Spurious Fit in Unidentified Asset Pricing Models

Nikolay Gospodinov
Atlanta Fed

Raymond Kan
U of Toronto

Cesare Robotti
Imperial College

All-Georgia Finance Conference
October 11, 2013

Outline of the talk

1. **Summary** of some seemingly anomalous results that arise in evaluation of possibly misspecified and unidentified linear asset pricing models estimated by maximum likelihood or optimal one-step GMM
2. **Empirical evidence** for some popular asset pricing models
3. **Simulation results** for parameter and specification tests, goodness-of-fit measures
4. **Theory**

General remarks and objectives

- Two observations consistently emerge in the empirical analysis of asset pricing models

General remarks and objectives

- Two observations consistently emerge in the empirical analysis of asset pricing models
  1. Asset pricing models appear to be misspecified

General remarks and objectives

- Two observations consistently emerge in the empirical analysis of asset pricing models
  1. Asset pricing models appear to be misspecified
  2. Asset returns are very noisy and only weakly correlated with the much less volatile macroeconomic (non-traded) factors.
     Some sample correlations: cg_{dur} (0.05), $labor$ (0.04), $cg \cdot cay$ (0.01)

General remarks and objectives

- Two observations consistently emerge in the empirical analysis of asset pricing models
  1. Asset pricing models appear to be misspecified
  2. Asset returns are very noisy and only weakly correlated with the much less volatile macroeconomic (non-traded) factors.
     Some sample correlations: cg_{dur} (0.05), *labor* (0.04), $cg \cdot cay$ (0.01)
- The goal of this research agenda is to assess the effect of these empirical regularities on the evaluation of the asset pricing models and, in particular, on

General remarks and objectives

- Two observations consistently emerge in the empirical analysis of asset pricing models
    1. Asset pricing models appear to be misspecified
    2. Asset returns are very noisy and only weakly correlated with the much less volatile macroeconomic (non-traded) factors.
       Some sample correlations: cg_{dur} (0.05), *labor* (0.04), $cg \cdot cay$ (0.01)
- The goal of this research agenda is to assess the effect of these empirical regularities on the evaluation of the asset pricing models and, in particular, on
    - The fit of the model computed as the squared correlation of the actual and the model-implied expected returns

General remarks and objectives

- Two observations consistently emerge in the empirical analysis of asset pricing models
  1. Asset pricing models appear to be misspecified
  2. Asset returns are very noisy and only weakly correlated with the much less volatile macroeconomic (non-traded) factors.
     Some sample correlations: cg_{dur} (0.05), *labor* (0.04), $cg \cdot cay$ (0.01)
- The goal of this research agenda is to assess the effect of these empirical regularities on the evaluation of the asset pricing models and, in particular, on
  - The fit of the model computed as the squared correlation of the actual and the model-implied expected returns
  - The behavior of the test for correct model specification

General remarks and objectives

- Two observations consistently emerge in the empirical analysis of asset pricing models
  1. Asset pricing models appear to be misspecified
  2. Asset returns are very noisy and only weakly correlated with the much less volatile macroeconomic (non-traded) factors.
     Some sample correlations: cg_{dur} (0.05), *labor* (0.04), $cg \cdot cay$ (0.01)
- The goal of this research agenda is to assess the effect of these empirical regularities on the evaluation of the asset pricing models and, in particular, on
  - The fit of the model computed as the squared correlation of the actual and the model-implied expected returns
  - The behavior of the test for correct model specification
  - The behavior of parameter tests (tests of a zero risk premium or if the risk factor is priced or not)

General remarks and objectives

- Two observations consistently emerge in the empirical analysis of asset pricing models
  1. Asset pricing models appear to be misspecified
  2. Asset returns are very noisy and only weakly correlated with the much less volatile macroeconomic (non-traded) factors.
     Some sample correlations: cg_{dur} (0.05), *labor* (0.04), $cg \cdot cay$ (0.01)
- The goal of this research agenda is to assess the effect of these empirical regularities on the evaluation of the asset pricing models and, in particular, on
  - The fit of the model computed as the squared correlation of the actual and the model-implied expected returns
  - The behavior of the test for correct model specification
  - The behavior of parameter tests (tests of a zero risk premium or if the risk factor is priced or not)
- The stochastic discount factor (SDF) and beta pricing representations of the model are estimated by optimal/invariant (maximum likelihood or continuously-updated GMM) methods

Pre-analysis (prior) beliefs/intuition

- When irrelevant/useless factors (factors that are uncorrelated with asset returns and do not contribute to the pricing ability of the model) are used to explain the cross-sectional variation in asset returns:

Pre-analysis (prior) beliefs/intuition

- When irrelevant/useless factors (factors that are uncorrelated with asset returns and do not contribute to the pricing ability of the model) are used to explain the cross-sectional variation in asset returns:
  1. The fit of the model should be poor

Pre-analysis (prior) beliefs/intuition

- When irrelevant/useless factors (factors that are uncorrelated with asset returns and do not contribute to the pricing ability of the model) are used to explain the cross-sectional variation in asset returns:
  1. The fit of the model should be poor
  2. The test for correct specification should detect model misspecification with high probability

Pre-analysis (prior) beliefs/intuition

- When irrelevant/useless factors (factors that are uncorrelated with asset returns and do not contribute to the pricing ability of the model) are used to explain the cross-sectional variation in asset returns:
  1. The fit of the model should be poor
  2. The test for correct specification should detect model misspecification with high probability
  3. The test of statistical significance should detect with high probability that the factors are not priced

Pre-analysis (prior) beliefs/intuition

- When irrelevant/useless factors (factors that are uncorrelated with asset returns and do not contribute to the pricing ability of the model) are used to explain the cross-sectional variation in asset returns:
  1. The fit of the model should be poor
  2. The test for correct specification should detect model misspecification with high probability
  3. The test of statistical significance should detect with high probability that the factors are not priced
  4. The inference on the other (useful) factors should not be affected (in a fundamental way) by the presence of useless factors

Pre-analysis (prior) beliefs/intuition

- When irrelevant/useless factors (factors that are uncorrelated with asset returns and do not contribute to the pricing ability of the model) are used to explain the cross-sectional variation in asset returns:
  1. The fit of the model should be poor
  2. The test for correct specification should detect model misspecification with high probability
  3. The test of statistical significance should detect with high probability that the factors are not priced
  4. The inference on the other (useful) factors should not be affected (in a fundamental way) by the presence of useless factors
- The intuition for the last two points comes from the regression analysis where the inclusion of irrelevant factors only inflates the variance of the parameter estimates but leaves the asymptotic inference (consistency and asymptotic normality) unchanged

Pre-analysis (prior) beliefs/intuition

- When irrelevant/useless factors (factors that are uncorrelated with asset returns and do not contribute to the pricing ability of the model) are used to explain the cross-sectional variation in asset returns:
  1. The fit of the model should be poor
  2. The test for correct specification should detect model misspecification with high probability
  3. The test of statistical significance should detect with high probability that the factors are not priced
  4. The inference on the other (useful) factors should not be affected (in a fundamental way) by the presence of useless factors
- The intuition for the last two points comes from the regression analysis where the inclusion of irrelevant factors only inflates the variance of the parameter estimates but leaves the asymptotic inference (consistency and asymptotic normality) unchanged
- We show that all of the above conjectures are wrong and the inference in the presence of useless factors is completely spurious!

Post-analysis (posterior) summary of results

- Our analysis reaches some striking conclusions.
  In particular, when useless factors are present and $T \to \infty$:

Post-analysis (posterior) summary of results

- Our analysis reaches some striking conclusions.
  In particular, when useless factors are present and $T \longrightarrow \infty$:
  1. The model exhibits perfect fit ($R^2 \longrightarrow 1$)!

Post-analysis (posterior) summary of results

- Our analysis reaches some striking conclusions.
  In particular, when useless factors are present and $T \to \infty$:
  1. The model exhibits perfect fit ($R^2 \to 1$)!
  2. The model is deemed to be correctly specified (using the LR and OIR tests) with high probability even when the degree of misspecification is arbitrarily large (i.e., the power of the specification tests is equal to their size)

Post-analysis (posterior) summary of results

- Our analysis reaches some striking conclusions.
  In particular, when useless factors are present and $T \to \infty$:
  1. The model exhibits perfect fit ($R^2 \to 1$)!
  2. The model is deemed to be correctly specified (using the LR and OIR tests) with high probability even when the degree of misspecification is arbitrarily large (i.e., the power of the specification tests is equal to their size)
  3. The risk factors, that are useless, are deemed to be priced with probability approaching one

Post-analysis (posterior) summary of results

- Our analysis reaches some striking conclusions.
  In particular, when useless factors are present and $T \to \infty$:
  1. The model exhibits perfect fit ($R^2 \to 1$)!
  2. The model is deemed to be correctly specified (using the LR and OIR tests) with high probability even when the degree of misspecification is arbitrarily large (i.e., the power of the specification tests is equal to their size)
  3. The risk factors, that are useless, are deemed to be priced with probability approaching one
  4. The risk factors, that are useful and priced, are likely to be deemed unpriced

Post-analysis (posterior) summary of results

- Our analysis reaches some striking conclusions.
  In particular, when useless factors are present and $T \to \infty$:
  1. The model exhibits perfect fit ($R^2 \to 1$)!
  2. The model is deemed to be correctly specified (using the LR and OIR tests) with high probability even when the degree of misspecification is arbitrarily large (i.e., the power of the specification tests is equal to their size)
  3. The risk factors, that are useless, are deemed to be priced with probability approaching one
  4. The risk factors, that are useful and priced, are likely to be deemed unpriced

- In summary, an arbitrarily bad model with factors that are independent of asset returns is concluded to be a correctly specified model with a spectacular fit and priced risk factors.

Post-analysis (posterior) summary of results

- Our analysis reaches some striking conclusions.
  In particular, when useless factors are present and $T \rightarrow \infty$:
  1. The model exhibits perfect fit ($R^2 \rightarrow 1$)!
  2. The model is deemed to be correctly specified (using the LR and OIR tests) with high probability even when the degree of misspecification is arbitrarily large (i.e., the power of the specification tests is equal to their size)
  3. The risk factors, that are useless, are deemed to be priced with probability approaching one
  4. The risk factors, that are useful and priced, are likely to be deemed unpriced
- In summary, an arbitrarily bad model with factors that are independent of asset returns is concluded to be a correctly specified model with a spectacular fit and priced risk factors.
- These surprising results bear some similarities to the spurious regression results for nonstationary time series

Some empirical evidence

- The test asset returns are monthly gross returns on the value-weighted 25 Fama-French size and book-to-market ranked portfolios from February 1959 until December 2012.

Some empirical evidence

- The test asset returns are monthly gross returns on the value-weighted 25 Fama-French size and book-to-market ranked portfolios from February 1959 until December 2012.
- Four asset pricing models:

Some empirical evidence

- The test asset returns are monthly gross returns on the value-weighted 25 Fama-French size and book-to-market ranked portfolios from February 1959 until December 2012.
- Four asset pricing models:
  1. Static CAPM: excess market return (vw) as a risk factor

Some empirical evidence

- The test asset returns are monthly gross returns on the value-weighted 25 Fama-French size and book-to-market ranked portfolios from February 1959 until December 2012.
- Four asset pricing models:
  1. Static CAPM: excess market return (*vw*) as a risk factor
  2. Fama-French (1993) three-factor model: *vw*, *smb* and *hml* as risk factors

Some empirical evidence

- The test asset returns are monthly gross returns on the value-weighted 25 Fama-French size and book-to-market ranked portfolios from February 1959 until December 2012.
- Four asset pricing models:
  1. Static CAPM: excess market return (*vw*) as a risk factor
  2. Fama-French (1993) three-factor model: *vw*, *smb* and *hml* as risk factors
  3. Jagannathan-Wang (1996) C-LAB model: *vw*, growth rate of per capita labor income (*labor*) and the lagged default premium (*prem*)

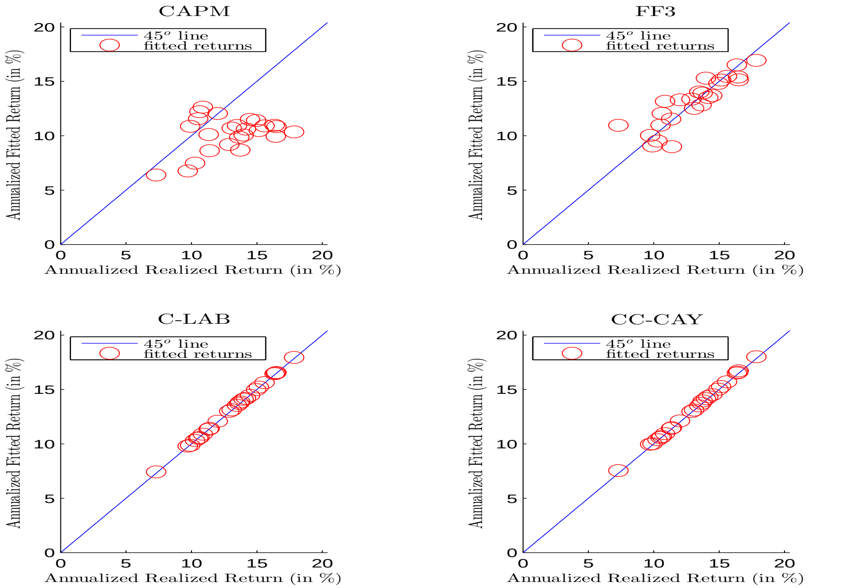
Some empirical evidence

- The test asset returns are monthly gross returns on the value-weighted 25 Fama-French size and book-to-market ranked portfolios from February 1959 until December 2012.
- Four asset pricing models:
    1. Static CAPM: excess market return (vw) as a risk factor
    2. Fama-French (1993) three-factor model: vw, smb and hml as risk factors
    3. Jagannathan-Wang (1996) C-LAB model: vw, growth rate of per capita labor income ($labor$) and the lagged default premium ($prem$)
    4. Lettau-Ludvigson (2001) CC-CAY: real per capita consumption growth (cg), the lagged consumption-aggregate wealth ratio (cay) and an interaction term between these two factors ($cg \cdot cay$)

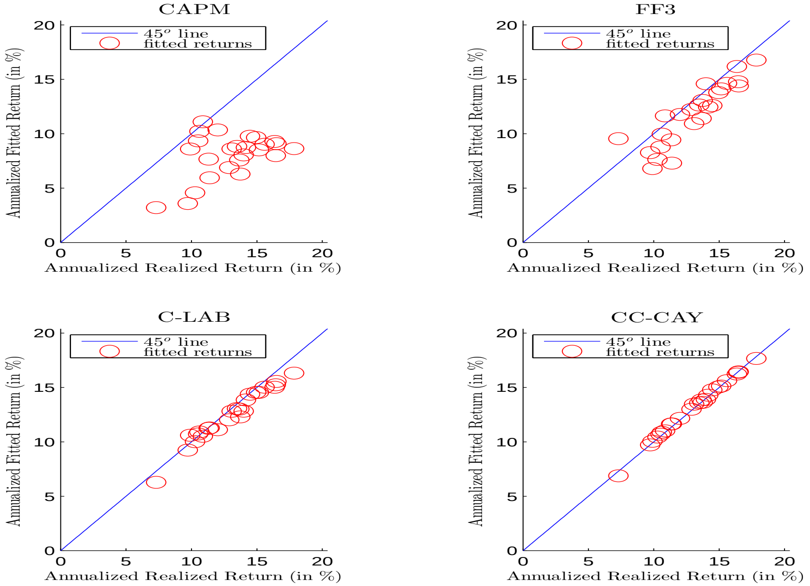
Some empirical evidence

- The test asset returns are monthly gross returns on the value-weighted 25 Fama-French size and book-to-market ranked portfolios from February 1959 until December 2012.
- Four asset pricing models:
  1. Static CAPM: excess market return (vw) as a risk factor
  2. Fama-French (1993) three-factor model: vw, smb and hml as risk factors
  3. Jagannathan-Wang (1996) C-LAB model: vw, growth rate of per capita labor income ($labor$) and the lagged default premium ($prem$)
  4. Lettau-Ludvigson (2001) CC-CAY: real per capita consumption growth (cg), the lagged consumption-aggregate wealth ratio (cay) and an interaction term between these two factors ($cg \cdot cay$)
- The beta and SDF representations of these models are estimated by ML and CU-GMM, respectively.

Realized vs. fitted returns (ML)



Realized vs. fitted returns (CU-GMM)



Empirical evidence

Preliminary evidence on model identification and specification

Rank and HJ-Distance Tests

	CAPM	FF3	C-LAB	CC-CAY
Rank	136.16	87.69	23.26	10.78
(p-value)	(0.0000)	(0.0000)	(0.3873)	(0.9778)
HJD	0.32	0.28	0.32	0.33
(p-value)	(0.0000)	(0.0024)	(0.0000)	(0.0005)

Notes: The null of the rank test is that is that the covariance matrix of the returns and the factors is of reduced rank. The null of the HJD test is that the model is correctly specified.

Empirical evidence

ML (beta representation)

	CAPM	FF3	C-LAB	CC-CAY
LR (p-value)	64.35 (0.0000)	47.29 (0.0009)	22.69 (0.3605)	11.48 (0.9527)
t_{vw}	-3.24	-3.43	-1.34	
t_{smb}		2.08		
t_{hml}		2.33		
t_{labor}			2.81	
t_{prem}			4.21	
t_{cg}				-0.90
t_{cay}				0.76
$t_{cg\cdot cay}$				3.45
R^2	0.1346	0.7677	0.9994	0.9997

Empirical evidence

CU-GMM (SDF framework)

	CAPM	FF3	C-LAB	CC-CAY
OIR (p-value)	64.58 (0.0000)	45.10 (0.0017)	20.58 (0.4848)	10.57 (0.9705)
t_{vw}	4.29	3.92	-0.93	
t_{smb}		-4.22		
t_{hml}		-2.01		
t_{labor}			4.26	
t_{prem}			2.81	
t_{cg}				1.46
t_{cay}				0.85
$t_{cg \cdot cay}$				-3.19
R^2	0.1999	0.7847	0.9595	0.9952

Simulation design

- Three linear models: (1) with a useful factor only, (2) with an irrelevant (useless) factor only and (3) with both factors

Simulation design

- Three linear models: (1) with a useful factor only, (2) with an irrelevant (useless) factor only and (3) with both factors
- The models can be correctly specified or misspecified

Simulation design

- Three linear models: (1) with a useful factor only, (2) with an irrelevant (useless) factor only and (3) with both factors
- The models can be correctly specified or misspecified
- The returns and the useful factor are drawn from a multivariate normal distribution with a covariance matrix set equal to the estimated covariance matrix from the 1959:2–2012:12 sample of monthly gross returns on the 25 Fama-French portfolios and the value-weighted market excess return.

Simulation design

- Three linear models: (1) with a useful factor only, (2) with an irrelevant (useless) factor only and (3) with both factors
- The models can be correctly specified or misspecified
- The returns and the useful factor are drawn from a multivariate normal distribution with a covariance matrix set equal to the estimated covariance matrix from the 1959:2–2012:12 sample of monthly gross returns on the 25 Fama-French portfolios and the value-weighted market excess return.
  - for misspecified models, the means of the simulated returns are set equal to the means of the actual returns

Simulation design

- Three linear models: (1) with a useful factor only, (2) with an irrelevant (useless) factor only and (3) with both factors
- The models can be correctly specified or misspecified
- The returns and the useful factor are drawn from a multivariate normal distribution with a covariance matrix set equal to the estimated covariance matrix from the 1959:2–2012:12 sample of monthly gross returns on the 25 Fama-French portfolios and the value-weighted market excess return.
  - for misspecified models, the means of the simulated returns are set equal to the means of the actual returns
  - for correctly specified models, the means of the returns are set such that the asset pricing model restrictions are satisfied (i.e., the pricing errors are zero)

Simulation design

- Three linear models: (1) with a useful factor only, (2) with an irrelevant (useless) factor only and (3) with both factors
- The models can be correctly specified or misspecified
- The returns and the useful factor are drawn from a multivariate normal distribution with a covariance matrix set equal to the estimated covariance matrix from the 1959:2–2012:12 sample of monthly gross returns on the 25 Fama-French portfolios and the value-weighted market excess return.
  - for misspecified models, the means of the simulated returns are set equal to the means of the actual returns
  - for correctly specified models, the means of the returns are set such that the asset pricing model restrictions are satisfied (i.e., the pricing errors are zero)
  - the mean of the simulated useful factor is calibrated to the mean of the market excess return

Simulation design

- Three linear models: (1) with a useful factor only, (2) with an irrelevant (useless) factor only and (3) with both factors
- The models can be correctly specified or misspecified
- The returns and the useful factor are drawn from a multivariate normal distribution with a covariance matrix set equal to the estimated covariance matrix from the 1959:2–2012:12 sample of monthly gross returns on the 25 Fama-French portfolios and the value-weighted market excess return.
    - for misspecified models, the means of the simulated returns are set equal to the means of the actual returns
    - for correctly specified models, the means of the returns are set such that the asset pricing model restrictions are satisfied (i.e., the pricing errors are zero)
    - the mean of the simulated useful factor is calibrated to the mean of the market excess return
- The useless factor is generated as a standard normal random variable which is uncorrelated with the returns and the useful factor

Simulation evidence (CU-GMM)

Rejection rates of specification test and t-tests of statistical significance

	t_1			t_2			OIR		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
(1)	0.953	0.936	0.889				1.00	1.00	0.999
(3)	0.171	0.096	0.024	1.00	1.00	1.00	0.085	0.040	0.007

Notes: (1) denotes the model with a useful factor only and (3) denotes the model with one useful and one irrelevant (useless) factors. The model is misspecified with a degree of misspecification calibrated to the CAPM estimated from actual data. The sample size is $T = 600$ and the number of Monte Carlo replications is 100,000.

Additional simulation results: Specification tests

Rejection Rates of Specification Tests

T	$\mathcal{J}$ Test (CU-GMM)			$\mathcal{LR}$ Test (ML)		
	10%	5%	1%	10%	5%	1%
Correctly Specified Model (Useful Factor)						
200	0.214	0.131	0.040	0.149	0.081	0.019
600	0.135	0.072	0.017	0.113	0.059	0.013
3600	0.105	0.054	0.011	0.103	0.052	0.011
Misspecified Model (Useful Factor)						
200	0.900	0.831	0.635	0.866	0.781	0.557
600	1.000	1.000	0.999	1.000	1.000	0.998
3600	1.000	1.000	1.000	1.000	1.000	1.000

Additional simulation results: Specification tests

Rejection Rates of Specification Tests

	$\mathcal{J}$ Test (CU-GMM)			$\mathcal{LR}$ Test (ML)		
T	10%	5%	1%	10%	5%	1%
Correctly Specified Model (Irrelevant Factor)						
200	0.030	0.011	0.001	0.014	0.004	0.000
600	0.010	0.003	0.000	0.007	0.002	0.000
3600	0.006	0.001	0.000	0.005	0.001	0.000
Misspecified Model (Irrelevant Factor)						
200	0.130	0.063	0.011	0.105	0.050	0.008
600	0.113	0.057	0.011	0.105	0.052	0.011
3600	0.103	0.052	0.010	0.103	0.052	0.010

Additional simulation results: R^2

Empirical Distribution of the R^2 coefficient (CU-GMM)

T	mean	1%	5%	10%	50%	90%	95%	99%
				Misspecified Model (Useful Factor)				
200	0.298	0.000	0.003	0.012	0.251	0.669	0.755	0.871
600	0.214	0.000	0.003	0.011	0.176	0.481	0.563	0.692
3600	0.172	0.012	0.041	0.062	0.164	0.293	0.332	0.404
				Misspecified Model (Irrelevant Factor)				
200	0.900	0.342	0.658	0.770	0.944	0.983	0.988	0.993
600	0.989	0.929	0.966	0.976	0.993	0.998	0.998	0.999
3600	1.000	0.999	0.999	0.999	1.000	1.000	1.000	1.000

Additional simulation results: R^2

Empirical Distribution of the R^2 coefficient (ML)

T	mean	1%	5%	10%	50%	90%	95%	99%
				Misspecified Model (Useful Factor)				
200	0.231	0.000	0.002	0.006	0.161	0.577	0.674	0.806
600	0.178	0.000	0.002	0.006	0.130	0.429	0.514	0.651
3600	0.143	0.006	0.026	0.043	0.133	0.256	0.294	0.367
				Misspecified Model (Irrelevant Factor)				
200	0.940	0.150	0.703	0.852	0.988	1.000	1.000	1.000
600	0.996	0.961	0.985	0.991	0.999	1.000	1.000	1.000
3600	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000

Additional simulation results: t-tests

Rejection Rates of t-tests (CU-GMM)

T	$H_0: \lambda = \lambda_*$ 10%	5%	1%	$H_0: \lambda = 0$ 10%	5%	1%
	Correctly Specified (Useful Factor)					
200	0.319	0.238	0.123	0.449	0.362	0.217
600	0.153	0.089	0.025	0.533	0.423	0.230
3600	0.109	0.056	0.012	0.987	0.973	0.904
	Misspecified (Useful Factor)					
200	0.632	0.565	0.442	0.849	0.814	0.732
600	0.459	0.377	0.245	0.953	0.936	0.889
3600	0.368	0.284	0.159	1.000	1.000	1.000

Note: λ_* is the (pseudo-) true value of the parameter

Additional simulation results: t-tests

Rejection Rates of t-tests (CU-GMM)

$H_0 : \lambda = 0$

T	10%	5%	1%
Correctly Specified (Irrelevant Factor)			
200	0.850	0.818	0.749
600	0.813	0.774	0.691
3600	0.800	0.758	0.668
Misspecified (Irrelevant Factor)			
200	0.997	0.996	0.994
600	1.000	1.000	1.000
3600	1.000	1.000	1.000

Simulation evidence (CU-GMM)

Rejection rates of specification test and t-tests of statistical significance

	$H_0 : \lambda_1 = 0$			$H_0 : \lambda_2 = 0$			$\mathcal{J}$ Test		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
(1)	0.953	0.936	0.889				1.00	1.00	0.999
(3)	0.171	0.096	0.024	1.00	1.00	1.00	0.085	0.040	0.007

Notes: (1) denotes the model with a useful factor only and (3) denotes the model with one useful and one irrelevant (useless) factors.

SDF approach

- Let

SDF approach

- Let
  - m_t be an admissible SDF at time t

SDF approach

- Let
  - m_t be an admissible SDF at time t
  - R_t be the payoffs (returns) of N assets at time t

SDF approach

- Let
  - m_t be an admissible SDF at time t
  - R_t be the payoffs (returns) of N assets at time t
  - $q \neq 0_N$ be a vector of costs ($q = 1_N$ in case of gross returns)

SDF approach

- Let
  - m_t be an admissible SDF at time t
  - R_t be the payoffs (returns) of N assets at time t
  - $q \neq 0_N$ be a vector of costs ($q = 1_N$ in case of gross returns)
- SDF approach to asset pricing (fundamental pricing equation)

$$E[R_t m_t] = q$$

SDF approach

- Let
  - m_t be an admissible SDF at time t
  - R_t be the payoffs (returns) of N assets at time t
  - $q \neq 0_N$ be a vector of costs ($q = 1_N$ in case of gross returns)
- SDF approach to asset pricing (fundamental pricing equation)

$$E[R_t m_t] = q$$

  - Asset prices are obtained by "discounting" future payoffs by m_t so that the expected present value of the payoffs is equal to their costs

SDF approach

- Let
  - m_t be an admissible SDF at time t
  - R_t be the payoffs (returns) of N assets at time t
  - $q \neq 0_N$ be a vector of costs ($q = 1_N$ in case of gross returns)
- SDF approach to asset pricing (fundamental pricing equation)

$$E[R_t m_t] = q$$

  - Asset prices are obtained by "discounting" future payoffs by m_t so that the expected present value of the payoffs is equal to their costs
  - Unifying approach to pricing stocks, bonds, and derivative products

SDF approach

- Let
  - m_t be an admissible SDF at time t
  - R_t be the payoffs (returns) of N assets at time t
  - $q \neq 0_N$ be a vector of costs ($q = 1_N$ in case of gross returns)
- SDF approach to asset pricing (fundamental pricing equation)

$$E[R_t m_t] = q$$

  - Asset prices are obtained by "discounting" future payoffs by m_t so that the expected present value of the payoffs is equal to their costs
  - Unifying approach to pricing stocks, bonds, and derivative products
  - Equivalent to a no-arbitrage principle, provided that $m_t > 0$

SDF approach

- Let
  - m_t be an admissible SDF at time t
  - R_t be the payoffs (returns) of N assets at time t
  - $q \neq 0_N$ be a vector of costs ($q = 1_N$ in case of gross returns)
- SDF approach to asset pricing (fundamental pricing equation)

$$E[R_t m_t] = q$$

  - Asset prices are obtained by "discounting" future payoffs by m_t so that the expected present value of the payoffs is equal to their costs
  - Unifying approach to pricing stocks, bonds, and derivative products
  - Equivalent to a no-arbitrage principle, provided that $m_t > 0$
  - Conditioning information can also be incorporated

SDF approach

- Let
  - m_t be an admissible SDF at time t
  - R_t be the payoffs (returns) of N assets at time t
  - $q \neq 0_N$ be a vector of costs ($q = 1_N$ in case of gross returns)
- SDF approach to asset pricing (fundamental pricing equation)

$$E[R_t m_t] = q$$

  - Asset prices are obtained by "discounting" future payoffs by m_t so that the expected present value of the payoffs is equal to their costs
  - Unifying approach to pricing stocks, bonds, and derivative products
  - Equivalent to a no-arbitrage principle, provided that $m_t > 0$
  - Conditioning information can also be incorporated
- Asset pricing models parameterize a candidate SDF y as a function of the data and parameters λ

SDF approach

- Let
  - m_t be an admissible SDF at time t
  - R_t be the payoffs (returns) of N assets at time t
  - $q \neq 0_N$ be a vector of costs ($q = 1_N$ in case of gross returns)
- SDF approach to asset pricing (fundamental pricing equation)

$$E[R_t m_t] = q$$

  - Asset prices are obtained by "discounting" future payoffs by m_t so that the expected present value of the payoffs is equal to their costs
  - Unifying approach to pricing stocks, bonds, and derivative products
  - Equivalent to a no-arbitrage principle, provided that $m_t > 0$
  - Conditioning information can also be incorporated
- Asset pricing models parameterize a candidate SDF y as a function of the data and parameters λ
  - nonlinear CCAPM: $\lambda_0 \frac{u'(c_t)}{u'(c_{t-1})}$ or $\lambda_0 \left(\frac{c_t}{c_{t-1}}\right)^{-\lambda_1}$ (CRRA)

SDF approach

- Let
  - m_t be an admissible SDF at time t
  - R_t be the payoffs (returns) of N assets at time t
  - $q \neq 0_N$ be a vector of costs ($q = 1_N$ in case of gross returns)
- SDF approach to asset pricing (fundamental pricing equation)

$$E[R_t m_t] = q$$

  - Asset prices are obtained by "discounting" future payoffs by m_t so that the expected present value of the payoffs is equal to their costs
  - Unifying approach to pricing stocks, bonds, and derivative products
  - Equivalent to a no-arbitrage principle, provided that $m_t > 0$
  - Conditioning information can also be incorporated
- Asset pricing models parameterize a candidate SDF y as a function of the data and parameters λ
  - nonlinear CCAPM: $\lambda_0 \frac{u'(c_t)}{u'(c_{t-1})}$ or $\lambda_0 \left(\frac{c_t}{c_{t-1}}\right)^{-\lambda_1}$ (CRRA)
  - linear models: $\lambda_0 + \lambda_1' f_t$, where f_t are risk factors

SDF approach

- Let
$$y_t(\lambda) = \tilde{f}_t'\lambda$$
be a proposed stochastic discount factor (SDF), where f_t is a (K-1)-vector of systematic risk factors, $\tilde{f}_t = [1,\ f_t']'$ and $\lambda = [\lambda_0,\ \lambda_1']'$

SDF approach

- Let
$$y_t(\lambda) = \tilde{f}_t'\lambda$$
be a proposed stochastic discount factor (SDF), where f_t is a (K-1)-vector of systematic risk factors, $\tilde{f}_t = [1,\ f_t']'$ and $\lambda = [\lambda_0,\ \lambda_1']'$
- Also, let $e_t(\lambda) = R_t\tilde{f}_t'\lambda - q$ denote the pricing errors of the model

SDF approach

- Let
$$y_t(\lambda) = \tilde{f}_t'\lambda$$
be a proposed stochastic discount factor (SDF), where f_t is a (K-1)-vector of systematic risk factors, $\tilde{f}_t = [1,\ f_t']'$ and $\lambda = [\lambda_0,\ \lambda_1']'$
- Also, let $e_t(\lambda) = R_t\tilde{f}_t'\lambda - q$ denote the pricing errors of the model
  - when the asset pricing model holds, $E[e_t(\lambda)] = E[R_t\tilde{f}_t']\lambda - q = D\lambda - q = 0_N$

SDF approach

- Let

$$y_t(\lambda) = \tilde{f}_t'\lambda$$

be a proposed stochastic discount factor (SDF), where f_t is a (K-1)-vector of systematic risk factors, $\tilde{f}_t = [1,\ f_t']'$ and $\lambda = [\lambda_0,\ \lambda_1']'$

- Also, let $e_t(\lambda) = R_t\tilde{f}_t'\lambda - q$ denote the pricing errors of the model
  - when the asset pricing model holds, $E[e_t(\lambda)] = E[R_t\tilde{f}_t']\lambda - q = D\lambda - q = 0_N$
- The continuously-updated GMM estimator of λ is defined as

$$\hat{\lambda} = \text{argmin}_\lambda \bar{e}(\lambda)'\hat{W}_e(\lambda)^{-1}\bar{e}(\lambda),$$

$\bar{e}(\lambda)=\frac{1}{T}\sum_{t=1}^{T} e_t(\lambda)$ and $\hat{W}_e(\lambda)=\frac{1}{T}\sum_{t=1}^{T}(e_t(\lambda)\text{-}\bar{e}(\lambda))(e_t(\lambda)\text{-}\bar{e}(\lambda))'$

SDF approach

- Let
$$y_t(\lambda) = \tilde{f}_t'\lambda$$
be a proposed stochastic discount factor (SDF), where f_t is a (K-1)-vector of systematic risk factors, $\tilde{f}_t = [1,\ f_t']'$ and $\lambda = [\lambda_0,\ \lambda_1']'$
- Also, let $e_t(\lambda) = R_t\tilde{f}_t'\lambda - q$ denote the pricing errors of the model
  - when the asset pricing model holds, $E[e_t(\lambda)] = E[R_t\tilde{f}_t']\lambda - q = D\lambda - q = 0_N$
- The continuously-updated GMM estimator of λ is defined as
$$\hat{\lambda} = \text{argmin}_\lambda \bar{e}(\lambda)'\hat{W}_e(\lambda)^{-1}\bar{e}(\lambda),$$
$\bar{e}(\lambda)=\frac{1}{T}\sum_{t=1}^T e_t(\lambda)$ and $\hat{W}_e(\lambda)=\frac{1}{T}\sum_{t=1}^T(e_t(\lambda)\text{-}\bar{e}(\lambda))(e_t(\lambda)\text{-}\bar{e}(\lambda))'$
- The over-identifying restriction test of the asset pricing model is
$$\mathcal{J}(\lambda) = T\min_\lambda \bar{e}(\lambda)'\hat{W}_e(\lambda)^{-1}\bar{e}(\lambda)$$
and $\mathcal{J}(\hat{\lambda}) \xrightarrow{d} \chi^2_{N-K}$ when the asset pricing model holds.

CU-GMM: Specification test

- Let P_q be an $N \times (N-1)$ orthonormal matrix whose columns are orthogonal to q such that $P_q P_q' = I_N - q(q'q)^{-1}q'$ and $P_q' q = 0_{N-1}$

CU-GMM: Specification test

- Let P_q be an $N \times (N-1)$ orthonormal matrix whose columns are orthogonal to q such that $P_q P_q' = I_N - q(q'q)^{-1}q'$ and $P_q' q = 0_{N-1}$
- Since $D\lambda = q$ holds if and only if $P_q' D\lambda = 0_{N-1}$, one can use the test

$$\mathcal{CD}(\lambda) = T \min_\lambda (P_q'\hat{D}\lambda)'[(\lambda' \otimes P_q')\hat{V}_d(\lambda \otimes P_q)]^{-1}(P_q'\hat{D}\lambda),$$

where $\hat{V}_d$ is the variance matrix of $\hat{D}$, to test if $\mathrm{rank}(P_q' D) = K - 1$

CU-GMM: Specification test

- Let P_q be an $N \times (N-1)$ orthonormal matrix whose columns are orthogonal to q such that $P_q P_q' = I_N - q(q'q)^{-1}q'$ and $P_q' q = 0_{N-1}$
- Since $D\lambda = q$ holds if and only if $P_q' D\lambda = 0_{N-1}$, one can use the test

  $$\mathcal{CD}(\lambda) = T \min_\lambda (P_q' \hat{D}\lambda)' [(\lambda' \otimes P_q') \hat{V}_d (\lambda \otimes P_q)]^{-1} (P_q' \hat{D}\lambda),$$

  where $\hat{V}_d$ is the variance matrix of $\hat{D}$, to test if $\mathrm{rank}(P_q' D) = K - 1$
- THEOREM. *Let $\mathcal{J}(\hat{\lambda})$ and $\mathcal{CD}(\tilde{\lambda})$ be the tests of $H_0 : D\lambda = q$ and $H_0 : \mathrm{rank}(P_q' D) = K - 1$, where $\hat{\lambda}$ and $\tilde{\lambda}$ denote their corresponding minimizers. Then, $\hat{\lambda} = c^* \tilde{\lambda}$ and $\mathcal{J}(\hat{\lambda}) = \mathcal{CD}(\tilde{\lambda})$*

CU-GMM: Specification test

- Let P_q be an $N \times (N-1)$ orthonormal matrix whose columns are orthogonal to q such that $P_q P_q' = I_N - q(q'q)^{-1}q'$ and $P_q' q = 0_{N-1}$
- Since $D\lambda = q$ holds if and only if $P_q' D\lambda = 0_{N-1}$, one can use the test

$$\mathcal{CD}(\lambda) = T \min_\lambda (P_q'\hat{D}\lambda)'[(\lambda' \otimes P_q')\hat{V}_d(\lambda \otimes P_q)]^{-1}(P_q'\hat{D}\lambda),$$

where $\hat{V}_d$ is the variance matrix of $\hat{D}$, to test if $\mathrm{rank}(P_q' D) = K - 1$
- THEOREM. *Let $\mathcal{J}(\hat{\lambda})$ and $\mathcal{CD}(\tilde{\lambda})$ be the tests of $H_0 : D\lambda = q$ and $H_0 : \mathrm{rank}(P_q' D) = K - 1$, where $\hat{\lambda}$ and $\tilde{\lambda}$ denote their corresponding minimizers. Then, $\hat{\lambda} = c^*\tilde{\lambda}$* and $\mathcal{J}(\hat{\lambda}) = \mathcal{CD}(\tilde{\lambda})$
- COROLLARY. *Suppose that the model contains an irrelevant factor so that $\mathrm{rank}(D) = K - 1$ and $D\lambda - q \neq 0_N$ (i.e., the model is misspecified). Let α denote the significance level of the test and c_α be the $100(1-\alpha)$-th percentile of χ^2_{N-K}. Then,*

$$\lim_{T\to\infty} \Pr\{\mathcal{J}(\hat{\lambda}) > c_\alpha\} = \alpha$$

CU-GMM: Specification test

- Let P_q be an $N \times (N-1)$ orthonormal matrix whose columns are orthogonal to q such that $P_q P_q' = I_N - q(q'q)^{-1}q'$ and $P_q' q = 0_{N-1}$
- Since $D\lambda = q$ holds if and only if $P_q' D\lambda = 0_{N-1}$, one can use the test

$$\mathcal{CD}(\lambda) = T \min_\lambda (P_q' \hat{D}\lambda)'[(\lambda' \otimes P_q')\hat{V}_d(\lambda \otimes P_q)]^{-1}(P_q'\hat{D}\lambda),$$

where $\hat{V}_d$ is the variance matrix of $\hat{D}$, to test if $\mathrm{rank}(P_q' D) = K - 1$
- Theorem. *Let $\mathcal{J}(\hat{\lambda})$ and $\mathcal{CD}(\tilde{\lambda})$ be the tests of $H_0 : D\lambda = q$ and $H_0 : \mathrm{rank}(P_q' D) = K - 1$, where $\hat{\lambda}$ and $\tilde{\lambda}$ denote their corresponding minimizers. Then, $\hat{\lambda} = c^* \tilde{\lambda}$ and $\mathcal{J}(\hat{\lambda}) = \mathcal{CD}(\tilde{\lambda})$*
- Corollary. *Suppose that the model contains an irrelevant factor so that $\mathrm{rank}(D) = K - 1$ and $D\lambda - q \neq 0_N$ (i.e., the model is misspecified). Let α denote the significance level of the test and c_α be the $100(1-\alpha)$-th percentile of χ^2_{N-K}. Then,*

$$\lim_{T\to\infty} \Pr\{\mathcal{J}(\hat{\lambda}) > c_\alpha\} = \alpha$$

- Similar results for LR and canonical correlation rank tests

Additional results: Misspecification-robust standard errors

- The CU-GMM estimator can be defined equivalently (see Newey and Smith, 2004) as the solution to the following saddle point problem:

$$\hat{\lambda} = \arg\min_{\lambda \in \Lambda} \max_{\eta \in \Psi(\lambda)} \frac{1}{T} \sum_{t=1}^{T} \rho\left(\eta' e_t(\lambda)\right),$$

where $\rho(v) = -\frac{1}{2}v^2 - v$ and η is an $N \times 1$ vector of Lagrange multipliers associated with the moment conditions $E\left[e_t(\lambda)\right] = 0_N$

Additional results: Misspecification-robust standard errors

- The CU-GMM estimator can be defined equivalently (see Newey and Smith, 2004) as the solution to the following saddle point problem:

$$\hat{\lambda} = \arg\min_{\lambda\in\Lambda}\max_{\eta\in\Psi(\lambda)} \frac{1}{T}\sum_{t=1}^{T}\rho\left(\eta' e_t(\lambda)\right),$$

  where $\rho(v) = -\frac{1}{2}v^2 - v$ and η is an $N\times 1$ vector of Lagrange multipliers associated with the moment conditions $E\left[e_t(\lambda)\right] = 0_N$
- Note that η characterizes the degree of model misspecification with $\eta_*(\lambda) = 0_N$ for correctly specified models and $\|\eta_*(\lambda)\| > 0$ for misspecified models

Additional results: Misspecification-robust standard errors

- The CU-GMM estimator can be defined equivalently (see Newey and Smith, 2004) as the solution to the following saddle point problem:

$$\hat{\lambda} = \arg\min_{\lambda \in \Lambda} \max_{\eta \in \Psi(\lambda)} \frac{1}{T} \sum_{t=1}^{T} \rho\left(\eta' e_t(\lambda)\right),$$

  where $\rho(v) = -\frac{1}{2}v^2 - v$ and η is an $N \times 1$ vector of Lagrange multipliers associated with the moment conditions $E\left[e_t(\lambda)\right] = 0_N$
- Note that η characterizes the degree of model misspecification with $\eta_*(\lambda) = 0_N$ for correctly specified models and $\|\eta_*(\lambda)\| > 0$ for misspecified models
- Then, under some regularity conditions,

$$\sqrt{T}(\hat{\lambda} - \lambda_*) \xrightarrow{d} N(0_K, E[l_t l_t']),$$

  where $l_t = (C - B'W(\lambda_*)^{-1}B)^{-1} c_t \left[D_t'\eta_* - B'W(\lambda_*)^{-1} e_t(\lambda_*)\right]$, $D_t = R_t \tilde{f}_t'$, $c_t = 1 + \eta_*' e_t(\lambda_*)$, $C = E[D_t'\eta_*\eta_*' D_t]$ and $B = E[c_t R_t \tilde{f}_t'] + E[e_t(\lambda_*)\eta_*' D_t]$

Additional results: Misspecification-robust standard errors

Empirical size and power of t-tests: misspecified model (useful factor)

T	Size: $H_0 : \lambda = \lambda_*$ 10%	5%	1%	Power: $H_0 : \lambda = 0$ 10%	5%	1%
			Panel A: t_c			
200	0.632	0.565	0.442	0.849	0.814	0.732
600	0.459	0.377	0.245	0.953	0.936	0.889
3600	0.368	0.284	0.159	1.000	1.000	1.000
			Panel B: t_m			
200	0.158	0.088	0.021	0.399	0.281	0.113
600	0.103	0.050	0.010	0.741	0.628	0.377
3600	0.099	0.049	0.009	1.000	1.000	0.999

Additional results: Misspecification-robust standard errors

Empirical size and power of t-tests: misspecified model (irrelevant factor)

	Size: $H_0 : \lambda = 0$		
T	10%	5%	1%
	Panel A: t_c		
200	0.997	0.996	0.994
600	1.000	1.000	1.000
3600	1.000	1.000	1.000
	Panel B: t_m		
200	0.135	0.070	0.014
600	0.082	0.038	0.007
3600	0.095	0.046	0.009

Beta pricing model

- It is often desirable to estimate and evaluate the asset pricing model in the beta pricing setup

Beta pricing model

- It is often desirable to estimate and evaluate the asset pricing model in the beta pricing setup
- Let $\mu = E[R_t]$, $\mu_f = E[f_t]$ and $V_f = \mathrm{Var}[f_t]$, and assume

$$R_t - \mu = \beta(f_t - \mu_f) + \epsilon_t, \quad t = 1, \ldots, T,$$

where $\beta = [\beta_1, ..., \beta_{K-1}]$ is an $N \times (K-1)$ matrix and $\epsilon_t \sim iid(0_N, \Sigma)$

Beta pricing model

- It is often desirable to estimate and evaluate the asset pricing model in the beta pricing setup
- Let $\mu = E[R_t]$, $\mu_f = E[f_t]$ and $V_f = \mathrm{Var}[f_t]$, and assume

$$R_t - \mu = \beta(f_t - \mu_f) + \epsilon_t, \quad t = 1, \ldots, T,$$

  where $\beta = [\beta_1, ..., \beta_{K-1}]$ is an $N \times (K-1)$ matrix and $\epsilon_t \sim iid(0_N, \Sigma)$
- The beta pricing model suggests that

$$\mu = 1_N \gamma_0 + \beta \gamma_1$$

Beta pricing model

- It is often desirable to estimate and evaluate the asset pricing model in the beta pricing setup
- Let $\mu = E[R_t]$, $\mu_f = E[f_t]$ and $V_f = \text{Var}[f_t]$, and assume

$$R_t - \mu = \beta(f_t - \mu_f) + \epsilon_t, \quad t = 1, \ldots, T,$$

  where $\beta = [\beta_1, \ldots, \beta_{K-1}]$ is an $N \times (K-1)$ matrix and $\epsilon_t \sim iid(0_N, \Sigma)$
- The beta pricing model suggests that

$$\mu = 1_N \gamma_0 + \beta \gamma_1$$

- Two main reasons why the beta pricing setup is often preferred in empirical asset pricing

Beta pricing model

- It is often desirable to estimate and evaluate the asset pricing model in the beta pricing setup
- Let $\mu = E[R_t]$, $\mu_f = E[f_t]$ and $V_f = \text{Var}[f_t]$, and assume

$$R_t - \mu = \beta(f_t - \mu_f) + \epsilon_t, \quad t = 1, \ldots, T,$$

  where $\beta = [\beta_1, ..., \beta_{K-1}]$ is an $N \times (K-1)$ matrix and $\epsilon_t \sim iid(0_N, \Sigma)$
- The beta pricing model suggests that

$$\mu = 1_N \gamma_0 + \beta \gamma_1$$

- Two main reasons why the beta pricing setup is often preferred in empirical asset pricing
  1. the parameters γ_1 have a direct interpretation of risk premium parameters

Beta pricing model

- It is often desirable to estimate and evaluate the asset pricing model in the beta pricing setup
- Let $\mu = E[R_t]$, $\mu_f = E[f_t]$ and $V_f = \text{Var}[f_t]$, and assume

$$R_t - \mu = \beta(f_t - \mu_f) + \epsilon_t, \quad t = 1, \ldots, T,$$

  where $\beta = [\beta_1, ..., \beta_{K-1}]$ is an $N \times (K-1)$ matrix and $\epsilon_t \sim iid(0_N, \Sigma)$
- The beta pricing model suggests that

$$\mu = 1_N \gamma_0 + \beta \gamma_1$$

- Two main reasons why the beta pricing setup is often preferred in empirical asset pricing
  1. the parameters γ_1 have a direct interpretation of risk premium parameters
  2. the beta representation allows for conveniently measuring and plotting the goodness-of-fit as model's expected returns versus realized returns

Beta pricing model (CU-GMM)

- The mapping between the SDF and beta pricing model parameters is given by

$$\gamma_0 = \frac{1}{\lambda_0 + \mu_f'\lambda_1}, \gamma_1 = \frac{-V_f\lambda_1}{\lambda_0 + \mu_f'\lambda_1},$$

  where V_f is the covariance matrix of f_t.

Beta pricing model (CU-GMM)

- The mapping between the SDF and beta pricing model parameters is given by
$$\gamma_0 = \frac{1}{\lambda_0 + \mu_f' \lambda_1}, \gamma_1 = \frac{-V_f \lambda_1}{\lambda_0 + \mu_f' \lambda_1},$$
where V_f is the covariance matrix of f_t.
- By augmenting $\bar{e}(\lambda)$ in the SDF representation with additional (just-identified) moment conditions for μ_f, V_f and β, the CU-GMM estimate of the augmented parameter vector $\theta = [\lambda_0, \lambda_1', \beta_1', \ldots, \beta_{K-1}', \mu_f', \mathrm{vech}(V_f)]'$ becomes numerically identical to the CU-GMM estimate of $[\gamma_0, \gamma_1', \beta_1', \ldots, \beta_{K-1}', \mu_f', \mathrm{vech}(V_f)]'$ in the beta pricing model

Beta pricing model (CU-GMM)

- The mapping between the SDF and beta pricing model parameters is given by
$$\gamma_0 = \frac{1}{\lambda_0 + \mu_f' \lambda_1}, \gamma_1 = \frac{-V_f \lambda_1}{\lambda_0 + \mu_f' \lambda_1},$$
where V_f is the covariance matrix of f_t.
- By augmenting $\bar{e}(\lambda)$ in the SDF representation with additional (just-identified) moment conditions for μ_f, V_f and β, the CU-GMM estimate of the augmented parameter vector $\theta = [\lambda_0,\, \lambda_1',\, \beta_1',\, \ldots,\, \beta_{K-1}',\, \mu_f',\, \mathrm{vech}(V_f)]'$ becomes numerically identical to the CU-GMM estimate of $[\gamma_0,\, \gamma_1',\, \beta_1',\, \ldots,\, \beta_{K-1}',\, \mu_f',\, \mathrm{vech}(V_f)]'$ in the beta pricing model
- However, the estimation of θ can be performed in a sequential manner which offers substantial computational advantages

Beta pricing model (CU-GMM)

- The mapping between the SDF and beta pricing model parameters is given by
$$\gamma_0 = \frac{1}{\lambda_0 + \mu_f' \lambda_1}, \gamma_1 = \frac{-V_f \lambda_1}{\lambda_0 + \mu_f' \lambda_1},$$
where V_f is the covariance matrix of f_t.
- By augmenting $\bar{e}(\lambda)$ in the SDF representation with additional (just-identified) moment conditions for μ_f, V_f and β, the CU-GMM estimate of the augmented parameter vector $\theta = [\lambda_0, \lambda_1', \beta_1', \ldots, \beta_{K-1}', \mu_f', \mathrm{vech}(V_f)]'$ becomes numerically identical to the CU-GMM estimate of $[\gamma_0, \gamma_1', \beta_1', \ldots, \beta_{K-1}', \mu_f', \mathrm{vech}(V_f)]'$ in the beta pricing model
- However, the estimation of θ can be performed in a sequential manner which offers substantial computational advantages
- The following theorem presents a general result for this sequential estimation

Beta pricing model (CU-GMM)

- THEOREM. *Let* $\theta = (\theta_1', \theta_2')'$, *where* θ_1 *is* $K_1 \times 1$ *and* θ_2 *is* $K_2 \times 1$, *and*
$$E[g_t(\theta)] = \begin{bmatrix} E[g_{1t}(\theta_1)] \\ E[g_{2t}(\theta_1, \theta_2)] \end{bmatrix} = \begin{bmatrix} 0_{N_1} \\ 0_{N_2} \end{bmatrix},$$
where $g_{1t}(\theta_1)$ *is* $N_1 \times 1$ *and* $g_{2t}(\theta)$ *is* $N_2 \times 1$, *with* $N_1 > K_1$ *and* $N_2 = K_2$. *Define the estimators*
$$\tilde{\theta}_1 = \text{argmin}_{\theta_1} \bar{g}_1(\theta_1)' \hat{W}_{11}(\theta_1)^{-1} \bar{g}_1(\theta_1),$$
$$\hat{\theta} \equiv \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix} = \text{argmin}_{\theta} \bar{g}(\theta)' \hat{W}(\theta)^{-1} \bar{g}(\theta).$$
Then, $\tilde{\theta}_1 = \hat{\theta}_1$, *and* $\mathcal{J}(\tilde{\theta}_1) = \mathcal{J}(\hat{\theta})$.

Beta pricing model (CU-GMM)

- THEOREM. *Let* $\theta = (\theta_1', \ \theta_2')'$, *where* θ_1 *is* $K_1 \times 1$ *and* θ_2 *is* $K_2 \times 1$, *and*
$$E[g_t(\theta)] = \begin{bmatrix} E[g_{1t}(\theta_1)] \\ E[g_{2t}(\theta_1, \theta_2)] \end{bmatrix} = \begin{bmatrix} 0_{N_1} \\ 0_{N_2} \end{bmatrix},$$
where $g_{1t}(\theta_1)$ *is* $N_1 \times 1$ *and* $g_{2t}(\theta)$ *is* $N_2 \times 1$, *with* $N_1 > K_1$ *and* $N_2 = K_2$. *Define the estimators*
$$\tilde{\theta}_1 = \mathrm{argmin}_{\theta_1} \bar{g}_1(\theta_1)' \hat{W}_{11}(\theta_1)^{-1} \bar{g}_1(\theta_1),$$
$$\hat{\theta} \equiv \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix} = \mathrm{argmin}_{\theta} \bar{g}(\theta)' \hat{W}(\theta)^{-1} \bar{g}(\theta).$$
Then, $\tilde{\theta}_1 = \hat{\theta}_1$, *and* $\mathcal{J}(\tilde{\theta}_1) = \mathcal{J}(\hat{\theta})$.
- The theorem establishes that for the CU-GMM, adding a new set of just-identified moment conditions to the original system does not alter the estimates of the original parameters as well as the test for overidentifying restrictions

Beta pricing model (CU-GMM)

- This implies that we can discard the subset of moment conditions that are exactly identified and only perform the estimation and the over-identifying restriction test on the remaining smaller set of moment conditions.

Beta pricing model (CU-GMM)

- This implies that we can discard the subset of moment conditions that are exactly identified and only perform the estimation and the over-identifying restriction test on the remaining smaller set of moment conditions.
- Then, we can solve for $\hat{\theta}_2 = [\hat{\beta}_1', \ldots, \hat{\beta}_{K-1}', \hat{\mu}_f', \mathrm{vech}(\hat{V}_f)]'$ after $\tilde{\theta}_1 = [\lambda_0, \lambda_1']'$ is obtained from the smaller system. In our linear setup, $\hat{\theta}_2$ has a closed-form solution

Beta pricing model (CU-GMM)

- This implies that we can discard the subset of moment conditions that are exactly identified and only perform the estimation and the over-identifying restriction test on the remaining smaller set of moment conditions.
- Then, we can solve for $\hat{\theta}_2 = [\hat{\beta}_1', \ldots, \hat{\beta}_{K-1}', \hat{\mu}_f', \mathrm{vech}(\hat{V}_f)]'$ after $\tilde{\theta}_1 = [\lambda_0, \lambda_1']'$ is obtained from the smaller system. In our linear setup, $\hat{\theta}_2$ has a closed-form solution
- The estimates of γ are obtained as

$$\hat{\gamma}_0 = \frac{1}{\hat{\lambda}_0 + \hat{\mu}_f'\hat{\lambda}_1}, \hat{\gamma}_1 = \frac{-\hat{V}_f\hat{\lambda}_1}{\hat{\lambda}_0 + \hat{\mu}_f'\hat{\lambda}_1},$$

and the expected returns as

$$\tilde{\mu} = 1_N\hat{\gamma}_0 + \hat{\beta}\hat{\gamma}_1$$

Beta pricing model (CU-GMM)

- This implies that we can discard the subset of moment conditions that are exactly identified and only perform the estimation and the over-identifying restriction test on the remaining smaller set of moment conditions.
- Then, we can solve for $\hat{\theta}_2 = [\hat{\beta}_1', \ldots, \hat{\beta}_{K-1}', \hat{\mu}_f', \mathrm{vech}(\hat{V}_f)]'$ after $\tilde{\theta}_1 = [\lambda_0, \lambda_1']'$ is obtained from the smaller system. In our linear setup, $\hat{\theta}_2$ has a closed-form solution
- The estimates of γ are obtained as

$$\hat{\gamma}_0 = \frac{1}{\hat{\lambda}_0 + \hat{\mu}_f'\hat{\lambda}_1}, \hat{\gamma}_1 = \frac{-\hat{V}_f\hat{\lambda}_1}{\hat{\lambda}_0 + \hat{\mu}_f'\hat{\lambda}_1},$$

  and the expected returns as

$$\tilde{\mu} = 1_N\hat{\gamma}_0 + \hat{\beta}\hat{\gamma}_1$$

- Finally we show that $R^2 = [\mathrm{Corr}(\tilde{\mu}, \hat{\mu})]^2$ converges, as $T \to \infty$, to 1.

Beta pricing model (CU-GMM)

- This implies that we can discard the subset of moment conditions that are exactly identified and only perform the estimation and the over-identifying restriction test on the remaining smaller set of moment conditions.
- Then, we can solve for $\hat{\theta}_2 = [\hat{\beta}_1', \ldots, \hat{\beta}_{K-1}', \hat{\mu}_f', \mathrm{vech}(\hat{V}_f)]'$ after $\tilde{\theta}_1 = [\lambda_0, \lambda_1']'$ is obtained from the smaller system. In our linear setup, $\hat{\theta}_2$ has a closed-form solution
- The estimates of γ are obtained as

$$\hat{\gamma}_0 = \frac{1}{\hat{\lambda}_0 + \hat{\mu}_f'\hat{\lambda}_1}, \hat{\gamma}_1 = \frac{-\hat{V}_f\hat{\lambda}_1}{\hat{\lambda}_0 + \hat{\mu}_f'\hat{\lambda}_1},$$

  and the expected returns as

$$\tilde{\mu} = 1_N\hat{\gamma}_0 + \hat{\beta}\hat{\gamma}_1$$

- Finally we show that $R^2 = [\mathrm{Corr}(\tilde{\mu}, \hat{\mu})]^2$ converges, as $T \to \infty$, to 1.
- The results for the ML estimator are similar

Concluding remarks

- We show that the results from many popular empirical asset pricing models may be spurious

Concluding remarks

- We show that the results from many popular empirical asset pricing models may be spurious
- The spurious results in these models (almost perfect fit and strong evidence of non-zero risk premium) arise from the combined effect of identification failure and model misspecification

Concluding remarks

- We show that the results from many popular empirical asset pricing models may be spurious
- The spurious results in these models (almost perfect fit and strong evidence of non-zero risk premium) arise from the combined effect of identification failure and model misspecification
- It is important to stress that this is not an isolated problem limited to a particular sample (data frequency), test assets and asset pricing models which suggests that the statistical evidence on the pricing ability of many macro factors and their usefulness in explaining the cross-section of asset returns should be interpreted with caution

Concluding remarks

- We show that the results from many popular empirical asset pricing models may be spurious
- The spurious results in these models (almost perfect fit and strong evidence of non-zero risk premium) arise from the combined effect of identification failure and model misspecification
- It is important to stress that this is not an isolated problem limited to a particular sample (data frequency), test assets and asset pricing models which suggests that the statistical evidence on the pricing ability of many macro factors and their usefulness in explaining the cross-section of asset returns should be interpreted with caution
- Some warning signs about this problem (for example, the outcome of a rank test) are often ignored by applied researchers

Concluding remarks

- We show that the results from many popular empirical asset pricing models may be spurious
- The spurious results in these models (almost perfect fit and strong evidence of non-zero risk premium) arise from the combined effect of identification failure and model misspecification
- It is important to stress that this is not an isolated problem limited to a particular sample (data frequency), test assets and asset pricing models which suggests that the statistical evidence on the pricing ability of many macro factors and their usefulness in explaining the cross-section of asset returns should be interpreted with caution
- Some warning signs about this problem (for example, the outcome of a rank test) are often ignored by applied researchers
- While non-invariant estimators (HJ-distance non-optimal GMM, OLS/GLS two-pass regression) also suffer from similar problems, the invariant estimators (CU-GMM, ML) turn out to be much more sensitive to model misspecification and lack of identification