Model Uncertainty in the Cross Section

Jiantao Huang (HKU)　　Ran Shi (Colorado Boulder)

Jan 5, 2025

AFA 2025, San Francisco

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); News implied volatility (Manela and Moreira, 2017) . . .
  - Time-varying and related to firms' investment/production/hiring activities
  - Focus on time-series dimension, e.g., the extent to which financial outcomes fluctuate
  - Insufficient to capture uncertainty related to asset allocation decisions

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); News implied volatility (Manela and Moreira, 2017) . . .
  - Time-varying and related to firms' investment/production/hiring activities
  - Focus on time-series dimension, e.g., the extent to which financial outcomes fluctuate
  - Insufficient to capture uncertainty related to asset allocation decisions

- Our 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
  - Two-model example: 99-1% (low uncertainty) vs. 50-50% (high uncertainty)

Main Findings

- Model uncertainty is sizable: heightened model uncertainty coincides with bad times
  - E.g., 2008 GFC, model uncertainty is maximized $\implies$ 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 $\implies$ persistent flows out of equity to government bonds
  - Outflows from small-cap and actively managed funds
  - NO such patterns for VIX & financial uncertainty

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: $\mathcal{D} = \{\boldsymbol{R}_1, \ldots, \boldsymbol{R}_T\} \stackrel{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
6. Model: a restriction on this distribution through the moment condition

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

The framework: Bayesian inference, g-prior

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

$$\text{Under model } \mathcal{M}_\gamma: \quad \boldsymbol{b}_\gamma \sim \mathcal{N}\left(\mathbf{0},\ g \cdot \mathrm{Var}(\boldsymbol{b}_\gamma)\right), \quad g > 0$$

- g: "confidence" in prior information

The framework: Bayesian inference, g-prior

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

$$\text{Under model } \mathcal{M}_\gamma: \quad \boldsymbol{b}_\gamma \sim \mathcal{N}\left(\mathbf{0},\ g \cdot \mathrm{Var}(\boldsymbol{b}_\gamma)\right), \quad g > 0$$

   - g: "confidence" in prior information

2. 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
   - $\mathrm{SR}^2_{\max}$ enters every model: Irrelevance of test assets (efficient GMM intuition)

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. $\text{entropy}[\mathcal{M}_\gamma \mid \mathcal{D}] \in [0, 1]$
2. minimized when $\mathcal{D}$ favors one dominant model
   - for one dominant model to exist: large SR_γ^2 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* $\boldsymbol{f}_{\gamma_0} \subset \boldsymbol{f}$ *(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 identify 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:

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

  - $g/(1+g)$: decides how posterior beliefs about the SDF should be updated
  - $g/(1+g) \sim \text{Uniform}(0,1)$: agnostic regarding this posterior updating process

- Bayes factor still has a closed-solution
  - Remains the same economic tradeoff: 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:

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

  - $g/(1+g)$: decides how posterior beliefs about the SDF should be updated
  - $g/(1+g) \sim \text{Uniform}(0,1)$: agnostic regarding this posterior updating process

- Bayes factor still has a closed-solution
  - Remains the same economic tradeoff: 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, $f_0 \not\subseteq f \implies$ 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

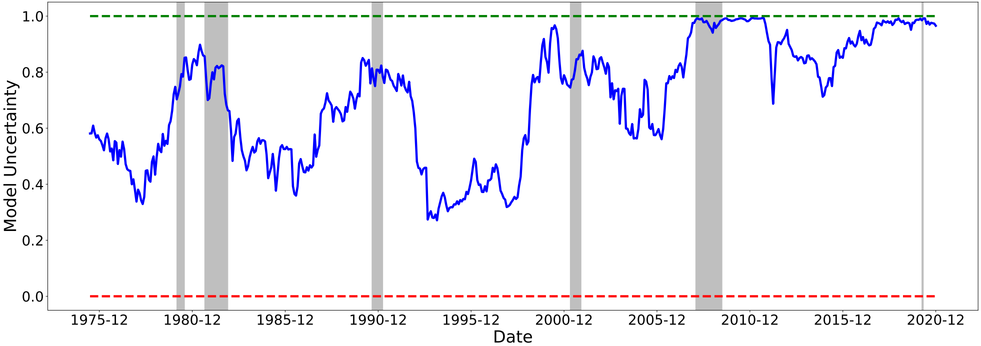
Assume that the observed return data are generated from a true linear SDF $m_0 = 1 - (f_0 - \mathbb{E}[f_0])^\top b_0$*. Let* $f_{\gamma_0} = f_0 \cap f$*; that is,* f_{γ_0} *is the subset of* f *that includes only the true pricing factors. As* $T \to \infty$*,*

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 (14 factors) ▸ Data ▸ Examples

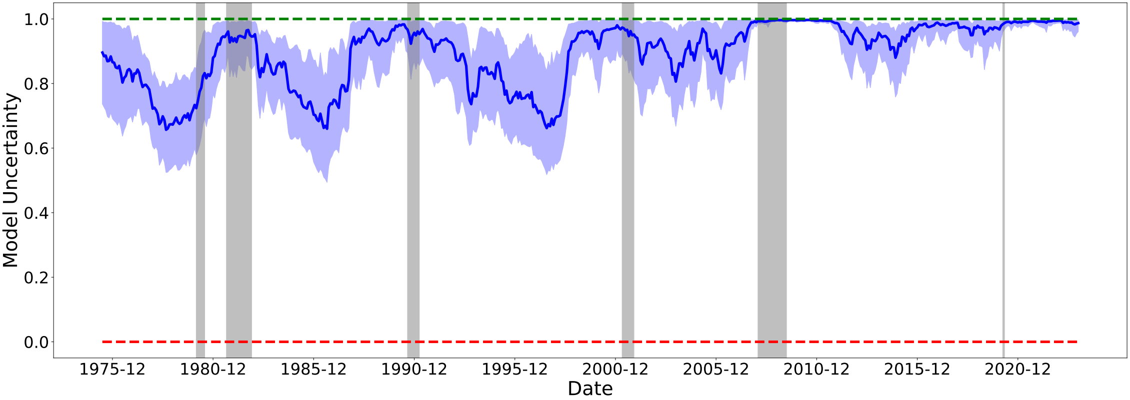


- Entropy $\in [0.27, 0.99]$: Average (standard dev) $= 0.70\ (0.21)$
- Heightened model uncertainty coincides with economic downturns / stock market crashes
- Model uncertainty hits upper bounds during 2008–2011 & 2018–2020
  - Marginal probs of all factors $\approx 50\% \implies$ Model/factor selection is elusive in these periods

Model uncertainty in the "Factor Zoo"

- A zoo of 153 factors in Jensen, Kelly, and Pedersen (2023)
  - Categorized into 13 clusters, e.g., accruals, investment, value, seasonality …
- Yet, handling simultaneously all 153 factors is infeasible
  - Theoretical property of our Bayesian approach based on small p and large T
  - Comparing 2^{153} models computationally impracticable
- A practical solution:
  - Randomly select one factor in each of the 13 cluster themes
  - Measure model uncertainty based on market factor + 13 randomly selected factors
  - Repeat this exercise 1000 times

Model uncertainty in the “Factor Zoo” – Continued



- Even in this exercise, model uncertainty measure
  - Is strongly time-varying, displaying clear business-cycle patterns
  - Has stayed at a very high level since the 2008 GFC

Is Model Uncertainty Priced? ▸ Robustness Check

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
- 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

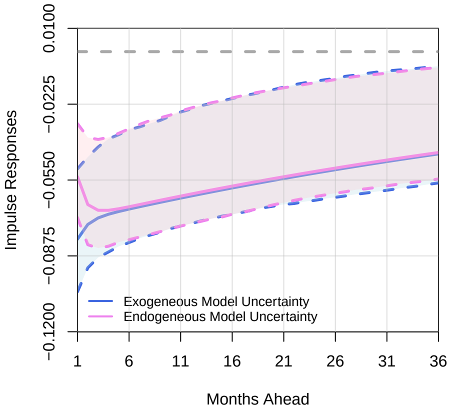
- Study the dynamic responses of fund flows to uncertainty shocks using VAR:

$$\boldsymbol{Y}_t = \boldsymbol{B}_0 + \boldsymbol{B}_1 \boldsymbol{Y}_{t-1} + \cdots + \boldsymbol{B}_l \boldsymbol{Y}_{t-l} + \text{controls}_t + \boldsymbol{S}\boldsymbol{\epsilon}_t$$

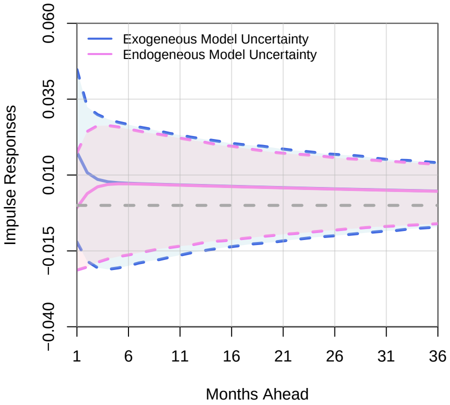
  l denotes the lag order and is chosen by AIC/BIC (equal to 1)

- Use Cholesky decomposition to identify the dynamic responses ($\boldsymbol{S}$):
  - 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

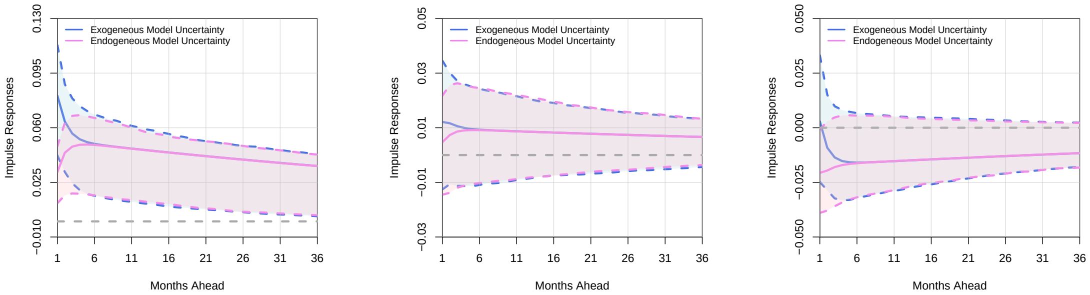


Equity Fund Flows



Fixed-Income Fund Flows

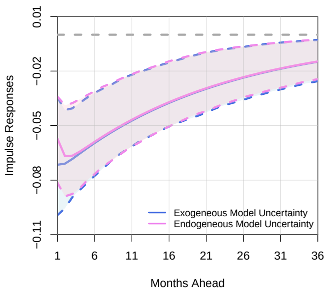
Fixed-income fund flows with different investment objective codes



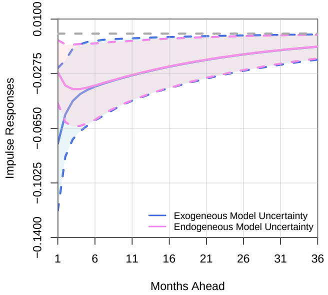
Government bonds — Money markets — Corporate bonds

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

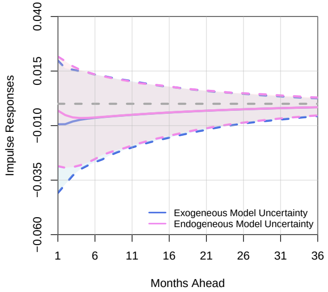
Equity fund flows with different investment objective codes



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

- Model uncertainty in cross-sectional asset pricing
  - 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

	Three years				Five years				50 years			
Scenario	g=2	g=4	g=16	*mix.* g	g=2	g=4	g=16	*mix.* g	g=2	g=4	g=16	*mix.* g
Posterior Probabilities of Factors $\mathbb{P}[\gamma_j = 1 \mid \mathcal{D}]$:												
$\gamma_{0,j} = 1$	0.98	0.98	0.98	0.98	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00
	(0.03)	(0.03)	(0.03)	(0.04)	(0.01)	(0.01)	(0.02)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)
$\gamma_{0,j} = 0$	0.49	0.45	0.34	0.25	0.48	0.44	0.34	0.21	0.48	0.44	0.33	0.10
	(0.07)	(0.09)	(0.10)	(0.10)	(0.07)	(0.08)	(0.10)	(0.10)	(0.07)	(0.08)	(0.09)	(0.07)
Posterior Probabilities of Models $\mathbb{P}[\mathcal{M}_\gamma \mid \mathcal{D}]$:												
$\mathcal{M}_\gamma = \mathcal{M}_{\gamma_0}$	0.07	0.09	0.18	0.29	0.07	0.10	0.19	0.39	0.08	0.10	0.20	0.67
	(0.04)	(0.06)	(0.12)	(0.17)	(0.04)	(0.06)	(0.12)	(0.19)	(0.04)	(0.06)	(0.12)	(0.22)
$\mathcal{M}_\gamma \supset \mathcal{M}_{\gamma_0}$	0.05	0.05	0.05	0.04	0.06	0.06	0.05	0.04	0.06	0.06	0.05	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)
Model Uncertainty Measure $\mathcal{E}$:												
	0.42	0.40	0.36	0.31	0.39	0.37	0.33	0.26	0.38	0.36	0.32	0.14
	(0.04)	(0.04)	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.04)	(0.02)	(0.03)	(0.03)	(0.04)

- true factors = {market, SMB, HML, MOM, RMW, CMA}
- f = {market, SMB, HML, MOM, RMW, CMA, QMJ, FIN, PEAD, BAB}
- 1,000 simulations; Sample sizes are three, five, and 50 years of daily observations

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

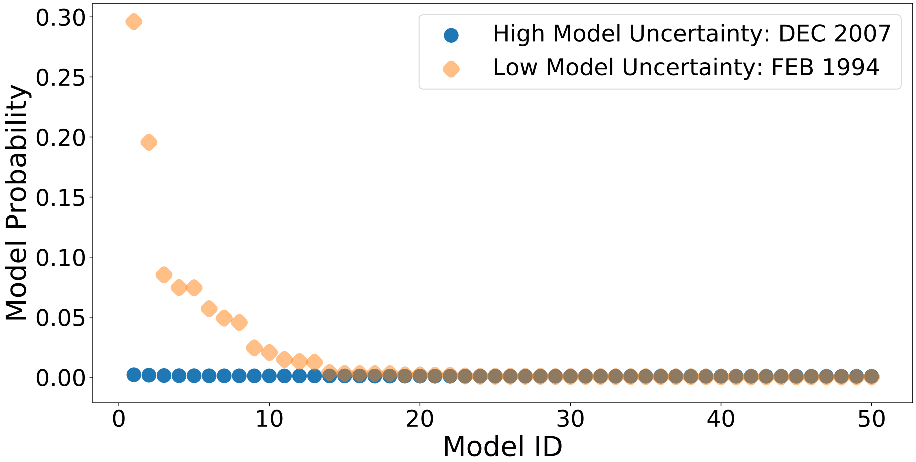
The table reports results from the following regression:

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

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



Is Model Uncertainty Priced? – Robustness Check ▸ Back

- Control for VIX & financial uncertainty in the AR(1) regression of model uncertainty
  $\Longrightarrow$ model uncertainty innovations now orthogonal to both two variables

Table A1: Risk Premia of Model Uncertainty Shocks: Monthly Estimates

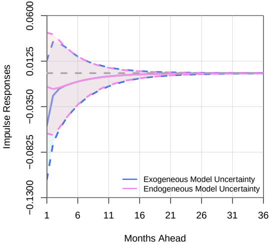
Number of PCs:	5	6	7	8	9	10
$\lambda_{\mathcal{E}}$	-0.058***	-0.059***	-0.059***	-0.057***	-0.055***	-0.051***
s.e.	0.016	0.017	0.017	0.017	0.017	0.018
Time-series R^2	5.0%	5.1%	5.1%	5.3%	5.4%	5.6%

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). We add the control for contemporaneous term spread and financial uncertainty to estimate the AR(1) innovations in model uncertainty. 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.

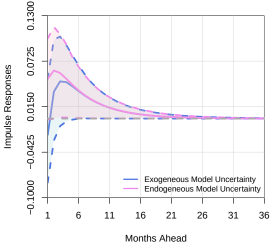
- Risk premia estimates, as well as the time-series fit, are similar to the baseline case

Equity fund flows to VIX shocks

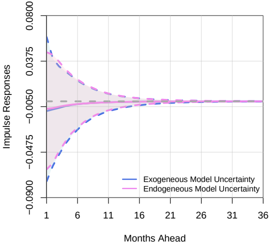
Back



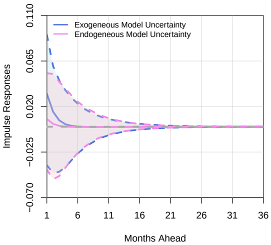
Style equity fund



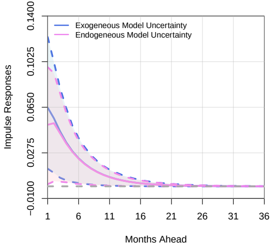
Small-cap funds



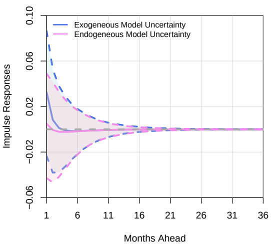
Large-cap funds



Government bonds

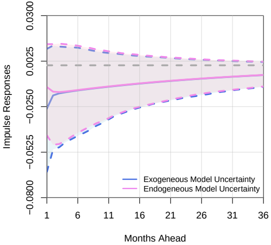


Money markets

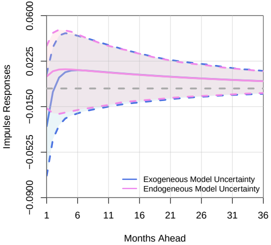


Corporate bonds

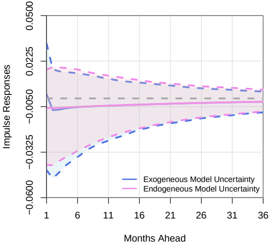
Equity fund flows to financial uncertainty shocks ▸ Back



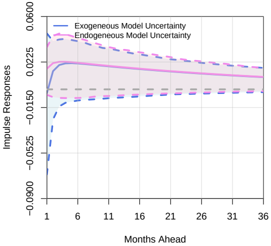
Style equity fund



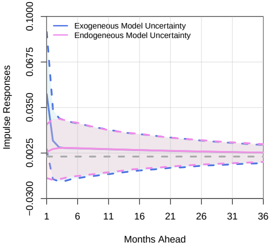
Small-cap funds



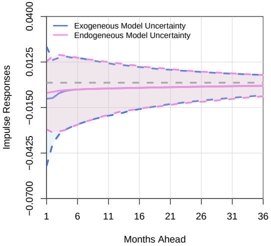
Large-cap funds



Government bonds



Money markets



Corporate bonds

Does Model Uncertainty Matter?

Table A2: Out-of-Sample Model Performance

	BMA	Top 1	Full Model	Carhart4	FF5	HXZ4	DHS3
Low Model Uncertainty	2.572	2.565	2.568	1.288	1.624	1.829	2.282
	-	-	-	***	***	***	-
Middle Model Uncertainty	1.717	1.653	1.771	0.450	0.677	1.232	1.818
	-	-	-	***	***	**	-
High Model Uncertainty	1.251	1.125	1.106	0.564	0.584	0.552	0.897
	-	*	*	***	***	***	**

First three columns: (1) BMA: Bayesian model averaging; (2) Top 1: the model with the highest model probability; (3) Full Model: include all 14 factors. We report the results on testing the null hypothesis that the Sharpe ratio of BMA is equal to the model γ, i.e., $H_0 : \text{SR}^2_{bma} = \text{SR}^2_{\gamma}$. We use the non-parametric Bootstrap to test the null hypothesis. *, ** and *** denote significance at the 90%, 95%, and 99% level, respectively.

Note: Model uncertainty matters only when it is heightened