

# What Drives Variation in Investor Portfolios?

## Estimating the Role of Beliefs and Risk Preferences

**Mark Egan**  
Harvard University  
& NBER

**Alexander MacKay**  
Harvard University

**Hanbin Yang**  
Harvard University

May 1, 2023

# Motivation

Why do individuals make different decisions with their investments?

# Motivation

Why do individuals make different decisions with their investments?

Claim: faced with a similar set of investment choices, investors vary widely in their allocations

# Motivation

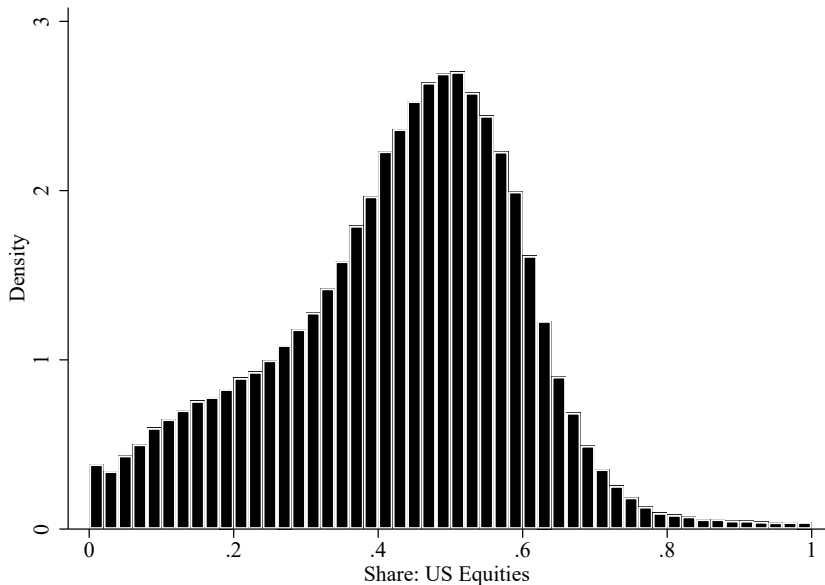
Why do individuals make different decisions with their investments?

Claim: faced with a similar set of investment choices, investors vary widely in their allocations

Key challenges for understanding demand for assets:

- Investors face a *portfolio problem*
- Behavior depends on both beliefs and preferences for risk
- Limited large-scale data about investment options and allocations

# Motivation: 401(k) Allocations



# This Paper

Use large-scale data to understand demand for assets

- 401k portfolio allocations over the period 2009-2019
- Investment menus and plan-level allocations for 70,000 plans
- \$7.3 trillion dollars held in 401(k) plans in 2021

# This Paper

Use large-scale data to understand demand for assets

- 401k portfolio allocations over the period 2009-2019
- Investment menus and plan-level allocations for 70,000 plans
- \$7.3 trillion dollars held in 401(k) plans in 2021

Estimate a portfolio demand model a la Markowitz (1952)

- Allow beliefs and risk aversion to vary arbitrarily across investors (i.e., nonparametric identification)
- Leverage exogenous variation in expense ratios for identification

# This Paper

Use large-scale data to understand demand for assets

- 401k portfolio allocations over the period 2009-2019
- Investment menus and plan-level allocations for 70,000 plans
- \$7.3 trillion dollars held in 401(k) plans in 2021

Estimate a portfolio demand model a la Markowitz (1952)

- Allow beliefs and risk aversion to vary arbitrarily across investors (i.e., nonparametric identification)
- Leverage exogenous variation in expense ratios for identification

Examine the determinants of expected returns (beliefs)

- Beliefs vary with sector of employment and demographics
  - Beliefs reflect recent fund performance, local economic conditions, and the performance of the investor's employer
- Beliefs (and investment decisions) depend on local information



## Some Details on Findings

Recover reasonable estimates of beliefs and risk aversion

- Avg. investor expected the market to go up by 10% annually
- Avg. risk aversion is 4.5

## Some Details on Findings

Recover reasonable estimates of beliefs and risk aversion

- Avg. investor expected the market to go up by 10% annually
- Avg. risk aversion is 4.5

Beliefs explain more variation in allocations than risk aversion

- Menus play a smaller role (average menu: 26 options)
- Investors are sensitive to fees and appear to rebalance

## Some Details on Findings

Recover reasonable estimates of beliefs and risk aversion

- Avg. investor expected the market to go up by 10% annually
- Avg. risk aversion is 4.5

Beliefs explain more variation in allocations than risk aversion

- Menus play a smaller role (average menu: 26 options)
- Investors are sensitive to fees and appear to rebalance

What determines beliefs and risk preferences?

- Vary with demographics and employment
  - ▷ Educated investors appear more optimistic and risk averse
  - ▷ Older investors appear more pessimistic and risk averse
  - ▷ Median expectations: Construction (8.5%) vs. Real Estate (10.8%)
- Beliefs are extrapolative
  - ▷ Beliefs about the market are correlated with employer performance and local economic conditions
  - ▷ Investors extrapolate beliefs even when setting up new plans

# Outline of Talk

1. **Data**
2. Model
3. Estimates of Beliefs and Risk Preferences
4. Counterfactual Allocations
5. Evidence on the Formation of Beliefs

## Data: 401 (k) Plans (Form 5500)

Asset allocations for the near universe of 401(k) plans

- Data cover 85% of plan assets (70,000 plans)
- Annual data over the period 2009-2019
  - ▷ 450k plan-by-year observations
  - ▷ 11m plan-by-investment option observations

## Data: 401 (k) Plans (Form 5500)

Asset allocations for the near universe of 401(k) plans

- Data cover 85% of plan assets (70,000 plans)
- Annual data over the period 2009-2019
  - ▷ 450k plan-by-year observations
  - ▷ 11m plan-by-investment option observations

For each plan-year we observe:

- Investment menu and holdings
- Other characteristics: participation rate, % retired, industry, etc.
- Match with ACS demographic data at the county  $\times$  year  $\times$  industry level

## Data: 401 (k) Plans (Form 5500)

Asset allocations for the near universe of 401(k) plans

- Data cover 85% of plan assets (70,000 plans)
- Annual data over the period 2009-2019
  - ▷ 450k plan-by-year observations
  - ▷ 11m plan-by-investment option observations

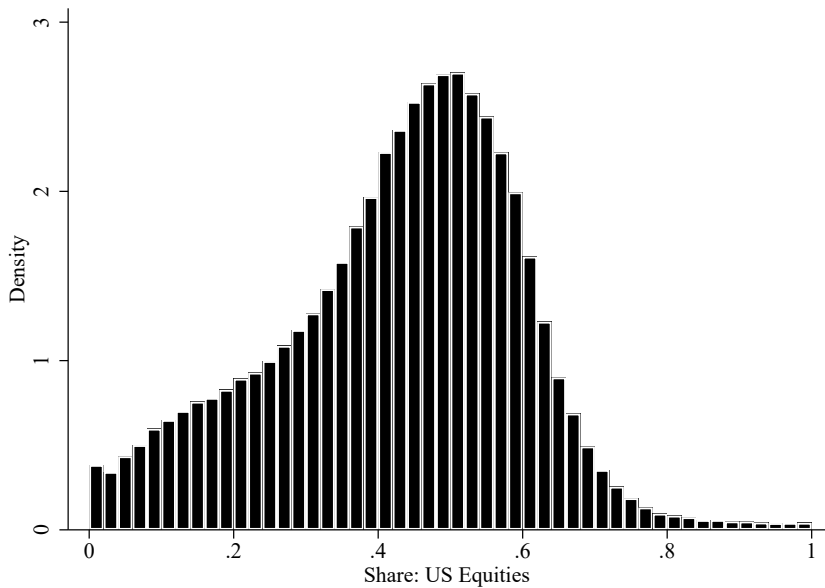
For each plan-year we observe:

- Investment menu and holdings
- Other characteristics: participation rate, % retired, industry, etc.
- Match with ACS demographic data at the county  $\times$  year  $\times$  industry level

At the fund/investment option level we observe:

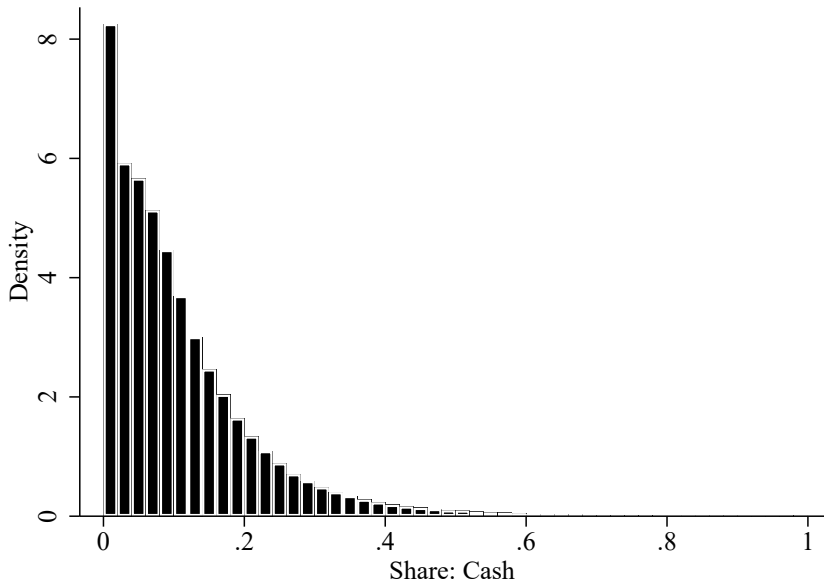
- Type of investment vehicle and investment category
- Expenses and historical returns (CRSP)

# Holdings: US Equities

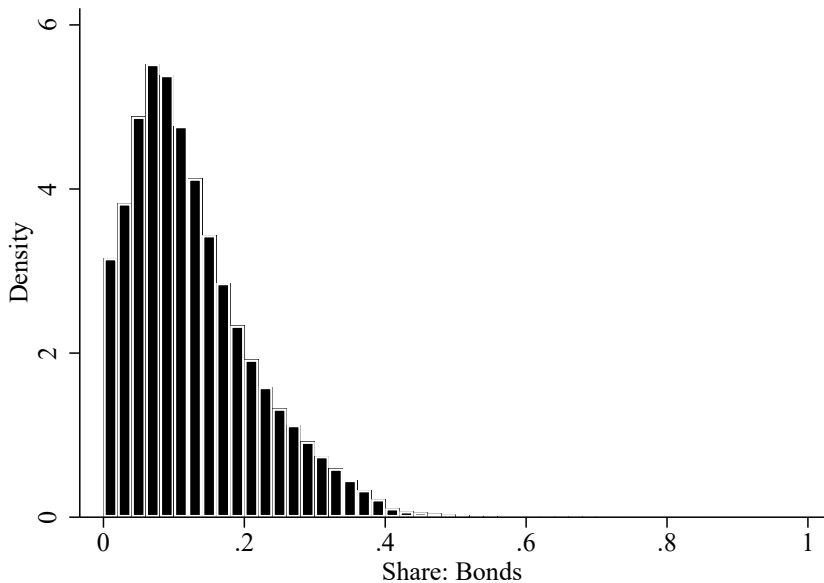




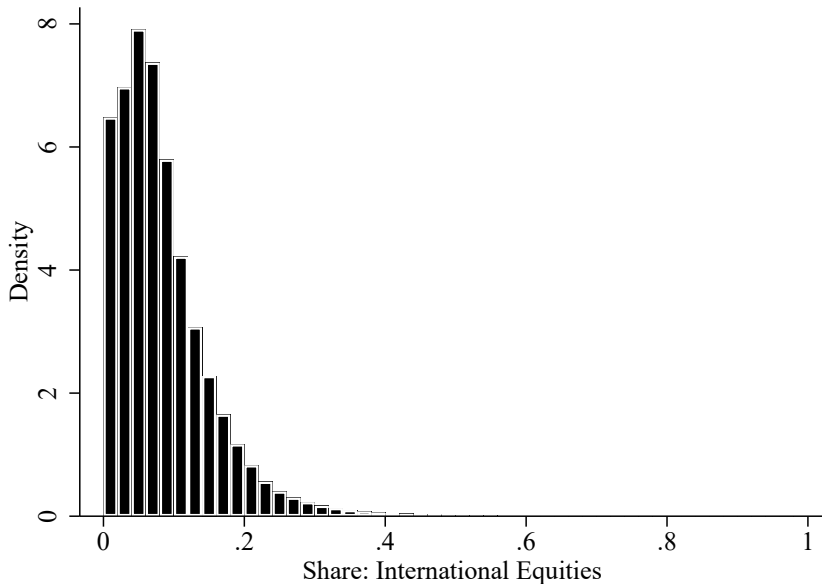
# Holdings: Cash



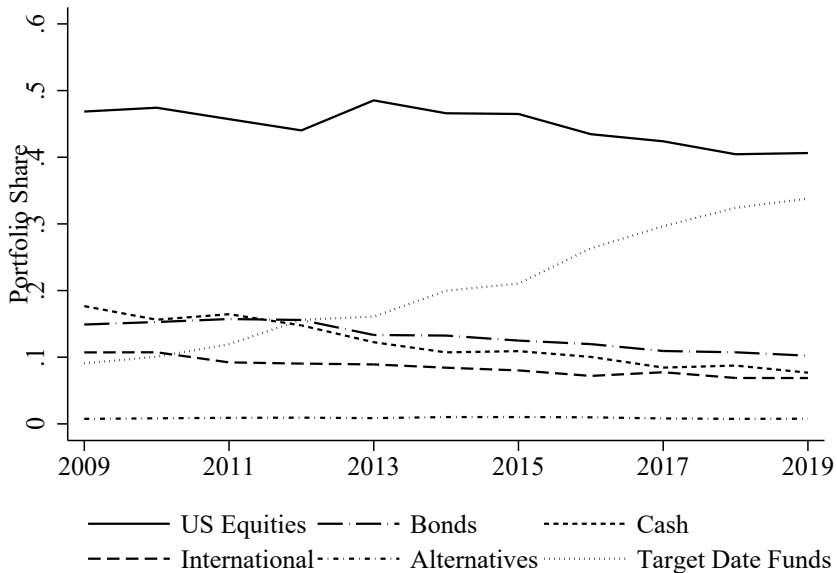
# Holdings: Bonds



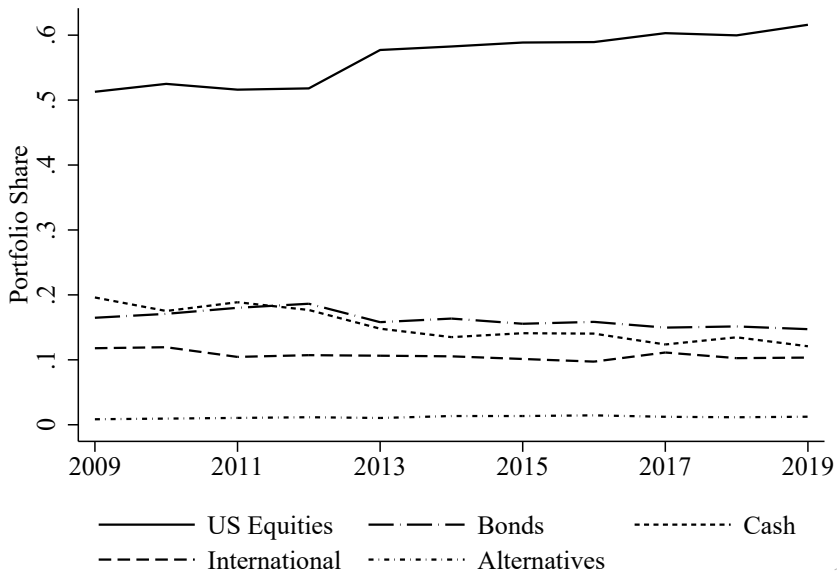
# Holdings: International Assets



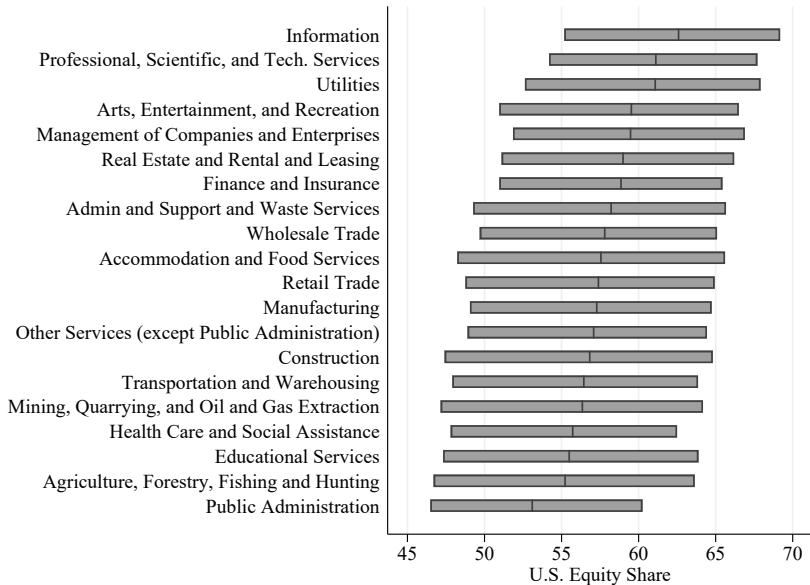
# Holdings Over Time



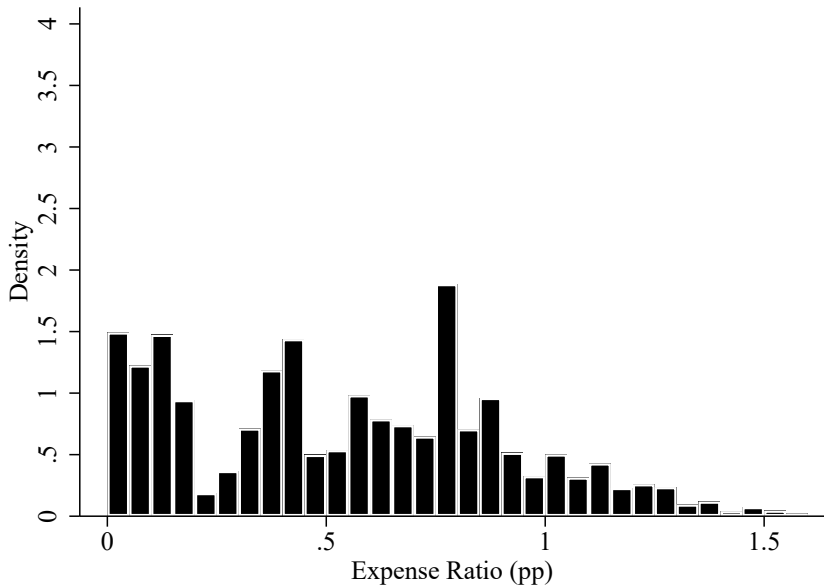
# Holdings Over Time - Excluding Target Date Funds



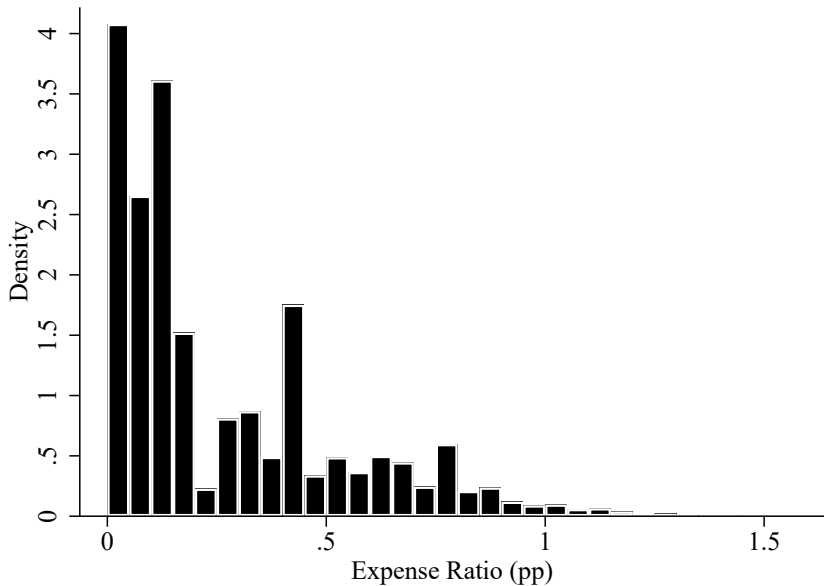
# Equity Allocation vs. Industry



# Fund/Investment Expenses



## Fund/Investment Expenses - Weighted by AUM





# Takeaways

## Substantial variation in portfolios

- Correlated with demographics and industry
- Equity exposure is positively correlated with:
  - ▷ Income/Wealth
  - ▷ Education
  - ▷ Employment
  - ▷ Non-minority

## Investors respond to fees

- Elasticity of demand: -0.4 (Berry 1994)

## Other results:

- Find similar patterns across other asset classes
- Autocorrelation in portfolio holdings is 0.8-0.9
- Menu effects appear less relevant than previously documented

# Outline of Talk

1. Data
2. **Model**
3. Estimates of Beliefs and Risk Preferences
4. Counterfactual Allocations
5. Evidence on the Formation of Beliefs

## Model

Heterogeneity in portfolio holdings could be driven by differences in:

- 401(k) investment menus/expenses
- Risk aversion
- Beliefs

# Model

Heterogeneity in portfolio holdings could be driven by differences in:

- 401(k) investment menus/expenses
- Risk aversion
- Beliefs

Develop a model to separately recover risk aversion and beliefs

- Identification:
  - ▷ Investors choose from a fixed menu
  - ▷ Exogenous variation in fund expenses
- Provides additional insight into holdings
  - ▷ Why do wealthy investors and educated investors have higher equity exposures?
  - ▷ Is it because of risk aversion or beliefs?
- Provides insight into belief formation

## Investor's Problem

Investor  $i$  chooses the  $K \times 1$  vector of weights  $\omega_i$  to maximize

$$\max_{\omega} \omega_i'(\boldsymbol{\mu}_i - \boldsymbol{p}) + (1 - \omega_i'\mathbf{1})R_F - \frac{\lambda_i}{2}\omega_i'\boldsymbol{\Sigma}\omega_i,$$

where

- $\boldsymbol{\mu}_i$  is a vector of investor  $i$ 's expectations of fund returns
- $\boldsymbol{p}$  is a vector of fund expenses
- $R_F$  is the risk-free return
- $\lambda_i$  is risk aversion
- $\boldsymbol{\Sigma}$  is the  $K \times K$  covariance matrix of expected fund returns

## Estimation Strategy

The investor's FOC is

$$\underbrace{\Sigma}_{\text{Obs./Estimated}} \times \underbrace{\omega_i}_{\text{Observed}} = - \underbrace{\frac{1}{\lambda_i}}_{\text{Est. Parameter}} \times \underbrace{\mathbf{p}}_{\text{Observed}} + \underbrace{\frac{1}{\lambda_i} (\boldsymbol{\mu}_i - \mathbf{1}R_F)}_{\text{Residual}}$$

- Given the covariance matrix  $\Sigma$  we can recover  $\lambda_i$  via OLS/IV
- Given  $\lambda_i$  we can directly recover  $\boldsymbol{\mu}_i$

# Estimation Strategy

The investor's FOC is

$$\underbrace{\Sigma}_{\text{Obs./Estimated}} \times \underbrace{\omega_i}_{\text{Observed}} = - \underbrace{\frac{1}{\lambda_i}}_{\text{Est. Parameter}} \times \underbrace{p}_{\text{Observed}} + \underbrace{\frac{1}{\lambda_i} (\mu_i - \mathbf{1}R_F)}_{\text{Residual}}$$

- Given the covariance matrix  $\Sigma$  we can recover  $\lambda_i$  via OLS/IV
- Given  $\lambda_i$  we can directly recover  $\mu_i$

Implementation:

- Estimate  $\Sigma$  using a 6 factor model with historical returns
- Need variation in  $p$  that is orthogonal to beliefs (Hausman IVs)
- Estimate the FOC using plan-level data (i.e., recover avg. plan beliefs)
- Risk aversion
  - ▷ Constant within a plan but heterogeneous across plans
  - ▷ Parameterized as a function of plan demographics and time

# Model Strengths and Limitations

## Strengths:

- Recover the distributions of beliefs and risk aversion across investors (e.g., structural parameters)
- Do not impose structure on beliefs (may not be rational)
- Transparent/analogous to demand estimation
- Substitution patterns disciplined by theory



# Model Strengths and Limitations

## Strengths:

- Recover the distributions of beliefs and risk aversion across investors (e.g., structural parameters)
- Do not impose structure on beliefs (may not be rational)
- Transparent/analogous to demand estimation
- Substitution patterns disciplined by theory

## Potential limitations:

- Measurement error in  $\Sigma$  or any optimization error will be captured in the residual  $\epsilon_{mkt}$  and treated as beliefs

# Model Strengths and Limitations

## Strengths:

- Recover the distributions of beliefs and risk aversion across investors (e.g., structural parameters)
- Do not impose structure on beliefs (may not be rational)
- Transparent/analogous to demand estimation
- Substitution patterns disciplined by theory

## Potential limitations:

- Measurement error in  $\Sigma$  or any optimization error will be captured in the residual  $\epsilon_{mkt}$  and treated as beliefs
- Investors may not actively re-balance their portfolios
  - ▷ Autocorrelation in holdings is 0.8-0.9
  - ▷ Robustness: use newly introduced 401(k) plans (e.g., active choice)
  - ▷ Focus on the cross section of beliefs

# Model Strengths and Limitations

## Strengths:

- Recover the distributions of beliefs and risk aversion across investors (e.g., structural parameters)
- Do not impose structure on beliefs (may not be rational)
- Transparent/analogous to demand estimation
- Substitution patterns disciplined by theory

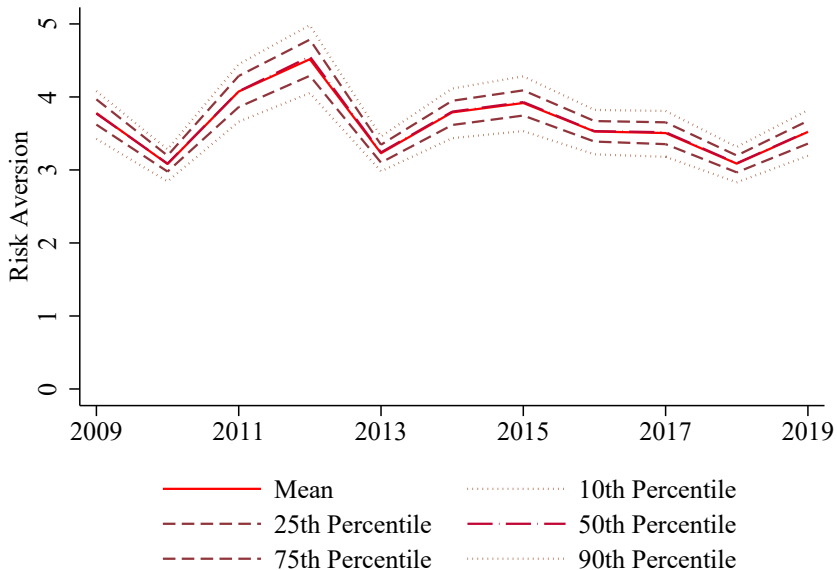
## Potential limitations:

- Measurement error in  $\Sigma$  or any optimization error will be captured in the residual  $\epsilon_{mkt}$  and treated as beliefs
- Investors may not actively re-balance their portfolios
  - ▷ Autocorrelation in holdings is 0.8-0.9
  - ▷ Robustness: use newly introduced 401(k) plans (e.g., active choice)
  - ▷ Focus on the cross section of beliefs
- Investors may not respond to fees
  - ▷ Insensitivity to fees translates to high risk aversion in the model
  - ▷ We find that investors do respond to fees

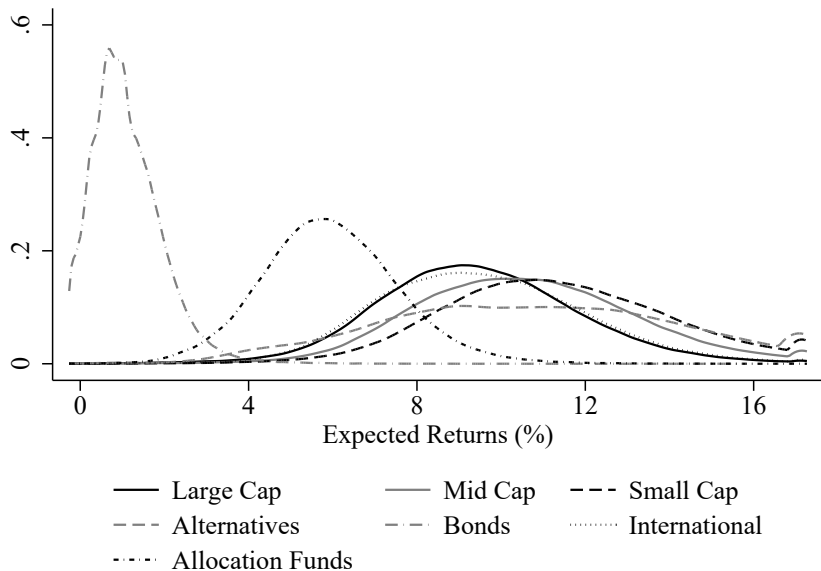
# Outline of Talk

1. Data
2. Model
3. **Estimates of Beliefs and Risk Preferences**
4. Counterfactual Allocations
5. Evidence on the Formation of Beliefs

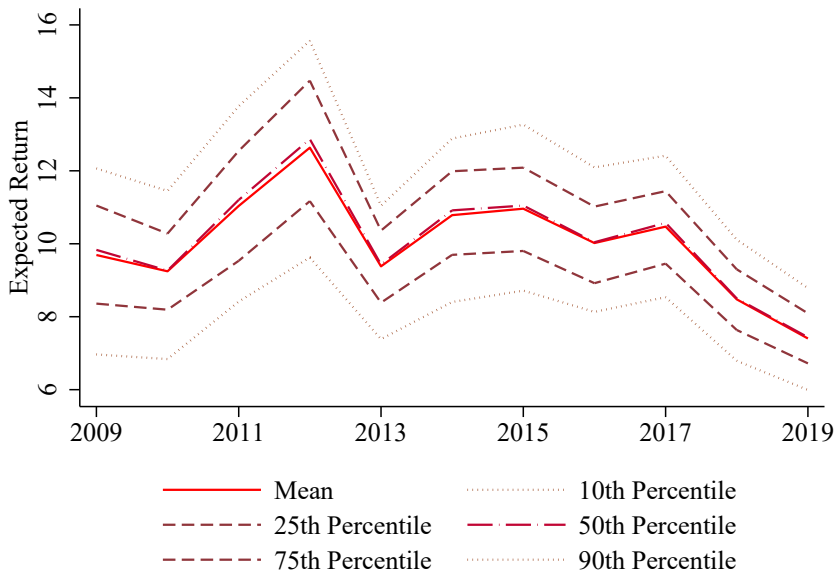
# Risk Aversion Over Time



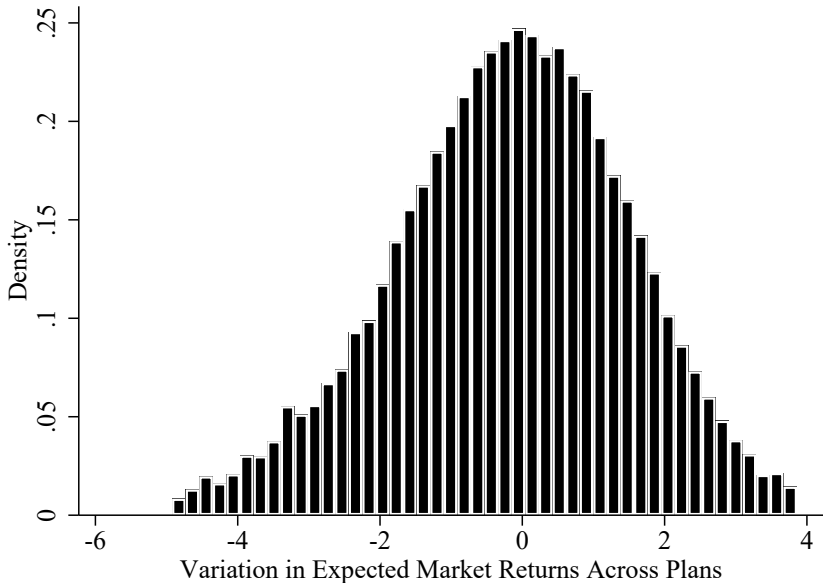
# Heterogeneity in Beliefs Across Asset Classes



# Stock Market Expectations Over Time

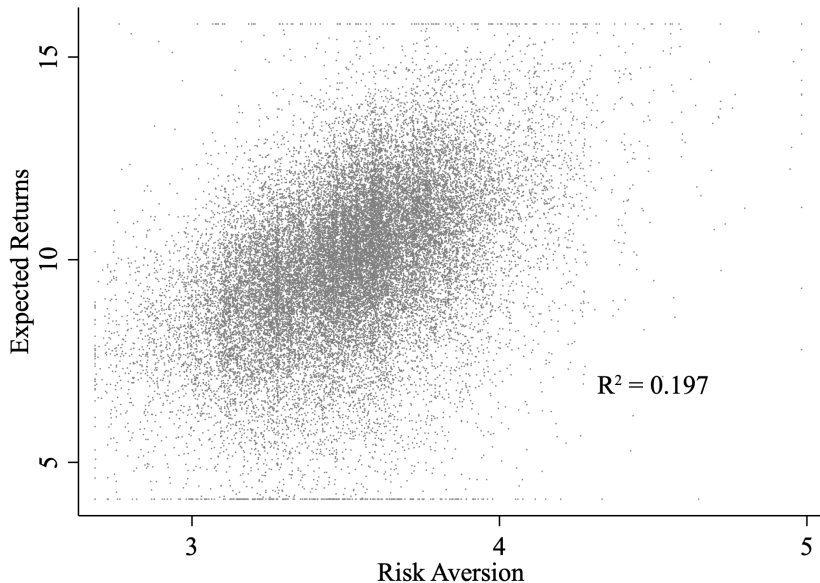


# Heterogeneity in Stock Market Expectations

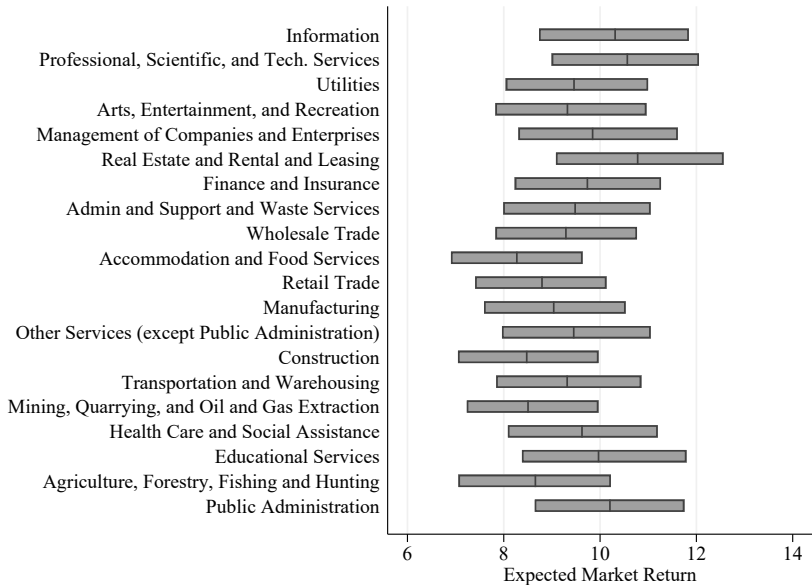




# Stock Market Expectations vs. Risk Aversion



# Beliefs by Sector of Employment



## What Explains Holdings? Beliefs vs. Risk Aversion

$$Equity\ Share_{mt} = \gamma Risk\ Aversion_{mt} + \phi Expected\ Returns_{mt}$$

	(1)	(2)
Risk Aversion (Std.)	-6.511*** (0.192)	-7.449*** (0.153)
Expected Returns (Std.)	9.974*** (0.367)	13.692*** (0.245)
Observations	243,268	243,268
R-squared	0.507	0.788
Year FE		X

# Outline of Talk

1. Data
2. Model
3. Estimates of Beliefs and Risk Preferences
4. **Counterfactual Allocations**
5. Evidence on the Formation of Beliefs

# Importance of Beliefs and Risk Aversion

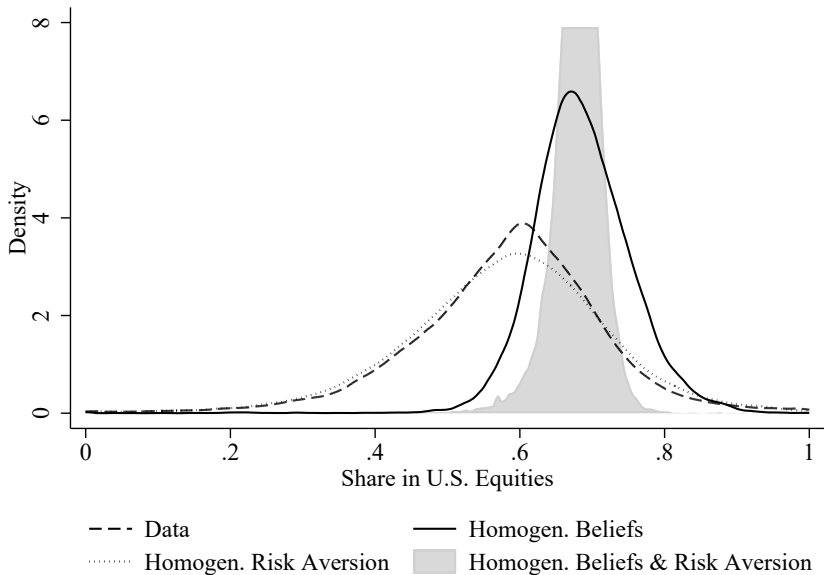
We consider counterfactual allocations where investors:

- Have identical beliefs
- Have identical risk preferences
- Have identical beliefs and risk preferences

We use the model to simulate chosen allocations

- Use the mean risk aversion parameter and the expected returns that would rationalize the aggregate portfolio

# Counterfactual Allocations



# Outline of Talk

1. Data
2. Model
3. Estimates of Beliefs and Risk Preferences
4. Counterfactual Allocations
5. **Evidence on the Formation of Beliefs**

# Understanding Beliefs and Risk Aversion

What explains the cross-sectional variation in beliefs and risk aversion?

How are beliefs formed?

- Are they extrapolative?
- Are they rational?



## Cross Section of Risk Aversion

	(1)
<b>Baseline Risk Aversion</b>	<b>3.558***</b>
× <b>Age</b>	<b>0.217**</b>
× Frac Black	0.061
× Frac Hispanic	-0.053
× <b>College</b>	<b>0.314**</b>
× <b>ln(Income)</b>	<b>-0.261*</b>
× ln(Home Value)	0.109
× Employed	-0.068
× Unionized	0.203
× Share Retired	-0.040
× ln(Avg. Acct. Balance)	-0.008
× <b>Existing 401(k) Plan</b>	<b>0.885**</b>
Observations	4,528,147

All independent variables are in units of standard deviations

## Cross Section of Beliefs About Market Returns

	(1)
<b>Age</b>	<b>-0.200***</b>
ln(Income)	0.076
ln(Home Value)	-0.002
College	0.114
Employed	0.021
<b>Black</b>	<b>-0.101***</b>
<b>Hispanic</b>	<b>-0.083**</b>
<b>Unionized</b>	<b>-0.412***</b>
<b>Sector Equity Beta</b>	<b>0.022***</b>
<b>Share Retired</b>	<b>-0.110***</b>
ln(Avg. Acct. Bal.)	0.076
Observations	243,268

All independent variables are in units of standard deviations

# Extrapolating Returns

Well documented that investors extrapolate from past returns

We look at two different types of extrapolation to provide more insight into how investors form beliefs

1. Does fund experience matter for extrapolation?
  - ▷ Do investors extrapolate about fund returns they never experienced?
  - ▷ Look at beliefs about new funds added to a plan
2. Do investors extrapolate from their personal experience?
  - ▷ Local economic conditions
  - ▷ Relationship between firm and market performance

## Extrapolating from Fund Returns

VARIABLES	(1)	(2)	(3)	(4)
Lag Fund Ret.	0.005*** (0.000)	0.016*** (0.001)	0.007*** (0.002)	0.005*** (0.000)
Lag Fund Ret. x New Investment				-0.000 (0.000)
Observations	4,499,736	672,910	79,041	4,499,736
R-squared	0.937	0.941	0.940	0.937
FE	X	X	X	X
New Funds		X		
New Plans			X	

Each specification includes Plan×Year and Morningstar Category×Passive×Year fixed effects

## Extrapolating from Local Economic Conditions

$$MarketBeliefs_{mt} = \Gamma EconomicConditions_{mt} + \varepsilon_{mt}$$

	(1)	(2)	(3)	(4)	(5)
Pop. Growth	0.125*** (0.016)				0.032** (0.013)
Home Price Growth		0.022*** (0.006)			0.003 (0.002)
Establishment Growth			0.039*** (0.011)		0.006 (0.004)
GDP Growth				0.036*** (0.004)	0.003** (0.002)
Observations	232,877	239,199	243,268	239,313	217,483
R-squared	0.357	0.344	0.343	0.344	0.872
Year FE	X	X	X	X	X
Plan FE					X

## Extrapolating from Firm Performance

$$MarketBeliefs_{mt} = \varphi FirmPerformance_t + \eta_{mt}$$

	(1)	(2)	(3)	(4)	(5)	(6)
Firm Return (1 years)	0.001*** (0.000)				0.001** (0.000)	0.001*** (0.000)
Firm Investment		0.018*** (0.003)			0.005** (0.002)	0.016*** (0.004)
Sales Growth			0.004*** (0.001)		-0.000 (0.000)	0.001** (0.001)
Employment Growth				0.005*** (0.001)	0.000 (0.001)	0.002* (0.001)
Observations	11,738	10,474	11,452	11,441	9,889	10,081
R-squared	0.510	0.521	0.510	0.510	0.890	0.519
Year FE					X	
Plan FE					X	
NAICS×Year FE	X	X	X	X		X

## Conclusion

We develop an empirical approach to nonparametrically identify beliefs and risk preferences from allocation data

- Model of investor portfolio problem
- Allow for arbitrary variation/correlation across plans

## Conclusion

We develop an empirical approach to nonparametrically identify beliefs and risk preferences from allocation data

- Model of investor portfolio problem
- Allow for arbitrary variation/correlation across plans

We explore portfolio allocations in 401(k) plans

- Substantial heterogeneity in asset allocations
- ⇒ Heterogeneity in beliefs and risk preferences



## Conclusion

We develop an empirical approach to nonparametrically identify beliefs and risk preferences from allocation data

- Model of investor portfolio problem
- Allow for arbitrary variation/correlation across plans

We explore portfolio allocations in 401(k) plans

- Substantial heterogeneity in asset allocations

⇒ Heterogeneity in beliefs and risk preferences

What drives variation in beliefs and preferences?

- Evidence in the cross section and over time:
  - ▷ Educated investors are more risk averse and more optimistic
  - ▷ Older investors tend to be more risk averse and pessimistic
  - ▷ Extrapolation from past investment performance, past employer performance, and local economic conditions

## Conclusion

We develop an empirical approach to nonparametrically identify beliefs and risk preferences from allocation data

- Model of investor portfolio problem
- Allow for arbitrary variation/correlation across plans

We explore portfolio allocations in 401(k) plans

- Substantial heterogeneity in asset allocations

⇒ Heterogeneity in beliefs and risk preferences

What drives variation in beliefs and preferences?

- Evidence in the cross section and over time:
  - ▷ Educated investors are more risk averse and more optimistic
  - ▷ Older investors tend to be more risk averse and pessimistic
  - ▷ Extrapolation from past investment performance, past employer performance, and local economic conditions

→ Information obtained from local environments/networks