Identification Robust Inference for the Risk Premium in Term Structure Models

Journal of Econometrics Invited lecture

on identification and asset pricing

Frank Kleibergen, Lingwei Kong and Zhaoguo Zhan

University of Amsterdam, University of Groningen

and Kennesaw State University

ASSA Philadelphia 2026

Motivation:

- Many many risk factors have been found to "explain" the cross-section of expected asset returns.
- It is hard to comprehend that so many factors can explain the cross section of asset returns which more than hints at problems with the employed two pass Fama-MacBeth (FM) 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 FM t-tests cannot be taken at face value.
- Awareness is currently emerging that this provides a major obstacle for empirical asset pricing.
- The identification issues similarly arise for bond pricing models.
- Identification strength is further reduced because asset pricing models are misspecified and identification conditions are more stringent in misspecified models compared to correctly specified models.
- Using a large cross-section of individuals asset returns alleviates identification issues compared to a small number of portfolio returns.

Review talk of papers on the topics of:

- Identification issues and robust inference for beta pricing with linear factor model: Kleibergen (JOE, 2009), Kleibergen and Zhan (JOE, 2015, JF 2020, JFEC 2018), Kleibergen, Kong and Zhan (JFEC, 2023)
- Risk premia from the cross-section of individual assets: Kleibergen and Zhan (JOE, 2025)
- Dynamic affine term structure model: Kleibergen and Kong (JOE, 2025)
- Misspecification and identification: Kleibergen and Zhan (QE, 2025 + working paper)

Outline:

- Beta asset pricing with linear factor model
- Fama-MacBeth two-pass procedure
- Cross-sectional R^2
- Risk premia from the cross-section of individual asset returns
- Different factors stand out when using assets vs portfolios
- Dynamic Affine Term Structure Model (DATSM)
- Misspecification and identification
- Conclusions

Beta asset pricing with linear factor model

Premia on risk factors result from linear factor model:

$$\mu_{R,i} = E(R_{i,t}) = \lambda_0 + \beta_i' \lambda_f, \ \ i = 1, ..., N,$$

where λ_0 is the zero-beta return, λ_f the k-dimensional risk premium vector on the k risk factors, β_i is a k-dimensional vector that results from a time-series regression of the i-th asset return on the k risk factors and $R_{i,t}$ is the excess return on the i-th asset at time t.

Linear asset pricing started with the Capital Asset Pricing Model (CAPM).

The one risk factor, market return, from the CAPM is too limited to explain the cross-section of asset returns.

Many many additional risk factors have therefore been discovered using the **Fama MacBeth (1973, FM) two pass procedure:**

Fama MacBeth (1973, FM) two pass procedure:

The FM two pass procedure estimates the risk premia on the betas associated with the $k \times 1$ dimensional risk factor f_t, $t = 1, \ldots, T$, in two steps (passes):

1. **Time-series regressions:** Estimate the betas, β_i, $i = 1, \ldots, N$, of the N factor models:

$$R_{i,t} = c_i + f_t' \beta_i + v_{i,t}, \qquad t = 1, \ldots, T,$$

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

2. **Cross-section regression:** For $\hat{\beta} = (\hat{\beta}_1 \ldots \hat{\beta}_N)'$, $\bar{R} = (\bar{r}_1 \ldots \bar{r}_N)'$, $\bar{r}_i = \frac{1}{T}\sum_{t=1}^{T} R_{i,t}$, estimate λ_0 and λ_F in:

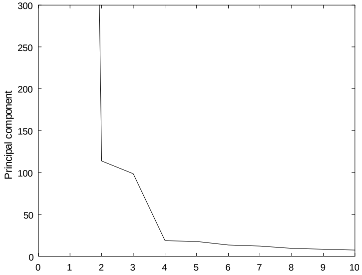
$$\bar{R} = \iota_N \lambda_0 + \hat{\beta}\lambda_f + \bar{v},$$

using linear regression, with ι_N the N-dimensional vector of ones.

For asset returns, it is common to use portfolio returns, like, for example, the 25 Fama-French size and book to market sorted portfolios and extensions thereof.

These portfolio returns exhibit a very strong factor structure: 95% of their variation is explained by the largest three principal components in data from Lettau and Ludvigson (2001):

Scree plot of the ten largest principal components of the twenty-five Fama-French size and book to market sorted portfolios (1-st PC$>$2K)



Factors thus have to correlate with the largest three principal components to have meaningful risk premia.

Risk premia from FM two-pass procedure are, however, gauged using:

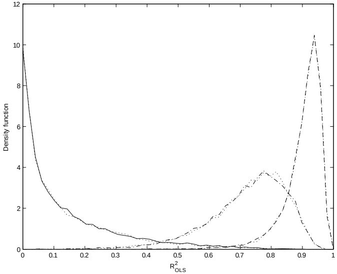
1. cross-sectional R^2 of the second-pass regression
2. t-statistics of the risk premia estimates

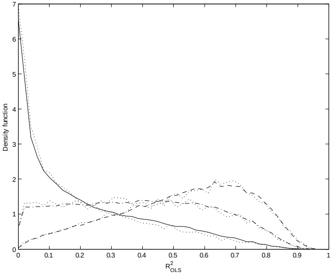
Large and significant values of these measures are considered to be indicative of relevant risk premia.

Cross-sectional R^2

The cross-sectional R^2 of the second pass regression is (very) sensitive to the factor structure:

Simulated density function of cross-sectional R^2 using data calibrated to 25 size and book-to-market sorted portfolio with three Fama-French factors





one, two, three authentic factors one, two, three irrelevant factors

The cross-sectional R^2 is large due to factor structure irrespective of whether the factors in the FM two-pass procedure capture it.

Lettau and Ludvigson (2001) find a significant risk premium for the interaction between the (lagged) consumption-wealth ratio and consumption growth:

Risk premium λ_f in multi-factor models for LL (2001)

	Risk factors			Risk factors		
	$\triangle c$	**cay**	$\triangle c \times$**cay**	R_m	**SMB**	**HML**
$\hat{\lambda}_f$	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-section R^2		0.698			0.803	
Pseudo-R^2		0.105			0.946	

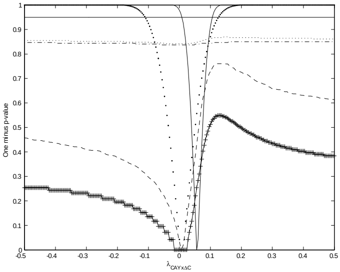
Pseudo-R^2 shows low percentage of the variation explained in the first pass regression when using $\triangle$**c**, cay, $\triangle$**c**$\times$cay, as risk factors.

Pseudo-R^2 is large when using the three Fama-French factors.

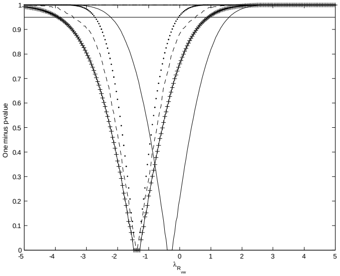
Low Pseudo-R^2 indicates that high cross-section R^2 and significant t-statistic on $\triangle c \times$**cay** are spurious despite their large values.

Large cross-section R^2 results from the strong factor structure shown by the scree plot.

one-minus p-value plots using Lettau-Ludvigson (2001) for:
Identification robust: FAR (dotted), FLR (dashed),
FLM (solid-plusses), FJLM (dash-dotted)
Non-robust: Shanken t (solid), MLE t (points)



risk premium on $\triangle\mathbf{c} \times \mathbf{cay}$ using ($\triangle\mathbf{c}$, cay, $\triangle\mathbf{c}\times$cay) as factors



risk premium of r_{vw} using (r_{vw}, SMB, HML) as factors

r_{vw} : similar p-values for identification robust and non-robust tests.

$\triangle\mathbf{c} \times \mathbf{cay}$: Shanken t-test rejects zero with 5% significance, identification robust tests do not reject any risk premium at 5%.

The cross-sectional R^2 and t-tests for Lettau-Ludvigson (2001) illustrate that they set a too low bar for gauging risk factors.

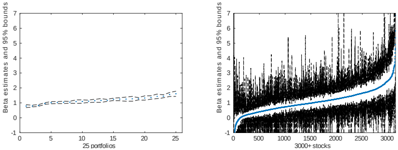
These results similarly apply to many other papers in asset pricing and partly explain the zoo of risk factors that have been proposed.

The identification robust tests show that just a few risk factors matter.

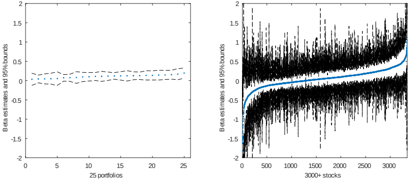
One of the reasons for the poor identification of risk premia is the usage of portfolios which average over the betas of the individual assets because of which there is often too little variation in the betas.

This is alleviated when using individuals assets instead of portfolios.

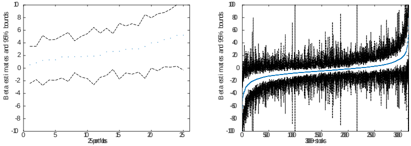
market betas estimated using portfolios vs. stocks



leverage betas estimated using portfolios vs. stocks



consumption betas estimated using portfolios vs. stocks



Risk premia from the cross-section of individual asset returns

For individual assets, it is convenient to use a short panel set up 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:

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

with $\bar{f}$ the average of the factors and $\lambda_{pop,f}$ the population risk premia.

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

Shanken (1992) bias-adjusted estimator:

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

with $\hat{\omega}^2$ estimator of the approximate common variance across assets, Q_{FF} factor covariance (matrix).

Raponi, Robotti, and Zaffaroni (2020) construct the covariance matrix of $\hat{\lambda}_{f,Shanken}$.

Continuous Updating Estimator

The continuous updating estimator (CUE) of λ_f can be written as (=so-called k-class notation of limited information maximum likelihood estimator):

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

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

$$\left| \mu \begin{pmatrix} 1 & 0 \\ 0 & Q_{FF}^{-1} \end{pmatrix} - \begin{pmatrix} \bar{R} & \hat{\beta} \end{pmatrix}' M_{\iota_N} \begin{pmatrix} \bar{R} & \hat{\beta} \end{pmatrix} \right| = 0.$$

Recall that:

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

Since we have many assets, the beta-pricing moment condition will only hold approximately and there will be misspecification because of unobserved priced risk factors.

The object of interest is then no longer the ex-post risk premia but the pseudo-true value which is the minimizer of the (conditional) population objective function.

For the CUE:

$$\hat{\beta}' M_{\iota_N} \hat{\beta} - \mu_{\min} Q_{FF}^{-1}$$

is guaranteed to be positive definite because $\mu_{\min}$ is always exceeded by the smallest eigenvalue of $Q_{FF} \hat{\beta}' M_{\iota_N} \hat{\beta}$.

For the Shanken estimator, it is, however, not ensured that

$$\hat{\beta}' M_{\iota_N} \hat{\beta} - \frac{N}{T} \hat{\omega}^2 Q_{FF}^{-1},$$

is positive definite so Raponi, Robotti, and Zaffaroni (2020) apply shrinkage when invertibility is at stake.

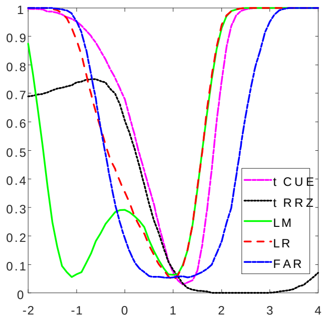
Five tests on ex post risk premium/pseudo true value and when they are valid:

- t test by Shanken (1992) and Raponi, Robotti, and Zaffaroni (2020):
  - Strong identification, no unobserved priced factors
- t test based on CUE:
  - Strong identification, unobserved priced factors
- FAR test based on the CUE objective function:
  - Weak identification, no unobserved priced factors
- LM test based on the score/derivative of the CUE objective function:
  - Weak identification, unobserved priced factors (double robust)
- LR test based on combining FAR and LM:
  - Weak identification, no unobserved priced factors.

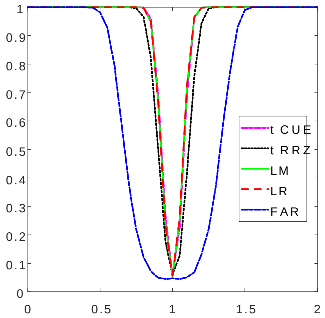
Simulation experiment illustrating size and power of these five tests on the pseudo-true value:

Rejection frequencies of tests of

$H_0 : \lambda_f = 1$ for correct beta pricing

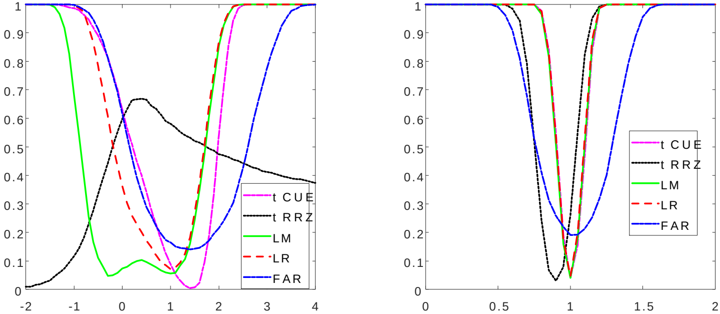


Weak identification of λ_f



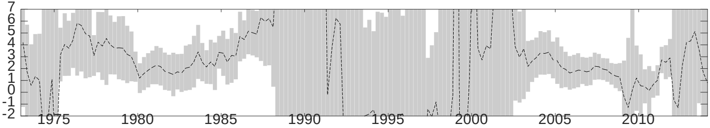
Strong identification of λ_f

Rejection frequencies of tests of $H_0 : \lambda_f^* = 1$ for misspecified beta pricing with priced unobserved risk factors

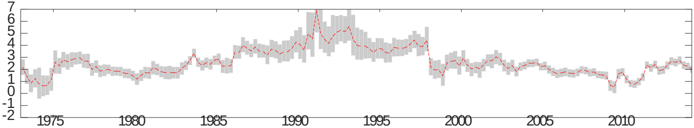


Weak identification of λ_f^* | Strong identification of λ_f^*

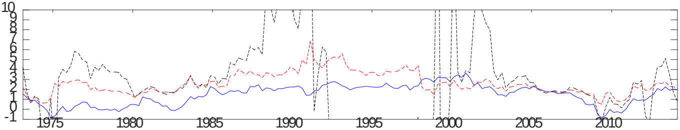
Market risk premium in Fama-French three factor model (FF3)



Bias-adjusted estimates and 95% confidence sets from the t-test



95% confidence sets from the LM test

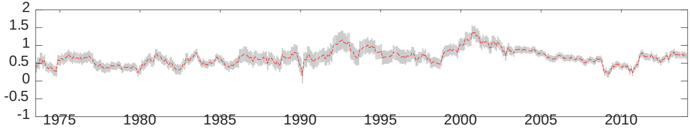


Point estimates vs. average market return

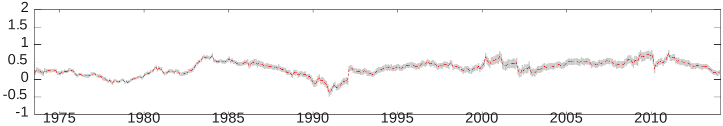
Note periods of indefiniteness of $\hat{\beta}' M_{\iota_N} \hat{\beta} - \frac{N}{T} \hat{\omega}^2 Q_{FF}^{-1}$ because of which signs of CUE and bias-adjusted estimator differ

Different risk factors stand out when using assets vs portfolios

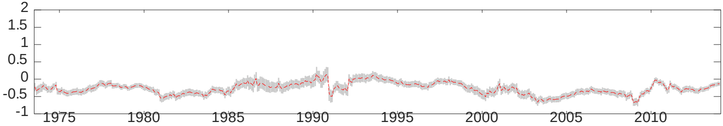
Risk premia from the CUE and the LM test in Fama-French (2018) 6 factor model, monthly data



Market

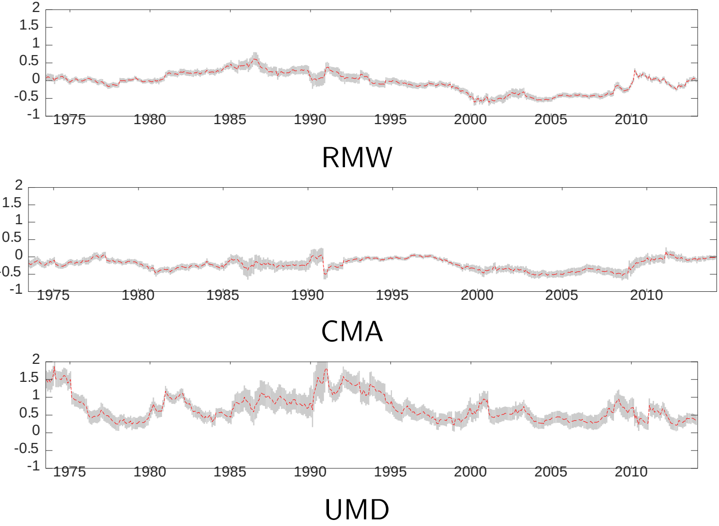


SMB



HML

Risk premia from the CUE and the LM test in FF6, monthly data



Positive risk premia mostly on market, SMB and UMD and much less so on HML, RMW and CMA

- Identification of (pseudo-true values of) risk premia markedly improves when using cross-section of assets compared to portfolios
- More research needed to determine which risk factors stand out in the cross-section of individual asset returns

Dynamic Affine Term Structure Model (DATSM)

Identification issues of risk premia also appear in DATSMs where, for r_t the one period short rate and λ_t the market price of risk, the pricing kernel is assumed exponential affine in innovation factors $v_t \sim_{i.i.d} N(0, \Sigma_v)$:

$$M_{t+1} = \exp\left(-r_t - \frac{1}{2}\lambda_t' \lambda_t - \lambda_t' \Sigma_v^{-\frac{1}{2}} v_{t+1}\right),$$

with the market price of risk λ_t an affine function of X_t :

$$\lambda_t = \Sigma_v^{-\frac{1}{2}} \left(\lambda_0 + \Lambda_1 v_t\right),$$

where λ_0 and Λ_1 are a k-dimensional vector and $k \times k$ dimensional matrix resp., and the k-dimensional vector of state variables X_t results from a VAR(1):

$$X_{t+1} = \mu + \Phi X_t + v_{t+1}.$$

For $r_{t+1,n}$ the one-period excess holding return of a n-period bond at $t+1$, $(r_{t+1,n},\ v_{t+1})$ jointly normal, DATSM implies, see Adrian et al (2013):

$$E_t\left(r_{t+1,n}\right) = \beta^{(n)\prime}\left(\lambda_0 + \Lambda_1 X_t\right) - \frac{1}{2} var(r_{t+1,n}),$$

with $\beta^{(n)} = \Sigma_v^{-1} cov(v_{t+1}, r_{t+1,n})$.

Decomposing, $r_{t+1,n} - E_t\left(r_{t+1,n}\right)$ into a component correlated with v_{t+1} and an uncorrelated component $e_{t+1,n}$, and subtracting the time-series average:

$$\bar{r}_{t+1,n} = \beta^{(n)\prime}\left(\Lambda_1 \bar{X}_t\right) + \beta^{(n)\prime}\bar{v}_{t+1} + \bar{e}_{t+1,n}.$$

Stacking the equations for N different maturities:

$$R_{t+1} = \beta\left(\Lambda_1 \bar{X}_t\right) + \beta \bar{v}_{t+1} + e_{t+1},$$

with $R_{t+1} = (\bar{r}_{t+1,1} \ldots \bar{r}_{t+1,N})'$, $\beta = (\beta^{(1)} \ldots \beta^{(n)})'$, $e_t = (\bar{e}_{t,1} \ldots \bar{e}_{t,N})'$, brings out the close resemblance with the beta-pricing model for the return on assets:

$$\mathrm{r}_{t+1} = \beta\lambda_f + \beta f_{t+1} + u_{t+1},$$

where r_t is an N-dimensional vector with asset returns, β the $N \times k$ dimensional beta matrix and F_t a k-dimensional vector of risk factors.

The similarity extends to the reduced rank structure that these models impose:

$$\text{DATSM:} \qquad R_{t+1} = \beta \left(\Lambda_1 \vdots I_k \right) \begin{pmatrix} \bar{X}_t \\ \bar{v}_{t+1} \end{pmatrix} + e_t$$

$$\text{beta-pricing:} \quad \mathrm{r}_{t+1} = \beta \left(\lambda_f \vdots I_k \right) \begin{pmatrix} 1 \\ f_{t+1} \end{pmatrix} + u_t,$$

where the $N \times 2k$ and $N \times (k+1)$ dimensional matrices $\beta \left(\Lambda_1 \vdots I_k \right)$ and $\beta \left(\lambda \vdots I_k \right)$ are each at most of rank k.

The reduced rank structures imply the probability limits:

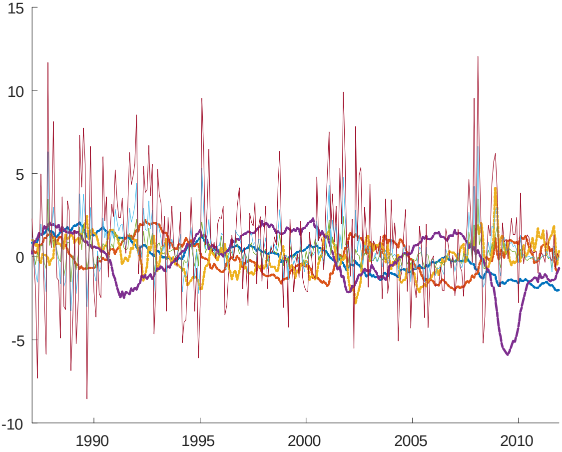
$$\text{DATSM:} \quad \frac{1}{T}\sum_{t=1}^{T} R_t \begin{pmatrix} \bar{X}_t \\ \bar{v}_{t+1} \end{pmatrix}' \left[\frac{1}{T}\sum_{t=1}^{T} \begin{pmatrix} \bar{X}_t \\ \bar{v}_{t+1} \end{pmatrix} \begin{pmatrix} \bar{X}_t \\ \bar{v}_{t+1} \end{pmatrix}' \right] \underset{p}{\rightarrow} \beta \left(\Lambda_1 \vdots I_K \right)$$

$$\text{beta-pricing:} \quad \frac{1}{T}\sum_{t=1}^{T} \mathrm{r}_{t+1} \begin{pmatrix} 1 \\ F_{t+1} \end{pmatrix}' \left[\frac{1}{T}\sum_{t=1}^{T} \begin{pmatrix} 1 \\ F_{t+1} \end{pmatrix} \begin{pmatrix} 1 \\ F_{t+1} \end{pmatrix}' \right] \underset{p}{\rightarrow} \beta \left(\lambda \vdots I_K \right)$$

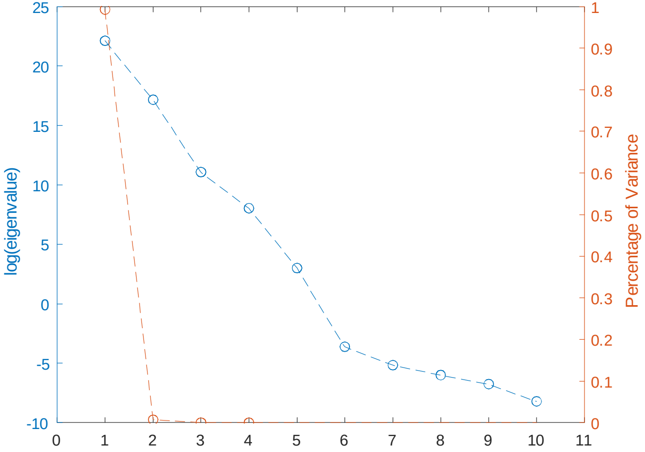
so resp. the k smallest and the smallest singular values of these matrices are zero.

Plots of

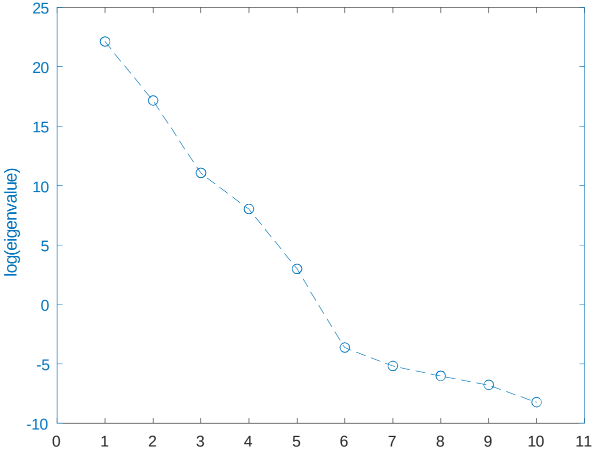
selected excess returns Adrian et al (2013) (multiplied by 1200); level, slope and curvature factors; and one macro factor (real economy, purple dashed curve)



The singular values of $\frac{1}{T}\sum_{t=1}^{T} R_t \begin{pmatrix} \bar{X}_t \\ \hat{v}_{t+1} \end{pmatrix}' \left[\frac{1}{T}\sum_{t=1}^{T} \begin{pmatrix} \bar{X}_t \\ \hat{v}_{t+1} \end{pmatrix} \begin{pmatrix} \bar{X}_t \\ \hat{v}_{t+1} \end{pmatrix}' \right]$ show that the first factor explains most of the variation.



The (log) singular values of $\frac{1}{T}\sum_{t=1}^{T} R_t \begin{pmatrix} \bar{X}_t \\ \hat{v}_{t+1} \end{pmatrix}' \left[\frac{1}{T}\sum_{t=1}^{T} \begin{pmatrix} \bar{X}_t \\ \hat{v}_{t+1} \end{pmatrix} \begin{pmatrix} \bar{X}_t \\ \hat{v}_{t+1} \end{pmatrix}'\right]$ show that from the fourth onwards they become rather small which indicates weak identification so the test for rank 4 does not reject.



$\hat{\beta}$ with t-statistics: five factors (Adrian et al (2013))

Kleibergen-Paap (2006) rank statistic testing H_0 : rank(β)=4: 1.6561 [0.9764].

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$
(1)	-0.0094	0.0031	-0.0008	0.0002	0.0000
	(-2.5289)	(1.7589)	(-1.0449)	(0.3049)	(0.0244)
(2)	-0.0213	0.0057	-0.0007	-0.0003	0.0002
	(-2.7029)	(1.4951)	(-0.4025)	(-0.2417)	(0.1914)
(3)	-0.0446	0.0070	0.0010	-0.0005	-0.0001
	(-2.5995)	(0.8482)	(0.2859)	(-0.2049)	(-0.0689)
(4)	-0.0656	0.0048	0.0024	0.0000	-0.0003
	(-2.4084)	(0.3670)	(0.4316)	(0.0037)	(-0.0919)
(5)	-0.0843	0.0003	0.0028	0.0007	-0.0001
	(-2.2751)	(0.0142)	(0.3704)	(0.1205)	(-0.0121)
(6)	-0.1011	-0.0059	0.0022	0.0010	0.0003
	(-2.2199)	(-0.2674)	(0.2370)	(0.1529)	(0.0611)
(7)	-0.1164	-0.0130	0.0008	0.0010	0.0006
	(-2.2279)	(-0.5139)	(0.0775)	(0.1265)	(0.0949)
(8)	-0.1305	-0.0206	-0.0011	0.0005	0.0005
	(-2.2842)	(-0.7455)	(-0.0913)	(0.0589)	(0.0831)
(9)	-0.1435	-0.0284	-0.0033	-0.0003	0.0002
	(-2.3772)	(-0.9723)	(-0.2623)	(-0.0386)	(0.0273)
(10)	-0.1556	-0.0361	-0.0056	-0.0015	-0.0005
	(-2.4936)	(-1.1969)	(-0.4306)	(-0.1578)	(-0.0685)
(11)	-0.1669	-0.0436	-0.0078	-0.0028	-0.0014
	(-2.6142)	(-1.4126)	(-0.5894)	(-0.2901)	(-0.1975)
$H_0 : \hat{\beta}_j = 0.$	33.8170	84.6685	8.7455	65.9716	23.0805
p-value	[0.0004]	[0.0000]	[0.6454]	[0.0000]	[0.0172]

Three step estimation procedure

Akin to the Fama-MacBeth two pass procedure, Adrian et al (2013) propose a three step estimation procedure:

1. Estimate:
$$X_{t+1} = \mu + \Phi X_t + v_{t+1},$$
by least squares to obtain $\hat{\mu}$, $\hat{\Phi}$, $\hat{v}_t$, $t = 1, \ldots, T$ and $\hat{\Sigma}_v$.
2. Use that $\hat{v}_t = X_t - \hat{\mu} - \hat{\Phi} X_{t-1}$, to estimate:
$$r_{t+1,n} = a^{(n)} + d^{(n)\prime} X_t + \beta^{(n)\prime} \hat{v}_{t+1} + e_{t+1,n},$$
by least squares to obtain $\hat{a}^{(n)}$, $\hat{d}^{(n)}$ and $\hat{\beta}^{(n)}$, $n = 1, \ldots, N$.
3. Construct $\hat{a} = (\hat{a}^{(1)} \ldots \hat{a}^{(N)})'$, $\hat{\beta} = (\hat{\beta}^{(1)} \ldots \hat{\beta}^{(N)})'$, $\hat{d} = (\hat{d}^{(1)} \ldots d^{(n)})'$, $\hat{g} = (\hat{g}^{(1)} \ldots \hat{g}^{(n)})$ for $\hat{g}^{(n)} = g^{(n)}(\hat{\beta}, \hat{\Sigma}_v, \hat{\Sigma}_e)$, $n = 1, \ldots, N$, and estimate λ_0 and Λ_1 using:
$$\begin{aligned} \hat{\lambda}_0 &= \left(\hat{\beta}'\hat{\beta}\right)^{\prime -1} \hat{\beta}'(\hat{a} + \hat{g} + \hat{\beta}\hat{\mu}) \\ \hat{\Lambda}_1 &= \left(\hat{\beta}'\hat{\beta}\right)^{\prime -1} \hat{\beta}'(\hat{d} + \hat{\beta}\hat{\Phi}). \end{aligned}$$

Identification robust test procedures

Because we can't reject a reduced rank value of the β matrix, t-tests resulting from the three step estimation procedure are unreliable.

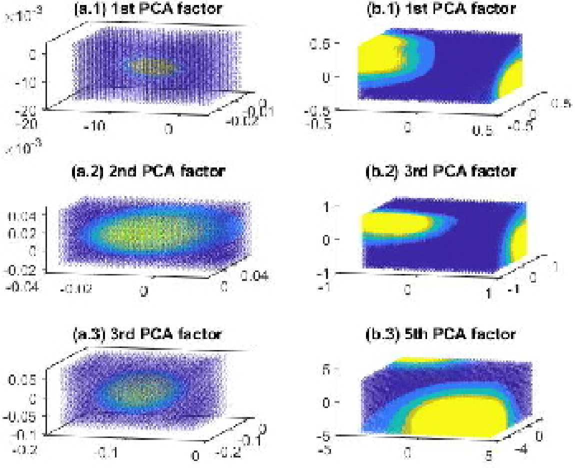
We therefore need identification robust test procedures for testing hypotheses on the risk premia in Λ_1.

Because Λ_1 is a $k \times k$ matrix of risk premia, the focus is on identification robust procedures for specific elements of them.

When the hypothesis of interest is phrased on all elements of a row or column of Λ_1, the distribution of the subset factor Anderson-Rubin (AR) statistic is bounded by a $\chi^2(K(N-(K-1))$ distributed random variable under a Kronecker product covariance structure.

For data from Adrian et al (2013), pre-tests for a Kronecker product structure covariance matrix do not reject at 5% significance.

Adrian et al (2013) use risk factors extracted using PCA. When using too few, the subset AR statistic always rejects at 5% significance so there is misspecification.



Joint confidence sets from the sFAR test for the (three) risk premia on one factor in three factor model. Left: 1-3 PCA risk factors, right: 1, 3 and 5 PCA risk factors.

We have bounded 95% confidence sets for the risk premia on the first three PCA factors.

Because the 5-th PCA factor is weak, we have unbounded 95% confidence sets when using the 1-st, 3-rd and 5-th PCAs as risk factors.

When we use four PCA factors, we need projection to show confidence sets in a visual manner.

For two PCAs as risk factors, we have mostly empty 95% confidence sets from subset FAR statistic since the model is then misspecified.

Misspecification and identification

Because risk factors do not explain the full cross-section of asset returns, the beta-pricing model is misspecified.

Identification condition turns out to be more stringent in misspecified model.

For

$$\text{MISS} = \min_{\theta \in \mathbb{R}^m} Q_{CUE}(\theta),$$

with $Q_{CUE}(\theta)$ continuous updating objective function using beta pricing moment conditons, and IS, the Cragg-Donald rank statistic which tests the rank of the β matrix:

$$\text{IS} > \text{MISS} > 0$$

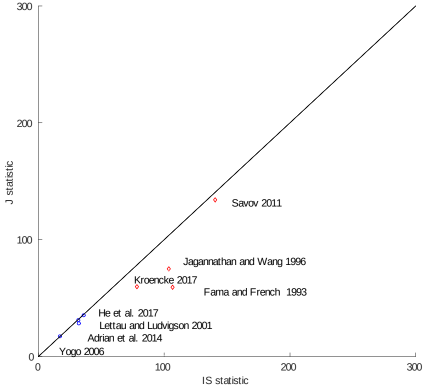
for pseudo-true value to be identified.

More stringent condition than in correctly specified models:

$$\text{IS} > 0.$$

Empirical relevance: linear asset pricing

We compute IS and MISS (=J) statistic for eight well-cited empirical studies in asset pricing



It shows the importance of considering IS−MISS instead of just IS.

Table: **MISS-statistic and IS-statistic** (*: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***

The MISS-statistics indicate that many models are misspecified while 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 is to be tested using IS$-$MISS which provides a quasi-likelihood ratio statistic to test no-identification.

The scatter plot of IS and MISS shows that the distribution of IS$-$MISS can be degenerate.

Kleibergen and Zhan (2025) "Testing for identifiπcation in potentially misspecified linear GMM" shows how to construct critical value functions.

Conclusions

- Many/most risk premia found in empirical studies are probably not 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 large number of individual asset returns compared to the default of a small number of portfolios
- Betas do not span the full cross-section of asset returns so procedures robust to both misspecification and weak identification should be employed
- Identification condition in misspecified models is more stringent than in correctly specified models