Realized semibetas: Disentangling "good" and "bad" downside risks

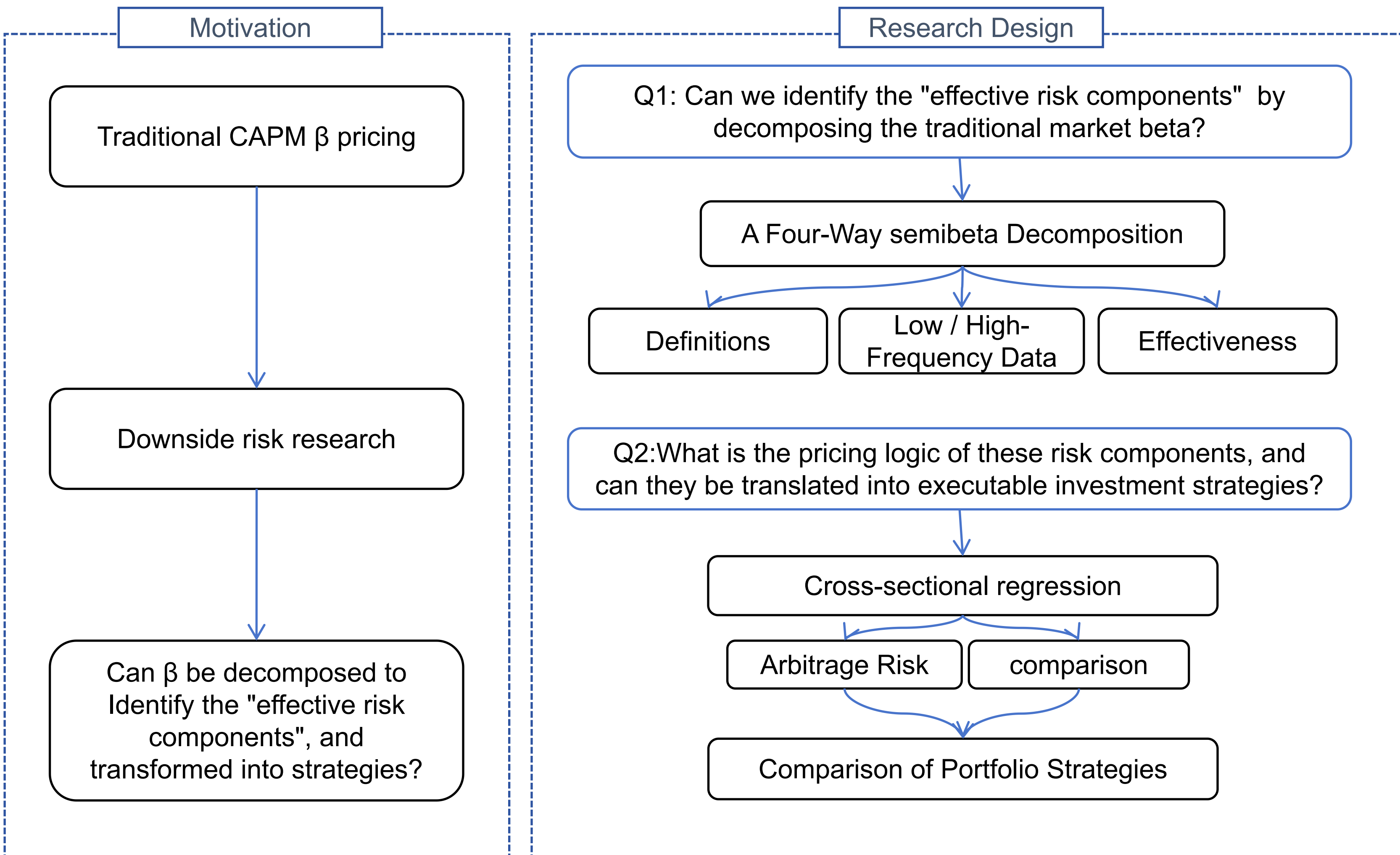
Journal of Financial Economics, 2022

Bollersleva, et al.

Presented by Xinye Cheng

November 9, 2025

Overview


Motivation
Traditional CAPM β pricing
Downside risk research
Can β be decomposed to Identify the "effective risk components", and transformed into strategies?
Research Design
Q1: Can we identify the "effective risk components" by decomposing the traditional market beta?
A Four-Way semibeta Decomposition
Definitions
Low / High-Frequency Data
Effectiveness
Q2:What is the pricing logic of these risk components, and can they be translated into executable investment strategies?
Cross-sectional regression
Arbitrage Risk
comparison
Comparison of Portfolio Strategies

Motivation

- Limitation of the Traditional CAPM’s Market Beta
    - Estimated risk premiums are inconsistent with expectations, often being too low, insignificant, or even negative.Cannot adequately explain the cross-sectional differences in returns across various assets.
    - Based on the mean-variance framework, it assumes investors are equally averse to volatility from gains and losses.Ignores the core characteristic of investors’ "loss aversion".
- Shortcomings of Existing Downside Beta Models
    - Downside beta’s ability to explain cross-sectional returns(Atilgan et al.,2018,JPM). Downside beta has no superior predictive power for cross-sectional returns compared to traditional beta (Levi and Welch,2020,RFS).

Semicovariances and Semibetas

- Obtain an exact decomposition of beta into semibetas:

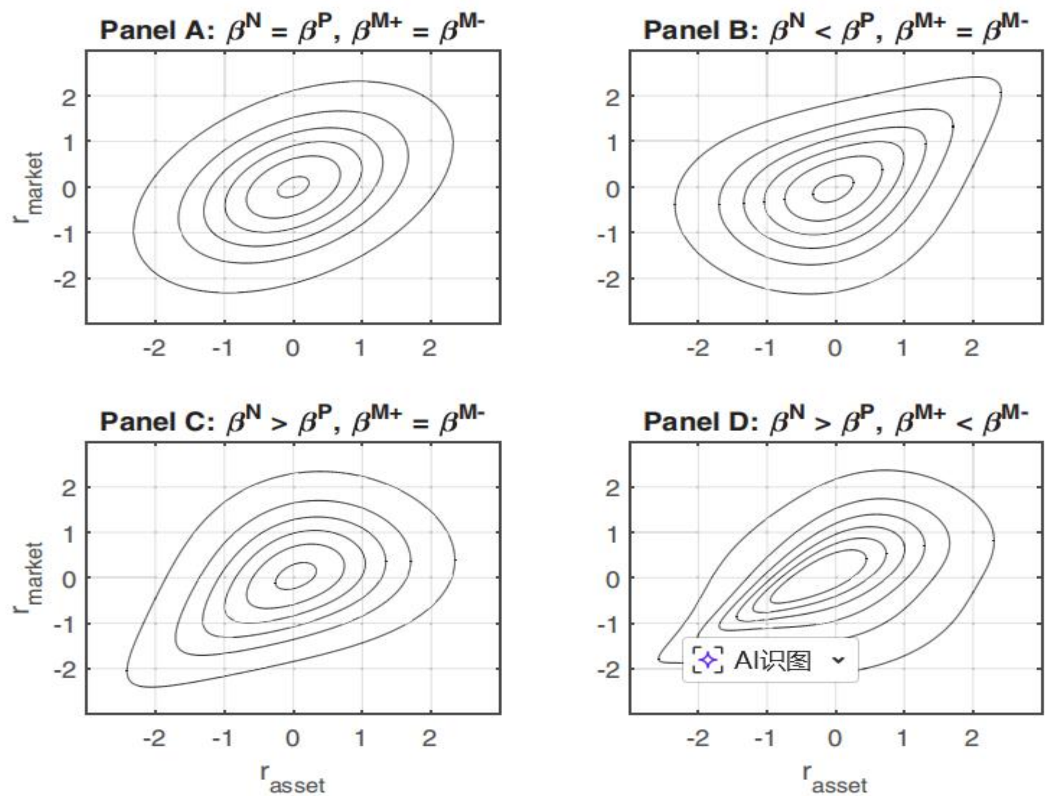
Letting r and f denote the returns on some risky asset and the aggregate market portfolio,respectively.

$$\beta = \frac{Cov(r,f)}{Var(f)} = \frac{N + P + M^{+} + M^{-}}{Var(f)} = \beta^{N} + \beta^{P} - \beta^{M^{+}} - \beta^{M^{-}}$$

- As M<0, switch the sign when defining β^{M} to ease interpretation.

Semicovariances and Semibetas

- Which asset would you preter? (CAPM beta equal to 1)



- The annual expected excess return:

asset A 9.45%;asset B 7.09%;asset C 11.91%;asset D 10.86%.

Question

- Q1: Can we identify the "effective risk components" that truly drive stock returns by decomposing the traditional market beta?

- Q2: What is the pricing logic of these risk components, and can they be translated into executable investment strategies?

Contribution

- Proposing a Four-Way semibeta Decomposition to Fill the Gap in Risk Decomposition
  - Uncovers asymmetric and differentiated pricing of non-linear dependencies between individual stocks and the market. Advancing the understanding of asset pricing beyond traditional beta and existing two-way downside beta frameworks.
- Revealing the Pricing Law of Semibetas
  - Identifies arbitrage risk as the core driver of pricing differences, leading to an imbalance in risk premiums between long and short positions.
- Simple and Operable, Empowering Investment Strategies
  - Verify that the strategy of betting on and against the "right" semibetas outperforms traditional beta and downside beta strategies.

Hypothesis: Q1

- Can we identify the "effective risk components" that truly drive stock returns by decomposing the traditional market beta?

  - H1: Traditional market beta can be decomposed into four semibetas ($\beta^{N}, \beta^{P}, \beta^{M^{+}}, \beta^{M^{-}}$) based on joint signs of asset-market returns, and only semibetas related to downside risk ($\beta^{N}, \beta^{M^{-}}$) are "effective risk components" that impact returns.

  - H2: Semibetas unrelated to downside risk ($\beta^{P}, \beta^{M^{+}}$) have no significant impact on cross-sectional returns.

Hypothesis: Q2

- What is the pricing logic of these risk components, and can they be translated into executable investment strategies?

  - H3: The pricing difference between β^N and $-\beta^{M^-}$ is driven by arbitrage risk (market frictions like short-sale constraints), with more significant differences in high-arbitrage-risk stocks.

  - H4: An investment strategy targeting effective semibetas (long β^N + short β^{M^-}) will generate superior risk-adjusted returns compared to traditional beta strategies.

Definitions

- $r_{t,k,i}$: the "high-frequency" return on asset i over the k^{th} time interval within some fixed time period t.
- $f_{t,k}$: the concurrent "high-frequency" return for the aggregate market.
- m: the number of higher-frequency return intervals within each time period.

$$\widehat{\beta}_{t,i}^{\mathcal{N}} \equiv \frac{\sum_{k=1}^{m} r_{t,k,i}^{-} f_{t,k}^{-}}{\sum_{k=1}^{m} f_{t,k}^{2}}, \quad \widehat{\beta}_{t,i}^{\mathcal{P}} \equiv \frac{\sum_{k=1}^{m} r_{t,k,i}^{+} f_{t,k}^{+}}{\sum_{k=1}^{m} f_{t,k}^{2}}$$

$$\widehat{\beta}_{t,i}^{\mathcal{M}^-} \equiv \frac{-\sum_{k=1}^{m} r_{t,k,i}^{+} f_{t,k}^{-}}{\sum_{k=1}^{m} f_{t,k}^{2}}, \quad \widehat{\beta}_{t,i}^{\mathcal{M}^+} \equiv \frac{-\sum_{k=1}^{m} r_{t,k,i}^{-} f_{t,k}^{+}}{\sum_{k=1}^{m} f_{t,k}^{2}} \qquad (2)$$

Definitions

- The realized semibetas consistently estimate the true semibetas (Bollerslev et al., 2020, Econometrica)

$$\widehat{\beta}_{t,i} \xrightarrow{p} \frac{\mathcal{COV}_{t,i}}{\mathcal{RV}_t}.$$

$$\widehat{\beta}^{\mathcal{N}}_{t,i} \xrightarrow{p} \frac{\mathcal{N}_{t,i}}{\mathcal{RV}_t}, \widehat{\beta}^{\mathcal{P}}_{t,i} \xrightarrow{p} \frac{\mathcal{P}_{t,i}}{\mathcal{RV}_t}, \widehat{\beta}^{\mathcal{M}^+}_{t,i} \xrightarrow{p} \frac{-\mathcal{M}^+_{t,i}}{\mathcal{RV}_t},$$
$$\widehat{\beta}^{\mathcal{M}^-}_{t,i} \xrightarrow{p} \frac{-\mathcal{M}^-_{t,i}}{\mathcal{RV}_t}.$$

RV_t and $COV_{t,I}$: the latent true period t market return variation and covariation between the market return and the return on the individual asset i.

$N_{t,i}, P_{t,i}, M^+_{t,i}, M^-_{t,i}$: true semicovariation measures .

Data

- Low-Frequency Data
  - Source: CRSP ; Jan 1963 – Dec 2019
  - Daily stock returns → monthly realized semibetas
  - Stocks with CRSP codes 10/11; exclude penny stocks (price < $5); 273,823 firm-month observations

- High-Frequency Data
  - Source: TAQ ; Jan 1993 – Dec 2019
  - 15-minute intraday returns → daily realized semibetas
  - All S&P 500 constituent stocks; 6,799 trading days, 1,182 unique stocks

Data

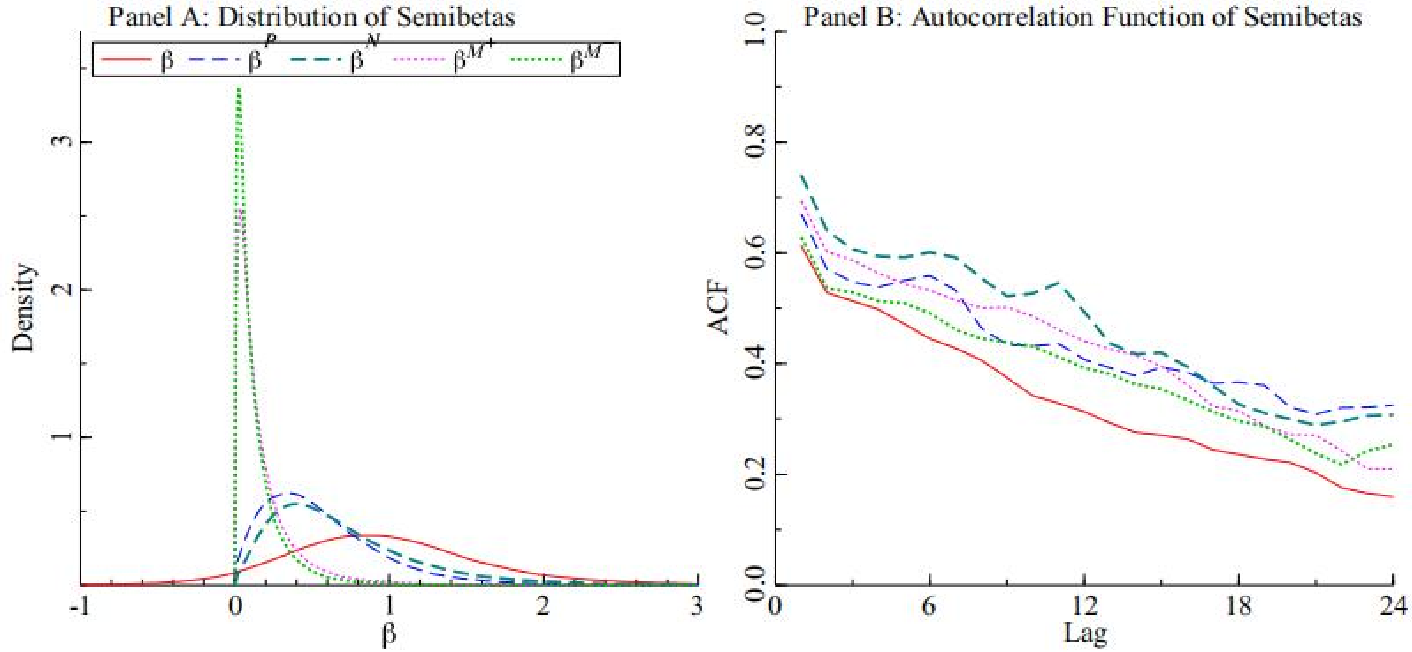


Fig. 2. Unconditional distributions and autocorrelations. Panel A displays kernel density estimates of the unconditional distribution of the monthly realized beta and semibetas averaged across time and stocks. Panel B reports the average autocorrelation functions for the monthly realized beta and semibetas averaged across stocks. The estimates are based on all of the common, non-penny, stocks in the CRSP database from January 1963 to December 2019.

Standard risk factors and controls

- Fama-MacBeth regressions:
  - month t = 1, ..., T − 1, stocks i = 1, ..., N_t.
  - we estimate the month t + 1 risk premiums (λs) for the different semibetas from the cross-sectional regression.

$$r_{t+1,i} = \lambda_{0,t+1} + \lambda^{\mathcal{N}}_{t+1}\left(\hat{\beta}^{\mathcal{N}}_{t,i} - \hat{\beta}^{\mathcal{M}^-}_{t,i}\right) + \lambda^{\mathcal{P}}_{t+1}\left(\hat{\beta}^{\mathcal{P}}_{t,i} - \hat{\beta}^{\mathcal{M}^+}_{t,i}\right) + \delta^{\mathcal{M}^+}_{t+1}\hat{\beta}^{\mathcal{M}^+}_{t,i} + \delta^{\mathcal{M}^-}_{t+1}\hat{\beta}^{\mathcal{M}^-}_{t,i} + \epsilon_{t+1,i}$$

  - Based on these T−1 cross-sectional estimates, the average risk premiums(lambdas):

$$\hat{\lambda}^j = \frac{1}{T-1}\sum_{t=2}^{T} \hat{\lambda}^j_t \; j = \mathcal{N}, \mathcal{P}, \mathcal{M}^+, \mathcal{M}^-$$

- Semibetas +FFC4 model(include ME, BM and MOM)
- Further incorporates REV, RV, IVOL and ILLIQ as additional controls.

Standard risk factors and controls

- Monthly Fama-MacBeth regressions

Table 2

Monthly Fama-MacBeth regressions. The table reports the estimated annualized risk premia and Newey-West robust t-statistics from overlapping monthly Fama-MacBeth cross-sectional predictive regressions. The monthly semibetas are calculated from daily data. All of the control variables are measured on the day prior to the monthly returns. The estimates are based on all of the common, non-penny, stocks in the CRSP database from January 1963 to December 2019.

β	β^{N}	β^{P}	β^{M+}	β^{M-}	ME	BM	MOM	REV	RV	IVOL	ILLIQ	R^2
4.27												2.33
3.96												
	10.54	1.84	4.59	-6.00								5.16
	4.51	1.17	1.32	-1.97								
	8.74	0.25	5.72	-13.55	-2.56	-0.64	0.05					10.55
	3.61	0.16	1.51	-3.68	-5.05	-0.57	2.05					
	9.78	2.56	8.01	-11.84	-4.82	-1.14	0.06	-0.12	-0.71	0.52	-2.81	13.85
	3.80	1.26	1.89	-3.03	-5.31	-1.00	2.22	-2.07	-1.87	0.39	-4.47	

- The semibeta-based pricing model reduces to the traditional CAPM model if the semibeta risk premiums satisfy $H_{0,t}^{CAPM}: \lambda_t^N = \lambda_t^P = -\lambda_t^{M^+} = -\lambda_t^{M^-}$.
- Reject this restriction at the 5% level for 46.1% of the 684 months .
- The significance of the risk premiums for the daily β^N and β^{M-} remain intact to the inclusion of the same set of controls.

Standard risk factors and controls

- Daily Fama-MacBeth regressions

Table 4

Daily Fama-MacBeth regressions The table reports the estimated annualized risk premia and Newey-West robust t-statistics from daily Fama-MacBeth cross-sectional predictive regressions. The daily semibetas are calculated from 15-minute intraday data. All of the control variables are measured prior to the daily returns, as detailed in Appendix A. The estimates are based on all of the S&P 500 constituent stocks and days in the January 1993 to December 2019 sample period.

β	$\beta^{\mathcal{N}}$	$\beta^{\mathcal{P}}$	$\beta^{\mathcal{M}+}$	$\beta^{\mathcal{M}-}$	ME	BM	MOM	REV	RV	IVOL	ILLIQ	R^2
4.49												2.57
3.42												
	18.10	-0.23	-4.49	-7.82								5.42
	5.40	-0.09	-1.14	-2.19								
	19.02	-4.10	-2.63	-11.41	-1.78	-2.01	0.09					8.62
	5.86	-1.64	-0.73	-3.27	-3.63	-1.94	3.36					
	18.79	-1.57	2.33	-7.18	-2.70	-2.64	0.08	-0.43	0.43	-3.26	-0.83	11.23
	5.87	-0.61	0.68	-2.02	-4.62	-2.55	2.83	-5.20	2.16	-4.45	-2.62	

- Reject this restriction at the 5% level for 70.0% of the 6,799 days .

- The significance of the risk premiums for the daily β^N and β^{M-} remain intact to the inclusion of the same set of controls.

Arbitrage risk and semibeta pricing

- Split the cross section of stocks into separate groups of stocks with high and low IVOLs, and compare the risk premium estimates for each group.

$$r_{t+1,i} = \lambda_{0,t+1} + \lambda_{t+1}^{\mathcal{N}}\left(\hat{\beta}_{t,i}^{\mathcal{N}} - \hat{\beta}_{t,i}^{\mathcal{M}^-}\right) + \lambda_{t+1}^{\mathcal{P}}\left(\hat{\beta}_{t,i}^{\mathcal{P}} - \hat{\beta}_{t,i}^{\mathcal{M}^+}\right) + \delta_{t+1}^{\mathcal{M}^+}\hat{\beta}_{t,i}^{\mathcal{M}^+} + \delta_{t+1}^{\mathcal{M}^-}\hat{\beta}_{t,i}^{\mathcal{M}^-} + \epsilon_{t+1,i}$$

- Hypothesis: $\lambda^N = -\lambda^{M-} \rightarrow \delta^{M^-} = 0$
- the 50% of stocks with the lowest IVOL in each of the months

  δ^{M^-} =3.84, with an insignificant t-statistic of 1.02.
- the 50% of stocks with the highest IVOL

  δ^{M^-} =10.98, with a significant t-statistic of 2.59.The different risk premiums for β^N and β^{M^-} may be attributed to arbitrage risk.

Arbitrage risk and semibeta pricing

- Split the cross section of stocks into separate groups of stocks with high and low TOs, and compare the risk premium estimates for each group.

$$r_{t+1,i} = \lambda_{0,t+1} + \lambda_{t+1}^{\mathcal{N}}\left(\hat{\beta}_{t,i}^{\mathcal{N}} - \hat{\beta}_{t,i}^{\mathcal{M}^-}\right) + \lambda_{t+1}^{\mathcal{P}}\left(\hat{\beta}_{t,i}^{\mathcal{P}} - \hat{\beta}_{t,i}^{\mathcal{M}^+}\right) + \delta_{t+1}^{\mathcal{M}^+}\hat{\beta}_{t,i}^{\mathcal{M}^+} + \delta_{t+1}^{\mathcal{M}^-}\hat{\beta}_{t,i}^{\mathcal{M}^-} + \epsilon_{t+1,i}$$

- Hypothesis: $\lambda^{N} = -\lambda^{M-} \rightarrow \delta^{M^-} = 0$
- the 50% of stocks with the lowest TO in each of the months

  $\delta^{M^-} = 8.38$, with an insignificant t-statistic of 2.30.
- the 50% of stocks with the highest TO

  $\delta^{M^-} = 5.60$, with an insignificant t-statistic of 1.22.

Arbitrage risk and semibeta pricing

High-frequency data:

- the 50% of stocks with the lowest IVOL in each of the months

  δ^{M^-} =2.58, with an insignificant t-statistic of 0.54.

- the 50% of stocks with the highest IVOL

δ^{M^-} =23.93, with a significant t-statistic of 3.82.The different risk premiums for β^N and β^{M^-} may be attributed to arbitrage risk.

- the 50% of stocks with the lowest TO in each of the months

  δ^{M^-} =8.04, with an insignificant t-statistic of 2.18.

- the 50% of stocks with the highest TO

  δ^{M^-} =-3.89, with an insignificant t-statistic of -0.79.

Comparison

- Upside and downside betas

$$\hat{\beta}^{+}_{t,i} = \left(\widehat{\beta}^{\mathcal{P}}_{t,i} - \widehat{\beta}^{\mathcal{M}^{+}}_{t,i}\right)\frac{\sum_{k=1}^{m} f^{2}_{t,k}}{\sum_{k=1}^{m}\left(f^{+}_{t,k}\right)^{2}} \qquad \hat{\beta}^{-}_{t,i} = \left(\widehat{\beta}^{\mathcal{N}}_{t,i} - \widehat{\beta}^{\mathcal{M}^{-}}_{t,i}\right)\frac{\sum_{k=1}^{m} f^{2}_{t,k}}{\sum_{k=1}^{m}\left(f^{-}_{t,k}\right)^{2}}$$

- Coskewness and cokurtosis

$$CSK_{t,i} = \frac{\frac{1}{m}\sum_{k=1}^{m}\left(r_{t,k,i} - \bar{r}_{t,i}\right)\left(f_{t,k} - \bar{f}_{t}\right)^{2}}{\sqrt{\frac{1}{m}\sum_{k=1}^{m}\left(r_{t,k,i} - \bar{r}_{t,i}\right)^{2}}\,\frac{1}{m}\sum_{j=1}^{m}\left(f_{t,k} - \bar{f}_{t}\right)^{2}}$$

$$CKT_{t,i} = \frac{\frac{1}{m}\sum_{k=1}^{m}\left(r_{t,k,i} - \bar{r}_{t,i}\right)\left(f_{t,k} - \bar{f}_{t}\right)^{3}}{\sqrt{\frac{1}{m}\sum_{k=1}^{m}\left(r_{t,k,i} - \bar{r}_{t,i}\right)^{2}}\left(\frac{1}{m}\sum_{k=1}^{m}\left(f_{t,k} - \bar{f}_{t}\right)^{2}\right)^{3/2}}$$

Results

- Upside and downside betas
  - A joint test that all of the semibeta risk premiums are zero, leaving only the up and downside betas with nonzero risk premiums, also strongly rejects the null (p-value< 0.01).
  - A joint test that both of the up and downside beta premiums are zero, leaving only the semibetas with nonzero premiums, fails to reject the null, with a p-value of 0.16.
- Coskewness and cokurtosis
  - Joint tests that the semibeta premiums, or the coskewness/cokurtosis premiums, are equal to zero are both rejected at the 5% level.
  - While coskewness and cokurtosis have substantially less cross-sectional explanatory power than semibetas, they do contain additional information about non-linear dependencies over and above the semibetas.
  - Coskewness and cokurtosis are primarily driven by joint dependencies in the tails.

Results

β^N	β^P	β^{M+}	β^{M-}	β^+	β^-	CSK	CKT	R^2
10.54	1.84	4.59	-6.00					5.16
4.51	1.17	1.32	-1.97					
				1.21	3.23			3.41
				1.85	3.84			
11.53	-2.30	2.82	-11.20	-6.21	1.84			5.48
3.50	-0.55	1.34	-2.90	-0.97	1.27			
						5.44	2.13	1.68
						3.05	2.77	
18.03	-1.59	3.70	-11.32			12.13	-2.64	6.40
5.02	-0.76	1.09	-3.31			4.36	-3.41	

- $H_{0,t}^{UP+DOWN}: \lambda_t^N = -\lambda_t^{M^-} \cap \lambda_t^P = -\lambda_t^{M^+}$
- Reject this restriction at the 5% level for 42.0% of the 684 months.

- $H_{0,t}^{DOWN}: \lambda_t^N = -\lambda_t^{M^-} \cap \lambda_t^P = -\lambda_t^{M^+} = 0$
- Reject this restriction at the 5% level for 50.5%% of the 684 months.

Results(High-frequency data)

Table 5

Daily Fama-MacBeth regressions on other measures The table reports the estimated annualized risk premia and Newey-West robust t-statistics from daily Fama-MacBeth cross-sectional predictive regressions. The daily semibetas, up and downside betas, and coskewness and cokurtosis measures are calculated from 15-minute intraday data based on all of the S&P 500 constituent stocks and days in the January 1993 to December 2019 sample period.

$\beta^{\mathcal{N}}$	$\beta^{\mathcal{P}}$	$\beta^{\mathcal{M}+}$	$\beta^{\mathcal{M}-}$	β^{+}	β^{-}	CSK	CKT	R^2
18.10	-0.23	-4.49	-7.82					5.42
5.40	-0.09	-1.14	-2.19					
				-0.37	5.71			3.60
				-0.38	5.32			
15.84	-8.69	7.76	-11.13	-4.27	2.10			6.48
3.89	-1.03	0.78	-2.16	-1.05	0.78			
						-1.86	0.88	1.51
						-0.71	0.96	
26.21	-2.19	-5.06	-14.99			10.93	-3.71	6.31
6.28	-0.73	-1.20	-4.00			3.18	-3.68	

- $H_{0,t}^{UP+DOWN}: \lambda_t^N = -\lambda_t^{M^-} \cap \lambda_t^P = -\lambda_t^{M^+}$
- Reject this restriction at the 5% level for 63.5% of the 6799 days.

- $H_{0,t}^{DOWN}: \lambda_t^N = -\lambda_t^{M^-} \cap \lambda_t^P = -\lambda_t^{M^+} = 0$
- Reject this restriction at the 5% level for 72.8% of the 6799 days.

Semibeta trading strategies

- Portfolios :
  - bet on β (traditional market beta)
  - bet on β^{N}
  - bet against β^{M-}
  - betting on β^{N} ” and “betting against β^{M-}” (semibeta)

- Value-weighted long/short positions in high/low quintile of S&P 500 stocks
  - Avoids smal and dificult to short micro-cap stocks, and the use of rank-weighted portfolios
- Sharpe ratios, alphas and factor loadings:
  - Four-factor Fama-French-Carhart model: MKT, SML, HML, MOM
  - Five-factor Fama-French model: MKT, SML, HML, RMW (profitability: robust minus weak), CMA (investment: conservative minus aggressive)

Results

	β		Semi-β		$\beta^{\mathcal{N}}$		$\beta^{\mathcal{M}^-}$	
Avg ret	5.62		8.17		10.02		5.56	
Std dev	15.37		8.86		15.78		7.80	
Sharpe	0.37		0.92		0.63		0.71	
α	2.52	3.94	6.84	7.52	6.89	8.59	6.02	5.68
	1.21	1.98	5.92	6.49	3.31	4.22	3.93	3.65
β_{MKT}	0.57	0.50	0.28	0.25	0.59	0.51	-0.02	-0.01
	75.03	62.98	67.31	53.06	76.91	62.03	-3.22	-2.28
β_{SMB}	0.27	0.15	0.31	0.24	0.39	0.26	0.23	0.22
	18.94	10.67	38.92	28.43	27.12	17.78	21.88	19.12
β_{HML}	-0.01	0.22	-0.01	0.16	-0.06	0.20	0.04	0.12
	-0.42	14.59	-1.10	17.85	-3.98	12.61	3.72	9.97
β_{MOM}	-0.21		-0.16		-0.22		-0.10	
	-20.47		-27.83		-21.16		-13.06	
β_{RMW}		-0.42		-0.25		-0.46		-0.03
		-21.65		-21.46		-22.80		-2.13
β_{CMA}		-0.33		-0.24		-0.40		-0.08
		-14.01		-17.42		-16.37		-4.48
R^2	56.11	58.20	55.83	56.68	58.98	61.96	10.19	8.02

Results

	β^- - β^+		β^-		β^+		CKT - CKS		CSK		CKT	
Avg ret	1.83		7.11		-5.12		-0.19		-1.91		1.96	
Std dev	5.34		14.46		13.58		6.11		7.49		9.17	
Sharpe	0.34		0.49		-0.38		0.03		-0.25		0.21	
α	1.46	1.78	4.15	5.64	-2.92	-3.75	-0.56	0.02	-2.12	-1.98	0.67	1.69
	1.33	1.63	1.95	2.70	-1.42	-1.87	-0.52	0.02	-1.42	-1.32	0.43	1.11
β_{MKT}	0.04	0.03	0.52	0.45	-0.45	-0.40	0.13	0.11	0.02	0.01	0.24	0.20
	8.87	5.91	66.70	53.85	-59.32	-49.73	33.19	25.54	2.86	2.11	42.74	33.18
β_{SMB}	0.01	0.01	0.25	0.15	-0.22	-0.13	-0.03	-0.06	0.02	0.03	-0.09	-0.14
	1.66	1.35	17.21	9.92	-15.90	-8.79	-4.50	-7.37	2.37	2.37	-8.43	-12.49
β_{HML}	-0.02	-0.03	-0.03	0.16	-0.02	-0.23	-0.08	-0.05	-0.02	-0.04	-0.13	-0.07
	-2.86	-3.86	-1.81	10.12	-1.21	-14.91	-10.22	-6.62	-2.20	-3.49	-11.91	-5.75
β_{MOM}	0.03		-0.15		0.22		0.00		0.03		-0.03	
	5.62		-14.73		21.42		-0.82		3.49		-4.41	
β_{RMW}		-0.01		-0.37		0.35		-0.10		0.00		-0.20
		-1.20	0.00	-18.07	0.00	17.54	0.00	-9.18	0.00	0.25	0.00	-12.92
β_{CMA}		-0.01		-0.30		0.27		-0.05		0.01		-0.12
		-1.13		-12.13		11.51		-4.38		0.67		-6.67
R^2	2.19	1.71	50.24	52.95	47.83	48.07	16.30	17.70	0.64	0.45	26.09	28.55

Extension

- Focuses mainly on the U.S. stock market, and cross-market applicability needs verification.
- The strategy’s performance during crisis periods has not been tested separately.

- Expand asset classes， bonds and options.
- Further quantify the extent to which arbitrage constraints affect pricing differences.