Dissecting Spurious Factors with Cross-Sectional Regressions

Valentina Raponi   Paolo Zaffaroni

Imperial College London   Imperial College London

December 30, 2018

Outline of the talk

1. Motivation and contribution
2. Methodology
3. Generalizations
4. Simulation evidence
5. Conclusion

Motivation

- Two-pass CSR methodology the most popular in empirical finance

Motivation

- Two-pass CSR methodology the most popular in empirical finance
    - Risk-premia estimation and inference

Motivation

- Two-pass CSR methodology the most popular in empirical finance
  - Risk-premia estimation and inference
  - Tests of asset-pricing restriction

Motivation

- Two-pass CSR methodology the most popular in empirical finance
  - Risk-premia estimation and inference
  - Tests of asset-pricing restriction
  - Risk versus characteristics

Motivation

- Two-pass CSR methodology the most popular in empirical finance
  - Risk-premia estimation and inference
  - Tests of asset-pricing restriction
  - Risk versus characteristics
  - Mutual funds/hedge funds applications

Motivation

- Two-pass CSR methodology the most popular in empirical finance
  - Risk-premia estimation and inference
  - Tests of asset-pricing restriction
  - Risk versus characteristics
  - Mutual funds/hedge funds applications
  - Corporate finance (cost of capital)

The two-pass methodology (Fama and MacBeth (1973))

- Let asset returns follow a latent factor structure:

$$R_{it} = ER_{it} + \underbrace{\beta_i' z_t + \epsilon_{it}}_{\text{unobserved innovation}},$$

and assume no-arbitrage and well-diversification of MV frontier.

The two-pass methodology (Fama and MacBeth (1973))

- Let asset returns follow a latent factor structure:

$$R_{it} = ER_{it} + \underbrace{\beta_i' z_t + \epsilon_{it},}_{\text{unobserved innovation}}$$

and assume no-arbitrage and well-diversification of MV frontier.

- Then exact-pricing: $ER_{it} = \gamma_0 + \gamma_1' \beta_i$.

The two-pass methodology (Fama and MacBeth (1973))

- Let asset returns follow a latent factor structure:

$$R_{it} = ER_{it} + \underbrace{\beta_i' z_t + \epsilon_{it}}_{\text{unobserved innovation}},$$

and assume no-arbitrage and well-diversification of MV frontier.

- Then exact-pricing: $ER_{it} = \gamma_0 + \gamma_1' \beta_i$.
- To estimate the risk-premium $\Gamma = (\gamma_0, \gamma_1')'$, two-pass methodology:

The two-pass methodology (Fama and MacBeth (1973))

- Let asset returns follow a latent factor structure:

$$R_{it} = ER_{it} + \underbrace{\beta_i' z_t + \epsilon_{it},}_{\text{unobserved innovation}}$$

and assume no-arbitrage and well-diversification of MV frontier.

- Then exact-pricing: $ER_{it} = \gamma_0 + \gamma_1' \beta_i$.
- To estimate the risk-premium $\Gamma = (\gamma_0, \gamma_1')'$, two-pass methodology:
- First-pass: run N regressions with OLS, one for every asset i,

$$R_{it} = \alpha_i + \beta_i' f_t + \epsilon_{it}, \ \ 1 \leq t \leq T.$$

The two-pass methodology (Fama and MacBeth (1973))

- Let asset returns follow a latent factor structure:

$$R_{it} = ER_{it} + \underbrace{\beta_i' z_t + \epsilon_{it},}_{\text{unobserved innovation}}$$

  and assume no-arbitrage and well-diversification of MV frontier.
- Then exact-pricing: $ER_{it} = \gamma_0 + \gamma_1' \beta_i.$
- To estimate the risk-premium $\Gamma = (\gamma_0, \gamma_1')'$, two-pass methodology:
- First-pass: run N regressions with OLS, one for every asset i,

$$R_{it} = \alpha_i + \beta_i' f_t + \epsilon_{it}, \ 1 \leq t \leq T.$$

- Second-pass: run one cross-sectional regression by OLS/GLS

$$\bar{R}_i = \gamma_0 + \gamma_1 \hat{\beta}_i + \eta_i, \ 1 \leq i \leq N.$$

  where $\bar{R}_i = \sum_{t=1}^{T} R_{it} / T$ and $\hat{\beta}_i$ is OLS estimator of β_i from first-pass.

The two-pass methodology (Fama and MacBeth (1973))

- Let asset returns follow a latent factor structure:

$$R_{it} = ER_{it} + \underbrace{\beta_i' z_t + \epsilon_{it},}_{\text{unobserved innovation}}$$

and assume no-arbitrage and well-diversification of MV frontier.

- Then exact-pricing: $ER_{it} = \gamma_0 + \gamma_1' \beta_i$.
- To estimate the risk-premium $\Gamma = (\gamma_0, \gamma_1')'$, two-pass methodology:
- First-pass: run N regressions with OLS, one for every asset i,

$$R_{it} = \alpha_i + \beta_i' f_t + \epsilon_{it}, \;\; 1 \leq t \leq T.$$

- Second-pass: run one cross-sectional regression by OLS/GLS

$$\bar{R}_i = \gamma_0 + \gamma_1 \hat{\beta}_i + \eta_i, \;\; 1 \leq i \leq N.$$

where $\bar{R}_i = \sum_{t=1}^{T} R_{it} / T$ and $\hat{\beta}_i$ is OLS estimator of β_i from first-pass.
- This gives the estimated risk-premium $\hat{\Gamma} = (\hat{\gamma}_0, \hat{\gamma}_1')'$.

The two-pass methodology: forms of misspecification

- Two-pass great results: simple, understood and works well (more on this later).

The two-pass methodology: forms of misspecification

- Two-pass great results: simple, understood and works well (more on this later).
- However...

The two-pass methodology: forms of misspecification

- Two-pass great results: simple, understood and works well (more on this later).
- However...
- ...one typically assumes that the model is correctly-specified.

The two-pass methodology: forms of misspecification

- Two-pass great results: simple, understood and works well (more on this later).
- However...
- ...one typically assumes that the model is correctly-specified.
- Several ways in which the exact-pricing restriction could be wrong, that is:

$$ER_{it} = \gamma_0 + \gamma_1' \beta_i + e_i$$

for some pricing-errors e_i.

The two-pass methodology: forms of misspecification

- Two-pass great results: simple, understood and works well (more on this later).
- However...
- ...one typically assumes that the model is correctly-specified.
- Several ways in which the exact-pricing restriction could be wrong, that is:

$$ER_{it} = \gamma_0 + \gamma_1' \beta_i + e_i$$

for some pricing-errors e_i.

- The $e_i \neq 0$ for many reasons:

The two-pass methodology: forms of misspecification

- Two-pass great results: simple, understood and works well (more on this later).
- However...
- ...one typically assumes that the model is correctly-specified.
- Several ways in which the exact-pricing restriction could be wrong, that is:

$$ER_{it} = \gamma_0 + \gamma_1' \beta_i + e_i$$

  for some pricing-errors e_i.
- The $e_i \neq 0$ for many reasons:
    - missing (pervasive) factors.

The two-pass methodology: forms of misspecification

- Two-pass great results: simple, understood and works well (more on this later).
- However...
- ...one typically assumes that the model is correctly-specified.
- Several ways in which the exact-pricing restriction could be wrong, that is:

$$ER_{it} = \gamma_0 + \gamma_1' \beta_i + e_i$$

  for some pricing-errors e_i.
- The $e_i \neq 0$ for many reasons:
  - missing (pervasive) factors.
  - mis-measured factors (Roll's critique).

The two-pass methodology: forms of misspecification

- Two-pass great results: simple, understood and works well (more on this later).
- However...
- ...one typically assumes that the model is correctly-specified.
- Several ways in which the exact-pricing restriction could be wrong, that is:

$$ER_{it} = \gamma_0 + \gamma_1' \beta_i + e_i$$

for some pricing-errors e_i.

- The $e_i \neq 0$ for many reasons:
  - missing (pervasive) factors.
  - mis-measured factors (Roll's critique).
  - deviations unrelated to common factors (sentiment/behavioural).

The two-pass methodology: forms of misspecification

- Two-pass great results: simple, understood and works well (more on this later).
- However...
- ...one typically assumes that the model is correctly-specified.
- Several ways in which the exact-pricing restriction could be wrong, that is:

$$ER_{it} = \gamma_0 + \gamma_1' \beta_i + e_i$$

for some pricing-errors e_i.
- The $e_i \neq 0$ for many reasons:
  - missing (pervasive) factors.
  - mis-measured factors (Roll's critique).
  - deviations unrelated to common factors (sentiment/behavioural).
- However...it might be that $e_i = 0$ and yet the model is wrong: useless factors.

The two-pass methodology: useless factors

- Consider special case when presumed beta-pricing models has two factors A and B. Then we think that:

$$ER_{it} = \gamma_0 + \gamma_{1A}\beta_{iA} + \gamma_{1B}\beta_{iB},$$

but in reality only factor A is priced:

$$ER_{it} = \gamma_0 + \gamma_{1A}\beta_{iA} + \gamma_{1B}\beta_{iB} = \gamma_0 + \gamma_{1A}\beta_{iA}.$$

The two-pass methodology: useless factors

- Consider special case when presumed beta-pricing models has two factors A and B. Then we think that:

$$ER_{it} = \gamma_0 + \gamma_{1A}\beta_{iA} + \gamma_{1B}\beta_{iB},$$

  but in reality only factor A is priced:

$$ER_{it} = \gamma_0 + \gamma_{1A}\beta_{iA} + \gamma_{1B}\beta_{iB} = \gamma_0 + \gamma_{1A}\beta_{iA}.$$

- Both equations holds in population!

The two-pass methodology: useless factors

- Consider special case when presumed beta-pricing models has two factors A and B. Then we think that:

$$ER_{it} = \gamma_0 + \gamma_{1A}\beta_{iA} + \gamma_{1B}\beta_{iB},$$

but in reality only factor A is priced:

$$ER_{it} = \gamma_0 + \gamma_{1A}\beta_{iA} + \gamma_{1B}\beta_{iB} = \gamma_0 + \gamma_{1A}\beta_{iA}.$$

- Both equations holds in population!
- So where is the problem?

The two-pass methodology: useless factors

- Problems arise when we try to estimate the risk premia, imposing factor f_{Bt} when in fact $\beta_{iB} = 0$!

The two-pass methodology: useless factors

- Problems arise when we try to estimate the risk premia, imposing factor f_{Bt} when in fact $\beta_{iB} = 0$!
- ...second-pass CSR:

$$\begin{pmatrix} \hat{\gamma}_0 \\ \hat{\gamma}_A \\ \hat{\gamma}_B \end{pmatrix} = \hat{\Gamma} = (\hat{X}'\hat{X})^{-1}\hat{X}'\bar{R},$$

where

$$\hat{X} = (1_N, \hat{\beta}_A, \hat{\beta}_B) \approx (1_N, \hat{\beta}_A, 0_N) \text{ when } T \text{ large .}$$

The two-pass methodology: useless factors

- Problems arise when we try to estimate the risk premia, imposing factor f_{Bt} when in fact $\beta_{iB} = 0$!
- ...second-pass CSR:

$$\begin{pmatrix} \hat{\gamma}_0 \\ \hat{\gamma}_A \\ \hat{\gamma}_B \end{pmatrix} = \hat{\Gamma} = (\hat{X}'\hat{X})^{-1}\hat{X}'\bar{R},$$

where

$$\hat{X} = \left(1_N, \hat{\beta}_A, \hat{\beta}_B\right) \approx \left(1_N, \hat{\beta}_A, 0_N\right) \text{ when } T \text{ large }.$$

- 'denominator' of $\hat{\Gamma}$ arbitrarily close to "zero" (that is, $(\hat{X}'\hat{X})$ becomes singular)!

The two-pass methodology: useless factors

- Problems arise when we try to estimate the risk premia, imposing factor f_{Bt} when in fact $\beta_{iB} = 0$!
- ...second-pass CSR:

$$\begin{pmatrix} \hat{\gamma}_0 \\ \hat{\gamma}_A \\ \hat{\gamma}_B \end{pmatrix} = \hat{\Gamma} = (\hat{X}'\hat{X})^{-1}\hat{X}'\bar{R},$$

where

$$\hat{X} = (1_N, \hat{\beta}_A, \hat{\beta}_B) \approx (1_N, \hat{\beta}_A, 0_N) \text{ when } T \text{ large .}$$

- 'denominator' of $\hat{\Gamma}$ arbitrarily close to "zero" (that is, $(\hat{X}'\hat{X})$ becomes singular)!
- Similar problem when say the $\beta_{iB} \approx$ constant cross-sectionally (documented when B is market factor).

The two-pass methodology: useless factors

- What happens to $\hat{\Gamma}$ as $T \to \infty$?

The two-pass methodology: useless factors

- What happens to $\hat{\Gamma}$ as $T \to \infty$?
- Non-standard behaviour arises.

The two-pass methodology: useless factors

- What happens to $\hat{\Gamma}$ as $T \to \infty$?
- Non-standard behaviour arises.
- Complicated: it depends on whether (netting out f_{tB}) the model is correctly specified
(that is, whether β_{iA} describes entirely the cross-section of expected returns).

The two-pass methodology: useless factors

- What happens to $\hat{\Gamma}$ as $T \to \infty$?
- Non-standard behaviour arises.
- Complicated: it depends on whether (netting out f_{tB}) the model is correctly specified
  (that is, whether β_{iA} describes entirely the cross-section of expected returns).
- Complicated: it depends on the fraction of assets for which $\beta_{iB} = 0$.

The two-pass methodology: useless factors

- Let' simplify even further and assume true model is zero-factor model but we insist and use f_{tB}.

The two-pass methodology: useless factors

- Let' simplify even further and assume true model is zero-factor model but we insist and use f_{tB}.
- Case I (Kan and Zhang (1999)): $\beta_{iB} = 0$

The two-pass methodology: useless factors

- Let' simplify even further and assume true model is zero-factor model but we insist and use f_{tB}.
- Case I (Kan and Zhang (1999)): $\beta_{iB} = 0$
- when model correctly specified ($E(R_{it}) = \gamma_0$)

$$\hat{\gamma}_B \to_d \frac{Z_1' M Z_2}{Z_1' M Z_1} \text{ for zero-mean normal r.v.s } Z_1, Z_2.$$

The two-pass methodology: useless factors

- Let' simplify even further and assume true model is zero-factor model but we insist and use f_{tB}.
- Case I (Kan and Zhang (1999)): $\beta_{iB} = 0$
- when model correctly specified ($E(R_{it}) = \gamma_0$)

$$\hat{\gamma}_B \to_d \frac{Z_1' M Z_2}{Z_1' M Z_1} \text{ for zero-mean normal r.v.s } Z_1, Z_2.$$

- when model is misspecified ($E(R_{it}) = \gamma_0 + e_i$ with non-zero e_i)

$$\hat{\gamma}_B \to_p \pm\infty.$$

The two-pass methodology: useless factors

- Let' simplify even further and assume true model is zero-factor model but we insist and use f_{tB}.
- Case I (Kan and Zhang (1999)): $\beta_{iB} = 0$
- when model correctly specified ($E(R_{it}) = \gamma_0$)

$$\hat{\gamma}_B \to_d \frac{Z_1' M Z_2}{Z_1' M Z_1} \text{ for zero-mean normal r.v.s } Z_1, Z_2.$$

- when model is misspecified ($E(R_{it}) = \gamma_0 + e_i$ with non-zero e_i)

$$\hat{\gamma}_B \to_p \pm\infty.$$

- In particular $\hat{\gamma}_B \approx \sqrt{T}\frac{Z_1' M c}{Z_1' M Z_1}$ where $c = \gamma_0 1_N + e$.

The two-pass methodology: useless factors

- Case II (Kleibergen (2009)): $\beta_{iB} = \beta/\sqrt{T}$ for some $\beta \neq 0$.

The two-pass methodology: useless factors

- Case II (Kleibergen (2009)): $\beta_{iB} = \beta/\sqrt{T}$ for some $\beta \neq 0$.
- when model correctly specified $(E(R_{it}) = \gamma_0)$

$$\hat{\gamma}_B \to_d \frac{(\beta + Z_1)' M(\beta + Z_1)}{(\beta + Z_1)' M(Z_1 + \beta)}$$ for zero-mean normal r.v.s Z_1, Z_2.

The two-pass methodology: useless factors

- Case II (Kleibergen (2009)): $\beta_{iB} = \beta/\sqrt{T}$ for some $\beta \neq 0$.
- when model correctly specified ($E(R_{it}) = \gamma_0$)

$$\hat{\gamma}_B \to_d \frac{(\beta + Z_1)' M (\beta + Z_1)}{(\beta + Z_1)' M (Z_1 + \beta)} \text{ for zero-mean normal r.v.s } Z_1, Z_2.$$

- when model is misspecified ($E(R_{it}) = \gamma_0 + e_i$ with non-zero e_i)

$$\hat{\gamma}_B \to_p \pm\infty.$$

The two-pass methodology: useless factors

- Case II (Kleibergen (2009)): $\beta_{iB} = \beta/\sqrt{T}$ for some $\beta \neq 0$.
- when model correctly specified ($E(R_{it}) = \gamma_0$)

$$\hat{\gamma}_B \to_d \frac{(\beta + Z_1)' M(\beta + Z_1)}{(\beta + Z_1)' M(Z_1 + \beta)} \text{ for zero-mean normal r.v.s } Z_1, Z_2.$$

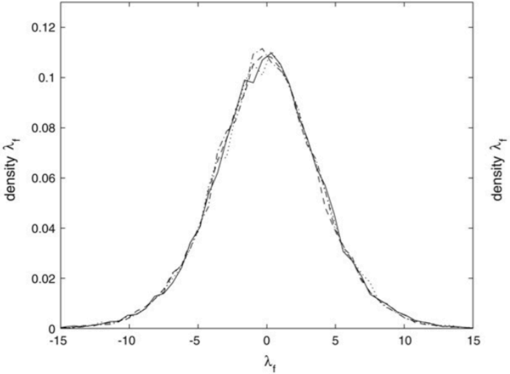
- when model is misspecified ($E(R_{it}) = \gamma_0 + e_i$ with non-zero e_i)

$$\hat{\gamma}_B \to_p \pm\infty.$$

- In particular $\hat{\gamma}_B \approx \sqrt{T}\frac{(\beta+Z_1)'Mc}{(\beta+Z_1)'M(\beta+Z_1)}$.

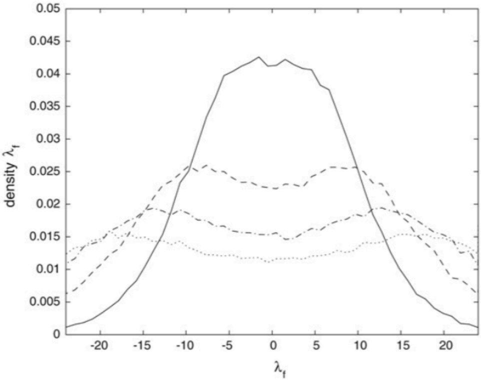
Useless factors (Fig 1.1 from Kleibergen (2009) JOE)

Case $\beta_{iB} = 0$ and correctly specified model



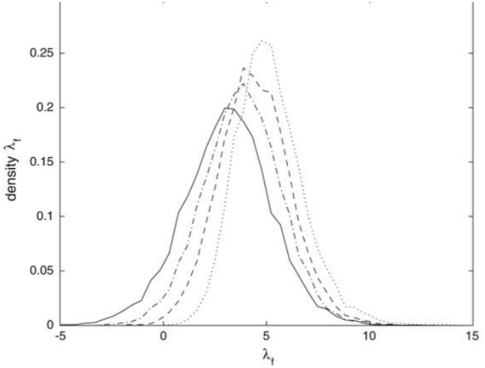
Useless factors (Fig 1.2 from Kleibergen (2009) JOE)

Case $\beta_{iB} = 0$ and misspecified model



Useless factors (Fig 1.3 from Kleibergen (2009) JOE)

Case $\beta_{1B} \neq 0$, $\beta_{iB} = 0$, $2 \leq i \leq N$ and correctly specified model



The two-pass methodology: useless factors

- Some more (peculiar) results.

The two-pass methodology: useless factors

- Some more (peculiar) results.
- Case I (Kan and Zhang 1999): $\beta_{iB} = 0$.

The two-pass methodology: useless factors

- Some more (peculiar) results.
- Case I (Kan and Zhang 1999): $\beta_{iB} = 0$.
- when model is misspecified ($E(R_{it}) = \gamma_0 + e_i$ with non-zero e_i)

$$\begin{aligned} R^2 &\to_d \xi \text{ (some random variable)}, \\ t_{\beta_B} &\to_p \pm\infty. \end{aligned}$$

The two-pass methodology: useless factors

- Some more (peculiar) results.
- Case I (Kan and Zhang 1999): $\beta_{iB} = 0$.
- when model is misspecified ($E(R_{it}) = \gamma_0 + e_i$ with non-zero e_i)

$$\begin{aligned} R^2 &\to_d \xi \text{ (some random variable)}, \\ t_{\beta_B} &\to_p \pm\infty. \end{aligned}$$

- IN SUMMARY: due useless factors inference on beta-pricing models is corrupted using standard CSRs methods valid for large-T.

The two-pass methodology: useless factors

- Some more (peculiar) results.
- Case I (Kan and Zhang 1999): $\beta_{iB} = 0$.
- when model is misspecified ($E(R_{it}) = \gamma_0 + e_i$ with non-zero e_i)

$$\begin{aligned} R^2 &\to_d \xi \text{ (some random variable)}, \\ t_{\beta_B} &\to_p \pm\infty. \end{aligned}$$

- IN SUMMARY: due useless factors inference on beta-pricing models is corrupted using standard CSRs methods valid for large-T.
- Gospodinov et al. (2017): GMM-tests of asset pricing restriction on SDF parameters have power equal to size when useless factors.

The two-pass methodology: useless factors

- Existing methodologies to tackle the effect of useless factors (all designed for large-T except double-asymptotics procedure of Anatolyev and Mikusheva (2018)): ingenious yet sophisticated approaches (non-standard).

The two-pass methodology: useless factors

- Existing methodologies to tackle the effect of useless factors (all designed for large-T except double-asymptotics procedure of Anatolyev and Mikusheva (2018)): ingenious yet sophisticated approaches (non-standard).
- Kleibergen (2009): confidence bounds for testing procedures robust to weak identification (near-zero betas) built on the MLE of Gibbons (1982).

The two-pass methodology: useless factors

- Existing methodologies to tackle the effect of useless factors (all designed for large-T except double-asymptotics procedure of Anatolyev and Mikusheva (2018)): ingenious yet sophisticated approaches (non-standard).
- Kleibergen (2009): confidence bounds for testing procedures robust to weak identification (near-zero betas) built on the MLE of Gibbons (1982).
- Burnside (2016): rank-tests on parameters of factor-SDF.

The two-pass methodology: useless factors

- Existing methodologies to tackle the effect of useless factors (all designed for large-T except double-asymptotics procedure of Anatolyev and Mikusheva (2018)): ingenious yet sophisticated approaches (non-standard).
- Kleibergen (2009): confidence bounds for testing procedures robust to weak identification (near-zero betas) built on the MLE of Gibbons (1982).
- Burnside (2016): rank-tests on parameters of factor-SDF.
- Bryzgalova (2016): penalized (LASSO) version of two-pass procedure.

The two-pass methodology: useless factors

- Existing methodologies to tackle the effect of useless factors (all designed for large-T except double-asymptotics procedure of Anatolyev and Mikusheva (2018)): ingenious yet sophisticated approaches (non-standard).
- Kleibergen (2009): confidence bounds for testing procedures robust to weak identification (near-zero betas) built on the MLE of Gibbons (1982).
- Burnside (2016): rank-tests on parameters of factor-SDF.
- Bryzgalova (2016): penalized (LASSO) version of two-pass procedure.
- Anatolyev and Mikusheva (2018): estimation procedure based on sample-splitting instrumental variables regression robust to weak identification (near-zero betas) and omitted weak factors.

Useless factors: motivation

- This paper: methodology to test for useless factors valid when T is fixed, possibly very small, and number of assets N large.

Useless factors: motivation

- This paper: methodology to test for useless factors valid when T is fixed, possibly very small, and number of assets N large.
- This paper: completely standard as it just uses the plain OLS CSR estimator.

Useless factors: motivation

- This paper: methodology to test for useless factors valid when T is fixed, possibly very small, and number of assets N large.
- This paper: completely standard as it just uses the plain OLS CSR estimator.
- This paper: traditional asymptotics (normal and chi-square limiting distributions immune of nuisance-parameters).

Useless factors: motivation

- This paper: methodology to test for useless factors valid when T is fixed, possibly very small, and number of assets N large.
- This paper: completely standard as it just uses the plain OLS CSR estimator.
- This paper: traditional asymptotics (normal and chi-square limiting distributions immune of nuisance-parameters).
- This paper: distinction between lack of identification (zero betas) and weak identification (quasi-zero betas) irrelevant.

Useless factors: motivation

- Idea: troubles with OLS CSR due to the fact it is a ‘good’ estimator for large T and model correctly specified.

Useless factors: motivation

- Idea: troubles with OLS CSR due to the fact it is a ‘good’ estimator for large T and model correctly specified.
- ....but OLS CSR is not a good estimator when T is fixed (first-order bias!).

Useless factors: motivation

- Idea: troubles with OLS CSR due to the fact it is a ‘good’ estimator for large T and model correctly specified.
- ....but OLS CSR is not a good estimator when T is fixed (first-order bias!).
- How does it behave when T is fixed but one only takes N large?

Useless factors: motivation

- Idea: troubles with OLS CSR due to the fact it is a 'good' estimator for large T and model correctly specified.
- ....but OLS CSR is not a good estimator when T is fixed (first-order bias!).
- How does it behave when T is fixed but one only takes N large?
- This sampling scheme empirically motivated as tens of thousands of assets traded every day (individual assets) but only short time-series used in practice (for data availability; for structural breaks; for time-variation of parameters, etc.)

Useless factors: motivation

- Idea: troubles with OLS CSR due to the fact it is a 'good' estimator for large T and model correctly specified.
- ....but OLS CSR is not a good estimator when T is fixed (first-order bias!).
- How does it behave when T is fixed but one only takes N large?
- This sampling scheme empirically motivated as tens of thousands of assets traded every day (individual assets) but only short time-series used in practice (for data availability; for structural breaks; for time-variation of parameters, etc.)
- Our result: OLS CSR is a powerful tool to dissect useless factors in a large-N environment!

Useless factors: base case

- From now on g_t denotes $K_g \times 1$ useless factor: $cov(g_t, R_{it}) = 0$ all i.

Useless factors: base case

- From now on g_t denotes $K_g \times 1$ useless factor: $cov(g_t, R_{it}) = 0$ all i.
- Assume correctly-specified zero-factor model:

$$ER_{it} = \gamma_0.$$

Useless factors: base case

- From now on g_t denotes $K_g \times 1$ useless factor: $cov(g_t, R_{it}) = 0$ all i.
- Assume correctly-specified zero-factor model:

$$ER_{it} = \gamma_0.$$

- ...but we estimate one-factor model:

$$R_{it} = \alpha_i + \beta_{ig} g_t + \epsilon_{it}.$$

Useless factors: base case

- From now on g_t denotes $K_g \times 1$ useless factor: $cov(g_t, R_{it}) = 0$ all i.
- Assume correctly-specified zero-factor model:

$$ER_{it} = \gamma_0.$$

- ...but we estimate one-factor model:

$$R_{it} = \alpha_i + \beta_{ig} g_t + \epsilon_{it}.$$

- Risk premia OLS CSR estimator:

$$\hat{\Gamma}_g = (\hat{X}_g' \hat{X}_g)^{-1} \hat{X}_g' \bar{R} \text{ with } \hat{X}_g = [1_N, \hat{B}_g].$$

Useless factors: base case

- From now on g_t denotes $K_g \times 1$ useless factor: $cov(g_t, R_{it}) = 0$ all i.
- Assume correctly-specified zero-factor model:

$$ER_{it} = \gamma_0.$$

- ...but we estimate one-factor model:

$$R_{it} = \alpha_i + \beta_{ig} g_t + \epsilon_{it}.$$

- Risk premia OLS CSR estimator:

$$\hat{\Gamma}_g = (\hat{X}_g' \hat{X}_g)^{-1} \hat{X}_g' \bar{R} \text{ with } \hat{X}_g = [1_N, \hat{B}_g].$$

- In particular

$$\hat{\beta}_{ig} = 0_{K_g} + (\tilde{G}'\tilde{G})^{-1} \tilde{G}'\epsilon_i \text{ where } \tilde{G} = G - 1_T \bar{g}'.$$

Useless factors: base case

Theorem

Under Assumptions 1-5 and correct specification:

(i)

$$\hat{\Gamma}_g - \begin{pmatrix} \gamma_0 \\ 0_{K_g} \end{pmatrix} = O_p\left(\frac{1}{\sqrt{N}}\right).$$

(ii)

$$\sqrt{N}\left(\hat{\Gamma}_g - \begin{pmatrix} \gamma_0 \\ 0_K \end{pmatrix}\right) \xrightarrow{d} \mathcal{N}\left(0_{K+1}, V\right)$$

where

$$V = \begin{pmatrix} \frac{\sigma^2}{T} & 0_K' \\ 0_K & \frac{1}{\sigma^4} C' U_\epsilon C \end{pmatrix}, \quad \text{with} \qquad C = \left(\frac{1_T}{T} \otimes \tilde{G}\right).$$

Useless factors: base case

- Remark: the risk premia associated with useless factors go to zero.

Useless factors: base case

- Remark: the risk premia associated with useless factors go to zero.
- Remark: consistent estimation of zero-beta rate.

Useless factors: base case

- Remark: the risk premia associated with useless factors go to zero.
- Remark: consistent estimation of zero-beta rate.
- Remark: asymptotic covariance matrix can be consistently estimated.

Useless factors: base case

- Remark: the risk premia associated with useless factors go to zero.
- Remark: consistent estimation of zero-beta rate.
- Remark: asymptotic covariance matrix can be consistently estimated.
- Remark: correctly-sized Wald test for $H_0 : \gamma_g = 0_{K_g}$.

Useless factors: base case

- Let $\hat{e}_g = \bar{R} - \hat{X}_g \hat{\Gamma}_g$ denote the vector of pricing errors (OLS CSR residuals).

Useless factors: base case

- Let $\hat{e}_g = \bar{R} - \hat{X}_g \hat{\Gamma}_g$ denote the vector of pricing errors (OLS CSR residuals).
- The t statistic for the k-th regression coefficient (where $c_{g,kk}$ denotes the $(k,k)-$th element of the matrix $(\hat{X}_g' \hat{X}_g)^{-1}$) is:

$$t_{g,k} = \frac{\hat{\gamma}_{g,k}}{s_g \cdot \sqrt{c_{g,kk}}}, \qquad 2 \leq k \leq K+1 \text{ with } s_g^2 = \frac{\hat{e}_g' \hat{e}_g}{N-K-1}. \quad (1)$$

Useless factors: base case

- Let $\hat{e}_g = \bar{R} - \hat{X}_g \hat{\Gamma}_g$ denote the vector of pricing errors (OLS CSR residuals).
- The t statistic for the k-th regression coefficient (where $c_{g,kk}$ denotes the $(k,k)-$th element of the matrix $(\hat{X}_g' \hat{X}_g)^{-1}$) is:

$$t_{g,k} = \frac{\hat{\gamma}_{g,k}}{s_g \cdot \sqrt{c_{g,kk}}}, \qquad 2 \leq k \leq K+1 \text{ with } s_g^2 = \frac{\hat{e}_g' \hat{e}_g}{N-K-1}. \tag{1}$$

- The R-squared (where we define $\mathcal{M} = I_N - \frac{1_N 1_N'}{N}$) is:

$$R^2_{CRS_g} = 1 - \frac{\hat{e}_g' \hat{e}_g}{\bar{R}' \mathcal{M} \bar{R}} \tag{2}$$

Useless factors: base case

- Let $\hat{e}_g = \bar{R} - \hat{X}_g \hat{\Gamma}_g$ denote the vector of pricing errors (OLS CSR residuals).
- The t statistic for the k-th regression coefficient (where $c_{g,kk}$ denotes the $(k,k)-$th element of the matrix $(\hat{X}_g' \hat{X}_g)^{-1}$) is:

$$t_{g,k} = \frac{\hat{\gamma}_{g,k}}{s_g \cdot \sqrt{c_{g,kk}}}, \qquad 2 \leq k \leq K+1 \text{ with } s_g^2 = \frac{\hat{e}_g' \hat{e}_g}{N-K-1}. \quad (1)$$

- The R-squared (where we define $\mathcal{M} = I_N - \frac{1_N 1_N'}{N}$) is:

$$R^2_{CRS_g} = 1 - \frac{\hat{e}_g' \hat{e}_g}{\bar{R}' \mathcal{M} \bar{R}} \quad (2)$$

- The F-statistic to test whether all the K coefficients except for the intercept are zero is:

$$F_{CSR_g} = \frac{R^2_{CSR} / K}{(1 - R^2_{CSR}) / (N-K-1)} \quad (3)$$

Useless factors: base case

Useless factors: base case

Theorem

Under Assumptions 1-5 and correct specification:

Useless factors: base case

Theorem

Under Assumptions 1-5 and correct specification:

- (i)

$$t_{g,k} \xrightarrow{d} \mathcal{N}\left(0, \frac{\frac{1}{T}k_4 + \sigma_4}{\sigma^4}\right)$$

Useless factors: base case

Theorem

Under Assumptions 1-5 and correct specification:

- (i)

$$t_{g,k} \xrightarrow{d} \mathcal{N}\left(0, \frac{\frac{1}{T}k_4 + \sigma_4}{\sigma^4}\right)$$

(ii)

$$R^2_{CRS_g} \to 0$$

Useless factors: base case

Theorem

Under Assumptions 1-5 and correct specification:

- (i)

$$t_{g,k} \xrightarrow{d} \mathcal{N}\left(0, \frac{\frac{1}{T}k_4 + \sigma_4}{\sigma^4}\right)$$

(ii)

$$R^2_{CRS_g} \to 0$$

(iii)

$$F_{CSR_g} \xrightarrow{d} \chi^2_K\left(\frac{\frac{k_4}{T} + \sigma_4}{\sigma^4}\right) / K$$

Useless factors: base case

- Inference can be carried out with t and F tests.

Useless factors: base case

- Inference can be carried out with t and F tests.
- When $\sigma_4 = \sigma^4$ and $k_4 = 0$, then $t_{g,k} \xrightarrow{d} \mathcal{N}(0,1)$ and $F_{CSR_g} \xrightarrow{d} \chi^2_K / K$

Useless factors: base case

- Inference can be carried out with t and F tests.
- When $\sigma_4 = \sigma^4$ and $k_4 = 0$, then $t_{g,k} \xrightarrow{d} \mathcal{N}(0,1)$ and $F_{CSR_g} \xrightarrow{d} \chi^2_K / K$
- When $\sigma_4 \neq \sigma^4$, limiting distributions not conventional but can be made so by estimating nuisance parameters: there exists $\hat{\sigma}^2, \hat{\sigma}_4$ such that

$$\hat{\sigma}_4 \rightarrow_p \sigma_4, \hat{\sigma}^2 \rightarrow_p \sigma^2.$$

Useless factors: base case

- Inference can be carried out with t and F tests.
- When $\sigma_4 = \sigma^4$ and $k_4 = 0$, then $t_{g,k} \xrightarrow{d} \mathcal{N}(0,1)$ and $F_{CSR_g} \xrightarrow{d} \chi^2_K / K$
- When $\sigma_4 \neq \sigma^4$, limiting distributions not conventional but can be made so by estimating nuisance parameters: there exists $\hat{\sigma}^2, \hat{\sigma}_4$ such that

$$\hat{\sigma}_4 \rightarrow_p \sigma_4, \hat{\sigma}^2 \rightarrow_p \sigma^2.$$

- R^2 is not inflated (goes to zero, as it should).

Useless factors (base case): misspecified case

- Let $1_N' c/N \to \mu_c$ and $c' M_{1_N} c/N \to \nu_c$ where $ER_{it} = c_i = \gamma_0 + e_i$.

Useless factors (base case): misspecified case

- Let $1_N' c / N \to \mu_c$ and $c' M_{1_N} c / N \to \nu_c$ where $ER_{it} = c_i = \gamma_0 + e_i$.

Theorem

Under Assumptions 1-5 and misspecification:

Useless factors (base case): misspecified case

- Let $1_N' c/N \to \mu_c$ and $c' M_{1_N} c/N \to \nu_c$ where $ER_{it} = c_i = \gamma_0 + e_i$.

Theorem

Under Assumptions 1-5 and misspecification:

(i)

$$\hat{\Gamma} - \begin{pmatrix} \mu_c \\ 0_K \end{pmatrix} = O_p\left(\frac{1}{\sqrt{N}}\right).$$

Useless factors (base case): misspecified case

- Let $1_N' c / N \to \mu_c$ and $c' M_{1_N} c / N \to \nu_c$ where $ER_{it} = c_i = \gamma_0 + e_i$.

Theorem

Under Assumptions 1-5 and misspecification:

(i)

$$\hat{\Gamma} - \begin{pmatrix} \mu_c \\ 0_K \end{pmatrix} = O_p\left(\frac{1}{\sqrt{N}}\right).$$

(ii)

$$\sqrt{N}\left(\hat{\Gamma}_g - \begin{pmatrix} \mu_c \\ 0_K \end{pmatrix}\right) \xrightarrow{d} \mathcal{N}\left(0_{K+1}, V + W\right),$$

where V as for correctly-specified case and $W = \begin{pmatrix} 0 & 0_K' \\ 0_K & \frac{\nu_c}{\sigma^2}\tilde{G}'\tilde{G} \end{pmatrix}$.

Useless factors: base case with misspecification

Useless factors: base case with misspecification

Theorem

Under Assumptions 1-5 and misspecification:

Useless factors: base case with misspecification

Theorem

Under Assumptions 1-5 and misspecification:

- (i)

$$t_{g,k} \xrightarrow{d} \mathcal{N}\left(0, \frac{\nu_c + \frac{\kappa_4 + T\sigma_4}{T^2\sigma^2}}{\nu_c + \frac{\sigma^2}{T}}\right)$$

Useless factors: base case with misspecification

Theorem

Under Assumptions 1-5 and misspecification:

- (i)

$$t_{g,k} \xrightarrow{d} \mathcal{N}\left(0, \frac{\nu_c + \frac{\kappa_4 + T\sigma_4}{T^2\sigma^2}}{\nu_c + \frac{\sigma^2}{T}}\right)$$

(ii)

$$R^2_{CRS_g} \to 0$$

Useless factors: base case with misspecification

Theorem

Under Assumptions 1-5 and misspecification:

- (i)

$$t_{g,k} \xrightarrow{d} \mathcal{N}\left(0, \frac{\nu_c + \frac{\kappa_4 + T\sigma_4}{T^2\sigma^2}}{\nu_c + \frac{\sigma^2}{T}}\right)$$

(ii)

$$R^2_{CRS_g} \rightarrow 0$$

(iii)

$$F_{CSR_g} \xrightarrow{d} \chi^2_K\left(\frac{\nu_c + \frac{\kappa_4 + T\sigma_4}{T^2\sigma^2}}{\nu_c + \frac{\sigma^2}{T}}\right)/K$$

Useless factors (base case): base case with misspecification

- Correctly-specified case obtained for $\nu_c = 0$.

Useless factors (base case): base case with misspecification

- Correctly-specified case obtained for $\nu_c = 0$.
- Qualitatively, the results do not differ from correctly-specified case.

Useless factors (base case): base case with misspecification

- Correctly-specified case obtained for $\nu_c = 0$.
- Qualitatively, the results do not differ from correctly-specified case.
- All quantities can be consistently estimated for $N \to \infty$:

$$\hat{\mu}_c = 1_N' \bar{R} / N, \hat{\nu}_c = 1_N' \bar{R}^2 / N - \hat{\mu}_c^2.$$

Useless factors (base case): FM t-ratios

- How are the traditional FM t-ratios behaving? Before we have seen non-traditional t-ratios.

Useless factors (base case): FM t-ratios

- How are the traditional FM t-ratios behaving? Before we have seen non-traditional t-ratios.
- Let $Z = (Z_0 \cdots Z_K)' \sim N(0_{K+1}, V + W)$ as defined before.

Useless factors (base case): FM t-ratios

- How are the traditional FM t-ratios behaving? Before we have seen non-traditional t-ratios.
- Let $Z = (Z_0 \cdots Z_K)' \sim N(0_{K+1}, V + W)$ as defined before.
- Let $\frac{\epsilon' 1_N}{\sqrt{N}} \to_d \xi \sim N(0, \sigma^2 I_T)$, $\frac{(\epsilon'\epsilon - N\sigma^2 I_T)}{\sqrt{N}} \to_d \Xi$ with $vec(\Xi) \sim N(0, U_\epsilon)$.

Useless factors (base case): FM t-ratios

- How are the traditional FM t-ratios behaving? Before we have seen non-traditional t-ratios.
- Let $Z = (Z_0 \cdots Z_K)' \sim N(0_{K+1}, V + W)$ as defined before.
- Let $\frac{\epsilon' 1_N}{\sqrt{N}} \to_d \xi \sim N(0, \sigma^2 I_T)$, $\frac{(\epsilon'\epsilon - N\sigma^2 I_T)}{\sqrt{N}} \to_d \Xi$ with $vec(\Xi) \sim N(0, U_\epsilon)$.
- Let

$$\Phi_k \equiv \frac{1}{(T-1)^{\frac{1}{2}}} \left(\iota_{k+1,K+1}' (\Sigma_X + \Lambda)^{-1} \left(\begin{pmatrix} \xi' A \xi & \xi' A \Xi \mathcal{P} \\ \mathcal{P}' \Xi A \xi & \mathcal{P}' \Xi A \Xi \mathcal{P} \end{pmatrix} \right) (\Sigma_X + \Lambda)^{-1} \iota_{k+1,K+1} \right)^{\frac{1}{2}}, \quad k = 1, \cdots, K.$$

for $A = I_T - \frac{1_T 1_T'}{T} - \tilde{G}(\tilde{G}'\tilde{G})^{-1}\tilde{G}'$.

Useless factors (base case): FM t-ratios

- How are the traditional FM t-ratios behaving? Before we have seen non-traditional t-ratios.
- Let $Z = (Z_0 \cdots Z_K)' \sim N(0_{K+1}, V + W)$ as defined before.
- Let $\frac{\epsilon' 1_N}{\sqrt{N}} \to_d \xi \sim N(0, \sigma^2 I_T)$, $\frac{(\epsilon'\epsilon - N\sigma^2 I_T)}{\sqrt{N}} \to_d \Xi$ with $vec(\Xi) \sim N(0, U_\epsilon)$.
- Let

$$\Phi_k \equiv \frac{1}{(T-1)^{\frac{1}{2}}} \left(\iota'_{k+1,K+1} (\Sigma_X + \Lambda)^{-1} \left(\begin{pmatrix} \xi' A \xi & \xi' A \Xi \mathcal{P} \\ \mathcal{P}' \Xi A \xi & \mathcal{P}' \Xi A \Xi \mathcal{P} \end{pmatrix} \right) (\Sigma_X + \Lambda)^{-1} \iota_{k+1,K+1} \right)^{\frac{1}{2}}, \quad k = 1, \cdots, K.$$

for $A = I_T - \frac{1_T 1_T'}{T} - \tilde{G}(\tilde{G}'\tilde{G})^{-1}\tilde{G}'$.

- These non-conventional quantities characterize the FM t-ratios when N is large.

Useless factors (base case): FM t-ratios

Theorem

Under Assumptions 1-5:

(i) *for the ex-ante risk premia*

$$|t_{FM}(\hat{\gamma}_0)| = \frac{|\hat{\gamma}_0 - \mu_c|}{SE_0^{FM}} \to_p \frac{Z_0}{\Phi_0} \quad \text{and} \quad \sqrt{N}|t_{FM}(\hat{\gamma}_{1k})| = \sqrt{N}\frac{|\hat{\gamma}_{1k}|}{SE_k^{FM}} \to_d \frac{Z_k}{\hat{\sigma}_k^2}.$$

(ii) *for the ex-post risk premia*

$$|t_{FM}(\hat{\gamma}_0)| = \frac{|\hat{\gamma}_0 - \mu_c|}{SE_0^{FM,P}} \to_d \frac{Z_0}{\Phi_0} \quad \text{and} \quad |t_{FM}(\hat{\gamma}_{1k})| = \frac{|\hat{\gamma}_{1k}|}{SE_k^{FM,P}} \to_d \frac{Z_k}{\Phi_k}.$$

Useless factors (base case): FM t-ratios

- Non-standard distributions arise for all cases.

Useless factors (base case): FM t-ratios

- Non-standard distributions arise for all cases.
- Shanken's correction vanishes as N diverges. The same applies for correctly-specified models without useless factors.

Useless factors (base case): FM t-ratios

- Non-standard distributions arise for all cases.
- Shanken's correction vanishes as N diverges. The same applies for correctly-specified models without useless factors.
- In fixed-T ex post risk premia $\Gamma^P = \Gamma + \bar{f} - Ef_t$ should be considered: however ex ante FM t-ratios goes to zero unlike ex post.

Useless factors (base case): FM t-ratios

- Non-standard distributions arise for all cases.
- Shanken's correction vanishes as N diverges. The same applies for correctly-specified models without useless factors.
- In fixed-T ex post risk premia $\Gamma^P = \Gamma + \bar{f} - Ef_t$ should be considered: however ex ante FM t-ratios goes to zero unlike ex post.
- Same results for correctly-specified (except that $\mu_c = \gamma_0$) and misspecified cases.

Useless factors (base case): FM t-ratios

- Non-standard distributions arise for all cases.
- Shanken's correction vanishes as N diverges. The same applies for correctly-specified models without useless factors.
- In fixed-T ex post risk premia $\Gamma^P = \Gamma + \bar{f} - Ef_t$ should be considered: however ex ante FM t-ratios goes to zero unlike ex post.
- Same results for correctly-specified (except that $\mu_c = \gamma_0$) and misspecified cases.
- Same results (obviously) for zero-beta rate t-ratios.

Useless factors: useful with useless

- The true model is

$$R_t = \alpha + B_f f_t + 0_{N,K_g} g_t + \epsilon_t = \alpha + B_f f_t + \epsilon_t.$$

Useless factors: useful with useless

- The true model is

$$R_t = \alpha + B_f f_t + 0_{N,K_g} g_t + \epsilon_t = \alpha + B_f f_t + \epsilon_t.$$

- Estimated risk premia (setting $\tilde{D} = (\tilde{F}, \tilde{G})$)

$$\hat{\Gamma}_{0fg} = (\hat{X}_{fg}' \hat{X}_{fg})^{-1} \hat{X}_{fg}' \bar{R} \text{ where } (\hat{B}_f, \hat{B}_g) = R' \tilde{D} (\tilde{D}' \tilde{D})^{-1}.$$

Useless factors: useful with useless

- The true model is

$$R_t = \alpha + B_f f_t + 0_{N,K_g} g_t + \epsilon_t = \alpha + B_f f_t + \epsilon_t.$$

- Estimated risk premia (setting $\tilde{D} = (\tilde{F}, \tilde{G})$)

$$\hat{\Gamma}_{0fg} = (\hat{X}_{fg}' \hat{X}_{fg})^{-1} \hat{X}_{fg}' \bar{R} \text{ where } (\hat{B}_f, \hat{B}_g) = R' \tilde{D} (\tilde{D}' \tilde{D})^{-1}.$$

- Under correct specification

$$ER_{it} = \gamma_0 + \gamma_{1f}' \beta_{if}.$$

Useless factors: useful with useless

- The true model is

$$R_t = \alpha + B_f f_t + 0_{N,K_g} g_t + \epsilon_t = \alpha + B_f f_t + \epsilon_t.$$

- Estimated risk premia (setting $\tilde{D} = (\tilde{F}, \tilde{G})$)

$$\hat{\Gamma}_{0fg} = (\hat{X}'_{fg}\hat{X}_{fg})^{-1}\hat{X}'_{fg}\bar{R} \text{ where } (\hat{B}_f, \hat{B}_g) = R'\tilde{D}(\tilde{D}'\tilde{D})^{-1}.$$

- Under correct specification

$$ER_{it} = \gamma_0 + \gamma'_{1f}\beta_{if}.$$

- Under misspecification

$$ER_{it} = \gamma_0 + \gamma'_{1f}\beta_{if} + e_i.$$

Useless factors: useful with useless (G and F orthogonal)

Theorem

When $\tilde{G}'\tilde{F} = 0$ (G and F orthogonal):

(i)

$$\hat{\Gamma}_{fg} - \begin{pmatrix} \gamma_0 + d_0 \\ \gamma_{1_f}^P + d_1 \\ 0_{K_g} \end{pmatrix} = O_p\left(\frac{1}{\sqrt{N}}\right).$$

(ii)

$$\left(\hat{\Gamma}_{fg} - \begin{pmatrix} \gamma_0 + d_0 \\ \gamma_{1_f}^P + d_1 \\ 0_{K_g} \end{pmatrix}\right) \xrightarrow{d} \mathcal{N}\left(0, \left(\Sigma_{X_{fg}} + \Lambda_{fg}\right)^{-1} (V_{fg} + W_{fg}) \left(\Sigma_{X_{fg}} + \Lambda_{fg}\right)^{-1}\right)$$

Useless factors: useful with useless (G and F orthogonal)

Theorem

When $\tilde{G}'\tilde{F} = 0$ (G and F orthogonal):

(i)

$$t_{g,k_g} \xrightarrow{d} \mathcal{N}\left(0, \frac{d_1'\tilde{\Sigma}_{\beta_f}d_1 + \sigma^{-2}W_{[k_g,k_g]}}{\frac{\sigma^2}{T} + \gamma_{1_f}^{P'}\sigma^2(\tilde{F}'\tilde{F})^{-1}D^{-1}\tilde{\Sigma}_{\beta_f}\gamma_{1_f}^{P}}\right).$$

(ii)

$$R^2_{CRS_{fg}} = 1 - \frac{\hat{e}_{fg}'\hat{e}_{fg}}{\bar{R}'\mathcal{M}_N\bar{R}} \to 1 - \frac{\frac{\sigma^2}{T} + \gamma_{1_f}^{P'}\sigma^2(\tilde{F}'\tilde{F})^{-1}D^{-1}\tilde{\Sigma}_{\beta_f}\gamma_{1_f}^{P}}{\frac{\sigma^2}{T} + \gamma_{1_f}^{P'}\tilde{\Sigma}_{\beta_f}\gamma_{1_f}^{P}}$$

Useless factors: useful with useless (G and F orthogonal)

Theorem

When $\tilde{G}'\tilde{F} = 0$ (G and F orthogonal):

(iii) *Let*

$$F_{CSR_{fg}} = \frac{\left(\hat{e}_f^{*'}\hat{e}_f^{*} - \hat{e}_{fg}'\hat{e}_{fg}\right)/K_g}{\hat{e}_{fg}'\hat{e}_{fg}/\left(N - (K_f + K_g + 1)\right)} \tag{4}$$

be the F-statistic to test the null hypotesis $\gamma_{1_g}^P = 0_{K_g}$. Then

$$F_{CSR_{fg}} \xrightarrow{d} \left(Z_1', \quad Z_2'\right) \frac{\mathcal{W}_{fg}/K_g}{\frac{\sigma^2}{T} - d_1'\tilde{\Sigma}_{\beta_f}\gamma_{1_f}^P} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$$

where $Z_1 \equiv \mathcal{N}\left(0_{T^2}, U_\epsilon\right)$ and $Z_2 \equiv \mathcal{N}\left(0_T, \sigma^2 d_1'\tilde{\Sigma}_{\beta_f} d_1 I_T\right)$ are two normally distributed vectors of dimension $T^2 \times 1$ and $T \times 1$, respectively, and where $\mathcal{W}_{fg}$ suitable matrix.

Useless factors: useful with useless (G and F orthogonal)

- Risk premia estimates for F (useful) are first-order biased.

Useless factors: useful with useless (G and F orthogonal)

- Risk premia estimates for F (useful) are first-order biased.
- Instead, risk premia for G (useless) converges to zero.

Useless factors: useful with useless (G and F orthogonal)

- Risk premia estimates for F (useful) are first-order biased.
- Instead, risk premia for G (useless) converges to zero.
- Results more complicated than previous case but similar spirit: all quantities can be consistently estimated and test with correct size and power be derived.

Useless factors: useful with useless (G and F not orthogonal)

- When G and F not orthogonal:

Useless factors: useful with useless (G and F not orthogonal)

- When G and F not orthogonal:
- $$\hat{\Gamma}^P_{fg} \xrightarrow{p} \begin{pmatrix} \gamma_0 - \mu'_{\beta_f}(I_{K_f} - E^{-1}\Sigma_{\beta_f})\gamma^P_{1_f} \\ E^{-1}\Sigma_{\beta_f}\gamma^P_{1_f} \\ AE^{-1}\Sigma_{\beta_f}\gamma^P_{1_f} \end{pmatrix} \quad (5)$$

Useless factors: useful with useless (G and F not orthogonal)

- When G and F not orthogonal:
-

$$\hat{\Gamma}^P_{fg} \quad \xrightarrow{p} \begin{pmatrix} \gamma_0 - \mu'_{\beta_f}(I_{K_f} - E^{-1}\Sigma_{\beta_f})\gamma^P_{1_f} \\ E^{-1}\Sigma_{\beta_f}\gamma^P_{1_f} \\ AE^{-1}\Sigma_{\beta_f}\gamma^P_{1_f} \end{pmatrix} \tag{5}$$

- Set

$$\theta_f = E^{-1}\Sigma_{\beta_f}\gamma^P_{1_f} \text{ and } \theta_g = -\frac{D}{\sigma^2}Q'_{fg}E^{-1}\Sigma_{\beta_f}\gamma^P_{1_f}.$$

Useless factors: useful with useless (G and F not orthogonal)

- When G and F not orthogonal:
-

$$\hat{\Gamma}^P_{fg} \quad \xrightarrow{p} \begin{pmatrix} \gamma_0 - \mu'_{\beta_f}(I_{K_f} - E^{-1}\Sigma_{\beta_f})\gamma^P_{1_f} \\ E^{-1}\Sigma_{\beta_f}\gamma^P_{1_f} \\ AE^{-1}\Sigma_{\beta_f}\gamma^P_{1_f} \end{pmatrix} \tag{5}$$

- Set

$$\theta_f = E^{-1}\Sigma_{\beta_f}\gamma^P_{1_f} \text{ and } \theta_g = -\frac{D}{\sigma^2}Q'_{fg}E^{-1}\Sigma_{\beta_f}\gamma^P_{1_f}.$$

- Under the null of useless factors, the following linear restriction holds:

$$H_0 : \theta_g = A\theta_f.$$

for an observed $A = (\tilde{G}'\tilde{G} - \tilde{G}'\tilde{F}(\tilde{F}'\tilde{F})^{-1}\tilde{F}'\tilde{G})(\tilde{F}'\tilde{F})^{-1}\tilde{F}'\tilde{G}D^{-1}$.

Useless factors: useful with useless (G and F not orthogonal)

- When G and F not orthogonal:
- $$\hat{\Gamma}_{fg}^{P} \quad \xrightarrow{p} \begin{pmatrix} \gamma_0 - \mu'_{\beta_f}(I_{K_f} - E^{-1}\Sigma_{\beta_f})\gamma_{1_f}^{P} \\ E^{-1}\Sigma_{\beta_f}\gamma_{1_f}^{P} \\ AE^{-1}\Sigma_{\beta_f}\gamma_{1_f}^{P} \end{pmatrix} \tag{5}$$
- Set
$$\theta_f = E^{-1}\Sigma_{\beta_f}\gamma_{1_f}^{P} \text{ and } \theta_g = -\frac{D}{\sigma^2}Q'_{fg}E^{-1}\Sigma_{\beta_f}\gamma_{1_f}^{P}.$$
- Under the null of useless factors, the following linear restriction holds:
$$H_0 : \theta_g = A\theta_f.$$
for an observed $A = (\tilde{G}'\tilde{G} - \tilde{G}'\tilde{F}(\tilde{F}'\tilde{F})^{-1}\tilde{F}'\tilde{G})(\tilde{F}'\tilde{F})^{-1}\tilde{F}'\tilde{G}D^{-1}$.
- Using the distribution of part of our theorem, derive the test.

Useless factors: useful with useless (G and F not orthogonal)

- When G and F not orthogonal:
-
$$\hat{\Gamma}^P_{fg} \quad \xrightarrow{p} \begin{pmatrix} \gamma_0 - \mu'_{\beta_f}(I_{K_f} - E^{-1}\Sigma_{\beta_f})\gamma^P_{1_f} \\ E^{-1}\Sigma_{\beta_f}\gamma^P_{1_f} \\ AE^{-1}\Sigma_{\beta_f}\gamma^P_{1_f} \end{pmatrix} \tag{5}$$
- Set
$$\theta_f = E^{-1}\Sigma_{\beta_f}\gamma^P_{1_f} \text{ and } \theta_g = -\frac{D}{\sigma^2}Q'_{fg}E^{-1}\Sigma_{\beta_f}\gamma^P_{1_f}.$$
- Under the null of useless factors, the following linear restriction holds:
$$H_0 : \theta_g = A\theta_f.$$
for an observed $A = (\tilde{G}'\tilde{G} - \tilde{G}'\tilde{F}(\tilde{F}'\tilde{F})^{-1}\tilde{F}'\tilde{G})(\tilde{F}'\tilde{F})^{-1}\tilde{F}'\tilde{G}D^{-1}$.
- Using the distribution of part of our theorem, derive the test.
- If F and G orthogonal in sample, then $H_0 : \theta_g = 0$.

Useless factors: useful with useless (G and F not orthogonal)

- When G and F not orthogonal:
- $$\hat{\Gamma}^P_{fg} \quad \xrightarrow{p} \begin{pmatrix} \gamma_0 - \mu'_{\beta_f}(I_{K_f} - E^{-1}\Sigma_{\beta_f})\gamma^P_{1_f} \\ E^{-1}\Sigma_{\beta_f}\gamma^P_{1_f} \\ AE^{-1}\Sigma_{\beta_f}\gamma^P_{1_f} \end{pmatrix} \tag{5}$$
- Set
$$\theta_f = E^{-1}\Sigma_{\beta_f}\gamma^P_{1_f} \text{ and } \theta_g = -\frac{D}{\sigma^2}Q'_{fg}E^{-1}\Sigma_{\beta_f}\gamma^P_{1_f}.$$
- Under the null of useless factors, the following linear restriction holds:
$$H_0 : \theta_g = A\theta_f.$$
for an observed $A = (\tilde{G}'\tilde{G} - \tilde{G}'\tilde{F}(\tilde{F}'\tilde{F})^{-1}\tilde{F}'\tilde{G})(\tilde{F}'\tilde{F})^{-1}\tilde{F}'\tilde{G}D^{-1}$.
- Using the distribution of part of our theorem, derive the test.
- If F and G orthogonal in sample, then $H_0 : \theta_g = 0$.
- Bias-adjusted estimator for $\gamma^P_{1_f}$ can be obtained (not the focus here).

Useless factors: useful with useless under misspecification

- The true model is still

$$R_t = \alpha + B_f f_t + 0_{N,K_g} g_t + \epsilon_t = \alpha + B_f f_t + \epsilon_t.$$

but we estimate

$R_t = \alpha + B_{f1} f_{1t} + B_g g_t + residual$ setting $F = (F_1, F_2)$(misspecification:

Useless factors: useful with useless under misspecification

- The true model is still

$$R_t = \alpha + B_f f_t + 0_{N,K_g} g_t + \epsilon_t = \alpha + B_f f_t + \epsilon_t.$$

  but we estimate

  $R_t = \alpha + B_{f1} f_{1t} + B_g g_t + residual$ setting $F = (F_1, F_2)$(misspecification:

- As a special case, we could miss out F entirely, so estimated model:

$$R_t = \alpha + B_g g_t + residual.$$

Useless factors: useful with useless and misspecification (G and F orthogonal)

Theorem (i)

$$\hat{\Gamma}_{f_1 g} - \begin{pmatrix} \gamma_0 + \tilde{d}_0 \\ \tilde{d}_{11}\gamma_{1_f}^{P[1]} + \tilde{d}_{12}\gamma_{1_f}^{P[2]} \\ 0_{K_g} \end{pmatrix} = O_p\left(\frac{1}{\sqrt{N}}\right)$$

Useless factors: useful with useless and misspecification (G and F orthogonal)

Theorem

[ii]

$$\sqrt{N}\left(\hat{\Gamma}_{f_1 g} - \begin{pmatrix} \gamma_0 + \tilde{d}_0 \\ \tilde{d}_{11}\gamma_{1_f}^{P[1]} + \tilde{d}_{12}\gamma_{1_f}^{P[2]} \\ 0_{K_g} \end{pmatrix}\right)$$

$$\xrightarrow{d} \mathcal{N}\left(0, \left(\Sigma_{X_{fg}}^{[1]} + \Lambda_{fg}^{[1]}\right)^{-1}\left(V_{fg} + W_{fg}\right)\left(\Sigma_{X_{fg}}^{[1]} + \Lambda_{fg}^{[1]}\right)^{-1}\right)$$

Problem is that acm is function of both F_1 and F_2 so not feasible!

Useless factors: useful with useless and misspecification (G and F orthogonal)

- In particular V_{fg} equal

$$\sigma^2 \left(\frac{1}{T} + (\tilde{d}_{11}\gamma_{1_f}^{P[1]} + \tilde{d}_{12}\gamma_{1_f}^{P[2]})'(\tilde{F}'\tilde{F})^{-1}(\tilde{d}_{11}\gamma_{1_f}^{P[1]} + \tilde{d}_{12}\gamma_{1_f}^{P[2]}) \right) \Sigma_{X_{fg}}^{[1]}$$
$$+\sigma^2\Omega_{fg},$$

with Ω_{fg} equal to

$$\begin{pmatrix} 0 & 0'_{K_{f_1}} & 0'_{K_g} \\ 0_{K_{f_1}} & \vartheta(\tilde{F}^{[1]'}\tilde{F}^{[1]})^{-1} - \left(\tilde{\Sigma}_{\beta_f}^{[1]}\tilde{d}_1 - \tilde{\Sigma}_{\beta_f}^{[1,2]}\gamma_{1_f}^{P[2]}\right) - \left(\tilde{d}_1'\tilde{\Sigma}_{\beta_f}^{[1]} - \gamma_{1_f}^{P[2]'}\tilde{\Sigma}_{\beta_f}^{[2,1]}\right) & 0_{K_{f_1}\times K_g} \\ 0_{K_g} & 0_{K_g\times K_{f_1}} & \vartheta(\tilde{G}'\tilde{G})^{-1} \end{pmatrix}$$

Useless factors: useful with useless and misspecification (G and F orthogonal)

- $W_{f_1 g}$ equal

$$
\begin{array}{cc}
0'_{K_{f_1}} & 0'_{K_g} \\
(Q_f^{[1,2]'} \otimes P_f^{[1]'}) U_\epsilon (Q_f^{[1,2]} \otimes P_f^{[1]}) & (Q_f^{[1,2]'} \otimes P_f^{[1]'}) U_\epsilon (Q_f^{[1,2]} \otimes P_g) \\
(Q_f^{[1,2]'} \otimes P_g') U_\epsilon (Q_f^{[1,2]} \otimes P_f^{[1]}) & (Q_f^{[1,2]'} \otimes P_g') U_\epsilon (Q_f^{[1,2]} \otimes P_g)
\end{array}
\right)
$$

with

$$
Q_f^{[1,2]} = \left(\frac{1_T}{T} - P_f^{[1]} \tilde{d}_{11} \gamma_{1_f}^{P[1]} - P_f^{[1]} \tilde{d}_{12} \gamma_{1_f}^{P[2]} \right).
$$

Useless factors: useful with useless and misspecification (G and F orthogonal)

- $$t_{f_1 g, k_g} = \frac{\hat{\gamma}_{1_g, k_g}}{s_{f_1 g} \cdot \sqrt{c_{g, k_g k_g}}}$$

  is the t-statistic for the k_g-th regression coefficient ($k_g = 1, ..., K_g$) and $c_{g, k_g k_g}$ is the (k_g, k_g)-th element of the matrix $(\hat{X}'_{f_1 g} \hat{X}_{f_1 g})^{-1}$.

Useless factors: useful with useless and misspecification (G and F orthogonal)

- $$t_{f_1 g, k_g} = \frac{\hat{\gamma}_{1_g, k_g}}{s_{f_1 g} \cdot \sqrt{c_{g, k_g k_g}}}$$

  is the t-statistic for the k_g-th regression coefficient ($k_g = 1, \ldots, K_g$) and $c_{g, k_g k_g}$ is the (k_g, k_g)-th element of the matrix $(\hat{X}'_{f_1 g} \hat{X}_{f_1 g})^{-1}$.

- $$R^2_{CSR_{f_1 g}} = 1 - \frac{\hat{e}'_{f_1 g} \hat{e}_{f_1 g}}{\bar{R}' \mathcal{M}_N \bar{R}}.$$

Useless factors: useful with useless and misspecification (G and F orthogonal)

- $$t_{f_1 g, k_g} = \frac{\hat{\gamma}_{1_g, k_g}}{s_{f_1 g} \cdot \sqrt{c_{g, k_g k_g}}}$$

  is the t-statistic for the k_g-th regression coefficient ($k_g = 1, \ldots, K_g$) and $c_{g, k_g k_g}$ is the (k_g, k_g)-th element of the matrix $(\hat{X}'_{f_1 g} \hat{X}_{f_1 g})^{-1}$.

- $$R^2_{CSR_{f_1 g}} = 1 - \frac{\hat{e}'_{f_1 g} \hat{e}_{f_1 g}}{\bar{R}' \mathcal{M}_N \bar{R}}.$$

- $$F_{CSR_{f_1 g}} = \frac{(\hat{e}'^{*}_{f_1} \hat{e}^{*}_{f_1} - \hat{e}'_{f_1 g} \hat{e}_{f_1 g}) / K_g}{\hat{e}'_{f_1 g} \hat{e}_{f_1 g} / (N - K_{f_1} - K_g - 1)},$$

  is the F-statistic to test $\gamma^P_{1_g} = 0_{K_g}$.

Useless factors: useful with useless and misspecification (G and F orthogonal)

Theorem

[i]

$$t_{f_1 g, k_g} \xrightarrow{d} \mathcal{N}\left(0, \frac{\vartheta + \sigma^{-2} W_{[k_g, k_g]}}{\frac{\sigma^2}{T} + \Gamma_{1_f}^{P'} \tilde{\Sigma}_{X_f} \Gamma_{1_f}^{P'}}\right),$$

where $W_{[k_g, k_g]}$ *denotes the* (k_g, k_g)*-th element of the matrix* $(Q_f^{[1,2]'} \otimes \tilde{G}')U_\epsilon(Q_f^{[1,2]} \otimes P_g)$, $\Gamma_{1_f}^P = \left[\gamma_{1_f}^{P[1]'},\ \gamma_{1_f}^{P[2]'}\right]'$ *and*

$$\tilde{\Sigma}_{X_f} = \begin{pmatrix} \tilde{\Sigma}_{\beta_f}^{[1]} - \tilde{\Sigma}_{\beta_f}^{[1]} D^{-1} \tilde{\Sigma}_{\beta_f}^{[1]} & \sigma^2 (\tilde{F}^{[1]'} \tilde{F}^{[1]})^{-1} D^{-1} \tilde{\Sigma}_{\beta_f}^{[1,2]} \\ \sigma^2 \tilde{\Sigma}_{\beta_f}^{[1,2]'} D^{-1} (\tilde{F}^{[1]'} \tilde{F}^{[1]})^{-1} & \tilde{\Sigma}_{\beta_f}^{[2]} - \tilde{\Sigma}_{\beta_f}^{[1,2]'} D^{-1} \tilde{\Sigma}_{\beta_f}^{[1,2]} \end{pmatrix}.$$

Useless factors: useful with useless and misspecification (G and F orthogonal)

Theorem

[ii]

$$R^2_{CSR_{f_1 g}} \xrightarrow{p} 1 - \frac{\frac{\sigma^2}{T} + \Gamma^{P'}_{1_f} \tilde{\Sigma}_{X_f} \Gamma^{P}_{1_f}}{\frac{\sigma^2}{T} + \Gamma^{P'}_{1_f} \tilde{\Sigma}_{\beta_f} \Gamma^{P}_{1_f}}$$

where $\Gamma^P_{1_f} = \left[\gamma^{P[1]'}_{1_f}, \ \gamma^{P[2]'}_{1_f} \right]'$ *and* $\tilde{\Sigma}_{\beta_f} = \begin{pmatrix} \tilde{\Sigma}^{[1]}_{\beta_f} & \tilde{\Sigma}^{[1,2]}_{\beta_f} \\ \tilde{\Sigma}^{[2,1]}_{\beta_f} & \tilde{\Sigma}^{[2]}_{\beta_f}. \end{pmatrix}.$

$$F_{CSR_{f_1 g}} \xrightarrow{d} \left(Z_1', \quad Z_2' \right) \frac{\mathcal{W}_{fg} / K_g}{\frac{\sigma^2}{T} + \Gamma^{P'}_{1_f} \tilde{\Sigma}_{X_f} \Gamma^{P'}_{1_f}} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix},$$

where $Z_1 \equiv \mathcal{N}(0_{T^2}, U_\epsilon)$ *and* $Z_2 \equiv \mathcal{N}(0_T, \vartheta\sigma^2 I_T)$ *are two normally distributed vectors of dimension* $T^2 \times 1$ *and* $T \times 1$.

Useless factors: useful with useless and misspecification

iii. Results extend to G and F not orthogonal.

Useless factors: useful with useless and misspecification

- Results extend to G and F not orthogonal.
- Problem: asymptotic distributions depend on F_2 which is not observed. Bounds can be derived but inaccurate for large N.

Useless factors: useful with useless and misspecification

- Results extend to G and F not orthogonal.
- Problem: asymptotic distributions depend on F_2 which is not observed. Bounds can be derived but inaccurate for large N.
- Solution: estimate the useful factors by PCA and derive asymptotics for useless factors based on the PCA distribution (along the idea of Giglio and Xiu (2017)).

Simulation results: base case

- The table reports the percentage bias (Bias) and root mean squared error (RMSE), all in percent, over 10,000 simulated data sets.

Simulation results: base case

- The table reports the percentage bias (Bias) and root mean squared error (RMSE), all in percent, over 10,000 simulated data sets.
- DGP

$$R_t = \gamma_0 1_N + \epsilon_t,$$

where $\epsilon_t \sim \mathcal{N}(0, \Sigma)$ and where we calibrate γ_0 as $\gamma_0 = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} R_{it}$.

Simulation results: base case

- The table reports the percentage bias (Bias) and root mean squared error (RMSE), all in percent, over 10,000 simulated data sets.
- DGP

$$R_t = \gamma_0 1_N + \epsilon_t,$$

where $\epsilon_t \sim \mathcal{N}(0, \Sigma)$ and where we calibrate γ_0 as $\gamma_0 = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} R_{it}$.

- Fitted Model is a One-Factor Model $R_{it} = a_i + b_i' g_t + u_{it}$, where g_t is the excess market return (from Kenneth French's website) from January 2008 to December 2010 for T=36, and the excess market return from January 2008 to December 2013 for T=72.

Simulation results: base case

- The table reports the percentage bias (Bias) and root mean squared error (RMSE), all in percent, over 10,000 simulated data sets.
- DGP

$$R_t = \gamma_0 1_N + \epsilon_t,$$

  where $\epsilon_t \sim \mathcal{N}(0, \Sigma)$ and where we calibrate γ_0 as $\gamma_0 = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} R_{it}$.
- Fitted Model is a One-Factor Model $R_{it} = a_i + b_i' g_t + u_{it}$, where g_t is the excess market return (from Kenneth French's website) from January 2008 to December 2010 for T=36, and the excess market return from January 2008 to December 2013 for T=72.
- The table also reports the R-squared (R^2) of the fitted model for different cross-sections of $N = 100, 500, 1000, 3000$ stocks.

Table I
Bias and RMSE of the OLS Estimator in a One-Factor Model with a useless factor (Σ scalar)

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
		Panel A: $T = 36$		
Bias($\hat{\gamma}_0$)	0.32%	0.18%	0.12%	0.11%
RMSE($\hat{\gamma}_0$)	0.184	0.083	0.058	0.035
Bias($\hat{\gamma}_1$)	0.000	0.000	0.000	0.000
RMSE($\hat{\gamma}_1$)	0.429	0.191	0.134	0.082
R^2	0.006	0.002	0.001	0.000

Table I
Bias and RMSE of the OLS Estimator in a One-Factor Model with a useless factor (Σ scalar)

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
		Panel B: $T = 72$		
Bias($\hat{\gamma}_0$)	0.05%	0.04%	0.03%	0.04%
RMSE($\hat{\gamma}_0$)	0.146	0.066	0.046	0.028
Bias($\hat{\gamma}_1$)	0.000	0.000	0.000	0.000
RMSE($\hat{\gamma}_1$)	0.379	0.166	0.119	0.072
R^2	0.002	0.002	0.001	0.000

Simulation results: base case (t-test) - Scalar Σ

- The table presents the size properties of t-tests of statistical significance.

Simulation results: base case (t-test) - Scalar Σ

- The table presents the size properties of t-tests of statistical significance.
- The null hypothesis is that the parameter of interest is equal to its true value.

Simulation results: base case (t-test) - Scalar Σ

- The table presents the size properties of t-tests of statistical significance.
- The null hypothesis is that the parameter of interest is equal to its true value.
- $t_{FM}(\cdot)$ denotes the t-statistic associated with the OLS estimator that uses the traditional Fama-MacBeth standard error.

Simulation results: base case (t-test) - Scalar Σ

- The table presents the size properties of t-tests of statistical significance.
- The null hypothesis is that the parameter of interest is equal to its true value.
- $t_{FM}(\cdot)$ denotes the t-statistic associated with the OLS estimator that uses the traditional Fama-MacBeth standard error.
- $t(\cdot)$ denotes the t-statistic associated with the OLS estimator.

Simulation results: base case (t-test) - Scalar Σ

- The table presents the size properties of t-tests of statistical significance.
- The null hypothesis is that the parameter of interest is equal to its true value.
- $t_{FM}(\cdot)$ denotes the t-statistic associated with the OLS estimator that uses the traditional Fama-MacBeth standard error.
- $t(\cdot)$ denotes the t-statistic associated with the OLS estimator.
- The t-statistics are compared with the critical values from a standard normal distribution.

Simulation results: base case (t-test) - Scalar Σ

Table II
Empirical size of t-tests in a One-Factor Model with a useless factor (Σ Scalar)

Panel A: $T = 36$

N	0.10	0.05	0.01	0.10	0.05	0.01
		$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$	
100	0.105	0.053	0.010	0.105	0.053	0.013
500	0.108	0.053	0.011	0.108	0.054	0.011
1000	0.105	0.051	0.011	0.103	0.053	0.011
3000	0.101	0.050	0.010	0.101	0.053	0.011
		$t(\hat{\gamma}_0)$			$t(\hat{\gamma}_1)$	
100	0.105	0.053	0.010	0.105	0.055	0.013
500	0.107	0.052	0.011	0.108	0.054	0.011
1000	0.106	0.051	0.011	0.103	0.053	0.011
3000	0.102	0.050	0.010	0.102	0.052	0.011

Simulation results: base case (t-test) - Scalar Σ

Table II
Empirical size of t-tests in a One-Factor Model with a useless factor (Σ Scalar)

Panel A: $T = 72$

N	0.10	0.05	0.01	0.10	0.05	0.01
		$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$	
100	0.105	0.054	0.011	0.101	0.052	0.011
500	0.105	0.052	0.011	0.0.98	0.049	0.009
1000	0.102	0.052	0.009	0.097	0.051	0.010
3000	0.102	0.051	0.010	0.099	0.050	0.009
		$t(\hat{\gamma}_0)$			$t(\hat{\gamma}_1)$	
100	0.103	0.053	0.009	0.098	0.055	0.011
500	0.103	0.051	0.010	0.098	0.046	0.009
1000	0.101	0.051	0.009	0.096	0.051	0.010
3000	0.101	0.051	0.010	0.099	0.050	0.009

Table III
Empirical size of F-tests in a One-Factor Model with a useless factor (Σ scalar)

The table presents the size properties of F-tests of statistical significance. The F-statistics are compared with the critical values from a $\chi^2_K\left(\frac{\sigma_4}{\sigma^4}/K\right)$.

	Panel A: $T = 36$			Panel A: $T = 72$		
N	0.10	0.05	0.01	0.10	0.05	0.01
100	0.107	0.056	0.012	0.108	0.056	0.012
500	0.101	0.052	0.011	0.104	0.053	0.011
1000	0.101	0.051	0.011	0.101	0.052	0.010
3000	0.100	0.049	0.010	0.101	0.051	0.010

Simulation results: base case (Bias and RMSE) - Diagonal Σ

Table IV
Bias and RMSE of the OLS Estimator One-Factor Model with a useless factor(Σ Diagonal).

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
		Panel A: $T = 36$		
Bias($\hat{\gamma}_0$)	-0.15%	0.15%	0.08%	0.02%
RMSE($\hat{\gamma}_0$)	0.923	0.425	0.308	0.190
Bias($\hat{\gamma}_1$)	0.000	0.000	0.000	0.000
RMSE($\hat{\gamma}_1$)	0.764	0.330	0.227	0.135
R^2	0.030	0.006	0.003	0.001

Simulation results: base case (Bias and RMSE) - Diagonal Σ

Table IV
Bias and RMSE of the OLS Estimator One-Factor Model with a useless factor(Σ Diagonal).

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
		Panel B: $T = 72$		
Bias($\hat{\gamma}_0$)	0.07%	0.03%	0.03%	0.02%
RMSE($\hat{\gamma}_0$)	0.400	0.160	0.127	0.075
Bias($\hat{\gamma}_1$)	0.000	0.000	0.000	0.000
RMSE($\hat{\gamma}_1$)	1.070	0.521	0.332	0.208
R^2	0.069	0.018	0.008	0.003

Table V
Empirical size of t-tests in a One-Factor Model with a useless factor (Σ Diagonal)

Panel A: $T = 36$

N	0.10	0.05	0.01	0.10	0.05	0.01
		$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$	
100	0.103	0.052	0.010	0.113	-0.060	0.015
500	0.101	0.050	0.010	0.101	0.053	0.011
1000	0.101	0.050	0.011	0.103	0.054	0.011
3000	0.100	0.050	0.009	0.102	0.050	0.011
		$t(\hat{\gamma}_0)$			$t(\hat{\gamma}_1)$	
100	0.095	0.048	0.010	0.113	0.059	0.014
500	0.098	0.048	0.009	0.102	0.050	0.011
1000	-0.099	0.050	0.011	0.103	0.052	0.011
3000	0.102	0.050	0.009	0.100	0.050	0.012

Table V
Empirical size of t-tests in a One-Factor Model with a useless factor (Σ Diagonal)

Panel B: $T = 72$

N	0.10	0.05	0.01	0.10	0.05	0.01
		$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$	
100	0.101	0.047	0.007	0.134	0.077	0.022
500	0.100	0.050	0.010	0.108	0.057	0.013
1000	0.098	0.049	0.010	0.103	0.053	0.011
3000	0.099	0.051	0.010	0.101	0.052	0.011
		$t(\hat{\gamma}_0)$			$t(\hat{\gamma}_1)$	
100	0.087	0.041	0.006	0.122	0.072	0.023
500	0.096	0.050	0.009	0.104	0.057	0.012
1000	0.097	0.047	0.010	0.108	0.053	0.013
3000	0.099	0.051	0.010	0.101	0.051	0.010

Simulation results: base case (F-test) - Diagonal Σ

Table VI
Empirical size of F-tests in a One-Factor Model with a useless factor (Σ Diagonal)

The table presents the size properties of F-tests of statistical significance. The F-s
the critical values from a $\chi^2_K\left(\frac{\sigma_4}{\sigma^4}/K\right)$.

	Panel A: $T = 36$			Panel A: $T = 72$		
N	0.10	0.05	0.01	0.10	0.05	0.01
100	0.113	0.060	0.015	0.134	0.077	0.022
500	0.101	0.053	0.011	0.108	0.057	0.013
1000	0.102	0.052	0.011	0.106	0.053	0.011
3000	0.102	0.050	0.011	0.101	0.052	0.011

Simulation results: base case (Bias and RMSE) - Full Σ $(\delta = 0.5)$

Table VII
Bias and RMSE of the OLS Estimator in a One-Factor Model with a useless factor(Σ Full - $\delta = 0.5$).

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
		Panel A: $T = 36$		
Bias$(\hat{\gamma}_0)$	-0.16%	0.13%	0.06%	0.05%
RMSE$(\hat{\gamma}_0)$	0.923	0.425	0.305	0.189
Bias$(\hat{\gamma}_1)$	0.000	0.000	0.000	0.000
RMSE$(\hat{\gamma}_1)$	1.253	0.474	0.349	0.196
R^2	0.031	0.006	0.003	0.001

Simulation results: base case (Bias and RMSE) - Full Σ ($\delta = 0.5$)

Table VII
Bias and RMSE of the OLS Estimator in a One-Factor Model with a useless factor(Σ Full - $\delta = 0.5$).

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
		Panel B: $T = 72$		
Bias($\hat{\gamma}_0$)	-0.03%	0.02%	0.05%	0.03%
RMSE($\hat{\gamma}_0$)	0.353	0.178	0.118	0.078
Bias($\hat{\gamma}_1$)	0.000	0.000	0.000	0.000
RMSE($\hat{\gamma}_1$)	0.764	0.329	0.230	0.138
R^2	0.090	0.015	0.009	0.003

Simulation results: base case (t-test) - Full Σ ($\delta = 0.5$)

Table VIII
Empirical size of t-tests in a One-Factor Model with a useless factor(Σ Full - $\delta = 0.5$)

Panel A: $T = 36$

N	0.10	0.05	0.01	0.10	0.05	0.01
		$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$	
100	0.103	0.053	0.010	0.113	0.060	0.015
500	0.102	0.050	0.010	0.101	0.053	0.011
1000	0.102	0.049	0.010	0.103	0.053	0.011
3000	0.099	0.050	0.010	0.101	0.052	0.011
		$t(\hat{\gamma}_0)$			$t(\hat{\gamma}_1)$	
100	0.096	0.048	0.010	0.113	0.059	0.014
500	0.097	0.048	0.010	0.101	0.051	0.011
1000	0.102	0.048	0.010	0.104	0.052	0.011
3000	0.099	0.050	0.010	0.103	0.051	0.010

Simulation results: base case (t-test) - Full Σ ($\delta = 0.5$)

Table VIII
Empirical size of t-tests in a One-Factor Model with a useless factor(Σ Full - $\delta = 0.5$)

Panel B: $T = 72$

N	0.10	0.05	0.01	0.10	0.05	0.01
		$t_{FM}(\hat{\gamma}_0)$			$t_{FM}(\hat{\gamma}_1)$	
100	0.102	0.046	0.005	0.138	0.080	0.030
500	0.107	0.053	0.009	0.106	0.055	0.011
1000	0.099	0.045	0.009	0.101	0.052	0.014
3000	0.101	0.049	0.011	0.097	0.049	0.012
		$t(\hat{\gamma}_0)$			$t(\hat{\gamma}_1)$	
100	0.087	0.044	0.008	0.129	0.073	0.022
500	0.102	0.050	0.009	0.105	0.052	0.011
1000	0.096	0.047	0.009	0.100	0.050	0.013
3000	0.100	0.049	0.010	0.097	0.049	0.011

Simulation results: base case (F-test) - Full Σ ($\delta = 0.5$)

Table IX
Empirical size of F-tests in a One-Factor Model with a useless factor (Σ Full - $\delta = 0.5$)

	Panel A: $T = 36$			Panel A: $T = 72$		
N	0.10	0.05	0.01	0.10	0.05	0.01
100	0.113	0.060	0.014	0.138	0.080	0.030
500	0.101	0.053	0.011	0.106	0.055	0.011
1000	0.103	0.053	0.011	0.101	0.052	0.014
3000	0.101	0.051	0.011	0.097	0.049	0.012

Simulation results: useful plus useless - scalar Σ

- DGP is

$$R_t = \gamma_0 1_N + B_f(\gamma_1 + f_t - E[f]) + \epsilon_t,$$

where $\epsilon_t \sim \mathcal{N}(0, \sigma^2 I_T)$ and where we calibrate γ_0 and γ_1 as the OLS estimates from the one factor model (CAPM).

- Fitted Model is a Two-Factor Model"

$$R_t = \alpha + B_f f_t + B_g g_t + \epsilon_t,$$

where g_t is an orthogonal (useless) factor to f_t.

- All factors are orthogonalized to each other such that $\tilde{F}'\tilde{G} = 0_{K_f \times K_g}$

Simulation results: useful plus useless (Bias and RMSE) - scalar Σ

Table XIII
Bias and RMSE of the OLS Estimator in a correctly specified model with useful and useless factors (Σ scalar).

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
		Panel A: $T = 36$		
Bias($\hat{\gamma}_0$)	0.78%	0.06%	-0.15%	0.10%
RMSE($\hat{\gamma}_0$)	0.291	0.132	0.071	0.047
Bias($\hat{\gamma}_{1_f}$)	0.43%	0.07%	0.08%	-0.03%
RMSE($\hat{\gamma}_{1_f}$)	0.211	0.102	0.053	0.039
Bias($\hat{\gamma}_{1_g}$)	0.000	0.000	0.000	0.000
RMSE($\hat{\gamma}_{1_g}$)	1.769	0.766	0.543	0.326
Bias(R^2)	4.66%	1.62%	0.38%	0.22%

Simulation results: useful plus useless (Bias and RMSE) - scalar Σ

Table XIII
Bias and RMSE of the OLS Estimator in a correctly specified model with useful and useless factors (Σ scalar).

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
		Panel B: $T = 72$		
Bias$(\hat{\gamma}_0)$	0.10%	0.05%	0.06%	0.07%
RMSE$(\hat{\gamma}_0)$	0.582	0.195	0.079	0.055
Bias$(\hat{\gamma}_{1_f})$	0.27%	0.08%	0.11%	0.03%
RMSE$(\hat{\gamma}_{1_f})$	0.278	0.125	0.052	0.033
Bias$(\hat{\gamma}_{1_g})$	0.000	0.000	0.000	0.000
RMSE$(\hat{\gamma}_{1_g})$	1.215	0.540	0.376	0.227
Bias(R^2)	4.00%	1.57%	0.65%	0.39%

Simulation results: useful plus useless (t-test) - scalar Σ

Table XIV
Empirical Size of t-tests in a correctly specified model with useful and useless factors (Σ Scalar)

Panel A: $T = 36$

N	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
		$t(\hat{\gamma}_0)$			$t(\hat{\gamma}_{1_f})$			$t(\hat{\gamma}_{1_g})$	
100	0.102	0.049	0.011	0.099	0.053	0.009	0.100	0.053	0.012
500	0.102	0.052	0.009	0.098	0.048	0.008	0.099	0.051	0.012
1000	0.100	0.051	0.011	0.098	0.048	0.009	0.100	0.051	0.010
3000	0.099	0.050	0.010	0.101	0.052	0.010	0.099	0.049	0.010

Table XIV
Empirical Size of t-tests in a correctly specified model with useful and useless factors (Σ Scalar)

Panel B: $T = 72$

N	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
		$t(\hat{\gamma}_0)$			$t(\hat{\gamma}_{1_f})$			$t(\hat{\gamma}_{1_g})$	
100	0.097	0.049	0.009	0.101	0.048	0.010	0.110	0.054	0.012
500	0.103	0.051	0.013	0.096	0.048	0.009	0.098	0.048	0.010
1000	0.099	0.052	0.010	0.095	0.049	0.010	0.098	0.048	0.011
3000	0.101	0.052	0.010	0.102	0.052	0.010	0.100	0.049	0.009

Simulation results: useful plus useless (Bias and RMSE) - diagonal Σ

Table XV
Bias and RMSE of the OLS Estimator in a correctly specified model with useful and useless factors (Σ Diagonal).

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
		Panel A: $T = 36$		
Bias($\hat{\gamma}_0$)	0.09%	0.06%	0.02%	0.04%
RMSE($\hat{\gamma}_0$)	0.052	0.034	0.029	0.021
Bias($\hat{\gamma}_{1_f}$)	0.11%	0.05%	0.03%	0.02%
RMSE($\hat{\gamma}_{1_f}$)	0.038	0.029	0.025	0.017
Bias($\hat{\gamma}_{1_g}$)	0.000	0.000	0.000	0.000
RMSE($\hat{\gamma}_{1_g}$)	1.820	1.203	0.958	0.543
Bias(R^2)	2.30%	0.90%	0.23%	0.18%

Simulation results: useful plus useless (Bias and RMSE) - diagonal Σ

Table XV
Bias and RMSE of the OLS Estimator in a correctly specified model with useful and useless factors (Σ Diagonal).

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
	Panel B: $T = 72$			
Bias($\hat{\gamma}_0$)	0.09%	0.03%	0.02%	0.02%
RMSE($\hat{\gamma}_0$)	0.036	0.032	0.029	0.023
Bias($\hat{\gamma}_{1_f}$)	0.13%	0.03%	0.03%	0.03%
RMSE($\hat{\gamma}_{1_f}$)	0.032	0.023	0.019	0.017
Bias($\hat{\gamma}_{1_g}$)	0.001	0.001	0.000	0.000
RMSE($\hat{\gamma}_{1_g}$)	1.807	0.922	0.653	0.392
Bias(R^2)	3.13%	1.07%	0.19%	0.17%

Table XVI

Empirical Size of t-tests in a correctly specified model with useful and useless factors (Σ Diagonal)

Panel A: $T = 36$

N	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
		$t(\hat{\gamma}_0)$			$t(\hat{\gamma}_{1_f})$			$t(\hat{\gamma}_{1_g})$	
100	0.101	0.051	0.011	0.109	0.068	0.016	0.112	0.056	0.016
500	0.101	0.051	0.011	0.084	0.045	0.009	0.106	0.049	0.009
1000	0.102	0.052	0.010	0.099	0.051	0.010	0.103	0.054	0.011
3000	0.099	0.049	0.010	0.099	0.051	0.010	0.099	0.049	0.009

Simulation results: useful plus useless (t-test) - diagonal Σ

Table XVI
Empirical Size of t-tests in a correctly specified model with useful and useless factors (Σ Diagonal)

Panel B: $T = 72$

N	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
		$t(\hat{\gamma}_0)$			$t(\hat{\gamma}_{1_f})$			$t(\hat{\gamma}_{1_g})$	
100	0.079	0.045	0.007	0.073	0.042	0.003	0.106	0.056	0.014
500	0.105	0.054	0.010	0.089	0.045	0.008	0.098	0.053	0.010
1000	0.097	0.049	0.009	0.097	0.048	0.010	0.098	0.049	0.010
3000	0.099	0.049	0.010	0.098	0.049	0.010	0.100	0.049	0.010

Simulation results: useful plus useless (Bias and RMSE) - Full Σ ($\delta = 0.5$)

Table XVII
Bias and RMSE of the OLS Estimator in a correctly specified model with useful and useless factors (Σ Full, $\delta = 0.5$).

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
	Panel A: $T = 36$			
Bias($\hat{\gamma}_0$)	**0.42%**	**0.38%**	**0.09%**	**0.06%**
RMSE($\hat{\gamma}_0$)	**0.086**	**0.042**	**0.035**	**0.021**
Bias($\hat{\gamma}_{1_f}$)	**0.06%**	**0.03%**	**0.04%**	**0.01%**
RMSE($\hat{\gamma}_{1_f}$)	**0.066**	**0.023**	**0.020**	**0.017**
$\hat{\gamma}_{1_g}$	**0.000**	**0.000**	**0.000**	**0.000**
RMSE($\hat{\gamma}_{1_g}$)	**1.211**	**0.916**	**0.903**	**0.543**
R^2	**2.90%**	**1.42%**	**0.49%**	**0.38%**

Simulation results: useful plus useless (Bias and RMSE) - Full Σ ($\delta = 0.5$)

Table XVII
Bias and RMSE of the OLS Estimator in a correctly specified model with useful and useless factors (Σ Full, $\delta = 0.5$).

Statistics	$N = 100$	$N = 500$	$N = 1000$	$N = 3000$
	Panel B: $T = 72$			
Bias$(\hat{\gamma}_0)$	**0.02%**	**0.04%**	**0.01%**	**0.02%**
RMSE$(\hat{\gamma}_0)$	**0.052**	**0.044**	**0.030**	**0.023**
Bias$(\hat{\gamma}_{1_f})$	**0.04%**	**0.07%**	**0.02%**	**0.02%**
RMSE$(\hat{\gamma}_{1_f})$	**0.056**	**0.036**	**0.028**	**0.021**
$\hat{\gamma}_{1_g}$	**0.000**	**0.000**	**0.000**	**0.000**
RMSE$(\hat{\gamma}_{1_g})$	**1.872**	**0.864**	**0.653**	**0.393**
R^2	**2.19%**	**1.29%**	**0.46%**	**0.29%**

Simulation results: useful plus useless (t-test) - Full Σ ($\delta = 0.5$)

Table XVIII
Empirical Size of t-tests in a correctly specified model with useful and useless factors (Σ Full - $\delta = 0.5$)

Panel A: $T = 36$

N	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
		$t(\hat{\gamma}_0)$			$t(\hat{\gamma}_{1_f})$			$t(\hat{\gamma}_{1_g})$	
100	0.127	0.069	0.016	0.126	0.072	0.019	0.102	0.053	0.010
500	0.107	0.055	0.014	0.110	0.054	0.013	0.102	0.052	0.009
1000	0.104	0.052	0.012	0.103	0.049	0.008	0.100	0.050	0.010
3000	0.099	0.048	0.010	0.099	0.051	0.010	0.100	0.050	0.010

Panel B: $T = 72$

N	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
		$t(\hat{\gamma}_0)$			$t(\hat{\gamma}_{1_f})$			$t(\hat{\gamma}_{1_g})$	
100	0.076	0.035	0.007	0.065	0.036	0.006	0.104	0.056	0.013
500	0.088	0.040	0.009	0.088	0.044	0.009	0.102	0.051	0.010
1000	0.095	0.047	0.009	0.096	0.046	0.008	0.102	0.051	0.010
3000	0.099	0.048	0.010	0.099	0.049	0.010	0.099	0.048	0.010

Conclusion

- Framework for testing useless factors within the context of beta-pricing models.

Conclusion

- Framework for testing useless factors within the context of beta-pricing models.
- Designed for when N is large and T is fixed, possibly very small ($T > K$ is enough).

Conclusion

- Framework for testing useless factors within the context of beta-pricing models.
- Designed for when N is large and T is fixed, possibly very small ($T > K$ is enough).
- Unlike the large-T methods, our approach is simple (based simply on the OLS CSR).

Conclusion

- Framework for testing useless factors within the context of beta-pricing models.
- Designed for when N is large and T is fixed, possibly very small ($T > K$ is enough).
- Unlike the large-T methods, our approach is simple (based simply on the OLS CSR).
- Unlike the large-T methods, our results do NOT depend on degree of misspecification.

Conclusion

- Framework for testing useless factors within the context of beta-pricing models.
- Designed for when N is large and T is fixed, possibly very small ($T > K$ is enough).
- Unlike the large-T methods, our approach is simple (based simply on the OLS CSR).
- Unlike the large-T methods, our results do NOT depend on degree of misspecification.
- Our results lead to conventional asymptotic distributions of OLS CSR estimator and test statistics.

THANKS