Model Uncertainty in the Cross Section

Jiantao Huang (HKU)　　Ran Shi (Colorado Boulder)

August 29, 2024

EEA-ESEM 2024

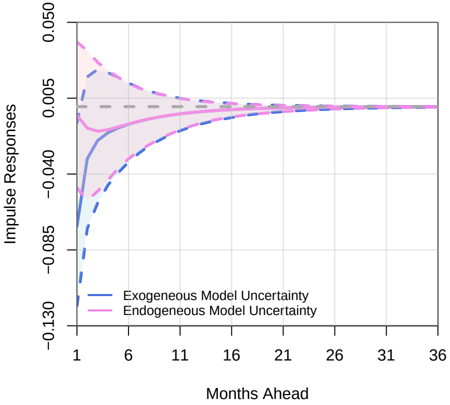
Motivation

- Existing uncertainty measures regarding macroeconomy and asset markets
  - VXO/VIX index: Bloom (2009)
  - Macro/real/financial uncertainty: Jurado, Ludvigson, and Ng (2015); Ludvigson, Ma, and Ng (2021)
  - Policy uncertainty: Baker, Bloom, and Davis (2016)

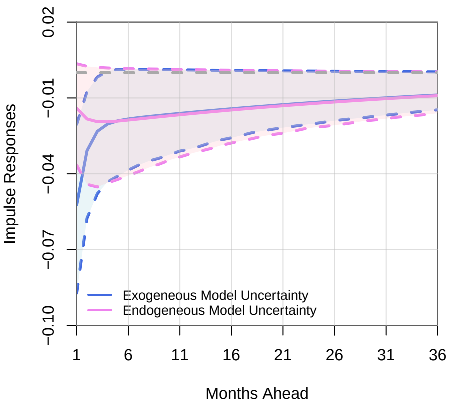
    . . .
  - Observation: time-varying and related to firms' investment and hiring activities
  - Focus on time-series dimension, e.g., the extent to which financial outcomes fluctuate?
  - Insufficient to capture uncertainty faced by equity investors

Motivation

- Existing uncertainty measures regarding macroeconomy and asset markets
  - Insufficient to capture uncertainty faced by equity investors



Equity mutual fund flows to VIX shock



Equity mutual fund flows to financial uncertainty shock

This paper

- Proposes cross-sectional uncertainty measure for Bayesian "investors"
  - Asset allocation (e.g., value or momentum funds) $\Longrightarrow$ Need an asset pricing (AP) model
  - Model uncertainty: Uncertainty regarding which AP model to use

This paper

- Proposes cross-sectional uncertainty measure for Bayesian "investors"
  - Asset allocation (e.g., value or momentum funds) $\Longrightarrow$ Need an asset pricing (AP) model
  - Model uncertainty: Uncertainty regarding which AP model to use
  - Model uncertainty is sizable, and heightened model uncertainty coincides with bad times
    - E.g., 2008 GFC, model uncertainty is maximized $\Longrightarrow$ A true AP model is elusive!
  - Model uncertainty shocks carry negative risk premium

This paper

- Proposes cross-sectional uncertainty measure for Bayesian “investors”
  - Asset allocation (e.g., value or momentum funds) $\Longrightarrow$ Need an asset pricing (AP) model
  - Model uncertainty: Uncertainty regarding which AP model to use
  - Model uncertainty is sizable, and heightened model uncertainty coincides with bad times
    - ★ E.g., 2008 GFC, model uncertainty is maximized $\Longrightarrow$ A true AP model is elusive!
  - Model uncertainty shocks carry negative risk premium

- Documents strong correlations between model uncertainty shocks and fund flows
  - Heightened model uncertainty $\Longrightarrow$ persistent flows out of equity to government bonds
    - ★ Outflows from small-cap and actively managed funds
  - NO such patterns for VIX & financial uncertainty in Ludvigson, Ma, and Ng (2021)

Econometric Theory

The framework: linear SDF models

1. N test assets, $\boldsymbol{R} \in \mathbb{R}^N$, all excess returns
2. p factors, $\boldsymbol{f} \in \mathbb{R}^p$, all tradable long-short portfolios : $\boldsymbol{f} \subseteq \boldsymbol{R}$
3. Linear SDF model: $\mathbb{E}[\boldsymbol{R} \times m] = \mathbf{0}$ in which

$$m = 1 - (\boldsymbol{f} - \mathbb{E}[\boldsymbol{f}])^\top \boldsymbol{b}$$

The framework: linear SDF models

1. N test assets, $\boldsymbol{R} \in \mathbb{R}^N$, all excess returns
2. p factors, $\boldsymbol{f} \in \mathbb{R}^p$, all tradable long-short portfolios : $\boldsymbol{f} \subseteq \boldsymbol{R}$
3. Linear SDF model: $\mathbb{E}[\boldsymbol{R} \times m] = \boldsymbol{0}$ in which

$$m = 1 - (\boldsymbol{f}_\gamma - \mathbb{E}[\boldsymbol{f}_\gamma])^\top \boldsymbol{b}_\gamma$$

4. Model uncertainty: which elements of $\boldsymbol{f}$ determine $\mathbb{E}[\boldsymbol{R}]$?
   - $\gamma_j = 1$: the j-th factor is included ($b_j \neq 0$)
   - $\gamma_j = 0$: $b_j \equiv 0$
5. Data: $\mathcal{D} = \{\boldsymbol{R}_1, \ldots, \boldsymbol{R}_T\} \overset{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
6. Model: a restriction on this distribution through the moment condition

$$\text{Under model } \mathcal{M}_\gamma: \quad \boldsymbol{\mu} = \text{Cov}[\boldsymbol{R}, \boldsymbol{f}_\gamma]\boldsymbol{b}_\gamma$$

The framework: Bayesian inference, g-prior

1. Prior: a generalized version of Arnold Zellner's g-prior for $\boldsymbol{b}_\gamma$:

$$\text{Under model } \mathcal{M}_\gamma: \quad \boldsymbol{b}_\gamma \sim \mathcal{N}\left(\mathbf{0}, \frac{g}{T}\left(\boldsymbol{C}_\gamma \boldsymbol{\Sigma}^{-1} \boldsymbol{C}_\gamma\right)^{-1}\right), \quad g > 0$$

   - Rotation invariant w.r.t. $\mathrm{Cov}[\boldsymbol{R}, \boldsymbol{f}_\gamma] \triangleq \boldsymbol{C}_\gamma$; g: "confidence" in prior information

The framework: Bayesian inference, g-prior

1. Prior: a generalized version of Arnold Zellner's g-prior for $\boldsymbol{b}_\gamma$:

$$\text{Under model } \mathcal{M}_\gamma: \quad \boldsymbol{b}_\gamma \sim \mathcal{N}\left(\mathbf{0}, \frac{g}{T}\left(\boldsymbol{C}_\gamma \boldsymbol{\Sigma}^{-1} \boldsymbol{C}_\gamma\right)^{-1}\right), \quad g > 0$$

   - Rotation invariant w.r.t. $\mathrm{Cov}[\boldsymbol{R}, \boldsymbol{f}_\gamma] \triangleq \boldsymbol{C}_\gamma$; g: "confidence" in prior information

2. Assume equal prior model probabilities: $\pi[\mathcal{M}_\gamma] = \pi[\mathcal{M}_{\gamma'}]$

3. The likelihood function:

$$\mathbb{P}[\mathcal{D} \mid \mathcal{M}_\gamma] \propto \exp\left\{-\frac{T}{2}\left(\mathrm{SR}^2_{\max} - \frac{g}{1+g}\mathrm{SR}^2_\gamma\right) - \frac{p_\gamma}{2}\log(1+g)\right\}$$

   - In-sample Sharpe ratio vs model dimensionality
   - GRS test (Gibbons, Ross, and Shanken, 1989) intuition in factor zoo

Model uncertainty in the cross section: definition

Our cross-sectional uncertainty measure:

$$\mathcal{E} = -\frac{1}{p\log 2}\sum_{\gamma}\left(\log \mathbb{P}[\mathcal{M}_\gamma \mid \mathcal{D}]\right)\mathbb{P}[\mathcal{M}_\gamma \mid \mathcal{D}]$$

1. entropy$[\mathcal{M}_\gamma \mid \mathcal{D}] \in [0, 1]$
2. minimized when $\mathcal{D}$ favors one dominate model
   - for one dominate model to exist: large SR^2_γ with small p_γ
3. maximized when $\mathcal{D}$ lends equal evidence across models

Posterior property: Pitfall of g-priors

Theorem

Assume that the observed return data are generated from a true linear SDF model $\mathcal{M}_{\gamma_0}$. If $\gamma_0 \neq \mathbf{0}$ (the SDF is not a constant) and $f_{\gamma_0} \subset f$ (the set of factors under consideration include all true factors), under the g-prior specification with $g \in (0, \infty)$, as $T \to \infty$,

1. *(factor selection consistency) if factor j belongs to the true model $\mathcal{M}_{\gamma_0}$, the posterior marginal probability of choosing it converges to one in probability:*
$$\mathbb{P}[\gamma_j = 1 \mid \mathcal{D}] \xrightarrow{p} 1;$$
2. *(model selection inconsistency) the posterior probability of the true model will always be strictly smaller than one, that is, $\mathbb{P}[\mathcal{M}_{\gamma_0} \mid \mathcal{D}] < 1$ with probability one.*

- g-priors can avoid discarding true factors, at the cost of incorporating redundant ones

Restoring model selection consistency: mixture of g-priors

- Adapt mixture of g-priors proposed by Liang et al. (2008) into SDF models:

$$\pi(g) = \frac{1}{(1+g)^2}, \quad g > 0$$

- Bayes factor for comparing model $\mathcal{M}_\gamma$ with the null model $\mathcal{M}_{\mathbf{0}}$

$$\mathrm{BF}(\gamma, \mathbf{0}) = \exp\left(\frac{T}{2}\mathrm{SR}_\gamma^2\right)\left(\frac{T}{2}\mathrm{SR}_\gamma^2\right)^{-\frac{p_\gamma+2}{2}} \underline{\Gamma}\left(\frac{p_\gamma+2}{2}, \frac{T}{2}\mathrm{SR}_\gamma^2\right)$$

  - $\mathrm{BF}(\gamma, \mathbf{0})$ is increasing in SR_γ^2 but decreasing in p_γ

Restoring model selection consistency: mixture of g-priors

- Adapt mixture of g-priors proposed by Liang et al. (2008) into SDF models:

$$\pi(g) = \frac{1}{(1+g)^2}, \quad g > 0$$

- Bayes factor for comparing model $\mathcal{M}_\gamma$ with the null model $\mathcal{M}_{\mathbf{0}}$

$$\mathrm{BF}(\gamma, \mathbf{0}) = \exp\left(\frac{T}{2}\mathrm{SR}_\gamma^2\right)\left(\frac{T}{2}\mathrm{SR}_\gamma^2\right)^{-\frac{p_\gamma+2}{2}} \underline{\Gamma}\left(\frac{p_\gamma+2}{2}, \frac{T}{2}\mathrm{SR}_\gamma^2\right)$$

  - $\mathrm{BF}(\gamma, \mathbf{0})$ is increasing in SR_γ^2 but decreasing in p_γ

Theorem

Under the mixture of g-priors specification, as $T \to \infty$, $\mathbb{P}[\mathcal{M}_{\gamma_0} \mid \mathcal{D}] \xrightarrow{p} 1.$

- Mixture of g-priors is to achieve posterior model selection consistency! ▸ Simulations

A misspecified set of factors

- f_0: the true set of factors, $f_0 \nsubseteq f \implies$ the studied factors omit some pricing factors

A misspecified set of factors

- f_0: the true set of factors, $f_0 \not\subseteq f \implies$ the studied factors omit some pricing factors

Theorem

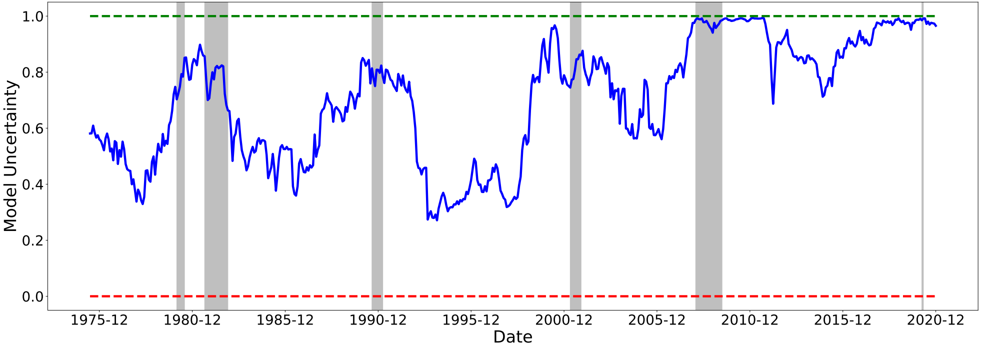
Assume that the observed return data are generated from a true linear SDF $m_0 = 1 - (f_0 - \mathbb{E}[f_0])^\top b_0$*. Let* $f_{\gamma_0} = f_0 \cap f$*; that is,* f_{γ_0} *is the subset of* f *that includes only the true pricing factors. As* $T \to \infty$*,*

1. *for all* j *such that* $\gamma_{0,j} = 1$, $\mathbb{P}[\gamma_j = 1 \mid \mathcal{D}] \xrightarrow{p} 1$;
2. *model uncertainty measure* $\mathcal{E}$ *satisfies* $\mathcal{E} \leq (p - p_{\gamma_0})/p$ *with probability one.*

- True factors in f can always be selected $\implies$ factor selection consistency ▸ Simulations
- If $\mathcal{E} \approx 1$, only two possibilities:
  - (1) $p_{\gamma_0} \approx 0$ (all factors under study are useless/no strong factors)
  - (2) Observed data entirely uninformative about the true SDF model

Empirics

Model uncertainty in US stock markets ▸ Data ▸ Examples



- Entropy $\in [0.27, 0.99]$: Average (standard dev) $= 0.70\ (0.21)$
- Heightened model uncertainty coincides with economic downturns and stock market crashes
- Model uncertainty hits upper bounds during 2008–2011 & 2018–2020
  - Marginal probs of all factors are low $\Longrightarrow$ Model/factor selection is elusive in these periods

Is Model Uncertainty Priced?

Table 1: Risk Premia of Model Uncertainty Shocks: Monthly Estimates

Number of PCs:	5	6	7	8	9	10
$\lambda_{\mathcal{E}}$	-0.066***	-0.067***	-0.067***	-0.065***	-0.062***	-0.060***
s.e.	0.017	0.017	0.017	0.017	0.018	0.018
Time-series R^2	5.8%	5.8%	5.8%	6.2%	6.2%	6.2%

The table reports the risk premia estimates of model uncertainty shocks ($\mathcal{E}_t^{ar1}$) based on the three-pass method of Giglio and Xiu (2021). In all estimations, we standardize $\mathcal{E}_t^{ar1}$ to have a unit variance. In particular, we project $\mathcal{E}_t^{ar1}$ onto the space of large PCs of 275 Fama-French characteristic-sorted portfolios in the US market. The number of latent factors ranges from five to 10. If the 90% (95%, 99%) confidence interval of the risk premium does not contain zero, the risk premium estimate will be highlighted by * (**, ***). We also report the time-series fit in each panel. Sample: 1975/07 - 2020/12.

- Investors willing to pay a premium to hedge against heightened model uncertainty
- To equate standard deviation of model uncertainty shocks to the market:

  Annualized risk premium equals $(\sqrt{12})\lambda_{\mathcal{E}} \times 17\% \approx -4\%$

Model Uncertainty and Mutual Fund Flows

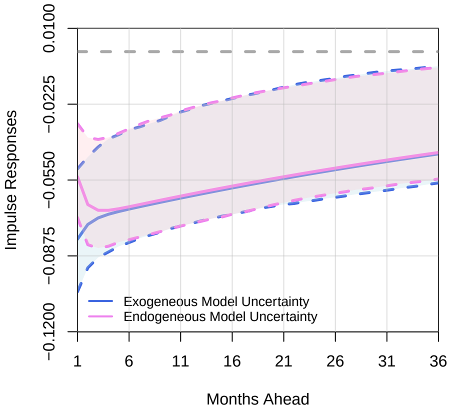
- Study the dynamic responses of fund flows to uncertainty shocks using VAR:

$$\boldsymbol{Y}_t = \boldsymbol{B}_0 + \boldsymbol{B}_1 \boldsymbol{Y}_{t-1} + \cdots + \boldsymbol{B}_l \boldsymbol{Y}_{t-l} + \text{controls}_t + \boldsymbol{S}\boldsymbol{\epsilon_t}, \quad \boldsymbol{\epsilon_t} \sim WN(\boldsymbol{0}, \boldsymbol{I})$$

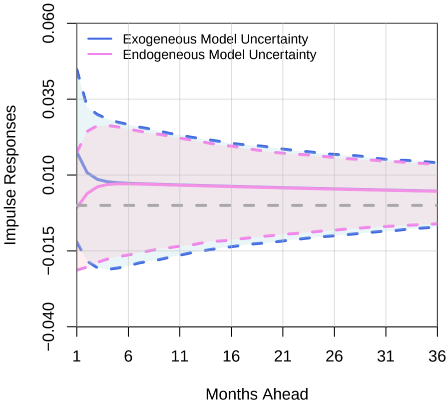
  l denotes the lag order and is chosen by AIC/BIC (always equal to 1)

- Use Cholesky decomposition to identify the dynamic responses (S):
  - exogenous cause: uncertainty measure as the first element in $\boldsymbol{Y_t}$
  - propagating mechanism: uncertainty measure as the last element

Relationships with aggregate fund flows

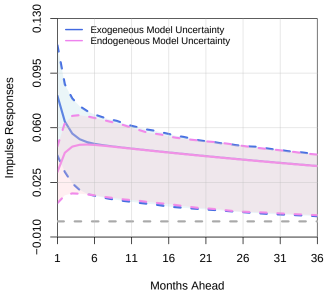


Equity Fund Flows

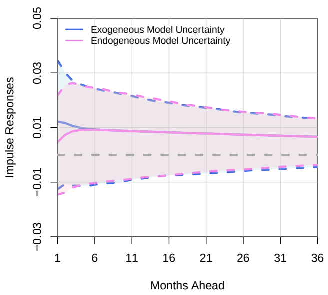


Fixed-Income Fund Flows

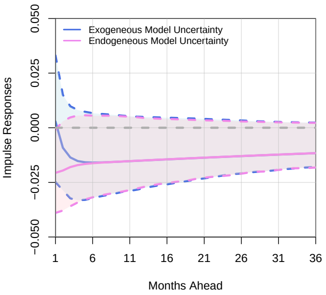
Fixed-income fund flows with different investment objective codes



Government bonds



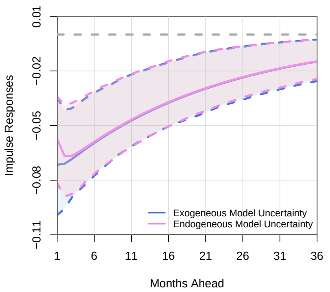
Money markets



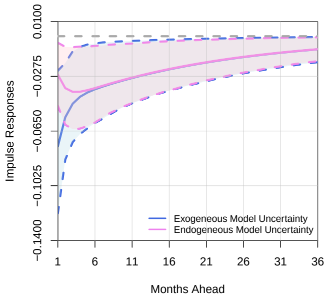
Corporate bonds

- Following high model uncertainty, sharp dynamic inflows to government bond funds
- No such patterns in money market or corporate bond funds

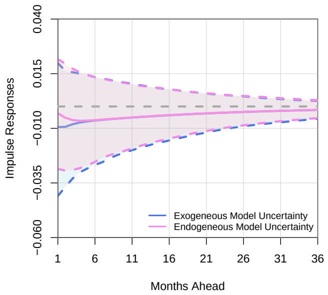
Equity fund flows with different investment objective codes



Style equity fund



Small-cap funds



Large-cap funds

- Following high model uncertainty shocks, equity outflows mainly from style and small-cap funds, instead of large-cap or sector funds

Conclusion

- Propose a measure of model uncertainty in the cross section that
  - is based on rigorous Bayesian econometric framework
  - is transparent with lower and upper bounds
  - varies significantly over time and is persistently high in bad times
  - commands a significantly negative risk premium
- Combined with low marginal factor probabilities: selecting SDF models is elusive
  - Example periods: 2008 GFC, recent years from 2018–2020
- Outflows from equity funds into US government bond funds under high uncertainty:
  - No such patterns detected using VIX or financial uncertainty ▸ VIX ▸ financial uncertainty

Appendix

Data
Back

- 14 prominent factors, from July 1972 to December 2020
  - Fama-French five factors (*Fama and French, 2016*)
  - Momentum factor (*Jegadeesh and Titman, 1993*)
  - Size, investment, and profitability factors (*Hou, Xue, and Zhang, 2015*)
  - Short and long-term behavioral factors (*Daniel, Hirshleifer, and Sun, 2020*)
  - HML devil (*Asness and Frazzini, 2013*)
  - Quality-minus-junk (*Asness et al., 2019*)
  - Betting-against-beta (*Frazzini and Pedersen, 2014*)
- Consider models that contain at most one factor from following categories:
  - Size (SMB or ME)
  - Profitability (RMW or ROE)
  - value (HML or HML Devil)
  - investment (CMA or IA)

Simulation study: posterior properties without misspecification

Back

scenario	$T = 750$ g=2	g=4	g=16	*mix.* g	$T = 1500$ g=2	g=4	g=16	*mix.* g	$T = 15000$ g=2	g=4	g=16	*mix.* g
Posterior Probabilities of Factors $\mathbb{P}[\gamma_j = 1 \mid \mathcal{D}]$:												
$\gamma_{0,j} = 1$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\gamma_{0,j} = 0$	0.48	0.44	0.33	0.22	0.48	0.44	0.33	0.18	0.47	0.43	0.32	0.08
	(0.13)	(0.16)	(0.18)	(0.17)	(0.14)	(0.16)	(0.18)	(0.17)	(0.13)	(0.15)	(0.18)	(0.12)
Posterior Probabilities of Models $\mathbb{P}[\mathcal{M}_\gamma \mid \mathcal{D}]$:												
$\mathcal{M}_\gamma = \mathcal{M}_{\gamma_0}$	0.28	0.32	0.45	0.61	0.28	0.32	0.45	0.67	0.28	0.33	0.46	0.85
	(0.10)	(0.13)	(0.17)	(0.19)	(0.10)	(0.13)	(0.18)	(0.20)	(0.10)	(0.13)	(0.17)	(0.16)
$\mathcal{M}_\gamma \supset \mathcal{M}_{\gamma_0}$	0.24	0.23	0.18	0.13	0.24	0.23	0.18	0.11	0.24	0.22	0.18	0.05
	(0.09)	(0.11)	(0.13)	(0.13)	(0.10)	(0.11)	(0.13)	(0.13)	(0.09)	(0.11)	(0.13)	(0.10)
Model Uncertainty Measure $\mathcal{E}$:												
	0.38	0.37	0.32	0.26	0.38	0.36	0.32	0.23	0.38	0.37	0.32	0.12
	(0.03)	(0.03)	(0.04)	(0.05)	(0.03)	(0.03)	(0.04)	(0.06)	(0.03)	(0.03)	(0.04)	(0.06)

- f = carhart four factors, true factors = FF3
- 1,000 simulations with $T = 750, 1,500, 15,000$ trading days

Simulation study: model uncertainty under misspecification

▸ Back

- f = $\{\text{MKT}, \text{SMB}, \text{HML}, \text{MOM}, \text{RMW}, \text{CMA}, \text{QMJ}, \text{FIN}, \text{PEAD}, \text{BAB}\}$ *excluding* the factor in each column
- True factors: FF3 plus the omitted factor (the name of which is at the top of each column)
- 1,000 simulations with $T = 750$ days, with standard deviations across simulations in parenthesis

Omitted factor:	*MOM*	*RMW*	*CMA*	*FIN*	*PEAD*	*QMJ*	*BAB*
Posterior Probabilities of Factors $\mathbb{P}[\gamma_j = 1 \mid \mathcal{D}]$:							
$\gamma_{0,j} = 1$	0.94	1.00	1.00	1.00	1.00	0.99	1.00
	(0.17)	(0.02)	(0.00)	(0.01)	(0.01)	(0.06)	(0.01)
$\gamma_{0,j} = 0$	0.49	0.35	0.27	0.43	0.29	0.36	0.38
	(0.33)	(0.27)	(0.21)	(0.31)	(0.22)	(0.25)	(0.27)
Model Uncertainty Measure $\mathcal{E}$:							
	0.46	0.46	0.46	0.43	0.46	0.49	0.47
	(0.08)	(0.06)	(0.05)	(0.07)	(0.05)	(0.06)	(0.06)
Upper bound $= (p - p_{\gamma_0})/p$	0.67	0.67	0.67	0.67	0.67	0.67	0.67

Note: Extremely high model uncertainty should not be driven by model misspecification.

Regressions of Model Uncertainty on Contemporaneous Variables

X	*FinU.*	*MacroU.*	*RealU.*	*EPUI*	*EPUII*	*VIX*	*TS*	*DS*
β	0.21	0.17	0.14	0.00	0.00	0.01	-0.03	-0.00
	(1.95)	(1.53)	(1.20)	(0.33)	(1.07)	(2.20)	(-3.44)	(-0.09)
#obs.	546	546	546	432	432	420	546	546

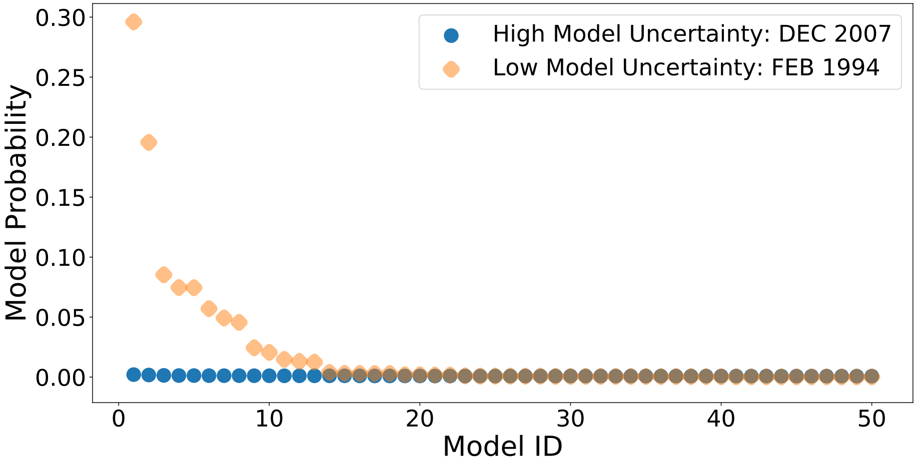
The table reports results from the following regression:

$$\mathcal{E}_t = \beta_0 + \beta X_t + \rho \mathcal{E}_{t-1} + \epsilon_t,$$

where the variable X_t represents a) macro, financial, and real uncertainty measures from Jurado et al. (2015) and Ludvigson et al. (2021) (Fin U, Macro U, and Real U); b) two economic policy uncertainty (EPU) indices from Baker et al. (2016) (EPU I and EPU II); c) the CBOE VIX index (VIX); d) the term spread between ten-year and three-month treasuries (TS), e) the default spread between BAA and AAA corporate bond yields (DS). The t-statistics in parenthesis are computed based on Newey-West standard errors with 36 lags.

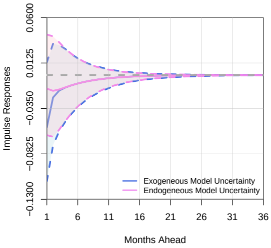
Model uncertainty in the cross section: two states

Back

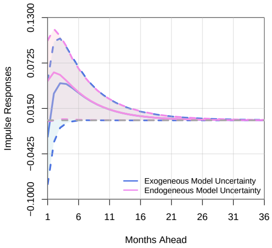


Equity fund flows to VIX shocks

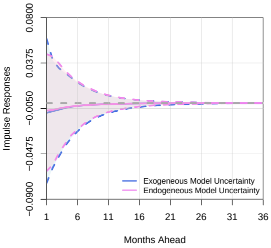
▸ Back



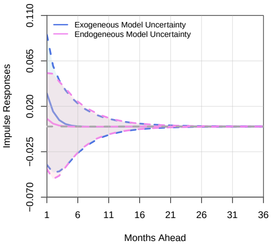
Style equity fund



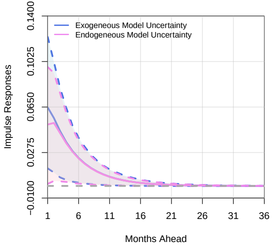
Small-cap funds



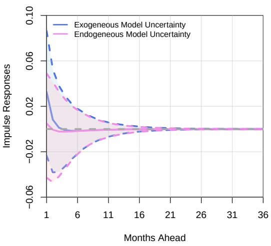
Large-cap funds



Government bonds



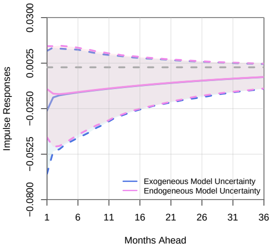
Money markets



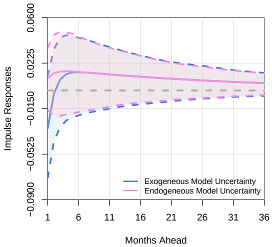
Corporate bonds

Equity fund flows to financial uncertainty shocks

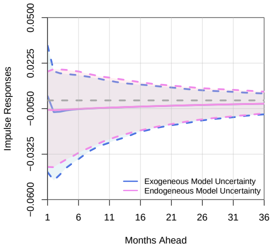
Back



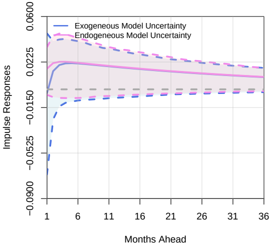
Style equity fund



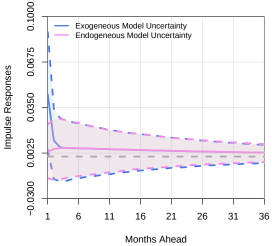
Small-cap funds



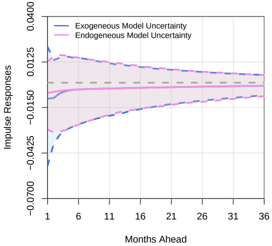
Large-cap funds



Government bonds



Money markets



Corporate bonds