Mutual Fund Families, Market Structure, and Fee Competition

Richard Grice
NUS

Ahmed Guecioueur
UW Foster

Four Corners
7 May 2026

Individual mutual funds belong to families

Individual mutual funds belong to families: why?

- One possibility: resource-sharing and economies of scope and scale
- Yet asset managers *incur* costs to maintain distinct identities; e.g.:
  - Advertising, even when it does not have a direct effect on flows (Gallaher, Kaniel, and Starks 2015)
  - Maintaining multiple brands (Franklin Resources, Natixis, ...)

This paper

- We argue that family identities shape and direct investors' consideration of various alternatives as they choose what to invest in
- Theoretically, we show such behavior provides individual funds with a (quantifiable) degree of pricing power in setting their fees
- Empirically, the mechanism predicts price dispersion among index mutual funds

Individual mutual funds belong to families: why?

- One possibility: resource-sharing and economies of scope and scale
- Yet asset managers *incur* costs to maintain distinct identities; e.g.:
  - Advertising, even when it does not have a direct effect on flows (Gallaher, Kaniel, and Starks 2015)
  - Maintaining multiple brands (Franklin Resources, Natixis, ...)

This paper

- We argue that family identities shape and direct investors' consideration of various alternatives as they choose what to invest in
- Theoretically, we show such behavior provides individual funds with a (quantifiable) degree of pricing power in setting their fees
- Empirically, the mechanism predicts price dispersion among index mutual funds

Contributions and findings

Theoretical

- We contribute a new oligopoly model with networked competition, admitting a closed-form Nash equilibrium for fees
- The model formalizes how any pattern of family consideration translates into a market structure...
- ... Thus providing individual funds with some fee-setting power

Empirical

- Investors consider a limited and heterogeneous subset of families
- Calibrating on consideration sets based on EDGAR usage, our model's fee predictions match actual S&P 500 index fund fees (1:1)
  - Supports our hypothesis that family identities influence fund fees
- Counterfactuals: observed fees are $\sim 1/3$ of monopoly but $\sim 3\times$ competitive benchmarks

Model of fee competition: investors / demand-side

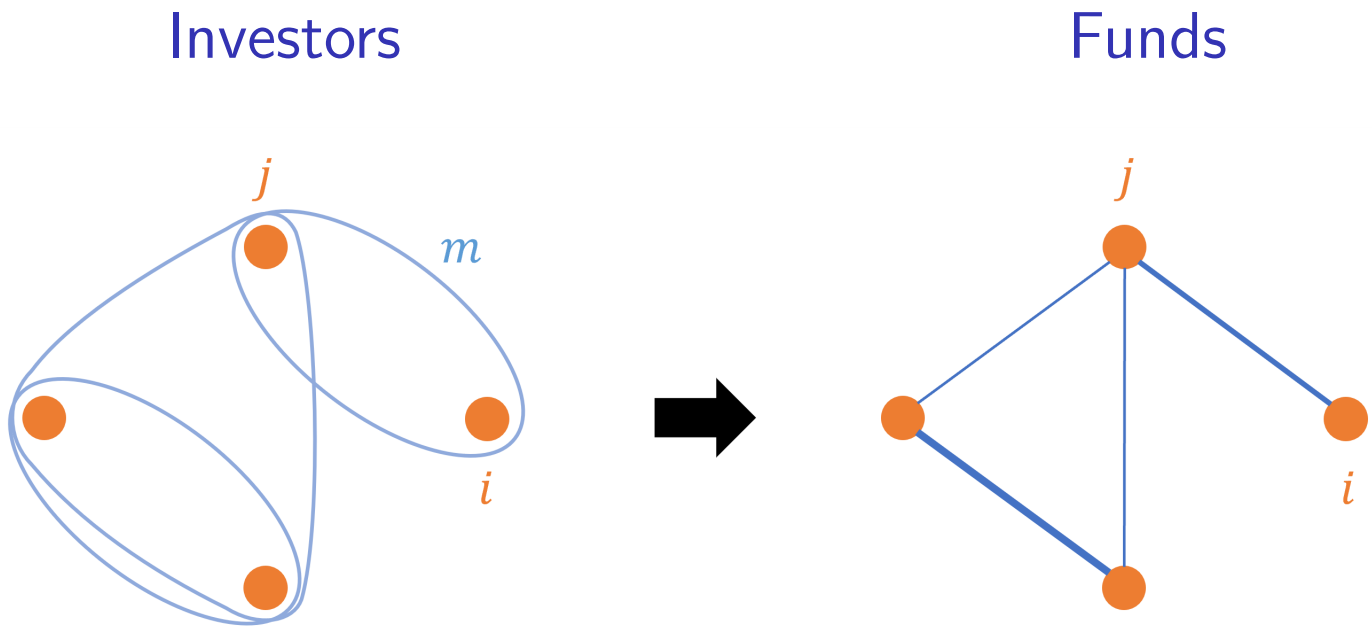
Single category of investment products

- Funds set fees and compete within a uniform product category
  - e.g. S&P 500 index mutual funds
- Each fund belongs to a family

Investors consider families, then allocate over funds

- Each investor considers funds from a subset of families
  - We stay agnostic about what determines an investor's consideration set (awareness, preferences, behavioral biases, ...)
  - Allow for heterogeneity in investors and their consideration sets
- Funds in considered families are imperfect substitutes for the investor's allocation problem
- Investors engage in mean-variance-like portfolio choice over considered funds & risk-free outside option, à la Merton (1987)

Consideration by investors ⇒ a Competition Network among funds



- Aggregating across investors' consideration sets yields a competition network between funds, with adjacency matrix $\boldsymbol{E}$
- $\boldsymbol{E}$ captures the competitive market structure of the industry
- Crucially, this works for any pattern of family consideration that investors exhibit; we make no assumptions about market structure

Model of fee competition: funds / supply-side

Funds set fees strategically

- Assume that mandates & family positioning are fixed over the short-run $\Rightarrow$ fees are the sole strategic choice
- As funds set fees, they internalize the fact that family identities shape competition (our hypothesized channel)
- Funds compete with one another when setting fees, but the competition may be imperfect, depending on $\boldsymbol{E}$

Equilibrium fees are a network centrality

- Fund fee-setting is a quadratic network game and we derive the unique closed-form Nash equilibrium $\boldsymbol{f}^*(\boldsymbol{E})$
  - (Jackson and Zenou 2015; Bramoullé and Kranton 2016)
- We show that each fund's fee is a weighted sum of all funds' performance (return moments), and is a form of centrality in $\boldsymbol{E}$
- Model-implied fees $\boldsymbol{f}^*$ capture each family's position in the competition network, not its attributes *per se*

Main idea

Hypothesis:

Fund family identities direct investors' consideration of alternatives, providing the offered funds with some pricing power when setting fees.

Main idea

Hypothesis:

Fund family identities direct investors' consideration of alternatives, providing the offered funds with some pricing power when setting fees.

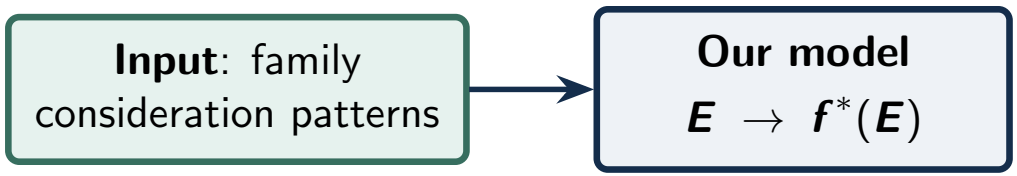
Our model

$$E \rightarrow f^*(E)$$

Main idea

Hypothesis:

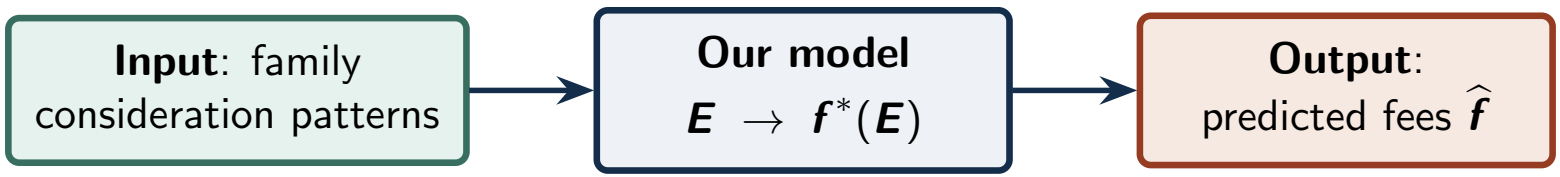
Fund family identities direct investors' consideration of alternatives, providing the offered funds with some pricing power when setting fees.



Main idea

Hypothesis:

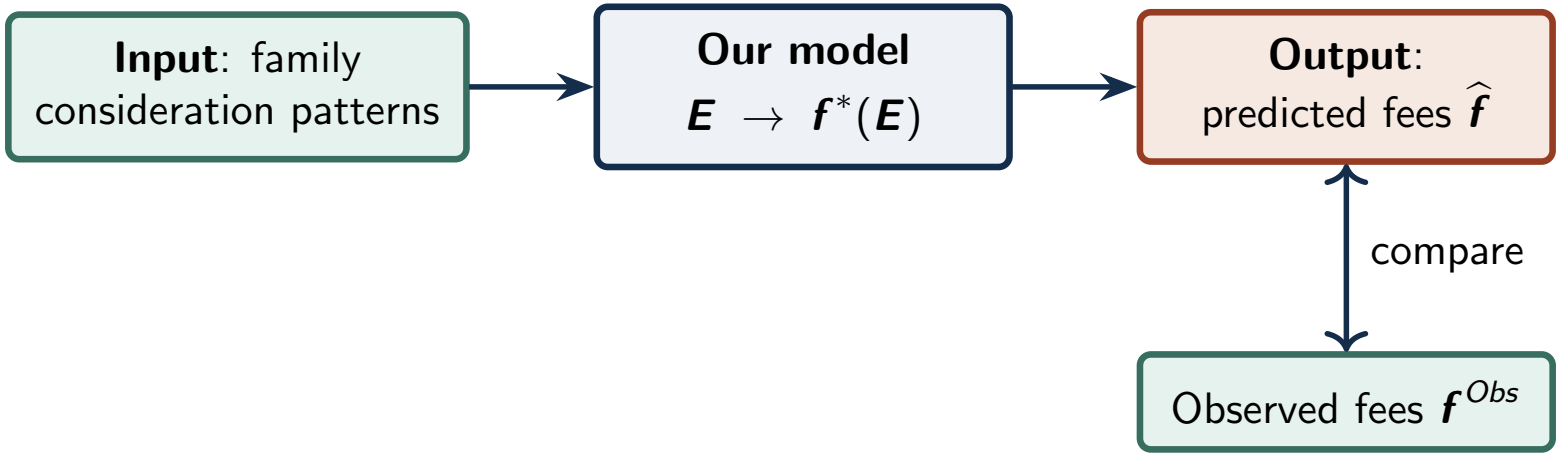
Fund family identities direct investors' consideration of alternatives, providing the offered funds with some pricing power when setting fees.



Main idea

Hypothesis:

Fund family identities direct investors' consideration of alternatives, providing the offered funds with some pricing power when setting fees.



Test:

When we calibrate our model to observed family consideration patterns (input), it should predict fees $\widehat{f}$ (output) that track observed fees f^{obs} in a homogeneous product category.

Data

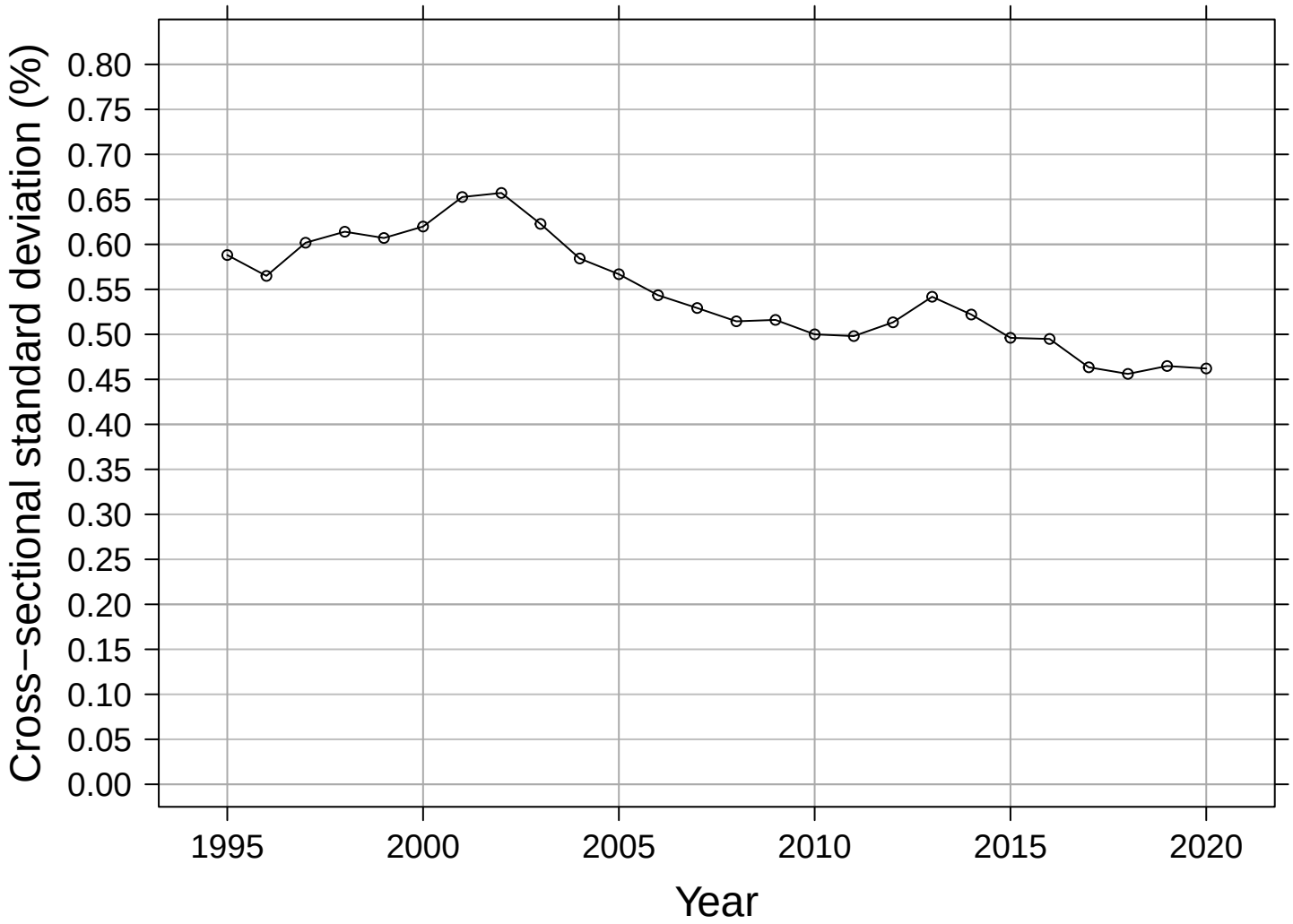
Fees & characteristics: CRSP Mutual Fund Database

- Focus on the homogeneous category of S&P 500 index funds, matching our model's setup
- Use non-institutional share classes, where fees are not set bilaterally

Consideration sets: SEC EDGAR usage logs (2006–2016)

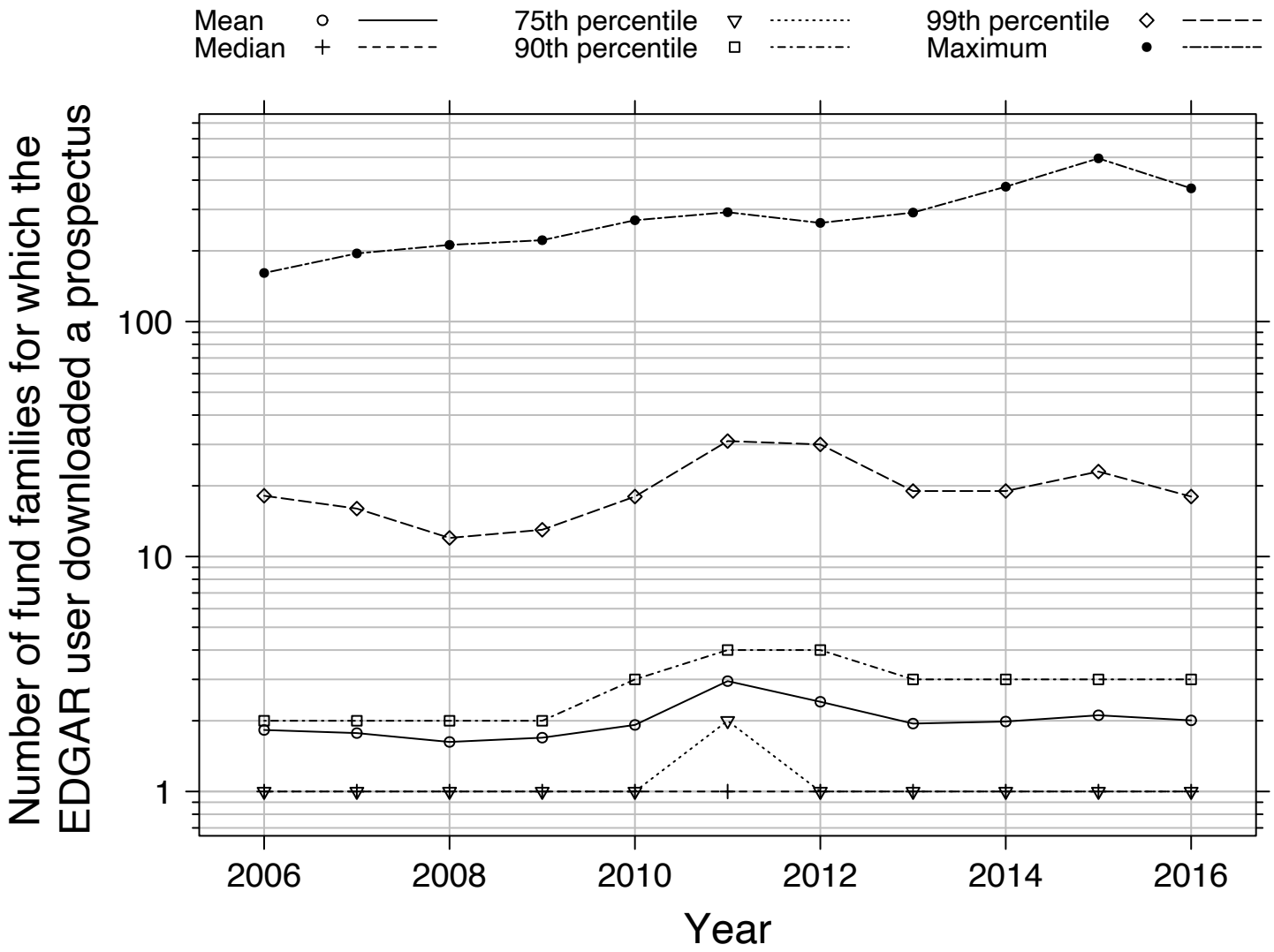
- A user (non-robot IP address) considers family i in year t if she downloaded any of its prospectus filings that year
- EDGAR is the only comprehensive (and free) source of prospectuses
  - Used even by non-professionals (Liaukonytė and Žaldokas 2022)
  - Users by state correlate with overall populations ($\rho \approx 0.89$)

Fee dispersion of observed S&P 500 index mutual funds



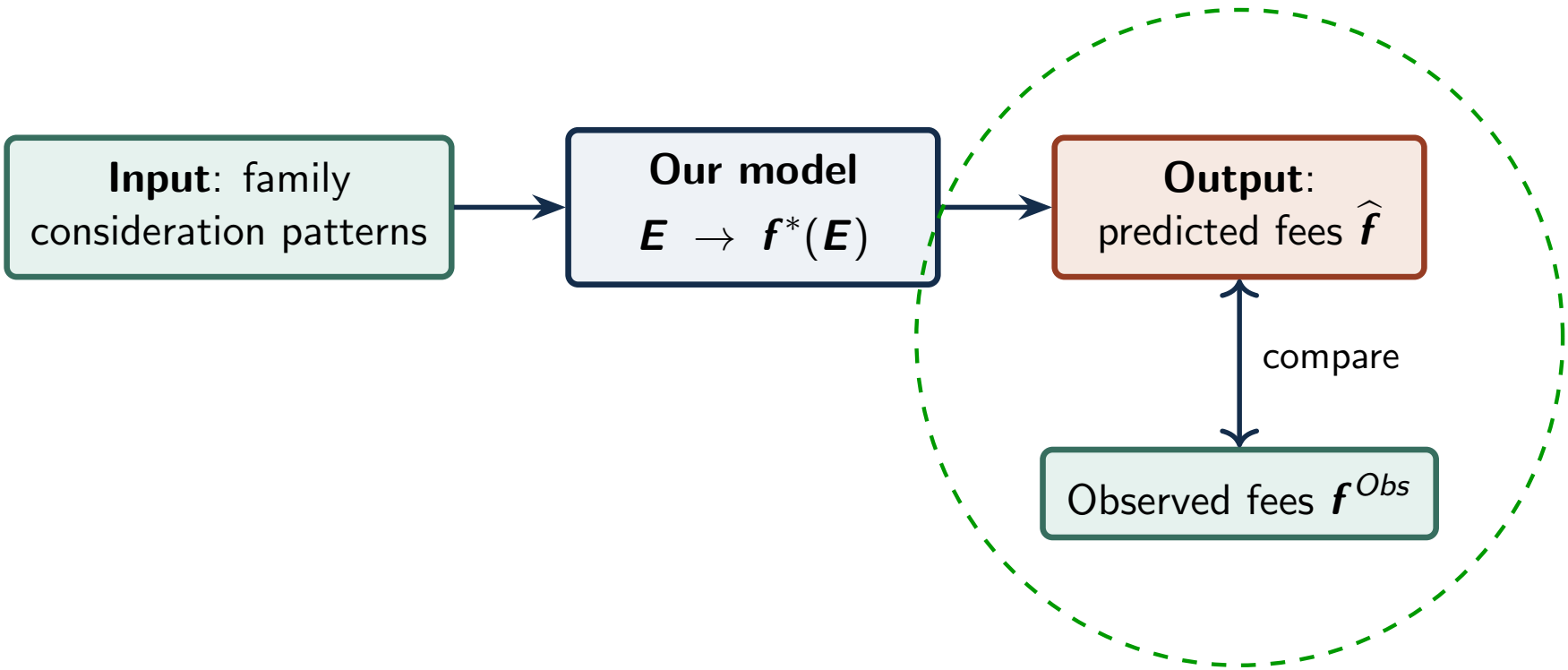
Annual cross-sectional standard deviation of S&P 500 index mutual fund total fees, expressed as an annualized percentage.

Family consideration sets from EDGAR: cardinality



Distribution of the number of fund families in each potential investor's consideration set, measured from her annual SEC EDGAR prospectus downloads. Y-axis: number of fund families (log scale).

Recap


Input: family consideration patterns
Our model
$\boldsymbol{E} \rightarrow \boldsymbol{f}^*(\boldsymbol{E})$
Output: predicted fees $\widehat{\boldsymbol{f}}$
compare
Observed fees $\boldsymbol{f}^{Obs}$

Model-implied fees $\widehat{f_{it}}$ correlate 1:1 with observed fees f_{it}^{obs}

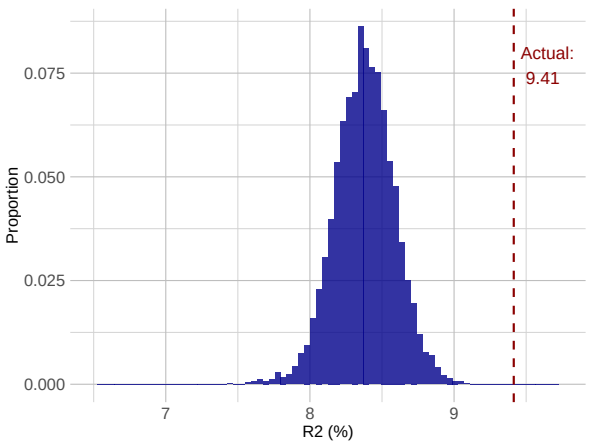
Dependent Variable:	f_{it}^{obs}		
	(1)	(2)	(3)
$\widehat{f_{it}}$	1.168**	1.283***	1.270***
	(0.3949)	(0.3539)	(0.3301)
Log Size$_{it}$		0.0002	0.0005
		(0.0003)	(0.0003)
Log Downloaders$_{it}$		-0.0007	-0.0002
		(0.0004)	(0.0006)
Log Family Size$_{it}$			-0.0009*
			(0.0005)
Family Equity Share$_{it}$			-0.00003
			(0.00002)
Family Passive Share$_{it}$			-0.0005
			(0.00003)
Family Fund Count$_{it}$			1.16×10^{-6}
			(8.52×10^{-6})
Family Asset Class Diversity$_{it}$			-3.27×10^{-7}
			(3.29×10^{-7})
Family Return Diversity$_{it}$			1.21×10^{-8}
			(3.87×10^{-8})
Year FEs	✓	✓	✓
N	693	693	475
R^2	0.07200	0.09414	0.14886

Double-clustered (Family & Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

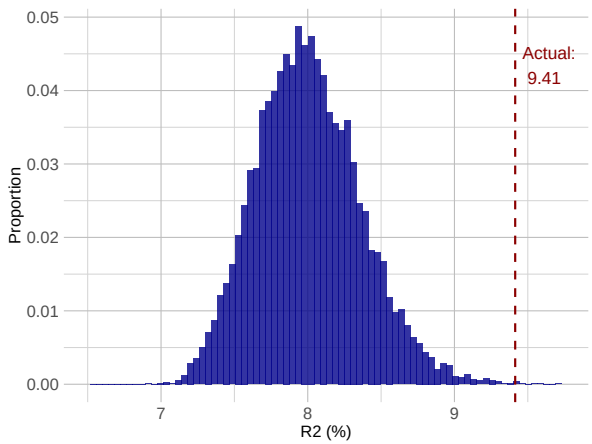
Permutation tests: consideration structure is informative

- Concern: EDGAR-derived consideration sets might be too noisy
- We run three permutation tests, each randomizing over one feature of the consideration data and re-running results
- If patterns were mere noise, baseline R^2 would sit in the *bulk* of each permutation distribution

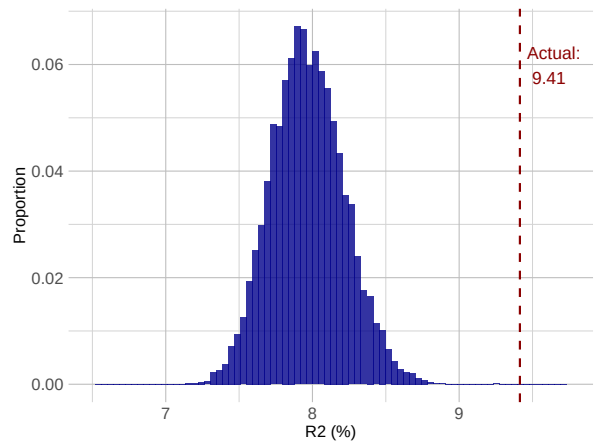
Re-drawing families per investor (p=0.00001)



Shuffling family identities (p=0.0007)



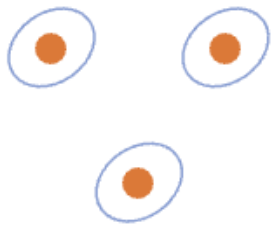
Re-drawing investors per family (p=0)



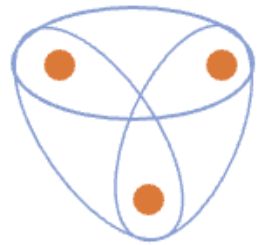
- Reassuringly, baseline R^2 lies in the extreme right tail (stat. sigf.)

Counterfactual market structures that E can represent

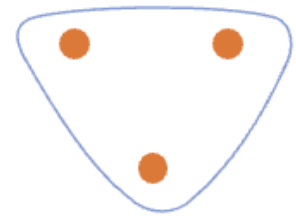
(a) Monopoly



(b) Uniform

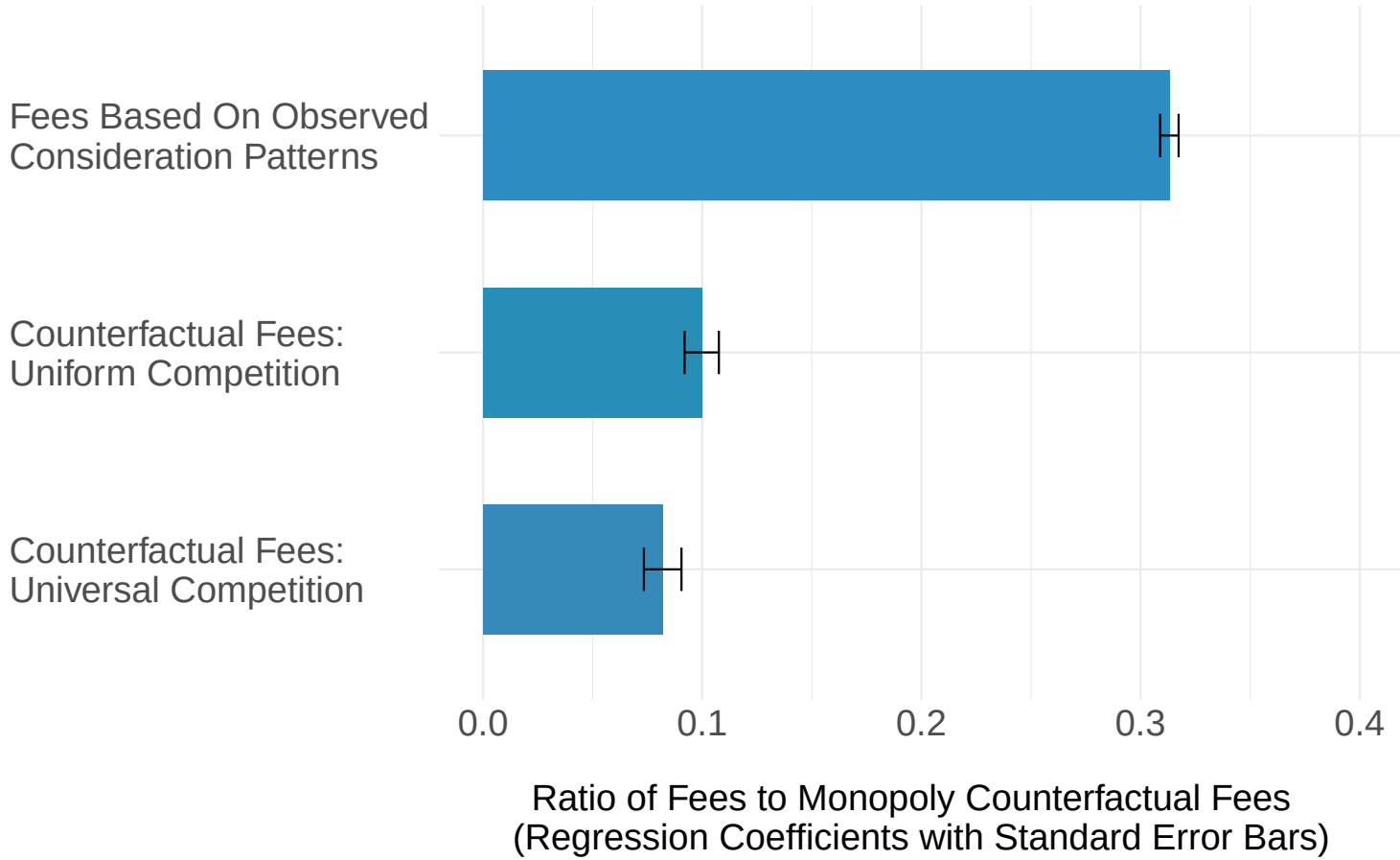


(c) Universal



- Monopoly is the absence of competition: all investors are "captured" and consider a single family only
- Uniform and Universal both model a high degree of competition

Different calibrations show that fees are 1/3 the monopoly level, but 3x higher than competitive benchmarks



Similar families are co-considered more frequently (Poisson regr.)

Dependent Variable:	Co-considerations$_{i,j,t}$		
	(1)	(2)	(3)
Total Fee Distance$_{i,j,t}$	-0.0774***	-0.0864***	-0.0609***
	(0.0243)	(0.0200)	(0.0161)
Age Distance$_{i,j,t}$	-0.2671***	-0.1650***	-0.1587***
	(0.0386)	(0.0179)	(0.0177)
Size Distance$_{i,j,t}$	-0.3106***	-0.2054***	-0.2091***
	(0.0571)	(0.0311)	(0.0267)
Fund Count Distance$_{i,j,t}$	-0.5880***	-0.5524***	-0.3953***
	(0.0534)	(0.0397)	(0.0389)
Equity Share Distance$_{i,j,t}$	-0.1226***	-0.0212	-0.0070
	(0.0417)	(0.0291)	(0.0117)
Fixed Income Share Distance$_{i,j,t}$	0.0596	0.0244	-0.0126
	(0.0418)	(0.0293)	(0.0189)
Mixed Assets Share Distance$_{i,j,t}$	-0.0824	-0.1248**	-0.0831**
	(0.0587)	(0.0543)	(0.0359)
Retail Share Distance$_{i,j,t}$	-0.1040***	-0.0724***	-0.0973***
	(0.0217)	(0.0156)	(0.0123)
Passive Share Distance$_{i,j,t}$	-0.4789***	-0.6423***	-0.6691***
	(0.1450)	(0.1140)	(0.1061)
Family i FEs	✓	✓	
Family j FEs	✓	✓	
Year t FEs		✓	
Family i × Year t FEs			✓
Family j × Year t FEs			✓
N	702,472	702,472	702,472
R^2	0.71280	0.81200	0.93206

Clustered (Family i & Family j & Year t) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Sophisticated investors tend to consider more alternatives

- EDGAR prospectus downloaders' IP addresses are decrypted (Chen, Cohen, Gurun, Lou, and Malloy 2020), geolocated, and matched to demographic attributes, some of which proxy for sophistication

Dependent Variable:	Consideration set size (mean)	
	(1)	(2)
Age (median)	-0.0240**	-0.0284**
	(0.0103)	(0.0126)
Household income (median)	0.0611**	0.0773**
	(0.0234)	(0.0249)
Home value (median)	-0.0205	-0.0172
	(0.0248)	(0.0230)
College graduates (share)	0.0854***	0.0711**
	(0.0196)	(0.0241)
Urban residents (share)	0.0552***	0.0423***
	(0.0128)	(0.0122)
Democrat voters (share)	0.1650***	0.1514***
	(0.0383)	(0.0309)
(Intercept)	1.560***	
	(0.0530)	
Year FEs		✓
N	18,209	18,209
R^2	0.01952	0.02876

Clustered (Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Conclusion

- Family identities are economically meaningful: they shape investor consideration, providing offered funds with a quantifiable degree of pricing power

- We measure investor consideration with a novel data source (EDGAR usage data), staying agnostic about what drives it

- We introduce a tractable framework for strategic networked price competition with consideration as a primitive, applicable to other financial settings beyond mutual funds

Thank you!

References

Bramoullé, Yann, and Rachel Kranton. 2016. "Games Played on Networks." In *Oxford Handbook of the Economics of Networks,* 83–112.

Chen, Huaizhi, Lauren Cohen, Umit Gurun, Dong Lou, and Christopher Malloy. 2020. "IQ from IP: Simplifying search in portfolio choice." *Journal of Financial Economics* 138 (1): 118–137.

Gallaher, Steven, Ron Kaniel, and Laura T Starks. 2015. "Advertising and Mutual Funds: From Families to Individual Funds." *CEPR Discussion Paper,* no. 10329.

Jackson, Matthew O, and Yves Zenou. 2015. "Games on networks." In *Handbook of Game Theory with Economic Applications,* 4:95–163.

Liaukonytė, Jūra, and Alminas Žaldokas. 2022. "Background noise? TV advertising affects real-time investor behavior." *Management Science* 68 (4): 2465–2484.

Merton, Robert. 1987. "A Simple Model of Capital Market Equilibrium with Incomplete Information." *Journal of Finance* 42 (3): 483–510.

Appendix:
Key Theoretical Results

Notation: primitives

Funds and investor consideration

- $\mathcal{N} = \{1, \ldots, N\}$ – families in the category (each offers one fund i)
- $\mathcal{N}_m \subseteq \mathcal{N}$ – consideration set common to investor segment m
- θ_m – mass of segment m, with $\sum_{m \in \mathcal{M}} \theta_m = 1$
- $\mathcal{P} \triangleq \{(\mathcal{N}_m, \theta_m)\}_{m \in \mathcal{M}}$ – pattern of consideration

Returns, fees, and beliefs

- $\eta_i \triangleq \mathbb{E}[r_i]$, $\sigma_i^2 \triangleq \mathbb{V}\mathrm{ar}(r_i)$ – mean & variance of fund i's return
- $\rho \in (0, 1)$ – common pairwise correlation: $\mathrm{Cov}(r_i, r_j) = \rho\, \sigma_i \sigma_j$
- r_0 – risk-free rate
- f_i – fee charged by fund i
- $R_i \triangleq (r_i - f_i)/\sigma_i$ – standardized ex-fee return; what investors compare

Notation: derived quantities

Competition network

- $\boldsymbol{E}$ – $N \times N$ adjacency matrix encoding $\mathcal{P}$ and beliefs:

$$e_{ij} \triangleq \sum_{m \in \mathcal{M}} \mathbb{1}\{i \in \mathcal{N}_m\}\, \mathbb{1}\{j \in \mathcal{N}_m\}\, \theta_m\, h_{mij}$$

- h_{mij}: mean-variance substitution coefficients (depend only on $|\mathcal{N}_m|$ and ρ)
- $e_{ij} < 0$: i and j directly compete (some segment co-considers them)
- $e_{ij} = 0$: no direct edge – may still compete indirectly via shared neighbors

Equilibrium and empirical counterpart

- $\boldsymbol{f}^*(\boldsymbol{E})$ – Nash equilibrium fee vector; closed-form function of $\boldsymbol{E}$
- $\widehat{\boldsymbol{E}}$ – competition network calibrated from observed EDGAR consideration sets
- $\widehat{\boldsymbol{f}} \triangleq \boldsymbol{f}^*(\widehat{\boldsymbol{E}})$ – model-implied fees; compared to $\boldsymbol{f}^{obs}$ in the test

Objective functions

Representative investor of segment m

Allocation across considered funds $\mathcal{N}_m$ + risk-free outside option:

$$\max_{\boldsymbol{x}_m} \; (\boldsymbol{\mu}_m - r_0 \mathbf{1}_m)^T \boldsymbol{x}_m \; - \; \tfrac{1}{2} \boldsymbol{x}_m^T \boldsymbol{V}_m \boldsymbol{x}_m$$

- $\boldsymbol{x}_m$ – allocations across considered funds
- $\boldsymbol{\mu}_m$ – expected standardized ex-fee returns: $\mu_i = (\eta_i - f_i)/\sigma_i$
- $\boldsymbol{V}_m$ – covariance of standardized returns: 1 on diagonal, $\rho \in (0,1)$ off-diagonal $\Rightarrow$ funds within $\mathcal{N}_m$ are imperfect substitutes

Fund i

Sets its fee f_i to max. fee revenue, taking competitors' fees $\boldsymbol{f}_{-i}$ as given:

$$\max_{f_i} \; \pi_i(f_i, \boldsymbol{f}_{-i}) \; \triangleq \; f_i \cdot x_i(\boldsymbol{f})$$

- MC normalized to zero $\Rightarrow$ focus on demand-side competition
- AD $x_i(\boldsymbol{f}) = \sum_{m \in \mathcal{M}} \theta_m \, x^*_{mi}(\boldsymbol{f})$ inherits network structure
- Best replies are strategic complements along edges of $\boldsymbol{E}$

Equilibrium fees (for any market structure $\boldsymbol{E}$)

Assumption 1 (Strict diagonal dominance)

The returns of funds co-considered by investors are such that the adjacency matrix $\boldsymbol{E}$ is strictly (row-wise) diagonally dominant with weights $\frac{\sigma_i}{\eta_i - \sigma_i r_0}$. That is:

$$|e_{ii}| \frac{\eta_i - \sigma_i r_0}{\sigma_i} > \sum_{j \neq i} |e_{ij}| \frac{\eta_j - \sigma_j r_0}{\sigma_j} \tag{1}$$

for all $i \in \mathcal{N}$.

Equilibrium fees (for any market structure $\boldsymbol{E}$)

Proposition 1 (Nash equilibrium)

There exists a unique Nash equilibrium to the fee-setting game between funds characterized by non-zero fees:

$$\boldsymbol{f}^* = \left(\mathbf{I} + \operatorname{diag}(\widetilde{\boldsymbol{E}})^{-1}\widetilde{\boldsymbol{E}}\right)^{-1} \operatorname{diag}(\widetilde{\boldsymbol{E}})^{-1}\widetilde{\boldsymbol{E}}\left(\boldsymbol{\eta} - r_0 \operatorname{diag}(\boldsymbol{S})\mathbf{1}\right), \quad (2)$$

where $\boldsymbol{\eta}$ *is a vector with elements* $\eta_i \triangleq \mathbb{E}[r_i]$, $\operatorname{diag}(\boldsymbol{S})$ *is the* $N \times N$ *diagonal matrix with the elements* σ_i *on the leading-diagonal and* 0 *on the off-diagonals, and* $\widetilde{\boldsymbol{E}} \triangleq \boldsymbol{E}\operatorname{diag}(\boldsymbol{S})^{-1}$ *is the* $N \times N$ *matrix with the elements* $\widetilde{e}_{ij} \triangleq \frac{e_{ij}}{\sigma_j}$, *and thus* $\operatorname{diag}(\widetilde{\boldsymbol{E}}) = \operatorname{diag}(\boldsymbol{E})\operatorname{diag}(\boldsymbol{S})^{-1}$.

Equilibrium fees (for any market structure $\boldsymbol{E}$)

Corollary

We can re-express the result in Eqn. (2) in terms of the adjacency graph $\boldsymbol{E}$ explicitly:

$$\boldsymbol{f}^*(\boldsymbol{E}) = \left(\mathbf{I}+\boldsymbol{Q}\odot\mathrm{diag}(\boldsymbol{E})^{-1}\boldsymbol{E}\right)^{-1}\left(\boldsymbol{Q}\odot\mathrm{diag}(\boldsymbol{E})^{-1}\boldsymbol{E}\right)\left(\boldsymbol{\eta}-r_0\,\mathrm{diag}(\boldsymbol{S})\mathbf{1}\right).$$

In the above, $\odot$ connotes the Hadamard product (elementwise multiplication), and we define $\boldsymbol{Q}$ as the $N \times N$ matrix consisting of elements $q_{ij} \triangleq \frac{\sigma_i}{\sigma_j}$.

Equilibrium fees for a monopoly market structure

- In a monopoly, every investor is captured: $\boldsymbol{E} = \boldsymbol{I}$
- Plugging into $\boldsymbol{f}^*(\boldsymbol{E})$, the equilibrium fee vector collapses to:

$$\boldsymbol{f}^{\text{Monopoly}} = \boldsymbol{f}^*(\boldsymbol{I}) = \tfrac{1}{2}\big(\boldsymbol{\eta} - r_0 \operatorname{diag}(\boldsymbol{S})\,\mathbf{1}\big)$$

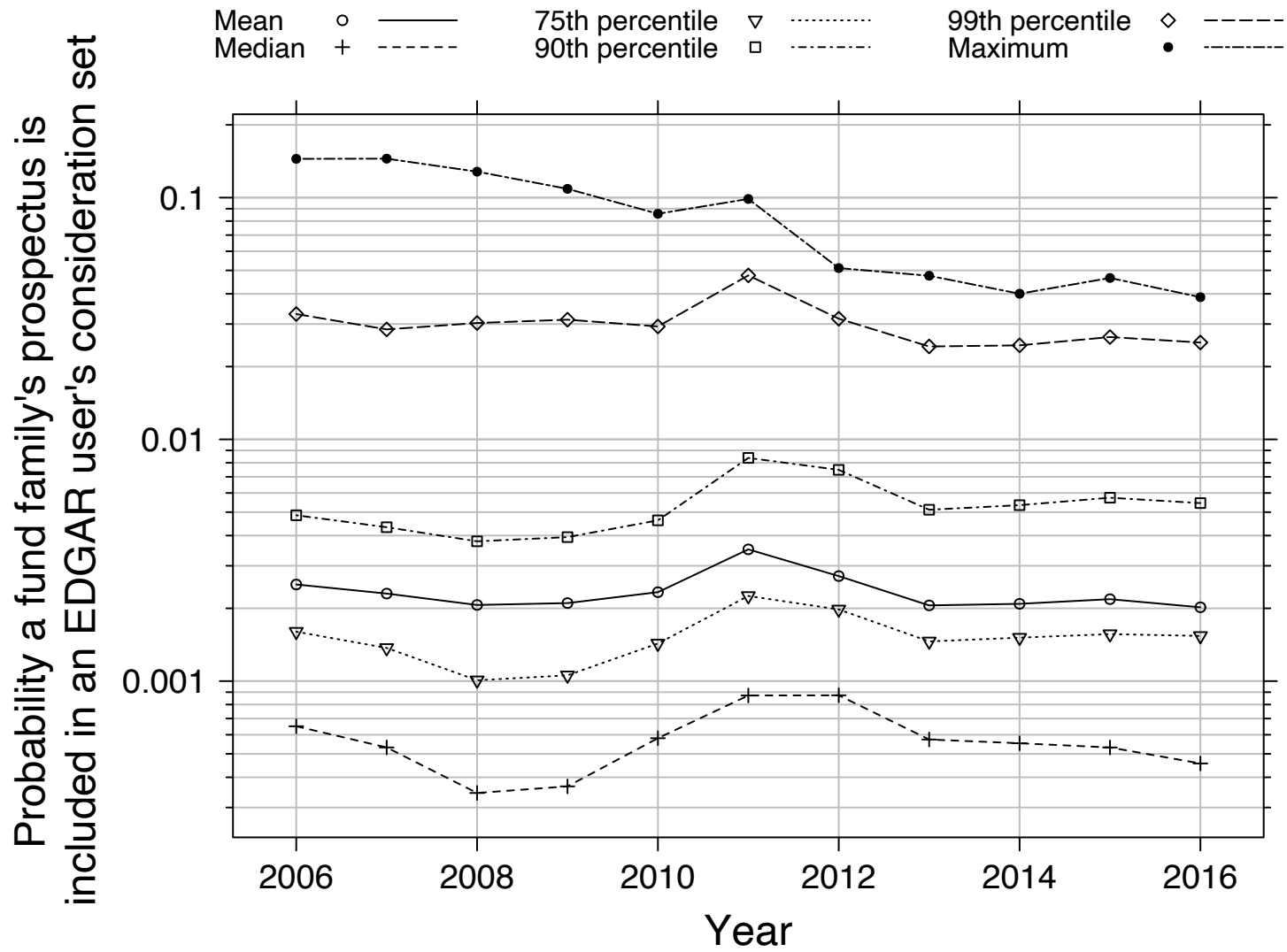
- Equivalently, in scalar form:

$$f_i^{\text{Monopoly}} = \tfrac{1}{2}\big(\eta_i - \sigma_i\, r_0\big)$$

$\Rightarrow$ fund i's fee depends only on its own expected return and risk, not on competitors'

Appendix:
Additional Empirical Results

Probability of inclusion in an investor's consideration set



Distribution of how frequently each fund family appears in any potential investor's consideration set, measured from annual SEC EDGAR prospectus downloads. Y-axis: prob. of consideration set membership, $\in [0, 1]$ (log scale).

Calibration results are similar when excluding singletons

- Columns 1–2 from baseline calibration; columns 3–4 exclude singletons

Dependent Variable:	f_{it}^{obs}			
Singletons:	Included		Excluded	
	(1)	(2)	(3)	(4)
$\widehat{f}_{it}$	1.168**	1.283***	1.260***	1.347***
	(0.3949)	(0.3539)	(0.3726)	(0.3340)
Log Size$_{it}$		0.0002		0.0002
		(0.0003)		(0.0003)
Log Downloaders$_{it}$		-0.0007		-0.0007
		(0.0004)		(0.0004)
Year FEs	✓	✓	✓	✓
N	693	693	693	693
R^2	0.07200	0.09414	0.07660	0.09778

Double-clustered (Family & Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Model fits best when families follow a direct distribution strategy

- Residual Expense Ratio$_{i,t}$ = Expense Ratio$_{i,t}$ − Actual 12b-1 Fee$_{i,t}$
- Includes an indicator for whether the family pays no distribution fees (thus following a purely direct distribution channel)

Dependent Variable:	Observed Residual Expense Ratio$_{i,t}$		
	(1)	(2)	(3)
$\widehat{f_{it}}$	0.8363***	0.6218**	0.6110**
	(0.2347)	(0.2486)	(0.2556)
Log Size$_{it}$	0.00004	0.00004	0.00004
	(0.0001)	(0.0002)	(0.0002)
Log Downloaders$_{it}$	-0.0008**	-0.0009**	-0.0009***
	(0.0003)	(0.0003)	(0.0003)
$\mathbb{1}\{\text{Purely Direct Distribution}\}_{i,t}$		-0.0018**	-0.0018**
		(0.0008)	(0.0008)
$\widehat{f_{it}} \times \mathbb{1}\{\text{Purely Direct Distribution}\}_{i,t}$		0.2998***	0.3484***
		(0.0096)	(0.0134)
Distribution Fee Definition		Actual 12b-1	Effective 12b-1
Year FEs	✓	✓	✓
N	693	693	693
R^2	0.11660	0.15026	0.15120

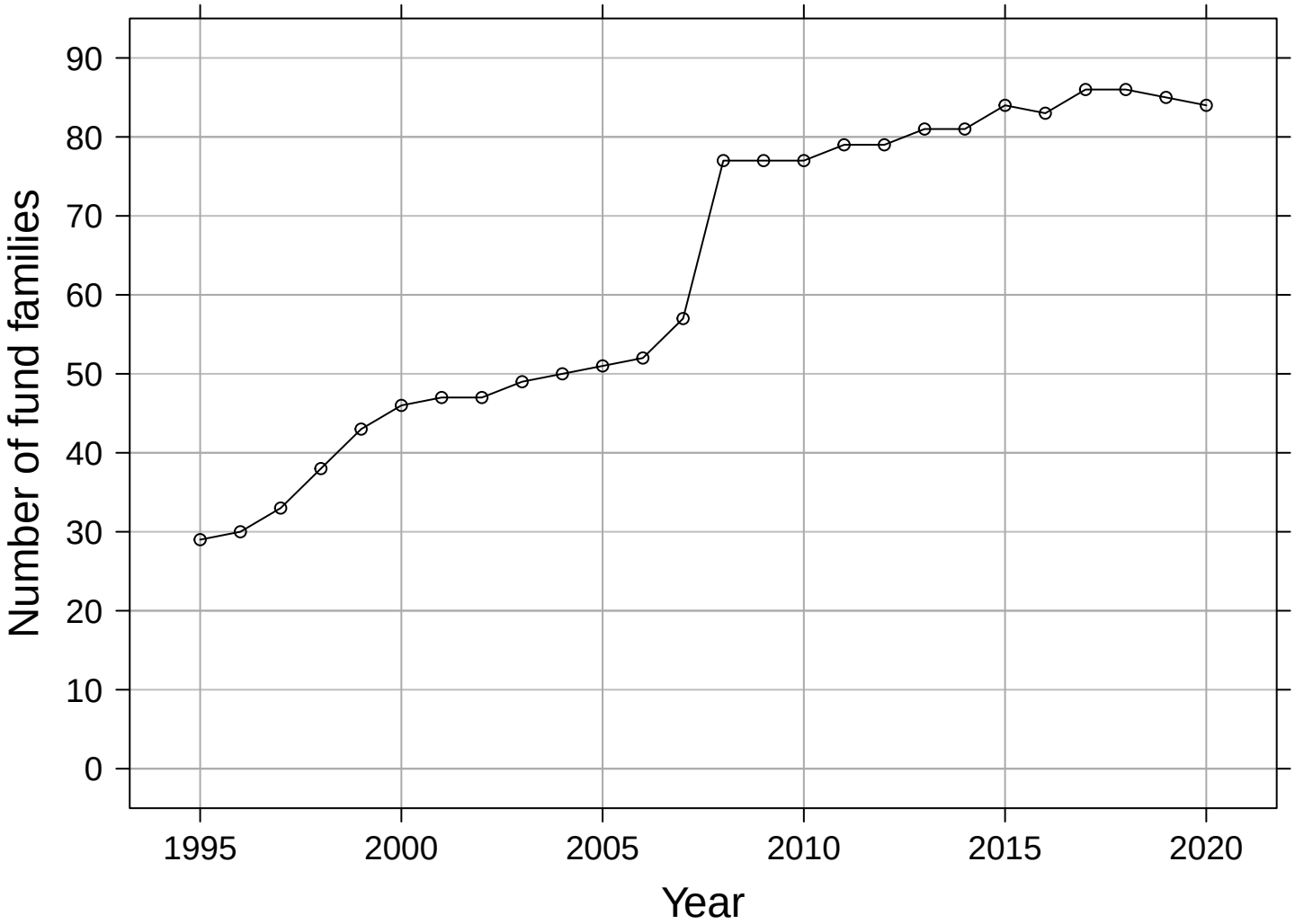
Double-clustered (Family & Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Appendix:
Summary Stats

S&P 500 fund and family characteristics

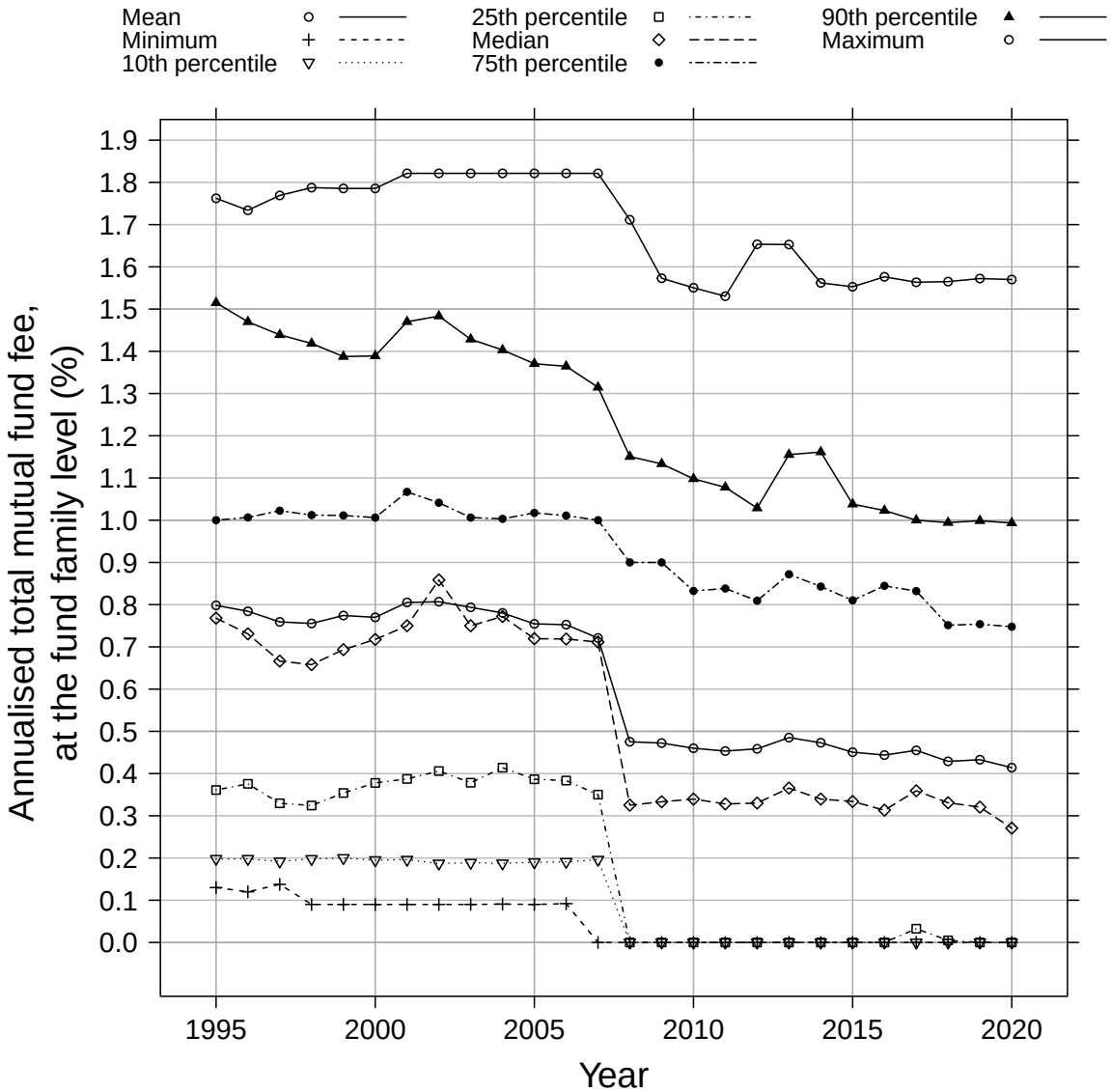
Statistic	N	Mean	S.D.	Lower quartile	Median	Upper quartile
Total Fee (%)	693	0.47	0.51	0.01	0.26	0.91
Expense Ratio (%)	693	0.37	0.39	0.01	0.21	0.60
Effective 12b-1 Fee (%)	693	0.17	0.27	0.00	0.02	0.18
Actual 12b-1 Fee (%)	693	0.07	0.12	0.00	0.01	0.09
Size ($ Mln)	693	5,613.54	19,952.50	186.50	918.30	3,096.50
Unique Downloaders	693	351.81	510.91	64	155	387
Family Size ($ Bln)	693	138.62	349.64	8.98	42.41	113.79
Family Equity Share (%)	693	56.59	20.91	45.05	57.61	69.61
Family Passive Share (%)	693	15.31	21.38	0.07	4.76	21.32
Family Fund Count	486	88.53	78.64	33	68	124
Family Asset Class Diversity (HHI)	693	2,271.69	1,450.48	1,289.49	1,755.25	2,664.15
Family Return Diversity (log-det)	680	12.10	1,250.58	−115.55	−1.14	100.86

Fund families offering S&P 500 index mutual funds



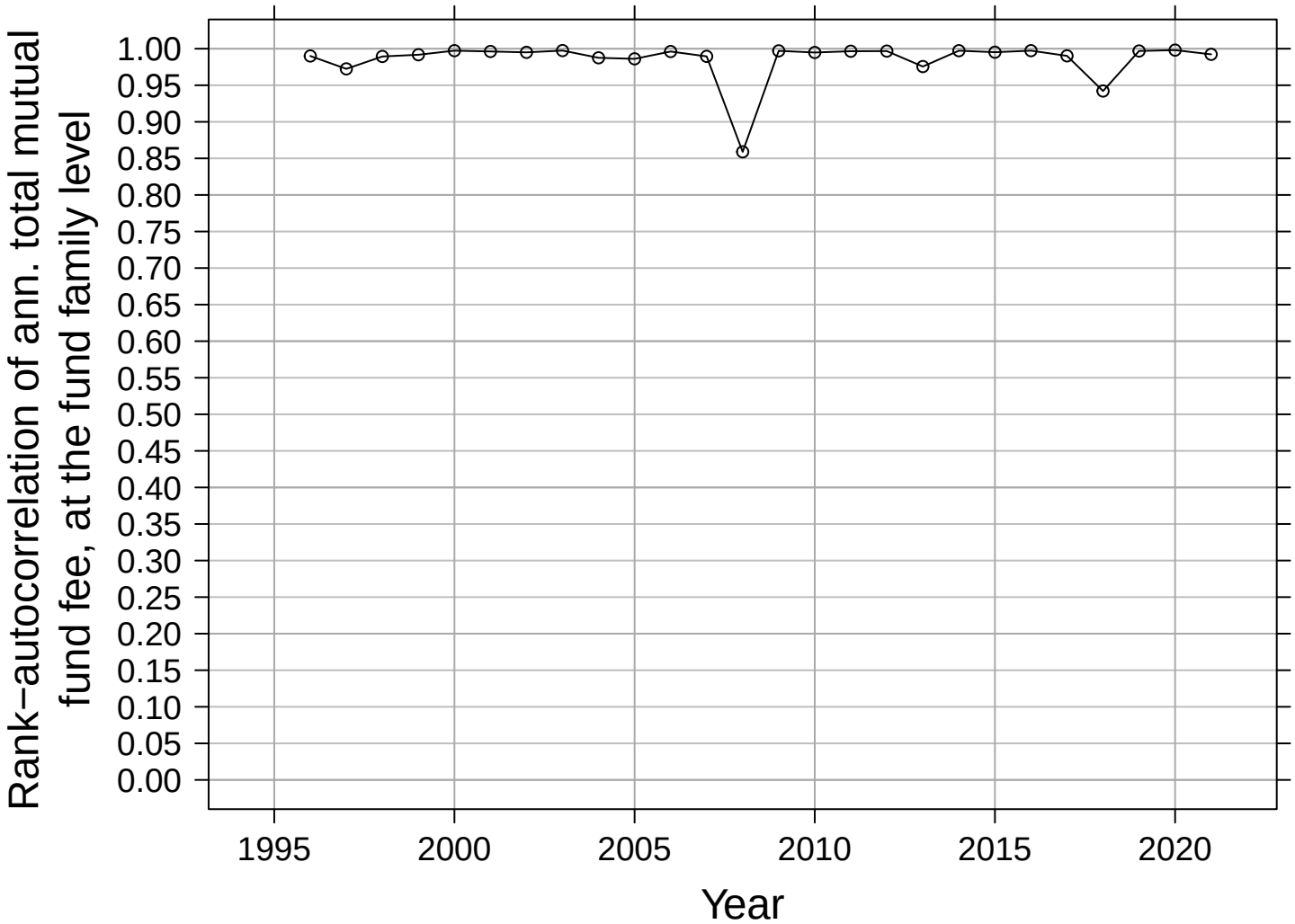
Number of mutual fund families in our sample offering S&P 500 index funds. Sample restricted to families filing prospectuses with the SEC and matched to the CRSP Mutual Fund Database.

S&P 500 index mutual fund total fees over time



Annual mean and quantiles of fund families' total fees (annualized %) on their S&P 500 index funds.

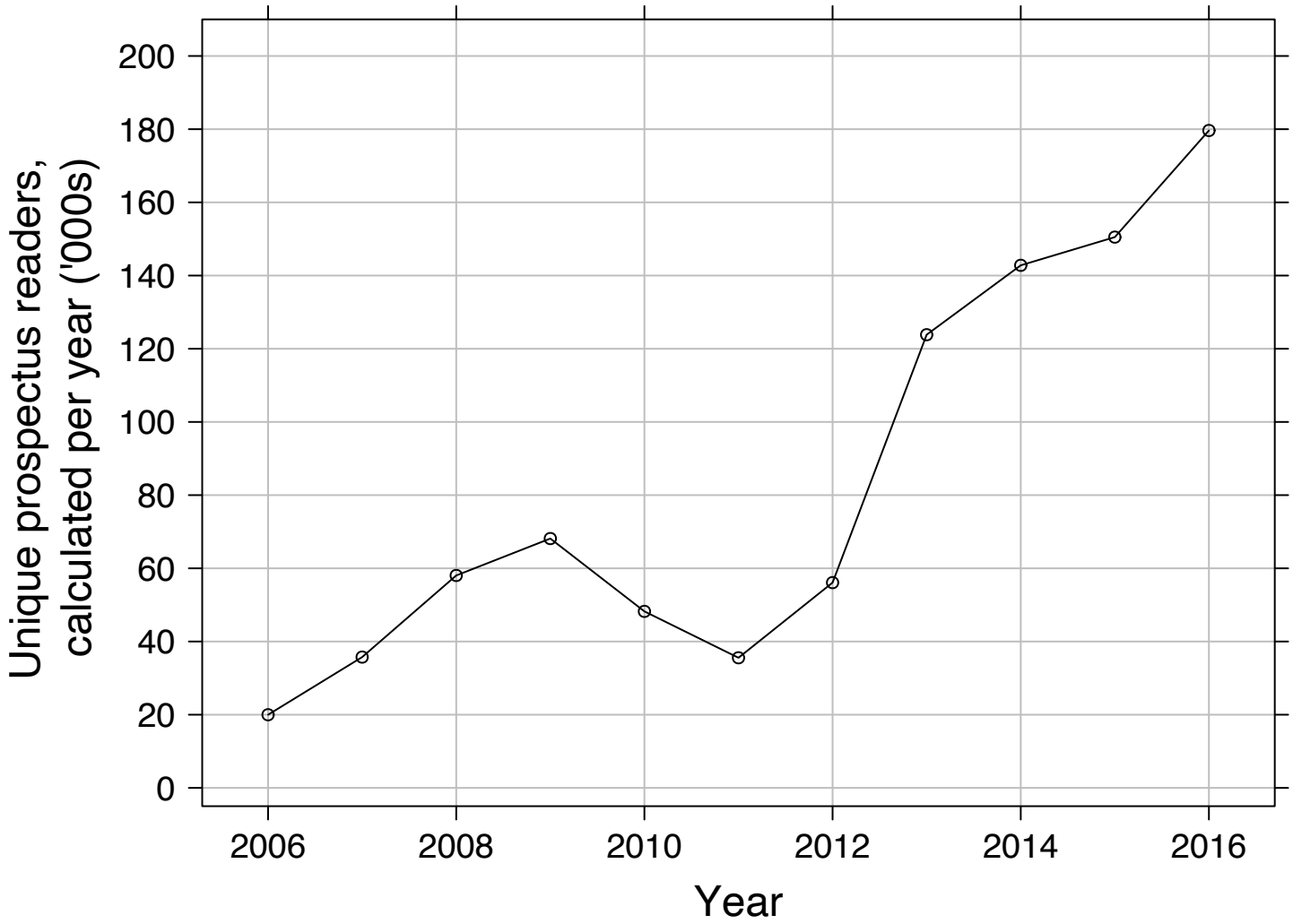
S&P 500 index fund total fee rank-autocorrelation (YoY)



Annual cross-sectional rank-correlation of S&P 500 index mutual fund total fees between year $t-1$ and year t.

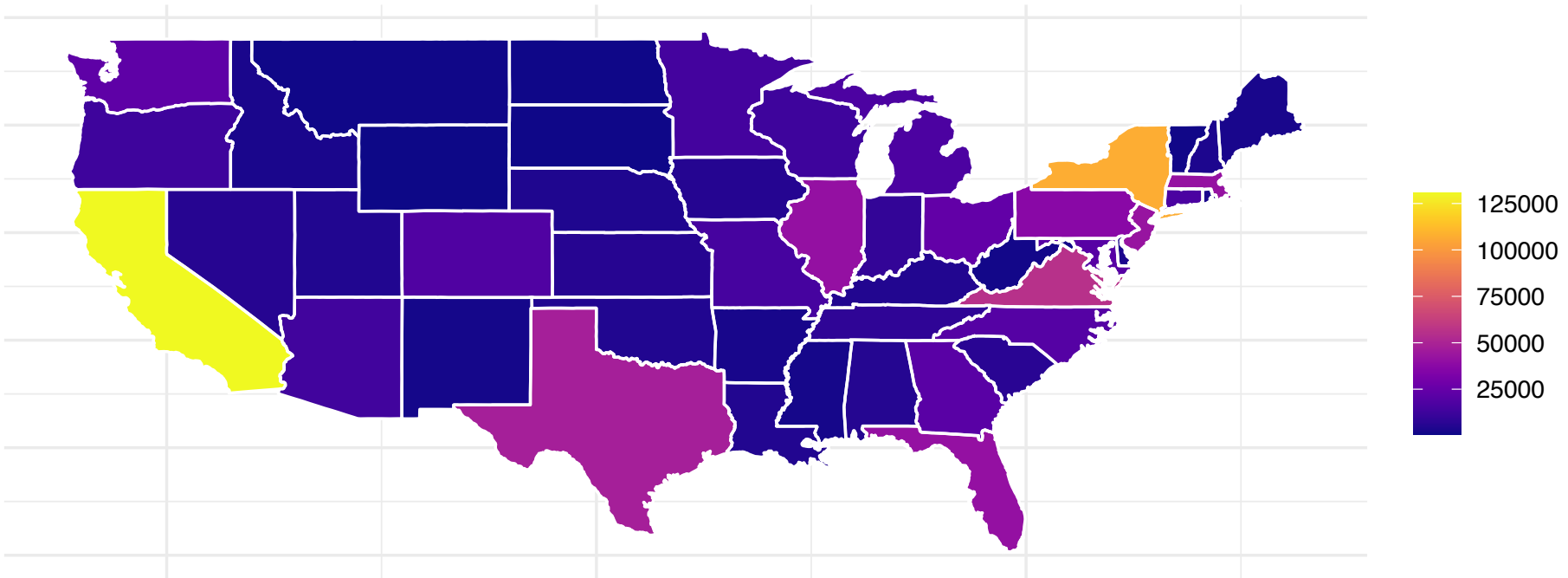
Appendix:
Further Details on EDGAR Prospectus Downloads

SEC EDGAR usage for prospectus downloads is widespread



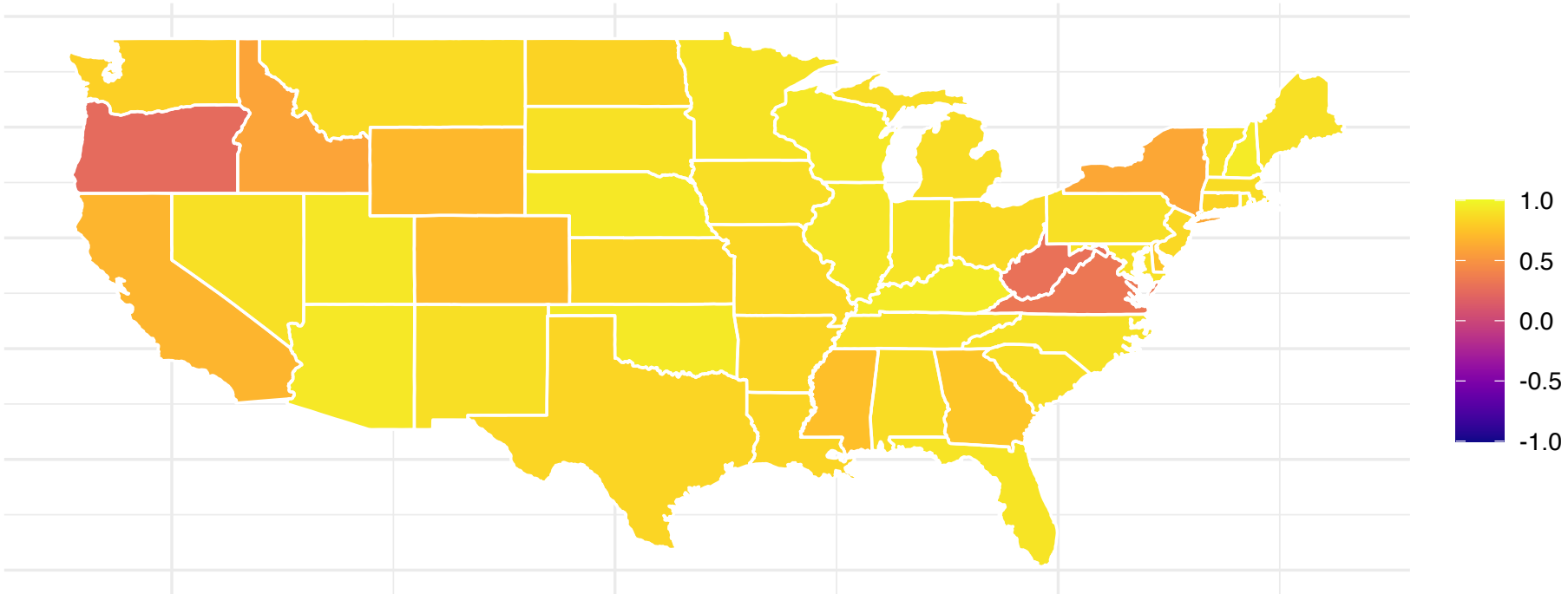
Annual number (in thousands) of unique EDGAR users who downloaded at least one prospectus during the year. Users are identified by IP address.

State-level unique user counts



Unique EDGAR prospectus downloaders per state, 2006–2016. Users geolocated by IP address; lighter states have more unique users. User counts are highly correlated with state population.

Within-state user-to-population correlations



Correlation between unique downloader counts and county population counts, computed within each state, 2006–2016. Lighter states exhibit higher correlations.