Internet Appendix for
“Robust Inference for Consumption-Based Asset Pricing”
FRANK KLEIBERGEN and ZHAOGUO ZHAN

In this appendix for the paper entitled “Robust Inference for Consumption-Based Asset Pricing”, we provide (I) technical details, (II) some additional numerical as well as (III) empirical results for linear models, and (IV) supplementary material for nonlinear models.

Contents

I Technical Details 2
A Notation 2
B Rank Test 2
C Closed-Form Expressions of GRS-FAR Confidence Sets 3
D Subset GRS-FAR Test 5
E Invariance to the Choice of the \(N\)-th Asset 6
F No Zero-Beta Return 8

II Additional Numerical Results for Linear Models 9
A Reliability of the FM \(t\)-test 9
B Sizes of the Rank Test when \(k = 2\) and \(k = 3\) 11

III Additional Empirical Results for Linear Models 11
A Robustness: Industry Portfolios as Test Assets 13
B Robustness: Including the Risk-Free Asset as a Test Asset 13
C Robustness: Seven Portfolio Returns 14
D \(p\)-value Plots for the Multifactor Models in Tables IV and V 14

IV Nonlinear Models 17
A Stochastic Discount Factors 17
B Tests on Relative Risk Aversion using GMM-AR and GMM \(t\) 18
C Additional Empirical Results for GMM-AR: \(N = 1\) 21
D DGPs that Satisfy both Linear and CRRA SDFs 24

* Citation format: Kleibergen, Frank and Zhaoguo Zhan, Internet Appendix for “Robust Inference for Consumption-Based Asset Pricing”, Journal of Finance, DOI:10.1111/jofi.12855. Please note: Wiley is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the authors of the article.
I. Technical Details

A. Notation

Consistent with the paper, we use the following notation: $R_{i,t}$ is the return of the $i$-th asset at time $t$, $i = 1, ..., N$, $t = 1, ..., T$; $f_t$ is the $k \times 1$ vector of risk factors, whose sample covariance is denoted by $\hat{Q}_{FF}$; $\beta_i = \text{var}(f_t)^{-1} \text{cov}(f_t, R_{i,t})$ is a $k$-dimensional vector, and $\beta = (\beta_1 \beta_2 \ldots \beta_N)'$ is an $N \times k$-dimensional matrix.

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

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

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

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

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

B. Rank Test

Consider $H_0 : \text{rank}(\mathbf{B}) = k - 1$. Let $rk$ be the smallest root of

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

which is identical to the smallest eigenvalue of the matrix $\hat{Q}_{FF} \hat{\mathbf{B}}' \hat{\Sigma}^{-1} \hat{\mathbf{B}}$. In the single factor case with $k = 1$, $rk = \hat{Q}_{FF} \hat{\mathbf{B}}' \hat{\Sigma}^{-1} \hat{\mathbf{B}}$.

The rank test $F$-statistic is

$$F(\mathbf{B} = 0) = \frac{T - N}{N - k} \times rk,$$
and under $H_0: \text{rank}(\mathbf{B}) = k - 1$, i.i.d. normal errors $\mathbf{u}_t$, and fixed factors $f_t$, it is bounded by an $F$-distributed random variable (see Kleibergen, Kong, and Zhan (2019)):

$$F(\text{rank}(\mathbf{B}) = k - 1) = \frac{T - N}{N - k} \times \text{rk} \leq F(N - k, T - N).$$

Alternatively, if $T$ is large, then the Wald rank statistic reads

$$W(\mathbf{B} = 0) = T \times \text{rk} \leq \chi^2_{N-k}.$$

C. Closed-Form Expressions of GRS-FAR Confidence Sets

Consider the second specification of the GRS-FAR statistic in the paper,

$$\text{GRS-FAR}(\lambda_{f,0}) = \frac{T}{1 + N_{f,0}Q_{FF}^{-1}f,0}(\mathbf{R} - \hat{\mathbf{B}}\lambda_{f,0})'\hat{\Sigma}^{-1}(\mathbf{R} - \hat{\mathbf{B}}\lambda_{f,0}),$$

which is related to the characteristic polynomial

$$\left| \begin{array}{cc}
1 & 0 \\
0 & \hat{Q}_{FF}^{-1}
\end{array} \right| - \left( \begin{array}{c}
\mathbf{R} \\
\hat{\mathbf{B}}
\end{array} \right)'\hat{\Sigma}^{-1}\left( \begin{array}{c}
\mathbf{R} \\
\hat{\mathbf{B}}
\end{array} \right) = 0.$$

The minimal value of the GRS-FAR statistic corresponds to $T$ times the smallest root of the characteristic polynomial above. The smallest root of this polynomial is equal to the smallest eigenvalue of the matrix

$$\left( \begin{array}{cc}
1 & 0 \\
0 & \hat{Q}_{FF}^{-1}
\end{array} \right) \left( \begin{array}{c}
\mathbf{R} \\
\hat{\mathbf{B}}
\end{array} \right)'\hat{\Sigma}^{-1}\left( \begin{array}{c}
\mathbf{R} \\
\hat{\mathbf{B}}
\end{array} \right).$$

The $100 \times (1 - \alpha)\%$ confidence set on $\lambda_f$ that results from the GRS-FAR test is

$$\text{CS}_{\lambda_f}(\alpha) = \left\{ \lambda_{f,0} : \frac{T - k - N + 1}{T(N - 1)} \times \text{GRS-FAR}(\lambda_{f,0}) \leq F_\alpha(N - 1, T - k - N + 1) \right\},$$

where $F_\alpha(N_1, N_2)$ is the upper $\alpha$-th quantile of the $F(N_1, N_2)$ distribution.
Using the expression of the GRS-FAR statistic above, the $100 \times (1 - \alpha) \%$ confidence set then results from the inequality
\[
\frac{T-k-N+1}{T(N-1)} \times \frac{T}{1+\lambda_f^T Q^{-1}_{FF} \lambda_f} (\hat{R} - \hat{B} \lambda_f) \hat{\Sigma}^{-1} (\hat{R} - \hat{B} \lambda_f) \leq F_\alpha(N-1, T-k-N+1).
\]

Multiplying by $1 + \lambda_f^T \hat{Q}_{FF}^{-1} \lambda_f$ on both sides and re-arranging, we obtain
\[
(\hat{R} - \hat{B} \lambda_f) \hat{\Sigma}^{-1} (\hat{R} - \hat{B} \lambda_f) \leq \frac{N-1}{T-k-N+1} F_\alpha(N-1, T-k-N+1)(1 + \lambda_f^T \hat{Q}_{FF}^{-1} \lambda_f),
\]
which can be further rewritten as $\lambda_f^T A \lambda_f + b' \lambda_f + c \leq 0$ with the following expressions for $A$, $b$, and $c$:
\[
A = \hat{B}' \hat{\Sigma}^{-1} \hat{B} - \frac{N-1}{T-k-N+1} F_\alpha(N-1, T-k-N+1) \hat{Q}_{FF}^{-1},
\]
\[
b = -2\hat{B}' \hat{\Sigma}^{-1} \hat{R},
\]
\[
c = \hat{R}' \hat{\Sigma}^{-1} \hat{R} - \frac{N-1}{T-k-N+1} F_\alpha(N-1, T-k-N+1).
\]

Depending on the values of $A$, $b$, and $c$, there exist an unbounded range, finite range, or no values of $\lambda_f$ such that $\lambda_f^T A \lambda_f + b' \lambda_f + c \leq 0$ is satisfied.

In the case of a single risk factor, $k = 1$, we have that $A$, $b$, $c$, $\lambda_f$ are all scalars, so $\lambda_f^T A \lambda_f + b' \lambda_f + c \leq 0$ is just an inequality based on a second-order polynomial. We then obtain the following expressions for the $100 \times (1 - \alpha) \%$ confidence set of $\lambda_f$.

1. If $A > 0$ and $\frac{1}{4} b' A^{-1} b - c \geq 0$ : \((\frac{-b - \sqrt{b^2 - 4Ac}}{2A}, \frac{-b + \sqrt{b^2 - 4Ac}}{2A})\).
2. If $A > 0$ and $\frac{1}{4} b' A^{-1} b - c < 0$ : it is empty, $\emptyset$.
3. If $A < 0$ and $\frac{1}{4} b' A^{-1} b - c < 0$ : \((\infty, \frac{-b + \sqrt{b^2 - 4Ac}}{2A}) \cup \left(\frac{-b - \sqrt{b^2 - 4Ac}}{2A}, \infty\right)\).
4. If $A < 0$ and $\frac{1}{4} b' A^{-1} b - c \geq 0$ : \((\infty, \infty)\).

Note that it is unlikely that the data-dependent $A$ is exactly equal to zero, but if $A = 0$, it is straightforward to derive $\lambda_f$ satisfying $b' \lambda_f + c \leq 0$. When $k = 1$,

5. If $A = 0$, $b = 0$, and $c \leq 0$ : \((\infty, \infty)\).
6. If $A = 0$, $b = 0$, and $c > 0$: it is empty, $\emptyset$.

7. If $A = 0$ and $b > 0$: $(-\infty, -\frac{c}{b})$.

8. If $A = 0$ and $b < 0$: $(-\frac{c}{b}, +\infty)$.

For the multifactor case with $k > 1$, see Dufour and Taamouti (2005) for analytical expressions of confidence sets similar to those provided above for the $k = 1$ case.

D. Subset GRS-FAR Test

When we have more than one risk premium (i.e., when we consider multifactor models), the GRS-FAR test can be used to test a joint hypothesis specified on all risk premia. When using $k$ risk premia, this GRS-FAR test leads to $k$-dimensional confidence sets. One-dimensional confidence sets for each risk premium can then be obtained by projecting out the joint confidence set.

Alternatively, one-dimensional confidence sets can be obtained by using the subset GRS-FAR test. The subset GRS-FAR test replaces every non-hypothesized risk premium by its maximum likelihood estimator under the null. When $H_0: \lambda_{f,1} = \lambda_{f,1}^0$ and $\lambda_2 = (\lambda_{f,2} \ldots \lambda_{f,k})'$ are left unspecified, the subset GRS-FAR (sGRS-FAR) statistic for testing $H_0$ reads

$$s\text{GRS-FAR}(\lambda_{f,1}) = \text{GRS-FAR}(\lambda_{f,1}^0, \hat{\lambda}_2^{(\lambda_{f,1}^0)}),$$

where $\hat{\lambda}_2^{(\lambda_{f,1}^0)}$ is the maximum likelihood estimator of $\lambda_2$ given $\lambda_{f,1} = \lambda_{f,1}^0$ (which equals the minimizer of GRS-FAR over $\lambda_2$ given $\lambda_{f,1} = \lambda_{f,1}^0$).

The $s\text{GRS-FAR}$ statistic can easily be computed using an eigenvalue problem. Specifically, the subset GRS-FAR statistic for testing the first element of $\lambda_f$ results from the characteristic polynomial

$$\left| \mu \left( 1 + \lambda_{f,1} \hat{Q}_{FF,11}^{-1} \lambda_{f,1} - \lambda_{f,1} \hat{Q}_{FF,12}^{-1} \right) - \left( \hat{B}_1 \lambda_{f,1} \hat{B}_2 \right) \right| \hat{\Sigma}^{-1} \left( \hat{B}_1 \lambda_{f,1} \hat{B}_2 \right) = 0,$$

where $\hat{Q}_{FF}^{-1} = \begin{pmatrix} \hat{Q}_{FF,11}^{-1} & \hat{Q}_{FF,12}^{-1} \\ \hat{Q}_{FF,21}^{-1} & \hat{Q}_{FF,22}^{-1} \end{pmatrix}$, $\hat{B} = (\hat{B}_1, \hat{B}_2)$, and the subset GRS-FAR statistic equals $T$ times the smallest root.
Under the usual i.i.d. (over time) settings, the distribution of the sGRS-FAR statistic is bounded in the limit by the $\chi_{N-k}^2$ distribution (see Guggenberger, et al. (2012) and Gospodinov, Kan, and Robotti (2017)):

$$s\text{GRS-FAR}(\lambda_{f,1}) \xrightarrow{d} \psi \preceq \chi_{N-k}^2.$$  

Under normal errors and fixed factors, the distribution of the sGRS-FAR statistic is bounded in finite sample according to (see Kleibergen, Kong, and Zhan (2019)):

$$\frac{T-N}{T(N-k)} \times s\text{GRS-FAR}(\lambda_{f,1}) \preceq F(N-k, T-N).$$

E. Invariance to the Choice of the N-th Asset

Our proposed rank and GRS-FAR test statistics are invariant with respect to the asset return that is subtracted, that is, the choice of the “N”-th asset. To prove this invariance, we use matrix notation. Let $R$ be the $T \times N$-dimensional matrix of asset returns,

$$R = \begin{pmatrix} R'_1 \\ \vdots \\ R'_T \end{pmatrix},$$

where $R_t$ is the $N \times 1$ vector of returns at time $t$, $t = 1, ..., T$. To switch the $j$-th asset and the $N$-th asset for $j = 1, 2, ..., N - 1$, we define the $N \times N$-dimensional matrix $M_1$,

$$M_1 = (v_1, ..., v_{j-1}, v_N, v_{j+1}, ..., v_{N-1}, v_j) = \begin{pmatrix} I_{j-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & I_{N-j-1} & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

where $v_i$ is the $N \times 1$ vector with one in its $i$-th entry, while the other entries are all zeros, $i = 1, \ldots, N$. Put differently, $M_1$ results from an identity matrix whose $j$-th and
$N$-th columns are switched. Therefore,

$$R \cdot M_1$$

switches the $j$-th asset and the (to be subtracted) last asset.

We further define $M_2$ as the $N \times (N - 1)$-dimensional subtraction matrix,

$$M_2 = \begin{pmatrix} I_{N-1} \\ -i'_{N-1} \end{pmatrix}.$$  

After switching and subtraction, we get the $T \times (N - 1)$-dimensional matrix $R$, where

$$R = R \cdot M_1 \cdot M_2 = R \cdot M_2 \cdot M_3$$

and $M_3$ is $(N - 1) \times (N - 1)$-dimensional matrix whose $j$-th row is made of negative ones and the remainder is the same as the identity matrix,

$$M_3 = \begin{pmatrix} I_{j-1} & 0 & 0 \\ -i'_{j-1} & -1 & -i'_{N-j-1} \\ 0 & 0 & I_{N-j-1} \end{pmatrix}.$$  

In effect, $M_1$ performs a row operation on $M_2$. This row operation is equivalent to the column operation performed by $M_3$. The invertible square matrix $M_3$ cancels out when we construct our test statistics. This implies that our test statistics are invariant with respect to which asset is subtracted.

For instance, the rank statistic is proportional to $\hat{B}'\hat{\Sigma}^{-1}\hat{B}$, where

$$\hat{B} = (RM_2M_3)'M_{\ell\tau}F(F'M_{\ell\tau}F)^{-1}$$

and

$$T\hat{\Sigma} = (RM_2M_3)'M_{F,\ell\tau}(RM_2M_3).$$
with $F = (f_1, \ldots, f_T)'$, the $T \times k$-dimensional matrix of factors, $M_{\iota T} = I_T - \iota_T (\iota_T' \iota_T)^{-1} \iota_T'$, the $T \times T$-dimensional demeaning matrix, and $M_{F,\iota T} = I_T - (F : \iota_T) [(F : \iota_T)' (F : \iota_T)]^{-1} (F : \iota_T)'$, the $T \times T$-dimensional projection matrix that is orthogonal to $F$ and $\iota_T$. It then follows that

$$
\hat{\Sigma}^{-1}\hat{B} / T
= (F' M_{\iota T} F)^{-1} F' M_{\iota T} R M_2 M_3 M_3^{-1} [(R M_2)' M_{F,\iota T} (R M_2)]^{-1} M_3^{-1} M_3 R' R M_{\iota T} F (F' M_{\iota T} F)^{-1}
= (F' M_{\iota T} F)^{-1} F' M_{\iota T} R M_2 [(R M_2)' M_{F,\iota T} (R M_2)]^{-1} M_2 R' R M_{\iota T} F (F' M_{\iota T} F)^{-1},
$$

so $M_3$ just cancels out. This implies that the choice of the (to be subtracted) $N$-th asset return does not affect the rank test statistic. Similarly, it can also be shown that the GRS-FAR statistic is invariant with respect to the choice of the $N$-th asset.

**F. No Zero-Beta Return**

In the case of no zero-beta return, so $\lambda_0 = 0$, the rank and GRS-FAR test statistics can be computed in a manner similar to the general case with a nonzero $\lambda_0$.

In particular, we no longer need to take the asset returns in deviation from the $N$-th asset return and subsequently remove the $N$-th return from the asset return vector. Estimation of the linear factor model $R_t = c + \beta f_t + u_t$ yields

$$
\hat{\beta} = \sum_{t=1}^{T} \tilde{R}_t \tilde{f}_t' \left( \sum_{t=1}^{T} \tilde{f}_t \tilde{f}_t' \right)^{-1}, \quad \hat{\Sigma} = \frac{1}{T} \sum_{t=1}^{T} \tilde{u}_t \tilde{u}_t', \quad \text{and} \quad \hat{Q}_{F F} = \frac{1}{T} \sum_{t=1}^{T} \tilde{f}_t \tilde{f}_t',
$$

where $\tilde{f}_t = f_t - \bar{f}$, $\bar{f} = \frac{1}{T} \sum_{t=1}^{T} f_t$, $\tilde{R}_t = R_t - \bar{R}$, $\bar{R} = \frac{1}{T} \sum_{t=1}^{T} R_t$, and $\tilde{u}_t = (R_t - \bar{R}) - \hat{\beta} (f_t - \bar{f})$ is the residual at time $t$.

Consider $H_0 : \text{rank}(\beta) = k - 1$. Let $rk$ be the smallest root of

$$
\left| \mu \hat{Q}_{F F}^{-1} \hat{\beta}^{\Sigma^{-1} \beta} \right| = 0,
$$

which is identical to the smallest eigenvalue of the matrix $\hat{Q}_{F F} \hat{\beta}^{\Sigma^{-1} \beta}$. In the single-factor case with $k = 1$, $rk = \hat{Q}_{F F} \hat{\beta}^{\Sigma^{-1} \beta}$.
The rank test $F$-statistic is

$$ F(\beta = 0) = \frac{T - N - 1}{N + 1 - k} \times \text{rk} $$

and under $H_0 : \text{rank}(\beta) = k - 1$, i.i.d. normal errors $u_t$, and fixed factors $f_t$,

$$ F(\text{rank}(\beta) = k - 1) = \frac{T - N - 1}{N + 1 - k} \times \text{rk} \leq F(N + 1 - k, T - N - 1). $$

Similarly, the GRS-FAR($\lambda_{f,0}$) statistic is

$$ \frac{T}{1 + \lambda_{f,0}^2 Q^{-1} \lambda_{f,0}} (\bar{R} - \hat{\beta} \lambda_{f,0}) \Sigma^{-1} (\bar{R} - \hat{\beta} \lambda_{f,0}) $$

and

$$ \frac{T - k - N}{T \times N} \times \text{GRS-FAR}(\lambda_{f,0}) \sim F(N, T - k - N). $$

II. Additional Numerical Results for Linear Models

A. Reliability of the FM $t$-test

Figure 3 in the paper plots power curves of 5% significance tests using the FM $t$-test of $H_0 : \lambda_f = 2$ with $N = 31$ and $T = 55$. We use the following linear factor model for the DGP, with $R_t : N \times 1$, $f_t : k \times 1$:

$$ R_t = \tau_N \lambda_0 + \beta (f_t + \lambda_f) + u_t $$

and $f_t \sim NID(0, V_{ff})$, $u_t \sim NID(0, \Omega)$, where $V_{ff}$, $\Omega$, $\lambda_0$, and $\beta$ are calibrated to actual data. Specifically, the $N = 31$ portfolios in Kroencke (2017) are used as test assets, while “Reported,” “P-J,” “Q4-Q4,” “Garbage,” “Unfiltered,” or “$R_m$” is used as the factor, so $k = 1$. We test $H_0 : \lambda_f = 2$ at the 5% significance level using the FM $t$-test.

In addition to the $N = 31$ and $T = 55$ case, we present three more cases in Figures IA1, IA2, IA3: $N = 31$ and $T = 200$, $N = 5$ and $T = 55$, $N = 5$ and $T = 200$, respectively. These figures are all similar to Figure 3 in the main text, so the reliability of the $t$-test remains questionable.
Figure IA1. Power curves of the $t$-test that tests $\lambda_f = 2$ at the 5% level, $N=31$, $T=200$. This figure plots power curves of the $t$-test with the Shanken (1992) correction. The null hypothesis is $\lambda_f = 2$. The significance level is 5%. The DGP is calibrated to the test assets in Kroencke (2017), with “Reported,” “P-J,” “Q4-Q4,” “Garbage,” “Unfiltered,” and “$R_m$” as the risk factor, in Panels A through F, respectively. The number of Monte Carlo replications is 5,000.

Figure IA2. Power curves of the $t$-test that tests $\lambda_f = 2$ at the 5% level, $N=5$, $T=55$. This figure plots power curves of the $t$-test with the Shanken (1992) correction. The null hypothesis is $\lambda_f = 2$. The significance level is 5%. The DGP is calibrated to the first 5 of the 31 test assets in Kroencke (2017), with “Reported,” “P-J,” “Q4-Q4,” “Garbage,” “Unfiltered,” and “$R_m$” as the risk factor, in Panels A through F, respectively. The number of Monte Carlo replications is 5,000.
Figure IA3. Power curves of the $t$-test that tests $\lambda_f = 2$ at the 5% level, $N=5$, $T=200$. This figure plots power curves of the $t$-test with the Shanken (1992) correction. The null hypothesis is $\lambda_f = 2$. The significance level is 5%. The DGP is calibrated to the first 5 of the 31 test assets in Kroencke (2017), with “Reported,” “P-J,” “Q4-Q4,” “Garbage,” “Unfiltered,” and “$R_m$” as the risk factor, in Panels A through F, respectively. The number of Monte Carlo replications is 5,000.

B. Sizes of the Rank Test when $k = 2$ and $k = 3$

Figure IA4 shows the rejection frequencies of Wald and $F$ identification tests at the 5% significance level when there are two and three factors ($k = 2$ or $k = 3$), respectively. The rejection frequencies for $k = 1$ are shown in Figure 2 in the main text.

III. Additional Empirical Results for Linear Models

For robustness, we employ a variety of test assets other than those used in the main paper. These test assets include:

- Ten industry portfolios from Savov (2011) in Section III.A below;
- Thirty portfolios sorted by size, value, and investment, plus the risk-free asset, in Section III.B below;
- Seven test assets chosen from the 31 assets used in Kroencke (2017) in Section III.C below. The seven test assets result from the top and bottom decile values for size, value, and investment sorts, which provides six different asset returns plus the market return.

In all of our robustness checks, we find wide or unbounded 95% confidence sets for the risk premium.
Figure IA4. Rejection frequencies of Wald and F rank tests that test \( \text{rank}(B) = k - 1 \) at the 5% level, \( k=2 \) and \( k=3 \). This figure plots the sizes of rank tests (Wald (blue); F(red)) used to test \( H_0 : \text{rank}(B) = k - 1 \) at the 5% level, where \( B \) is the centered version of \( \beta \) (i.e., remove the \( N \)-th element). In the DGP, the last column of \( \beta \) is set to \( \iota_N \cdot d \), with \( d \in [-3, 3] \). The DGP is calibrated to the first 5 or 30 of the 31 test assets in Kroencke (2017) and the garbage consumption growth in Savov (2011) plus the market factor. The number of Monte Carlo replications is 5,000.
A. Robustness: Industry Portfolios as Test Assets

Table IAI contains our estimation and test results using the 10 industry portfolios.

Table IAI. Risk Premium $\lambda_f$ with 10 Industry Portfolio Returns

The test assets are the 10 industry portfolios over 1960 to 2006 (yearly data) taken from Savov (2011). The estimate of $\lambda_f$ and $t$-statistics result from the Fama-MacBeth (1973) two-pass procedure. The cross-sectional $R^2$ results from the regression of $\bar{R}$ on $(\iota_N ; \beta)$. The pseudo-$R^2$ is a goodness of fit measure that captures the percentage of the variation in the asset returns that is explained by the risk factor. FACCHECK equals the percentage of the variation in the residuals explained by the three largest principal components. FACCHECK of the test assets is 81%. The GRS-FAR test uses the critical value from the $F$-distribution. The rank test $p$-value is based on the $F$-test of $H_0 : \text{rank}(\iota_N ; \beta) = 1$ (or equivalently, $B = 0$).

<table>
<thead>
<tr>
<th></th>
<th>Reported</th>
<th>P-J</th>
<th>Q4-Q4</th>
<th>Garbage</th>
<th>Unfiltered</th>
<th>$R_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate of $\lambda_f$</td>
<td>-0.17</td>
<td>-0.44</td>
<td>-0.33</td>
<td>-0.04</td>
<td>-0.90</td>
<td>-0.45</td>
</tr>
<tr>
<td>FM $t$</td>
<td>-0.27</td>
<td>-0.45</td>
<td>-0.43</td>
<td>-0.04</td>
<td>-0.68</td>
<td>-0.09</td>
</tr>
<tr>
<td>Shanken $t$</td>
<td>-0.27</td>
<td>-0.45</td>
<td>-0.42</td>
<td>-0.04</td>
<td>-0.65</td>
<td>-0.09</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.040</td>
<td>0.168</td>
<td>0.130</td>
<td>0.001</td>
<td>0.281</td>
<td>0.007</td>
</tr>
<tr>
<td>Pseudo-$R^2$</td>
<td>0.007</td>
<td>0.042</td>
<td>0.079</td>
<td>0.224</td>
<td>0.041</td>
<td>0.583</td>
</tr>
<tr>
<td>FACCHECK</td>
<td>81%</td>
<td>80%</td>
<td>79%</td>
<td>76%</td>
<td>80%</td>
<td>69%</td>
</tr>
</tbody>
</table>

95% C.S. of $\lambda_f$

<table>
<thead>
<tr>
<th></th>
<th>(-\infty, \infty)</th>
<th>(-\infty, \infty)</th>
<th>(-\infty, \infty)</th>
<th>(-\infty, \infty)</th>
<th>(-\infty, \infty)</th>
<th>(-16.1, 61.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank Test $p$-value</td>
<td>0.924</td>
<td>0.534</td>
<td>0.887</td>
<td>0.229</td>
<td>0.796</td>
<td>0.033</td>
</tr>
</tbody>
</table>

B. Robustness: Including the Risk-Free Asset as a Test Asset

We use the risk-free asset to augment the 30 portfolios sorted by size, value, and investment from Kroencke (2017), and all portfolio returns are adjusted to be raw returns. We then use these $N = 31$ test assets and rerun our rank and GRS-FAR tests. The findings reported in Table IAI suggest that adding the risk-free asset helps improve identification of the risk premia, especially for the “Garbage” consumption measure. Including the risk-free asset as the $(N + 1)$-th asset is identical to removing the zero-beta return and taking all test asset returns in deviation from the risk-free rate, that is, to using excess returns. The difference between Tables III and IAI is explained by the difference in the test assets since Table III in the paper uses the market return as one of its test assets while Table IAI does not. If identical test assets besides the risk-free rate are used, these tables would be identical. Table IAI shows that the identification problem
is not fully resolved, that is, for some consumption measures, we still have large $p$-values from the rank test, and consequently unbounded confidence sets from the GRS-FAR test.

**Table IIAII. Risk Premium $\lambda_f$ with 30 Portfolios Augmented by the Risk-Free Asset**

The rank test $p$-value results from the $F$-test of $H_0 : \text{rank}(\iota_N : \beta) = 1$ (or equivalently, $B = 0$). The $N = 31$ test assets are the 30 portfolios sorted by size, value, and investment taken from Kroencke (2017), plus the risk-free asset, over 1960 to 2014 (yearly data). Instead of excess returns, all portfolio returns are adjusted to be raw returns for this table.

<table>
<thead>
<tr>
<th></th>
<th>Reported</th>
<th>P-J</th>
<th>Q4-Q4</th>
<th>Garbage</th>
<th>Unfiltered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank Test $p$-value</td>
<td>0.361</td>
<td>0.643</td>
<td>0.266</td>
<td>0.003</td>
<td>0.173</td>
</tr>
<tr>
<td>95% C.S. of $\lambda_f$</td>
<td>$(\infty, \infty)$</td>
<td>$(\infty, \infty)$</td>
<td>$(\infty, \infty)$</td>
<td>(-0.6, 6.8)</td>
<td>$(-\infty, -8.4) \cup (-0.4, \infty)$</td>
</tr>
<tr>
<td>GRS-FAR</td>
<td>$(-\infty, \infty)$</td>
<td>$(-\infty, \infty)$</td>
<td>$(-\infty, 6.8)$</td>
<td>$(-\infty, -8.4) \cup (-0.4, \infty)$</td>
<td>$(-\infty, \infty)$</td>
</tr>
</tbody>
</table>

**C. Robustness: Seven Portfolio Returns**

Table IIAIII reports our findings when using the seven test assets that result from the top and bottom decile values for size, value, and investment sorts, which provides six different asset returns plus the market return. The GRS-FAR test results in unbounded 95% confidence sets for the risk premium for all five consumption measures.

**Table IIAIII. Risk Premium $\lambda_f$ with the $N = 7$ Portfolio Returns**

The $N = 7$ assets are labeled (1), (10), (11), (20), (21), (30), and (31) in Table I of the paper. The rank test $p$-value results from the $F$-test of $H_0 : \text{rank}(\iota_N : \beta) = 1$ (or equivalently, $B = 0$). The various consumption measures and the time periods considered are the same as in Table I.

<table>
<thead>
<tr>
<th></th>
<th>Reported</th>
<th>P-J</th>
<th>Q4-Q4</th>
<th>Garbage</th>
<th>Unfiltered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank Test $p$-value</td>
<td>0.516</td>
<td>0.528</td>
<td>0.669</td>
<td>0.545</td>
<td>0.338</td>
</tr>
<tr>
<td>95% C.S. of $\lambda_f$</td>
<td>$(-\infty, -1.4) \cup (-\infty, -4.8) \cup (-\infty, -2.1) \cup (-\infty, -7.5) \cup (0.5, \infty) \cup (1.0, \infty) \cup (0.4, \infty) \cup (-\infty, \infty) \cup (0.3, \infty)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRS-FAR</td>
<td>$(-\infty, -1.4) \cup (-\infty, -4.8) \cup (-\infty, -2.1) \cup (-\infty, -7.5) \cup (0.5, \infty) \cup (1.0, \infty) \cup (0.4, \infty) \cup (-\infty, \infty) \cup (0.3, \infty)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**D. $p$-value Plots for the Multifactor Models in Tables IV and V**

Figures IIA5 and IIA6 correspond to Table IV in the paper. Similarly, Figures IIA7 and IIA8 correspond to Table V in the paper.
Figure IA5. $p$-values of the subset GRS-FAR test with $R_m$ and $\Delta c$, $T = 55$, $N = 30$. This figure plots $p$-values of the subset GRS-FAR test (red, with $F$-critical values), together with the 5% line (black). The test assets are the 30 portfolios sorted by size, value, and investment over the period 1960 to 2014 (yearly data) taken from Kroencke (2017). The subset GRS-FAR test is provided in the Internet Appendix Section I.D. The null hypothesis is that $\lambda_f$ is equal to the value on the horizontal axis.

Figure IA6. $p$-values of the subset GRS-FAR test with three Fama-French factors, $T = 55$, $N = 30$. This figure plots $p$-values of the subset GRS-FAR test (red, with $F$-critical values), together with the 5% line (black). The test assets are the 30 portfolios sorted by size, value, and investment over the period 1960 to 2014 (yearly data) taken from Kroencke (2017). The subset GRS-FAR test is provided in the Internet Appendix Section I.D. The null hypothesis is that $\lambda_f$ is equal to the value on the horizontal axis.
Figure IA7. $p$-values of the subset GRS-FAR test with $\triangle c$, $cay$, $\triangle c \times cay$, $T = 141$, $N = 25$. This figure plots $p$-values of the subset GRS-FAR test (red, with $F$-critical values), together with the 5% line (black). The test assets are the 25 Fama-French portfolios over the period 1963Q3 to 1998Q3 (quarterly data) taken from Lettau and Ludvigson (2001). The subset GRS-FAR test is provided in the Internet Appendix Section I.D. The null hypothesis is that $\lambda_f$ is equal to the value on the horizontal axis.

Figure IA8. $p$-values of the subset GRS-FAR test with three Fama-French factors, $T = 141$, $N = 25$. This figure plots $p$-values of the subset GRS-FAR test (red, with $F$-critical values), together with the 5% line (black). The test assets are the 25 Fama-French portfolios over the period 1963Q3 to 1998Q3 (quarterly data) taken from Lettau and Ludvigson (2001). The subset GRS-FAR test is provided in the Internet Appendix Section I.D. The null hypothesis is that $\lambda_f$ is equal to the value on the horizontal axis.
IV. Nonlinear Models

A. Stochastic Discount Factors

A generic way to define linear and nonlinear asset pricing models is to use the SDF,

\[ E_t[M_{t+1}(\theta)(1 + R_{i,t+1})] = 1 \iff E_t[M_{t+1}(\theta)R_{i,t+1}^e] = 0, \tag{IA1} \]

where \( M_{t+1}(\theta) \) is the SDF that depends on the \( p \)-dimensional parameter \( \theta \), and \( R_{i,t+1} \) and \( R_{i,t+1}^e \) are the return and excess return on the \( i \)-th asset, respectively. The beta representation in (1) of the paper is referred to as a representation of the linear SDF.\(^1\)

More generally, there exist various consumption-based asset pricing models with SDFs specified in nonlinear forms; see, for example, Epstein and Zin (1991).

With the \( N \) test assets in \( R_{t+1} = (R_{1,t+1}, R_{2,t+1}, ..., R_{N,t+1})' \), we have the \( N \) moment conditions

\[ E_t[M_{t+1}(\theta)(\iota_N + R_{t+1})] = \iota_N, \tag{IA2} \]

which can be used in Hansen’s (1982) GMM to estimate \( \theta \) if \( N \geq p \). The GMM \( t \)-test is commonly used to conduct statistical inference on the parameter of interest. Identical to the FM two-pass \( t \)-test, the distribution of the GMM \( t \)-test is approximated by a normal distribution. We have argued that this approximation is inadequate for the FM two-pass \( t \)-test in many empirically relevant settings. The same applies to the GMM \( t \)-test. An elaborate literature has therefore emerged that proposes inference procedures that remain reliable in such settings; see, for example, Stock and Wright (2000) and Kleibergen (2005).

We use one of the statistics this literature proposes, which is a straightforward extension of the GRS-FAR test: the so-called GMM-AR test; see Stock and Wright (2000). When \( H_0 : \theta = \theta_0 \) holds, the moment condition in (IA2) is satisfied at the hypothesized value \( \theta_0 \), so the sample moment converges to a normally distributed random variable with mean zero under mild conditions. The GMM-AR statistic is then just the quadratic form of

\[ \text{Using the linear SDF, } M_t = \frac{1}{1+\lambda_0}[1 - \lambda_t'var(f_t)^{-1}f_t] \text{ results in the beta representation in (1).} \]

Equation (IA1) results from an optimizing agent in a utility-maximization framework.

\(^1\)
this sample moment weighted by its covariance matrix:

\[
\text{GMM-AR}(\theta_0) = T \bar{f}(\theta_0)' \mathbb{V}(\theta_0)^{-1} \bar{f}(\theta_0)
\]

\[
\overset{d}{\to} \chi^2_N, \text{ under } H_0 \text{ and as } T \to \infty,
\]

where \( \bar{f}(\theta_0) = \frac{1}{T} \sum_{t=1}^{T} f_t(\theta_0) \), \( f_t(\theta_0) = M_{t+1}(\theta_0)(\iota_N + R_{t+1}) - \iota_N \), and \( \mathbb{V}(\theta_0) \) is the \( N \times N \)-dimensional covariance estimator of the covariance matrix of the GMM sample moment that uses the hypothesized value of \( \theta, \theta_0 \). The GMM-AR statistic equals the GMM objective function evaluated at \( H_0 \) using the covariance matrix as the weight matrix, and it can be written in a similar manner if excess returns are used instead of raw returns.\(^2\)

Compared to the GMM \( t \)-test, the GMM-AR test is more reliable, particularly in settings in which the conventional \( t \)-test does not function well, as we will show below.

### B. Tests on Relative Risk Aversion using GMM-AR and GMM \( t \)

**CRRA.** Following existing consumption-based asset pricing literature, we consider the constant relative risk aversion (CRRA) or power utility SDF (see, for example, Cochrane (2001)),

\[
M_{t+1}(\theta) = \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma}, \text{ so } \theta = (\delta, \gamma)
\]

and

\[
E_t \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} (\iota_N + R_{t+1}) \right] = \iota_N,
\]

where \( \delta \) is the discount factor, \( \gamma \) is the rate of relative risk aversion, and \( C_t \) is consumption at time \( t \). The discount factor \( \delta \) is often fixed at 0.95 for convenience, while \( \frac{C_{t+1}}{C_t} \) is set to one plus the rate of consumption growth. The value of \( \gamma \) is of particular interest, since a high value of \( \gamma \) implies implausibly large risk-free interest rates; see, for example, Savov (2011) and Kroencke (2017).

To succinctly compare the GMM-AR test with the commonly used \( t \)-test, we just use simulated power curves, which are shown in Figure IA9. The data for the nonlinear model (IA5) are simulated along the lines of Gospodinov, Kan, and Robotti (2013). Specifically,\(^2\)

\( R_{t+1} \) will be replaced by \( R^e_{t+1} \) while \( \iota_N \) reduces to zero in (IA2) and (IA3).
consumption growth and asset returns are simulated from a joint (log-) normal distribution calibrated to the data of Kroencke (2017). The mean of the asset returns is set to satisfy (IA5); see Gospodinov, Kan, and Robotti (2013). Using the simulated data, we test $H_0 : \gamma = 15$ using the $t$-test and the GMM-AR test, respectively. The discount factor $\delta$ is fixed at its true value of 0.95, so $\gamma$ is the only unknown parameter in the SDF. We consider various values of $N$ and $T$ in Figure IA9, and the $N$ test assets used for calibration are taken from Kroencke (2017) (following their order in the 31 portfolios, starting with the market portfolio). For the $t$-test, the GMM estimator that we employ is the conventional two-step GMM estimator.

Figure IA9 shows that the GMM-AR test performs well for nonlinear settings with $N = 1$ or 2 and $T = 55$, since its rejection frequency is then close to the nominal size of 5%. In contrast, the $t$-test overrejects the null $H_0 : \gamma = 15$ when $T$ is small. When the time span gets large, both tests perform equally well. Therefore, for the empirical settings of Kroencke (2017) with $T = 55$ and $N = 1$ or 2, the GMM-AR test is preferred.

Figure IA9 also shows that when $N$ gets large (e.g., $N = 5$, 10, or 25) while $T$ is only 55 as in Kroencke (2017), the GMM-AR test becomes size-distorted. This results from the large number of elements of the $N \times N$ covariance matrix that have to be estimated. The asymptotics for consistent estimation of covariance matrices when both the time span $T$ and cross-section dimension $N$ become large are such that the time span has to exceed the square of the cross-section dimension, $N^2/T \to 0$; see, for example, Newey and Windmeijer (2009). These conditions are clearly not met for the large cross-section dimensions in Figure IA9, where we observe that the GMM-AR test is size-distorted. For the beta representation of the SDF, we also find such size distortions when using $\chi^2$ critical values (see Figure 4 in the paper), but we could overcome them using the GRS-FAR test, which has an $F$-distribution when the time span exceeds the cross-section dimension, $T > N$. The $F$-distribution of the GRS-FAR test results under i.i.d. normal errors in the linear factor model (15) in the paper. Since the SDF for CRRA is nonlinear, we cannot construct a baseline DGP for the returns on the test assets that implies a well-established finite-sample distribution for the GMM-AR statistic like we did for the GRS-FAR test.
Figure IA9. Power curves of 5% GMM-AR and $t$ tests of $H_0 : \gamma = 15$. This figure plots power curves of GMM-AR (blue) and $t$ (red) at the 5% significance level (black). The moment condition is $E_t \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( \iota_N + R_{t+1} \right) \right] = \iota_N$ with $\delta = 0.95$. The null hypothesis is $H_0 : \gamma = 15$. The DGP is calibrated to the test assets and the unfiltered consumption growth in Kroencke (2017). The $N$ test assets used for calibration are taken from the 31 portfolios in their order starting from the market portfolio. GMM estimation for the $t$-test is conducted using two-step GMM. The number of Monte Carlo replications is 5,000.
Thus, we cannot make assumptions that lead to a known distribution of the GMM-AR test statistic when the time span minorly exceeds the cross-section dimension. Hence, we always have to approximate the distribution of the GMM-AR statistic using asymptotic results. These asymptotic results, however, are still more general than those used for the GMM $t$-test. The GMM $t$-test also requires a so-called rank condition for identification of the SDF parameters that requires the derivative of the moment equation in (IA5) with respect to the SDF parameters to be a full rank matrix. This condition is identical to the full rank condition of $(c_N : \beta)$ for the linear SDF discussed in Section I.A of the main paper. The size distortions of the GMM $t$-test in Figure IA9 show that the asymptotic approximation for the distribution of the GMM $t$-test does not work well even when $N$ is small, that is, $N = 1$ or 2. These size distortions are similar to those of the FM two-pass $t$-test shown in Figure 4 in the paper. In both cases, they result from issues with the rank condition for identification of the parameters in the SDF.

**Epstein-Zin.** In addition to CRRA, we consider the Epstein and Zin (1991) SDF. The simulated power curves are provided in Figure IA10, which is qualitatively similar to Figure IA9.

### C. Additional Empirical Results for GMM-AR: $N = 1$

Using the data depicted by Figure 1, Figure IA11 plots the sample GMM objective function, which is just the GMM-AR statistic in (IA3) using the excess return for different values of $\gamma$, based on moment condition (28) in the paper. The rather flat objective function for large values of $\gamma$ indicates that $\gamma$ is weakly identified. In particular, as $\gamma$ increases, the GMM objective function is close to one, so large values of $\gamma$ also approximately satisfy (26) in a statistical sense.

The 95% confidence sets for $\gamma$ for the different consumption measures then result from the intersection of the GMM sample objective function in Figure IA11, which equals the GMM-AR statistic, at 3.84, the 95% critical value of the $\chi^2_1$ distribution. These results

---

3We note that we can also use the bootstrap to approximate the distribution of the GMM-AR statistic, but it relies on asymptotic arguments that require $N^2/T \to 0$. 

21
Figure IA10. Power curves of GMM-AR and Wald that test $H_0 : \vartheta = 1$ and $\gamma = 20$ with an Epstein-Zin SDF. This figure plots power curves of GMM-AR (blue) and Wald (red) tests at the 5% significance level. The moment condition is $E_t \left[ \delta^\vartheta R_{m,t+1}^{\vartheta-1} \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} (\vartheta N + R_{t+1}) \right] = \vartheta N$. The null hypothesis is $H_0 : \vartheta = 1$ and $\gamma = 20$, while $\delta$ is fixed at 0.95. The data of the market factor, consumption growth and test assets are simulated from the joint (log-) normal distribution. The DGP is calibrated to the test assets in Kroencke (2017) and the garbage measure in Savov (2011). The test assets used for calibration are taken from the first $N$ out of 30 portfolios sorted by size, value and investment. The GMM estimates used for the Wald test on $\vartheta$ and $\gamma$ result from the continuous updating estimator. The number of Monte Carlo replications is 5,000.
Figure IA11. **GMM objective function** (≡GMM-AR statistic) with five consumption measures. This figure plots the GMM objective function based on the condition $E_t \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} R_{m,t+1}^c \right] = 0$ with $\delta = 0.95$. The excess market return is over the period 1960 to 2014 (yearly data), and taken from Kroencke (2017).

Figure IA12. **$p$-values of GMM-AR with five consumption measures.** This figure plots the $p$-value of the GMM-AR test (red), together with the 5% line (black). The GMM-AR test is taken from Stock and Wright (2000), using the condition $E_t \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} R_{m,t+1}^c \right] = 0$ with $\delta = 0.95$. The null hypothesis is that $\gamma$ is equal to the value on the horizontal axis. The excess market return for the GMM-AR test is over the period 1960 to 2014 (yearly data), and taken from Kroencke (2017).
are reported in the last row of Table VI. For any of the consumption measures, the 95% confidence sets do not have an upper bound, so little information on $\gamma$ is available in the different consumption measures. The 95% confidence sets can alternatively be obtained from the $p$-value plots for testing $\gamma$ equal to the value on the horizontal axis in Figure IA12. A $p$-value larger than 0.05 then implies that the null is not rejected at the 5% level. The largest $p$-value in Figure IA12 corresponds to the GMM estimate of $\gamma$ reported in Table VI.

D. DGPs that Satisfy both Linear and CRRA SDFs

Consider the $N \times 1$ vector of moment conditions,

$$E \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} (\iota_N + R_{t+1}) \right] = \iota_N. \tag{IA6}$$

Let $\Delta c_{t+1} = \ln \left( \frac{C_{t+1}}{C_t} \right)$ and $r_{t+1} = \ln(\iota_N + R_{t+1})$, so the moment conditions can also be specified as

$$E \left[ e^{\ln(\delta) - \gamma \Delta c_{t+1} + r_{t+1}} \right] = \iota_N. \tag{IA7}$$

**DGP: log-normal returns.** When log-consumption growth and the log-asset returns are i.i.d. normally distributed,

$$\begin{bmatrix} \Delta c_{t+1} \\ r_{t+1} \end{bmatrix} \sim NID(\mu, V) \equiv NID \begin{bmatrix} 0 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}, \tag{IA8}$$

the pricing error reads

$$e = E \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} (\iota_N + R_{t+1}) \right] - \iota_N \tag{IA9}$$

$$= E \left[ e^{\ln(\delta) - \gamma \Delta c_{t+1} + r_{t+1}} \right] - \iota_N \tag{IA10}$$

$$= e^{\ln(\delta) + \mu_2 + \frac{V_{11} + \gamma^2 V_{12} - 2\gamma V_{21}}{2}} - \iota_N, \tag{IA11}$$

where the last equation results from the moment-generating function of the normal dis-
tribution with \( V_r = \text{Diag}(V_{22}) \) an \( N \times 1 \) vector that contains the diagonal elements of \( V_{22} \). To equate the pricing error to zero, so the moment conditions hold, we set

\[
\mu_2 = -\ln(\delta) - \frac{V_r + \gamma^2 V_{11} - 2\gamma V_{21}}{2}.
\] (IA12)

With \( \beta = V_{21} V_{11}^{-1} \), we then get the linear condition

\[
\mu_2 = \left[ -\ln(\delta) - \frac{V_r + \gamma^2 V_{11}}{2} \right] + \beta \gamma V_{11}.
\] (IA13)

This linear condition implies a misspecified linear SDF (unless \( V_r \propto \iota_N \)) with \( \lambda_f = \gamma V_{11} \), if we use \( r_{t+1} \) as returns and \( \Delta c_{t+1} \) as the risk factor. In contrast, if we use \( R_{t+1} \) (instead of \( r_{t+1} \)) and \( \Delta c_{t+1} \) for the linear SDF, then it is approximately correctly specified. Since \( R_{t+1} = e^{r_{t+1}} - \iota_N \),

\[
E(R_{t+1}) = E(e^{r_{t+1}} - \iota_N) = e^{\left[ -\ln(\delta) - \frac{V_r + \gamma^2 V_{11}}{2} \right] + \beta \gamma V_{11}} - \iota_N = e^{\left[ -\ln(\delta) - \frac{\gamma^2 V_{11}}{2} \right] + \beta \gamma V_{11}} - \iota_N,
\] (IA14)

where the expression of \( E(e^{r_{t+1}}) \) results from the moment generating function of the normal distribution and (IA13).

When \( \left[ -\ln(\delta) - \frac{\gamma^2 V_{11}}{2} \right] + \beta \gamma V_{11} \) is close to zero, we have

\[
E(R_{t+1}) = e^{\left[ -\ln(\delta) - \frac{\gamma^2 V_{11}}{2} \right] + \beta \gamma V_{11}} - \iota_N \approx \left[ -\ln(\delta) - \frac{\gamma^2 V_{11}}{2} \right] \iota_N + \beta \gamma V_{11}.
\] (IA15)

Similarly, if we let \( \beta_R = \text{Cov}(R_{t+1}, \Delta c_{t+1}) V_{11}^{-1} \), then we have

\[
\beta_R = E[(e^{r_{t+1}} - \iota_N) \Delta c_{t+1}] V_{11}^{-1} = E[e^{\beta \Delta c_{t+1}} \Delta c_{t+1}] V_{11}^{-1} = \beta \odot e^{\frac{\gamma}{2} \beta \odot \beta} \approx \beta,
\] (IA16)

where \( \odot \) denotes the element-by-element multiplication, and the approximation holds if \( \frac{\gamma}{2} \beta \odot \beta \) is close to zero, so the linear SDF approximately holds with \( \lambda_0 = -\ln(\delta) - \frac{\gamma^2 V_{11}}{2} \) and \( \lambda_f = \gamma \text{var}(\Delta c_t) \). \(^4\)

\(^4\)Note that if \( x \) is a normal random variable with mean zero, variance \( V \), and constant \( d \), then \( E(x e^{d x}) = d V e^{\frac{d^2}{2}} \). This result helps derive several (messy) expressions (e.g., \( \beta_R, \tilde{\epsilon} \)) in this section.
The above two expressions therefore indicate that the linear SDF with \( R_{t+1} \) (instead of \( r_{t+1} \)) and \( \Delta c_{t+1} \) is approximately correctly specified. As a result, the GRS-FAR test appears to have correct size, as we observe in Figure 11 in the paper.

**DGP: normal returns.** When log-consumption growth and the asset returns are i.i.d. normally distributed,

\[
\begin{bmatrix}
\Delta c_{t+1} \\
R_{t+1}
\end{bmatrix} \sim NID(\tilde{\mu}, \tilde{V}) \equiv NID\left(\begin{bmatrix} 0 \\ \tilde{\mu}_2 \end{bmatrix}, \begin{bmatrix} \tilde{V}_{11} & \tilde{V}_{12} \\
\tilde{V}_{21} & \tilde{V}_{22} \end{bmatrix}\right),
\]

(IA17)

where we use the "\( \sim \)" symbol to indicate that we use the specifications from the DGP: normal returns instead of from the DGP: log-normal returns, so \( \tilde{\beta} = \tilde{V}_{21}\tilde{V}_{11}^{-1} \) in the DGP: normal returns. The pricing error for the nonlinear model now reads

\[
\tilde{e} = E\left[\frac{\delta \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}}{\left(t_N + R_{t+1}\right)} - t_N\right] - t_N
\]

(IA18)

\[
= E\left[\delta e^{-\gamma \Delta c_{t+1}}\right] + E\left[\delta e^{-\gamma \Delta c_{t+1}} R_{t+1}\right] - t_N
\]

(IA19)

\[
= \delta e^{\tilde{\gamma}^2 \tilde{V}_{11}/2} (t_N + \tilde{\mu}_2 - \tilde{\beta} \tilde{V}_{11}) - t_N,
\]

(IA20)

where the last step follows from \( R_{t+1} = \tilde{\mu}_2 + \tilde{\beta} \Delta c_{t+1} + error \) and the properties of the normal distribution.

To equate the pricing error to zero, we thus need \( \tilde{\mu}_2 \) to equal

\[
\tilde{\mu}_2 = \left[ e^{-\gamma^2 \tilde{V}_{11}/2}/\delta - 1 \right] t_N + \tilde{\beta} \tilde{V}_{11}.
\]

(IA21)

The above equation implies a correctly specified linear SDF for \( R_{t+1} \) and \( \Delta c_{t+1} \), with

\[
\lambda_f = \gamma \tilde{V}_{11}, \text{ and } \lambda_0 = \left[ e^{-\gamma^2 \tilde{V}_{11}/2}/\delta - 1 \right].
\]

(IA22)

Put differently, a linear SDF satisfying the above equations is equivalent to a nonlinear CRRA SDF with

\[
\gamma = \frac{\lambda_f}{\tilde{V}_{11}}, \text{ and } \delta = e^{-\gamma^2 \tilde{V}_{11}/2}/(1 + \lambda_0).
\]

(IA23)
REFERENCES


Kleibergen, Frank, 2005, Testing parameters in GMM without assuming that they are identified, Econometrica 73, 1103–1123.

Kleibergen, Frank, Lingwei Kong, and Zhaoguo Zhan, 2019, Identification-robust testing of risk premia in finite samples, Working paper, University of Amsterdam.


