

# A Practitioner's Guide to Identification-Robust Methods in Asset Pricing

Frank Kleibergen\*      Zhaoguo Zhan<sup>†</sup>

September 9, 2026

## Abstract

A variety of identification-robust inference methods have been proposed in the econometrics literature, many of which are explicitly developed for asset pricing. In the classic two-pass procedure for inferring risk premia, weak or no identification arises when the observed risk factors are of poor quality, while the commonly used  $t$ -statistic may spuriously favor such factors. The identification issue similarly jeopardizes the generalized method of moments (GMM), making conventional inference on the structural parameter of interest unreliable. We show that for both the two-pass procedure and GMM in asset pricing, identification-robust methods are available, which can be further extended to account for model misspecification.

**Keywords:** Identification-robust, two-pass regression, GMM.

---

\*Email: f.r.kleibergen@uva.nl. Amsterdam School of Economics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands.

<sup>†</sup>Email: zzhan@kennesaw.edu. Department of Economics, Finance and Quantitative Analysis, Coles College of Business, Kennesaw State University, GA 30144, USA.

# 1 Introduction

Over the past decades, the large and growing literature on weak identification has provided a variety of so-called identification-robust inference methods; see, for example, Staiger and Stock (1997), Dufour (1997), Stock and Wright (2000), Kleibergen (2002, 2005), Moreira (2003), Andrews and Cheng (2012), Andrews and Mikusheva (2016a, 2016b), and Han and McCloskey (2019). Much of this literature focuses on the generalized method of moments (GMM) framework of Hansen (1982), where identification of the structural parameter of interest relies on the GMM rank condition. Weak identification arises in this framework when the rank condition is at risk, and a leading example that has attracted sizable attention is the weak instrument problem in the linear instrumental variables (IV) regression; see also Stock, Wright, and Yogo (2002). To address this issue, identification-robust methods therefore aim to provide uniformly reliable inference regardless of the identification strength as reflected by the rank condition.

This guide is motivated by the fact that the identification issue is also prevalent in the asset pricing literature, where hundreds of risk factors have been proposed to explain the cross-section of expected asset returns; see, for example, Harvey, Liu, and Zhu (2016). It is hard to comprehend that so many factors can explain the cross-section of asset returns, which more than hints at the reliability of the commonly used methodologies for establishing the validity of these factors, such as the popular Fama-MacBeth (FM, 1973) two-pass procedure. It has been documented for about twenty years, parallel to the weak instrument literature, that the risk premia on many risk factors are not well identified, so their associated FM  $t$ -statistic cannot be taken at face value; see, for example, Kan and Zhang (1999), and Kleibergen (2009). The widely used cross-sectional  $R^2$  in the two-pass procedure can also be spuriously large, to the extent that it may favor irrelevant factors that provide no identification, as shown by Kleibergen and Zhan (2015). Awareness that

weak or no identification provides a major obstacle for empirical asset pricing is currently emerging. Correspondingly, identification-robust methods that remain valid regardless of the quality of risk factors have been developed; see, for example, Kleibergen (2009), Beaulieu, Dufour, and Khalaf (2013), and Kleibergen and Zhan (2020).

The identification problem in asset pricing is further aggravated because asset pricing models are widely considered misspecified; see, for example, Kan, Robotti, and Shanken (2013), Gospodinov, Kan, and Robotti (2014). Given that asset pricing models are approximations to reality, misspecification, together with possibly weak identification, is a general concern that needs to be addressed. The identification condition then becomes more stringent in misspecified models compared to correctly specified models, *i.e.*, the full rank condition is insufficient to identify the structural parameter of interest in possibly misspecified models; see Kleibergen and Zhan (2026). The existing identification-robust inference methods, however, can be extended to account for model misspecification, so the resulting methodology becomes doubly robust, *i.e.*, it takes both weak identification and misspecification into account; see, for example, Kleibergen and Zhan (2025a).

In addition to the common practice of using risk factors and portfolio returns for conducting asset pricing tests, empirical asset pricing studies often involve variations such as using the constructed mimicking portfolios to replace the observed risk factors, or using a large cross-section of individual assets instead of a small number of portfolios. These variations may help alleviate the identification issue, but do not fully resolve it. Therefore, identification-robust methods remain relevant in these scenarios; see, for example, Kleibergen and Zhan (2018, 2025b). Another practically relevant issue is that the typical sample size associated with macroeconomic factors, such as consumption growth, jeopardizes asymptotically valid inference methods, which therefore need to be adjusted for small-sample applications; see, for example, Kleibergen and Zhan (2020), and Kleibergen, Kong, and Zhan (2023).

Overall, identification-robust inference methods do not require the parameter of interest (such as risk premia in asset pricing) to be strongly identified, so they provide reliable inference regardless of the strength of identification. In contrast, the traditional two-pass and GMM methodology generally assume strong identification, so a consistent and asymptotically normal estimator for the parameter of interest exists, and consequently the commonly used  $t$ -test applies. In the practically relevant scenario where the identification condition is possibly at risk, identification-robust methods therefore remain reliable and can lead to empirical findings that are substantially different from conventional ones.

This guide proceeds as follows. In Section 2, we start from the identification issue in beta pricing with the linear factor model. Section 3 extends to risk premia from the cross-section of individual assets. Section 4 discusses the co-existence of misspecification and weak identification. Section 5 concludes. The Appendix contains the related implementation details for practitioners.

## 2 Inference on risk premia in beta pricing

### 2.1 Beta pricing

Let  $R_{i,t}$  be the return on the  $i$ -th asset at time  $t$ , with  $i = 1, \dots, N$ , and  $t = 1, \dots, T$ . The beta pricing equation relates the expected asset return to a linear function of betas:

$$E(R_{i,t}) = \lambda_0 + \beta_i' \lambda_f, \quad (1)$$

where  $\lambda_0$  is the zero-beta return,  $\lambda_f$  the  $k$ -dimensional risk premium vector,  $\beta_i$  is a  $k$ -dimensional vector of betas that results from a time-series regression of  $R_{i,t}$  on the  $k$  risk factors in  $f_t$ , *i.e.*,

$$\beta_i = \text{Var}(f_t)^{-1} \text{Cov}(f_t, R_{i,t}). \quad (2)$$

Linear asset pricing as described by (1) and (2) started with the classic Capital Asset Pricing Model (CAPM), for which the single risk factor, market return, has been found to be too limited to explain the cross-section of asset returns. Similarly, for the classic consumption CAPM (C-CAPM), the single risk factor, consumption growth, does not appear to work well in existing studies either. See, for example, Lettau and Ludvigson (2001), Kleibergen and Zhan (2020).

## 2.2 FM two-pass procedure

Given the unsatisfactory performance of CAPM and C-CAPM, many additional risk factors have therefore been discovered using the popular Fama-MacBeth (1973, FM) two-pass procedure. The FM two-pass procedure estimates the risk premia on the  $k \times 1$  dimensional risk factor  $f_t$  in two steps (passes):

1. **Time-series regression:** Estimate the betas,  $\beta_i$ ,  $i = 1, \dots, N$ , of the  $N$  models:

$$R_{i,t} = c_i + f_t' \beta_i + v_{i,t}, \quad t = 1, \dots, T, \quad (3)$$

with  $v_{i,t}$  the error term, using linear regression to obtain  $\hat{\beta}_i$ .

2. **Cross-section regression:** For  $\bar{R} = \frac{1}{T} \sum_{t=1}^T R_t$ , with  $R_t = (R_{1,t} \dots R_{N,t})'$ , use  $\hat{\beta} = (\hat{\beta}_1 \dots \hat{\beta}_N)'$  for  $\beta = (\beta_1 \dots \beta_N)'$  to estimate  $\lambda_0$  and  $\lambda_f$  in the cross-section regression:

$$\bar{R} = \iota_N \lambda_0 + \beta \lambda_f + \bar{v}, \quad (4)$$

with  $\iota_N$  the  $N$ -dimensional vector of ones,  $\bar{v} = \frac{1}{T} \sum_{t=1}^T v_t$ ,  $v_t = (v_{1,t} \dots v_{N,t})'$ . Thus, the two-pass risk premia estimator is:

$$\hat{\lambda}_f = \left( \hat{\beta}' M_{\iota_N} \hat{\beta} \right)^{-1} \hat{\beta}' M_{\iota_N} \bar{R}, \quad (5)$$

with the projection matrix  $M_{\iota_N} = I_N - \iota_N(\iota_N'\iota_N)^{-1}\iota_N'$ .

For asset returns, it is common to use portfolio returns, like, for example, the  $N = 25$  Fama-French size and book-to-market sorted portfolios and extensions thereof. These portfolio returns exhibit a very strong factor structure: for example, Kleibergen and Zhan (2015) show that about 95% of their variation is explained by the largest three principal components in the data taken from Lettau and Ludvigson (2001). Factors thus have to correlate with the largest three principal components to have meaningful risk premia.

Practitioners, however, typically report two conventional statistics from the FM two-pass procedure:

1. cross-sectional  $R^2$  of the second pass regression;
2.  $t$ -statistic on the risk premia estimates.

Large and significant values of these measures are considered to be indicative of relevant risk premia, so the proposed risk factors are favored.

Despite its widespread use, researchers have questioned the reliability of the FM two-pass procedure; see, for example, Lewellen, Nagel, and Shanken (2010). An early study by Kan and Zhang (1999) considers a scenario of no-identification: when the proposed risk factors are useless (independent of asset returns), the betas are zero, so the risk premia  $\lambda_f$  are unidentified. The FM  $t$ -statistic on  $\lambda_f$ , however, can be spuriously large to favor useless factors. In line with Kan and Zhang (1999), Kleibergen (2009) further considers local-to-zero betas, and warns that the  $t$ -statistic does not provide valid inference under this weak identification scenario. On the other hand, Kleibergen and Zhan (2015) show that the cross-sectional  $R^2$  can also be spuriously large under useless or weak factors. Therefore, the cross-sectional  $R^2$  and  $t$ -statistic commonly used by practitioners just set too low a bar for gauging risk factors, because they can be spuriously large even if the identification of risk premia is at risk.

## 2.3 Diagnostic statistics

The identification issue in the two-pass procedure can be revealed if additional diagnostic statistics are reported in practice, such as the pseudo- $R^2$ , and the rank test outcome. For illustration, we take an empirical study from Kleibergen and Zhan (2020), which is presented in Table 1 below.

As reported in Table 1, Lettau and Ludvigson (2001) find a significant risk premium for the interaction between the (lagged) consumption-wealth ratio  $\text{cay}$  and consumption growth  $\Delta c$  in a three-factor model: both the FM  $t$ -statistic and Shanken (1992)  $t$ -statistic on  $\Delta c \times \text{cay}$  are significantly positive. The cross-sectional  $R^2$  is also large.

Table 1: Risk premia in multi-factor models

|                       | <b>LL3</b> |            |                              | <b>FF3</b> |            |            |
|-----------------------|------------|------------|------------------------------|------------|------------|------------|
|                       | $\Delta c$ | <b>cay</b> | $\Delta c \times \text{cay}$ | $R_m$      | <b>SMB</b> | <b>HML</b> |
| $\hat{\lambda}_f$     | 0.02       | -0.13      | 0.06                         | 1.33       | 0.47       | 1.46       |
| FM $t$                | 0.20       | -0.43      | 3.13                         | 0.83       | 0.94       | 3.24       |
| Shanken $t$           | 0.15       | -0.31      | 2.25                         | 0.78       | 0.94       | 3.22       |
| Cross-sectional $R^2$ |            | 0.698      |                              |            | 0.803      |            |
| Diagnostic statistics |            |            |                              |            |            |            |
| Pseudo- $R^2$         |            | 0.105      |                              |            | 0.946      |            |
| Rank test $p$ -value  |            | 0.274      |                              |            | 0.030      |            |

Notes: This table is taken from Kleibergen and Zhan (2020). LL3 uses the three-factor model of Lettau and Ludvigson (2001), while FF3 uses the three-factor model of Fama and French (1993). The test assets are the 25 Fama-French portfolios over 1963Q3 to 1998Q3. We refer to the published articles for further details of the adopted data and FM two-pass estimation procedure.

The cross-sectional  $R^2$  of the second pass regression, however, is sensitive to the factor structure of asset returns. In fact, the cross-sectional  $R^2$  can be large due to the factor structure irrespective of whether the factors in the FM two-pass procedure capture it. Kleibergen and Zhan (2015) therefore suggest to report the first pass pseudo- $R^2$  to accompany the cross-sectional  $R^2$ .

The pseudo- $R^2$  reported in Table 1 shows the low percentage of the variation explained

in the first pass regression when using  $\Delta c$ ,  $\text{cay}$ ,  $\Delta c \times \text{cay}$ , as risk factors. In contrast, the pseudo- $R^2$  is large when using the three Fama and French (1993) factors. The low pseudo- $R^2$  indicates that the high cross-sectional  $R^2$  and significant  $t$ -statistic on  $\Delta c \times \text{cay}$  are spurious despite their large values.

The identification issue can also be evaluated by conducting a rank test on  $[\iota_N; \beta]$ . This matrix needs to have full rank in order to identify risk premia in the cross-section regression, yet this full rank condition is violated if betas are (close to) zero, or exhibit little variation. Kleibergen and Zhan (2020) suggest a rank test for testing the reduced rank of  $[\iota_N; \beta]$ , which leads to a large  $p$ -value in Table 1 for the Lettau and Ludvigson (2001) application. This large  $p$ -value implies that we cannot reject the reduced rank of  $[\iota_N; \beta]$  in Lettau and Ludvigson (2001); in contrast, we can do so at the 5% significance level for Fama and French (1993) given the small  $p$ -value reported in Table 1. Such rank test  $p$ -values therefore indicate the identification issue in Lettau and Ludvigson (2001), but much less so in Fama and French (1993). Kleibergen, Kong, and Zhan (2023) further provide the insignificant beta estimates on  $\Delta c \times \text{cay}$ , which help explain why the full rank condition of  $[\iota_N; \beta]$  is possibly violated. For more discussions on rank tests, see, for example, Cragg and Donald (1997), Robin and Smith (2000), and Kleibergen and Paap (2006).

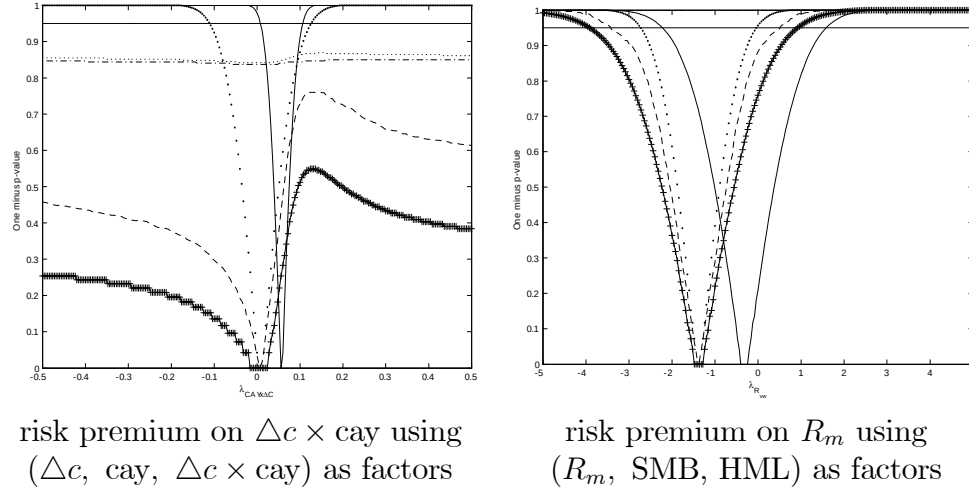
The misleading FM two-pass results illustrated by Table 1 similarly apply to many other papers in asset pricing (see, *e.g.*, Kleibergen (2009), Kleibergen and Zhan (2020)), and partly explain the zoo of risk factors that have been proposed in this literature.

## 2.4 Identification-robust tests on risk premia

Unlike the commonly used  $t$ -test, the so-called identification-robust tests on risk premia do not depend on the quality of risk factors, *i.e.*, they remain valid no matter whether the rank condition for identifying risk premia is satisfied or not. For poor quality risk

factors that jeopardize the rank condition, inverting identification-robust tests lead to wide or even unbounded confidence sets of risk premia, which just reflect that risk premia cannot be precisely identified in this scenario. In contrast, for risk factors of better quality, inverting identification-robust tests can yield tight confidence sets.

Figure 1. One minus  $p$ -value plots for the identification robust: FAR (dotted), CLR (dashed), GLS-LM (solid-plusses), JGLS (dash-dotted); non-robust: Shanken  $t$  (solid), MLE  $t$  (points).



Notes: We refer to Kleibergen (2009) for the identification-robust tests on risk premia used for this figure.

Figure 1 presents an example by inverting the identification-robust tests proposed in Kleibergen (2009). To be in line with Table 1, we use these tests to construct the confidence sets for the risk premium on  $\Delta c \times \text{cay}$  and  $R_m$ , respectively. The adopted FAR, CLR, GLS-LM, JGLS tests are the asset pricing counterparts of identification-robust tests in the GMM framework; see, for example, Stock and Wright (2000), Kleibergen (2002, 2005), and Moreira (2003). For comparison, the outcome based on the non-robust Shanken (1992)  $t$ -test and the MLE (maximum likelihood estimator)  $t$ -test are also presented.

For  $\Delta c \times \text{cay}$ : The left panel of Figure 1 shows that the adopted identification-robust tests do not reject any risk premium at the 5% significance level. Therefore, the result-

ing 95% confidence sets are unbounded. Because Table 1 shows that the Lettau and Ludvigson (2001) application does not pass the rank test, we cannot rule out that the risk premium on  $\Delta c \times \text{cay}$  is unidentified. When the risk premium is unidentified, it is as expected that the identification-robust tests should yield unbounded confidence sets; see, for example, Dufour (1997). The Shanken (1992)  $t$ -test rejects zero risk premium with 5% significance, which is in line with Table 1. However, the  $t$ -test becomes unreliable when the identification is at risk.

For the market factor  $R_m$ : The right panel of Figure 1 shows similar  $p$ -values for identification-robust and non-robust tests, which lead to bounded 95% confidence sets for the market risk premium. This finding is thus in line with the rank test outcome presented in Table 1 for Fama and French (1993). When the risk premium is identified, we expect to get bounded confidence sets from identification-robust tests.

As illustrated by Figure 1 and Table 1, the commonly used  $t$ -test produces bounded confidence sets for risk premia, regardless of the quality of risk factors. Such confidence sets are therefore unreliable, and may overstate the pricing performance of risk factors. For this reason, identification-robust inference methods are recommended over the conventional  $t$ -test.

Gospodinov, Kan, and Robotti (2014) show that their proposed robust GLS  $t$ -test, however, can be used for testing zero risk premia on useless factors; see also Gospodinov and Robotti (2021). While the robust  $t$ -test is reliable in this special case, it does not serve as a uniformly valid test which can be inverted to produce size-correct confidence sets of risk premia.

## 2.5 Inference on risk premia in finite samples

Kleibergen and Zhan (2020) further relate the identification-robust FAR test in Kleibergen (2009) to the Gibbons, Ross, and Shanken (GRS, 1989) test, and develop the GRS-FAR

test for testing hypotheses on risk premia. Like the GRS statistic, the GRS-FAR test statistic has a finite-sample  $F$ -distribution, so it is suitable for small-sample applications where  $T$  is not much larger than  $N$ . By using the GRS-FAR test, Kleibergen and Zhan (2020) find that the confidence sets on the consumption risk premium in C-CAPM are typically unbounded, reflecting that there is little information in commonly used data sets for identifying the consumption risk premium.

Likewise, Kleibergen, Kong, and Zhan (2023) construct the finite-sample counterparts of CLR, GLS-LM, JGLS tests, so that they can be applied to the scenario of small samples encountered in practice. On the other hand, Beaulieu, Dufour, and Khalaf (2013) focus on the zero-beta rate, and propose exact simulation-based procedures for conducting identification-robust inference.

## 2.6 Inference on risk premia of mimicking portfolios

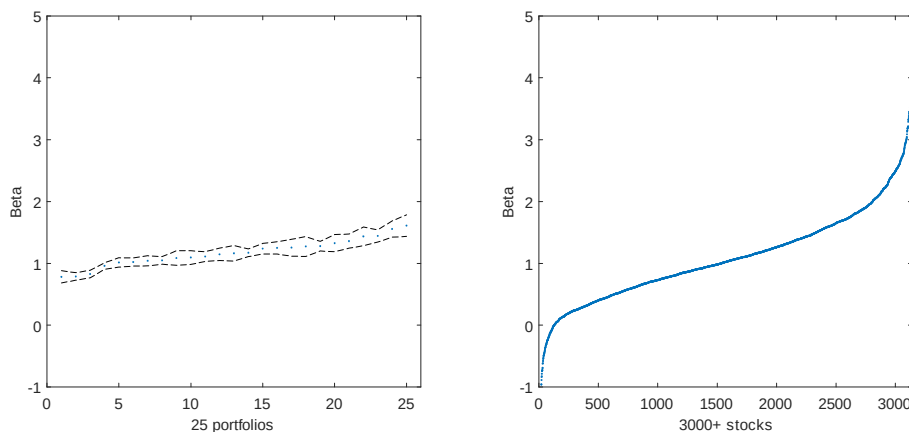
The observed risk factors can be replaced by their mimicking portfolios for inferring risk premia; see, for example, Adrian, Etula, and Muir (2014), who construct mimicking portfolios by regressing their proposed leverage factor on seven base assets: excess returns of the six Fama-French portfolios on size and book-to-market, plus the momentum factor. Kleibergen and Zhan (2018) reveal the consequences of using mimicking portfolios in the FM two-pass procedure. In particular, the estimation errors associated with the weights on mimicking portfolios are taken into account, and an identification-robust test labeled as MPAR (Mimicking Portfolio Anderson-Rubin) test is thus developed.

By inverting the MPAR test, Kleibergen and Zhan (2018) obtain the confidence sets for the risk premium on the mimicking portfolios of the leverage factor. These confidence sets are mostly unbounded, which is consistent with the fact that the leverage factor is a statistically weak factor.

### 3 Risk premia from individual assets

As indicated by (1), identification of factor risk premia relies on the variation of betas. One of the reasons for the poor identification of risk premia is the usage of portfolios, which average over the betas of the individual assets, so there is often too little variation in the betas.

Figure 2. Comparison of market betas estimated using portfolios vs. stocks.



Notes: The blue dots are for the beta estimates, while the dashed black lines in the left panel are for the 95% confidence intervals when portfolios are used. The data of the market factor are taken from Lettau and Ludvigson (2001). For portfolios, we use the twenty-five size and book-to-market sorted portfolios over 1963Q3 to 1998Q3 taken from Lettau and Ludvigson (2001). For stocks, we use over three thousand CRSP stocks which have quarterly returns available during the ten-year time window of 1988Q4 - 1998Q3. We report the estimated market betas sorted in ascending order with respect to the adopted portfolios (left panel) or stocks (right panel).

To illustrate, Figure 2 presents the estimated market betas by using the twenty-five size and book-to-market sorted portfolios and over three thousand individual stocks from the Center for Research in Security Prices (CRSP), respectively. The left panel of Figure 2 shows that the market betas of the adopted portfolios are all close to one, so their cross-sectional variation appears to be limited. This issue is alleviated when using individual assets instead of portfolios, as shown by the right panel of Figure 2. Therefore, Figure 2

suggests that inference on risk premia can be improved by using many individual assets instead of a limited number of portfolios.

For individual assets, it is convenient to use a short panel setup where  $N$  is large and  $T$  is small, to capture time variation of the risk premia. We then care about the ex-post risk premia:

$$\lambda_f = \lambda_{pop,f} + \bar{f} - E(f_t), \quad (6)$$

with  $\bar{f} = \frac{1}{T} \sum_{t=1}^T f_t$ , and  $\lambda_{pop,f}$  the population risk premia; see, for example, Shanken (1992).

Since  $T$  is small,  $\hat{\beta}$  from the first pass of the FM two-pass procedure is not consistent, and similarly the two-pass estimator of  $\lambda_f$ . We therefore have to correct for estimation noise for consistent ex-post risk premia estimation.

### 3.1 Bias correction

Shanken (1992) proposes a bias-adjusted estimator for the ex-post risk premia:

$$\hat{\lambda}_{f,Shanken} = \left( \hat{\beta}' M_{\iota_N} \hat{\beta} - \frac{N}{T} \hat{\omega}^2 \hat{Q}_{FF}^{-1} \right)^{-1} \hat{\beta}' M_{\iota_N} \bar{R}, \quad (7)$$

with  $\hat{\omega}^2$  the estimator of the approximate common variance across assets, and  $\hat{Q}_{FF} = \frac{1}{T} \sum_{t=1}^T \bar{f}_t \bar{f}_t'$ ,  $\bar{f}_t = f_t - \bar{f}$ , the factor covariance (matrix). Raponi, Robotti, and Zaffaroni (2020) construct the asymptotic variance of  $\hat{\lambda}_{f,Shanken}$ , so the  $t$ -test can be conducted.

More recently, Kleibergen and Zhan (2025b) propose the continuous updating estimator (CUE) of  $\lambda_f$ , which can be written as:

$$\hat{\lambda}_{f,CUE} = \left( \hat{\beta}' M_{\iota_N} \hat{\beta} - \mu_{\min} \hat{Q}_{FF}^{-1} \right)^{-1} \hat{\beta}' M_{\iota_N} \bar{R}, \quad (8)$$

with  $\mu_{\min}$  the smallest root of the characteristic polynomial:

$$\left| \mu \begin{pmatrix} 1 & 0 \\ 0 & \hat{Q}_{FF}^{-1} \end{pmatrix} - \begin{pmatrix} \bar{R} & \hat{\beta} \end{pmatrix}' M_{\ell_N} \begin{pmatrix} \bar{R} & \hat{\beta} \end{pmatrix} \right| = 0. \quad (9)$$

For the CUE,  $\hat{\beta}' M_{\ell_N} \hat{\beta} - \mu_{\min} \hat{Q}_{FF}^{-1}$  is guaranteed to be positive definite, because  $\mu_{\min}$  is always exceeded by the smallest eigenvalue of  $\hat{Q}_{FF}^{-1} \hat{\beta}' M_{\ell_N} \hat{\beta}$ . For the Shanken (1992) bias-adjusted estimator, it is, however, not ensured that  $\hat{\beta}' M_{\ell_N} \hat{\beta} - \frac{N}{T} \hat{\omega}^2 \hat{Q}_{FF}^{-1}$  is positive definite, so Raponi, Robotti, and Zaffaroni (2020) apply a shrinkage procedure when invertibility is at stake.

The simulation study in Kleibergen and Zhan (2025b) shows that the CUE can outperform both the Shanken (1992) bias-adjusted estimator and the IV estimator (see, *e.g.*, Jegadeesh et al. (2019)) for addressing the bias in  $\hat{\beta}$  due to short panels. In addition, Kleibergen and Zhan (2025b) develop the LM test based on the score of the objective function of CUE, which can be used for conducting identification-robust inference on the ex-post risk premia.

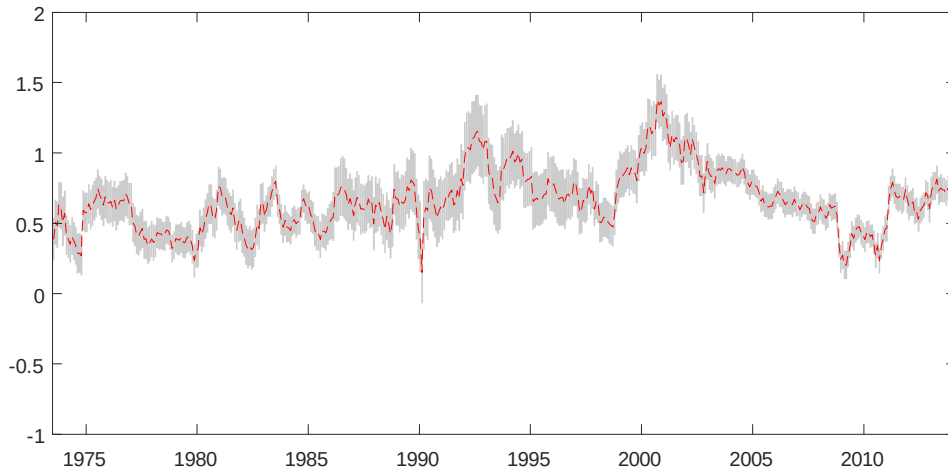
### 3.2 Identification-robust LM test in short panels

The LM test proposed by Kleibergen and Zhan (2025b) is an identification-robust test on the ex-post risk premia. Since we have many assets, the beta pricing moment condition will only hold approximately, *i.e.*, there will be misspecification because of unobserved priced risk factors. The object of interest is then the pseudo-true value which is the minimizer of the CUE population objective function. The LM test is, however, also robust to misspecification, so it can be used for constructing confidence sets of the ex-post risk premia, which contain the estimate of the CUE.

Resulting from the LM test, Figure 3 presents the time-varying 95% confidence sets

of the market risk premium in the Fama and French (2018) six-factor model, where the CUE (dashed red) is also provided. The test assets are the individual stocks taken from CRSP. Because of the large number of individual assets, the confidence sets from the LM test for the market risk premium are tight, which therefore show the improvement in inference. These confidence sets contain mostly positive values, reflecting that the market risk premium is significantly positive. Instead of the Fama and French (2018) six-factor model, similar findings for the market risk premium in the CAPM and Fama and French (1993) three-factor model are presented in Kleibergen and Zhan (2025b).

Figure 3. Market risk premium in the Fama and French (2018) six-factor model from the LM test using CRSP stocks.



Notes: This figure is taken from Kleibergen and Zhan (2025b). The shaded areas are the 95% confidence sets of the market risk premium from the LM test based on the CUE (dashed red) in the Fama and French (2018) six-factor model. The ten-year rolling window is used. We refer to the published article for the details of the adopted monthly data.

Recall that when portfolios are used instead of individual stocks, Table 1 shows that the market risk premium is positive but insignificant. Figure 3 therefore shows that identification of risk premia improves when using the cross-section of individual assets compared to portfolios, which is consistent with the variation of betas presented in Figure

2.

Other than the market factor, Kleibergen and Zhan (2025b) use the identification-robust LM test to show that a few other risk factors also matter, including the size and momentum factors, while other well-established factors (such as value, investment, profitability) do much less so; see also Gagliardini, Ossola, and Scaillet (2016). More research is therefore needed to determine which risk factors stand out in the cross-section of individual asset returns.

## 4 Misspecification and weak identification

A major concern in asset pricing is that model misspecification is generally in place. For beta pricing in (1), because betas of the specified risk factors are unlikely to fully explain the cross-section of asset returns, (1) is rewritten as:

$$E(R_{i,t}) = \lambda_0 + \beta_i' \lambda_f + e_i, \quad (10)$$

where  $e_i$  is the model misspecification error. The misspecification issue similarly applies to the GMM setting, where the moment equations may not exactly hold. See, for example, Kan, Robotti, and Shanken (2013), Gospodinov, Kan, and Robotti (2014).

### 4.1 Identification under misspecification

The severity of misspecification described above is often evaluated by using the  $J$ -test in the GMM framework:

$$\text{MISS} = J = \min Q_{CUE}, \quad (11)$$

with  $Q_{CUE}$  the continuous updating objective function using beta pricing moment conditions.

Correspondingly, we use IS as the population counterpart of the Cragg and Donald (1997) rank statistic which tests the full rank condition for beta pricing. The explicit expressions for IS and MISS statistics are provided in the Appendix.

Kleibergen and Zhan (2026) show that we need

$$IS > MISS \quad (12)$$

for the pseudo-true value of risk premia to be identified. The identification condition therefore turns out to be more stringent in misspecified models, while for correctly specified models, the condition is just  $IS > 0$ .

Table 2: **MISS and IS statistics** (\*:10%, \*\*:5%, \*\*\*:1%)

|     | (A) Impose $\lambda_0 = 0$ : No |           | (B) Impose $\lambda_0 = 0$ : Yes |           |
|-----|---------------------------------|-----------|----------------------------------|-----------|
|     | MISS                            | IS        | MISS                             | IS        |
| FF  | 59.34***                        | 106.81*** | 87.47***                         | 974.39*** |
| JW  | 75.07                           | 103.54    | 86.46                            | 103.56    |
| LL  | 31.11*                          | 31.75*    | 37.15**                          | 40.90**   |
| Y   | 17.14                           | 17.34     | 19.42                            | 19.60     |
| S   | 134.27***                       | 140.68*** | 268.60***                        | 296.78*** |
| AEM | 28.42                           | 31.97     | 30.41                            | 42.03**   |
| K   | 59.84***                        | 78.47***  | 60.03***                         | 102.77*** |
| HKM | 35.32**                         | 35.88**   | 44.44***                         | 59.74***  |

Notes: This table is taken from Kleibergen and Zhan (2026). FF, JW, LL, Y, S, AEM, K, HKM are eight asset pricing models, for which we report their IS and MISS statistics, with and without the  $\lambda_0 = 0$  restriction. We refer to Kleibergen and Zhan (2026) for further details.

Table 2 presents the MISS and IS statistics for eight asset pricing studies. The MISS statistics indicate that many models are misspecified, while the IS statistics indicate identification. It is, however,  $IS - MISS$  which is indicative of identification, not just IS.

Identification on the pseudo-true values of the risk premia can be tested using  $IS - MISS$ , which provides a quasi-likelihood ratio statistic to test no-identification. Kleibergen

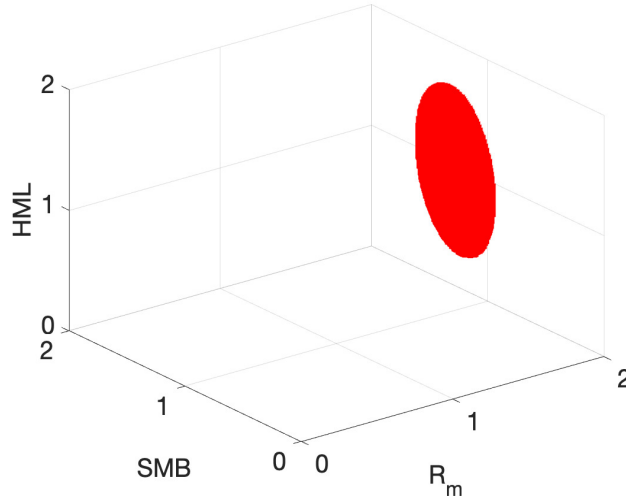
and Zhan (2026) show how to construct critical value functions for such a test. Because MISS and IS are often close as we observe in Table 2, weak identification and misspecification likely occur simultaneously, so they need to be jointly addressed.

## 4.2 Double robust inference

The double robust inference for asset pricing aims to take both weak identification and misspecification into account, so it provides valid inference on the pseudo-true value of the parameter of interest regardless of the strength of identification and/or misspecification.

For beta pricing with many individual assets so  $N \gg T$ , the LM test in Kleibergen and Zhan (2025b) is doubly robust, and Figure 3 has shown that it can produce informative confidence sets for the pseudo-true value of ex-post risk premia. In a similar fashion, for beta pricing with a limited number of portfolios so  $T \gg N$ , the DRLM (Double Robust LM) test in Kleibergen and Zhan (2025a) is applicable, as shown by Figure 4.

Figure 4. Joint 95% confidence set for risk premia by DRLM.



Notes: The shaded area is the joint 95% confidence set of the risk premia from the DRLM test in the Fama and French (1993) three-factor model. The adopted data of twenty-five Fama-French portfolios are over 1963Q3 - 2013Q4, and the  $\lambda_0 = 0$  restriction is imposed.

Figure 4 presents the joint 95% DRLM confidence set for the risk premia on the three Fama and French (1993) factors, for which the twenty-five Fama-French portfolios are used as test assets. Projecting the joint confidence set on each axis produces the confidence interval for each risk premium, which is significantly positive for all the three factors.

### 4.3 Nonlinear GMM

It is worth noting that the DRLM test can also be applied to more general GMM settings with, for example, nonlinear moment conditions that result from the stochastic discount factor framework for asset pricing:

$$E \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} (\iota_N + R_{t+1}) - \iota_N \right] = 0, \quad (13)$$

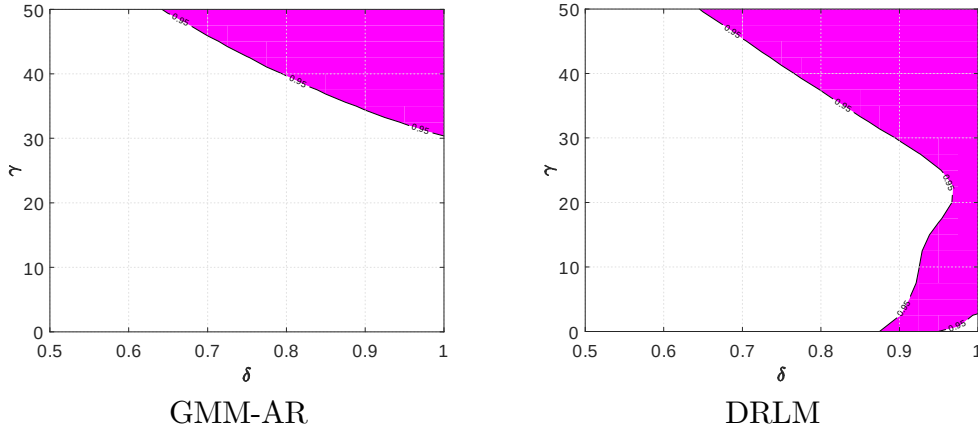
with  $\delta$  the discount factor,  $\gamma$  the relative rate of risk aversion,  $C_t$  consumption at time  $t$ ,  $R_{t+1}$  the  $N$ -dimensional vector of asset returns, and  $\iota_N$  the  $N$ -dimensional vector of ones; see, for example, Hansen and Singleton (1982).

Misspecification in GMM occurs when moment conditions like (13) do not hold exactly, while weak identification occurs if the Jacobian of the moment function is close to reduced rank. These two issues are jointly taken into account by the DRLM test, while many other identification-robust tests such as the GMM-AR test do not account for misspecification.

Figure 5 presents the joint 95% confidence set for the discount rate and risk aversion based on the moment condition (13). The left panel of Figure 5 shows that the GMM-AR test leads to large values for  $\gamma$ , which are generally considered unrealistic. In contrast, the right panel of Figure 5 shows the joint confidence set of  $(\delta, \gamma)$  by the DRLM test. It leads to values of  $\gamma$  that are economically more reasonable, *i.e.*, the 95% confidence set resulting from the DRLM test contains much smaller values of  $\gamma$ . This difference occurs, because the DRLM test allows for misspecification while the GMM-AR test does not. When we

account for misspecification by using the DRLM test, the additional uncertainty leads to a relatively larger confidence set for  $(\delta, \gamma)$ , which therefore contains economically meaningful values.

Figure 5. 95% confidence regions of  $(\delta, \gamma)$  from identification-robust tests.



Notes: This figure is taken from Kleibergen and Zhan (2025a).  $\delta$ , the discount factor;  $\gamma$ , the relative rate of risk aversion. We focus on  $\delta \in [0.5, 1]$ ,  $\gamma \in [0, 50]$ . The shaded areas contain the values of  $(\delta, \gamma)$  that are not rejected at the 5% level by the GMM-AR and DRLM tests, respectively. The six portfolios sorted by size and book-to-market are used for test assets. We refer to the published article for further details.

## 5 Conclusions

Many risk premia documented in empirical asset pricing studies are probably not well identified. Identification-robust testing procedures should therefore generically be used to test hypotheses on risk premia. Identification of risk premia is greatly improved when using a large number of individual assets compared to the default of a small number of portfolios. Therefore, more informative confidence sets of risk premia can be constructed by inverting identification-robust tests under individual assets. In addition, since betas

do not fully span the cross-section of asset returns, procedures robust to both misspecification and weak identification should be employed. The same argument applies to the GMM framework, where the identification condition in misspecified models is more stringent than in correctly specified models, so satisfying the rank condition alone does not necessarily imply strong identification of the parameter of interest. Doubly robust inference procedures are available for both the two-pass procedure and GMM.

We conclude by making three practical suggestions. First, if practitioners continue to use the traditional two-pass procedure and GMM methodology for asset pricing, then an accompanying identification test needs to be provided. Empirical findings based on the traditional methodology without the evidence of identification should be interpreted with caution. Second, we recommend using doubly robust inference procedures, since both weak identification and misspecification are prevalent in asset pricing studies. Third, depending on the empirical asset pricing settings of portfolios vs. individual assets, linear vs. nonlinear moment conditions, the implementation of doubly robust inference procedures differs, for which we provide further details in the Appendix.

## References

- [1] Adrian, T., E. Etula and T. Muir. Financial Intermediaries and the Cross-Section of Asset Returns. *Journal of Finance*, **69**:2557–2596, 2014.
- [2] Andrews, D.W.K. and X. Cheng. Estimation and inference with Weak, Semi-Strong and Strong Identification. *Econometrica*, **80**:2153–2211, 2012.
- [3] Andrews, I. and A. Mikusheva. A Geometric Approach to Nonlinear Econometric Models. *Econometrica*, **84**:1249–1264, 2016a.
- [4] Andrews, I. and A. Mikusheva. Conditional inference with a functional nuisance parameter. *Econometrica*, **84**:1571–1612, 2016b.
- [5] Beaulieu, M.-C., J.-M. Dufour and L. Khalaf. Identification-Robust Estimation and Testing of the Zero-Beta CAPM. *Review of Economic Studies*, **80**:892–924, 2013.
- [6] Cragg, J.C. and S.G. Donald. Inferring the rank of a matrix. *Journal of Econometrics*, **76**:223–250, 1997.
- [7] Dufour, J.-M. Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models. *Econometrica*, **65**:1365–388, 1997.
- [8] Fama, E.F. and J.D. MacBeth. Risk, Return and Equilibrium: Empirical Tests. *Journal of Political Economy*, **81**:607–636, 1973.
- [9] Fama, E.F. and K.R. French. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, **33**:3–56, 1993.
- [10] Fama, E.F. and K.R. French. Choosing factors. *Journal of Financial Economics*, **128**(2):234–252, 2018.
- [11] Gagliardini, P. and Ossola, E. and Scaillet, O. Time-varying risk premium in large cross-sectional equity data sets. *Econometrica*, **84**(3):985–1046, 2016.
- [12] Gibbons, M., S. Ross and J. Shanken. A Test of the Efficiency of a Given Portfolio. *Econometrica*, **57**:1121–1152, 1989.

- [13] Gospodinov, N. and Robotti, C. Common pricing across asset classes: Empirical evidence revisited. *Journal of Financial Economics*, 140(1):292–324, 2021.
- [14] Gospodinov, N. R. Kan and C. Robotti. Misspecification-Robust Inference in Linear Asset-Pricing Models with Irrelevant Factors. *Review of Financial Studies*, **27**:2139–2170, 2014.
- [15] Han, S. and A. McCloskey. Estimation and inference with a (nearly) singular jacobian. *Quantitative Economics*, 10(3):1019–1068, 2019.
- [16] Hansen, L.P. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, **50**:1029–1054, 1982.
- [17] Hansen, L.P. and K. Singleton. Generalized Instrumental Variable Estimation of Nonlinear Rational Expectations Models. *Econometrica*, **50**:1269–1286, 1982.
- [18] Harvey, C. R. and Liu, Y. and Zhu, H. ... and the cross-section of expected returns. *Review of Financial Studies*, 29(1):5–68, 2016.
- [19] Jegadeesh, N. and Noh, J. and Pukthuanthong, K. and Roll, R. and Wang, J. Empirical tests of asset pricing models with individual assets: Resolving the errors-in-variables bias in risk premium estimation. *Journal of Financial Economics*, 133(2):273–298, 2019.
- [20] Kan, R. and C. Zhang. Two-Pass Tests of Asset Pricing Models with Useless Factors. *Journal of Finance*, **54**:203–235, 1999.
- [21] Kan, R., C. Robotti and J. Shanken. Pricing Model Performance and the Two-Pass Cross-Sectional Regression Methodology. *Journal of Finance*, **68**:2617–2649, 2013.
- [22] Kleibergen, F. Pivotal Statistics for testing Structural Parameters in Instrumental Variables Regression. *Econometrica*, **70**:1781–1803, 2002.
- [23] Kleibergen, F. Testing Parameters in GMM without assuming that they are identified. *Econometrica*, **73**:1103–1124, 2005.
- [24] Kleibergen, F. Tests of Risk Premia in Linear Factor Models. *Journal of Econometrics*, **149**:149–173, 2009.

- [25] Kleibergen, F. and R. Paap. Generalized Reduced Rank Tests using the Singular Value Decomposition. *Journal of Econometrics*, **133**:97–126, 2006.
- [26] Kleibergen F. and Z. Zhan. Unexplained factors and their effects on second pass R-squared’s. *Journal of Econometrics*, **189**:101–116, 2015.
- [27] Kleibergen, F. and Z. Zhan. Identification-robust inference on risk premia of mimicking portfolios of non-traded factors. *Journal of Financial Econometrics*, 16(2):155–190, 2018.
- [28] Kleibergen, F. and Z. Zhan. Robust Inference for Consumption-Based Asset Pricing. *Journal of Finance*, **75**:507–550, 2020.
- [29] Kleibergen, F. and Z. Zhan. Double robust inference for continuous updating gmm. *Quantitative Economics*, 16(1):295–327, 2025a.
- [30] Kleibergen, F. and Z. Zhan. Risk premia from the cross-section of individual assets. *Journal of Econometrics*, 252:106108, 2025b.
- [31] Kleibergen, F. and Z. Zhan. Testing for identification in potentially misspecified linear GMM. *Unpublished Working paper, University of Amsterdam*, 2026.
- [32] Kleibergen, F., L. Kong and Z. Zhan. Identification robust testing of risk premia in finite samples. *Journal of Financial Econometrics*, 21(2):263–297, 2023.
- [33] Lettau, M. and S. Ludvigson. Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia are Time-Varying. *Journal of Political Economy*, **109**:1238–1287, 2001.
- [34] Lewellen, J., S. Nagel and J. Shanken. A Skeptical Appraisal of Asset-Pricing Tests. *Journal of Financial Economics*, **96**:175–194, 2010.
- [35] Moreira, M.J.,. A Conditional Likelihood Ratio Test for Structural Models. *Econometrica*, **71**:1027–1048, 2003.
- [36] Raponi, V., Robotti, C. and Zaffaroni, P. Testing beta-pricing models using large cross-sections. *Review of Financial Studies*, 33(6):2796–2842, 2020.

- [37] Robin, J.-M. and R.J. Smith. Tests of Rank. *Econometric Theory*, **16**:151–175, 2000.
- [38] Shanken, J. On the Estimation of Beta-Pricing Models. *Review of Financial Studies*, **5**:1–33, 1992.
- [39] Staiger, D. and J.H. Stock. Instrumental Variables Regression with Weak Instruments. *Econometrica*, **65**:557–586, 1997.
- [40] Stock, J.H. and J.H. Wright. GMM with Weak Identification. *Econometrica*, **68**:1055–1096, 2000.
- [41] Stock, J.H., J.H. Wright and M. Yogo. A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business and Economic Statistics*, **20**:518–529, 2002.

## Appendix

In this appendix, we continue to use the following notation:  $R_{i,t}$  is the return on the  $i$ -th asset at time  $t$ ,  $i = 1, \dots, N$ ,  $t = 1, \dots, T$ ;  $f_t$  is the  $k \times 1$  vector of risk factors, whose sample covariance is  $\hat{Q}_{FF} = \frac{1}{T} \sum_{t=1}^T \bar{f}_t \bar{f}_t'$  with  $\bar{f}_t = f_t - \bar{f}$ ,  $\bar{f} = \frac{1}{T} \sum_{t=1}^T f_t$ ;  $\beta_i = Var(f_t)^{-1} Cov(f_t, R_{i,t})$  is the  $k$ -dimensional vector of factor loadings,  $\beta = (\beta_1 \dots \beta_N)'$  is the  $N \times k$  dimensional matrix with the least squares estimator  $\hat{\beta} = (\hat{\beta}_1 \dots \hat{\beta}_N)'$ .

### A1: Pseudo- $R^2$

The pseudo- $R^2$  is defined as:

$$\text{pseudo-}R^2 = 1 - \frac{\sum_{i=1}^N \lambda_{i,residual}}{\sum_{i=1}^N \lambda_{i,asset}}$$

where  $\lambda_{i,asset}$  is the  $i$ -th characteristic root of the  $N \times N$  covariance matrix of asset returns, and  $\lambda_{i,residual}$  is the  $i$ -th characteristic root of the  $N \times N$  covariance matrix of residuals in the time-series regression of asset returns on risk factors.

### A2: Rank test

By removing the  $N$ -th asset return, we obtain the  $(N-1) \times 1$  vector  $\mathbf{R}_t$  and the  $(N-1) \times k$  dimensional matrix  $\mathbf{B}$ :

$$\mathbf{R}_t = (R_{1,t} \dots R_{(N-1),t})' - \iota_{N-1} R_{N,t}, \quad \mathbf{B} = (\beta_1 \dots \beta_{N-1})' - \iota_{N-1} \beta_N'.$$

The full rank condition of  $[\iota_N; \beta]$  is the same as the full rank of  $\mathbf{B}$ .

Estimation of the auxiliary regression model  $\mathbf{R}_t = \boldsymbol{\alpha} + \mathbf{B}f_t + \mathbf{u}_t$  yields

$$\hat{\mathbf{B}} = \sum_{t=1}^T \bar{\mathbf{R}}_t \bar{f}_t' \left( \sum_{t=1}^T \bar{f}_t \bar{f}_t' \right)^{-1}, \quad \hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T \hat{\mathbf{u}}_t \hat{\mathbf{u}}_t',$$

with  $\bar{\mathbf{R}}_t = \mathbf{R}_t - \bar{\mathbf{R}}$ ,  $\bar{\mathbf{R}} = \frac{1}{T} \sum_{t=1}^T \mathbf{R}_t$ ,  $\hat{\mathbf{u}}_t = (\mathbf{R}_t - \bar{\mathbf{R}}) - \hat{\mathbf{B}}(f_t - \bar{f})$  is the residual at time  $t$ .

Let  $rk$  be the smallest root of

$$\left| \mu \hat{Q}_{FF}^{-1} - \hat{\mathbf{B}}' \hat{\Sigma}^{-1} \hat{\mathbf{B}} \right| = 0.$$

The rank test  $F$ -statistic for testing  $H_0 : \text{rank}(\mathbf{B}) = k - 1$  is

$$\frac{T-N}{N-k} \times \text{rk} \preceq F(N-k, T-N).$$

When  $T$  is large, the corresponding Wald rank statistic (also the IS statistic) reads:

$$\text{IS} = T \times rk \preceq \chi_{N-k}^2.$$

### A3: Identification-robust tests on risk premia

Consider the CUE objective function for  $H_0 : \lambda_f = \lambda_{f,0}$ :

$$Q_{CUE}(\lambda_{f,0}) = \frac{T}{1 + \lambda_{f,0}' \hat{Q}_{FF}^{-1} \lambda_{f,0}} (\bar{\mathbf{R}} - \hat{\mathbf{B}} \lambda_{f,0})' \hat{\Sigma}^{-1} (\bar{\mathbf{R}} - \hat{\mathbf{B}} \lambda_{f,0}),$$

which is related to the characteristic polynomial:

$$\left| \mu \begin{pmatrix} 1 & 0 \\ 0 & \hat{Q}_{FF}^{-1} \end{pmatrix} - \begin{pmatrix} \bar{\mathbf{R}} & \hat{\mathbf{B}} \end{pmatrix}' \hat{\Sigma}^{-1} \begin{pmatrix} \bar{\mathbf{R}} & \hat{\mathbf{B}} \end{pmatrix} \right| = 0.$$

The minimal value of  $Q_{CUE}(\lambda_{f,0})$  (also the MISS =  $J$  statistic) corresponds with  $T$  times the smallest root of the above characteristic polynomial:

$$\text{MISS} = J = T \times \mu_{\min}.$$

The GRS-FAR statistic adjusts  $Q_{CUE}(\lambda_{f,0})$  for an  $F$ -distributed random variable:

$$\frac{T-k-N+1}{T(N-1)} \times Q_{CUE}(\lambda_{f,0}) \sim F(N-1, T-k-N+1),$$

which can be inverted to produce confidence sets of  $\lambda_f$ ; see Kleibergen and Zhan (2020).

The FAR test is the asymptotic counterpart of GRS-FAR. As  $T$  goes large,

$$\text{FAR}(\lambda_{f,0}) = \frac{T}{1+\lambda'_{f,0}\hat{Q}_{FF}^{-1}\lambda_{f,0}}(\bar{\mathbf{R}} - \hat{\mathbf{B}}\lambda_{f,0})'\hat{\Sigma}^{-1}(\bar{\mathbf{R}} - \hat{\mathbf{B}}\lambda_{f,0}) \xrightarrow{d} \chi^2_{N-1},$$

where Kleibergen (2009) used  $\hat{\Sigma} = \frac{1}{T-k-1} \sum_{t=1}^T \hat{\mathbf{u}}_t \hat{\mathbf{u}}_t'$ .

One dimensional confidence sets can be obtained by using the subset GRS-FAR test (similarly, the subset FAR test). Consider (without loss of generality) the partition of  $\lambda_f = (\lambda_{f1} \dots \lambda_{fk})' = (\lambda_{f1}, \lambda'_2)'$ , where  $\lambda_2 = (\lambda_{f2} \dots \lambda_{fk})'$ . When  $H_0 : \lambda_{f1} = \lambda_{f1}^0$  and  $\lambda_2$  are left unspecified, the subset GRS-FAR (sGRS-FAR) statistic for testing  $H_0 : \lambda_{f1} = \lambda_{f1}^0$  results from the characteristic polynomial:

$$\left| \mu \begin{pmatrix} 1 + \lambda'_{f1} \hat{Q}_{FF,11}^{-1} \lambda_{f1} & -\lambda'_{f1} \hat{Q}_{FF,12}^{-1} \\ -\hat{Q}_{FF,21}^{-1} \lambda_{f1} & \hat{Q}_{FF,22}^{-1} \end{pmatrix} - \begin{pmatrix} \bar{\mathbf{R}} - \hat{B}_1 \lambda_{f1} & \hat{B}_2 \end{pmatrix}' \hat{\Sigma}^{-1} \begin{pmatrix} \bar{\mathbf{R}} - \hat{B}_1 \lambda_{f1} & \hat{B}_2 \end{pmatrix} \right| = 0,$$

when  $\hat{Q}_{FF}^{-1} = \begin{pmatrix} \hat{Q}_{FF,11}^{-1} & \hat{Q}_{FF,12}^{-1} \\ \hat{Q}_{FF,21}^{-1} & \hat{Q}_{FF,22}^{-1} \end{pmatrix}$ ,  $\hat{\mathbf{B}} = (\hat{B}_1, \hat{B}_2)$  and equals  $T$  times the smallest root.

Under the i.i.d. setting, the distribution of the sGRS-FAR statistic is in the limit bounded by the  $\chi^2_{N-k}$  distribution,

$$\text{sGRS-FAR}(\lambda_{f1}) \preceq \chi^2_{N-k}.$$

In finite samples, the distribution of the sGRS-FAR statistic is bounded according to:

$$\frac{T-N}{T(N-k)} \times \text{sGRS-FAR}(\lambda_{f1}) \preceq F(N-k, T-N).$$

Kleibergen (2009) further decomposes the FAR statistic into the GLS-LM statistic and the JGLS statistic, both can be used for testing  $H_0 : \lambda_f = \lambda_{f,0}$ :

$$\text{FAR}(\lambda_{f,0}) = \text{GLS-LM}(\lambda_{f,0}) + \text{JGLS}(\lambda_{f,0}).$$

The FAR statistic can be rewritten as:

$$\text{FAR}(\lambda_{f,0}) = \frac{T}{1 - \lambda'_{f,0} \widehat{\mathcal{Q}}_{FF}^{-1}(\lambda_{f,0}) \lambda_{f,0}} \left( \overline{\mathbf{R}} - \widetilde{\mathbf{B}} \lambda_{f,0} \right)' \widehat{\Sigma}^{-1} \left( \overline{\mathbf{R}} - \widetilde{\mathbf{B}} \lambda_{f,0} \right)$$

with

$$\widetilde{\mathbf{B}} = \sum_{t=1}^T \mathbf{R}_t (\bar{f}_t + \lambda_{f,0})' \left[ \sum_{t=1}^T (\bar{f}_t + \lambda_{f,0})(\bar{f}_t + \lambda_{f,0})' \right]^{-1}$$

and  $\widehat{\mathcal{Q}}_{FF}(\lambda_{f,0}) = \frac{1}{T} \sum_{t=1}^T (\bar{f}_t + \lambda_{f,0})(\bar{f}_t + \lambda_{f,0})'$ .

The GLS-LM statistic results from projecting  $\widehat{\Sigma}^{-\frac{1}{2}} \left( \overline{\mathbf{R}} - \widetilde{\mathbf{B}} \lambda_{f,0} \right)$  on the column space spanned by  $\widehat{\Sigma}^{-\frac{1}{2}} \widetilde{\mathbf{B}}$ , so

$$\begin{aligned} & \text{GLS-LM}(\lambda_{f,0}) \\ &= \frac{T}{1 - \lambda'_{f,0} \widehat{\mathcal{Q}}_{FF}^{-1}(\lambda_{f,0}) \lambda_{f,0}} \left( \overline{\mathbf{R}} - \widetilde{\mathbf{B}} \lambda_{f,0} \right)' \widehat{\Sigma}^{-\frac{1}{2}'} P_{\widehat{\Sigma}^{-\frac{1}{2}} \widetilde{\mathbf{B}}} \widehat{\Sigma}^{-\frac{1}{2}} \left( \overline{\mathbf{R}} - \widetilde{\mathbf{B}} \lambda_{f,0} \right) \xrightarrow{d} \chi_k^2 \end{aligned}$$

where  $P_{\widehat{\Sigma}^{-\frac{1}{2}} \widetilde{\mathbf{B}}} = \widehat{\Sigma}^{-\frac{1}{2}} \widetilde{\mathbf{B}} \left( \widetilde{\mathbf{B}}' \widehat{\Sigma}^{-1} \widetilde{\mathbf{B}} \right)^{-1} \widetilde{\mathbf{B}}' \widehat{\Sigma}^{-\frac{1}{2}'}$ ,  $\widehat{\Sigma}^{-1} = \widehat{\Sigma}^{-\frac{1}{2}'} \widehat{\Sigma}^{-\frac{1}{2}}$ . Thus, the GLS-LM statistic can also be viewed as a quadratic form of the derivative of the FAR statistic.

Similarly, the JGLS statistic results from projecting  $\widehat{\Sigma}^{-\frac{1}{2}} \left( \overline{\mathbf{R}} - \widetilde{\mathbf{B}} \lambda_{f,0} \right)$  on the space orthogonal to the one spanned by  $\widehat{\Sigma}^{-\frac{1}{2}} \widetilde{\mathbf{B}}$ , so  $\text{JGLS}(\lambda_{f,0}) = \text{FAR}(\lambda_{f,0}) - \text{GLS-LM}(\lambda_{f,0}) \xrightarrow{d}$

$$\chi_{N-k-1}^2.$$

The conditional likelihood ratio (CLR) test resembles the CLR test in Moreira (2003) for linear instrumental variables regression. In the single factor case, the CLR statistic has a closed form expression:

$$\begin{aligned} & \text{CLR}(\lambda_{f,0}) \\ &= \frac{1}{2} \left\{ \text{FAR}(\lambda_{f,0}) - r(\lambda_{f,0}) + \sqrt{(\text{FAR}(\lambda_{f,0}) + r(\lambda_{f,0}))^2 - 4r(\lambda_{f,0})\text{JGLS}(\lambda_{f,0})} \right\} \end{aligned}$$

where  $r(\lambda_{f,0}) = \widehat{\mathbf{Q}}_{FF}(\lambda_{f,0})\widetilde{\mathbf{B}}'\widehat{\Sigma}^{-1}\widetilde{\mathbf{B}}$ . The conditional critical values of the CLR statistic can be simulated in the manner of Moreira (2003).

For CLR, GLS-LM and JGLS above, Kleibergen, Kong, and Zhan (2023) discuss their finite-sample counterparts. On the other hand, Kleibergen and Zhan (2018) provide the MPAR test for mimicking portfolios. Because these tests are constructed along the same lines as FAR and GRS-FAR, we refer readers to the corresponding papers for further details.

#### A4: Doubly robust LM test for short panels

Consider the setting of short panels with  $N \gg T$ . Let  $\lambda_f^*$  be the pseudo-true value of ex-post risk premia, which coincides the ex-post risk premia  $\lambda_f$  under correct specification.

The LM test statistic for testing  $H_0 : \lambda_f^* = \lambda_{f,0}^*$  reads:

$$\text{LM}(\lambda_{f,0}^*) = \frac{N}{T} s^*(\lambda_{f,0}^*)' \hat{V}_s^*(\lambda_{f,0}^*)^{-1} s^*(\lambda_{f,0}^*) \xrightarrow{d} \chi_k^2$$

with  $s^*(\lambda_f) = \frac{1}{N} \sum_{i=1}^N s_i^*(\lambda_f) = \frac{1}{N} \sum_{i=1}^N \tilde{\beta}_i(\lambda_f)(\bar{R}_i - \lambda_f' \bar{\beta}_i)$ ,  $\hat{V}_s^*(\lambda_f) = \frac{1}{N} \sum_{i=1}^N (s_i^*(\lambda_f) - s^*(\lambda_f))$

$(s_i^*(\lambda_f) - s^*(\lambda_f))'$ , and  $\bar{\beta}_i = \hat{\beta}_i - \frac{1}{N} \sum_{i=1}^N \hat{\beta}_i$ ,  $R_t = (R_{1,t} \dots R_{N,t})'$ ,  $\bar{R} = \frac{1}{T} \sum_{t=1}^T R_t = (\bar{R}_1 \dots \bar{R}_N)'$ ,  $\bar{R}_i = \bar{R}_i - \frac{1}{N} \sum_{i=1}^N \bar{R}_i$ ,  $\tilde{\beta}_i(\lambda_f) = \bar{\beta}_i + \hat{Q}_{FF}^{-1} \lambda_f \frac{(\bar{R}_i - \lambda_f' \bar{\beta}_i)}{1 + \lambda_f' \hat{Q}_{FF}^{-1} \lambda_f}$ .

For testing a single ex-post risk premium in a multi-factor model, consider the partition of  $\lambda_f = (\lambda_{f1} \dots \lambda_{fk})' = (\lambda_{f1}, \lambda_2')'$ , where  $\lambda_2 = (\lambda_{f2} \dots \lambda_{fk})'$ . For testing  $H_0 : \lambda_{f1}^* = \lambda_{f1,0}^*$ , the resulting subset LM test statistic  $LM(\lambda_{f1,0})$  reads:

$$LM(\lambda_{f1,0}^*) = \frac{N}{T} s^*((\lambda_{f1,0}^*, \hat{\lambda}_2(\lambda_{f1,0}^*))' )' \hat{V}_s^*((\lambda_{f1,0}^*, \hat{\lambda}_2(\lambda_{f1,0}^*))' )^{-1} s^*((\lambda_{f1,0}^*, \hat{\lambda}_2(\lambda_{f1,0}^*))' )' \\ \xrightarrow{d} \chi_1^2,$$

with  $\hat{\lambda}_2(\lambda_{f1,0}^*)$  the restricted CUE of  $\lambda_2$  under  $H_0 : \lambda_{f1}^* = \lambda_{f1,0}^*$ :

$$\hat{\lambda}_{f2}(\lambda_{f1,0}^*) = \left( \hat{B}_2' M_{\ell_N} \hat{B}_2 - \mu_{\min} \hat{Q}_{FF,22}^{-1} \right)^{-1} (\hat{B}_2' M_{\ell_N} (\bar{R} - \hat{B}_1 \lambda_{f1,0}^*) + \mu_{\min} \hat{Q}_{FF,21}^{-1} \lambda_{f1,0}^*),$$

where  $\mu_{\min}$  is the smallest root of the characteristic polynomial:

$$\left| \mu \begin{pmatrix} 1 + \lambda_{f1,0}^{*'} \hat{Q}_{FF,11}^{-1} \lambda_{f1,0}^* & -\lambda_{f1,0}^{*'} \hat{Q}_{FF,12}^{-1} \\ -\hat{Q}_{FF,21}^{-1} \lambda_{f1,0}^* & \hat{Q}_{FF,22}^{-1} \end{pmatrix} - (\bar{R} - \hat{B}_1 \lambda_{f1,0}^* : \hat{B}_2)' M_{\ell_N} (\bar{R} - \hat{B}_1 \lambda_{f1,0}^* : \hat{B}_2) \right| \\ = 0,$$

for  $\hat{Q}_{FF}^{-1} = \begin{pmatrix} \hat{Q}_{FF,11}^{-1} & \hat{Q}_{FF,12}^{-1} \\ \hat{Q}_{FF,21}^{-1} & \hat{Q}_{FF,22}^{-1} \end{pmatrix}$ ,  $\hat{\beta} = (\hat{B}_1, \hat{B}_2)$  corresponding to the partition  $\lambda_f = (\lambda_{f1}, \lambda_2')'$ .

## A5: DRLM for GMM

The DRLM test is proposed for the over-identified GMM setting with linear or nonlinear moment conditions. Consider the  $m \times 1$  parameter vector  $\theta = (\theta_1 \dots \theta_m)'$ , for which the moment function  $f(\theta, X_t)$  leads to the GMM population moment function:

$$E_X(f(\theta, X_t)) = \mu_f(\theta).$$

Let  $V(\theta^*)$  be the asymptotic covariance matrix of the sample moment  $f_T(\theta) = \frac{1}{T} \sum_{t=1}^T f(\theta, X_t)$  and its derivative  $q_T(\theta) = \frac{\partial f_T(\theta)}{\partial \theta'}|_{\theta} = (q_{1T}(\theta) \dots q_{mT}(\theta))$  at the pseudo-true value  $\theta^*$ :

$$V(\theta^*) = \lim_{T \rightarrow \infty} \text{var} \left( \sqrt{T} \begin{pmatrix} f_T(\theta^*) \\ \text{vec}(q_T(\theta^*)) \end{pmatrix} \right) = \begin{pmatrix} V_{ff}(\theta^*) & V_{fq}(\theta^*) \\ V_{qf}(\theta^*) & V_{qq}(\theta^*) \end{pmatrix},$$

with  $\hat{V}_{ff}(\theta)$  a consistent estimator of  $V_{ff}(\theta)$ ,  $\hat{V}_{ff}(\theta) \xrightarrow{p} V_{ff}(\theta)$ , and similarly,  $\hat{V}_{qq}(\theta) \xrightarrow{p} V_{qq}(\theta)$ ,  $\hat{V}_{qf}(\theta) \xrightarrow{p} V_{qf}(\theta)$ ,  $\hat{V}_{qf}(\theta) = (\hat{V}_{q1f}(\theta)' \dots \hat{V}_{qm f}(\theta)')'$ .

The double robust Lagrange multiplier (DRLM) statistic for testing  $H_0 : \theta^* = \theta_0^*$ , with  $\theta_0^*$  the hypothesized pseudo-true value, is:

$$DRLM(\theta_0^*) = T \times \hat{s}(\theta_0^*) \hat{W}(\theta_0^*)^{-1} \hat{s}(\theta_0^*)' \preceq \chi_m^2,$$

where  $\hat{s}(\theta_0^*)$  is the sample score:

$$\hat{s}(\theta_0^*) = f_T(\theta_0^*)' \hat{V}_{ff}(\theta_0^*)^{-1} \hat{D}(\theta_0^*),$$

$\hat{D}(\theta_0^*)$  is the recentered Jacobian:

$$\hat{D}(\theta_0^*) = q_T(\theta_0^*) - \left[ \hat{V}_{q1f}(\theta_0^*) \hat{V}_{ff}(\theta_0^*)^{-1} f_T(\theta_0^*) \dots \hat{V}_{qm f}(\theta_0^*) \hat{V}_{ff}(\theta_0^*)^{-1} f_T(\theta_0^*) \right],$$

and  $\hat{W}(\theta_0^*)$  is the covariance matrix of the sample score:

$$\begin{aligned} \hat{W}(\theta_0^*) = & \left( I_m \otimes \hat{V}_{ff}(\theta_0^*)^{-1} f_T(\theta_0^*) \right)' \hat{V}_{\theta\theta}(\theta_0^*) \left( I_m \otimes \hat{V}_{ff}(\theta_0^*)^{-1} f_T(\theta_0^*) \right) + \\ & \hat{D}(\theta_0^*)' \hat{V}_{ff}(\theta_0^*)^{-1} \hat{D}(\theta_0^*), \end{aligned}$$

with  $\hat{V}_{\theta\theta}(\theta) = \hat{V}_{qq}(\theta) - \hat{V}_{qf}(\theta) \hat{V}_{ff}(\theta)^{-1} \hat{V}_{qf}(\theta)'$ .

## A6: DRLM for beta pricing

The DRLM test proposed in the GMM setting is applicable to beta pricing, so it can be used for testing hypotheses on the pseudo-true value of the risk premia on the  $k \times 1$  vector of risk factors. Under the hypothesis of interest,  $H_0 : \lambda_f^* = \lambda_{f,0}^*$ , the limiting distribution of the DRLM statistic under  $T \gg N$  is bounded by a  $\chi_k^2$  distribution. The DRLM statistic then reads:

$$\begin{aligned} & DRLM(\lambda_{f,0}^*) \\ = & T \times (\bar{\mathbf{R}} - \hat{\mathbf{B}}\lambda_{f,0}^*)' \hat{V}_{(\bar{\mathbf{R}} - \hat{\mathbf{B}}\lambda_{f,0}^*)}(\lambda_{f,0}^*)^{-1} \hat{D}(\lambda_{f,0}^*) \left[ \hat{D}(\lambda_{f,0}^*)' \hat{V}_{(\bar{\mathbf{R}} - \hat{\mathbf{B}}\lambda_{f,0}^*)}(\lambda_{f,0}^*)^{-1} \hat{D}(\lambda_{f,0}^*) + \right. \\ & \left. \left( I_k \otimes \hat{V}_{(\bar{\mathbf{R}} - \hat{\mathbf{B}}\lambda_{f,0}^*)}(\lambda_{f,0}^*)^{-1} (\bar{\mathbf{R}} - \hat{\mathbf{B}}\lambda_{f,0}^*) \right)' \hat{V}_{\hat{D}(\lambda_{f,0}^*)}(\lambda_{f,0}^*) \left( I_k \otimes \hat{V}_{(\bar{\mathbf{R}} - \hat{\mathbf{B}}\lambda_{f,0}^*)}(\lambda_{f,0}^*)^{-1} (\bar{\mathbf{R}} - \hat{\mathbf{B}}\lambda_{f,0}^*) \right) \right]^{-1} \\ & \hat{D}(\lambda_{f,0}^*)' \hat{V}_{(\bar{\mathbf{R}} - \hat{\mathbf{B}}\lambda_{f,0}^*)}(\lambda_{f,0}^*)^{-1} (\bar{\mathbf{R}} - \hat{\mathbf{B}}\lambda_{f,0}^*), \end{aligned}$$

where  $\hat{V}_{(\bar{\mathbf{R}} - \hat{\mathbf{B}}\lambda_{f,0}^*)}(\lambda_{f,0}^*)$  is the covariance matrix estimator of the sample pricing error  $\bar{\mathbf{R}} - \hat{\mathbf{B}}\lambda_{f,0}^*$ ,  $\hat{V}_{\hat{D}(\lambda_{f,0}^*)}$  is the estimator of the covariance matrix of  $\text{vec}(\hat{D}(\lambda_{f,0}^*))$ , with:

$$\hat{D}(\lambda_{f,0}^*) = -\frac{1}{T} \sum_{t=1}^T \mathbf{R}_t(\bar{f}_t + \lambda_{f,0}^*)' \left[ \frac{1}{T} \sum_{t=1}^T (\bar{f}_t + \lambda_{f,0}^*)(\bar{f}_t + \lambda_{f,0}^*)' \right]^{-1}.$$

For the i.i.d. setting, the DRLM statistic above simplifies to:

$$DRLM(\lambda_{f,0}^*) = \hat{\mu}(\lambda_{f,0}^*)^*{}' \hat{D}(\lambda_{f,0}^*)^* \left[ \hat{\mu}(\lambda_{f,0}^*)^*{}' \hat{\mu}(\lambda_{f,0}^*)^* I_k + \hat{D}(\lambda_{f,0}^*)^*{}' \hat{D}(\lambda_{f,0}^*)^* \right]^{-1} \hat{D}(\lambda_{f,0}^*)^*{}' \hat{\mu}(\lambda_{f,0}^*)^*,$$

with  $\hat{\mu}(\lambda_{f,0}^*)^* = \sqrt{T} \hat{\Sigma}^{-\frac{1}{2}} (\bar{\mathbf{R}} - \hat{\mathbf{B}}\lambda_{f,0}^*) (1 + \lambda_{f,0}^{*'} \hat{Q}_{FF}^{-1} \lambda_{f,0}^*)^{-\frac{1}{2}}$ , and  $\hat{D}(\lambda_{f,0}^*)^* = \sqrt{T} \hat{\Sigma}^{-\frac{1}{2}} \hat{D}(\lambda_{f,0}^*) (\hat{Q}_{FF} + \lambda_{f,0}^* \lambda_{f,0}^{*'})^{\frac{1}{2}}$ .