sim NID(0, 1)$ fixed across N, and z_t follows an AR(1) with $\phi_z = 1.00$. Panels show $T_{\text{win}} = 12$ (top) and $T_{\text{win}} = 24$ (bottom).



(a) $\phi = 1$, $T = 12$.



(b) $\phi = 1$, $T = 24$.

Table OA.III: Autocorrelation function (ACF, lags 1–12) of the simulated series $B_{i,t}(\tau, N)$. Entries are computed on the monthly series formed by concatenating all non-overlapping windows. Within each window of length T_{win}, the baseline $B_{0,\tau,i}$ is set equal to the the OLS estimated beta on Mkt–RF; the simulated series is made as $B_{i,t}(\tau, N) = B_{0,\tau,i} + B_{1,i}(N)z_t$ with $z_t \sim \text{AR}(1;\phi)$ and $B_{1,i}(N) = u_{1,i}/N^{1/4+\varepsilon}$, with $\varepsilon = 0.05$ and the draws $u_{1,i}$ held fixed across N). Panels vary $\phi_z \in \{0.7, 0.9, 1.0\}$; within each panel the left (right) block reports $T_{\text{win}} = 12$ ($T_{\text{win}} = 24$). Values are rounded to two decimals.

	T=12												**T=24**											
N	**Lag 1**	**Lag 2**	**Lag 3**	**Lag 4**	**Lag 5**	**Lag 6**	**Lag 7**	**Lag 8**	**Lag 9**	**Lag 10**	**Lag 11**	**Lag 12**	**Lag 1**	**Lag 2**	**Lag 3**	**Lag 4**	**Lag 5**	**Lag 6**	**Lag 7**	**Lag 8**	**Lag 9**	**Lag 10**	**Lag 11**	**Lag 12**
	Panel A: $\varphi = 0.70$																							
50	0.90	0.82	0.73	0.65	0.57	0.50	0.42	0.35	0.28	0.22	0.16	0.11	0.85	0.74	0.64	0.54	0.48	0.44	0.42	0.38	0.36	0.33	0.29	0.26
100	0.91	0.83	0.75	0.67	0.59	0.52	0.44	0.37	0.30	0.23	0.17	0.12	0.88	0.78	0.70	0.61	0.55	0.51	0.48	0.45	0.42	0.39	0.34	0.31
250	0.92	0.84	0.76	0.69	0.61	0.53	0.46	0.38	0.31	0.24	0.18	0.12	0.91	0.83	0.76	0.68	0.62	0.59	0.55	0.52	0.49	0.45	0.40	0.36
500	0.92	0.84	0.77	0.69	0.62	0.54	0.47	0.39	0.32	0.25	0.18	0.12	0.92	0.85	0.79	0.72	0.67	0.63	0.59	0.55	0.52	0.48	0.43	0.39
1000	0.92	0.85	0.77	0.70	0.62	0.55	0.47	0.40	0.32	0.25	0.19	0.12	0.93	0.87	0.81	0.75	0.70	0.66	0.62	0.58	0.54	0.50	0.45	0.41
	Panel B: $\varphi = 0.90$																							
N	**Lag 1**	**Lag 2**	**Lag 3**	**Lag 4**	**Lag 5**	**Lag 6**	**Lag 7**	**Lag 8**	**Lag 9**	**Lag 10**	**Lag 11**	**Lag 12**	**Lag 1**	**Lag 2**	**Lag 3**	**Lag 4**	**Lag 5**	**Lag 6**	**Lag 7**	**Lag 8**	**Lag 9**	**Lag 10**	**Lag 11**	**Lag 12**
50	0.90	0.81	0.73	0.64	0.56	0.48	0.40	0.33	0.26	0.21	0.15	0.11	0.89	0.79	0.68	0.58	0.50	0.44	0.39	0.34	0.30	0.26	0.21	0.18
100	0.91	0.82	0.74	0.66	0.58	0.50	0.43	0.35	0.28	0.22	0.17	0.11	0.90	0.82	0.72	0.63	0.56	0.51	0.46	0.41	0.37	0.33	0.28	0.24
250	0.92	0.84	0.76	0.68	0.60	0.52	0.45	0.37	0.30	0.24	0.18	0.12	0.92	0.85	0.77	0.69	0.63	0.58	0.54	0.49	0.45	0.41	0.36	0.32
500	0.92	0.84	0.76	0.69	0.61	0.53	0.46	0.38	0.31	0.25	0.18	0.12	0.93	0.87	0.80	0.73	0.67	0.63	0.58	0.54	0.50	0.45	0.40	0.36
1000	0.92	0.85	0.77	0.69	0.62	0.54	0.47	0.39	0.32	0.25	0.18	0.12	0.94	0.88	0.82	0.76	0.70	0.66	0.62	0.57	0.53	0.49	0.43	0.39
	Panel C: $\varphi = 1.00$																							
N	**Lag 1**	**Lag 2**	**Lag 3**	**Lag 4**	**Lag 5**	**Lag 6**	**Lag 7**	**Lag 8**	**Lag 9**	**Lag 10**	**Lag 11**	**Lag 12**	**Lag 1**	**Lag 2**	**Lag 3**	**Lag 4**	**Lag 5**	**Lag 6**	**Lag 7**	**Lag 8**	**Lag 9**	**Lag 10**	**Lag 11**	**Lag 12**
50	0.90	0.81	0.73	0.64	0.56	0.48	0.41	0.33	0.27	0.21	0.15	0.11	0.88	0.78	0.67	0.57	0.49	0.43	0.38	0.33	0.29	0.24	0.20	0.17
100	0.91	0.83	0.74	0.66	0.58	0.51	0.43	0.35	0.29	0.23	0.17	0.11	0.90	0.81	0.71	0.63	0.56	0.51	0.46	0.41	0.37	0.32	0.28	0.24
250	0.92	0.84	0.76	0.68	0.60	0.53	0.45	0.38	0.31	0.24	0.18	0.12	0.92	0.85	0.77	0.70	0.64	0.59	0.54	0.50	0.45	0.41	0.36	0.32
500	0.92	0.84	0.77	0.69	0.61	0.54	0.46	0.39	0.32	0.25	0.18	0.12	0.93	0.87	0.80	0.74	0.68	0.63	0.59	0.55	0.51	0.46	0.41	0.37
1000	0.92	0.85	0.77	0.70	0.62	0.54	0.47	0.39	0.32	0.25	0.19	0.12	0.94	0.88	0.82	0.77	0.71	0.67	0.63	0.58	0.54	0.50	0.45	0.40

OA.7 Two-Pass Methodology and Anomalies: the Conventional Approach - Asymptotics

This section provides the formal derivations of the main results established in Section 3 of the paper.[37] Particularly, we derive the limiting behavior of the *conventional* anomaly premia estimator (see Fama and French (2008)), which consists of running, for every period, a CSR OLS regression of asset returns $\mathbf{R}_t$ on the anomaly variables $\mathbf{Z}_{t-1}$, yielding the time-varying estimator in (21). The T premia estimates are then averaged across time, resulting in the average premia estimator (22). This approach coincides exactly with the second step of the two-pass Fama and MacBeth (1973) procedure, where, however, one *excludes* the betas from the model, to avoid EIV-related issues.

In the following, we analyze the limiting behavior of the conventional estimators in (21) and (22), together with the corresponding conventional t-ratio in (23), under three different sampling schemes: (i) when $T \to \infty$ and N is fixed, (ii) when $N \to \infty$ and T is fixed, and (iii) when both N and T are allowed to diverge. We first present the results in the univariate-regression setting, that is, when one considers only one anomaly at a time ($K_{\mathrm{z}} = 1$) in the regression model. Extensions to the multivariate case are presented in the final remarks.

Consider the following model

$$\mathbf{R}_t = \gamma_{0,t-1}\mathbf{1}_N + \mathbf{Z}_{t-1}\gamma_{\mathrm{z},t-1} + \mathbf{B}\boldsymbol{\delta}_{\mathrm{f,t-1}} + \boldsymbol{\epsilon}_t, \tag{OA.106}$$

with the objective of estimating the anomaly premium $\gamma_{\mathrm{z},t-1}$. Following Fama and French (2008), at each point in time, we run a cross-sectional OLS regression of (OA.106), using the anomaly $\mathbf{Z}_{t-1}$, but *excluding* the betas $\mathbf{B}$ from the model. This gives the time-varying OLS estimator

$$\tilde{\gamma}_{\mathrm{z},t-1} = \frac{\mathbf{Z}'_{t-1}\mathbb{M}_{1_N}\mathbf{R}_t}{\mathbf{Z}'_{t-1}\mathbb{M}_{1_N}\mathbf{Z}_{t-1}}, \tag{OA.107}$$

which satisfies

$$\tilde{\gamma}_{\mathrm{z},t-1} = \gamma_{\mathrm{z},t-1} + \frac{\mathbf{Z}'_{t-1}\mathbb{M}_{1_N}\left(\mathbf{B}\boldsymbol{\delta}_{\mathrm{f,t-1}} + \boldsymbol{\epsilon}_t\right)}{\mathbf{Z}'_{t-1}\mathbb{M}_{1_N}\mathbf{Z}_{t-1}}.$$

Then, averaging the time-varying estimator across time yields the *average* OLS estimator

$$\bar{\tilde{\gamma}}_{\mathrm{z}} = \frac{1}{T-1}\sum_{t=2}^{T}\tilde{\gamma}_{\mathrm{z},t-1} = \frac{1}{T}\sum_{t=1}^{T}\frac{\mathbf{Z}'_{t-1}\mathbb{M}_{1_N}\mathbf{R}_t}{\mathbf{Z}'_{t-1}\mathbb{M}_{1_N}\mathbf{Z}_{t-1}} \tag{OA.108}$$

comma

[37] As for the previous sections, we develop the proofs for the more general case of absence of the risk-free asset, i.e., when $\gamma_{0,t-1}$ needs to be estimated. Section 3 of the paper refers to the special case when the test assets are excess returns, and $\gamma_{0,t-1} = 0$.

which satisfies

$$\bar{\tilde{\gamma}}_{\mathrm{z}} = \bar{\gamma}_{\mathrm{z}} + \frac{1}{T-1}\sum_{t=2}^{T}\left(\frac{\mathbf{Z}'_{t-1}\mathbb{M}_{1_N}(\mathbf{B}\boldsymbol{\delta}_{\mathrm{f},t-1}+\boldsymbol{\epsilon}_t)}{\mathbf{Z}'_{t-1}\mathbb{M}_{1_N}\mathbf{Z}_{t-1}}\right), \tag{OA.109}$$

with $\bar{\gamma}_{\mathrm{z}} = \frac{1}{T-1}\sum_{t=2}^{T}\gamma_{\mathrm{z},t-1}$. and with estimated variance (squared standard error)

$$\widehat{\mathrm{Var}}[\bar{\tilde{\gamma}}_z] = \frac{1}{(T-1)^2}\sum_{t=2}^{T}(\tilde{\gamma}_{\mathrm{z},t-1} - \bar{\tilde{\gamma}}_{\mathrm{z}})^2 = \frac{\tilde{\Sigma}_{\gamma_{\mathrm{z}}}}{(T-1)}, \tag{OA.110}$$

with $\tilde{\Sigma}_{\gamma_{\mathrm{z}}}$ denoting the sample average of the time-varying estimates.

In the following, we want to derive the limiting behaviour of the time-varying estimator (OA.107) and locally-averaged estimator (OA.108) under the three sampling schemes mentioned above. Before introducing our results, we state below the set of assumptions required (not necessarily at the same time) to derive our results.

Assumption OA.15 (Finite-N Orthogonality)**.** For a given fixed N, the $\boldsymbol{\beta}_i$ and $Z_{i,t-1}$ are cross-sectionally orthogonal in the sample:

$$\mathbf{B}'\mathbb{M}_{1_N}\mathbf{Z}_{t-1} = \mathbf{0}_{K_{\mathrm{f}}\times K_{\mathrm{z}}}.$$

Assumption OA.16 (Large-N Orthogonality)**.** As $N\to\infty$, the $\boldsymbol{\beta}_i$ and $Z_{i,t-1}$ are asymptotically cross-sectionally orthogonal:

$$\frac{1}{N}\mathbf{B}'\mathbb{M}_{1_N}\mathbf{Z}_{t-1} \to_p \mathbf{0}_{K_{\mathrm{f}}\times K_{\mathrm{z}}}.$$

Assumption OA.17 (Uncorrelatedness of risk factors and asset returns)**.** For every $t = 1, ..., T$, the risk factors $\mathbf{f}_t$ and the asset returns $\mathbf{R}_t$ are uncorrelated:

$$\mathbf{B} = \mathbf{0}_{N\times K_{\mathrm{f}}}.$$

Assumption OA.18 (Constant anomaly premia)**.** The anomaly premia $\boldsymbol{\gamma}_{\mathrm{z},t-1}$ are constant over time:

$$\boldsymbol{\gamma}_{\mathrm{z},t-1} = \boldsymbol{\gamma}_{\mathrm{z}}, \qquad \textit{for every } t = 2, .., T$$

OA.7.1 The large-T–fixed-N case

Under the large-T–fixed-N sampling scheme, clearly the time-varying OLS estimator in (OA.107) remains unchanged, and no meaningful asymptotic property can be established. The results regarding the locally-averaged OLS estimator in (OA.108) are summarized in the following theorem.[38]

Theorem OA.8. *(Locally-Averaged OLS Estimator — large T, fixed N) Assume the* finite-N *orthogonality condition* (OA.15)*, and the following regularity conditions:* $\mathrm{E}[\boldsymbol{\epsilon}_t \mid \mathbf{Z}_{t-1}] = \mathbf{0}_N$, $\mathrm{E}[\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t' \mid \mathbf{Z}_{t-1}] = \boldsymbol{\Sigma}$*, and*

$$\frac{1}{T-1}\sum_{t=2}^{T}\frac{\mathbf{Z}_{t-1}'\mathbb{M}_{1_N}\boldsymbol{\Sigma}\,\mathbb{M}_{1_N}\mathbf{Z}_{t-1}}{(\mathbf{Z}_{t-1}'\mathbb{M}_{1_N}\mathbf{Z}_{t-1})^2} \ \to_p \ V_N, \qquad 0 < V_N < \infty.$$

Then, as $T \to \infty$ with N fixed:

(i) Let $\bar{\gamma}_{\mathrm{z}}^0 = \lim_{T\to\infty}\bar{\gamma}_{\mathrm{z}}$. Then

$$\bar{\tilde{\gamma}}_{\mathrm{z}} \ \to_p \ \bar{\gamma}_{\mathrm{z}}^0.$$

(ii) When, in addition, $\frac{1}{\sqrt{T-1}}\sum_{t=2}^{T}\mathbf{c}_{t-1}'\boldsymbol{\epsilon}_t \ \to_d \ \mathcal{N}(0, V_N)$, with $\mathbf{c}_{t-1} = \frac{\mathbb{M}_{1_N}\mathbf{Z}_{t-1}}{\mathbf{Z}_{t-1}'\mathbb{M}_{1_N}\mathbf{Z}_{t-1}}$, we have

$$\sqrt{T}\big(\bar{\tilde{\gamma}}_{\mathrm{z}} - \bar{\gamma}_{\mathrm{z}}\big) \ \to_d \ \mathcal{N}(0, V_N).$$

(iii) When, in addition, $\frac{1}{T-1}\sum_{t=2}^{T}\left(\gamma_{\mathrm{z},t-1} - \bar{\gamma}_{\mathrm{z}}\right)^2 \ \to \ \sigma_{\gamma_{\mathrm{z}}}^2$, then

$$\tilde{\Sigma}_{\gamma_{\mathrm{z}}} = \frac{1}{T-1}\sum_{t=2}^{T}\left(\tilde{\gamma}_{\mathrm{z},t-1} - \bar{\tilde{\gamma}}_{\mathrm{z}}\right)^2 \ \to_p \ \sigma_{\gamma_{\mathrm{z}}}^2 + V_N.$$

Proof. Parts (i) and (ii) follow immediately from (OA.109) and the assumptions made above. To prove part (iii), notice that, using (OA.108) and (OA.109), we have that

$$\tilde{\gamma}_{\mathrm{z},t-1} - \bar{\tilde{\gamma}}_{\mathrm{z}} = \gamma_{\mathrm{z},t-1} - \bar{\gamma}_{\mathrm{z}} + \frac{\mathbf{Z}_{t-1}'\mathbb{M}_{1_N}\boldsymbol{\epsilon}_t}{\mathbf{Z}_{t-1}\mathbb{M}_{1_N}\mathbf{Z}_{t-1}} - \frac{1}{T-1}\sum_{s=2}^{T}\frac{\mathbf{Z}_{s-1}'\mathbb{M}_{1_N}\boldsymbol{\epsilon}_s}{\mathbf{Z}_{s-1}\mathbb{M}_{1_N}\mathbf{Z}_{s-1}},$$

implying that

$$\left(\tilde{\gamma}_{\mathrm{z},t-1} - \bar{\tilde{\gamma}}_{\mathrm{z}}\right)^2 = \left(\gamma_{\mathrm{z},t-1} - \bar{\gamma}_{\mathrm{z}}\right)^2 + \left(\frac{\mathbf{Z}_{t-1}'\mathbb{M}_{1_N}\boldsymbol{\epsilon}_t}{\mathbf{Z}_{t-1}\mathbb{M}_{1_N}\mathbf{Z}_{t-1}}\right)^2 + o_p(1).$$

[38] We continue to define (OA.108) as locally-averaged for consistency with our definition even though, in this case, we let $T \to \infty$.

Therefore, using the definition in (OA.110) and the assumptions above, we get

$$
\begin{aligned}
(T-1)\widehat{\mathrm{Var}}(\tilde{\gamma}_{\mathrm{z},t-1}) &= \frac{1}{T-1}\sum_{t=2}^{T}(\tilde{\gamma}_{\mathrm{z},t-1}-\bar{\tilde{\gamma}}_{\mathrm{z}})^2 \\
&= \frac{1}{T-1}\sum_{t=2}^{T}\left(\gamma_{\mathrm{z},t-1}-\bar{\gamma}_{\mathrm{z}}\right)^2+\frac{1}{T-1}\sum_{t=2}^{T}\frac{\mathbf{Z}_{t-1}'\mathbb{M}_{1_N}\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t'\mathbb{M}_{1_N}\mathbf{Z}_{t-1}}{\left(\mathbf{Z}_{t-1}\mathbb{M}_{1_N}\mathbf{Z}_{t-1}\right)^2}+o_p(1) \\
&\to_p \sigma^2_{\gamma_{\mathrm{z}}}+V_N
\end{aligned}
$$

■

Remark OA.23. *Unless Assumption OA.18 is satisfied, namely that $\gamma_{\mathrm{z},t-1}$ is constant, the conventional standard error of the average OLS estimator from the Fama and MacBeth (1973) regression contains an upward bias given by $\sigma^2_{\gamma_{\mathrm{z}}}>0$. This implies that the associated t-ratio is smaller than its correct value, leading to potential under-rejections.*

Remark OA.24. *Theorem OA.8 easily extends to the multivariate case $K_{\mathrm{z}}>1$. In this case:*

(i) Let $\bar{\boldsymbol{\gamma}}^0_{\mathrm{z}}=\lim_{T\to\infty}\bar{\boldsymbol{\gamma}}_z$, with $\bar{\boldsymbol{\gamma}}_z=\frac{1}{T-1}\sum_{t=2}^{T}\boldsymbol{\gamma}_{\mathrm{z},t-1}$. Then

$$\bar{\tilde{\boldsymbol{\gamma}}}_{\mathrm{z}} \quad \to_p \quad \bar{\boldsymbol{\gamma}}^0_{\mathrm{z}}.$$

(ii) Assume that

$$\frac{1}{T-1}\sum_{t=2}^{T}\left(\mathbf{Z}_{t-1}'\mathbb{M}_{1_N}\mathbf{Z}_{t-1}\right)^{-1}\mathbf{Z}_{t-1}'\mathbb{M}_{1_N}\boldsymbol{\Sigma}\mathbb{M}_{1_N}\mathbf{Z}_{t-1}\left(\mathbf{Z}_{t-1}'\mathbb{M}_{1_N}\mathbf{Z}_{t-1}\right)^{-1} \quad \to_p \quad \mathbf{V}_N,$$

with $\mathbf{V}_N$ symmetric and positive-definite, and that

$$\frac{1}{\sqrt{T-1}}\sum_{t=2}^{T}\mathbf{C}_{t-1}'\boldsymbol{\epsilon}_t \quad \to_d \quad \mathcal{N}(\mathbf{0}_{K_{\mathrm{z}}},\mathbf{V}_N),$$

where $\mathbf{C}_{t-1}=\mathbb{M}_{1_N}\mathbf{Z}_{t-1}\left(\mathbf{Z}_{t-1}'\mathbb{M}_{1_N}\mathbf{Z}_{t-1}\right)^{-1}$. Then

$$\sqrt{T}\left(\bar{\tilde{\boldsymbol{\gamma}}}_{\mathrm{z}}-\bar{\boldsymbol{\gamma}}_{\mathrm{z}}\right) \quad \to_d \quad \mathcal{N}(\mathbf{0}_{K_{\mathrm{z}}},\mathbf{V}_N).$$

(iii) When, in addition,

$$\frac{1}{T-1}\sum_{t=2}^{T}\left(\boldsymbol{\gamma}_{\mathrm{z},t-1}-\bar{\boldsymbol{\gamma}}_{\mathrm{z}}\right)\left(\boldsymbol{\gamma}_{\mathrm{z},t-1}-\bar{\boldsymbol{\gamma}}_{\mathrm{z}}\right)' \quad \to_p \quad \boldsymbol{\Sigma}_{\gamma_{\mathrm{z}}},$$

with $\boldsymbol{\Sigma}_{\gamma_z}$ *symmetric and positive-definite, then*

$$(T-1)\,\widehat{\mathrm{Var}}\big[\bar{\tilde{\gamma}}_z\big] = \frac{1}{T-1}\sum_{t=2}^{T}\big(\tilde{\gamma}_{z,t-1}-\bar{\tilde{\gamma}}_z\big)\big(\tilde{\gamma}_{z,t-1}-\bar{\tilde{\gamma}}_z\big)' \;\;\to_p\;\; \boldsymbol{\Sigma}_{\gamma_z}+\mathbf{V}_N.$$

OA.7.2 The fixed-T–large-N case

We now establish the asymptotic properties of the conventional estimators in the fixed-T–large-N setting. In this framework, the time-varying OLS estimator (OA.107) depends on N and its limiting behavior can be studied. The following theorem summarizes the main results for both the conventional time-varying OLS estimator and the average OLS estimator.

Theorem OA.9. *(Locally-Averaged and Time-Varying OLS Estimators — fixed T, large N) Assume that the* large-N *orthogonality condition in Assumption* (OA.16) *holds. Assume also that the following regularity conditions are satisfied:* $\mathrm{E}[\boldsymbol{\epsilon}_t|\mathbf{Z}_{t-1}]=\mathbf{0}_N$, $\mathrm{E}[\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t' \mid \mathbf{Z}_{t-1}]=\boldsymbol{\Sigma}$, $N^{-1}\mathbf{Z}_{t-1}'\mathbb{M}_{1_N}\mathbf{Z}_{t-1}\to_p a_{t-1}>0$, *and* $N^{-1}(\mathbf{Z}_{t-1}'\mathbb{M}_{1_N}\boldsymbol{\Sigma}\mathbb{M}_{1_N}\mathbf{Z}_{t-1})\to_p\sigma^2 a_{t-1}$, *with* $\mathbf{1}_N'\boldsymbol{\Sigma}\mathbf{1}_N/N\to\sigma^2$. *Then, as* $N\to\infty$ *and* T *is fixed:*

(i)

$$\tilde{\gamma}_{z,t-1}\;\to_p\;\gamma_{z,t-1},\qquad \bar{\tilde{\gamma}}_z\;\to_p\;\bar{\gamma}_z^0.$$

(ii) *Let* $V_{t-1}=\sigma^2/a_{t-1}$. *When Assumption OA.16 is strengthened with*

$$\frac{1}{\sqrt{N}}\mathbf{B}'\mathbb{M}_{1_N}\mathbf{Z}_{t-1}\;\to_p\;\mathbf{0}_{K_f\times K_z},$$

and assuming that

$$\frac{1}{\sqrt{N}}\mathbf{Z}_{t-1}'\mathbb{M}_{1_N}\boldsymbol{\epsilon}_t\;\to_d\;\mathcal{N}(0,\sigma^2 a_{t-1}),$$

then

$$\sqrt{N}\big(\tilde{\gamma}_{z,t-1}-\gamma_{z,t-1}\big)\;\;\to_d\;\;\mathcal{N}(0,V_{t-1}),$$

$$\sqrt{N}\big(\bar{\tilde{\gamma}}_z-\bar{\gamma}_z\big)\;\;\to_d\;\;\mathcal{N}(0,\bar{V}),$$

with

$$\bar{V}=\frac{1}{(T-1)^2}\sum_{t=2}^{T}V_{t-1}.$$

(iii)

$$\widehat{\mathrm{Var}}[\bar{\tilde{\gamma}}_{\mathrm{z}}] = \frac{1}{(T-1)^2}\sum_{t=2}^{T}\left(\tilde{\gamma}_{\mathrm{z},t-1} - \bar{\tilde{\gamma}}_{\mathrm{z}}\right)^2 \to_p \frac{1}{(T-1)^2}\sum_{t=2}^{T}\left(\gamma_{\mathrm{z},t-1} - \bar{\gamma}_{\mathrm{z}}\right)^2.$$

Proof. Parts (i) and (ii) follow immediately from (OA.108) and (OA.109), together with the assumptions stated above. To prove part (iii), notice that, using (OA.108) and (OA.109), we obtain:

$$\tilde{\gamma}_{\mathrm{z},t-1} - \bar{\tilde{\gamma}}_{\mathrm{z}} = \gamma_{\mathrm{z},t-1} - \bar{\gamma}_{\mathrm{z}} + \frac{\mathbf{Z}'_{t-1}\mathbb{M}_{1_N}(\mathbf{B}\boldsymbol{\delta}_{\mathrm{f},t-1} + \boldsymbol{\epsilon}_t)}{\mathbf{Z}_{t-1}\mathbb{M}_{1_N}\mathbf{Z}_{t-1}} - \frac{1}{T-1}\sum_{s=2}^{T}\frac{\mathbf{Z}'_{s-1}\mathbb{M}_{1_N}(\mathbf{B}\boldsymbol{\delta}_{\mathrm{f},s-1} + \boldsymbol{\epsilon}_s)}{\mathbf{Z}_{s-1}\mathbb{M}_{1_N}\mathbf{Z}_{s-1}}$$

Then, under the assumptions stated above, as $N \to \infty$

$$\frac{1}{(T-1)^2}\sum_{t=2}^{T}(\tilde{\gamma}_{\mathrm{z},t-1} - \bar{\tilde{\gamma}}_{\mathrm{z}})^2 = \frac{1}{(T-1)^2}\sum_{t=2}^{T}(\gamma_{\mathrm{z},t-1} - \bar{\gamma}_{\mathrm{z}})^2 + o_p(1).$$

■

Remark OA.25. Theorem OA.9 (iii) shows that the sample variance of the time-varying OLS estimates converges to a positive constant, different from V_{t-1} or $\bar{V}$. However, one can still obtain a consistent estimation of the asymptotic variance of both $\tilde{\gamma}_{\mathrm{z},t-1}$ and $\bar{\tilde{\gamma}}_{\mathrm{z},t-1}$, by imposing further orthogonality conditions, as we show below. Assume that Assumption OA.17 holds, so that risk factors and returns are orthogonal to each other. Let

$$\tilde{\sigma}^2_{zt} = \frac{\tilde{\boldsymbol{\epsilon}}'_t\tilde{\boldsymbol{\epsilon}}_t}{N - K_{\mathrm{z}} - 1} \tag{OA.111}$$

with $\tilde{\boldsymbol{\epsilon}}_t = \boldsymbol{R}_t - \mathbf{1}_N\bar{R}_t - (\boldsymbol{Z}_{t-1} - \mathbf{1}_N\bar{Z}_{t-1})\tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1}$. Then, under the above assumptions, $\tilde{\sigma}^2_{zt} - \sigma^2 \to_p 0$, implying that

$$\tilde{V}_{t-1} \equiv \frac{\tilde{\sigma}^2_{zt}}{\frac{1}{N}\left(\mathbf{Z}'_{t-1}\mathbb{M}_{1_N}\mathbf{Z}_{t-1}\right)} \to_p V_{t-1} \quad \text{and} \quad \frac{1}{(T-1)^2}\sum_{t=2}^{T}\tilde{V}_{t-1} \to_p \bar{V}.$$

Remark OA.26. Theorem OA.9 extends to the multivariate case ($K_{\mathrm{z}} > 1$) as follows.

(i) Let $\bar{\boldsymbol{\gamma}}_z \equiv \frac{1}{T-1}\sum_{t=2}^{T}\boldsymbol{\gamma}_{\mathrm{z},t-1}$. Then, under the same assumptions of Theorem OA.9,

$$\tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1} \to_p \boldsymbol{\gamma}_{\mathrm{z},t-1} \quad \text{and} \quad \bar{\tilde{\boldsymbol{\gamma}}}_{\mathrm{z}} \to_p \bar{\boldsymbol{\gamma}}^0_{\mathrm{z}}.$$

(ii) Let $\boldsymbol{V}_{t-1} = \sigma^2\boldsymbol{A}^{-1}_{t-1}$, with $\frac{1}{N}\mathbf{Z}'_{t-1}\mathbb{M}_{1_N}\mathbf{Z}_{t-1} \to_p \mathbf{A}_{t-1}$ as $N \to \infty$. Then, under the same assumptions of Theorem (OA.9),

$$\sqrt{N}\left(\tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1} - \boldsymbol{\gamma}_{\mathrm{z},t-1}\right) \to_d \mathcal{N}\left(\mathbf{0}_{K_{\mathrm{z}}}, \boldsymbol{V}_{t-1}\right) \quad \text{and}$$

comma

$$\sqrt{N}\left(\bar{\tilde{\boldsymbol{\gamma}}}_{\mathrm{z}} - \bar{\boldsymbol{\gamma}}_{\mathrm{z}}\right) \to_d \mathcal{N}\left(\mathbf{0}_{K_{\mathrm{z}}}, \bar{\boldsymbol{V}}\right) \quad \text{with} \quad \bar{\boldsymbol{V}} = \frac{1}{(T-1)^2}\sum_{t=2}^{T}\boldsymbol{V}_{t-1}.$$

comma

(iii)

$$(T-1)\widehat{\mathrm{Var}}\left[\bar{\tilde{\gamma}}_{\mathrm{z}}\right] \to_p \frac{1}{T-1}\sum_{t=2}^{T}\left(\tilde{\gamma}_{\mathrm{z},t-1}-\bar{\tilde{\gamma}}_{\mathrm{z}}\right)\left(\tilde{\gamma}_{\mathrm{z},t-1}-\bar{\tilde{\gamma}}_{\mathrm{z}}\right)'$$

full stop

OA.7.3 The large-T–large-N case

In this section we generalize all the above results to the case of both $N, T \to \infty$. We show below that the limiting properties of the time-varying OLS estimator $\tilde{\gamma}_{\mathrm{z},t-1}$ remain the same as the ones obtained in the fixed-T–large-N case described in Section OA.7.2. Similar results hold also for the locally-averaged estimator, even though it now benefits from the faster rate of convergence, given that both N and T are now allowed to diverge jointly. The main results are summarized in the following theorem.

Theorem OA.10. *(Locally-Averaged and Time-Varying OLS Estimators - large N–large T) Assume the all the conditions stated in Theorems OA.9 and OA.8 are satisfied. In addition, assume that $(T-1)^{-1}\sum_{t=2}^{T}\frac{\sigma^2}{a_{t-1}} = (T-1)^{-1}\sum_{t=2}^{T}V_{t-1} \to_p \bar{\mathcal{V}}$. Then, as $N, T \to \infty$,*

(i)

$$\tilde{\gamma}_{\mathrm{z},t-1} \to_p \gamma_{\mathrm{z},t-1} \quad \text{comma} \quad \textit{and} \qquad \bar{\tilde{\gamma}}_{\mathrm{z}} \to_p \bar{\gamma}^0_{\mathrm{z}}.$$

(ii)

$$\sqrt{N}\left(\tilde{\gamma}_{\mathrm{z},t-1}-\gamma_{\mathrm{z},t-1}\right) \to_d \mathcal{N}\left(0, V_{t-1}\right) \qquad \textit{and} \ \text{comma}$$

$$\sqrt{NT}\left(\bar{\tilde{\gamma}}_{\mathrm{z}}-\bar{\gamma}_{\mathrm{z}}\right) \to_d \mathcal{N}\left(0, \bar{\mathcal{V}}\right).$$

(iii)

$$T\widehat{\mathrm{Var}}\left[\bar{\tilde{\gamma}}_{\mathrm{z}}\right] = \frac{1}{T-1}\sum_{t=2}^{T}(\tilde{\gamma}_{\mathrm{z},t-1}-\bar{\tilde{\gamma}}_{\mathrm{z}})^2 \to_p \boldsymbol{\Sigma}^2_{\gamma_{\mathrm{z}}}.$$

Proof. The proof follows immediately by combining all the results obtained in the above theorems.

∎

Remark OA.27. Theorem OA.10(iii) shows that, even when both $N, T \to \infty$, the sample variance of the time-varying OLS estimates converges to a positive quantity that differs from $\bar{\mathcal{V}}$. Therefore, it would not be appropriate to use it for inferential conclusions on $\tilde{\gamma}_{\mathrm{z},t-1}$ or $\bar{\tilde{\gamma}}_z$. However, it is still possible to obtain a consistent estimator of the asymptotic variance of the estimators. Let $\mathrm{tr}(\cdot)$ denote the trace operator and define, for $\tilde{\boldsymbol{\epsilon}} = (\tilde{\boldsymbol{\epsilon}}_2, \cdots, \tilde{\boldsymbol{\epsilon}}_T)'$,

$$\tilde{\sigma}^2 = \mathrm{tr}\left(\frac{\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}'}{(T-1)(N-K_\mathrm{z}-1)}\right).$$

Then, under the assumptions stated above, $\tilde{\sigma}^2 \to_p \sigma^2$, implying that

$$\tilde{\mathcal{V}}_{t-1} = \frac{\tilde{\sigma}^2}{\frac{1}{N}\left(\mathbf{Z}'_{t-1}\mathbb{M}_{1_N}\mathbf{Z}_{t-1}\right)} \;\to_p\; V_{t-1}, \qquad \text{and} \qquad \frac{1}{T-1}\sum_{t=2}^{T}\tilde{\mathcal{V}}_{t-1} \;\to_p\; \bar{\mathcal{V}}.$$

Remark OA.28. Theorem OA.10 extends to the multivariate case ($K_\mathrm{z} > 1$) as follows.

(i)

$$\tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1} \;\to_p \boldsymbol{\gamma}_{\mathrm{z},t-1} \qquad \text{and} \qquad \bar{\tilde{\boldsymbol{\gamma}}}_\mathrm{z} \to_p \bar{\boldsymbol{\gamma}}^0_\mathrm{z}.$$

(ii) Let $\frac{\sigma^2}{(T-1)}\sum_{t=2}^{T}\left(\frac{\mathbf{Z}'_{t-1}\mathbb{M}_{1_N}\mathbf{Z}_{t-1}}{N}\right)^{-1} \to_p \bar{\mathbf{V}}$. Then,

$$\sqrt{N}\left(\tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1} - \boldsymbol{\gamma}_{\mathrm{z},t-1}\right) \to_d \mathcal{N}\left(\mathbf{0}_{K_\mathrm{z}}, \boldsymbol{V}_{t-1}\right) \qquad \text{and}$$

$$\sqrt{NT}\left(\bar{\tilde{\boldsymbol{\gamma}}}_\mathrm{z} - \bar{\boldsymbol{\gamma}}_\mathrm{z}\right) \to_d \mathcal{N}\left(\mathbf{0}_{K_\mathrm{z}}, \bar{\mathbf{V}}\right).$$

(iii)

$$(T-1)\widehat{\mathrm{Var}}\left[\bar{\tilde{\boldsymbol{\gamma}}}_\mathrm{z}\right] = \frac{1}{T-1}\sum_{t=2}^{T}\left(\tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1} - \bar{\tilde{\boldsymbol{\gamma}}}_\mathrm{z}\right)\left(\tilde{\boldsymbol{\gamma}}_{\mathrm{z},t-1} - \bar{\tilde{\boldsymbol{\gamma}}}_\mathrm{z}\right)' \to_p \boldsymbol{\Sigma}_{\gamma_\mathrm{z}}.$$

OA.8 Granularity

The relevance of granularity can be understood from the following result, reported without proof, which extends Adler and Rosalsky (1991).

Proposition OA.2 (limiting behavior of weighted averages)**.** *For an iid sequence Y_i with* $\mathrm{E}[|Y|^2] < \infty$*, assume the following granularity conditions hold, for some finite constant C,*

$$\sum_{i=1}^{N} a_i^2 = o(b_N^2), \quad \frac{1}{b_N} \sum_{i=1}^{N} a_i \to C.$$

(i) Then

$$\frac{1}{b_N} \sum_{i=1}^{N} a_i Y_i \to_p C\mathrm{E}[Y] < \infty.$$

(ii) If, in addition, for some finite $n < N$ and some constant $0 < C_i < \infty$, $1 \leq i \leq n$,

$$\frac{a_i}{b_N} \to C_i \textit{ for every } 1 \leq i \leq n < \infty,$$

then

$$\frac{1}{b_N} \sum_{i=1}^{N} a_i Y_i \to_p C\mathrm{E}[Y] + \sum_{i=1}^{n} C_i (Y_i - \mathrm{E}[Y]).$$

Case (i) is the granular case, which is implicit in our regularity assumptions, setting $a_i = w_{i,t-1}$ and $b_N = N$, with $C = 1$, such that $w_{i,t}/N = O_p(N^{-1})$. Note that only when $C = 1$, the simple and weighted average converge to the same limit, namely, $\mathrm{E}[Y]$. Case (ii) is the non-granular case, which leads to a random limit of the weighted average, i.e., the second term $\sum_{i=1}^{n} C_i (Y_i - \mathrm{E}[Y])$.

OA.9 No-Arbitrage with Anomalies

We show how the presence of anomalies does not necessarily rule out no arbitrage with the following proposition, reported without proof.

Proposition OA.3 (no-arbitrage with anomalies)**.** *The asset pricing restriction (3) does* not *violate conditional no-arbitrage whenever (OA.5) holds with*

$$\sup_N \boldsymbol{\gamma}'_{t-1,z} \mathbf{Z}'_{t-1} (\mathrm{E}[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}'_t | I_{t-1}, \boldsymbol{\Pi}])^{-1} \mathbf{Z}_{t-1} \boldsymbol{\gamma}'_{t-1,z} \leq C < \infty \textit{ almost surely,} \tag{OA.112}$$

for some constant C, where we set the $N \times K_z$ matrix $\mathbf{Z}_t = (\mathbf{z}_{1,t}, \cdots, \mathbf{z}_{N,t})'$.

As practical examples when Proposition OA.3 does and does not hold, consider the case when $\mathrm{E}[\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t'|I_{t-1},\mathbf{\Pi}]$ exhibits a limited degree of cross-dependence of the $\boldsymbol{\epsilon}_t$, such as when equal to $\sigma^2\mathbf{I}_N$ for some constant scalar σ^2. Then, condition (OA.112) is satisfied when either $\boldsymbol{\gamma}_z = O(N^{-\frac{1}{2}})$ or $\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1} = O(1)$. These restrictions are needed because the $\mathbf{Z}_{t-1}$ affect the mean but not the variances and covariances of the returns. In contrast, when instead $\mathrm{E}[\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t'|I_{t-1},\mathbf{\Pi}] = \mathbf{Z}_{t-1}\mathbf{C}_z\mathbf{Z}_{t-1}' + \sigma^2\mathbf{I}_N$, for some $K_z \times K_z$ constant non-singular matrix $\mathbf{C}_z$, then (OA.112) is redundant as it imposes no cross-sectional constraint. In fact, by the Sherman-Morrison decomposition, whenever $\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}$ diverges as $N \to \infty$, one obtains $\mathbf{Z}_{t-1}'(\mathrm{E}[\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t'|I_{t-1},\mathbf{\Pi}])^{-1}\mathbf{Z}_{t-1} = \mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}(\sigma^2\mathbf{C}_z^{-1} + \mathbf{Z}_{t-1}'\mathbf{Z}_{t-1})^{-1}\mathbf{C}_z^{-1} \to_p \mathbf{C}_z^{-1}$, that is, bounded even for large N. No restriction arises because the $\mathbf{Z}_{t-1}$ affect the mean, variance, and covariances of the returns in the same way, making their effect neutral in terms of the risk-return trade-off.

OA.10 Supplementary Material to Section 3 and Section 4: Monte Carlo Experiments

In this Section, we demonstrate the finite-sample properties of the estimators presented in Section 3 and Section 4, i.e., the conventional estimators and our bias-adjusted estimators, using Monte Carlo exercises. These simulations confirm that, regardless of the regime considered, the conventional approach appears to be ill-suited, either because it is inconsistent or because, when consistent, the inferential theory is invalid. All the challenges related to the conventional approach can be instead resolved using a new OLS *bias-adjusted* estimator defined in Section 4, which is valid when N is large while T is kept fixed.

For our Monte Carlo experiment, the return-generating process is given by

$$\mathbf{R}_t = \gamma_{0,t-1}\mathbf{1}_N + \mathbf{Z}_{t-1}\boldsymbol{\gamma}_{\mathrm{z,t-1}} + \mathbf{B}\delta_{\mathrm{f,t-1}} + \boldsymbol{\epsilon}_t, \tag{OA.113}$$

where $\mathbf{1}_N$ denotes a $N \times 1$ vector of ones, $\mathbf{Z}_{t-1} = (\mathbf{z}_{1,t-1}, \cdots, \mathbf{z}_{N,t-1})'$ represents the $N \times K_{\mathrm{z}}$ matrix of anomalies at time $t-1$, where we consider $K_{\mathrm{z}} = 2$ (two anomalies), and one risk factor, i.e.

$$\delta_{\mathrm{f,t-1}} = \gamma_{\mathrm{f,t-1}} + f_t - \mathrm{E}\left[f_t | I_{t-1}, \mathbf{\Pi}\right]. \tag{OA.114}$$

We calibrate f_t in (OA.114) using the excess market return (we refer to the paper for the full description of the data). We set $\gamma_{0,t} = 0.5$ for every t, while we generate $\boldsymbol{\gamma}_{\mathrm{f,t-1}}$ and $\boldsymbol{\gamma}_{\mathrm{z,t-1}}$ in (OA.113)-(OA.114) as

$$\begin{aligned}
\boldsymbol{\gamma}_{\mathrm{f,t}} &= \mu_{\mathrm{f}}(1-\phi_{\mathrm{f}}) + \phi_{\mathrm{f}}\boldsymbol{\gamma}_{\mathrm{f,t-1}} + u_{\mathrm{f,t}}, && \text{(OA.115)}\\
\boldsymbol{\gamma}_{\mathrm{z_1,t}} &= \mu_{\mathrm{z_1}}(1-\phi_{\mathrm{z_1}}) + \phi_{\mathrm{z_1}}\boldsymbol{\gamma}_{\mathrm{z_1,t-1}} + u_{\mathrm{z_1,t}}, && \text{(OA.116)}\\
\boldsymbol{\gamma}_{\mathrm{z_2,t}} &= 0, && \text{(OA.117)}
\end{aligned}$$

with $\mu_{\mathrm{f}} = 0.5$, $\phi_{\mathrm{f}} = 0.1$, $\mu_{\mathrm{z_1}} = 0.2$, $\phi_{\mathrm{z_1}} = 0.35$, and where $u_{\mathrm{f,t}}$, and $u_{\mathrm{z_1,t}}$ are i.i.d. normally distributed with zero mean and standard deviation equal to 0.01. The N-vector of betas is generated from a normal distribution $\mathcal{N}(1, 0.5)$.

To define the two anomalies in (OA.113), we first generate $\mathbf{Z}^{\dagger}_{1,t-1}$ and $\mathbf{Z}^{\dagger}_{2,t-1}$ from i.i.d. normally distributed random variables with mean 0.04 and standard deviation 0.02, respectively. We then orthogonalize both $\mathbf{Z}^{\dagger}_{1,t-1}$ and $\mathbf{Z}^{\dagger}_{2,t-1}$ with the market factor f_t, that is:

$$\mathbf{Z}_{1i} = \mathbf{M}_D \mathbf{Z}^{\dagger}_{1i} + \mathbf{1}_{T-1} \bar{Z}^{\dagger}_{1i} \quad \text{comma} \tag{OA.118}$$

$$\mathbf{Z}_{2i} = \mathbf{M}_D \mathbf{Z}^{\dagger}_{2i} + \mathbf{1}_{T-1} \bar{Z}^{\dagger}_{2i} \quad \text{comma} \tag{OA.119}$$

where $\mathbf{Z}_{1i} = [Z_{1i,1}, \ldots, Z_{1i,T-1}]'$, $\mathbf{Z}_{2i} = [Z_{2i,1}, \ldots, Z_{2i,T-1}]'$, $\mathbf{M}_D = \mathbf{I}_{T-1} - \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'$, with $\mathbf{D} = (\mathbf{1}_{T-1}, \mathbf{f})$ and where $\bar{Z}^{\dagger}_{ik} = \frac{1}{T-1}\sum_{t=1}^{T-1} Z^{\dagger}_{ki,t}$, with $k = 1, 2$. In this way, we ensure that $\mathrm{cov}(Z_{1i,t-1}, f_t) = \mathrm{cov}(Z_{2i,t-1}, f_t) = 0$ and that the sample averages of both $\mathbf{Z}_{1i}$ and $\mathbf{Z}_{2i}$ match those of the original data. We then use (OA.118) and (OA.119) in the data-generating process in (OA.113). In our simulation design, both the factor and the anomaly realizations are taken as given and kept fixed throughout.

For the calibration of the error term, we allow the model disturbances to be weakly cross-sectionally correlated with the anomalies. In particular, we consider the following data-generating process for the error terms:

$$\epsilon_{it} \quad = \quad \frac{\sigma}{\tau}\left(u_{it} + \frac{\mathbf{1}'_{K_z}\boldsymbol{\eta}_{it}}{N^{\delta}}\right) \quad \text{comma} \tag{OA.120}$$

where $\tau = 1 + \frac{K_z}{N^{2\delta}}$, u_{it} is generated from an i.i.d. standard normal random variable, $\sigma^2 = (0.1)^2$, while the $K_z \times 1$ vector $\boldsymbol{\eta}_{it}$ is calibrated using the standardized residuals obtained by fitting a vector autoregressive (VAR) process of order 1 on the two anomalies. The parameter δ controls the strength of the cross-sectional correlation between the shocks and the anomalies: the higher the value of δ is, the weaker the cross-sectional correlation is. For our theoretical results to hold, we require $\delta \geq 0.5$. In the simulation experiment, we set $\delta = 0.8$.

To comprehensively assess the performance of the different estimators under different sampling schemes, we analyze two distinct settings. In the first setting, where T is large and N is fixed, we fix the number of assets at $N = 25$ and progressively increase the number of periods from $T = 24, 36, 72, 120$, and 240. In the second setting, instead, we fix $T = 36$ and consider the following increasing sequence of values for $N = 25, 50, 100, 500, 1000$, and 2000.

To facilitate the reading and highlight the complementarity with Sections 3 and 4, we begin by presenting the results for the conventional approach. For simplicity, we restrict attention to estimators where both risk factors and anomalies are included in the model, focusing on the performance of both the time-varying estimator in (19) and average estimator defined in (28), under the two asymptotic regimes.

The main results are summarized in Table OA.IV below, which mirrors Table I of Section 3 and includes references to the corresponding figures that illustrate the performance of each case.

Estimator / Regime	**Conventional Time-varying (Eq. (19))**	**Coventional Average (Eq. (28))**
(i) As $T \to \infty$ (fixed N)	• See Figures OA.4 and OA.5. The figures illustrate the behavior of the conventional t-ratio—computed under the null hypothesis that the parameter equals its true value—based on estimator (19), for both $\gamma_{z_1,t-1}$ and $\gamma_{z_2,t-1}$, as T increases with N fixed at 25. In both cases, although the distribution becomes increasingly centered around zero as T grows, the estimator remains inconsistent because the estimated variance is invalid, leading to incorrect inference.	• See Figures OA.6 and OA.7. The figures illustrate the behavior of the conventional t-ratio—computed under the null hypothesis that the parameter equals its true value—based on the average estimator (28), for both $\bar{\gamma}_{z_1}$ and $\bar{\gamma}_{z_2}$, as T increases with N fixed at 25. In both figures, the distribution becomes centered around the true long-run value (indicated by the red dotted line in Figure OA.6 and the black dotted line in Figure OA.7). However, inference remains invalid due to the EIV problem and the time-varying nature of the premia parameters.
(ii) As $N \to \infty$ (fixed T)	• See Figures OA.8 and OA.9. The figures illustrate the behavior of the conventional t-ratio—computed under the null hypothesis that the parameter equals its true value—based on estimator (19), for both $\gamma_{z_1,t-1}$ and $\gamma_{z_2,t-1}$, as N increases with T fixed at 36. The estimator is both biased and inconsistent, due to EIV and time-variation sources of bias which cannot vanish when T is fixed.	• See Figures OA.10 and OA.11. The figures illustrate the behavior of the conventional t-ratio—computed under the null hypothesis that the parameter equals its true value—based on the average estimator (28), for both $\bar{\gamma}_{z_1}$ and $\bar{\gamma}_{z_2}$, as N increases with T fixed at 36. The distributions remain always biased, because the EIV bias does not vanish when T is kept fixed.

Table OA.IV: The behaviour of the conventional estimators under different regimes.

As expected, Table OA.IV confirms that, regardless of the asymptotic regime considered, the conventional approach is generally ill-suited—either due to inconsistency or, when consistency holds, because the inferential theory breaks down.

To validate our new approach of Section 4, we then repeat the same exercise considering the time-varying bias-adjusted estimator defined in (35). We summarize the main results in Table OA.V, which includes references to the corresponding figures that illustrate the performance of the estimators under the two different regimes. In this exercise, we omit the Monte Carlo results for the average bias-adjusted estimator and refer the reader to Raponi et al. (2020) for a partial analysis.

Regime \ **Estimator**	**Bias-adjusted Time-varying (Eq.** (35))	**Bias-adjusted Average (Eq.** (44))
(i) As $T \to \infty$ (fixed N)	• See Figure OA.12. The figure shows the distribution of the $\boldsymbol{\gamma}_{z,t-1}$ estimates obtained using the conventional approach (yellow) and the proposed bias-adjusted estimator (blue) introduced in (35)), across increasing values of T (6, 12, 24, 36, 120, 360), with N fixed at 25. As predicted by Remark 2, the two estimators produce different results when T is small, but their distributions gradually converge as T increases.	• See Raponi, Robotti, and Zaffaroni (2020), Section 3.3
(ii) As $N \to \infty$ (fixed T)	• See Figures OA.13 and OA.14. The figures illustrate the behavior of the t-ratio - computed under the null hypothesis that the parameter equals its true value—based on the new time-varying bias adjusted estimator defined in (35), for both $\gamma_{z_1,t-1}$ and $\gamma_{z_2,t-1}$, as N increases with T fixed at 36. The estimated variance is obtained based on the results of Theorem 2. The estimator exhibits good performance and follows a normal distribution.	• See Raponi, Robotti, and Zaffaroni (2020), Section 3.3

Table OA.V: The behaviour of the bias-adjusted estimators under different regimes.

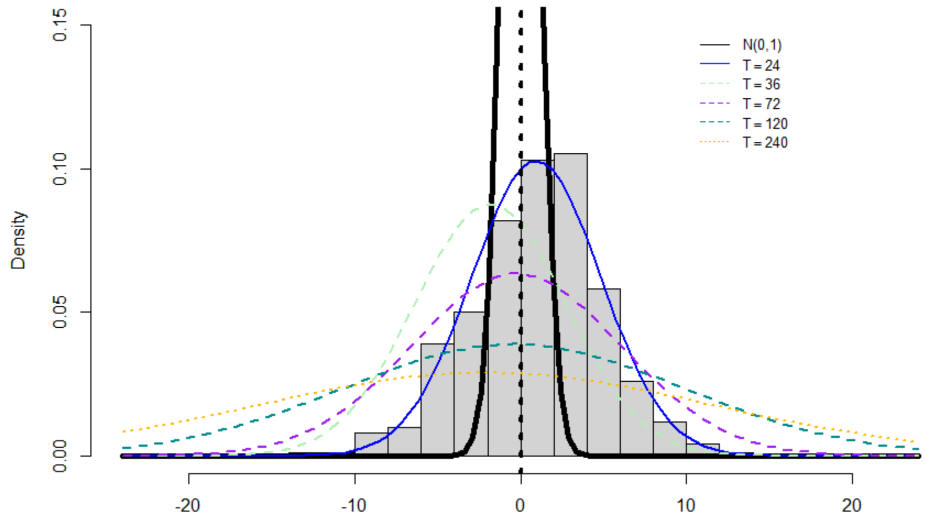


Figure OA.4 Performance of the conventional time-varying estimator for $\gamma_{z_1,t}$ as T increases with $N = 25$. The figures illustrate the behavior of the conventional t-ratio—computed under the null hypothesis that the parameter equals its true value—based on the estimator (19) for $\gamma_{z_1,t-1}$ as T increases from 24 to 240, with N fixed at 25. The black curve represents the standard Normal distribution. Data are generated following the model in (OA.113).

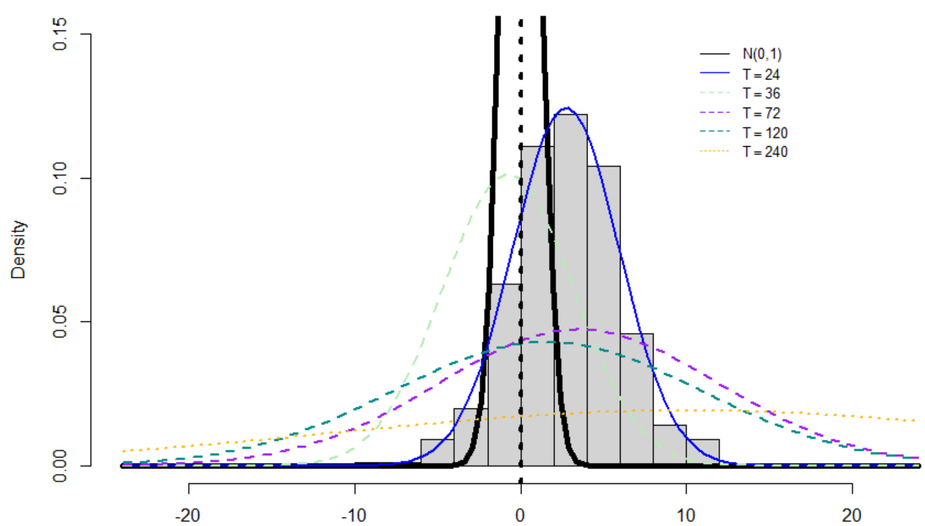


Figure OA.5 Performance of the conventional time-varying estimator for $\gamma_{z_2,t}$ as T increases with $N = 25$. The figures illustrate the behavior of the conventional t-ratio—computed under the null hypothesis that the parameter equals its true value—based on the estimator (19) for $\gamma_{z_2,t-1}$ as T increases from 24 to 240, with N fixed at 25. The black curve represents the standard Normal distribution. Data are generated following the model in (OA.113).

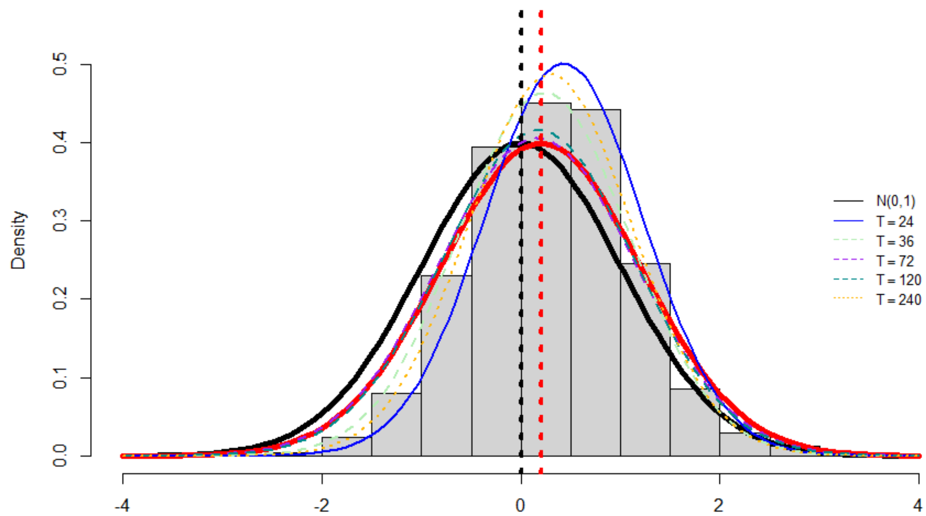


Figure OA.6 Performance of the conventional local average estimator for $\gamma_{z_1,t}$ as T increases with $N = 25$. The figures illustrate the behavior of the conventional t-ratio—computed under the null hypothesis that the parameter equals its true value—based on the average estimator (28) for $\bar{\gamma}_{z_1}$ as T increases from 24 to 240, with N fixed at 25. The black solid curve represents the standard Normal distribution, while the red curve denotes a Normal distribution centered around the long-run value (indicated by the red dotted line). Data are generated following the model in (OA.113).

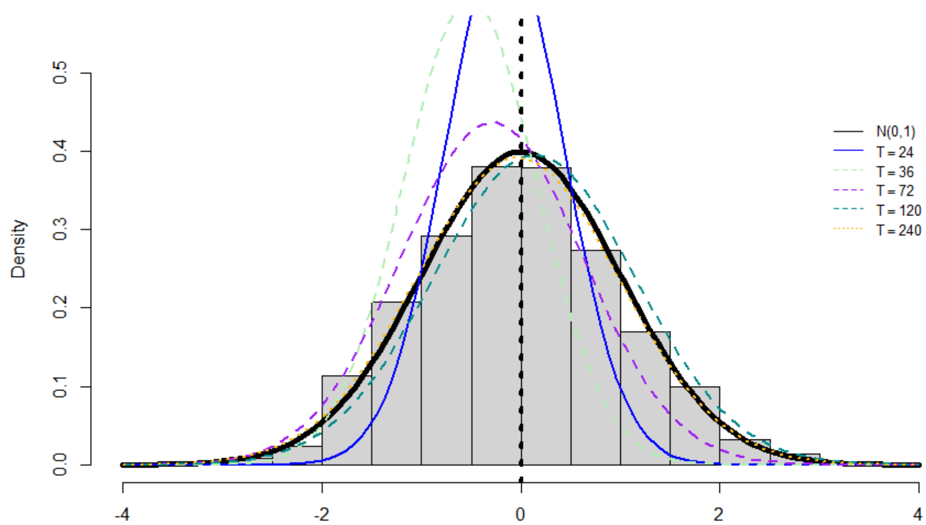


Figure OA.7 Performance of the conventional local average estimator for $\gamma_{z_2,t}$ as T increases with $N = 25$. The figures illustrate the behavior of the conventional t-ratio—computed under the null hypothesis that the parameter equals its true value—based on the average estimator (28) for $\bar{\gamma}_{z_2}$ as T increases from 24 to 240, with N fixed at 25. The black solid curve represents the standard Normal distribution. In this case, the long-run value coincides with the zero value (indicated by the black dotted line). Data are generated following the model in (OA.113).

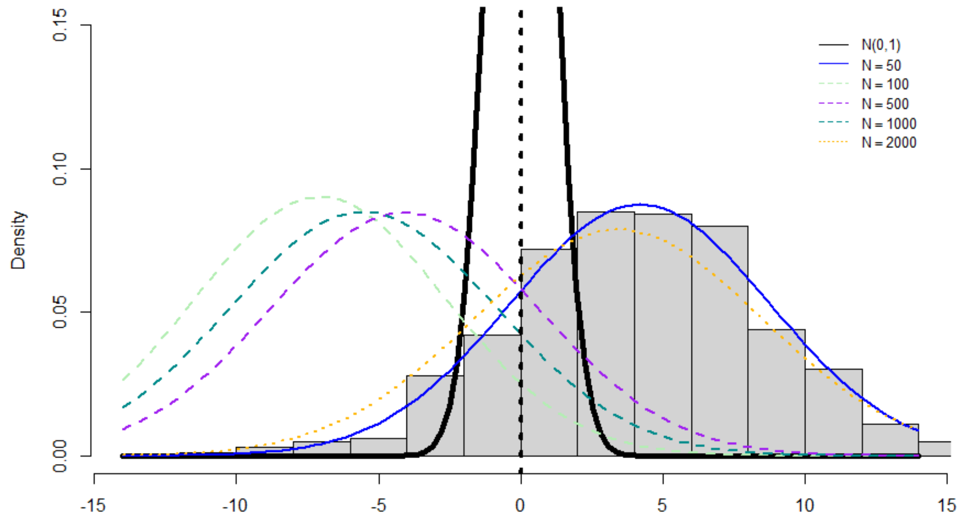


Figure OA.8 Performance of the conventional time-varying estimator for $\gamma_{z_1,t}$ as N increases with $T = 36$. The figures illustrate the behavior of the conventional t-ratio—computed under the null hypothesis that the parameter equals its true value—based on the estimator (19) for $\gamma_{z_1,t-1}$ as N increases from 50 to 2,000, with T fixed at 36. The black curve represents the standard Normal distribution. Data are generated following the model in (OA.113).

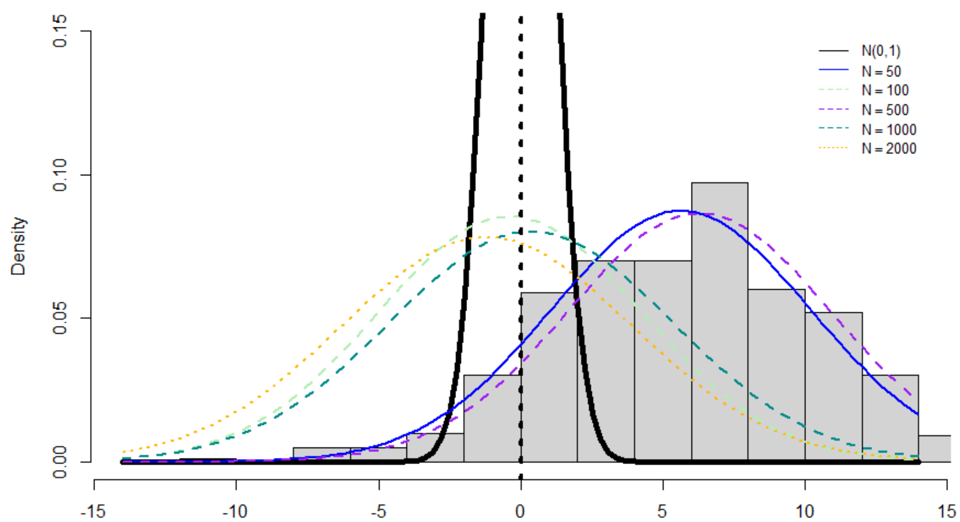


Figure OA.9Performance of the conventional time-varying estimator for $\gamma_{z_2,t}$ as N increases with $T = 36$. The figures illustrate the behavior of the conventional t-ratio—computed under the null hypothesis that the parameter equals its true value—based on the estimator (19) for $\gamma_{z_2,t-1}$ as N increases from 50 to 2,000, with T fixed at 36. The black curve represents the standard Normal distribution. Data are generated following the model in (OA.113).

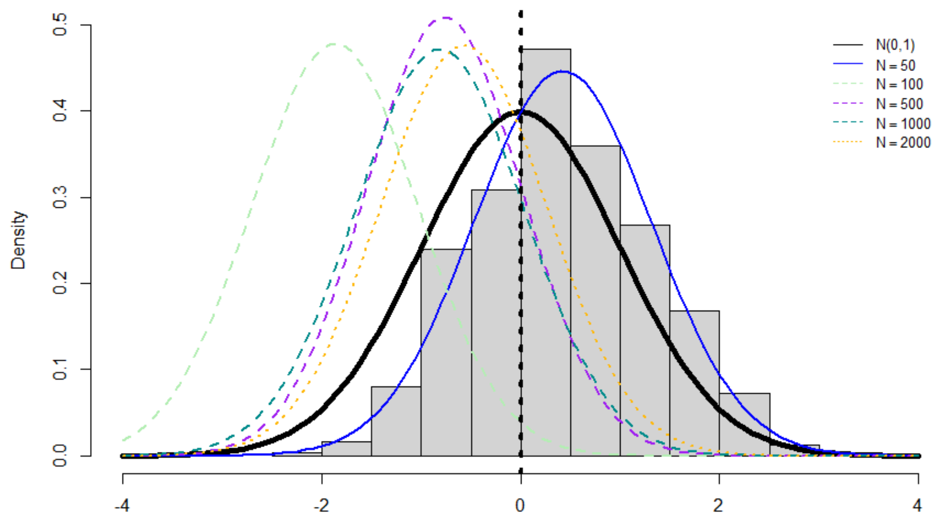


Figure OA.10 Performance of the conventional local average estimator for $\gamma_{z_1,t}$ as N increases with $T = 36$. The figures illustrate the behavior of the conventional t-ratio—computed under the null hypothesis that the parameter equals its true value—based on the average estimator (28) for $\bar{\gamma}_{z_1}$ as N increases from 50 to 2,000, with T fixed at 36. The black curve represents the standard Normal distribution. Data are generated following the model in (OA.113).

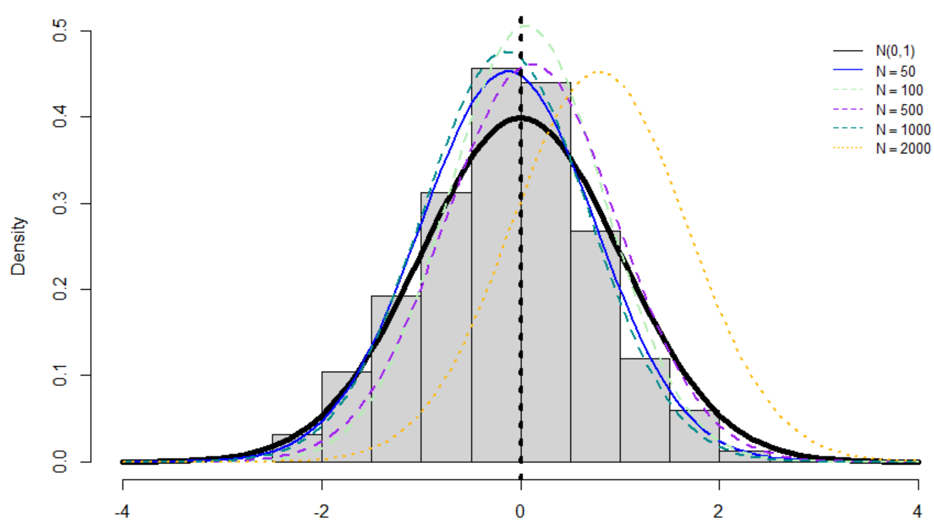


Figure OA.11 Performance of the conventional local average estimator for $\gamma_{z_2,t}$ as N increases with $T = 36$. The figures illustrate the behavior of the conventional t-ratio—computed under the null hypothesis that the parameter equals its true value—based on the average estimator (28) for $\bar{\gamma}_{z_2}$ as N increases from 50 to 2,000, with T fixed at 36. The black curve represents the standard Normal distribution. Data are generated following the model in (OA.113).

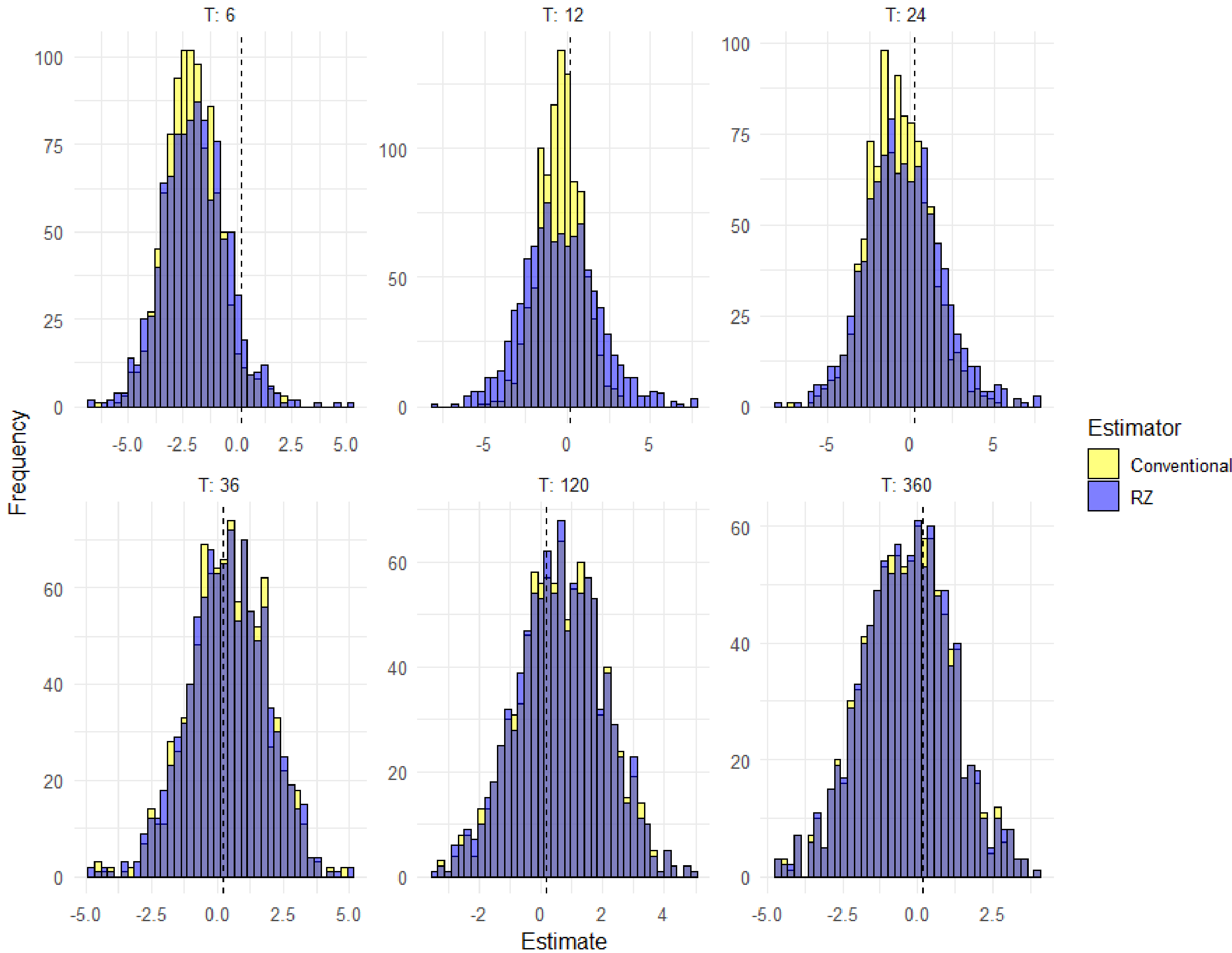


Figure OA.12 Performance of the bias-adjusted (RZ) time-varying estimator for $\gamma_{z_1,t}$ as T increases with $N = 25$. The figure shows the distribution of the $\boldsymbol{\gamma}_{z,t-1}$ estimates obtained using the conventional approach (yellow) and the proposed bias-adjusted estimator (blue) introduced in (35)), across increasing values of $T = 6, 12, 24, 36, 120$, and 360, with N fixed at 25. Data are generated following the model in (OA.113).

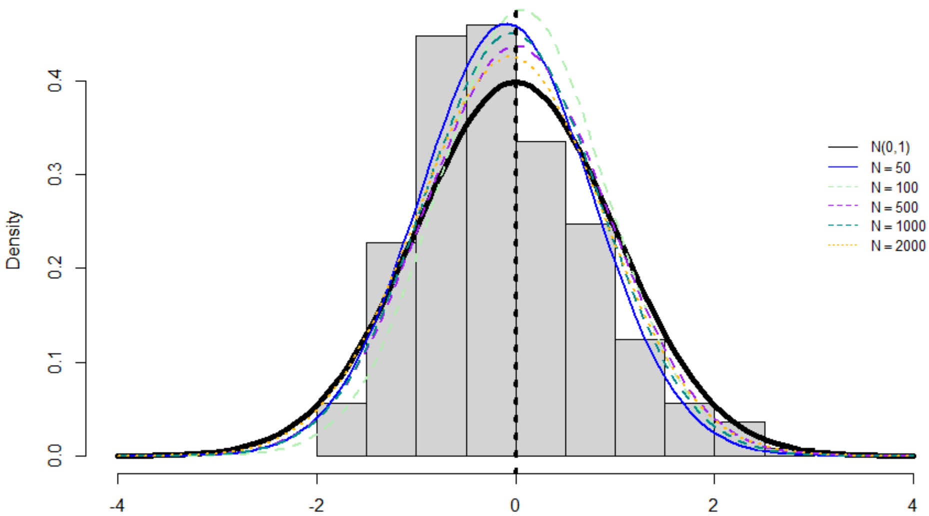


Figure OA.13 Performance of the bias-adjusted (RZ) time-varying estimator for $\gamma_{z_1,t}$ as N increases with $T = 36$. The figures illustrate the behavior of the t-ratio - computed under the null hypothesis that the parameter equals its true value—based on the new time-varying bias adjusted estimator defined in (35) for $\gamma_{z_1,t-1}$ as N increases from 50 to 2,000, with T fixed at 36. The estimated variance is obtained based on the results of Theorem 2. The black solid curve denotes the standard Normal distribution. Data are generated following the model in (OA.113).

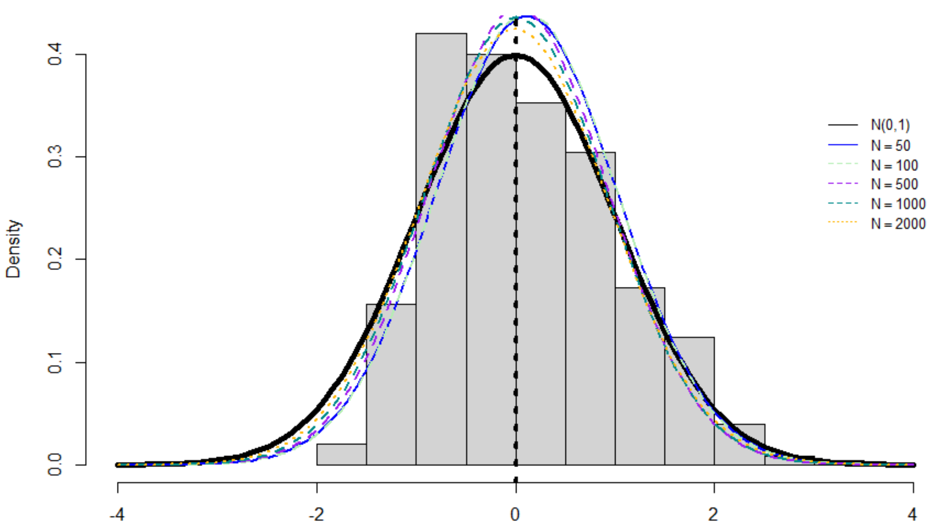


Figure OA.14 Performance of the bias-adjusted (RZ) time-varying estimator for $\gamma_{z_2,t}$ as N increases with $T = 36$. The figures illustrate the behavior of the t-ratio - computed under the null hypothesis that the parameter equals its true value—based on the new time-varying bias-adjusted estimator defined in (35) for $\gamma_{z_2,t-1}$ as N increases from 50 to 2,000, with T fixed at 36. The estimated variance is obtained based on the results of Theorem 2. The black solid curve denotes the standard Normal distribution. Data are generated following the model in (OA.113).

OA.11 Supplementary Material to Section 6: Empirical Application

In this section we provide supplementary material related to the empirical application presented in Section 6 of the main manuscript. The additional results are organized in three parts. First, we report a set of preliminary analyses on identification, locally-constant risk exposures, and the time-variation of anomaly premia, which complement and extend the discussion in the main text. Second, we present further results obtained using the FHNW dataset (Freyberger, Hoeppner, Neuhierl, and Weber (2025)), with the aim of validating the robustness of our findings across alternative empirical settings, like the FF3 and FF5 models. Finally, we replicate the same analyses using the BLLP dataset (Bryzgalova, Lerner, Lettau, and Pelger (2025), including univariate and multivariate specifications as well as WLS estimation. These supplementary results are intended to reinforce the empirical evidence in favor of our methodology and to document its robustness to different data sources and estimation strategies.

OA.11.1 Preliminary Analysis: Identification, Locally-Constant Risk Exposures and Time-Variation

In this section, we provide the preliminary analyses that support our main empirical application. Specifically, we verify that anomalies are genuinely driving pricing errors, rather than merely capturing time variation in risk exposures (betas), and we show that the key assumptions underlying the conventional two-pass approach are systematically violated in the data. Taken together, this evidence highlights the necessity of departing from the conventional framework and makes our proposed methodology compelling for dissecting the role of anomalies in explaining the cross-section of asset returns.

In the following, Sections OA.11.1.1, OA.11.1.2, and OA.11.1.3 provide key empirical evidence supporting our time-varying approach and its key assumptions. Section OA.11.1.1 validates the identification strategy, showing that anomalies (not betas) drive pricing errors. Section OA.11.1.2 supports the assumption of locally constant betas, while Section OA.11.1.3 shows that core assumptions of the conventional approach appear violated in the data.

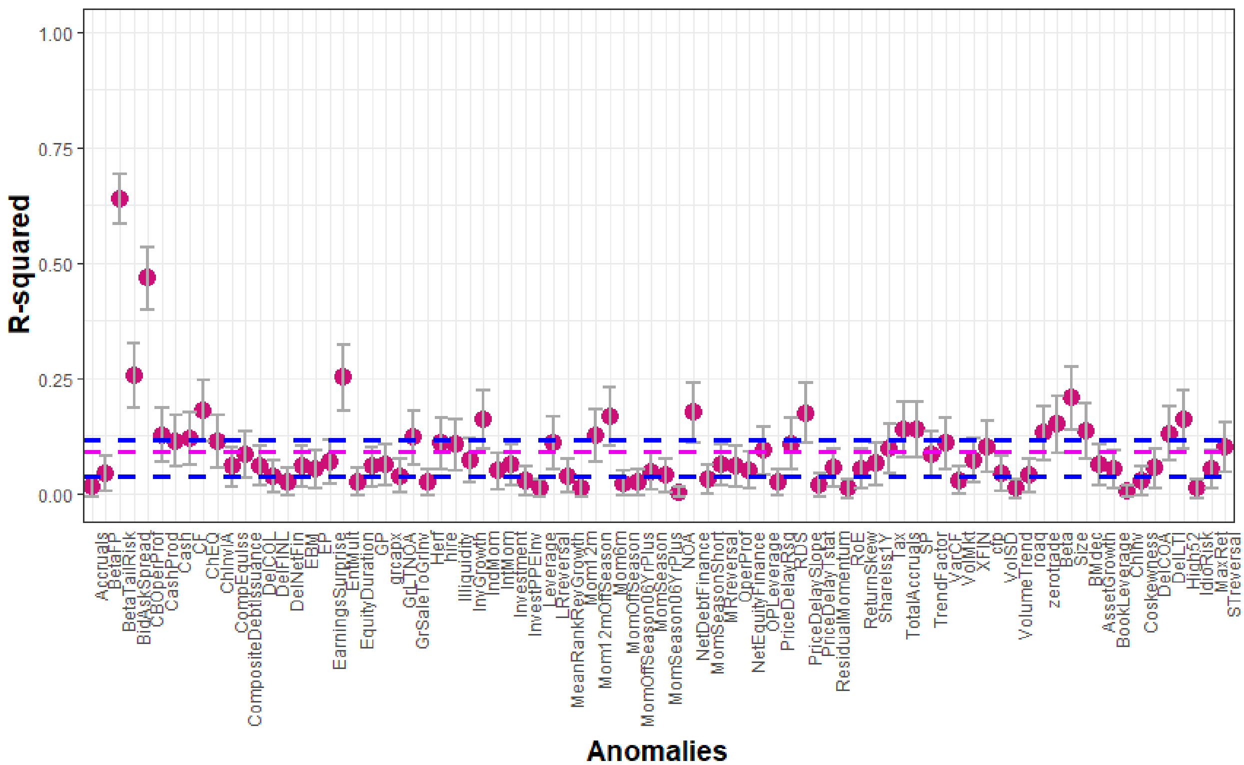
OA.11.1.1 Preliminary Analysis I: Identification

Before presenting our main empirical results, we first empirically evaluate the restriction in (13) to ensure the identification of the anomalies' effect. That is, we want to assess whether time-varying pricing errors are genuinely driven by the anomalies, without masking any time-varying effect in the risk exposures. To this end, we initially estimate the parameters $\gamma_{\mathrm{z},t-1}$ and $\boldsymbol{\delta}_{\mathrm{f},t-1}$ from the regression in (34) using our time-varying estimator in (35). In each regression, we consider one anomaly at the time (i.e., $K_{\mathrm{z}} = 1$), while the estimated betas $\hat{\mathbf{B}}$ are obtained using (20), in which the matrix $\mathbf{F}$ contains the Fama-French five-factors (i.e., $K_{\mathrm{f}} = 5$). Then, for each anomaly, we regress the estimated time series of $\hat{\boldsymbol{\gamma}}_{\mathrm{z},t-1}$ against the time series of $\hat{\boldsymbol{\delta}}_{\mathrm{f},t-1}$, and report the corresponding R-squared, along with its confidence interval. The results are displayed in Figure OA.15.[39] The figure shows that in 76.7% of the cases, the R^2 confidence intervals do not include zero, and in all cases, they lie well below one. Following Section 2.1.1, these results allow us to empirically rule out Case (ii) of *observational equivalence.* Instead, the evidence supports Case (i), where anomalies span both pricing errors and betas, in approximately three-quarters of the cases. The remaining cases are consistent with Case (iii) where anomalies span only pricing errors. This establishes the foundation for our empirical analysis, which we will conduct in the following sections.

[39]For simplicity, we display the results of this section using Freyberger, Neuhierl, and Weber (2020) dataset but the same findings hold using the Bryzgalova, Lerner, Lettau, and Pelger (2025) dataset. Details are available upon request.

space

Figure OA.15**R-squared across anomalies in the regression of $\boldsymbol{\gamma}_{\mathrm{z}}$ against $\boldsymbol{\delta}_{\mathrm{f}}$.** R-squared (95% confidence intervals in grey) of the time-series regression of each $\boldsymbol{\gamma}_{\mathrm{z}}$ against all elements of $\boldsymbol{\delta}_{\mathrm{f}}$ for each anomaly. Mean, 25-th, and 75-th percentiles (across anomalies) in pink and blue lines, respectively (Freyberger, Hoeppner, Neuhierl, and Weber (2025) dataset).



OA.11.1.2 Preliminary Analysis II: Locally-Constant Risk Exposures

To justify Assumption 1, we provide evidence that the (locally-constant) estimated betas for the five Fama-French risk factors exhibit meaningful time variation. To do so, we regress the cross-sectional averages of the estimated betas—both in levels (*Level*) and absolute values (Abs)—on four macro-financial variables (i.e., instruments) commonly used to capture time-variation of the risk exposures to systematic risk: the S&P500 index return (SP500), market dividend yield (DIVYIELD), term spread (TERMSPREAD), and the cyclically adjusted price-to-earnings ratio (CAPE).[40]

We repeat this analysis across different time windows ($T = 12, 24, 36, 72, 120$ months). Table OA.VI reports the first-order correlations between each beta and each instrument, and Table OA.VII shows the corresponding R^2 from regressions using all four instruments jointly. These results confirm that even when betas are estimated over longer windows, their time-variation persists, and is systematically related to the instruments. This supports the conclusion that the locally constant beta assumption in our methodology is not overly restrictive, and that betas were indeed given a fair opportunity to contribute to the model's explanatory power.

[40] S&P500, DIVYIELD, and CAPE are from Robert Shiller's website; TERMSPREAD is from the FRED database.

Table OA.VI: **First-Order Correlations Between Betas and Macro-Financial Variables**. The table reports the first-order correlations between the cross-sectional averages of the estimated betas—both in levels (*Level*) and absolute values (*Abs*)—and four macro-financial instruments, namely the S&P500 index return, the market dividend yield (DIVYIELD), the term spread (TERMSPREAD) between 10-year and 3-month Treasury rates, and the cyclically adjusted price-to-earnings (CAPE) ratio. The estimated betas correspond to the five Fama-French risk factors and are obtained using different time windows ($T = 12, 24, 36, 72, 120$ months). In bold, we highlight the highest correlation values for each risk factor and time window, across the two averaging methods.

	$T=12$		$T=24$		$T=36$		$T=72$		$T=120$	
	Level	Abs	Level	Abs	Level	Abs	Level	Abs	Level	Abs
Panel A: SP500										
MKT	0.07	**0.09**	0.13	**0.19**	0.27	**0.32**	0.27	**0.42**	0.40	**0.48**
SMB	**-0.36**	-0.14	**-0.56**	-0.33	**-0.63**	-0.41	**-0.70**	-0.49	**-0.54**	-0.36
HML	0.04	**-0.26**	0.07	**-0.45**	0.06	**-0.49**	0.18	**-0.41**	0.23	**-0.36**
RMW	0.02	**-0.13**	0.08	**-0.30**	0.11	**-0.38**	0.23	**-0.31**	**0.17**	0.00
CMA	-0.04	**-0.18**	-0.06	**-0.27**	-0.06	**-0.27**	-0.24	**-0.26**	**-0.47**	-0.11
Panel B: DIVYIELD										
MKT	0.08	**-0.28**	0.13	**-0.44**	0.04	**-0.58**	0.10	**-0.32**	**0.19**	0.02
SMB	**0.32**	0.04	**0.45**	0.15	**0.46**	0.25	**0.43**	0.40	0.14	**0.26**
HML	**-0.11**	0.01	**-0.21**	0.14	**-0.29**	0.18	**-0.61**	-0.10	**-0.74**	-0.34
RMW	**0.04**	-0.02	0.03	**0.08**	-0.07	**0.19**	**-0.49**	0.12	**-0.69**	-0.03
CMA	-0.01	**-0.05**	**0.07**	0.05	0.08	**0.09**	**0.17**	0.17	**0.07**	-0.05
Panel C: TERMSPREAD										
MKT	**-0.06**	-0.02	**-0.06**	-0.05	**-0.12**	-0.09	**0.23**	0.08	**0.28**	0.25
SMB	**0.03**	0.00	0.05	**0.23**	0.08	**0.23**	0.05	**0.11**	**-0.19**	-0.09
HML	0.03	**-0.05**	-0.03	**-0.24**	0.03	**-0.33**	0.15	**-0.12**	-0.06	**-0.41**
RMW	0.09	**-0.20**	-0.08	**-0.20**	**-0.15**	-0.14	**-0.09**	0.07	**-0.15**	-0.10
CMA	-0.03	**0.17**	-0.11	**0.18**	**-0.22**	0.07	**-0.29**	0.13	-0.12	**-0.15**
Panel D: CAPE										
MKT	-0.13	**0.26**	-0.19	**0.41**	-0.12	**0.48**	**-0.26**	0.13	**-0.37**	-0.24
SMB	**-0.26**	-0.06	**-0.37**	-0.25	**-0.37**	-0.31	**-0.28**	-0.30	**0.17**	0.07
HML	**0.04**	-0.03	**0.23**	-0.06	**0.31**	-0.07	**0.53**	0.25	**0.52**	0.34
RMW	**-0.06**	0.01	0.03	-0.03	**0.24**	-0.10	**0.56**	-0.02	**0.58**	0.27
CMA	**0.08**	-0.00	0.04	-0.04	**0.07**	-0.06	**0.09**	-0.03	0.07	**0.36**

Table OA.VII: **Beta Variation Explained by Macro-Financial Factors.** The table reports the R^2 values from regressions of the cross-sectional averages of the estimated betas—both in levels (*Level*) and absolute values (*Abs*)—on four macro-financial instruments: the S&P500 index return, the market dividend yield (DIVYIELD), the term spread (TERMSPREAD) between 10-year and 3-month Treasury rates, and the cyclically adjusted price-to-earnings (CAPE) ratio. The estimated betas correspond to the five Fama-French risk factors and are obtained using different time windows ($T = 12, 24, 36, 72, 120$ months). In bold, we highlight the highest R^2 value for each risk factor and time window.

	$T = 12$		$T = 24$		$T = 36$		$T = 72$		$T = 120$	
	Level	Abs	Level	Abs	Level	Abs	Level	Abs	Level	Abs
MKT	0.074	**0.107**	0.171	**0.236**	0.277	**0.383**	**0.381**	0.313	**0.611**	0.565
SMB	**0.171**	0.027	**0.383**	0.183	**0.455**	0.212	**0.566**	0.287	**0.650**	0.463
HML	0.030	**0.123**	0.068	**0.359**	0.150	**0.457**	0.535	**0.576**	0.623	**0.742**
RMW	0.018	**0.101**	0.032	**0.184**	0.164	**0.213**	**0.349**	0.171	**0.550**	0.226
CMA	0.041	**0.105**	0.061	**0.150**	**0.121**	0.110	**0.343**	0.200	0.351	**0.414**

OA.11.1.3 Preliminary Analysis III: Time-Variation and Cross-Correlations

This section provides empirical evidence on the shortcomings of the conventional approach in estimating time-varying anomaly premia. As discussed in Section 3, this method relies on orthogonality between factor betas and anomalies, and on constant premia over time—two assumptions that, as we show below, are often violated in the data.

In the first exercise, we assess the correlation between estimated betas from standard asset pricing models (CAPM, FF3, FF5) and anomaly variables. Betas are estimated via a first-pass regression on a rolling window of size $T = 240$, and at each point in time, we compute the R^2 from regressing each column of $\mathbf{Z}_t$ on the estimated $\hat{\mathbf{B}}$. This produces a time series of R^2 values, which we average and aggregate by category. Table OA.VIII summarizes the results: the first column reports average R^2 values (with min and max in parentheses), while the second column shows the percentage of times the correlation is statistically significant.[41] A clear pattern emerges: while the correlation between betas and anomalies may often be small in magnitude, in most of cases it appears to be statistically significant—highlighting the risk of bias in conventional two-pass estimators when this dependence is ignored.

To further support our results, we compute the canonical correlation between the full set of FF5 betas and all anomalies jointly. As shown in Figure OA.16, the average canonical correlation is 0.93, further challenging the orthogonality assumption inherent in the conventional approach.

In the next part of this section, our objective is to evaluate whether there is any supporting (or contradicting) evidence regarding the time-varying nature of the premia coefficients. To achieve this, we perform several exercises. In the first exercise, we estimate the time series of $\gamma_{\mathrm{z},t-1}$ using (21), one anomaly at a time, and fit ARMA(p, q) models (with p, q from 0 to 6) to each series. A best-fit white noise model ($p+q = 0$) would indicate no autocorrelation and minimal time variation. Using AIC (Akaike Information Criterion) for model selection, Table OA.IX (Column 1) reports the percentage of anomalies in each category where a time-varying model fits better ($p + q > 0$) than a white noise model. Columns 2–4 repeat the exercise, controlling for risk exposures using CAPM, FF3, and FF5 betas. Across all specifications, over 80% of the series exhibit significant

[41] To assess whether the correlation coefficient is statistically different from zero, we use the standard Pearson correlation test in the CAPM specification, while we use the F-test for the FF3 and FF5 models. We report the results using a 5% confidence level.

time variation—indicating that the anomaly premia are far from constant across time.

To further quantify this time variation, we estimate the first-order autocorrelation of each anomaly premium in a model augmented with the FF5 factors. As shown in Figure OA.17, the autocorrelations are significantly different from zero in nearly all cases (as indicated by the corresponding confidence intervals), with a mean of 0.121. These results confirm the presence of persistent time dependence and motivate the calibration used in our simulations in Section 3.[42]

To deepen our investigation on the strength of time variation, we test for structural breaks in their time series using the CUSUM test (see Ploberger, Krämer, and Kontrus (1989)).[43] Table OA.X reports the percentage of anomalies in each category where the null of no structural change is rejected. We find that structural breaks are widespread, especially in the *Investment* (60–70%) and *Momentum* (33–40%) categories.

As a final exercise, we test for significant shifts in the mean of anomaly premia by applying paired t-tests to all possible sub-sample pairs (three- and five-year windows) using the estimated $\gamma_{\mathrm{z},t}$. For each anomaly, we compute mean differences across sub-periods and test whether they differ significantly from zero. Table OA.XI reports the average percentage of sub-sample pairs (by category) with significant mean shifts at the 10% level. Panel A refers to three-year windows, Panel B to five-year windows. Column 1 uses estimates from regressions on anomalies only, while Columns 2–4 control for betas from CAPM, FF3, and FF5, respectively. Across all categories, we find frequent and significant shifts, reinforcing the presence of time variation in anomaly premia.

In summary, our results show that anomalies and risk exposures are cross-sectionally correlated and that anomaly premia vary over time—challenging the core assumptions of the traditional approach and providing the ground for our novel methodology. Moreover, we demonstrate that anomalies affect pricing errors independently of the risk exposures.

[42] For comparison, we also estimate the first-order autocorrelation of the FF5 risk premia, using models augmented with one anomaly at a time. We find that the time variation in risk premia is generally much lower than that observed for anomaly premia. We present these results in OA, see Figure OA.18 and Table OA.XIII.

[43] As robustness checks, we also consider MOSUM and ME tests from Chu, Hornik, and Kaun (1995), which yield similar results. Full details are available upon request.

Figure OA.16 **Canonical Correlations.** Time series of the canonical correlations between the FF5 betas and all 82 anomalies jointly, full sample (Freyberger, Hoeppner, Neuhierl, and Weber (2025) dataset).

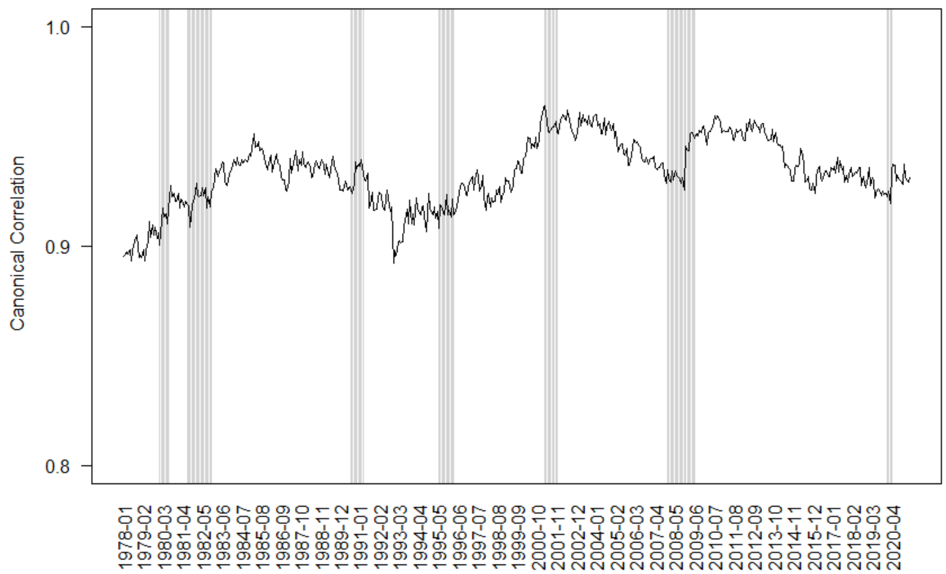


Figure OA.17 **First-order autocorrelation coefficients for each anomaly premium.** The figure shows the first-order autocorrelation coefficients (red dots) for each anomaly premium obtained from a model augmented with the Fama-French five factors (FF5). The figure also includes the corresponding 95% confidence intervals, represented by the gray bands.

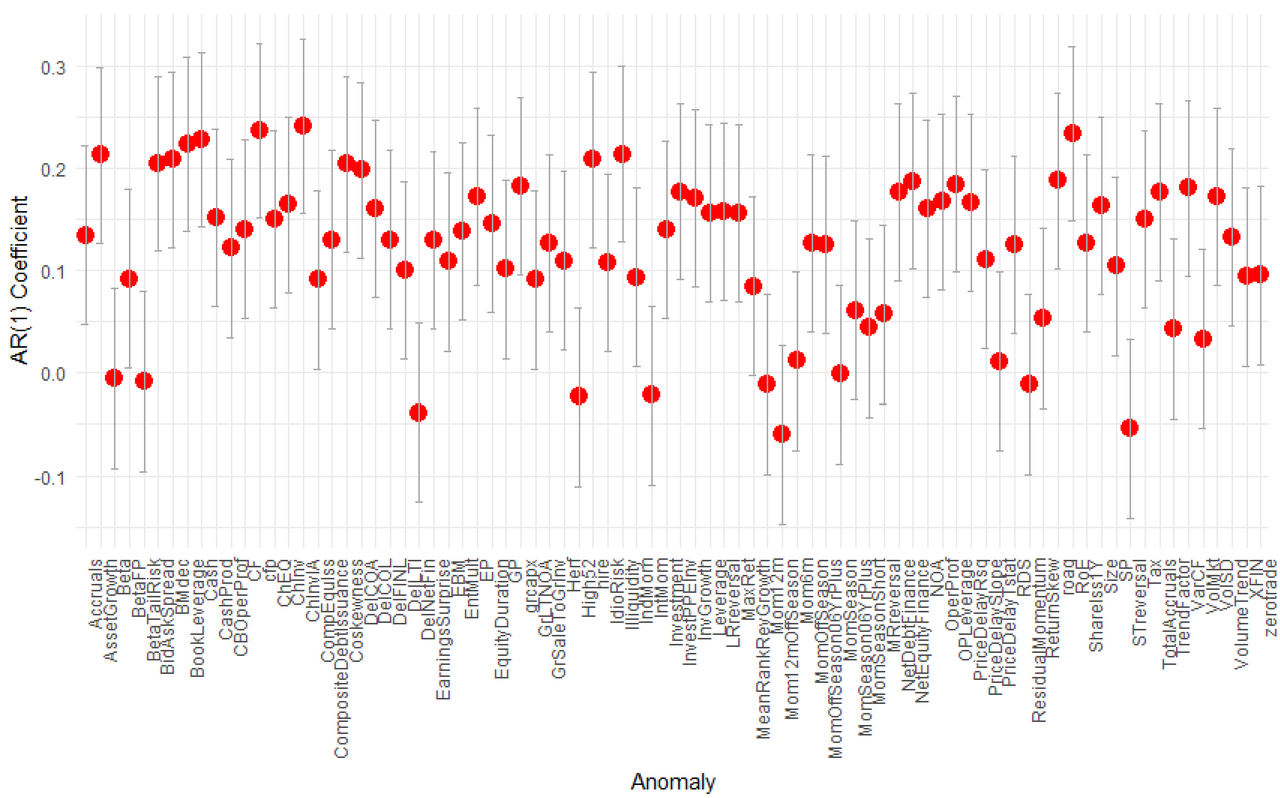


Table OA.VIII: **Correlation between Anomalies and Risk Factors**. The table provides evidence of the correlation between each anomaly and the betas estimated from a given asset pricing factor model. The betas are estimated using a first-pass regression, using the CAPM, the FF3 and FF5 specifications. The first column of the table shows the average R-squared (i.e., the square of the correlation coefficient), together with its minimum and maximum value (in parenthesis), for each of the six categories and each of the three model specifications (CAPM in Panel A, FF3 in Panel B and FF5 in panel C). The second column reports the average percentage of times in which the correlation coefficient is statistically different from zero. To assess whether the correlation coefficient is statistically different from zero, we use the standard Pearson correlation test in the CAPM specification, while we use the p-value of the F-test for the FF3 and FF5 models. We report the results using a 5% confidence level.

Panel A: Average correlation between $\mathbf{Z}_t$ and B (CAPM)

	R-squared (min, max)	% of significance
Momentum	0.059 (0.000, 0.333)	68.29%
Value VS Growth	0.034 (0.000, 0.201)	55.99%
Investment	0.014 (0.000, 0.096)	38.05%
Profitability	0.028 (0.003, 0.120)	48.92%
Intangibles	0.040 (0.001, 0.229)	59.13%
Trading Frictions	0.162 (0.046, 0.339)	79.97%

Panel B: Average correlation between $\mathbf{Z}_t$ and B (FF3)

	R-squared (min, max)	% of significance
Momentum	0.104 (0.001, 0.429)	81.78%
Value vs Growth	0.144 (0.049, 0.366)	95.07%
Investment	0.043 (0.003, 0.154)	64.93%
Profitability	0.125 (0.039, 0.270)	92.50%
Intangibles	0.116 (0.025, 0.330)	83.81%
Trading Frictions	0.339 (0.168, 0.537)	94.94%

Panel C: Average correlation between $\mathbf{Z}_t$ and B (FF5)

	R-squared (min, max)	% of significance
Momentum	0.125 (0.004, 0.453)	85.14%
Value vs Growth	0.167 (0.063, 0.388)	96.26%
Investment	0.062 (0.010, 0.193)	70.03%
Profitability	0.171 (0.073, 0.318)	95.84%
Intangibles	0.146 (0.038, 0.358)	87.04%
Trading Frictions	0.364 (0.193, 0.559)	95.19%

Table OA.IX: **Time-varying anomaly premia: ARMA models versus White Noise**. The table reports the percentage of anomalies (within each category) for which we find a statistically superior fitting ability of a time-series model (among ARMA(p, q) models, with p and q ranging from 1 to 6) over a white noise (i.e., an ARMA$(0, 0)$). Column 1 of the table refers to the case in which we obtain the time series of the $\gamma_{\mathrm{z},t}$ estimates in (21), considering one anomaly at the time in the regression of $\mathbf{R}_t$ on $\mathbf{Z}_t$. Columns 2–4 use instead the time series of the $\gamma_{\mathrm{z},t}$ obtained by regressing the asset returns $\mathbf{R}_t$ on each of the anomaly $\mathbf{Z}_t$ at the time, and including also the betas estimated from the CAPM (Column 2), the FF3 (Column 3) and FF5 (Column 4) in the regression specification, using T=120 months.

Category	$\mathbf{Z}_t$	$\mathbf{Z}_t$+CAPM	$\mathbf{Z}_t$+FF3	$\mathbf{Z}_t$+FF5
Momentum	80.0	93.3	86.7	86.7
Value vs Growth	100.0	100.0	96.6	96.6
Investment	84.2	89.5	84.2	94.7
Profitability	93.3	96.7	93.3	86.7
Intangibles	89.8	95.9	89.8	91.8
Trading Frictions	92.6	100.0	100.0	96.3

Table OA.X: **Time-varying anomaly premia: Structural breaks**. The table reports the average percentage of anomalies (within each category) for which we reject the null hypothesis of no structural breaks in the time series of the premia estimates. For each anomaly, we estimate the regression $\Delta\gamma_{\mathrm{z},t} = \delta_1 + \delta_2\gamma_{\mathrm{z},t} + u_t$ and test for structural changes by analyzing the cumulative sums of OLS-residuals (CUSUM) from the estimated linear model (see Ploberger, Krämer, and Kontrus (1989)). In the first column of the table, the time series of the $\gamma_{\mathrm{z},t}$ is obtained by regressing the asset returns $\mathbf{R}_t$ on each of the anomaly $\mathbf{Z}_t$ at the time. Columns 2–4 use instead the time series of the $\gamma_{\mathrm{z},t}$ obtained by regressing the asset returns $\mathbf{R}_t$ on each of the anomaly $\mathbf{Z}_t$ at the time, but including also the betas estimated from the CAPM (Column 2), the FF3 (Column 3) and FF5 (Column 4) in the regression specification, using T=120 months.

Category	$\mathbf{Z}_t$	$\mathbf{Z}_t$+CAPM	$\mathbf{Z}_t$+FF3	$\mathbf{Z}_t$+FF5
Momentum	40.0	40.0	40.0	33.3
Value vs Growth	13.8	17.2	20.7	20.7
Investment	70.0	63.3	66.7	60.0
Profitability	15.8	26.3	21.1	15.8
Intangibles	20.4	24.5	28.6	28.6
Trading Frictions	22.2	22.2	44.4	48.2

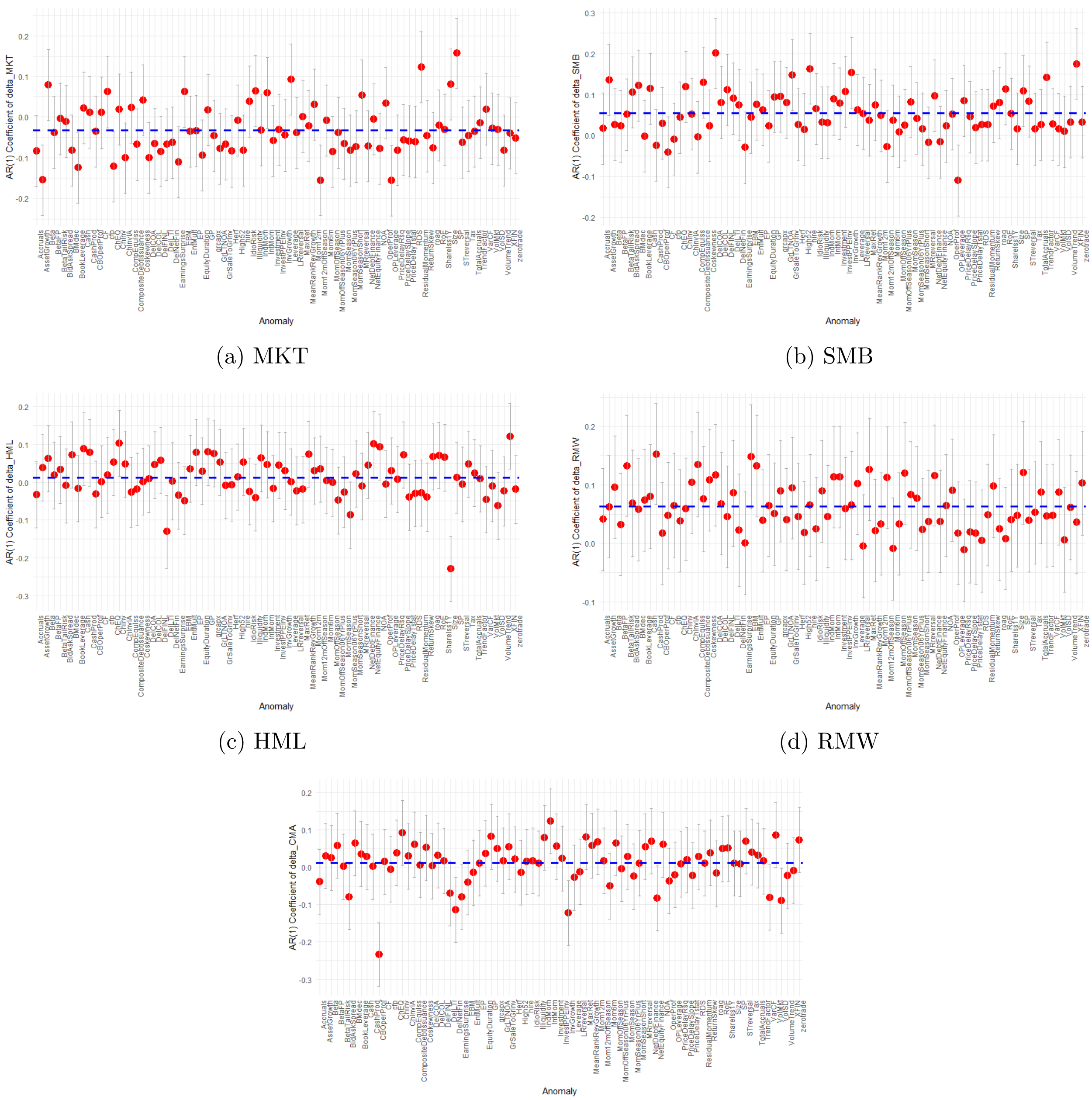
Table OA.XI: **Time-varying anomaly premia: mean differences in anomaly premia over different sub-periods**. The table reports the average percentage of paired sub-periods (aggregated at the category level) for which we find a statistically significant mean difference. Specifically, using the time series of estimated $\gamma_{\mathrm{z},t}$ (one for each anomaly at the time), we consider all the possible pairs of sub-samples (not necessarily consecutive) of length equal to either three years (Panel A) or five years (Panel B), and calculate the difference in the estimates between the corresponding values in each pair. We then calculate the mean difference between each pair of sub-samples and derive the corresponding t-statistic under the null hypothesis that there is no shift in the mean of the two sub-periods (i.e., the mean difference is equal to zero). Column 1 of the table refers to the time series of $\gamma_{\mathrm{z},t}$ which is obtained by regressing the asset returns $\mathbf{R}_t$ on each of the anomaly $\mathbf{Z}_t$ at the time. Columns 2–4 use instead the time series of $\gamma_{\mathrm{z},t}$ obtained by regressing the asset returns $\mathbf{R}_t$ on each of the anomaly $\mathbf{Z}_t$ at the time, but including also the betas estimated from the CAPM (Column 2), the FF3 (Column 3) and FF5 (Column 4) in the regression specification.

Category	$\mathbf{Z}_t$	$\mathbf{Z}_t$+CAPM	$\mathbf{Z}_t$+FF3	$\mathbf{Z}_t$+FF5
Panel A: Paired sub-period of length $T = 36$ (3 years)				
Momentum	15.3	21.1	21.6	23.8
Value vs Growth	21.3	28.8	32.7	35.4
Investment	29.6	34.1	31.2	31.3
Profitability	12.2	21.4	22.9	26.1
Intangibles	14.1	22.4	26.0	28.4
Trading Frictions	10.6	27.3	35.4	34.0
Panel B: Paired sub-period of length $T = 60$ (5 years)				
Momentum	28.0	29.3	22.7	20.9
Value vs Growth	27.1	34.3	30.6	32.9
Investment	43.3	37.6	29.9	31.8
Profitability	29.1	31.6	28.0	31.1
Intangibles	20.0	26.3	25.8	28.2
Trading Frictions	18.6	44.5	43.0	43.6

OA.11.2 Supplementary material using the FHNW Dataset (Freyberger, Hoeppner, Neuhierl, and Weber (2025))

Table OA.XII: List of variables in FHNW Dataset (Freyberger, Hoeppner, Neuhierl, and Weber (2025))

Acronym	**Description**	**Acronym**	**Description**
Accruals	Accruals	AssetGrowth	Asset growth
Beta	CAPM beta	BetaFP	Frazzini-Pedersen Beta
BetaTailRisk	Tail risk beta	BidAskSpread	Bid-ask spread
BMdec	Book to market using December ME	BookLeverage	Book leverage (annual)
Cash	Cash to assets	CashProd	Cash Productivity
CBOperProf	Cash-based operating profitability	CF	Cash flow to market
cfp	Operating Cash flows to price	ChEQ	Growth in book equity
ChInv	Inventory Growth	ChInvIA	Change in capital inv (ind adj)
CompEquIss	Composite equity issuance	CompositeDebtIssuance	Composite debt issuance
Coskewness	Coskewness	DelCOA	Change in current operating assets
DelCOL	Change in current operating liabilities	DelFINL	Change in financial liabilities
DelLTI	Change in long-term investment	DelNetFin	Change in net financial assets
EarningsSurprise	Earnings Surprise	EBM	Enterprise component of BM
EntMult	Enterprise Multiple	EP	Earnings-to-Price Ratio
EquityDuration	Equity Duration	GP	Gross profits / total assets
grcapx	Change in capex (2 years)	GrLTNOA	Growth in long-term operating assets
GrSaleToGrInv	Sales growth over inventory growth	Herf	Industry concentration (sales)
High52	52 week high	hire	Employment growth
IdioRisk	Idiosyncratic risk	Illiquidity	Amihud's illiquidity
IndMom	Industry Momentum	IntMom	Intermediate Momentum
Investment	Investment to revenue	InvestPPEInv	Change in ppe and inv/assets
InvGrowth	Inventory Growth	Leverage	Market leverage
LRreversal	Long-run reversal		
MaxRet	Maximum return over month	MeanRankRevGrowth	Revenue Growth Rank
Mom12m	Momentum (12 month)	Mom12mOffSeason	Momentum without the seasonal part
Mom6m	Momentum (6 month)	MomOffSeason	Off season long-term reversal
MomOffSeason06YrPlus	Off season reversal years 6 to 10	MomSeason	Return seasonality years 2 to 5
MomSeason06YrPlus	Return seasonality years 6 to 10	MomSeasonShort	Return seasonality last year
MRreversal	Medium-run reversal	NetDebtFinance	Net debt financing
NetEquityFinance	Net equity financing	NOA	Net Operating Assets
OperProf	Operating profits / book equity	OPleverage	Operating leverage
PriceDelayRsq	Price delay r square	PriceDelaySlope	Price delay coef
PriceDelayTstat	Price delay SE adjusted	RDS	Real dirty surplus
ResidualMomentum	Momentum based on FF3 residuals	ReturnSkew	Return skewness
roaq	Return on assets (qtrly)	RoE	Net income / book equity
ShareIss1Y	Share issuance (1 year)	Size	Size
SP	Sales-to-price	STreversal	Short-term reversal
Tax	Taxable income to income	TotalAccruals	Total accruals
TrendFactor	Trend Factor	VarCF	Cash-flow to price variance
VolMkt	Volume to market equity	VolSD	Volume Variance
VolumeTrend	Volume Trend	XFIN	Net external financing
zerotrade	Days with zero trades		



(a) MKT

(b) SMB

(c) HML

(d) RMW

(e) CMA

Figure OA.18Estimated first-order autocorrelation coefficients for the five Fama-French risk premia with single-anomaly augmentations. Each point corresponds to the AR(1) coefficient obtained by estimating an FF5 model augmented with one anomaly at a time; the horizontal axis indexes the different anomalies used in each model specification. Data are from Freyberger, Hoeppner, Neuhierl, and Weber (2025). Panel differences: Panel (a) reports results for the market (MKT) risk premium; Panel (b) for size (SMB); Panel (c) for value (HML); Panel (d) for profitability (RMW); and Panel (e) for investment (CMA). All panels are constructed using the same estimation procedure and sample to make cross-panel comparisons directly interpretable.

Figure OA.19**Anomalies' contribution using time-varying multivariate OLS regressions - FHNW Dataset**. The figure shows the time series of the total variance decomposition obtained in each multivariate regression (asset returns on the market factor and the 82 firms characteristics from Freyberger, Hoeppner, Neuhierl, and Weber (2025), $T = 24$).

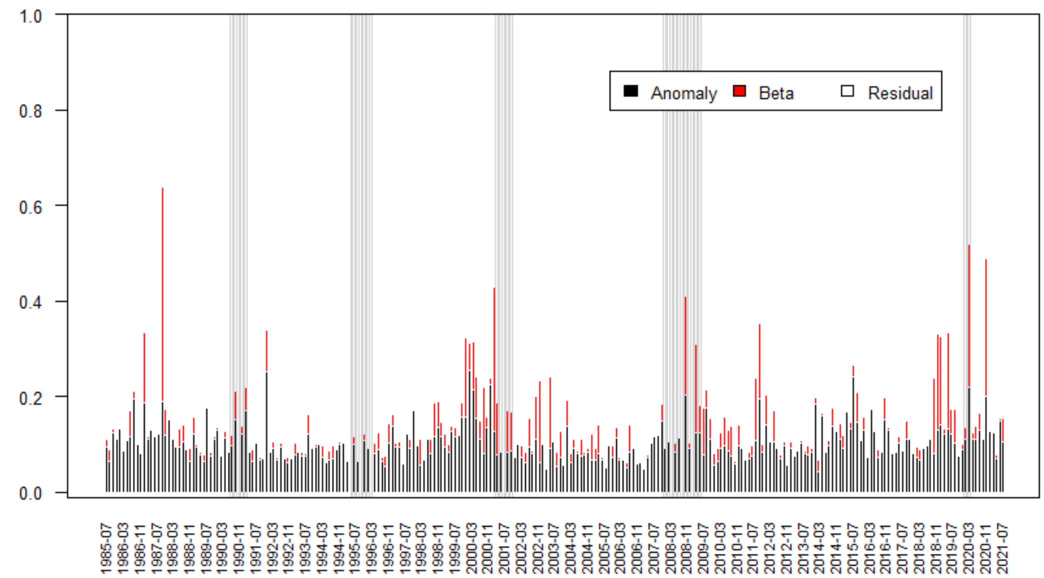


Figure OA.20**Anomalies' Contribution Using PCA risk factors - FHNW Dataset.** The figure displays the time series of the local R^2 decomposition obtained from multivariate regressions of asset returns on the full set of 82 firm characteristics and one Principal Component. Each regression is estimated using a rolling window of $T = 120$ months over balanced panels from January 1978 to August 2021, following Freyberger, Hoeppner, Neuhierl, and Weber (2025). The bars show the contributions of anomalies (black), PC betas (red), and residuals (white) to the explained cross-sectional variance.

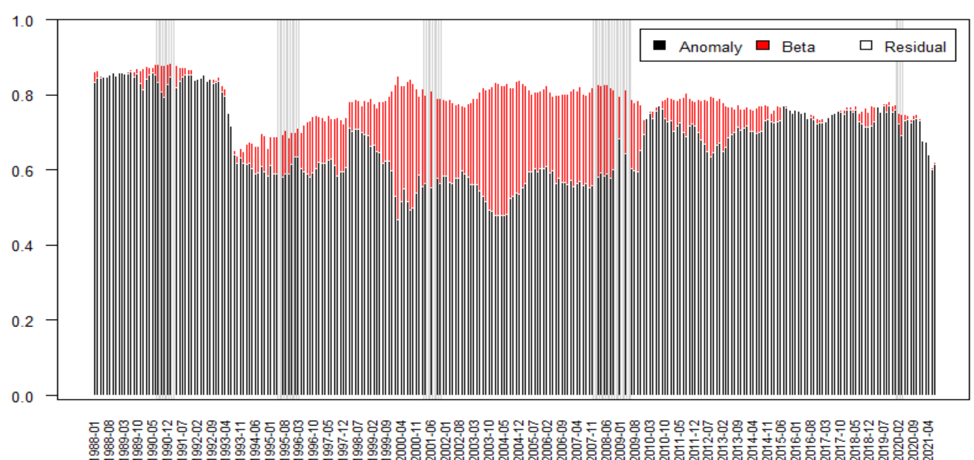
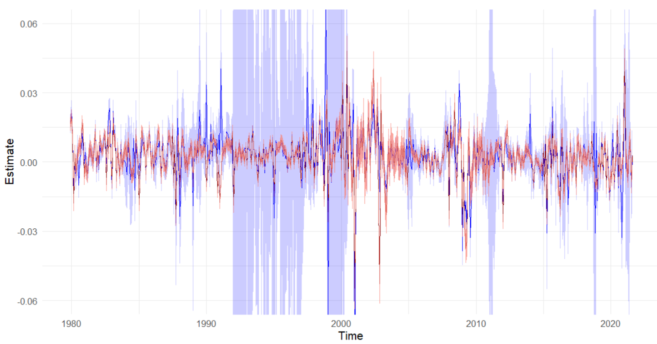
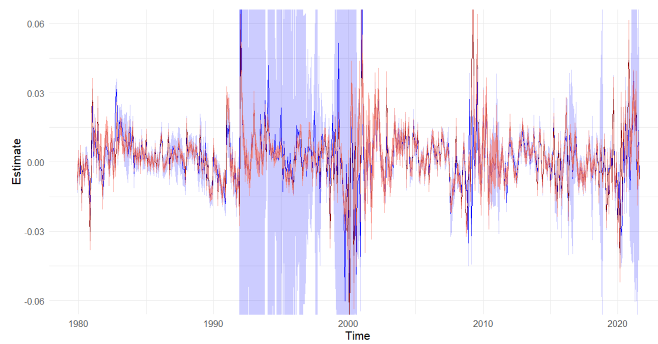


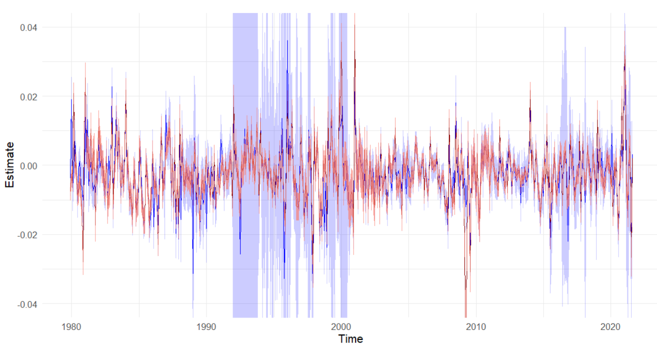
Figure OA.21**Time series of six estimated anomaly premia.** The figure shows the time series of estimated anomaly premia with 95% confidence intervals for six representative anomalies—one from each category: (a) ResidualMomentum, (b) SP, (c) CompEquISS, (d) OperProf, (e) Leverage, and (f) BidAskSpread. The blue line shows the estimates obtained with our novel methodology of Section 4, and the red line represents the time-varying conventional estimates based on the conventional estimator (21). The shaded areas indicate the respective confidence intervals (blue area for our methodology, red area for the conventional approach).



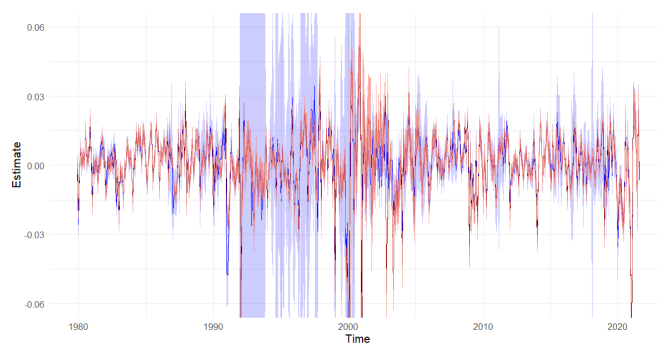
(a) ResidualMomentum



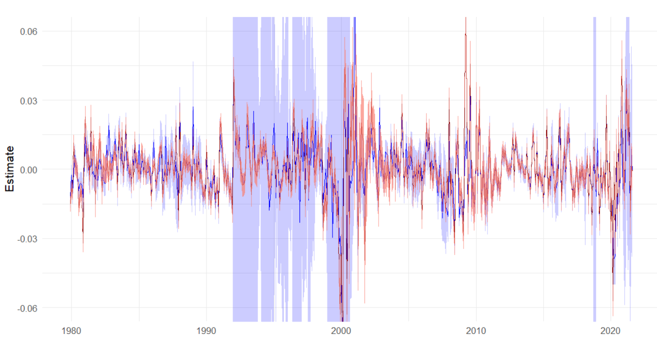
(b) SP



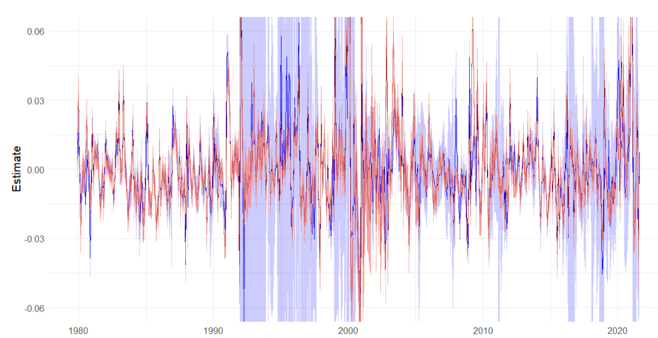
(c) CompEquISS



(d) OperProf



(e) Leverage



(f) BidAskSpread

Table OA.XIII: Summary Statistics of the series of AR(1) coefficients for the five risk premia (δ_{MKT}, δ_{SMB}, δ_{HML}, δ_{RMW}, δ_{CMA}). For each series, the table reports the minimum, first quartile, median, mean, third quartile, and maximum of the estimated AR(1) coefficients, along with the percentage of coefficients found to be not statistically different from zero.

Statistic	MKT	SMB	HML	RMW	CMA
Min.	-0.156	-0.110	-0.228	-0.012	-0.233
1st Qu.	-0.075	0.022	-0.019	0.036	-0.014
Median	-0.038	0.047	0.010	0.059	0.018
Mean	-0.032	0.055	0.013	0.063	0.011
3rd Qu.	0.008	0.087	0.052	0.091	0.051
Max.	0.157	0.201	0.121	0.153	0.123
% not significant	85.4	73.2	92.7	72.0	92.7

Table OA.XIV: **The impact of financial crises on the statistical significance of anomalies (FHNW Dataset).** This table reports fixed-effect panel regressions parallel to Table III, now controlling for FF3 and FF5 betas. Outcomes are: $d_{|t|}$, $d_{|t|\text{Bonferroni}}$, $d_{|t|_{\text{Holm}}}$, and $|t|$. Coefficients on d_{crisis} are reported with t-ratios in parentheses. Additional controls include $fund\_rate$, $fund\_shadow$ (Wu and Xia (2016)), and mp_{shock} (Bu, Rogers, and Wu (2021)).

Panel A: FF3

	$\mathbf{d\_\|t\|}$		$\mathbf{d\_\|t\|}_{\text{Bonferroni}}$		$\mathbf{d\_\|t\|}_{\text{Holm}}$		$\mathbf{\|t\|}$	
d_crisis	**0.153**	(-28.64)	**0.064**	(23.54)	**0.068**	(24.69)	**0.608**	(36.34)
d_crisis	**0.152**	(28.51)	**0.063**	(23.43)	**0.068**	(24.57)	**0.602**	(36.21)
fund_rate	**-0.006**	(-10.54)	**-0.002**	(-7.72)	**-0.002**	(-8.15)	**-0.034**	(-19.23)
d_crisis	**0.172**	(33.44)	**0.067**	(25.97)	**0.073**	(27.09)	**0.693**	(42.52)
fund_shadow	**-0.014**	(-24.14)	**-0.047**	(-12.93)	**-0.004**	(-13.19)	**-0.069**	(-37.76)
d_crisis	**0.200**	(32.92)	**0.082**	(26.31)	**0.088**	(27.46)	**0.762**	(39.78)
mp_shock	**-0.409**	(-8.38)	**-0.272**	(-10.82)	**-0.309**	(-12.01)	**-1.569**	(-10.21)
d_crisis	**0.154**	(28.93)	**0.064**	(23.64)	**0.068**	(24.79)	**0.613**	(37.2)
fund_rate	**-0.031**	(-16.65)	**-0.009**	(-10.05)	**-0.010**	(-10.3)	**-0.177**	(-30.93)
fund_rate^2	**0.003**	(14.11)	**0.001**	(8.08)	**0.001**	(8.21)	**0.019**	(26.28)
d_crisis	**0.175**	(33.8)	**0.069**	(25.91)	**0.073**	(26.99)	**0.707**	(43.3)
fund_shadow	**-0.020**	(-16.58)	**-0.004**	(-6.71)	**-0.004**	(-6.48)	**-0.102**	(-27.33)

Panel B: FF5

	$\mathbf{d\_\|t\|}$		$\mathbf{d\_\|t\|}_{\text{Bonferroni}}$		$\mathbf{d\_\|t\|}_{\text{Holm}}$		$\mathbf{\|t\|}$	
d_crisis	**0.088**	(18.38)	**0.024**	(9.92)	**0.025**	(10.02)	**0.408**	(25.32)

Continued on next page

Table OA.XIV continued from previous page

d_crisis	**0.087**	(18.19)	**0.024**	(9.78)	**0.024**	(9.87)	**0.403**	(25.13)
fund_rate	**-0.007**	(-14.43)	**-0.002**	(-8.96)	**-0.002**	(-9.58)	**-0.031**	(-18.3)
d_crisis	**0.096**	(19.85)	**0.026**	(10.4)	**0.026**	(10.49)	**0.465**	(28.74)
fund_shadow	**-0.012**	(-22.81)	**-0.003**	(-11.95)	**-0.004**	(-12.46)	**-0.065**	(-35.83)
d_crisis	**0.075**	(13.62)	**0.022**	(7.77)	**0.023**	(7.87)	**0.366**	(19.5)
mp_shock	**-0.234**	(-5.29)	**-0.080**	(-3.50)	**-0.088**	(-3.77)	**-1.109**	(-7.36)
d_crisis	**0.090**	(19.11)	**0.024**	(10.16)	**0.025**	(10.25)	**0.422**	(27.2)
fund_rate	**-0.052**	(-31.9)	**-0.013**	(-16.12)	**-0.014**	(-16.56)	**-0.285**	(-52.92)
fund_rate^2	**0.006**	(28.83)	**0.001**	(14.06)	**0.002**	(14.32)	**0.033**	(49.51)
d_crisis	**0.107**	(22.26)	**0.028**	(11.48)	**0.029**	(11.56)	**0.519**	(32.8)
fund_shadow	**-0.037**	(-34.06)	**-0.010**	(-16.8)	**-0.010**	(-17.03)	**-0.190**	(-52.66)

Table OA.XV: **Anomalies' significance: OLS vs WLS in the FHNW Dataset.** The table presents the proportion of significant anomalies as identified by our methodology, comparing outcomes from OLS (Columns 1–3) and WLS (Columns 4–6) approaches, and using three different model specifications (Capm, FF3, and FF5). Columns 7–9 and 10–12 report the absolute and relative changes between the OLS and WLS results, respectively. The different panels report the breakdown for each specific category.

capital letters

	OLS (%)			WLS (%)			WLS-OLS (%)			WLS/OLS-1 (%)		
Category	Capm	FF3	FF5	Capm	FF3	FF5	Capm	FF3	FF5	Capm	FF3	FF5
Panel A: Momentum												
EarningsSurprise	16.8	7.1	8.7	10.2	3.9	6.7	-6.6	-3.2	-2.1	-39.3	-45.2	-23.7
IndMom	28.5	12.6	7.8	23.0	10.8	6.2	-5.6	-1.8	-1.6	-19.6	-14.5	-20.6
IntMom	20.6	12.4	7.1	14.8	8.0	4.4	-5.8	-4.4	-2.8	-28.2	-35.2	-38.7
Mom12m	28.9	17.2	10.8	23.2	11.5	9.0	-5.8	-5.7	-1.8	-20.0	-33.3	-17.0
Mom6m	25.9	12.4	8.5	19.6	9.4	7.1	-6.4	-3.0	-1.4	-24.6	-24.1	-16.2
ResidualMomentum	29.7	19.3	8.7	23.2	15.6	7.6	-6.6	-3.7	-1.1	-22.1	-19.0	-13.2
TrendFactor	21.6	10.6	8.5	14.8	9.2	7.6	-6.8	-1.4	-0.9	-31.5	-13.0	-10.8
High52	27.3	17.2	12.0	19.8	10.6	10.3	-7.6	-6.7	-1.6	-27.7	-38.7	-13.5
Panel B: Value versus Growth												
CF	28.9	9.9	8.3	15.0	4.4	5.3	-14.0	-5.5	-3.0	-48.3	-55.8	-36.1
EBM	25.0	13.6	13.1	16.8	7.8	10.1	-8.2	-5.7	-3.0	-32.8	-42.4	-22.8
EP	29.7	12.2	8.5	16.0	4.4	4.6	-13.8	-7.8	-3.9	-46.3	-64.2	-45.9
EntMult	31.3	9.9	9.2	17.0	6.9	7.1	-14.4	-3.0	-2.1	-45.9	-30.2	-22.5
EquityDuration	20.8	7.1	7.1	12.4	2.5	2.5	-8.4	-4.6	-4.6	-40.4	-64.5	-64.5
LRreversal	29.7	17.2	6.7	18.8	8.3	4.6	-11.0	-9.0	-2.1	-36.9	-52.0	-31.0
MeanRankRevGrowth	19.8	7.8	6.7	13.2	4.1	1.8	-6.6	-3.7	-4.8	-33.3	-47.1	-72.4
MRreversal	19.8	10.6	7.1	14.2	3.9	2.8	-5.6	-6.7	-4.4	-28.3	-63.0	-61.3
RDS	24.8	8.0	2.5	15.8	4.1	0.7	-9.0	-3.9	-1.8	-36.3	-48.6	-72.7
SP	35.7	12.6	9.2	28.1	5.7	2.8	-7.6	-6.9	-6.4	-21.2	-54.5	-70.0
cfp	31.7	10.3	10.1	20.2	5.1	4.4	-11.6	-5.3	-5.7	-36.5	-51.1	-56.8
BMdec	20.0	6.2	7.6	16.0	3.0	4.6	-4.0	-3.2	-3.0	-20.0	-51.9	-39.4

Continued on next page

Table OA.XV – continued from previous page

Category	OLS (%) Capm	FF3	FF5	WLS (%) Capm	FF3	FF5	WLS-OLS (%) Capm	FF3	FF5	WLS/OLS-1 (%) Capm	FF3	FF5
STreversal	22.4	10.1	5.7	13.0	8.3	6.0	-9.4	-1.8	0.2	-42.0	-18.2	4.0
Panel C: Investment												
Accruals	18.6	10.3	11.0	14.0	5.5	6.2	-4.6	-4.8	-4.8	-24.7	-46.7	-43.7
ChInvIA	15.2	8.7	12.6	7.0	3.2	7.8	-8.2	-5.5	-4.8	-53.9	-63.2	-38.2
CompEquIss	31.7	15.9	12.0	21.4	10.6	6.2	-10.4	-5.3	-5.7	-32.7	-33.3	-48.1
CompositeDebtIssuance	17.4	7.6	5.7	10.0	3.7	3.4	-7.4	-3.9	-2.3	-42.5	-51.5	-40.0
DelCOL	15.2	8.0	9.4	8.2	3.0	4.6	-7.0	-5.1	-4.8	-46.1	-62.9	-51.2
DelFINL	16.2	5.7	7.1	7.2	1.8	3.9	-9.0	-3.9	-3.2	-55.6	-68.0	-45.2
DelNetFin	15.0	6.2	9.9	6.8	2.5	6.4	-8.2	-3.7	-3.4	-54.7	-59.3	-34.9
grcapx	20.4	9.0	8.3	10.2	4.1	3.7	-10.2	-4.8	-4.6	-50.0	-53.8	-55.6
GrLTNOA	17.6	9.0	9.7	12.4	6.7	5.3	-5.2	-2.3	-4.4	-29.5	-25.6	-45.2
InvGrowth	18.2	10.3	8.3	10.2	4.1	3.9	-8.0	-6.2	-4.4	-44.0	-60.0	-52.8
Investment	16.0	7.8	6.9	6.0	3.7	2.8	-10.0	-4.1	-4.1	-62.5	-52.9	-60.0
InvestPPEInv	23.0	11.0	5.3	16.4	6.2	3.2	-6.6	-4.8	-2.1	-28.7	-43.7	-39.1
NOA	31.7	14.5	7.4	23.4	9.9	5.1	-8.4	-4.6	-2.3	-26.4	-31.7	-31.3
NetDebtFinance	15.0	6.7	7.8	7.0	2.8	5.3	-8.0	-3.9	-2.5	-53.3	-58.6	-32.4
NetEquityFinance	21.2	11.7	9.9	13.8	6.4	8.0	-7.4	-5.3	-1.8	-34.9	-45.1	-18.6
ShareIss1Y	18.2	8.0	7.4	13.2	4.8	5.5	-5.0	-3.2	-1.8	-27.5	-40.0	-25.0
TotalAccruals	21.8	13.8	8.7	11.0	5.5	5.5	-10.8	-8.3	-3.2	-49.5	-60.0	-36.8
XFIN	20.0	8.7	7.4	7.0	3.9	4.1	-13.0	-4.8	-3.2	-65.0	-55.3	-43.8
AssetGrowth	23.8	14.7	6.7	13.6	6.2	3.9	-10.2	-8.5	-2.8	-42.9	-57.8	-41.4
ChInv	17.0	10.6	7.6	9.0	3.2	5.1	-8.0	-7.4	-2.5	-47.1	-69.6	-33.3
DelCOA	17.6	14.0	7.6	8.4	4.1	5.1	-9.2	-9.9	-2.5	-52.3	-70.5	-33.3
DelLTI	9.2	5.3	5.7	5.2	3.2	2.8	-4.0	-2.1	-3.0	-43.5	-39.1	-52.0
Panel D: Profitability												
CBOperProf	36.5	12.6	11.3	24.6	2.8	6.7	-12.0	-9.9	-4.6	-32.8	-78.2	-40.8
CashProd	25.9	8.3	7.8	19.4	4.4	5.3	-6.6	-3.9	-2.5	-25.4	-47.2	-32.4
ChEQ	22.4	13.8	8.7	11.4	7.4	6.4	-11.0	-6.4	-2.3	-49.1	-46.7	-26.3
GP	35.5	11.5	9.2	24.8	4.8	4.6	-10.8	-6.7	-4.6	-30.3	-58.0	-50.0
OperProf	36.9	12.6	7.4	22.0	5.7	2.8	-15.0	-6.9	-4.6	-40.5	-54.5	-62.5
RoE	30.5	13.1	6.4	18.6	6.2	4.1	-12.0	-6.9	-2.3	-39.2	-52.6	-35.7
Tax	29.7	13.1	8.3	18.6	4.8	5.7	-11.2	-8.3	-2.5	-37.6	-63.2	-30.6
VarCF	24.2	10.3	7.1	13.8	2.3	2.5	-10.4	-8.0	-4.6	-43.0	-77.8	-64.5
VolumeTrend	20.0	10.6	8.7	11.4	5.1	5.1	-8.6	-5.5	-3.7	-43.0	-52.2	-42.1
roaq	38.3	26.0	12.4	25.5	11.0	9.9	-12.8	-14.9	-2.5	-33.3	-57.5	-20.4
BookLeverage	28.5	3.7	3.7	21.0	1.1	2.1	-7.6	-2.5	-1.6	-26.6	-68.8	-43.7
Panel E: Intangibles												
Cash	19.0	7.1	5.3	13.6	5.1	3.7	-5.4	-2.1	-1.6	-28.4	-29.0	-30.4
GrSaleToGrInv	15.0	7.4	8.5	8.0	4.1	7.8	-7.0	-3.2	-0.7	-46.7	-43.8	-8.1
Herf	22.0	14.7	7.4	12.0	8.0	4.8	-10.0	-6.7	-2.5	-45.5	-45.3	-34.4
hire	20.2	10.8	7.4	11.2	4.8	4.4	-9.0	-6.0	-3.0	-44.6	-55.3	-40.6
Leverage	32.5	5.7	6.7	24.2	2.1	4.1	-8.4	-3.7	-2.5	-25.8	-64.0	-37.9
Mom12mOffSeason	28.3	15.6	11.0	20.2	11.3	9.0	-8.2	-4.4	-2.1	-28.9	-27.9	-18.8
MomOffSeason	24.0	11.5	4.6	17.6	4.1	2.5	-6.4	-7.4	-2.1	-26.7	-64.0	-45.0

Continued on next page

Table OA.XV – continued from previous page

Category	OLS (%) Capm	FF3	FF5	WLS (%) Capm	FF3	FF5	WLS-OLS (%) Capm	FF3	FF5	WLS/OLS-1 (%) Capm	FF3	FF5
MomOffSeason06YrPlus	19.0	6.2	5.7	11.6	2.5	1.1	-7.4	-3.7	-4.6	-38.9	-59.3	-80.0
MomSeason	20.4	6.2	4.8	9.0	2.8	2.5	-11.4	-3.4	-2.3	-55.9	-55.6	-47.6
MomSeason06YrPlus	16.8	6.7	6.9	7.8	3.9	3.9	-9.0	-2.8	-3.0	-53.6	-41.4	-43.3
MomSeasonShort	16.8	9.2	5.1	12.0	6.7	3.7	-4.8	-2.5	-1.4	-28.6	-27.5	-27.3
OPLeverage	25.3	7.1	4.8	19.4	6.2	1.8	-6.0	-0.9	-3.0	-23.6	-12.9	-61.9
PriceDelayRsq	24.2	12.4	5.3	14.2	3.7	3.9	-10.0	-8.7	-1.4	-41.3	-70.4	-26.1
PriceDelaySlope	22.8	13.8	7.4	18.0	4.8	4.1	-4.8	-9.0	-3.2	-21.1	-65.0	-43.8
PriceDelayTstat	5.8	4.4	7.8	3.4	3.7	3.7	-2.4	-0.7	-4.1	-41.4	-15.8	-52.9
Panel F: Trading Frictions												
BetaFP	3.2	3.4	4.4	3.0	1.1	2.8	-0.2	-2.3	-1.6	-6.3	-66.7	-36.8
BetaTailRisk	8.6	2.8	2.8	6.4	1.4	0.9	-2.2	-1.4	-1.8	-25.6	-50.0	-66.7
BidAskSpread	31.7	11.3	7.8	14.4	3.0	5.5	-17.4	-8.3	-2.3	-54.7	-73.5	-29.4
Illiquidity	41.7	9.2	6.4	26.5	0.7	1.6	-15.2	-8.5	-4.8	-36.4	-92.5	-75.0
ReturnSkew	13.6	6.9	9.0	6.6	3.4	6.0	-7.0	-3.4	-3.0	-51.5	-50.0	-33.3
VolMkt	10.2	5.7	6.9	6.2	0.9	6.0	-4.0	-4.8	-0.9	-39.2	-84.0	-13.3
VolSD	16.6	2.5	3.2	11.6	0.5	1.4	-5.0	-2.1	-1.8	-30.1	-81.8	-57.1
zerotrade	12.0	6.0	7.1	6.6	1.1	4.4	-5.4	-4.8	-2.8	-45.0	-80.8	-38.7
Beta	2.8	0.5	2.8	1.4	0.2	0.2	-1.4	-0.2	-2.5	-50.0	-50.0	-91.7
Coskewness	21.6	4.8	3.7	15.2	2.3	2.3	-6.4	-2.5	-1.4	-29.6	-52.4	-37.5
IdioRisk	22.2	12.6	9.4	9.6	3.0	5.1	-12.6	-9.7	-4.4	-56.8	-76.4	-46.3
MaxRet	18.0	10.1	9.7	6.4	2.8	4.1	-11.6	-7.4	-5.5	-64.4	-72.7	-57.1

Table OA.XVI: **WLS estimation: the impact of financial crises on the statistical significance of anomalies - FHNW Dataset.** The table reports fixed-effect panel regression results designed to explore the impact of financial crises on the emergence and significance of anomalies. In the table, $d\_|t|_{it}$ is a dummy variable (one if the anomaly i in month t is priced - according to the t-statistic derived using our WLS estimation of Section 4.1 - and zero otherwise). The variable $|t|_{it}$ quantifies the absolute t-statistic value of the premium for each anomaly i at time t. As the weighting matrix $\mathbf{W_{t-1}}$, we use the ranked variable *size* measured as market equity. We tackle multiple testing distortions through Bonferroni and Holm methods ($d_{|t|\text{Bonferroni}}$ and $d|t|_{\text{Holm}}$) to address potential type I errors. Panel A focuses on results derived from univariate regressions controlling for market factor betas under the CAPM framework. Panels B and C extend this analysis by incorporating the betas for the FF3 and FF5 models, respectively. For each dependent variable, the table reports the coefficient associated with the crisis dummy variable (d_{crisis}), indicating whether a period is classified as an NBER-defined crisis period. The coefficients are accompanied by their respective t-ratios in parentheses.

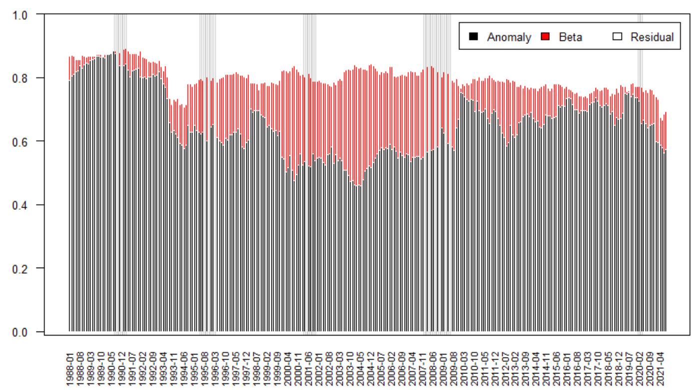
	$\mathbf{d\_|t|}$		$\mathbf{d\_|t|_{\text{Bonferroni}}}$		$\mathbf{d\_|t|_{\text{Holm}}}$		$\mathbf{|t|}$	
Panel A: CAPM								
d_crisis	**0.190**	(38.32)	**0.099**	(33.65)	**0.102**	(34.24)	**0.805**	(50.02)
Panel B: FF3								

Continued on next page

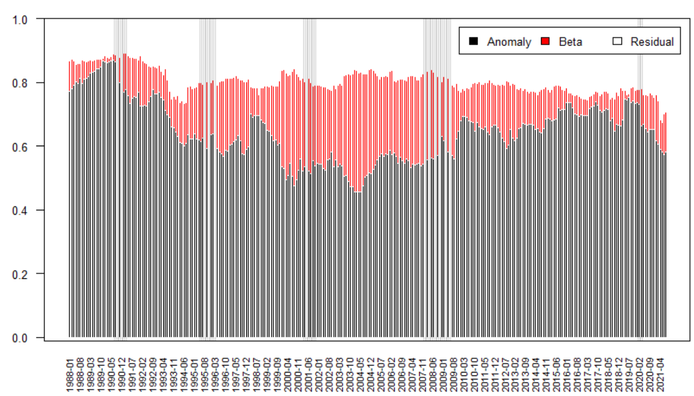
Table OA.XVI continued from previous page

d_crisis	**0.126**	(32.61)	**0.042**	(24.97)	**0.042**	(24.91)	**0.531**	(43.38)
Panel C: FF5								
d_crisis	**0.097**	(25.41)	**0.022**	(13.57)	**0.023**	(13.89)	**0.487**	(40.42)

Figure OA.22**Anomalies' contribution using local-average approach - FHNW Dataset**. The figure shows the time series of the total variance decomposition obtained in each multivariate regression using rolling window of $T = 120$ months. The results are obtained by performing multivariate regressions of asset returns on all the 82 firms characteristics together with (a) the three Fama-French factors betas, and (b) the five Fama-French factors betas. The application uses balanced panels for each rolling sample, with a reference period ranging from January 1978 to August 2021 (from Freyberger, Hoeppner, Neuhierl, and Weber (2025)).

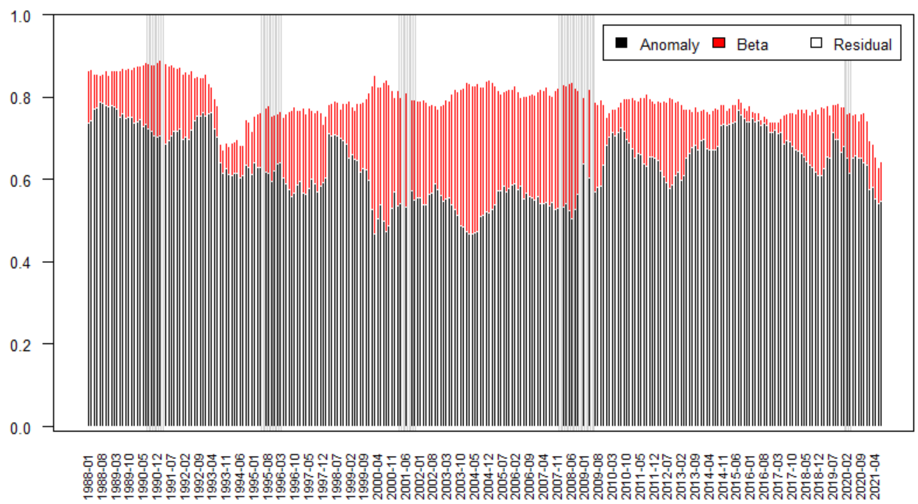


(a) FF3 Model

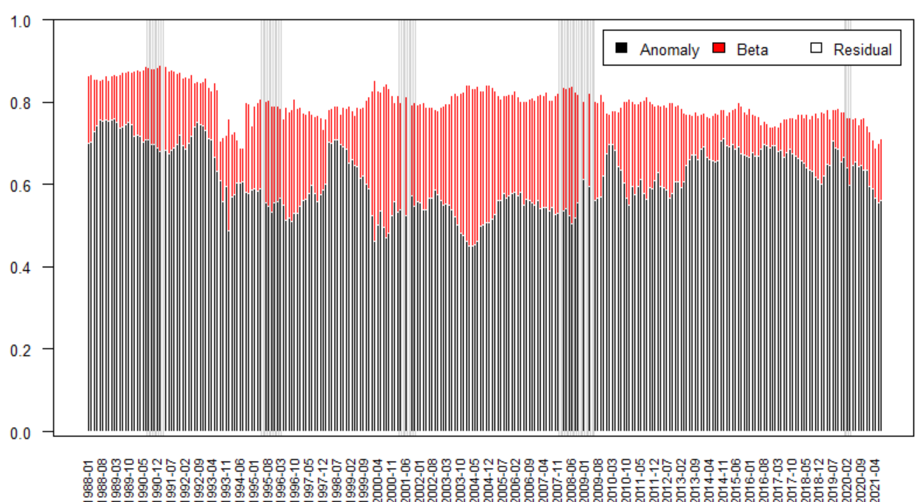


(b) FF5 Model

Figure OA.23**Anomalies' Contribution Using PCA risk factors - FHNW Dataset.** The figure displays the time series of the local R^2 decomposition obtained from multivariate regressions of asset returns on the full set of 82 firm characteristics and different numbers of principal components: (a) three (PC3), and (b) five (PC5). Each regression is estimated using a rolling window of $T = 120$ months over balanced panels from January 1978 to August 2021, following Freyberger, Hoeppner, Neuhierl, and Weber (2025). The bars show the contributions of anomalies (black), PC betas (red), and residuals (white) to the explained cross-sectional variance.



(a) PC3



(a) PC5

Figure OA.24**Heatmap of the t-statistics using the conventional time-varying estimator.** The figure shows the heatmap of the t-statistics distribution obtained in each of the 82 univariate models (vertical axis) and for each month (horizontal axis). Each cell in the map represents the degree of statistical significance of the t-statistics with a different colour, from grey (non-significant t-stat), to yellow (significance at only 10% level), orange (significance at 5% level), and red (significance at 1% level). The results are obtained by performing univariate regressions of asset returns on the market factor and each of the 82 anomalies, using the time-varying estimator defined in equation (21), where the standard errors are computed as $\sqrt{\tilde{\Sigma}_{\gamma_z}/(T-1)}$ within each window. The data is from from Freyberger, Hoeppner, Neuhierl, and Weber (2025). The analysis uses balanced panels for each month, with a reference period ranging from January 1978 to August 2021. At each month t, the market beta is obtained by running a first-pass regression using a rolling window on the past two years of data ($T = 24$).

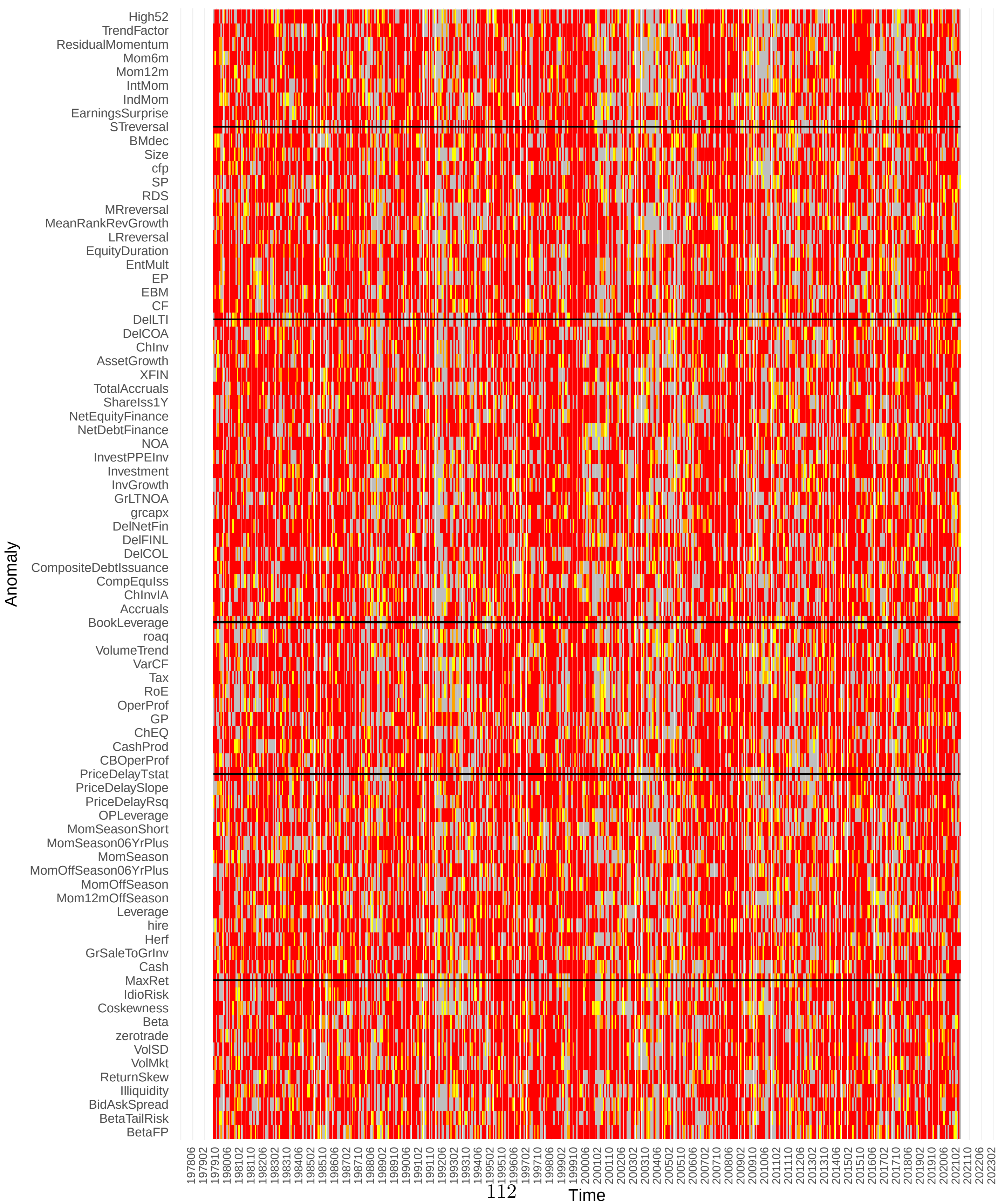


Figure OA.25**Heatmap of the t-statistics using the conventional average estimator**. The figure shows the heatmap of the t-statistics distribution obtained in each of the 82 univariate models (vertical axis) and for each month (horizontal axis). Each cell in the map represents the degree of statistical significance of the t-statistics with a different colour, from grey (non-significant t-stat), to yellow (significance at only 10% level), orange (significance at 5% level), and red (significance at 1% level). The results are obtained by performing univariate regressions of asset returns on the market factor and each of the 82 anomalies, using the average estimator defined in equation (22), with the t-ratio computed as in (23) within each window. The data is from from Freyberger, Hoeppner, Neuhierl, and Weber (2025). The analysis uses balanced panels for each month, with a reference period ranging from January 1978 to August 2021. At each month t, the market beta is obtained by running a first-pass regression using a rolling window on the past two years of data ($T = 24$).

OA.11.3 Supplementary material using the BLLP Dataset (Bryzgalova, Lerner, Lettau, and Pelger (2025))

In this section, we wish to replicate our previous analysis using the alternative dataset of Bryzgalova, Lerner, Lettau, and Pelger (2025), hereafter referred to as BLLP Dataset, to validate the robustness of our initial findings. The coplete list of variables is reported below in Table OA.XVII. By employing a distinct set of firm characteristics, which are subject to different data transformations and imputed using different methodologies for handling missing values, we intend to assess whether the conclusions drawn in the previous sections still hold.

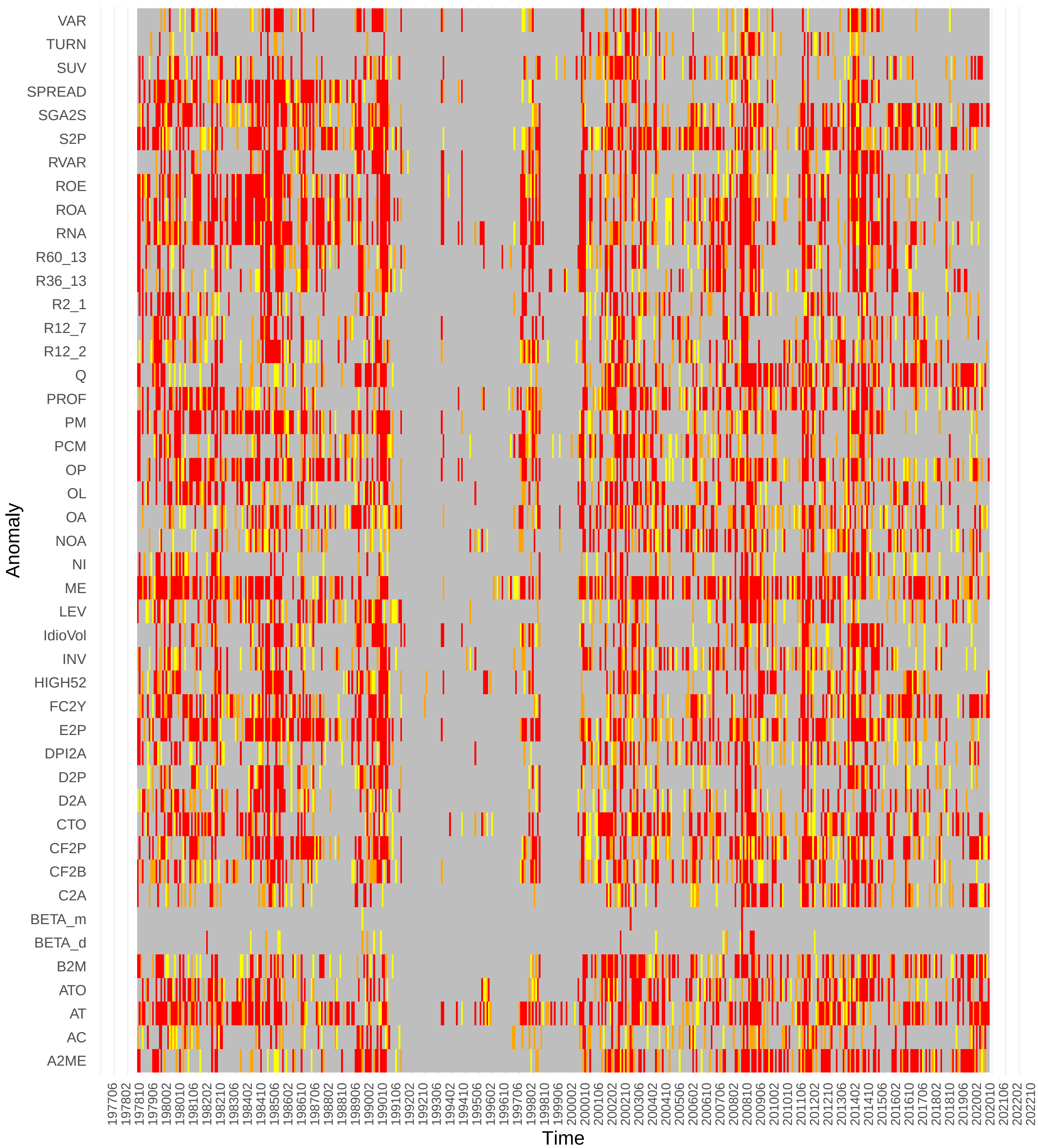
Table OA.XVII: List of variables in BLLP Dataset (Bryzgalova, Lerner, Lettau, and Pelger (2025))

Acronym	**Description**	**Acronym**	**Description**
A2ME	Assets to market cap	AC	Accrual
AT	Total Assets	ATO	Net sales over lagged net operating assets
B2M	Book to Market Ratio	Beta_d	CAPM Beta
Beta_m	Market Beta	C2A	Ratio of cash and short-term investments to total assets
CF2B	Free Cash Flow to Book Value	CF2P	Cashflow to price
CTO	Capital turnover	D2A	Capital intensity
D2P	Dividend Yield	DPI2A	Change in property, plants, and equipment
E2P	Earnings to price	FC2Y	Fixed costs to sales
HIGH52	Closeness to past year high	IdioVol	Idiosyncratic volatility
INV	Investment	LEV	Leverage
ME	Size	LT_Rev	Long-term reversal
TURN	Turnover	NI	Net Share Issues
NOA	Net operating assets	OA	Operating accruals
OL	Operating leverage	OP	Operating profitability
PCM	Price to cost margin	PM	Profit margin
PROF	Profitability	Q	Tobin's Q
R12_2	Momentum	R12_7	Intermediate momentum
R36_13	Long-term reversal	R2_1	Short-term reversal
RNA	Return on net operating assets	ROA	Return on assets
ROE	Return on equity	RVAR	Residual Variance
S2P	Sales to price	SGA2S	Selling, general and administrative expenses to sales
SPREAD	Bid-ask spread	SUV	Standard unexplained volume
VAR	Variance		

OA.11.3.1 Univariate Analysis

We begin with univariate regressions of asset returns on each anomaly, including market betas estimated via the CAPM. The monthly t-statistics are displayed in Figure OA.26.

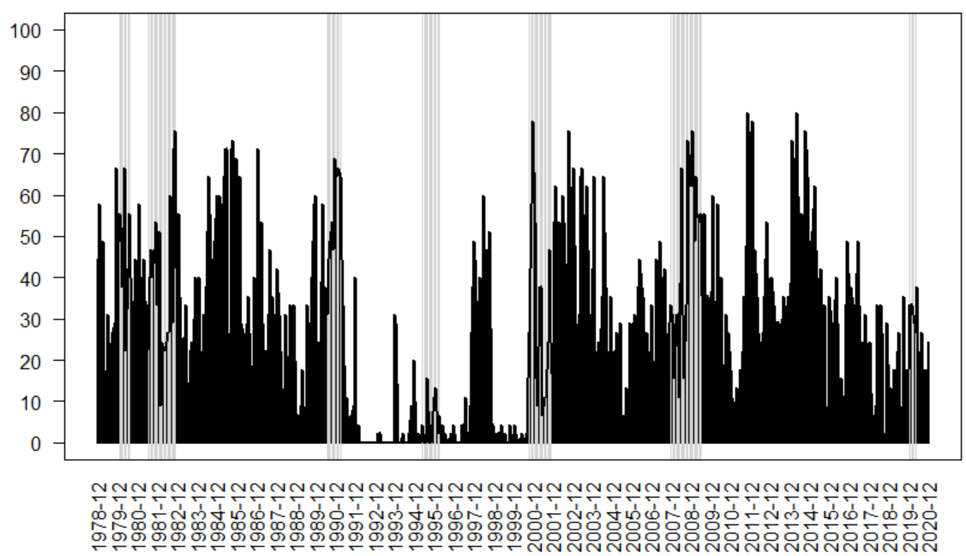
Figure OA.26 **Heatmap of the t-statistics distribution in the augmented-CAPM - BLLP Dataset**. The figure shows the heatmap of the t-statistics distribution obtained in each of the 45 univariate model (vertical axis) and for each month (horizontal axis) using data from Bryzgalova, Lerner, Lettau, and Pelger (2025). Each cell in the map represents the degree of statistical significance of the t-statistics with a different color, from gray (non-significant t-stat), to yellow (significance at only 10% level), orange (significance at 5% level), and red (significance at 1% level). The results are obtained by performing univariate regressions of asset returns on the market factor and each of the 45 anomalies, using the theoretical results of Section 4. The analysis uses balanced panels at each month, with a reference period ranging from December 1978 to December 2020. At each month t, the market beta is obtained by running a first-pass regression using a rolling window on the past two years of data ($T = 24$).



Consistent with our previous results, on average, 26.7% of anomalies are statistically significant across time and characteristics. Adding FF3 and FF5 factors reduces this percentage to 11.3% and 6.6%, respectively.[44] Notably, the two anomalies *Beta_m* (Fama and MacBeth, 1973) and *Beta_d* (Frazzini and Pedersen, 2014) show almost no explanatory power—consistent with expectations, as their information is already embedded in the CAPM beta included in our model.

To reaffirm the reliability of the temporal patterns and crisis effects identified in the previous sections, in Figure OA.27 we depict the percentage of anomalies found to be statistically significant at the 5% confidence level over time. In line with our previous findings, the figure confirms that higher percentages of statistically significant anomalies tend to correlate with periods of higher market uncertainty, with a peak of almost 80% during the 2007-2009 financial crisis.

Figure OA.27**Anomalies and financial crises in the augmented-CAPM - BLLP Dataset**. The figure shows the percentage of anomalies found to be significant at 5% (or lower) confidence level, by means of univariate regressions of asset returns on the market factor and each of the 45 anomalies. NBER recession dates and economic crises in grey bands (Bryzgalova, Lerner, Lettau, and Pelger (2025) dataset, $T = 24$).



To statistically re-establish the significance of such a relationship, we also replicate the fixed-effect regressions in (49) using the new dataset. Our results offer interesting insights into how financial crises impact the statistical significance of anomalies, even after accounting for multiple testing. In line with our previous results, anomalies are not just more likely to be significant; they are also significantly stronger in statistical terms during crises, with this effect being moderated when additional monetary policy factors are accounted for in the regression models. The results are reported in Table OA.XVIII.

[44]Details are available upon request.

Table OA.XVIII: **The impact of financial crises on the statistical significance of anomalies - BLLP Dataset.** Fixed-effect panel regression results designed to explore the impact of financial crises on the emergence and significance of anomalies. Specifically, it quantifies the influence of crisis periods on two key dimensions: the likelihood of an anomaly becoming statistically significant ($d_{|t|}$), and the degree of its significance as measured by the absolute t-statistic value ($|t|$). We tackle multiple testing distortions through Bonferroni and Holm methods ($d_{|t|\text{Bonferroni}}$ and $d|t|_{\text{Holm}}$) to address potential biases in type I errors. Panel A reports univariate regressions controlling for market factor betas under the CAPM framework. Panels B and C extend this analysis by incorporating the betas for the FF3 and FF5 models, respectively. For each dependent variable, the table reports the coefficient associated with the crisis dummy variable (d_{crisis}), indicating whether a period is classified as an NBER-defined crisis period. The coefficients are accompanied by their respective t-ratios in parentheses. Furthermore, the table explores the impact of additional explanatory variables, such as the federal funds rate ($fund\_rate$), its shadow rate ($fund\_shadow$, Wu and Xia (2016)), and unexpected monetary policy actions (mp_{shock}, Bu, Rogers, and Wu (2021)). The results are obtained using the firm characteristics data of Bryzgalova, Lerner, Lettau, and Pelger (2025).

Panel A: CAPM

	d_|t|		**d_|t|Bonferroni**		**d_|t|Holm**		**|t|**	
d_crisis	**0.166**	(19.47)	**0.122**	(18.53)	**0.125**	(18.73)	**0.873**	(23.70)
d_crisis	**0.167**	(19.14)	**0.123**	(18.24)	**0.125**	(18.35)	**0.889**	(23.60)
fund_rate	0.000	(-0.53)	0.000	(-0.67)	0.000	(-0.20)	**-0.006**	(-2.07)
d_crisis	**0.202**	(19.91)	**0.153**	(19.82)	**0.158**	(20.15)	**1.143**	(25.95)
fund_rate_shadow	**-0.027**	(-23.31)	**-0.016**	(-17.57)	**-0.016**	(-18.07)	**-0.137**	(-27.21)
d_crisis	**0.128**	(10.90)	**0.096**	(10.80)	**0.097**	(10.87)	**0.690**	(13.68)
mp_shock	**-0.545**	(-5.74)	**-0.333**	(-4.65)	**-0.366**	(-5.05)	**-2.894**	(-7.09)
d_crisis	**0.145**	(16.42)	**0.113**	(16.44)	**0.115**	(16.50)	**0.788**	(20.66)
fund_rate	**-0.023**	(-12.95)	**-0.011**	(-8.12)	**-0.011**	(-8.16)	**-0.111**	(-14.47)
fund_rate^2	**0.002**	(13.88)	**0.001**	(8.55)	**0.001**	(8.80)	**0.008**	(14.87)
d_crisis	**0.216**	(21.35)	**0.163**	(21.08)	**0.167**	(21.46)	**1.213**	(27.78)
fund_rate_shadow	**-0.061**	(-25.72)	**-0.039**	(-21.31)	**-0.040**	(-21.98)	**-0.313**	(-30.48)

Panel B: FF3

	d_|t|		**d_|t|Bonferroni**		**d_|t|Holm**		**|t|**	
d_crisis	**0.117**	(15.36)	**0.057**	(13.05)	**0.064**	(14.25)	**0.497**	(20.68)
d_crisis	**0.117**	(15.33)	**0.058**	(13.16)	**0.065**	(14.35)	**0.491**	(20.44)
fund_rate	0.000	(-0.53)	**0.001**	(2.87)	**0.001**	(2.63)	**-0.014**	(-5.82)
d_crisis	**0.143**	(19.92)	**0.069**	(17.68)	**0.077**	(19.03)	**0.603**	(26.95)
fund_rate_shadow	**-0.013**	(-15.79)	**-0.004**	(-9.74)	**-0.005**	(-10.27)	**-0.066**	(-25.67)
d_crisis	**0.175**	(20.56)	**0.086**	(18.48)	**0.096**	(19.88)	**0.710**	(26.95)
mp_shock	**-0.226**	(-3.28)	**-0.157**	(-4.15)	**-0.184**	(-4.72)	**-1.129**	(-5.30)

Continued on next page

Table OA.XVIII continued from previous page

d_crisis	**0.121**	(16.00)	**0.060**	(13.70)	**0.067**	(14.92)	**0.511**	(21.55)
fund_rate	**-0.037**	(-15.57)	**-0.016**	(-12.02)	**-0.018**	(-12.69)	**-0.178**	(-24.24)
fund_rate^2	**0.004**	(16.27)	**0.002**	(13.69)	**0.002**	(14.31)	**0.020**	(23.60)
d_crisis	**0.144**	(20.00)	**0.069**	(17.56)	**0.077**	(18.88)	**0.608**	(27.10)
fund_rate_shadow	**-0.016**	(-9.22)	**-0.004**	(-4.07)	**-0.004**	(-4.06)	**-0.079**	(-14.97)

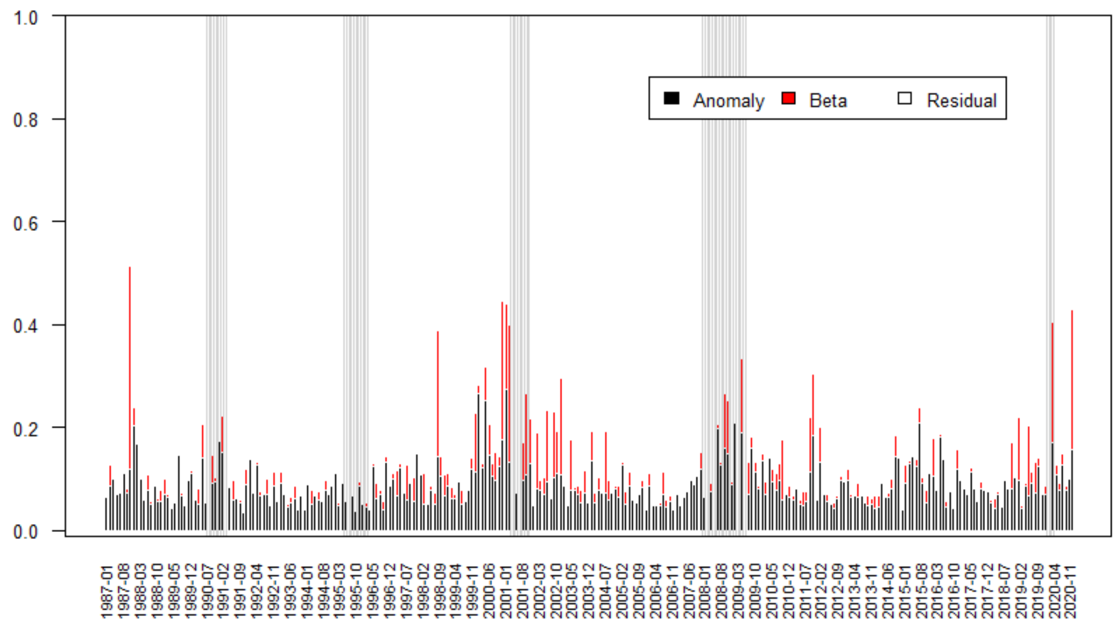
Panel C: FF5

	$\mathbf{d\_\|t\|}$		$\mathbf{d\_\|t\|}_{\text{Bonferroni}}$		$\mathbf{d\_\|t\|}_{\text{Holm}}$		$\mathbf{\|t\|}$	
d_crisis	**0.027**	(3.90)	**0.008**	(1.67)	0.006	(1.38)	**0.101**	(3.91)
d_crisis	-0.005	(-0.66)	-0.009	(-1.90)	**-0.010**	(-2.23)	-0.023	(-0.89)
fund_rate	**0.012**	(21.17)	**0.006**	(16.63)	**0.006**	(16.88)	**0.047**	(22.33)
d_crisis	**0.016**	(1.89)	0.004	(0.69)	0.002	(0.47)	**0.065**	(2.09)
fund_rate_shadow	**0.008**	(8.47)	**0.002**	(4.23)	**0.003**	(4.49)	**0.031**	(8.74)
d_crisis	**0.018**	(1.89)	0.005	(0.77)	0.004	(0.60)	**0.071**	(1.94)
mp_shock	-0.077	(-1.02)	-0.020	(-0.41)	-0.019	(-0.40)	-0.515	(-1.74)
d_crisis	-0.005	(-0.72)	0.005	(1.03)	**-0.015**	(-3.13)	-0.028	(-1.04)
fund_rate	**0.012**	(8.02)	-0.002	(-1.56)	0.002	(1.89)	**0.043**	(7.99)
fund_rate^2	0.000	(0.40)	**0.001**	(4.14)	**0.000**	(5.22)	0.000	(0.93)
d_crisis	**0.020**	(2.41)	0.005	(1.03)	0.004	(0.82)	**0.082**	(2.64)
fund_rate_shadow	-0.003	(-1.37)	-0.002	(-1.56)	-0.002	(-1.53)	-0.012	(-1.61)

OA.11.3.2 Multivariate Analysis

When analyzing the 45 anomalies of BLLP Dataset altogether, our time-varying R_z^2 measure again identifies a significant temporal variation in the explanatory power of anomalies. As shown in Figure OA.28, anomalies account for an average of 8.9% of return variance, while market beta contributes just 2.8%.[45]

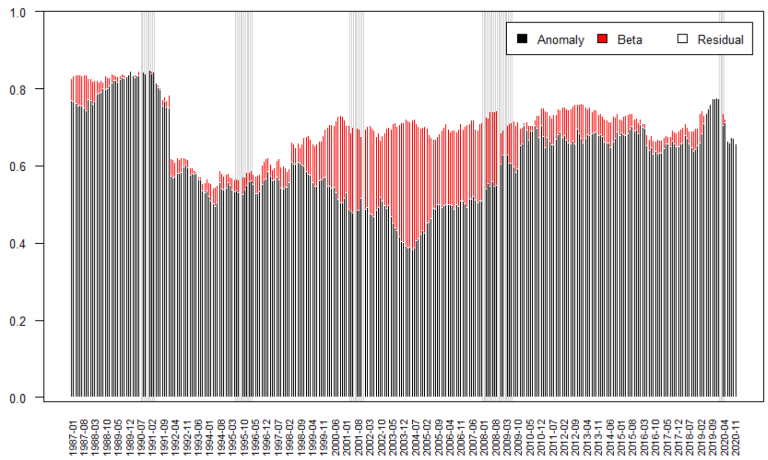
Figure OA.28**Anomalies' contribution using time-varying multivariate OLS regressions - BLLP Dataset** . The figure shows the time series of the total variance decomposition obtained in each multivariate regression of asset returns on the market factor and the 45 firms characteristics (Bryzgalova, Lerner, Lettau, and Pelger (2025) dataset, $T = 24$).



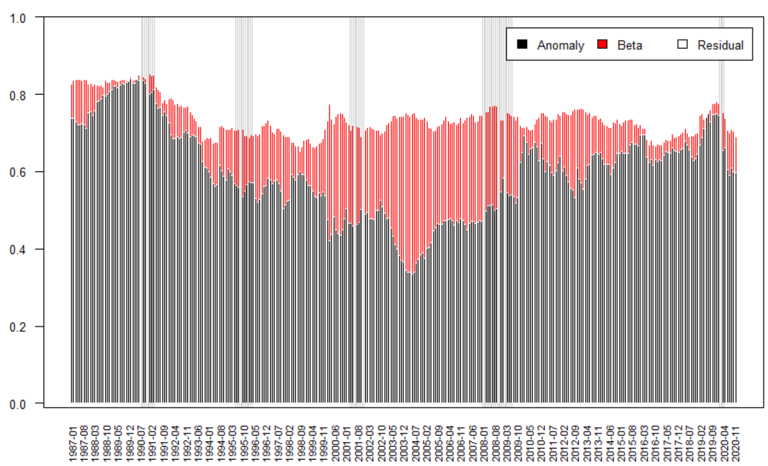
Compared to this modest 8.9%, the *local-average* method (using a $T = 120$-month window), as in the previous case, reveals a substantial increase in the explanatory power of anomalies. As shown in Figure OA.29, anomalies explain up to 61.6% of return variation under the CAPM, 60.0% with FF3, and 59.1% with FF5. In contrast, the contribution of betas, while slightly higher, remains limited.

[45]This result is consistent even when the model includes the FF3 and FF5 factors, in which betas and anomalies together account for 13.7% and 16.6% of the total variance, respectively. The anomalies still lead, contributing between 61.8% and 53.1% of such combined explanatory power. Details are available upon request.

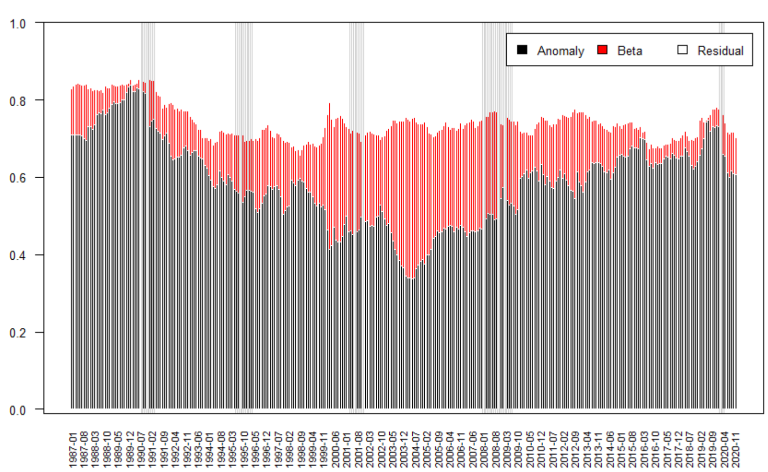
Figure OA.29**Anomalies' contribution using local-average approach - BLLP Dataset**. The figure shows the time series of the total variance decomposition obtained in each multivariate regression using rolling window of $T = 120$ months. The results are obtained by performing multivariate regressions of asset returns on all the 82 firms characteristics together with (a) the market factor betas, (b) the FF3 factors betas, and (c) the FF5 factors betas. The application uses balanced panels for each rolling sample, with a reference period ranging from December 1978 to December 2020 (Bryzgalova, Lerner, Lettau, and Pelger (2025) dataset).



(a) CAPM Model



(b) FF3 Model



(c) FF5 Model

OA.11.3.3 WLS Estimation

To enhance the robustness of our analysis against the potential distorting influence of microcaps, we reapply our WLS methodology to the BLLP dataset, using market equity (ME) as the matrix of weights.[46] The key results are detailed in Table OA.XIX, which displays the proportion of significant anomalies detected through WLS and compares these figures in both absolute and relative terms to the results achieved with OLS. Consistent with our previous findings, accounting for microcaps in the analysis leads to a significant change in the overall statistical significance of anomalies, showing a relative reduction of 36.8%—from 26.7% in the OLS analysis to 16.9% using WLS. This pattern persists when extending the analysis to include the three- and five-factor Fama-French models.

Table OA.XIX: **Anomalies' significance: OLS vs WLS in the BLLP Dataset.** The table presents the proportion of significant anomalies as identified by our methodology, comparing outcomes from OLS (Columns 1–3) and WLS (Columns 4–6) approaches, and using three different model specifications (CAPM, FF3, and FF5). Columns 7–9 and 10–12 report the absolute and relative changes between the OLS and WLS results, respectively.

	OLS (%)			WLS (%)			WLS-OLS (%)			WLS/OLS-1 (%)		
Anomaly	CAPM	FF3	FF5	CAPM	FF3	FF5	CAPM	FF3	FF5	CAPM	FF3	FF5
A2ME	33.27	6.38	4.16	22.77	2.96	3.87	-10.5	-3.4	-0.29	-31.5	-53.6	-6.9
AC	20.20	9.57	6.53	6.14	2.73	8.43	-14.1	-6.8	1.89	-69.6	-71.4	29.0
AT	45.15	2.96	5.74	26.53	0.46	0.91	-18.6	-2.5	-4.83	-41.2	-84.6	-84.1
ATO	31.49	14.12	4.95	22.18	8.88	5.47	-9.3	-5.2	0.52	-29.6	-37.1	10.4
B2M	30.30	9.79	6.93	21.39	4.10	6.15	-8.9	-5.7	-0.78	-29.4	-58.1	-11.3
BETA_d	1.98	4.33	3.17	1.19	0.91	2.73	-0.8	-3.4	-0.43	-40.0	-78.9	-13.7
BETA_m	0.40	0.68	9.90	0.20	0.23	0.46	-0.2	-0.5	-9.45	-50.0	-66.7	-95.4
C2A	19.41	6.61	3.96	11.68	4.10	2.73	-7.7	-2.5	-1.23	-39.8	-37.9	-31.0
CF2B	28.12	14.35	6.93	17.03	8.20	6.61	-11.1	-6.2	-0.32	-39.4	-42.9	-4.7
CF2P	33.47	14.12	8.71	15.45	5.47	5.01	-18.0	-8.7	-3.70	-53.8	-61.3	-42.5
CTO	32.67	13.21	4.55	24.36	8.20	2.28	-8.3	-5.0	-2.28	-25.5	-37.9	-50.0
D2A	21.39	8.43	5.35	20.99	6.61	2.73	-0.4	-1.8	-2.61	-1.9	-21.6	-48.9
D2P	18.42	8.66	5.94	7.72	3.19	4.56	-10.7	-5.5	-1.38	-58.1	-63.2	-23.3
DPI2A	21.39	11.85	6.93	16.04	9.57	7.52	-5.3	-2.3	0.59	-25.0	-19.2	8.5
E2P	39.21	21.41	10.69	21.78	10.71	11.62	-17.4	-10.7	0.92	-44.4	-50.0	8.6
FC2Y	32.28	6.15	6.14	25.94	5.24	2.51	-6.3	-0.9	-3.63	-19.6	-14.8	-59.2
HIGH52	24.95	13.90	4.95	14.85	8.20	7.97	-10.1	-5.7	3.02	-40.5	-41.0	61.0
INV	24.36	19.36	5.35	12.87	11.16	5.47	-11.5	-8.2	0.12	-47.2	-42.4	2.3
IdioVol	21.78	11.62	7.52	10.89	3.87	4.10	-10.9	-7.7	-3.42	-50.0	-66.7	-45.5
LEV	25.54	4.33	2.77	16.83	1.37	1.37	-8.7	-3.0	-1.41	-34.1	-68.4	-50.7
NI	15.25	8.66	4.55	9.70	5.24	5.92	-5.5	-3.4	1.37	-36.4	-39.5	30.0
NOA	24.55	10.71	6.14	22.77	8.66	2.73	-1.8	-2.1	-3.41	-7.3	-19.1	-55.5
OA	32.08	20.73	10.30	8.12	4.78	7.52	-24.0	-15.9	-2.78	-74.7	-76.9	-27.0
OL	25.15	7.97	4.75	20.20	6.15	0.91	-5.0	-1.8	-3.84	-19.7	-22.9	-80.8
OP	36.44	12.53	7.52	24.36	8.20	5.01	-12.1	-4.3	-2.51	-33.2	-34.5	-33.4
PCM	23.96	6.38	9.50	14.65	2.51	4.78	-9.3	-3.9	-4.72	-38.8	-60.7	-49.7
PM	32.08	12.07	6.53	20.40	6.15	2.96	-11.7	-5.9	-3.57	-36.4	-49.1	-54.7
PROF	32.87	18.00	4.55	22.77	10.93	6.61	-10.1	-7.1	2.05	-30.7	-39.2	45.0
Q	31.09	8.20	3.56	21.98	3.19	3.87	-9.1	-5.0	0.31	-29.3	-61.1	8.6

Continued on next page

[46] In this case, we recenter the original variable to fit within the interval $[0, 1]$, ensuring that the weights are positive.

Table OA.XIX – continued from previous page

	OLS (%)			WLS (%)			WLS-OLS (%)			WLS/OLS-1 (%)		
Anomaly	CAPM	FF3	FF5	CAPM	FF3	FF5	CAPM	FF3	FF5	CAPM	FF3	FF5
R12_2	25.74	12.53	3.56	20.59	10.02	6.38	-5.1	-2.5	2.81	-20.0	-20.0	78.9
R12_7	21.19	10.48	3.76	14.06	8.88	6.15	-7.1	-1.6	2.39	-33.6	-15.2	63.5
R2_1	20.40	16.40	3.96	14.85	8.88	7.74	-5.5	-7.5	3.78	-27.2	-45.8	95.6
R36_13	22.18	15.26	4.36	17.23	9.79	3.64	-5.0	-5.5	-0.71	-22.3	-35.8	-16.3
R60_13	26.34	16.86	3.96	19.41	9.79	2.73	-6.9	-7.1	-1.23	-26.3	-41.9	-31.0
RNA	40.20	20.73	9.70	27.52	12.53	11.39	-12.7	-8.2	1.69	-31.5	-39.6	17.4
ROA	36.63	22.78	5.54	24.55	10.93	6.15	-12.1	-11.8	0.61	-33.0	-52.0	10.9
ROE	30.69	13.90	7.13	21.19	9.79	6.15	-9.5	-4.1	-0.98	-31.0	-29.5	-13.7
RVAR	21.39	11.62	7.33	10.69	3.87	3.64	-10.7	-7.7	-3.68	-50.0	-66.7	-50.3
S2P	41.19	10.02	6.53	29.90	4.33	3.87	-11.3	-5.7	-2.66	-27.4	-56.8	-40.7
SGA2S	33.27	5.69	6.34	22.97	3.64	2.51	-10.3	-2.1	-3.83	-31.0	-36.0	-60.5
SPREAD	26.14	7.97	7.52	12.87	2.28	4.10	-13.3	-5.7	-3.42	-50.8	-71.4	-45.5
SUV	22.97	11.16	6.34	11.88	5.24	7.52	-11.1	-5.9	1.18	-48.3	-53.1	18.6
TURN	11.88	8.20	2.77	5.74	2.73	4.33	-6.1	-5.5	1.56	-51.7	-66.7	56.1
VAR	16.63	10.02	6.73	7.92	3.19	3.19	-8.7	-6.8	-3.54	-52.4	-68.2	-52.6
Overall	26.74	11.34	6.06	16.89	6.09	4.83	-9.8	-5.3	-1.23	-36.8	-46.3	-20.3

Interestingly, our findings align with those we previously obtained in Section 6.2.3 using the FHNW dataset. Along with accrual-related variables (OA and AC), which show the largest relative drops in significance (–74.7% and –69.6%), our analysis continues to highlight significant reductions for characteristics associated with firms' liquidity and profitability. For example, the bid-ask spread (SPREAD) falls by 13.3% (–50.8%), cash flow (CF2P) by 18% (–53%), operating profitability (OP) by 12.1% (–33.2%), earnings to price (E2P) by 17.4% (–44.4%), and return on assets (ROA) by 12.1% (–33%).

Despite the overall decline in anomaly significance, the impact of financial crises remains strong. Using WLS, we replicate the fixed-effects panel regressions and continue to find a robust, positive, and statistically significant effect of NBER-defined crises, even after multiple testing adjustments (See results in Table OA.XX)

Table OA.XX: **WLS estimation: the impact of financial crises on the statistical significance of anomalies - BLLP Dataset.** See the description of Table OA.XVI.

	$\mathbf{d\_\vert t\vert}$		$\mathbf{d\_\vert t\vert}_{\text{Bonferroni}}$		$\mathbf{d\_\vert t\vert}_{\text{Holm}}$		$\mathbf{\vert t\vert}$	
Panel A: CAPM								
d_crisis	**0.152**	(20.75)	**0.092**	(18.54)	**0.094**	(18.83)	**0.713**	(28.35)
Panel B: FF3								
d_crisis	**0.114**	(19.72)	**0.136**	(19.88)	**0.144**	(20.79)	**0.473**	(26.36)
Panel C: FF5								
d_crisis	**0.090**	(17.30)	**0.136**	(19.88)	**0.144**	(20.79)	**0.473**	(28.19)

References

different style from biblio of MAIN??

Acharya, V. V., and L. H. Pedersen (2005): "Asset pricing with liquidity risk," *Journal of Financial Economics*, 77, 375–410.

Adler, A., and A. Rosalsky (1991): "On the weak law of large numbers for normed weighted sums of I.I.D. random variables," *International Journal of Mathematics and Mathematical Sciences*, 14, 191–202.

Ahn, S. C., A. R. Horenstein, and N. Wang (2018): "Beta matrix and common factors in stock returns," *Journal of Financial and Quantitative Analysis*, 53, 1417–1440.

Anatolyev, S., and A. Mikusheva (2022): "Factor models with many assets: Strong factors, weak factors, and the two-pass procedure," *Journal of Econometrics*, 229, 103–126.

Ang, A., R. Hodrick, Y. Xing, and X. Zhang (2006): "The cross section of volatility and expected return," *The Journal of Finance*, 11, 259–299.

Ang, A., J. Liu, and K. Schwarz (2020): "Using stocks or portfolios in tests of factor models," *Journal of Financial and Quantitative Analysis*, 55, 709–750.

Avarucci, M., and P. Zaffaroni (2022): "Robust estimation of large panels with factor structures," *Journal of the American Statistical Association*, 118, 1–33.

Avramov, D., and T. Chordia (2006): "Asset pricing models and financial market anomalies," *The Review of Financial Studies*, 19, 1001–1040.

Barras, L., O. Scaillet, and R. Wermers (2010): "False discoveries in mutual fund performance: Measuring luck in estimated alphas," *The Journal of Finance*, 65, 179–216.

Brennan, M., T. Chordia, and A. Subrahmanyam (1998): "Alternative factor specifications, security characteristics, and the cross-section of expected stock returns," *Journal of Financial Economics*, 49, 345–373.

Bryzgalova, S. (2014): "Spurious factors in linear asset pricing models," *Working Paper, Stanford University*.

Bryzgalova, S., S. Lerner, M. Lettau, and M. Pelger (2025): "Missing financial data," *The Review of Financial Studies*, 38, 803–882.

Bu, C., J. Rogers, and W. Wu (2021): "A unified measure of FED monetary policy shocks," *Journal of Monetary Economics*, 118, 331–349.

Burnside, C. (2016): "Identification and inference in linear stochastic discount factor models with excess returns," *Journal of Financial Econometrics*, 14, 295–330.

Carhart, M. M. (1997): "On persistence in mutual fund performance," *The Journal of Finance*, 52, 57–82.

Chamberlain, G., and M. Rothschild (1983): "Arbitrage, factor structure, and mean-variance analysis on large asset markets," *Econometrica*, 51, 1281–1304.

Chen, A. (2021): "The limits of p-hacking: Some thought experiments," *Journal of Finance*, 76, 2447–2480.

Chordia, T., A. Goyal, and J. Shanken (2015): "Cross-sectional asset pricing with individual stocks: Betas versus characteristics," *Working Paper, Emory University*.

Chu, C.-S. J., K. Hornik, and C.-M. Kaun (1995): "MOSUM tests for parameter constancy," *Biometrika*, 82, 603–617.

Connor, G., and R. Korajczyk (2024): "Semi-strong factors in asset returns," *Journal of Financial Econometrics*, 22, 70–93.

Connor, G., and O. Linton (2007): "Semiparametric estimation of a characteristic-based factor model of common stock returns," *Journal of Empirical Finance*, 14, 694–717.

Conrad, J., M. Cooper, and G. Kaul (2003): "Value versus glamour," *The Journal of Finance*, 58, 1969–1995.

Daniel, K., and S. Titman (1997): "Evidence on the characteristics of cross sectional variation in stock returns," *The Journal of Finance*, 52, 1–33.

DeMiguel, V., A. Martin-Utrera, F. J. Nogales, and R. Uppal (2020): "A transaction-cost perspective on the multitude of firm characteristics," *The Review of Financial Studies*, 33, 2180–2222.

Dong, X., L. Yan, D. E. Rapach, and G. Zhou (2022): "Anomalies and the expected market return," *The Journal of Finance*, 77, 639–681.

Fama, E. F. (1976): *Foundations of Finance*. New York: Basic Books.

Fama, E. F., and K. R. French (1993): "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 33, 3–56.

——— (2008): "Dissecting anomalies," *The Journal of Finance*, 63, 1653–1678.

——— (2015): "A five-factor asset pricing model," *Journal of Financial Economics*, 116, 1–22.

Fama, E. F., and J. D. MacBeth (1973): "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy*, 81, 607–636.

Fan, J., Y. Liao, and W. Wang (2016): "Projected principal component analysis in factor models," *The Annals of Statistics*, 44, 219–254.

Frazzini, A., and L. H. Pedersen (2014): "Betting against beta," *Journal of Financial Economics*, 111, 1–25.

Freyberger, J., B. Hoeppner, A. Neuhierl, and M. Weber (2025): "Missing data in asset pricing panels," *The Review of Financial Studies*, 38, 760–802.

Freyberger, J., A. Neuhierl, and M. Weber (2020): "Dissecting characteristics nonparametrically," *The Review of Financial Studies*, 33, 2326–2377.

Gagliardini, P., E. Ossola, and O. Scaillet (2016): "Time-varying risk premium in large cross-sectional equity data sets," *Econometrica*, 84, 985–1046.

Giglio, S., D. Xiu, and D. Zhang (2025): "Test assets and weak factors," *The Journal of Finance*, 80, 259–319.

Gospodinov, N., R. Kan, and C. Robotti (2014): "Misspecification-robust inference in linear asset-pricing models with irrelevant risk factors," *The Review of Financial Studies*, 27, 2139–2170.

Hansen, L. P., and R. Jagannathan (2007): "Assessing specification errors in stochastic discount factor models," *The Journal of Finance*, 52, 557–590.

Hansen, L. P., and S. F. Richard (1987): "The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models," *Econometrica*, 55, 587–613.

Harvey, C. R., Y. Liu, and H. Zhu (2016): "... and the cross-section of expected returns," *The Review of Financial Studies*, 29, 5–68.

Hou, K., X. Chen, and L. Zhang (2015): "Digesting anomalies: An investment approach," *The Review of Financial Studies*, 28, 650–705.

——— (2020): "Replicating anomalies," *The Review of Financial Studies*, 12, 2019–2133.

Hou, K., and R. Kimmel (2006): "On the estimation of risk premia in linear factor models," *Working Paper, Ohio State University.*

Jagannathan, R., and Z. Wang (1998): "An asymptotic theory for estimating beta-pricing models using cross-sectional regression," *The Journal of Finance*, 53, 1285–1309.

Jegadeesh, N., and S. Titman (1993): "Returns to buying winners and selling losers: Implications for stock market efficiency," *The Journal of Finance*, 48, 65–91.

Kan, R., C. Robotti, and J. Shanken (2013): "Pricing model performance and the two-pass cross-sectional regression methodology," *The Journal of Finance*, 68, 2617–2649.

Kan, R., and C. Zhang (1999a): "GMM tests of stochastic discount factor models with useless factors," *Journal of Financial Economics*, 54, 103–127.

——— (1999b): "Two-pass tests of asset pricing models with useless factors," *The Journal of Finance*, 54, 203–235.

Kelly, B. T., S. Pruitt, and Y. Su (2019): "Characteristics are covariances: A unified model of risk and return," *Journal of Financial Economics*, 134, 501–524.

Kim, S., R. A. Korajczyk, and A. Neuhierl (2021): "Arbitrage portfolios," *The Review of Financial Studies*, 34, 2813–2856.

Kim, S., V. Raponi, and P. Zaffaroni (2023): "Testing for weak factors in asset pricing," Discussion paper, KAIST, IESE, Imperial College London.

Kleibergen, F. (2009): "Tests of risk premia in linear factor models," *Journal of Econometrics*, 149, 149–173.

Kleibergen, F., and Z. Zhan (2014): "Mimicking portfolios of macroeconomic factors," *Working Paper, Aarhus University.*

——— (2020): "Robust inference for consumption-based asset pricing," *The Journal of Finance*, 75, 507–550.

Kozak, S., S. Nagel, and S. Santosh (2020): "Shrinking the cross-section," *Journal of Financial Economics*, 135, 271–292.

Kuersteiner, G. M., and I. R. Prucha (2013): "Limit theory for panel data models with cross–sectional dependence and sequential exogeneity," *Journal of Econometrics*, 174, 107–126.

Lettau, M., and M. Pelger (2020): "Factors that fit the time series and cross-section of stock returns," *The Review of Financial Studies*, 33, 2274–2325.

Lewellen, J. (2015): "The cross-section of expected stock returns," *Critical Finance Review*, 4, 1–44.

Lo, A. W., and A. C. MacKinlay (1990): "Data-snooping biases in tests of financial asset pricing models," *The Review of Financial Studies*, 3, 431–467.

McLean, R., and J. Pontiff (2016): "Does academic research destroy stock return predictability?," *The Journal of Finance*, 71, 5–32.

Pástor, L., and R. F. Stambaugh (2003): "Liquidity risk and expected stock returns," *Journal of Political Economy*, 111, 642–685.

Ploberger, W., W. Krämer, and K. Kontrus (1989): "A new test for structural stability in the linear regression model," *Journal of Econometrics*, 40, 307–318.

Raponi, V., C. Robotti, and P. Zaffaroni (2020): "Testing beta-pricing models using large cross-sections," *The Review of Financial Studies*, 33, 2796–2842.

Shanken, J. (1992): "On the estimation of beta-pricing models," *The Review of Financial Studies*, 5, 1–33.

Shanken, J., and G. Zhou (2007): "Estimating and testing beta pricing models: Alternative methods and their performance in simulations," *Journal of Financial Economics*, 84, 40–86.

Stambaugh, R. F., and Y. Yuan (2017): "Mispricing factors," *The Review of Financial Studies*, 30, 1270–1315.

Wu, J. C., and F. D. Xia (2016): "Measuring the macroeconomic impact of monetary policy at the zero lower bound," *Journal of Money, Credit and Banking*, 48, 253–291.