Quantity, Risk, and Return

Yu An — *Johns Hopkins*
Yinan Su — *Johns Hopkins*
Chen Wang — *Notre Dame*

AFA Annual Meetings
January 5, 2026, Philadelphia

Research question

Expected stock returns: why different stocks earn different returns?

- In theory: risk
  investors are risk averse, require compensation for bearing risk
  $\Rightarrow$ high-risk high-return

- Empirical challenges:
  - high-risk high-return is elusive in data (e.g., flat SML)
  - risk-based models (β) hardly predict stock returns
    vs. machine learning + characteristics: unstructured predictions

What is missing in factor pricing?

Integrate quantity into risk-return modeling

- APT: expected stock return driven by factor exposures (β)

$$\mathbb{E}_t r_{i,t+1} = \sum_k \mu_{k,t} \beta_{i,k,t}$$

Integrate quantity into risk-return modeling

- APT: expected stock return driven by factor exposures (β)

$$\mathbb{E}_t r_{i,t+1} = \sum_k \mu_{k,t} \beta_{i,k,t}$$

- Add quantity ($q_{k,t}$, factor-level time series)

  - model: $$\mu_{k,t} = \lambda_k q_{k,t}$$

  - $q \uparrow$: sophisticated investors buying factor risk recently
    constructed as retail selling via mutual fund flow-induced trading (FIT)

  - finding: strong q-μ positive association (for almost all factors)

  - interpretation: hold more quantity $\Rightarrow$ greater risk compensation

Integrate quantity into risk-return modeling

- APT: expected stock return driven by factor exposures (β)

$$\mathbb{E}_t r_{i,t+1} = \sum_k \mu_{k,t} \beta_{i,k,t}$$

- Add quantity ($q_{k,t}$, factor-level time series)

  - model: $\mu_{k,t} = \lambda_k q_{k,t}$

  - $q \uparrow$: sophisticated investors buying factor risk recently
    constructed as retail selling via mutual fund flow-induced trading (FIT)

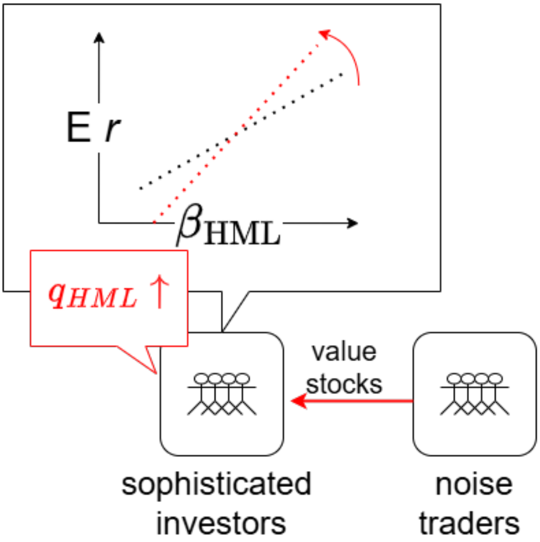
  - finding: strong q-μ positive association (for almost all factors)

  - interpretation: hold more quantity $\Rightarrow$ greater risk compensation

Together:

- "β-times-quantity" (BTQ) predicts stock returns (OOS $R^2 \approx 1\% \gtrapprox$ ML sota)

$$r_{i,t+1} \sim \beta_{i,k,t} q_{k,t} \quad \text{vs. canonical} \quad r_{i,t+1} \sim \beta_{i,k,t}$$



Expected stock return $\mathbb{E}_t r_{i,t+1}$ depends on:

- not only factor loading $\beta_{i,k,t}$,
- but also $q_{k,t}$

Construct $q_{k,t}$

the quantity of factor risk absorbed by sophisticated investors recently

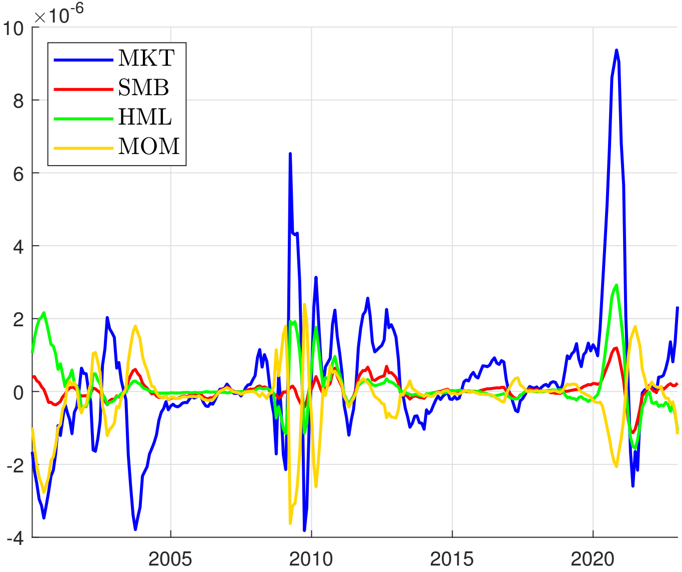
- Stock-level flows:
  $\$\text{flow}_{i,t}^{\text{stock}} = -$ mutual fund flow-induced trading of stock i at month t
  $\$\text{flow}^{\text{stock}} \uparrow$: retail selling or sophisticated buying
- Aggregate stock-level flow to factor-level

$$\text{flow}_{k,t}^{\text{factor}} := \sum_{\text{stock } i} \$\text{flow}_{i,t}^{\text{stock}} \, \text{cov}_{i,k,t}$$

↑ stock's exposure to factor k

- Accumulate flow in recent six months, with normalization

Construction result: $\widetilde{q}_{k,t}$ time-series plot



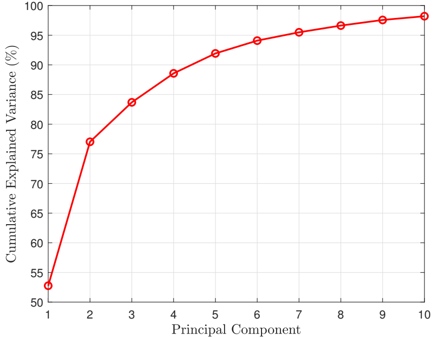
$q's$ are not highly correlated across factors

robust evidence across different factors

Correlation matrix for q's of FFC4

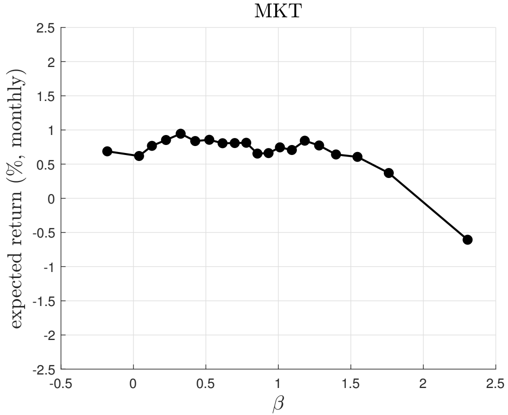
	MKT	SMB	HML	MOM
MKT	1			
SMB	0.55	1		
HML	0.47	0.57	1	
MOM	-0.47	-0.23	-0.75	1

PC variances for q's of 153 JKP factors



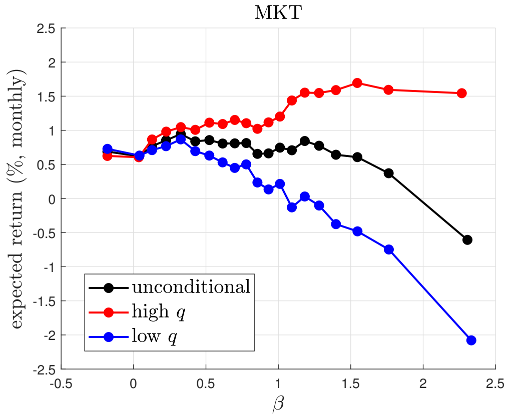
Baseline: security market line (SML) is flat

contradicts "high risk, high return"



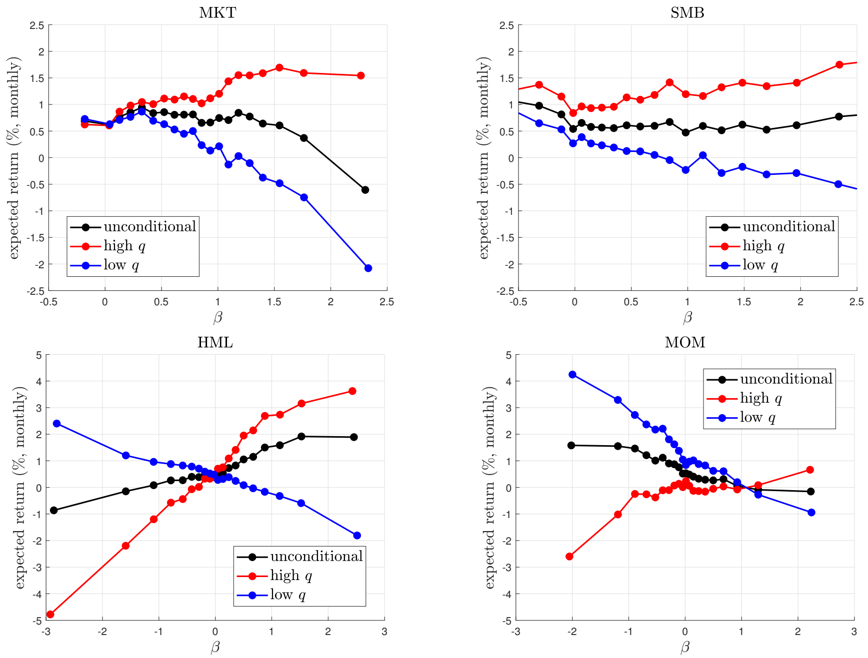
SML as non-parametric regression: $\mathbb{E}_t[r_{i,t+1}] = Er(\beta_{i,k,t})$ for the stock-month panel

Risk-return tradeoff (SML) conditioning on q



SML as non-parametric regression: $\mathbb{E}_t[r_{i,t+1}] = Er(\beta_{i,k,t})$

upgraded SML: one more input: $\mathbb{E}_t[r_{i,t+1}] = Er(\beta_{i,k,t}, q_{k,t})$



MKT
SMB
HML
MOM
expected return (%, monthly)
β
unconditional
high q
low q

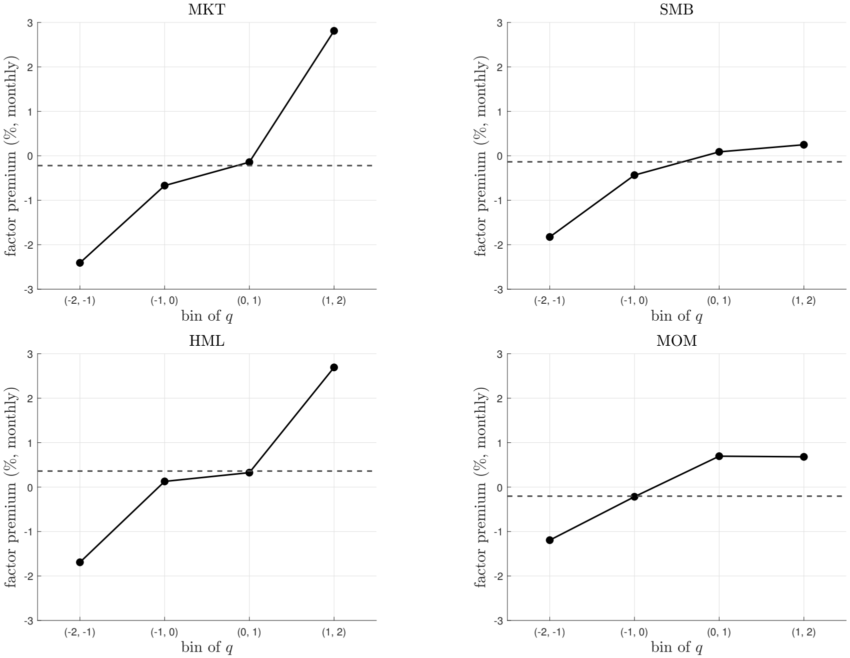
Fama-MacBeth factor premium increases with $q_{k,t}$

Estimation:

- Fama-MacBeth: cross-sectional reg $r_{i,t+1}$ on $\widehat{\beta}_{k,i,t}$, get coef. $\gamma_{k,t+1}$
  - Canonical: $\mu_k =$ time-series average of $\gamma_{k,t+1}$
  - Upgraded: varying $\mu_{k,t} = \mu_k(q_{k,t})$ conditional on $q_{k,t}$

Model:

$$\mathbb{E}_t[r_{i,t+1}] = Er(\beta_{i,k,t}, q_{k,t}) = \beta_{i,k,t}\mu_k(q_{k,t})$$


MKT
SMB
HML
MOM
factor premium (%, monthly)
bin of q
(-2, -1)
(-1, 0)
(0, 1)
(1, 2)
3
2
1
0
-1
-2
-3

BTQ (beta-times-quantity) predicts stock returns

- Factor pricing (APT):

$$\mathbb{E}_t[r_{i,t+1}] = \sum_k \beta_{i,k}\mu_{k,t}$$

- Factor premium is constant vs. linear function of $q_{k,t}$:

$$\mu_{k,t} = \mu_k \quad \text{vs.} \quad \lambda_k q_{k,t}$$

- Plug in:

$$\mathbb{E}_t[r_{i,t+1}] = \sum_k \mu_k \beta_{i,k,t} \quad \text{vs.} \quad \sum_k \lambda_k \beta_{i,k,t} q_{k,t}$$

BTQ (beta-times-quantity) predicts stock returns

- Factor pricing (APT):

$$\mathbb{E}_t[r_{i,t+1}] = \sum_k \beta_{i,k}\mu_{k,t}$$

- Factor premium is constant vs. linear function of $q_{k,t}$:

$$\mu_{k,t} = \mu_k \quad \text{vs.} \quad \lambda_k q_{k,t}$$

- Plug in:

$$\mathbb{E}_t[r_{i,t+1}] = \sum_k \mu_k \beta_{i,k,t} \quad \text{vs.} \quad \sum_k \lambda_k \beta_{i,k,t} q_{k,t}$$

- Estimation: BTQ predictive regression (stock-month panel)

$$r_{i,t+1} = \sum_k \lambda_k \left(\widehat{\beta}_{i,k,t} q_{k,t}\right) + error_{i,t+1}, \qquad \forall i,t$$

vs. "β-only"

$$r_{i,t+1} = \sum_k \mu_k \widehat{\beta}_{i,k,t} + error_{i,t+1}, \qquad \forall i,t$$

BTQ vs. β-only, single factor

	Fama-French-Carhart factors				Across 153 JKP factors		
	MKT	SMB	HML	MOM	Q25	Median	Q75
Panel A: IS R^2 comparison, full sample 2000-2022 (%)							
BTQ	1.01	0.30	1.00	0.91	0.39	0.62	0.95
β-only	0.05	0.05	0.12	0.06	0.02	0.06	0.10
Panel B: OOS R^2 comparison, evaluation window 2010-2022 (%)							
BTQ	0.75	0.60	0.84	0.65	0.20	0.38	0.67
β-only	0.05	-0.10	0.15	0.02	-0.03	0.04	0.11

BTQ vs. β-only, single factor

	Fama-French-Carhart factors				Across 153 JKP factors		
	MKT	SMB	HML	MOM	Q25	Median	Q75
Panel A: IS R^2 comparison, full sample 2000-2022 (%)							
BTQ	1.01	0.30	1.00	0.91	0.39	0.62	0.95
β-only	0.05	0.05	0.12	0.06	0.02	0.06	0.10
Panel B: OOS R^2 comparison, evaluation window 2010-2022 (%)							
BTQ	0.75	0.60	0.84	0.65	0.20	0.38	0.67
β-only	0.05	-0.10	0.15	0.02	-0.03	0.04	0.11

BTQ vs. β-only, single factor, coefficients

	Fama-French-Carhart factors				Across 153 JKP factors		
	MKT	SMB	HML	MOM	Q25	Median	Q75
Panel C: full-sample coefficient comparison: 2000-2022							
BTQ							
λ_k	1.80	0.72	1.48	1.77	0.62	0.99	1.48
t-stat	(4.18)	(2.76)	(3.52)	(3.38)	(2.24)	(2.96)	(3.69)
β-only							
μ_k	0.38	0.31	0.56	-0.50	-0.33	-0.14	0.22
t-stat	(1.07)	(1.25)	(1.71)	(-1.23)	(-1.52)	(-0.71)	(1.11)

BTQ vs. β-only, multi-factor

	CAPM	FF3	FF3C	FF5	FF5C
	$K = 1$	3	4	5	6
Panel A: IS R^2, full sample 2000-2022 (%)					
BTQ	1.01	1.17	1.19	1.17	1.21
β-only	0.05	0.17	0.21	0.18	0.22
Panel B: OOS R^2, evaluation window 2010-2022 (%)					
BTQ	0.75	1.03	1.07	0.44	0.65
β-only	0.05	0.15	0.22	-0.26	-0.05

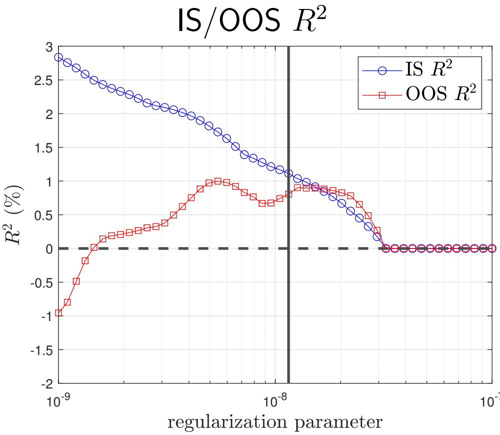
coefficents

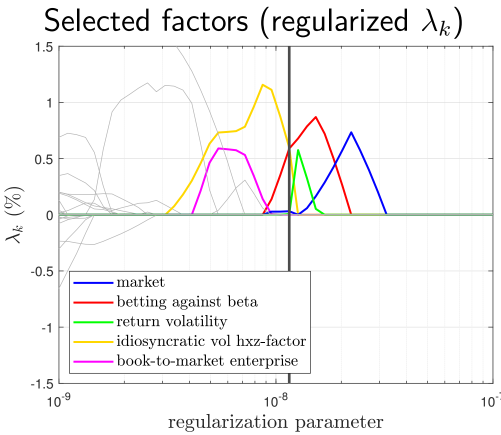
“Taming the factor zoo” with BTQ

- So many proposed factors, which are fundamental?
- New perspective to discipline factors with quantity
  - old question: $\mu_k > 0$? is there factor premium?
  - new question: $\lambda_k > 0$? does factor premium vary with investor risk holdings?
- Method:
  - BTQ prediction with 159 FF+JKP factors
  - factor selection with Lasso

BTQ, selecting from factor zoo

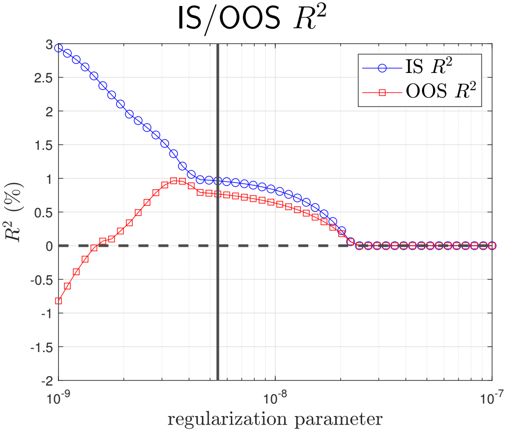
OOS predictive $R^2 \approx$ 1%, 5 factors selected, positive coefficients

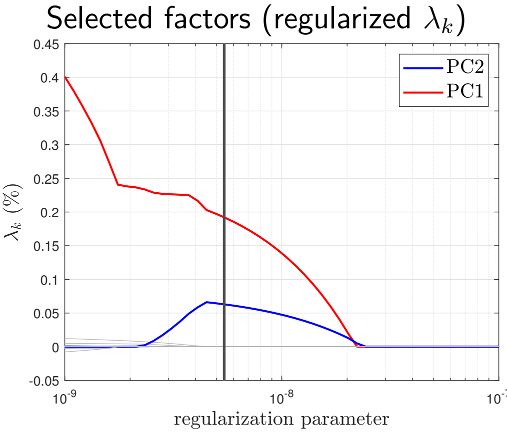




BTQ, selecting from PC factors

PC1 and PC2 selected, positive coefficients, high OOS R^2





Quantity-premium association is stronger when intermediary risk-bearing capacity is lower

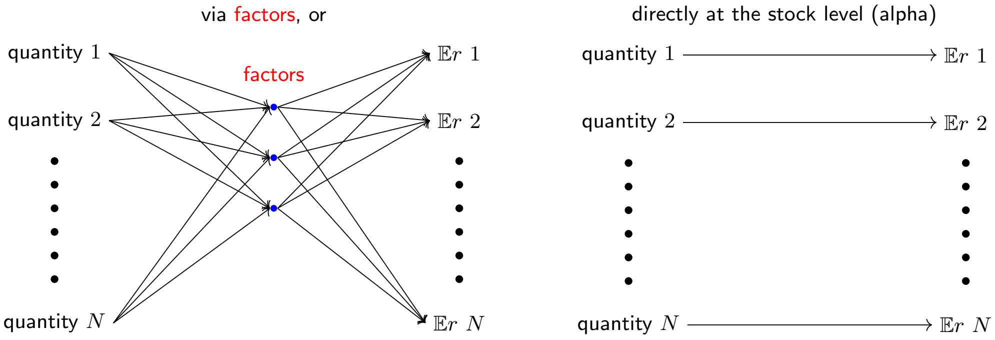
support risk-based interpretation of quantity-premium relation

$$r_{i,t+1} = \lambda_{k,\text{const}} \widehat{\beta}_{i,k,t} q_{k,t} + \lambda_{k,\text{slope}} \widehat{\beta}_{i,k,t} q_{k,t} \times \text{risk-bearing capacity}_t + error_{i,t+1}$$

	baseline BTQ	BTQ $\times$ risk-bearing capacity	
risk-bearing capacity proxy used	none	ΔICR	BKX return
A. Market factor			
$\lambda_{\text{mkt,const}}$ (%)	1.80	2.49	1.21
t-stat	(4.18)	(4.17)	(2.76)
$\lambda_{\text{mkt,slope}}$ (%)		**−1.11**	**−0.90**
t-stat		**(−2.24)**	**(−3.12)**
full-sample R^2 (%)	1.01	1.21	1.37
OOS R^2 (%)	0.75	0.62	0.80

Alpha model with quantity at individual stock level

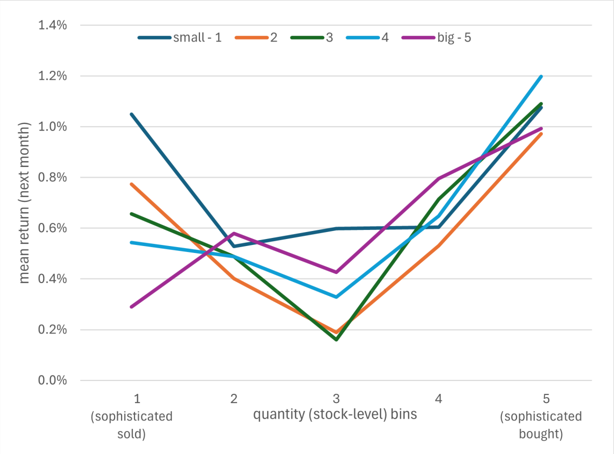
- Quantity affects expected returns via factors or also directly at stock level?



- Yes, quantity-driven alpha complements BTQ
- U-shaped quantity$_{i,t}$-$\mathbb{E}_t r_{i,t+1}$ relation, mostly in small stocks

Alpha model with quantity at individual stock level
(preliminary results)

- U-shaped quantity$_{i,t}$-$\mathbb{E}_t r_{i,t+1}$ relation, mostly in small stocks
  potential trend-following of extreme mutual fund inflows (maybe meme stocks)
- $q_{i,t}^{stock}$–size 5×5 double sort:



More results

- $q_{\text{mkt},t}$ negatively correlated with sentiment measures
  support interpretation of q direction: $q \uparrow=$ sophisticated buy / noise sell
- β_k and q_k cannot mis-match
  a factor's q_k is only relevant to risk-return trade-off along that factor's β_k
  suggest factor risk structure is essential
- beta-times-[factor momentum] does not work
  suggest "flow chasing past performance" is not an explanation
- beta-times-[macro variables] does not work
  suggest q is not repackaging known factor return predictors

- Robust results to size groups, time periods, and alternative q construction specifications

Quantity, Risk, and Return

factor risk + quantity to explain expected stock returns

Findings:

- Risk-return tradeoff (β-$\mathbb{E}r$ relation) depends on quantity
- BTQ predicts stock returns
- A new perspective to the "factor zoo" problem with quantity

Quantity, Risk, and Return

Yu An
Johns Hopkins

Yinan Su
Johns Hopkins

Chen Wang
Notre Dame

AFA Annual Meetings
January 5, 2026, Philadelphia