DISCUSSION PAPER SERIES

DP16748

Ambiguity with Machine Learning: An Application to Portfolio Choice

Eric Ghysels, Yan Qian and Steve Raymond

FINANCIAL ECONOMICS

CEPR

ISSN 0265-8003

Ambiguity with Machine Learning: An Application to Portfolio Choice

Eric Ghysels, Yan Qian and Steve Raymond

Discussion Paper DP16748
Published 22 November 2021
Submitted 15 November 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Ambiguity with Machine Learning: An Application to Portfolio Choice

Abstract

To characterize ambiguity we use machine learning to impose guidance and discipline on the formulation of expectations in a data-rich environment. In addition, we use the bootstrap to generate plausible synthetic samples of data not seen in historical real data to create statistics of interest pertaining to uncertainty. While our approach is generic we focus on robust portfolio allocation problems as an application and study the impact of risk versus uncertainty in a dynamic mean-variance setting. We show that a mean-variance optimizing investor achieves economically meaningful wealth gains (33%) across our sample from 1996-2019 by internalizing our uncertainty measure during portfolio formation.

JEL Classification: N/A

Keywords: N/A

Eric Ghysels - eghysels@unc.edu
University of North Carolina and CEPR

Yan Qian - yanqian@live.unc.edu
University of North Carolina Chapel Hill

Steve Raymond - smr9@live.unc.edu
University of North Carolina Chapel Hill

Ambiguity with Machine Learning: An Application to Portfolio Choice*

Eric Ghysels† Yan Qian‡ Steve Raymond§

First draft: January 11, 2020

This version: October 27, 2021

Abstract

To characterize ambiguity we use machine learning to impose guidance and discipline on the formulation of expectations in a data-rich environment. In addition, we use the bootstrap to generate plausible synthetic samples of data not seen in historical real data to create statistics of interest pertaining to uncertainty. While our approach is generic we focus on robust portfolio allocation problems as an application and study the impact of risk versus uncertainty in a dynamic mean-variance setting. We show that a mean-variance optimizing investor achieves economically meaningful wealth gains (33%) across our sample from 1996-2019 by internalizing our uncertainty measure during portfolio formation.

*We thank seminar participants at UNC Chapel Hill for valuable comments and feedback.

†Department of Economics, University of North Carolina Chapel Hill, Department of Finance, Kenan-Flagler Business School and CEPR.

‡Department of Economics, University of North Carolina Chapel Hill.

§Kenan-Flagler Business School, University of North Carolina Chapel Hill.

1 Introduction

Decision makers who are averse to ambiguity often start with a family of candidate models they might consider as reasonable or plausible. In this paper we propose to use machine learning to impose guidance and discipline on the formulation of model families. Moreover, in a data rich environment it is fair to say that model uncertainty might be even more of an acute problem compared to the traditional setting with only a few predictors. For this reason, machine learning is also a natural way to proceed.

While the ideas explored in our paper are general, we focus on the cross-section of asset returns because of the abundance of both data and potential models. More specifically, we consider an investor who is intrigued by the appeal of artificial intelligence as guidance to portfolio allocation but is worried about model uncertainty, despite the sophistication of the statistical learning methods being used.[1] Hence, our investor sees opportunities with high-dimensional data but acts as an uncertainty averse Knightian decision maker using the concepts of ambiguity (or uncertainty) in the sense of Knight (1921) and aversion towards such uncertainty.[2] According to Knight (1921) there is *risk*, which corresponds to situations in which all relevant events are associated with a (objectively or subjectively) uniquely determined probability assignment, and *uncertainty* – often called Knightian uncertainty – which corresponds to situations in which some events do not have an obvious probability assignment.

There already exists a substantial body of literature pertaining to the decision-theoretic foundations of (portfolio choice with) ambiguity (see Section 2 for further discussion). To make the ideas operational one has to solicit and formulate preferences, meaning (a) attitudes towards risk (for a given distribution) and (b) concerns about ambiguity. Besides the preferences, one also has to characterize the menu of probability distributions, i.e. models,

[1] There are now a number of papers documenting the benefits of machine learning to predict, in particular stock returns, see e.g. Rapach, Strauss, and Zhou (2010), Kim and Swanson (2014), Gu, Kelly, and Xiu (2020), D'Hondt, De Winne, Ghysels, and Raymond (2020), among others.

[2] We will use the terms (Knightian) uncertainty and ambiguity interchangeably in this paper.

to be considered. Our paper has a number of innovations. The first, already mentioned, is the use of machine learning. The second is the use of bootstrap histories - in our application asset returns (and covariates) similar to Sullivan, Timmermann, and White (1999) and White (2000) - to construct empirical measures of uncertainty.

One can think of uncertainty in terms of (a) parameter uncertainty, (b) variable selection uncertainty and (c) functional form uncertainty. We cover all three. Suppose we fix the functional form to be linear, but don't know which variables to select and the value of the associated parameters. The machine learning-bootstrap approach involving regularized linear models accounts for both model uncertainty and variable selection uncertainty in a data rich environment. If we throw functional form uncertainty into the mix, we need to augment to class of models to including procedures involving deep learning or say random forests. We cover all these possibilities. Moreover, particularly practical and appealing is the fact that our methodology applies to any combination of assets considered in the portfolio exercise and produces in the multi-asset case a matrix-based uncertainty measure. Moreover, the approach put forward in our paper applies to any data-rich environment beyond the financial application considered. For example, a number of attempts were made to capture model uncertainty using survey data, see e.g. Anderson, Ghysels, and Juergens (2009) and Bhandari, Borovička, and Ho (2016), among others. The use of survey data has limitations, however, as such data only covers a small set of predictions. Our approach could notably be used for various macroeconomic applications where concerns about robust policy making have been widely discussed (see e.g. Hansen and Sargent (2011)).

Intuitively speaking the procedure we propose creates plausible synthetic bootstrap samples of data not seen in historical real data. The machine learning procedure can be thought of as creating reduced form prediction models, which with a sufficient number of bootstrap iterations allows us to create statistics of interest pertaining to uncertainty. While the procedure is intuitively appealing, we also note that there are a number of technical issues which remain unresolved and, although briefly discussed in the paper, are left to future research. The contribution of our paper is to lay out the procedure and show that it provides economically plausible and significant gains in asset allocation decisions.

The remainder of the paper is organized as follows. In Section 2 we discuss ambiguity in the context of portfolio choice, which is the application of interest in the paper. Sections 3 and 4 cover machine learning methods and the bootstrap procedure respectively. Both sections pertain to the main contributions of our paper. The empirical application appears in Section 5 followed by conclusions.

2 Portfolio choice and ambiguity

The thought-provoking paradox formulated by Ellsberg (1961) highlighted the distinction between risk and uncertainty and prompted researchers to search for new preference classes to accommodate Knightian uncertainty. The focus of our paper is portfolio choice, which of course is a dynamic decision-making problem.[3] Starting from the idea that models are approximations, Epstein and Wang (1994) use the Ellsberg (1961) paradox to motivate a decision theory for dynamic settings based on the minimax principle with multiple priors of Gilboa and Schmeidler (1989).

More precisely, Gilboa and Schmeidler (1989) built the axiomatic foundations for minmax ideas formulated by, among others, Wald (1950) and Hurwicz (1951) which can be written as follows. Let $f \succsim g$ mean that the decision maker weakly prefers decision f to g, then:

$$f \succsim g \text{ if and only if } \min_{p \in \mathcal{P}} \mathbb{E}_p[U(f)] \geq \min_{p \in \mathcal{P}} \mathbb{E}_p[U(g)], \tag{1}$$

where $\mathcal{P}$ is a convex set of probability measures over outcomes, whose size can be interpreted as representing the level of perceived ambiguity. The intuition is that – because of ambiguity aversion – agents are considering as valid and relevant the prior which is most unfavorable. Preferences represented by (1) have became commonly known as multiple prior preferences,

[3]Material in this section draws on the survey by Guidolin and Rinaldi (2013).

given the multi-valued nature of the set $\mathcal{P}$.

Anderson, Hansen, and Sargent (2003) and Hansen and Sargent (2001) noted the connection between Knightian uncertainty and the robust control theory used in engineering.[4] In that literature the set of probabilities $\mathcal{P}$ is characterized by taking a single "approximating model" and statistically perturbing it, or more precisely:

$$f \succsim g \text{ if and only if } \min_{q \in \Delta(\Omega)} \mathbb{E}_q[U(f) + \varrho R(q \| p)] \geq \min_{q \in \Delta(\Omega)} \mathbb{E}_q[U(g) + \varrho R(q \| p)], \tag{2}$$

where p is the reference or baseline approximate model, $\Delta(\Omega)$ is the simplex (i.e. probability densities) based on Ω which represents the states of the world. The Kullback and Leibler (1951) (KL) divergence measure $R(q\|p)$ is used to assess the distance between any candidate density/model q vis-à-vis p. The parameter ϱ controls the appetite for robustness or model uncertainty and can be thought of as an ambiguity aversion index since it measures the fear of model misspecification. The lower is ϱ, the higher is the degree of ambiguity aversion and $\varrho R(q\|p)$ is a penalty term. The decision maker considers a range of models $q \in \Delta(\Omega)$ with a weight assigned according to the KL distance to the baseline model p. Hansen and Sargent (2001) also show that the decision ordering scheme appearing in (2) can be reformulated as a constrained optimization of the following type:

$$f \succsim g \text{ if and only if } \min_{q \in \mathcal{B}} \mathbb{E}_q[U(f)] \geq \min_{q \in \mathcal{B}} \mathbb{E}_q[U(g)] \qquad \mathcal{B} = \{q \in \Delta(\Omega) : R(q\|p) \leq \eta\}, \tag{3}$$

where $\mathcal{B}$ is a "ball" centered at the baseline model p and containing models with maximal KL divergence of η. These robust preferences have been axiomatized by Maccheroni, Marinacci, and Rustichini (2006) who show that they are in fact a specific sub-class of so called variational preferences (VP). VP nest many of the known ambiguity aversion preferences structures, including the Monotone Mean Variance Preferences (MMVP) that extend outside their domain of monotonicity the classical Mean Variance Preferences of Markowitz

[4]See Hansen and Sargent (2011) for an elegant and detailed overview of their many papers on the topic.

(1959).

In this paper f and g pertain to portfolio choices. Assuming there are N risky assets and a risk-free one, this means decisions are represented by w_t, the $N \times 1$ vector of asset allocations among the risky assets at time t where $w_t \in \mathbb{R}^N \; \forall \; t$ with the constraint that $\iota_N^\top w_t = 1$, with ι_N the $N \times 1$ vector of ones.[5] This leaves us with the characterization of: (a) the utility function $U(\cdot)$ and (b) the set of probability functions $\mathcal{P}$ and resulting expectations $\mathbb{E}_p[U(\cdot)]$ for each $p \in \mathcal{P}$. Regarding the probability densities $\mathcal{P}$ we rely on a commonly used assumption in the literature, which relates to a discrete time version of the often used diffusion setting with uncertainty regarding the drift, see e.g. the textbook coverage by Hansen and Sargent (2011):

Assumption 2.1. *Let r_t be the $N \times 1$ vector of excess returns for N assets in period t and let $\mathcal{F}_{t-1}$ be the σ-field information filtration, then all $p \in \mathcal{P}$ are conditionally Gaussian $N(\mu_t^p, \Sigma_t)$.*

Assumption 2.1 tells us that all densities in the set $\mathcal{P}$ are conditionally Gaussian and share a common covariance matrix Σ_t. Therefore, ambiguity is concentrated on the vector of expected returns μ_t^p. Following the seminal work by Markowitz (1959) we use quadratic utility functions and solve for the so called optimal mean variance (MV) portfolio weights for a given density p :

$$\max_{w_t} \left[w_t^\top \mu_t^p - \frac{\gamma}{2} w_t^\top \Sigma_t w_t \right] \qquad \forall t. \tag{4}$$

Such preferences might be either perceived as reflecting attitudes towards risk, as in the original setup of Markowitz, or with suitable modifications discussed shortly may represent more sophisticated implementations which account for ambiguity.

Garlappi, Uppal, and Wang (2007) study an investor with multiple priors and aversion to ambiguity. They characterize the multiple priors by a "confidence interval" around the estimated expected returns and model ambiguity aversion via a minimization over such pri-

[5] For expositional purpose we ignore here short-selling constraints, whereas in the empirical implementation we impose them.

ors. Although confidence intervals or significance levels are often associated with hypothesis testing in statistics, Bewley (1988 and 2011) shows that they can also be interpreted as a measure of the level of uncertainty associated with the parameters estimated.

The empirical analysis in our paper follows more closely Maccheroni, Marinacci, and Ruffino (2013) who exploit the smooth model of decision making under ambiguity of Klibanoff, Marinacci, and Mukerji (2005) to derive a tractable mean-variance model adjusted for ambiguity. In their setting, risk aversion determines the decision maker's reaction to expected volatility and model uncertainty aversion determines her reaction to the variance of μ_t^p's, $p \in \mathcal{P}$. In their framework, the optimal MV portfolio allocation with Knightian uncertainty takes the form (see equation (23) in Maccheroni, Marinacci, and Ruffino (2013)):

$$\max_{w_t} \left[w_t^\top \mu_t^p - \frac{\gamma}{2} w_t^\top \Sigma_t w_t - \frac{\delta}{2} w_t^\top \Lambda_t w_t \right], \tag{5}$$

where (a) Λ_t is the variance matrix of expected returns, measuring the dispersion among all probability densities q considered, (b) γ is the relative risk-aversion parameter and (c) δ is uncertainty-aversion parameter. The solution to the above optimization problem is:

$$w_t = (\gamma \Sigma_t + \delta \Lambda_t)^{-1} \mu_t^p, \tag{6}$$

and the challenge is how to compute Λ_t. As an aside, we will estimate Σ_t via standard methods, such as for example Engle (2002), and denote its estimates by $\hat{\Sigma}_t$. In the remainder of the paper, our main focus is how to estimate Λ_t.

3 Machine learning methods

The investor operates in a data-rich environment. This means that there is potentially ambiguity about (a) the economic variables to select and (b) the functional form used to predict returns given these variables and finally (c) the parameters to use in the selected models. We focus exclusively on univariate models for each asset i in our portfolio. To simplify notation, we drop the subscript i referring to the asset and proceed with a single asset case.

We start with machine learning methods which assume a linear regression functional form, which means that ambiguity pertains to variable selection and parameter uncertainty. We consider a number linear machine learning methods, namely: Ridge, LASSO, Elastic Net, SCAD, and MCP. Next, we endeavor into functional form ambiguity, in addition to variable selection and parameter uncertainty. Here we consider three nonlinear machine learning models: random forests, gradient boosted trees and artificial neural network. A subsection is devoted to each.

3.1 Linear models

We will use a generic notation $\mu_t^m(z_t, \theta_m)$ to denote the expected returns where m identifies the method/model being used, z_t are the covariates, and θ_m the model parameters, including regularization and/or tuning parameters, required in the specification of m. For example, $\mu_t^{\ell_1}(z_t, (\beta(\lambda_{\ell_1}), \lambda_{\ell_1}))$ for LASSO with λ_{ℓ_1} regularization parameter involving a linear regression with parameters $\beta(\lambda_{\ell_1})$ which determine both variable selection and estimation of parameters for selected regressors.

The linear models we estimate for each asset across time periods $t = k, \ldots, T-1$ are of the form:

$$r_{t+1} = z_t'\beta + \epsilon_{t+1},$$

where z_t is a P-dimensional vector of predictors. Instead of estimating this model using OLS, we append a penalty to the original loss function:

$$\mathcal{L}(\beta; \cdot) = \frac{1}{T} \sum_{t=1}^{T} \epsilon_t^2 + \sum_{j=1}^{P} \phi(\beta_j; \cdot).$$

There are several choices for the penalty function $\phi(\beta_j; \cdot)$. We consider three popular types of penalty: Elastic Net, SCAD and MCP. The Elastic Net penalty, proposed by Zou and Hastie (2005), takes the form: $\phi_E(\beta_j; \lambda, \alpha) = \frac{1}{2}\lambda(1-\alpha)\sum_{j=1}^{P}|\beta_j| + \lambda\alpha\sum_{j=1}^{P}\beta_j^2$. It involves two non-negative hyperparameters, λ and α, and includes two regularizers as special cases. The $\alpha = 0$ case corresponds to LASSO (Tibshirani (1996)) and uses an ℓ_1 parameter penalization. LASSO imposes sparsity and sets coefficients on a subset of predictors to exactly zero. The $\alpha = 1$ case corresponds to Ridge (Hoerl and Kennard (1970)), which uses an ℓ_2 parameter penalization. Ridge regression draws all coefficient estimates closer to zero but does not impose exact zeros. For intermediate values of α, the Elastic Net encourages simple models through both shrinkage and variable selection.

The smoothly clipped absolute deviation (SCAD), proposed by Fan and Li (2001), has a penalty function given by:

$$\phi_S(\beta_j; \lambda, \alpha) = \begin{cases} \lambda|\beta_j|, & \text{if } |\beta_j| \leq \lambda \\ \frac{2\alpha\lambda|\beta_j| - |\beta_j|^2 - \lambda^2}{2(\alpha-1)}, & \text{if } \lambda < |\beta_j| \leq \alpha\lambda \\ \frac{\lambda^2(\alpha+1)}{2}, & \text{if } |\beta_j| > \alpha\lambda \end{cases}$$

for $\alpha > 2$ and $\lambda > 0$. For small signals $|\beta_i| < \lambda$, it acts like LASSO; while for larger signals $|\beta_i| \geq \alpha\lambda$, the penalty flattens and leads to the unbiasedness of the regularized estimate.

The minimax concave penalty (MCP), developed by Zhang (2010), takes the form:

$$\phi_M(\beta_j; \lambda, \alpha) = \begin{cases} \lambda|\beta_j| - \frac{|\beta_j|^2}{2\alpha}, & \text{if } |\beta_j| \leq \alpha\lambda \\ \frac{\alpha\lambda^2}{2}, & \text{if } |\beta_j| > \alpha\lambda \end{cases}$$

where $\lambda > 0$ amd $\alpha > 1$. The penalty function is quadratic on $[0, \alpha\lambda]$ and flattens beyond $\alpha\lambda$. MCP starts with the same rate of penalization as LASSO but relaxes the penalization rate as the absolute value of the coefficient increase. Finally, we estimate the regularized regression models on the training data (details appear later) and optimize all respective tuning parameters via 10-fold cross validation. The collection of predictions from the class of regularized linear models will be denoted by $\mu_t^{\ell_1}$ for LASSO, $\mu_t^{\ell_2}$ for Ridge, μ_t^{ENET} for Elastic Net, and finally μ_t^{MCP} and μ_t^{SCAD} for MCP and SCAD.

3.2 Nonlinear models

Regression tree type models Regression tree is a popular nonparametric machine learning model designed to incorporate nonlinear interactions between covariates. The prediction of a tree with K terminal nodes (called 'leaves') and depth L, can be written as: $g(z_t; \beta, K, L) = \sum_{k=1}^{K} \beta_k \mathbf{1}_{\{z_t \in C_k(L)\}}$, where $C_k(L)$ is the k-th partition that has at most L different branches. A set of branches for a given partition can be represented as a product of indicators for sequential branches. For a given partition, then $\hat{\beta}_k$ is the average of returns for all members of that given partition. We follow the algorithm of Breiman, Friedman, Stone, and Olshen (1984). At each step of the tree, we choose a sorting variable from the set of predictors and the split value to maximize the information (impurity) gained. The recursive binary splitting algorithm continues until a set of stopping criterion are met, which typically rely on the maximal additional information gained from a split being less than a threshold, or a max number of leaves and/or depth of a tree being reached.

Regression trees often suffer from high variance. We consider two regularization methods for improving the predictions: Random Forest and Gradient Boosting. Random Forest averages forecasts from many different trees. We employ the algorithm of Breiman (2001) to estimate random forest models: each tree is built from a bootstrapped training sample (bagging). In addition, the method de-correlates trees by only choosing a randomly drawn subset of predictors for splitting at each potential branch. For the random forest model, we fix the number of bootstrapped trees so that the key hyper parameters are the predictor subset size S and the depth of each tree L.

Boosting recursively combines forecasts from many weak learners (small regression trees) to form a 'strong learner'. In boosting, trees are grown sequentially. We adopt the Friedman (2001) algorithm to fit gradient boosted trees: at each new step b, a shallow tree is fitted to the residuals from the model with $b-1$ trees, and its residual forecast is added to the total with a shrinkage weight. For the gradient boosting models, we tune the number of trees, the depth of each tree and learning rate λ through cross validation.

Deep learning We focus our deep learning analysis on traditional "feed-forward" neural networks. These consists of an "input layer" of raw predictors, one or more "hidden layers" and an "output layer". Let $K^{(l)}$ denote the number of neurons in each layer $l = 1, \ldots,$ L and $x_k^{(l)}$ denote the output of neuron k in layer l. Also define the vector of outputs for layer l as $x^{(l)} = (1, x_1^{(l)}, ...x_{K^{(l)}}^{(l)})'$. The network is initialized using the raw predictors, $x^{(0)} = (1, z_1, \ldots, z_N)'$. Each neuron in the "hidden layer" draws information linearly from all units in the previous layer and applies a nonlinear "activation function" to its aggregated signal before sending its output to the next layer: $x_k^{(l)} = h(x^{(l-1)\prime}\beta_k^{(l-1)})$, where h is an activation function. In the last layer, the results from each neuron are linearly aggregated into the ultimate output forecast: $g(z;\beta) = x^{(L-1)\prime}\beta^{(L-1)}$. The number of weight parameter in each hidden layer l is $K^{(k)}(1 + K^{(l-1)})$, plus another $1 + K^{(L-1)}$ weights for the output layer.

Our network architecture has one hidden layer and four neurons. All architectures are fully connected so each unit receives an input from all units in the layer below. We use rectified linear unit (ReLU) as the activation function and gradient descent with adaptive learning rate backpropagation algorithm to minimize a regularized ℓ_1 objective function:

$$\mathcal{L}(\beta;\lambda) = (1-\lambda)\frac{1}{T}\sum_{t=1}^{T}\epsilon_{i,t+1}^2 + \lambda\frac{1}{P'}\sum_{j=1}^{P'}\beta_j^2 \tag{7}$$

where $0 \leq \lambda < 1$ and P' is the total number of weight parameters. In addition to ℓ_1 penalization of the weight parameters, we simultaneously employ three other regularization techniques in our estimation: learning rate shrinkage, early stopping, and ensembles. For each deep learning tree, we tune the penalization parameter λ through 10-fold cross validation. The class of nonlinear models yields predictions μ_t^{RF} for random forest, μ_t^{GBRT} for gradient boosting and finally μ_t^{NN} for neural nets.

4 Bootstrap estimation of model uncertainty

Hastie, Tibshirani, and Friedman (2017, Section 7.11) in their widely used textbook on machine learning characterize bootstrap as a general tool used for assessing statistical accuracy used in particular to understand the properties of model prediction errors. Not surprisingly, several authors have suggested the use of bootstrap methods in machine learning settings, recent examples include Bach (2008), Chatterjee and Lahiri (2011), Pearce, Zaki, Brintrup, and Neely (2018), Li, Luo, Ferrari, Hu, and Qin (2019), Kostrikov and Nachum (2020), among others. It is also worth noting that bootstrapping is one of many resampling methods used in the machine learning literature with the widely used cross-validation method as a close cousin.

Here we use it to construct empirical measures of uncertainty, a novel feature of our paper. To that end we draw $B = 1{,}000$ bootstrap samples of size T for the entire cross-section of size N. Given the nature of the data, we follow the stationary bootstrap of Politis and Romano (1994), commonly used to generate financial returns data, see e.g. Sullivan, Timmermann, and White (1999). More precisely, for each $b = 1, \ldots, B$ we generate the panel of individual asset returns r_{it} and covariates Z_{it} for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. In our empirical study, we set the average resampling block to be 12 months. Let $i_m = 1, \ldots, M$ be the set of models. The models are grouped in G groups with $i_g = 1, \ldots, G$. In this paper, we will work with the entire set of machine learning models as well as two groups. One group consists of all regularized linear regression methods yielding $\mu_t^{\ell_1}$, $\mu_t^{\ell_2}$, μ_t^{ENET}, μ_t^{MCP}, μ_t^{SCAD}, the other group consists of nonlinear methods μ_t^{RF}, μ_t^{GBRT}, μ_t^{NN}.

Recall that, following Anderson, Hansen, and Sargent (2003), Hansen and Sargent (2001), among others, we are looking for the set of probabilities $\mathcal{P}$ characterized by statistical perturbations of a single approximating model. Given Assumption 2.1, we care about a set of conditional mean perturbations with respect to a reference model. This is achieved via the bootstrap. More specifically, for each draw $b = 1, \ldots, B$ we estimate across all N assets for $t = T_0$, $\ldots, T$, all models – which means applying the training and validation sample schemes across all expanding samples, assets, and methods. This yields $\left(\mu_{it}^{i_m,b}; t = T_0, \ldots, T, i = 1, \ldots, N\right)$, panels of expected returns for models $i_m = 1, \ldots, M$ and bootstrap draws $b = 1, \ldots, B$. To make this clear, each model is estimated using the bootstrap sample and the real data is used to generate $\mu_{it}^{i_m,b} = \mu_{it}^{i_m,b}(Z_t, \hat{\theta}_m^b)$.

Using these data sets, and collecting for all assets i the expected returns $\mu_{it}^{i_m,b}$ into an N-dimensional vector $\mu_t^{i_m,b}$, we compute the following:

$$
\begin{aligned}
\hat{\Lambda}_{i_m,t}^B &= \left[\frac{1}{B}\sum_{b=1}^{B}(\mu_{it}^{i_m,b} - \bar{\mu}_{it}^{i_m,B})(\mu_{jt}^{i_m,b} - \bar{\mu}_{jt}^{i_m,B}) + \left[\mu_{it}^{i_m,p} - \bar{\mu}_{it}^{i_m,B}\right]\left[\mu_{jt}^{i_m,p} - \bar{\mu}_{jt}^{i_m,B}\right]\right]_{i,j=1}^{N} \\
&= \frac{1}{B}\sum_{b=1}^{B}(\mu_t^{i_m,b} - \bar{\mu}_t^{i_m,B})(\mu_t^{i_m,b} - \bar{\mu}_t^{i_m,B})^\top + \left[\mu_t^{i_m,p} - \bar{\mu}_t^{i_m,B}\right]\left[\mu_t^{i_m,p} - \bar{\mu}_t^{i_m,B}\right]^\top \\
\bar{\mu}_{it}^{i_m,B} &= \frac{1}{B}\sum_{b=1}^{B}\mu_{it}^{i_m,b}, i = 1, \ldots, N \qquad (8)
\end{aligned}
$$

for a specific model class i_m. The second term on the right hand side of the above equation is a bias correction which re-centers the uncertainty at the expected returns $\mu_{it}^{i_m,p}$ and $\mu_{jt}^{i_m,p}$ for assets i and j for the reference model p for the class of models i_m. In our empirical study, we choose the reference model p such that its asset i expected return $\mu_{jt}^{i_m,p}$ is the prediction with model i_m obtained from the real data sample. Hence, the uncertainty measures are centered around the expected return series used in the portfolio optimization problem.

Take for example an investor who only looks at regularized regression models using LASSO with a large set of covariates Z_t. For $\mu_t^{\ell_1}$ we explore model uncertainty associated with variable selection as well as parameter estimation. The empirical sample yields an estimate, in this case using LASSO, for the expected return for each asset i, namely $\mu_{it}^{\ell_1,p}$. Using the LASSO procedure applied to the B bootstrap samples for each asset i yields a distribution of expected returns $[\mu_{it}^{\ell_1,b}, b = 1, \ldots, B]$ and yields the empirical measure $\hat{\Lambda}_{\ell_1,t}^B$ for the case of LASSO. All other model procedures proceed along the same lines.

If our investor uses a particular group of machine learning models, she would compute:

$$\hat{\Lambda}^{B}_{i_g,t} = \left[\frac{1}{B\times N_{i_g}}\sum_{i_m\in i_g}\sum_{b=1}^{B}(\mu^{i_m,b}_{it}-\bar{\mu}^{i_g,B}_{it})(\mu^{i_m,b}_{jt}-\bar{\mu}^{i_g,B}_{jt})+\left[\mu^{i_g,p}_{it}-\bar{\mu}^{i_g,B}_{it}\right]\left[\mu^{i_g,p}_{jt}-\bar{\mu}^{i_g,B}_{jt}\right]\right]^{N}_{i,j=1}$$

$$\bar{\mu}^{i_g,B}_{it} = \frac{1}{B\times N_{i_g}}\sum_{i_m\in i_g}\sum_{b=1}^{B}\mu^{i_m,b}_{it}, i=1,\ldots,N \qquad (9)$$

where N_{i_g} is the number of models in group i_g. Here we re-center the uncertainty at the reference measure p expected returns $\mu^{i_m,p}_{it}$ using the real data averaged across all models in i_g, yielding $\mu^{i_g,p}_{it}$. The fact that there are more linear models in the mix may tilt the measure defined in equation (9) toward that particular group. In Appendix Section OA.3, we show that our main findings continue to hold when the group of linear models and the group of nonlinear models contribute equally to the expected return and composite uncertainty estimates.

In this paper, the linear uncertainty estimator is calculated only using linear models and the nonlinear uncertainty estimator is calculated only using nonlinear models. Finally, an agnostic investor taking every model under consideration would consider the composite uncertainty estimator with the N-vector $\bar{\mu}^{M,B}_{t}$ with elements $\bar{\mu}^{M,B}_{it} = \frac{1}{B\times M}\sum_{i_m=1}^{M}\sum_{b=1}^{B}\mu^{i_m,b}_{it}$:

$$\hat{\Lambda}^{B}_{M,t} = \frac{1}{B\times M}\sum_{i_m=1}^{M}\sum_{b=1}^{B}(\mu^{i_m,b}_{t}-\bar{\mu}^{M,B}_{t})(\mu^{i_m,b}_{t}-\bar{\mu}^{M,B}_{t})^{\top}+\left[\mu^{M,p}_{t}-\bar{\mu}^{M,B}_{t}\right]\left[\mu^{M,p}_{t}-\bar{\mu}^{M,B}_{t}\right]^{\top} \qquad (10)$$

It is quite standard in the machine learning literature to introduce a procedure or algorithm, show that it works, but not necessarily establish all its theoretical properties. The theoretical foundations of our procedure starts with the work of White (2000). While his emphasis is data snooping and hypothesis testing, his framework shares features with our procedure. In particular, he is interested in estimating the covariance matrix for a number of model predictions (see section 2.b in White (2000)). His setting is different, as he is interested in extreme values of correlated normally distributed random variables (cfr. our Assumption 2.1). His theoretical framework is inspired by the standard GMM asymptotics of Hansen (1982). It is a serious theoretical undertaking to establish the validity of the same procedure in a high-dimensional data setting in particular for all the machine learning procedures covered in Section 3. For a starter, the development of GMM-type estimators in a machine learning context is still not fully developed, see e.g. Belloni, Chernozhukov, Chetverikov, Hansen, and Kato (2018) for a recent discussion. Notably challenging is the role of the weighting matrix (see e.g. Caner and Kock (2019) for results pertaining to linear GMM with LASSO regularization), an object similar to our measures of uncertainty. It would require many novel contributions beyond the scope of the current paper. The main attraction of the procedure is its versatility and intuitive appeal and more importantly, as we show in the remainder of the paper, yields economically plausible estimates of uncertainty as well as economically significant improvements in asset allocation with internalized concern for model uncertainty.

5 Empirical Study

We start with a description of the data used in our analysis, which is the subject of a first subsection. A second subsection documents the performance of the empirical models. Next we study the properties of the uncertainty measures obtained through the bootstrap procedure presented. A fourth subsection covers portfolio allocation with uncertainty. Finally we conclude with a subsection on the impact of uncertainty and risk shocks on the portfolio weights.

5.1 Data and sampling

The asset return and predictors data are obtained from a variety of sources. All the series are monthly, and the details are provided in Section OA.1 of the Online Appendix. The sample period spans from January 1986 to December 2019. The predictions are computed via an expanding window scheme. Effectively, the models are re-trained annually to predict the monthly asset returns next year. The set of predictors considered for the purpose of model selection, can be categorized along the following lines.

Macroeconomic series: Macroeconomic data are typically subject to future revisions and released with delay. Ghysels, Horan, and Moench (2018) find that a sizable fraction of the predictive information contained in final macro data is carried by data revisions. In order to avoid introducing forward-looking bias, we use real-time macro series vintages from the ALFRED database, with the exception of interest rates and oil prices.

Financial series: We use a set of financial predictors. Namely, asset-pricing factors related to the Fama and French (2015) five factor models are from the data library of Kenneth French. In addition, we use the financial predictors constructed by Welch and Goyal (2008) and the CBOE Volatility Index (VIX) of Chicago Board Options Exchange.

Textual data: We add news attention data based on textual analysis available from Nick Bloom's data library. More specifically, components of the Economic Policy Uncertainty Index and Categorial EPU data constructed by Baker, Bloom, and Davis (2016) are candidate predictors, as well as category-specific Equity Market Volatility trackers provided by Baker, Bloom, Davis, and Kost (2019).

Asset returns: A portfolio of eight assets is considered: Monthly S&P 500 Index Return and Fama 60 month-120 month bond portfolio return data from the Center for Research in Security Prices (CRSP), Fama-French 5 industry portfolio returns from the data library of Kenneth French, and Gold return series obtained from FRED.

We have a total of 134 regressors which are used to build the various prediction models. The first models are trained using data from January 1986 until December 1995 and used to predict asset returns in 1996, so there are 120 observations in the first training sample and the prediction sample is 12 months, and the predictions are computed according to the data of December 2005. Moving forward with yearly increments, we calculate predictions until the sample is exhausted in December 2019.

5.2 Prediction model performance

It is of independent interest to report on the empirical performance of the various models in predicting the returns of the eight assets in the portfolio. The bulk of the literature has focused on equity return predictions using machine learning. In our mix of assets we have fixed income and gold which have received less attention in the literature. In Table 1 we report the out-of-sample (OOS) mean squared errors (MSE) of the linear and nonlinear prediction models. Ignoring for the moment the last row labeled *EW Average* we note that for the S&P 500 index the linear models are better than the nonlinear ones and among the linear ones LASSO, MCP and SCAD are the best, although overall the differences in OOS MSE are small. The last row pertains to the equally weighted average of all the models and is used as the reference model in the portfolio optimization around which we build the model uncertainty measures. Continuing with equities, we note that the health industry portfolio OOS MSE are similar to those of the market index, Consumer has slightly higher prediction errors whereas the Hi Tech industry returns are the most difficult to predict. In all cases, we find that linear models yield the best OOS return predictions. Among the nonlinear models Deep Learning in particular produces the worst whereas Random Forest and Gradient Boosting are typically only slightly worse than the linear models. Fixed income returns are in comparison easy to predict and gold returns are easier to predict than equities. Linear models turn out to be the best again. It is also worth paying attention to the last row in Table 1 where we report the OOS MSE of linear models estimated via OLS. We note that the performance is orders of magnitude worse that any of the machine learning methods. For this reason we opt to ignore OLS estimation in the remainder of the paper.

Table 2 provides the OOS MSE statistics for the bootstrap data. We note that the prediction performance in the synthetic data is much worse than with the real data and nonlinear models turn out to be better than linear models. In fact the OOS MSE for the real data and bootstrap histories are comparable for nonlinear models, but much worse for linear models. This should not be of major concern as for the uncertainty measures we care mostly about deviation from average performance for a certain type or class of model(s).

Table 3 provides some insights about the top 10 most selected regressors by the linear machine learning models for each of the eight assets considered in the portfolio. We are reminded of the fact that machine learning methods are black box procedures and may uncover some surprising patterns. This being said, among the top 10 most frequently selected regressors for each of the eight assets there are many that make perfect sense. For example real personal consumption, price dividend ratio, stock variance, some Fama-French factors are among the top regressor for the equity market index as well as for the industry equity portfolios. The lagged bond yield predict long term bond returns and mine industry returns predict gold prices is another example. Gold is also predicted by the CPI and a measure of inflation-related volatility. There are a few unexpected ones, perhaps most prominently is the infectious disease EMV. This is a volatility tracker (EMV stands for equity market volaility) based on textual analysis constructed by Baker, Bloom, and Davis (2016). Note that the pandemic is not part of our sample and therefore this is not caused by the unusual market movements in recent history. Online Appendix Table OA.3 provides a list of variable definitions for the textual data and it shows that there are in fact more than 30 EMV variables. It is therefore somewhat peculiar that this particular tracker is being selected so often for predicting S&P 500 returns. Note that it also appears prominently among the health industry covariates, which may seem less anomalous. Several of the other EMV trackers appear among the selected regressors, as well, including the top predictor for the S&P 500 index. Overall, though it appears that the prediction models obtained via the bootstrap procedure appear as being genuine. Table 4 displays the same information for the bootstrap samples. Overall, it seems that covariates are generally picked more often in the synthetic data as they are in the real data, judging by the percentages appearing next to each of the selected predictors. The top 10 most selected regressors by the linear models, with the real data share many predictors with the top 10 most selected regressors with the bootstrap data.

5.3 Uncertainty measures and their properties

Figure 1 shows the time series plots of the Frobenius norm of the uncertainty measure $\|\hat{\Lambda}^B_{i_g,t}\|_F$ appearing in equation (9) for i_g representing the group of all regularized linear prediction models against (a) the Economic Policy Uncertainty Index (EPU), (b) the financial uncertainty index (JLN), (c) the CBOE Volatility Index (VIX), and finally (d) the Frobenius norm of the DCC(1,1) covariance matrix $\|\Sigma_t\|_F$ following the estimation procedure of Engle (2002). The series in the top and bottom panels are normalized to unit length. The sample period is from January 1996 to December 2019.

More specifically, in the top panel, we compare the Frobenius norm of the composite uncertainty matrix with the Economic Policy Uncertainty Index (EPU) by Baker, Bloom, and Davis (2016), the financial uncertainty Index (JLN) by Jurado, Ludvigson, and Ng (2015) and the CBOE Volatility Index (VIX). Each uncertainty measure is normalized to be comparable. We find a close link between our uncertainty estimator and the time series behavior of the other measures. All four appear to be counter-cyclical and peak around the Great Recession of 2007-2009. According to our estimator, the uncertainty is also more pronounced during the dot-com bubble years.

Note that EPU is based on textual analysis, JLN is based on residual variation in reference to a common factor model for macro panel data and finally the VIX is based on derivative pricing. It is therefore quite remarkable that all measures feature such strong co-variation. The appeal of our measure, of course, is that it can be computed for any collection of assets, and can therefore target specific sectors of the economy. With the eight assets selected for our portfolio analysis, we cover a rather wide range of different asset classes, which may explain why it matches closely the broadly defined uncertainty measures.

In the central panel, we compare the Frobenius norm of the uncertainty matrix $\|\hat{\Lambda}^B_{i_g,t}\|$ for different groups of models, namely: (a) the group of regularized linear regression models as in the top panel, (b) only using the group of nonlinear models and (c) $\hat{\Lambda}^B_{M,t}$ appearing in equation (10) using the entire set of machine learning models. We observe that the uncertainty implied by the group of linear models is mostly larger and more volatile than for the nonlinear models. Moreover, the uncertainty for both the linear and nonlinear models tend to respond simultaneously, especially during financial crisis, with an overall correlation of 0.88.[6]

The lower panel of Figure 1 compares the Frobenius norm of $\|\hat{\Lambda}^B_{M,t}\|_F$ against that of $\hat{\Sigma}_t$, or put differently compares uncertainty versus risk with both normalized to be comparable. The time series pattern in the lower panel suggests that uncertainty tends to lead volatility. The uncertainty typically spikes one month ahead of volatility, with the maximum conditional autocorrelation for $k > 0$ equal to 0.58 with $k = 1$.

Table 5 reports the results regarding simple linear regressions of the three uncertainty measures, namely $\hat{\Lambda}^B_{M,t}$ and $\hat{\Lambda}^B_{i_g,t}$ for i_g respectively linear and nonlinear models, projected on the various measures using in the figure - EPU, JLN, VIX, Volatility - in addition to Industrial Production Growth (IP). We also report a pure time series AR(1) regression - which yields the highest R^2's. Panel A covering $\hat{\Lambda}^B_{M,t}$ shows that the slope coefficient estimates are mostly positive and statistically significant at the 5% level, except for the negative beta with IP as regressor and the insignificant slope estimate with EPU. The results for IP confirm the counter-cyclical nature of Knightian uncertainty. Panels B and C pertain to the model group-specific measures $\hat{\Lambda}^B_{i_g,t}$. The linear regression results confirm a strong relationship between our uncertainty estimators and other uncertainty measures. It is worth noting though that across all our measures, the regressions involving EPU result in near-zero R^2's.

To appraise the difference between risk and uncertainty we plot some of the off-diagonal

[6] In Appendix Section OA.3 we display the same plots as in Figure 1 but adjust the uncertainty measures for the unbalanced number of linear versus nonlinear models, using the correction appearing in equation (OA.2). Figure OA.1 indicates that our findings appear robust to the imbalance.

elements of $\hat{\Lambda}^B_{M,t}$ versus those of $\hat{\Sigma}_t$. In Figure 2 we plot a twelve month moving average of the former against its risk counterpart (which is smooth already as a filtered estimate). We express those off-diagonal elements in terms of correlations so that the scales are comparable. The shaded areas correspond to NBER recessions (in particular the Great Recession). The top panel covers all equity assets (namely S&P 500 and Fama-French five industry portfolios specified in Online Appendix Table OA.4) combined against the LT bond risk versus uncertainty. The latter is either positive or close to zero. This means that uncertainty about bond and stock returns typically co-move together. Risk starts out positive but turns negative for most of the sample. Hence, the conditional correlations between equity and bond return fluctuations tend to go in opposite direction. Uncertainty trends towards zero, however, except for upticks during recessions, which means stock and bond uncertainty appear to become more disentangled.

Moving the the second panel, we see that both uncertainty between equities and gold and risk are mostly negative. The third panel shows that uncertainty can flip signs as well, as we see that happen for bonds and gold. In this case conditional risk correlations are with only a few exceptions always positive.

In Table 6, we report the summary statistics of the Frobenius norm of our uncertainty measure for the entire sample and three subsamples: pre-crisis (from January 1996 to November 2007), crisis (from December 2007 to June 2009), and post-crisis (July 2009 to December 2019). Not surprisingly, uncertainty is high during the crisis, almost three times the overall sample mean. In contrast, post-crisis uncertainty is extremely low. The volatility of uncertainty is high during the crisis and low afterwards as well.

The comparisons with alternative measures, namely the Economic Policy Uncertainty Index of Baker, Bloom, and Davis (2016), the financial uncertainty Index of Jurado, Ludvigson, and Ng (2015) and the CBOE Volatility Index, may leave the impression that uncertainty is driven by a single factor. The plot in Figure 3 suggests otherwise, however. We plot the Frobenius norm of our uncertainty measure against a single uncertainty factor extracted from off-diagonal elements of $\hat{\Lambda}^B_{M,t}$, following Bai and Ng (2002). A single factor appears inadequate capturing the temporal variation in uncertainty.

5.4 Portfolio allocation with ambiguity

A number of papers have studied portfolio choice with uncertainty, including Garlappi, Uppal, and Wang (2007) and Anderson and Cheng (2016), among others. Our approach is distinct from the previous literature in important ways and offers advantages along several dimensions. First, our machine learning-bootstrap approach accounts for both model uncertainty and variable selection uncertainty in a data rich environment. Second, our methodology produces a matrix-based uncertainty measure. Third, we pin down the portfolio weights through a clean optimization problem by extending Maccheroni, Marinacci, and Ruffino (2013). Although the problem can only be solved numerically with weight constraints (short sale constraint for example), the unconstrained problem has a closed form solution with both a risk-return and uncertainty-return trade-off. Fourth, our uncertainty estimator can be applied to any asset class with time series data for the variable of interest and predictors.

Our prime focus is the extent of uncertainty changing a representative investor's optimal asset allocation in a mean-variance-uncertainty setting. Let us first recall the setting in equation (5), namely the investor solves portfolio weights vector w_t according to,

$$\max_{w_t} \left[w_t^\top \hat{\mu}_t - \frac{\gamma}{2} w_t^\top \hat{\Sigma}_t w_t - \frac{\delta}{2} w_t^\top \hat{\Lambda}_t w_t \right]$$

subject to $w_t \geq 0$, $i = 1, \ldots, N$ and $\sum_i^N w_t = 1$ where we add an additional short-sale constraint. The portfolio is re-balanced every month.

While we write the above equation in terms of generic estimators, we consider a rich class of models for expected returns, and associated model uncertainty measures while adopting a standard approach for risk. To streamline the analysis, we focus on uncertainty estimators calculated from (a) only using the group of linear machine learning models, (b) only using the group of nonlinear machine learning models, and (c) using both the linear and nonlinear models. For each of three cases, we apply the same group of machine learning models with the real data and then estimate $\hat{\mu}_t$ by averaging the return predictions.

In Tables 7 and 8, we compare the out-of-sample performance of the Markowitz with uncertainty portfolio ("UNC") with two benchmark portfolios: the standard Markowitz mean-variance ("MV") portfolio and the equally weighted ("EW") portfolio. The out-of-sample tests use the utility specification appearing in equation (5), hence the investor cares about both risk and uncertainty when (s)he optimizes portfolio weights. All portfolio strategies are based on the same out-of-sample inputs: (a) asset returns r_{t+1}, (b) conditional covariance matrices Σ_t and (c) uncertainty matrices Λ_t, while the optimal weights of course are different. Panel A reports the ratio of average out-of-sample utility with the composite uncertainty estimator ("Composite UNC") computed as in equation (10), using the entire set of machine learning models. In Table 7, the Composite UNC portfolio has a higher average utility than the MV portfolio under all relative risk aversion (γ) and uncertainty aversion parameter (δ) pairs, except the first column. This is because when $\delta = 0$, the investor is indifferent to the Knightian uncertainty so the Composite UNC is equivalent to the MV. In Panel A of Table 8 we present the ratio of average utility between the Composite UNC and the EW portfolios. It has been well documented that the EW portfolio often outperforms more complicated rules, especially for cases involving a large dimension of assets (see DeMiguel, Garlappi, and Uppal (2009)). With our uncertainty measure, the Composite UNC portfolio beats the EW portfolio in terms of the average out-of-sample utility, except when $\gamma = 2$ and $\delta < 2$.

The gains from internalizing uncertainty in the portfolio optimization are both economically meaningful and statistically significant. For example, the average utility of the Composite UNC portfolio is about 63% higher than of the MV portfolio and 38% higher than of the EW portfolio when $\gamma = 4$ and $\delta = 4$. For investors who are less uncertain averse ($\gamma = 4$ and $\delta = 2$) for example, they still benefit from the Composite UNC portfolio by having 27%

higher utility than the MV portfolio and 21% higher utility than the EW portfolio. The Diebold and Mariano (2002) statistic is used in Tables 7 and 8 to assess whether the utility differences are statistically significant. The null hypothesis - meaning equal average utility - is rejected at the 5% level for all specifications in Panel A of Table 7. We also reject the null hypothesis at 10% level for most values of γ and δ in Panel A of Table 8.

Next we turn our attention to Figure 4 where in the first panel we find a heat map for the equity, bond and gold weight differences between the "UNC" and the "MV" portfolio for the case where we set both γ and δ equal to four. The two other panels show results already discussed but repeated here for the purpose of facilitating the presentation and interpretation. The second panel displays the the Frobenius norm of $\|\hat{\Lambda}^B_{M,t}\|_F$ as well as the part that is orthogonal to the norm of $\hat{\Sigma}_t$. The third panel displays the off-diagonal elements of $\|\hat{\Lambda}^B_{M,t}\|_F$ appearing in each of the three subplots of Figure 2 into a single plot. The color code in the first panel is such that hot (red) means tilting toward a particular asset when comparing the "MV" versus "UNC" allocations. Looking across the three panels we see that bonds appear as hot when uncertainty spikes, and in particular when uncertainty not correlated with risk spikes. The shift occurs between equities and bonds, as the top equity panel turns cold (blue). There is also some movement into gold, but this looks more as short-lived bursts. From the third panel we learn, as noted earlier, that uncertainty comovements between stocks and bonds tend to tapper off toward zero at the end of the sample. This translates into the heatmap showing less frequent occurrences of shifts out of the equity markets into fixed income.

After documenting the superior performance of Markowitz with uncertainty portfolio, we now turn to addressing the following question: is there any difference in the portfolio performance when linear or nonlinear uncertainty estimators are used? In other words, the exercise needs to be repeated by substituting the uncertainty estimator with the one calculated using only the linear models or only the nonlinear models. We conduct this analysis for two reasons. First, given the substantial difference in the magnitude of linear uncertainty estimator and nonlinear uncertainty estimator, we wish to study whether our result is driven by a particular group of machine learning models. Second, investors may have a preference over one group of machine learning models over the other. For example, some investors may believe in regularized regressions while others may think that the nonlinear regression trees and deep learning methods are more suitable.

As seen in Panels B of Tables 7 and 8, using the linear uncertainty estimator does not affect our main results in any meaningful way. The "Linear UNC" portfolio has higher average out-of-sample utility under all (γ, δ) pairs, compared to the MV and the EW portfolio. In Panel C Tables 7 and 8, we find that the "Nonlinear UNC" portfolio outperforms the MV portfolio under all of (γ, δ) pairs and beats EW portfolio when the investor has a high degree of risk aversion ($\gamma > 6$). Therefore, the results in Panel A continue to hold in most cases when we restrict our model choices to one particular group of machine learning models.

In Figure 5 we plot the ratio of average out-of-sample utility between the UNC portfolio and the benchmarks over a grid of (γ, δ). In the top panel, the utility ratio between the "Composite UNC" portfolio and benchmarks is always greater than 1, indicating that the "Composite UNC" portfolio has a better risk and uncertainty adjusted performance. From the top left panel, we find that the utility ratio between the "Composite UNC" portfolio and the MV portfolio monotonically increases when the investor is more uncertainty-averse. On the other hand, when the investor is more risk-averse, the utility ratio between the "Composite UNC" portfolio and the MV portfolio decreases because the risk hedging incentives start to dominate in the portfolio optimization. In the top right panel, the utility ratio between the "Composite UNC" portfolio and the EW portfolio rises both when the δ and γ is higher. The result suggests that both the risk-return and the uncertainty-return trade-off channel bring benefits for the investor in terms of relative utility. For a suffciently high risk and uncertainty averse investor, the average utility of the "Composite UNC" portfolio can be almost 5 times higher than the EW portfolio. In the middle and bottom panel, we plot the utility ratio between the "Linear UNC" and "Nonlinear UNC" portfolio and the benchmark portfolios over a grid of (γ,δ), and the result is consistent with the findings with the composite uncertainty estimator. The only notable difference is that the utility ratio between the "Nonlinear UNC" portfolio and the "MV" portfolio first increases and then decreases when the investor is more risk-averse.

Finally, although our main result is robust to both different group of machine learning models, we do uncover differences across the different types of uncertainty estimators regarding the portfolio performance. In Table 8 we find that the "Linear UNC" performs the best, the "Nonlinear UNC" has the worst performance, and the "Composite UNC" is, not surprisingly, in between. In our case, the inclusion of nonlinear machine learning models actually lower the average out-of-sample utility of the investor.

In Figure 6, the cumulative return of three portfolios rules: the Markowitz with uncertainty (UNC) portfolio, the Markowitz mean variance (MV) portfolio and the Equal Weight (EW) portfolio are plotted. We set both the relative risk aversion γ and uncertain aversion δ to four, but our main conclusions hold for a relatively wide range of parameters. By investing \$100 in the initial period, Figure 6 shows the wealth accumulation of each portfolio from January 1996 to December 2019.

Compared with EW portfolio, the introduction of risk and uncertainty management alters the picture dramatically. The MV portfolio has a higher cumulative return in the sample period, compared to the EW portfolio. In addition, an investor can further benefit from accounting for uncertainty in portfolio allocation: the terminal market value of UNC portfolio is about 33% higher than that of MV portfolio. The wealth growth of UNC portfolio is similar to the MV portfolio in general. The UNC portfolio is more conservative before 2008 because of the relatively high and volatile uncertainty but drops less during the financial crisis, compared to the MV portfolio.

5.5 Portfolio Weights and Uncertainty Shock

How do portfolio weights respond to uncertainty versus risk shocks? To uncover the relationship between uncertainty (risk) shocks and portfolio weights, we estimate a structural VAR model with four variables: the natural log of equity weight in the UNC portfolio, the natural log of the bond weight in the same portfolio, the Frobenius norm of composite uncertainty matrix, and the Frobenius norm of the DCC(1,1) covariance matrix. We also experimented with other specifications but use this particular case as the most illustrative and at the same time representative of our findings.[7] The structural VAR takes the form:

$$Y_t^i = \sum_{s=1}^{L} B_s^i Y_{t-s}^i + C\epsilon_t^i \tag{11}$$

for $i = 1$ and 2 corresponding to different orders of the entries to the vectors, namely: $Y_t^1 = (\|\hat{\Sigma}_t\|_F,\ \|\Lambda_t\|_F,\ \ln(w_t^{bond}),\ \ln(w_t^{equity}))^\top$ or $Y_t^2 = (\|\Lambda_t\|_F,\ \|\hat{\Sigma}_t\|_F,\ \ln(w_t^{bond}),\ \ln(w_t^{equity}))^\top$ and with the following structural matrix C :

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{12}$$

In the VAR with Y_t^1, we isolate the uncertainty shock that is orthogonal to the risk since in the Cholesky-type C matrix we have a shock to $\|\Lambda_t\|_F$ placed second to the risk shock. When we flip the order of the first two variables, i.e. we use Y_t^2 in the VAR, we obtain the reverse which is a risk shock that is orthogonal to uncertainty. We estimate both VAR models with one lag (selected according to BIC) using the optimal composite UNC portfolio weights when the relative risk aversion parameter γ and the uncertainty aversion parameter

[7] We sum up the portfolio weight of S&P500 ETF and five Fama-French industry portfolios to get the total weight of equity.

δ are both equal to four (our findings are robust to various parameter specifications).

We compute impulse response functions (IRF) for both VAR specifications and only look at two cases which are of direct interest to us: (a) from the VAR with Y_t^1 we examine the IRF of a shock to uncertainty orthogonal to risk and its impact on equity and bond holdings, and (b) from the VAR with Y_t^2 we examine the IRF for a shock to risk orthogonal to uncertainty and its impact on equity and bond holdings. In Figure 7 we report the IRFs as well as 90% confidence intervals. Starting with the IRF for a shock to uncertainty we see that equity initially increase but then decrease and remain negative for 12 lags/months. In contrast, the same shock has a immediate positive impact on bond holdings. These findings are consistent with the heat map of Figure 4. The IRF to a risk shock yields a decline in equity holdings which tapers off after 12 lags/months and a short-lived negative impact on bonds. The scales in both plots indicate that the response to risk shocks are larger in magnitude than the responses to uncertainty shocks. Using the confidence intervals, we clearly see a significant difference between uncertainty versus risk shocks on bond holdings. For equity holdings the distinction is not so clear.

6 Conclusion

We presented a general methodology for dealing with ambiguity guided by model selection procedures popularized by the machine learning literature and demonstrated its use in the context of optimal portfolio allocation with risk and uncertainty. We build a set of statistical perturbations of a model via bootstrap samples to characterize model uncertainty associated with variable selection as well as parameter estimation in a high-dimensional date environment. In the specific application of portfolio selection this means we are interested in expected return uncertainty. Our empirical study involves an allocation problem involving eight assets and we find model perturbations featuring a selection of commonly used predictors in the finance literature. Moreover, we find that the comovements of the different pairs of asset expected return uncertainties behave very differently from their corresponding joint risk profiles. Interestingly we find that equity market index versus government bonds uncertainty trends towards zero during our sample, except for upticks during recessions, which means stock and bond uncertainty appear to become more disentangled. Broadly speaking, across all eight assets, our findings suggest that the time series pattern of uncertainty tends to lead that of risk.

We compare the out-of-sample performance of the Markowitz with uncertainty portfolio with the standard Markowitz mean-variance portfolio. Overall the portfolios taking into account uncertainty have a higher average out-of-sample utility. The gains from internalizing uncertainty in the portfolio optimization are both economically meaningful and statistically significant. Comparing the allocations we find that the ambiguity-sensitive investor moves out of equities into fixed income and to a certain degree gold when uncertainty spikes. In terms of terminal wealth we find that over our sample period a $100 initial investment has a 33% higher terminal wealth for a robust investor compared to the MV portfolio. Finally, we also construct a VAR model to study the impact of shocks to risk and uncertainty onto portfolio weights. A shock to uncertainty results in reduced equity allocations and a positive impact on bond holdings. Responses to risk shocks, however, are in comparison larger in magnitude than the responses to uncertainty shocks.

We limited our attention to eight representative assets. Our methodology can easily be adopted to handle a larger set of assets, or for that matter any prediction problem beyond that of expected returns. The computational burdens increase as well and progress on numerical implementations is certainly an area where more research is needed.

References

Anderson, E. W., and A.-R. Cheng (2016): "Robust bayesian portfolio choices," Review of Financial Studies, 29(5), 1330–1375.

Anderson, E. W., E. Ghysels, and J. L. Juergens (2009): "The impact of risk and uncertainty on expected returns," Journal of Financial Economics, 94(2), 233–263.

Anderson, E. W., L. P. Hansen, and T. J. Sargent (2003): "A quartet of semigroups for model specification, robustness, prices of risk, and model detection," Journal of the European Economic Association, 1(1), 68–123.

Bach, F. R. (2008): "Bolasso: model consistent LASSO estimation through the bootstrap," in Proceedings of the 25th international conference on Machine learning, pp. 33–40.

Bai, J., and S. Ng (2002): "Determining the number of factors in approximate factor models," Econometrica, 70(1), 191–221.

Baker, S. R., N. Bloom, and S. J. Davis (2016): "Measuring economic policy uncertainty," The quarterly journal of economics, 131(4), 1593–1636.

Baker, S. R., N. Bloom, S. J. Davis, and K. J. Kost (2019): "Policy news and stock market volatility," Discussion paper, National Bureau of Economic Research.

Belloni, A., V. Chernozhukov, D. Chetverikov, C. Hansen, and K. Kato (2018): "High-dimensional econometrics and regularized GMM," *Handbook of Econometrics*, forthcoming and arXiv preprint arXiv:1806.01888.

Bewley, T. F. (1988): "Knightian decision theory and econometric inference," Cowles Foundation for Research in Economics, Yale University.

——— (2011): "Knightian decision theory and econometric inferences," Journal of Economic Theory, 146(3), 1134–1147.

Bhandari, A., J. Borovička, and P. Ho (2016): "Identifying ambiguity shocks in business cycle models using survey data," Discussion paper, National Bureau of Economic Research.

Breiman, L. (2001): "Random forests," Machine learning, 45(1), 5–32.

Breiman, L., J. Friedman, C. J. Stone, and R. A. Olshen (1984): Classification and regression trees. CRC press.

Caner, M., and A. B. Kock (2019): "High Dimensional Linear GMM," arXiv preprint arXiv:1811.08779.

Chatterjee, A., and S. N. Lahiri (2011): "Bootstrapping LASSO estimators," Journal of the American Statistical Association, 106(494), 608–625.

DeMiguel, V., L. Garlappi, and R. Uppal (2009): "Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?," The review of Financial studies, 22(5), 1915–1953.

D'Hondt, C., R. De Winne, E. Ghysels, and S. Raymond (2020): "Artificial Intelligence Alter Egos: Who might benefit from robo-investing?," Journal of Empirical Finance, 59, 278–299.

Diebold, F. X., and R. S. Mariano (2002): "Comparing predictive accuracy," Journal of Business and Economic Statistics, 20(1), 134–144.

Ellsberg, D. (1961): "Risk, ambiguity, and the Savage axioms," Quarterly Journal of Economics, pp. 643–669.

Engle, R. F. (2002): "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," Journal of Business and Economic Statistics, 20, 339–350.

Epstein, L. G., and T. Wang (1994): "Intertemporal asset pricing under Knightian uncertainty," Econometrica, 62, 283–322.

Fama, E. F., and K. R. French (2015): "A five-factor asset pricing model," Journal of financial economics, 116(1), 1–22.

Fan, J., and R. Li (2001): "Variable selection via nonconcave penalized likelihood and its oracle properties," Journal of the American statistical Association, 96(456), 1348–1360.

Friedman, J. H. (2001): "Greedy function approximation: a gradient boosting machine," Annals of statistics, pp. 1189–1232.

Garlappi, L., R. Uppal, and T. Wang (2007): "Portfolio selection with parameter and model uncertainty: A multi-prior approach," Review of Financial Studies, 20(1), 41–81.

Ghysels, E., C. Horan, and E. Moench (2018): "Forecasting through the rearview mirror: Data revisions and bond return predictability," Review of Financial Studies, 31(2), 678–714.

Gilboa, I., and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," Journal of Mathematical Economics, 18(2), 141–153.

Goodfellow, I., Y. Bengio, and A. Courville (2016): Deep learning. MIT Press.

Gu, S., B. Kelly, and D. Xiu (2020): "Empirical asset pricing via machine learning," Review of Financial Studies, 33(5), 2223–2273.

Guidolin, M., and F. Rinaldi (2013): "Ambiguity in asset pricing and portfolio choice: A review of the literature," Theory and Decision, 74(2), 183–217.

Hansen, L. P. (1982): "Large sample properties of generalized method of moments estimators," Econometrica, pp. 1029–1054.

Hansen, L. P., and T. J. Sargent (2001): "Robust control and model uncertainty," American Economic Review, 91(2), 60–66.

——— (2011): Robustness. Princeton University Press.

Hastie, T., R. Tibshirani, and J. Friedman (2017): The elements of statistical learning: Data mining, inference, and prediction - Second edition. Springer, New York.

Hoerl, A. E., and R. W. Kennard (1970): "Ridge regression: Biased estimation for nonorthogonal problems," Technometrics, 12(1), 55–67.

Hurwicz, L. (1951): "Optimality criteria for decision making under ignorance. Cowles Commun," Discuss. Paper Statistics No. 370, Cowles Commission.

Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring uncertainty," American Economic Review, 105(3), 1177–1216.

Kim, H. H., and N. R. Swanson (2014): "Forecasting financial and macroeconomic variables using data reduction methods: New empirical evidence," Journal of Econometrics, 178, 352–367.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A smooth model of decision making under ambiguity," Econometrica, 73(6), 1849–1892.

Knight, F. H. (1921): Risk, uncertainty and profit. Houghton Mifflin, Boston.

Kostrikov, I., and O. Nachum (2020): "Statistical bootstrapping for uncertainty estimation in off-policy evaluation," arXiv preprint arXiv:2007.13609.

Kullback, S., and R. A. Leibler (1951): "On information and sufficiency," Annals of Mathematical Statistics, 22, 79–86.

Li, Y., Y. Luo, D. Ferrari, X. Hu, and Y. Qin (2019): "Model confidence bounds for variable selection," Biometrics, 75(2), 392–403.

Maccheroni, F., M. Marinacci, and D. Ruffino (2013): "Alpha as Ambiguity: Robust Mean-Variance Portfolio Analysis," Econometrica, 81(3), 1075–1113.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Ambiguity aversion, robustness, and the variational representation of preferences," Econometrica, 74(6), 1447–1498.

Markowitz, H. (1959): Portfolio selection: Efficient diversification of investments. John Wiley New York.

Pearce, T., M. Zaki, A. Brintrup, and A. Neely (2018): "High-quality prediction intervals for deep learning: A distribution-free, ensembled approach," arXiv preprint arXiv:1802.07167.

Politis, D. N., and J. P. Romano (1994): "The stationary bootstrap," Journal of the American Statistical Association, 89(428), 1303–1313.

Rapach, D. E., J. K. Strauss, and G. Zhou (2010): "Out-of-sample equity premium prediction: Combination forecasts and links to the real economy," Review of Financial Studies, 23(2), 821–862.

Sullivan, R., A. Timmermann, and H. White (1999): "Data-snooping, technical trading rule performance, and the bootstrap," Journal of Finance, 54(5), 1647–1691.

Tibshirani, R. (1996): "Regression shrinkage and selection via the lasso," Journal of the Royal Statistical Society: Series B (Statistical methodology), 58(1), 267–288.

Wald, A. (1950): Statistical decision functions. Wiley.

Welch, I., and A. Goyal (2008): "A comprehensive look at the empirical performance of equity premium prediction," Review of Financial Studies, 21(4), 1455–1508.

White, H. (2000): "A reality check for data snooping," Econometrica, 68(5), 1097–1126.

Zhang, C.-H. (2010): "Nearly unbiased variable selection under minimax concave penalty," Annals of statistics, 38(2), 894–942.

Zou, H., and T. Hastie (2005): "Regularization and variable selection via the elastic net," Journal of the Royal Statistical Society: Series B (Statistical methodology), 67(2), 301–320.

Table 1: Out-Of-Sample Model MSE with Real Data

	S&P 500	LT Bond	Gold	Cnsmr	Manuf	HiTec	Hlth	Other
LASSO	18.22	1.96	11.88	15.92	18.93	40.49	17.93	33.70
RIDGE	19.49	1.97	13.60	15.68	18.98	41.39	18.04	28.69
Elastic Net	19.50	1.94	11.93	15.85	19.46	41.34	18.23	28.91
MCP	18.22	1.96	11.88	15.93	18.93	40.49	17.93	33.46
SCAD	18.22	1.96	11.88	15.92	18.93	40.49	17.93	33.70
Random Forest	18.95	1.92	12.69	15.41	19.67	42.87	18.40	28.32
Gradient Boosting	19.08	2.04	12.54	15.79	19.74	42.50	18.58	28.24
Deep Learning	22.61	2.12	14.29	19.25	22.84	52.83	23.29	33.79
ML Model EW Average	18.74	1.92	12.03	15.69	19.08	41.85	18.05	30.12
OLS	108.62	19.37	50.51	103.67	207.83	149.75	161.55	139.86

Notes - The table reports the Out-Of-Sample (OOS) Mean Squared Error (MSE) of asset return predictions for each asset and each machine learning model (including the equal weighted average of all machine learning model predictions) and the ordinary least squares (OLS) model. The models are trained with the real data. The sample period is between January 1996 and December 2019 for the machine learning models and is between January 1998 and December 2019 for the OLS.

Table 2: Out-Of-Sample Model MSE with Bootstrap Data

	S&P 500	LT Bond	Gold	Cnsmr	Manuf	HiTec	Hlth	Other
LASSO	36.36	4.27	17.85	31.70	34.49	78.44	39.71	55.39
RIDGE	34.20	3.81	18.64	28.34	31.58	74.50	34.91	50.50
Elastic Net	36.38	4.07	18.28	30.84	34.52	78.24	37.55	53.40
MCP	39.39	4.54	18.37	34.57	37.32	83.58	42.82	58.76
SCAD	44.28	4.80	19.50	39.41	42.24	95.45	49.05	62.78
Random Forest	19.89	1.98	13.15	15.98	20.57	44.71	19.24	29.04
Gradient Boosting	23.75	2.36	14.09	19.78	24.38	53.64	23.86	33.81
Deep Learning	33.81	3.25	18.13	28.44	31.56	73.33	34.84	48.33

Notes - The table reports the Out-Of-Sample (OOS) Mean Squared Error (MSE) of asset return predictions for each asset and each machine learning model. The models are trained with the bootstrap data and the numbers in the table are averaged across bootstrap repetitions. The sample period is between January 1996 and December 2019.

Table 3: Most Selected Regressors For Linear Machine Learning Models - Real Data

	S&P 500		Long Term Bond	
1	Government Sponsored Enterprises EMV	0.275	Long Term Yield	0.717
2	Real Personal Consumption	0.250	Oil Price	0.700
3	Industrial Production	0.242	Regulation EPU	0.700
4	Dividend Price Ratio	0.242	Restaurants Hotels Industry Return	0.675
5	Stock Variance	0.242	Oil Industry Return	0.675
6	Robust Minus Weak Profitability Factor	0.242	Mines Industry Return	0.667
7	Conservative Minus Aggressive Factor	0.242	Momentum	0.517
8	Infectious Disease EMV	0.242	Market Portfolio Minus Risk-Rree Factor	0.500
9	Intellectual Property Matters EMV	0.242	Conservative Minus Aggressive Factor	0.475
10	Food Drug Policy EMV	0.242	Consumer Price Index	0.458

	Consumer Goods Industry		Manufacturing Industry	
1	Government Sponsored Enterprises EMV	0.417	Government Sponsored Enterprises EMV	0.258
2	Real Personal Consumption	0.408	Real Personal Consumption	0.233
3	1 Year Treasury Minus Federal Funds Rate	0.408	Real Personal Income	0.208
4	Stock Variance	0.408	Housing Starts	0.208
5	Small Minus Big Factor	0.408	Consumer Price Index	0.208
6	Infectious Disease EMV	0.408	Earnings Price Ratio	0.208
7	Policy-related EPU	0.408	Dividend Payout Ratio	0.208
8	Dividend Yield	0.392	Stock Variance	0.208
9	Conservative Minus Aggressive Factor	0.392	Default Return Spread	0.208
10	Industrial Production	0.383	Trading Industry Return	0.208

	High Technology Industry		Healthcare Industry	
1	Infectious Disease EMV	0.275	Term Spread	0.317
2	Stock Variance	0.258	Earnings Price Ratio	0.292
3	Industrial Production	0.250	Real Personal Consumption	0.283
4	Real Personal Consumption	0.250	M1	0.283
5	Dividend Yield	0.250	5 Year Treasury Minus Federal Funds Rate	0.283
6	Robust Minus Weak Profitability Factor	0.250	10 Year Treasury Minus Federal Funds Rate	0.283
7	Steel Industry Return	0.250	Stock Variance	0.283
8	Utilities Industry Return	0.250	Conservative Minus Aggressive Factor	0.283
9	Wholesale Industry Return	0.250	Momentum	0.283
10	Labor Markets EMV	0.250	Oil Price	0.283

	Other Industry		Gold	
1	Oil Price	0.358	Mines Industry Return	1.000
2	Conservative Minus Aggressive Factor	0.350	Finance Industry Return	0.900
3	Earnings Price Ratio	0.342	Inflation EMV Indicator	0.900
4	Stock Variance	0.342	Dollar Index	0.867
5	Default Return Spread	0.342	Consumer Price Index	0.775
6	Exchange Rates EMV	0.342	Food Products Industry Return	0.700
7	Monetary Policy EPU	0.342	Labor Regulations EMV	0.608
8	Agricultural Policy EMV	0.333	Long Term Yield	0.600
9	Fiscal Policy EPU	0.325	Energy Environmental Regulation EMV	0.600
10	Dividend Payout Ratio	0.317	Retail Industry Return	0.575

Notes - The table summarizes the top 10 most selected covariates in the linear machine learning models described in Section 3.1 for each of the 8 assets considered in the portfolio. The first column of each asset reports the name of the covariates and second column reports the percentage of time periods where the covariate is selected. Models are trained with the real data. The variable description is in the online Appendix OA.1, OA.2, OA.3, and OA.4.

Table 4. Most Selected Regressors For Linear Machine Learning Models - Bootstrap Data

	S&P 500		**Long Term Bond**	
1	Real Personal Consumption	0.882	Long Term Yield	0.916
2	Infectious Disease EMV	0.821	Utility Industry Return	0.914
3	Small Minus Big Factor	0.798	Government Sponsored Enterprises EMV	0.881
4	Agricultural Policy EMV	0.785	Mines Industry Return	0.870
5	Default Return Spread	0.776	Oil Orice	0.860
6	Oil Price	0.763	Trans Infras Util EMV	0.851
7	Industrial Production	0.759	Industrial Production	0.838
8	Conservative Minus Aggressive Factor	0.757	M2	0.836
9	Steel Industry Return	0.757	Housing Land Management EMV	0.830
10	Earnings Price Ratio	0.749	Agricultural Policy EMV	0.819
	Consumer Goods Industry		**Manufacturing Industry**	
1	Real Personal Consumption	0.846	Real Personal Consumption	0.828
2	Default Return Spread	0.824	Default Return Spread	0.794
3	Conservative Minus Aggressive Factor	0.813	Agricultural Policy EMV	0.739
4	Agricultural Policy EMV	0.812	Intellectual Property Matters EMV	0.731
5	Small Minus Big Factor	0.803	Infectious Disease EMV	0.719
6	Infectious Disease EMV	0.790	Government Sponsored Enterprises EMV	0.713
7	Food Drug Policy EMV	0.784	Oil Price	0.713
8	Steel Industry Return	0.780	Momentum	0.712
9	Oil Price	0.776	Dollar Index	0.711
10	Earnings Price Ratio	0.773	Stock Variance	0.705
	High Technology Industry		**Healthcare Industry**	
1	Infectious Disease EMV	0.820	Agricultural Policy EMV	0.916
2	Real Personal Consumption	0.779	Dollar Index	0.872
3	Robust Minus Weak Factor	0.775	Real Personal Consumption	0.860
4	Intellectual Property Matters EMV	0.762	Conservative Minus Aggressive Factor	0.826
5	Energy Environmental Regulation EMV	0.751	Earnings Price Ratio	0.820
6	Labor Regulation EMV	0.750	Infectious Disease EMV	0.811
7	Agricultural Policy EMV	0.746	Tobacco Products Industry Return	0.810
8	Oil Price	0.737	Momentum	0.782
9	Stock Variance	0.735	Stock Variance	0.778
10	Industrial Production	0.735	Food Drug Policy EMV	0.777
	Other Industry		**Gold**	
1	Conservative Minus Aggressive Factor	0.912	Mines Industry Return	0.999
2	Real Personal Consumption	0.858	Finance Industry Return	0.909
3	Oil Price	0.818	Inflation EMV	0.907
4	Default Return Spread	0.818	Dollar Index	0.880
5	Agricultural Policy EMV	0.808	Trans Infras Util EMV	0.823
6	Dollar Index	0.802	Energy Environmental Regulation EMV	0.810
7	Manuf And Trade Industry Sales	0.797	CPI	0.805
8	Infectious Disease EMV	0.796	Other Financial Indicator EMV	0.799
9	Small Minus Big Factor	0.792	Coal Industry Return	0.792
10	Intellectual Property Matters EMV	0.767	Agricultural Policy EMV	0.774

Notes - The table summarizes the top 10 most selected covariates in the linear machine learning models described in Section 3.1 for each of the 8 assets considered in the portfolio. The first column of each asset reports the name of the covariates and second column reports the percentage of bootstrap repetitions where the covariate is selected. The variable descriptions appear in the Online Appendix Tables OA.1 through OA.4.

Table 5: Knightian Uncertainty and EPU, JLN, VIX, Volatility and IP

	$AR(1)$	EPU	JLN	VIX	$\|\hat{\Sigma}\|_F$	IP
			Panel A: $\|\hat{\Lambda}^B_{M,t}\|_F$			
Slope	0.63***	4.47*	0.11***	4.78***	0.51***	−0.07*
	(0.05)	(2.53)	(0.01)	(0.47)	(0.04)	(0.04)
R^2	0.40	0.01	0.25	0.26	0.34	0.08
			Panel B: $\|\hat{\Lambda}^B_{i_g,t}\|_F$ *with* i_g *the group of linearized regressions*			
Slope	0.61***	2.58	0.07***	3.09***	0.34***	−0.04**
	(0.05)	(1.73)	(0.01)	(0.33)	(0.03)	(0.02)
R^2	0.37	0.01	0.23	0.24	0.32	0.06
			Panel C: $\|\hat{\Lambda}^B_{i_g,t}\|_F$ *with* i_g *the group of non-linear models*			
Slope	0.74***	18.90***	0.29***	13.32***	1.29***	−0.22**
	(0.04)	(6.31)	(0.03)	(1.14)	(0.11)	(0.11)
R^2	0.55	0.03	0.26	0.32	0.35	0.12

Notes - The table reports the statistics from regressing the Frobenius norm of uncertainty measures $\hat{\Lambda}^B_{M,t}$ and $\hat{\Lambda}^B_{i_g,t}$ for i_g respectively linear and non-linear models, projected on the Economic Policy Uncertainty Index (EPU) by Baker, Bloom, and Davis (2016), the financial uncertainty Index (JLN) by Jurado, Ludvigson, and Ng (2015) and the CBOE Volatility Index (VIX), (d) the Frobenius norm of the DCC(1,1) covariance matrix $\hat{\Sigma}_t$, and (e) industrial production growth (IP).
The uncertainty measures are multiplied by one hundred, except for the autoregression and the regression involving $\|\hat{\Sigma}\|_F$. Standard errors are reported in brackets, all data is monthly frequency. One, two, and three stars denote statistical significance at the 1%, 5%, and 10% level respectively.

Table 6: Summary Statistics of Uncertainty

	Mean	**Volatility**	**Corr(-1)**
Pre-Crisis	78.72	64.36	0.49
Crisis	182.68	186.27	0.68
Post-Crisis	44.56	56.85	0.43
All	70.63	82.43	0.63

Notes - The table reports the mean, volatility and the lag one autocorrelation of the Frobenius norm of uncertainty measures $\|\hat{\Lambda}^B_M\|_F$ in the entire sample and three subsamples: pre-crisis (January 1996 to November 2007), crisis (December 2007 to June 2009) and post-crisis (July 2009 to December 2019). The uncertainty measures are scaled by ten thousand.

Table 7: Out-of-Sample Utility: UNC Portfolio versus MV Portfolio

	$\delta = 0$	$\delta = 2$	$\delta = 4$	$\delta = 6$	$\delta = 8$
			Panel A: UNC with $\hat{\Lambda}^B_{M,t}$ / MV		
$\gamma = 2$	1.00	1.30**	1.80***	2.81***	6.84***
$\gamma = 4$	1.00	1.27**	1.63***	2.26***	3.67***
$\gamma = 6$	1.00	1.18**	1.40***	1.70***	2.20***
$\gamma = 8$	1.00	1.15***	1.29***	1.48***	1.77***
$\gamma = 10$	1.00	1.11***	1.21***	1.35***	1.53***
			Panel B: UNC with $\hat{\Lambda}^B_{i_g,t}$, i_g linear models / MV		
$\gamma = 2$	1.00	1.23***	1.46***	1.99***	3.52***
$\gamma = 4$	1.00	1.12***	1.33***	1.73***	2.61***
$\gamma = 6$	1.00	1.11***	1.29***	1.60***	2.19***
$\gamma = 8$	1.00	1.08***	1.21***	1.41***	1.72***
$\gamma = 10$	1.00	1.07***	1.18***	1.33***	1.54***
			Panel C: UNC with $\hat{\Lambda}^B_{i_g,t}$, i_g nonlinear models / MV		
$\gamma = 2$	1.00	1.01***	1.06***	1.14***	1.24***
$\gamma = 4$	1.00	1.05***	1.18***	1.39***	1.66***
$\gamma = 6$	1.00	1.14**	1.30***	1.50***	1.74***
$\gamma = 8$	1.00	1.10**	1.20***	1.32***	1.45***
$\gamma = 10$	1.00	1.06***	1.13***	1.21***	1.29***

Notes - The table compares the average out-of-sample utility for an uncertainty/risk averse versus risk averse portfolio under different relative risk aversion parameters γ and uncertainty aversion parameters δ, where the investor uses different machine learning models and associated uncertainty measures $\hat{\Lambda}^B_{M,t}$ and $\hat{\Lambda}^B_{i_g,t}$ for i_g respectively linear and nonlinear models. Out-of-sample returns are at monthly frequency. One, two, and three stars denote statistical significance of Diebold and Mariano (2002) test at the 1%, 5%, and 10% level respectively.

Table 8: Out-of-Sample Utility: UNC Portfolio versus EW Portfolio

	$\delta = 0$	$\delta = 2$	$\delta = 4$	$\delta = 6$	$\delta = 8$
			Panel A: UNC with $\hat{\Lambda}^B_{M,t}$ / EW		
$\gamma = 2$	0.91***	1.02***	1.14***	1.26**	1.44
$\gamma = 4$	1.04***	1.21**	1.38	1.58	1.85
$\gamma = 6$	1.29*	1.50	1.74	2.05**	2.53**
$\gamma = 8$	1.59	1.91**	2.30***	2.93***	4.12***
$\gamma = 10$	2.10**	2.65***	3.54***	5.42***	12.31***
			Panel B: UNC with $\hat{\Lambda}^B_{i_g,t}$, i_g linear models / EW		
$\gamma = 2$	1.30***	1.47***	1.55	1.72	2.04
$\gamma = 4$	1.44***	1.56	1.76	2.11	2.74*
$\gamma = 6$	1.61	1.83	2.20**	2.89***	4.43***
$\gamma = 8$	1.91	2.31***	3.06***	4.84***	12.98***
$\gamma = 10$	2.43***	3.26***	5.34***	17.32***	–
			Panel C: UNC with $\hat{\Lambda}^B_{i_g,t}$, i_g nonlinear models / EW		
$\gamma = 2$	0.86***	0.82***	0.80***	0.80***	0.80***
$\gamma = 4$	0.74***	0.73***	0.76***	0.81***	0.87***
$\gamma = 6$	0.83***	0.90***	0.98***	1.05**	1.13
$\gamma = 8$	1.10	1.20	1.30	1.41	1.53
$\gamma = 10$	1.49	1.63	1.79*	2.00**	2.25***

Notes - The table compares the average out-of-sample utility for an uncertainty/risk averse versus equally weighted portfolio under different relative risk aversion parameters γ and uncertainty aversion parameters δ, where the investor uses different machine learning models and associated uncertainty measures $\hat{\Lambda}^B_{M,t}$ and $\hat{\Lambda}^B_{i_g,t}$ for i_g respectively linear and nonlinear models. Out-of-sample returns are at monthly frequency. One, two, and three stars denote statistical significance of Diebold and Mariano (2002) test at the 1%, 5%, and 10% level respectively. In Panel B, the EW portfolio has negative utility value when $\gamma = 10$ and $\delta = 8$. We omit the negative utility ratio from the table under this (γ, δ) specification.

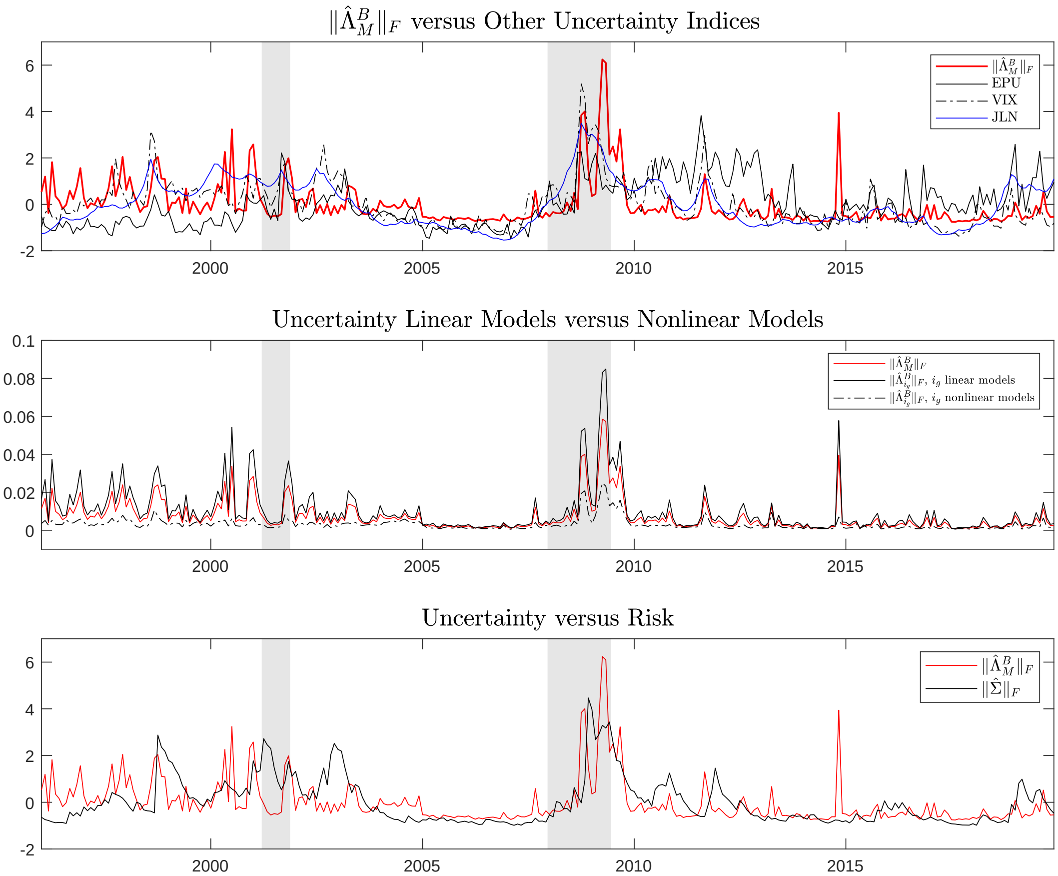

Fig. 1. **Knightian Uncertainty versus Other Uncertainty Indices and Volatility.** The top and bottom panel of the figure plot the Frobenius norm $\|\hat{\Lambda}^B_{M,t}\|_F$ appearing in equation (10) against (a) the Economic Policy Uncertainty Index (EPU), (b) the Jurado et al. financial uncertainty index (JLN), (c) the CBOE Volatility Index (VIX), and finally (d) the Frobenius norm of the DCC(1,1) covariance matrix $\|\Sigma_t\|_F$ following the estimation procedure of Engle (2002). The series in the top and bottom panel are normalized. The middle panel plots the Frobenius norm $\|\hat{\Lambda}^B_{M,t}\|_F$ appearing in equation (10) against the Frobenius norm $\|\hat{\Lambda}^B_{i_g,t}\|_F$ appearing in equation (9) for i_g representing the group of all regularized linear models or the group of all nonlinear machine learning models. The sample period is from January 1996 to December 2019.

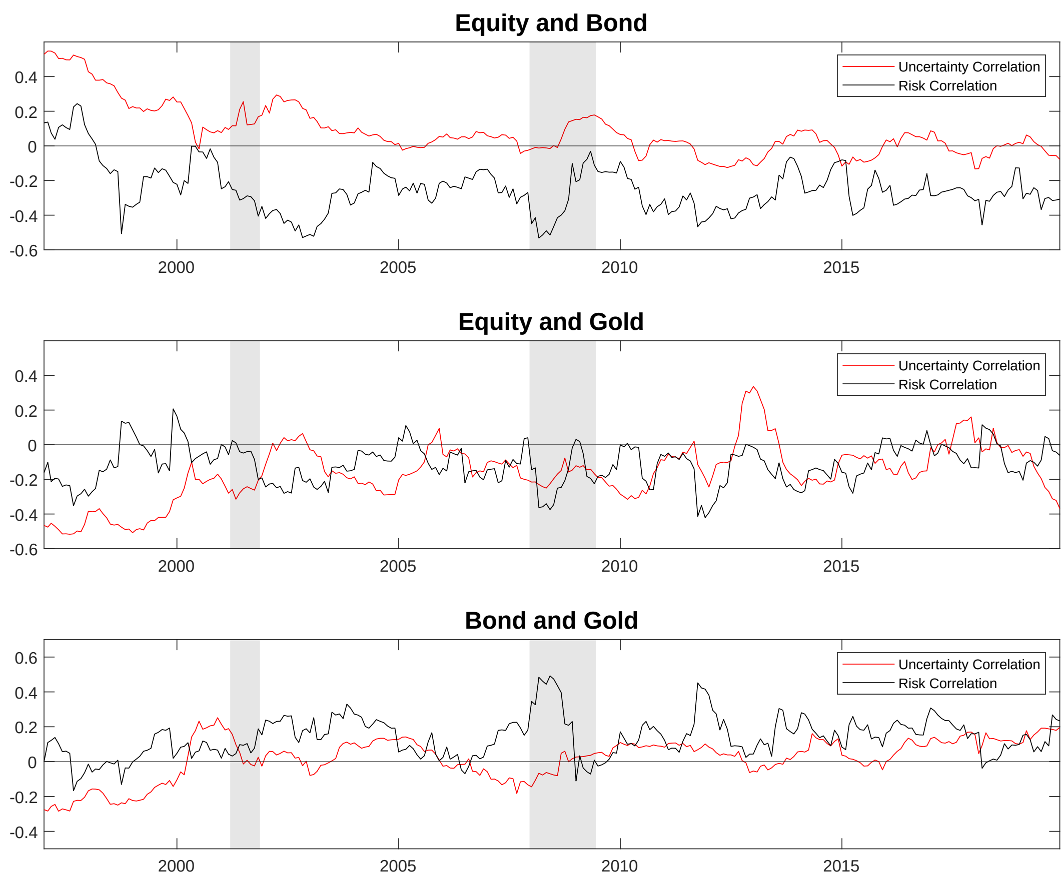

Fig. 2. **Off-diagonal elements of $\hat{\Lambda}^{B}_{M,t}$ versus Σ_t expressed as correlations.** Equities refers to the combination of all equity assets, namely S&P 500 and Fama-French five industry portfolios specified in Online Appendix Table OA.4.

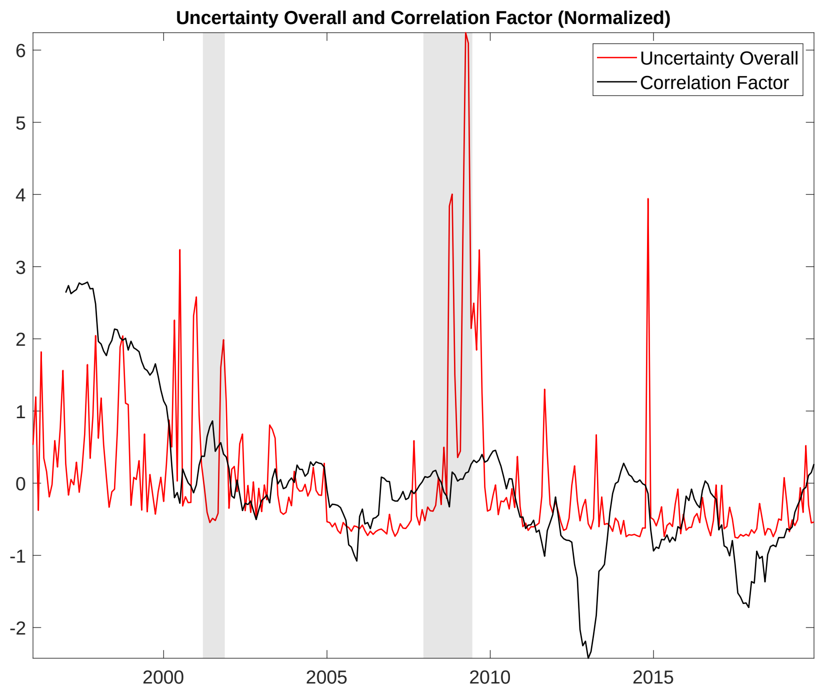

Fig. 3. **Overall Uncertainty versus Uncertainty Correlation Factor** The figure plots the Frobenius norm $\|\hat{\Lambda}^B_{M,t}\|_F$ appearing in equation (10) against the uncertainty correlation factor, which is extracted from off-diagonal elements of $\hat{\Lambda}^B_{M,t}$, following Bai and Ng (2002). Both series are normalized. The sample period is from January 1996 to December 2019.

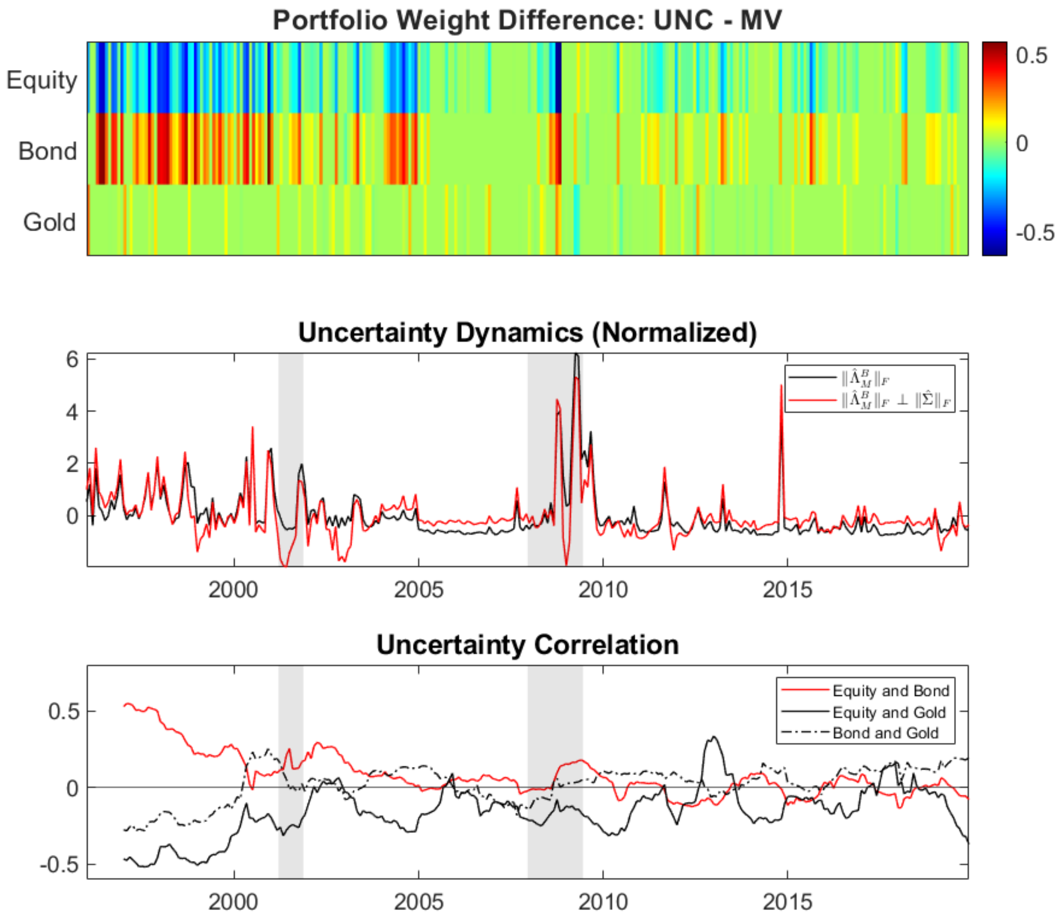

Fig. 4. **Portfolio Weight Difference Heatmap** The figure plots the equity, bond and gold weight difference between the "UNC" and the "MV" portfolio for the case where we set both γ and δ equal to four. The three subplots share the same x-axis range. The sample period is from January 1996 to December 2019.

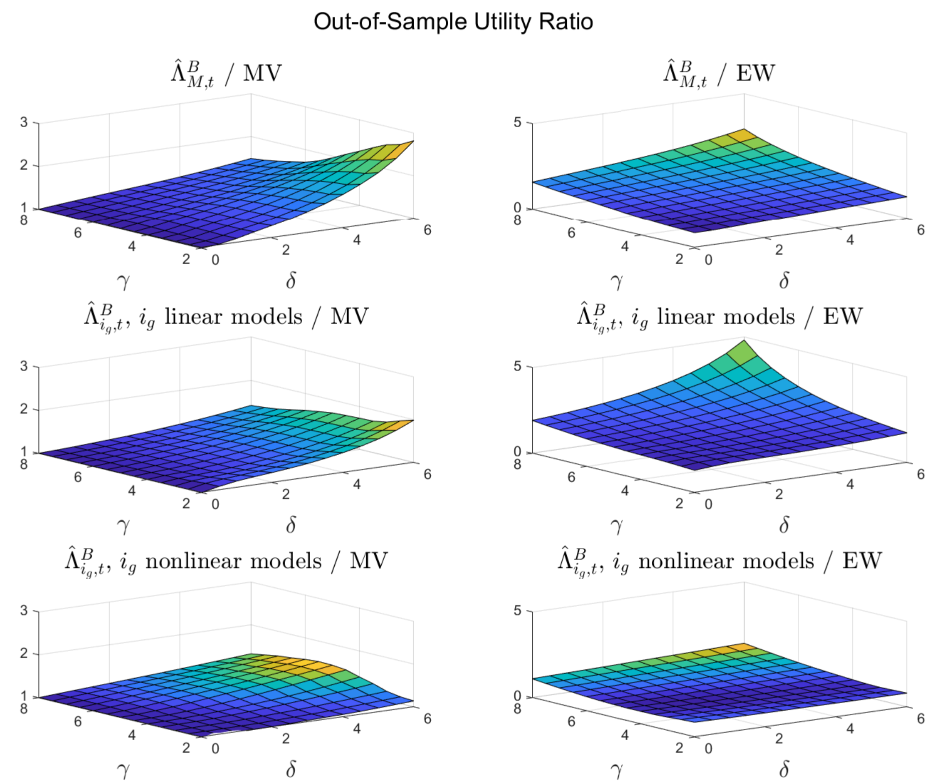

Fig. 5. **Out-of-Sample Utility: UNC versus MV and EW** The figure reports the ratio of average out-of-sample utility between portfolio rules over a grid of γ and δ. "Composite UNC", "Linear UNC" and "Nonlinear UNC" refers to the Markowitz with uncertainty portfolio allocation with composite, linear model and nonlinear model uncertainty estimator. MV refers to the mean-variance portfolio, and EW refers to the equal-weight portfolio. γ refers to the relative risk aversion parameter and δ refers to the uncertainty aversion parameter. The out-of-sample spans from January 1996 to December 2019.

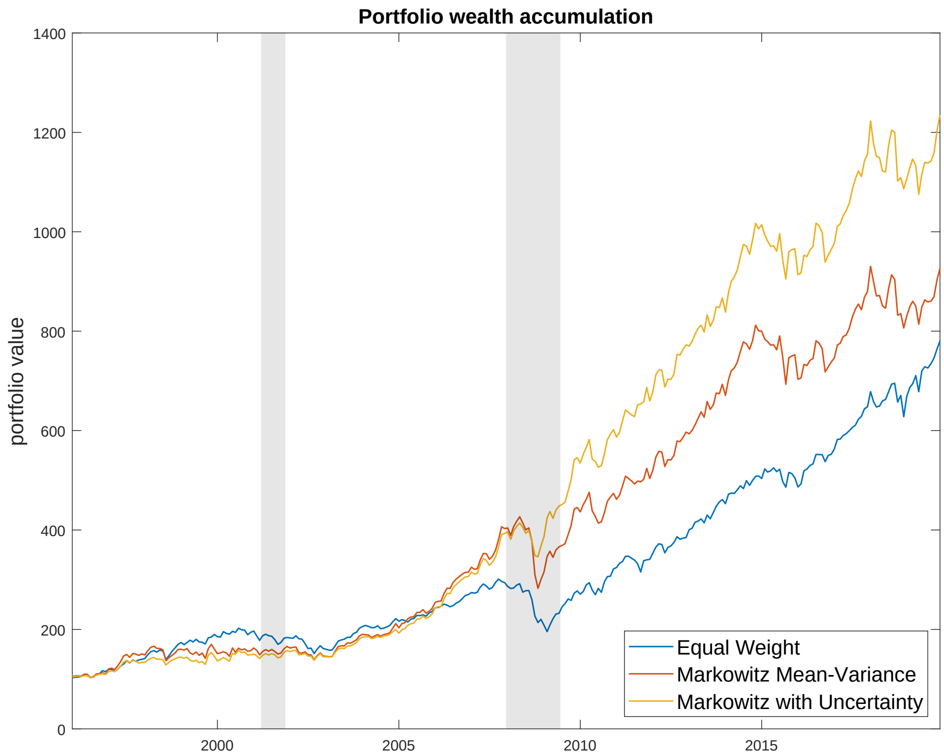

Fig. 6. **Portfolio Cumulative Return** The figure shows the wealth accumulation process of three portfolio rules between January 1996 and December 2019. Portfolios are rebalanced every month. The yellow line represents the Markowitz with uncertainty (UNC) portfolio, the red line represents the Markowitz mean-variance (MV) portfolio and the blue line represents the Equal Weight (EW) portfolio. We set both the relative risk aversion γ and uncertain aversion δ to four.

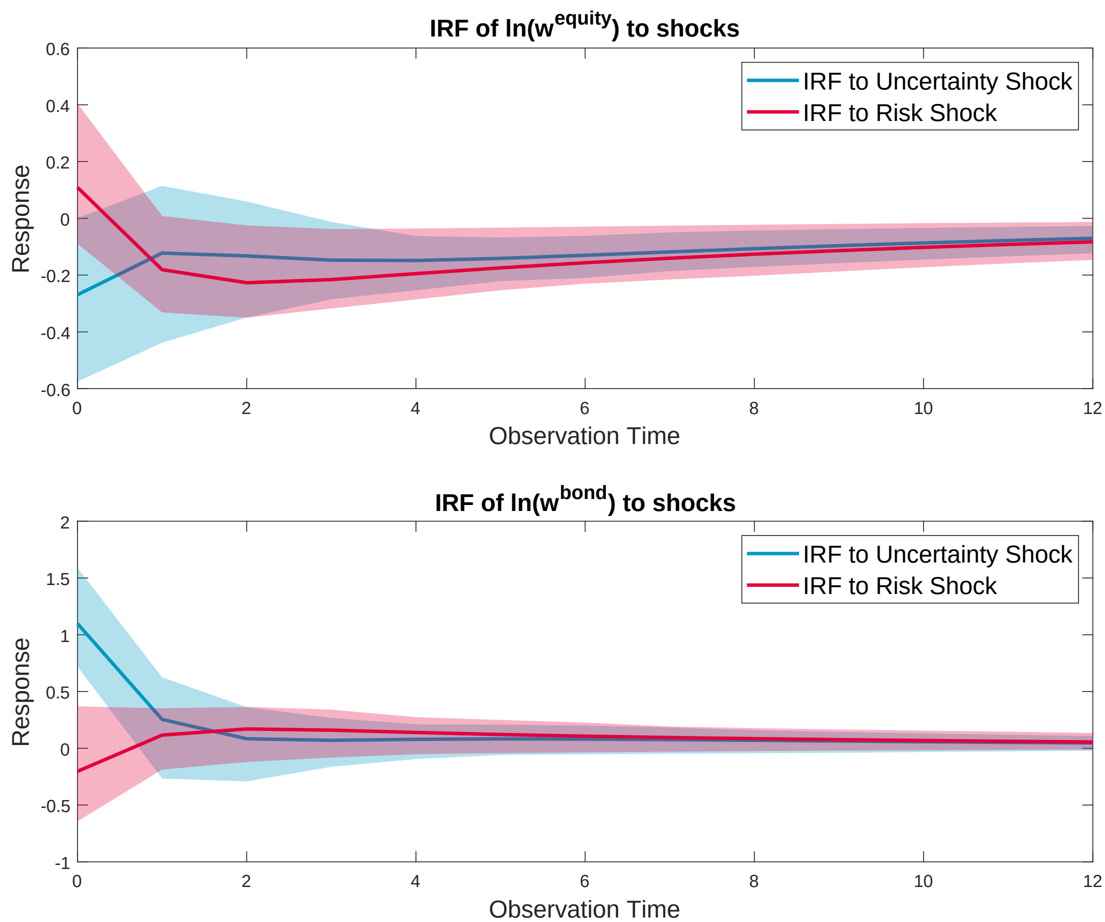

Fig. 7. **IRF of portfolio weights to shocks** The figure reports the impulse response of portfolio weights to the uncertainty and risk shocks along with the 90% confidence intervals. We set both the relative risk aversion γ and uncertain aversion δ to four. The structural VARs take the form appearing in equation (11) with different orders of the entries to the vectors, namely: $Y_t^1 = (\|\hat{\Sigma}_t\|_F, \|\Lambda_t\|_F, \ln(w_t^{bond}), \ln(w_t^{equity}))^\top$ or $Y_t^2 = (\|\Lambda_t\|_F, \|\hat{\Sigma}_t\|_F, \ln(w_t^{bond}), \ln(w_t^{equity}))^\top$. We compute impulse response functions (IRF) for both VAR specifications and only look at two cases which are of direct interest to us: (a) from the VAR with Y_t^1 we examine the IRF of a shock to uncertainty orthogonal to risk and its impact on equity and bond holdings, and (b) from the VAR with Y_t^2 we examine the IRF for a shock to risk orthogonal to uncertainty and its impact on equity and bond holdings.The number of lags in the VARs is set to one.

Online Appendix

Ambiguity with Machine Learning: An Application to Portfolio Choice

Eric Ghysels

Department of Economics, University of North Carolina Chapel Hill,
Department of Finance, Kenan-Flagler Business School and CEPR

Yan Qian

Department of Economics, University of North Carolina Chapel Hill

Steve Raymond

Kenan-Flagler Business School, University of North Carolina Chapel Hill

OA.1 Variable Description and Construction

Table OA.1: Macroeconomic series

No.	Description	Source	T-code
1	Chicago Fed National Activity Index	ALFRED	1
2	All Employees, Total Nonfarm	ALFRED	5
3	Industrial Production Index	ALFRED	5
4	Real Personal Income	ALFRED	5
5	Real Manufacturing and Trade Industries Sales	ALFRED	5
6	Real Personal Consumption	ALFRED	5
7	Housing Starts: Total New Privately Owned	ALFRED	4
8	Consumer Price Index: All Item	ALFRED	5
9	Producer Price Index for All Commodities	ALFRED	5
10	M1 Money Stock	ALFRED	5
11	M2 Money Stock	ALFRED	5
12	Trade Weighted U.S. Dollar Index	ALFRED	5
13	Effective Federal Funds Rate	FRED	2
14	3-Month Treasury Minus FEDFUNDS	FRED	1
15	6-Month Treasury Minus FEDFUNDS	FRED	1
16	1-Year Treasury Minus FEDFUNDS	FRED	1
17	5-Year Treasury Minus FEDFUNDS	FRED	1
18	10-Year Treasury Minus FEDFUNDS	FRED	1
19	Spot Crude Oil Price: West Texas Intermediate	FRED	5

Note: The column T-code denotes the data transformation applied to a time-series, which are: (1) Not transformed, (2) Δx_t, (3) $\Delta^2 x_t$, (4) $log(x_t)$, (5) $\Delta log(x_t)$.

Table OA.2: Finance series

No.	Description	Source	T-code
	Fama-French Factors		
1	Market Portfolio Minus Risk-Rree Bond	Ken French Data Library	1
2	Small Minus Big	Ken French Data Library	1
3	Value Minus Growth	Ken French Data Library	1
4	Robust Minus Weak Operating Profitability	Ken French Data Library	1
5	Conservative Minus Aggressive Investment	Ken French Data Library	1
6	Risk Free Rate	Ken French Data Library	1
7	Momentum Factor	Ken French Data Library	1
	Welch-Goyal Financial Predictors		
8	Dividend Price Ratio	Welch & Goyal (2008)	1
9	Dividend Yield	Welch & Goyal (2008)	1
10	Earnings Price Ratio	Welch & Goyal (2008)	1
11	Dividend Payout Ratio	Welch & Goyal (2008)	1
12	Stock Market Variance	Welch & Goyal (2008)	1
13	Book-to-Market Ratio	Welch & Goyal (2008)	1
14	Net Equity Expansion	Welch & Goyal (2008)	1
15	3 Month Treasury Bill Rate	Welch & Goyal (2008)	1
16	Long Term Yield	Welch & Goyal (2008)	1
17	Long Term Rate of Return	Welch & Goyal (2008)	1
18	Term Spread	Welch & Goyal (2008)	1
19	Default Yield Spread	Welch & Goyal (2008)	1
20	Default Return Spread	Welch & Goyal (2008)	1
21	CBOE S&P 100 Volatility Index	FRED	1
	Fama-French 49 Industries Portfolio Return		
22	Agriculture	Ken French Data Library	1
23	Food Products	Ken French Data Library	1
24	Candy & Soda	Ken French Data Library	1
25	Beer & Liquor	Ken French Data Library	1
26	Tobacco Products	Ken French Data Library	1
27	Recreation	Ken French Data Library	1
28	Entertainment	Ken French Data Library	1
29	Printing and Publishing	Ken French Data Library	1
30	Consumer Goods	Ken French Data Library	1
31	Apparel	Ken French Data Library	1
32	Healthcare	Ken French Data Library	1
33	Medical Equipment Industry Return	Ken French Data Library	1
34	Pharmaceutical Products	Ken French Data Library	1
35	Chemicals	Ken French Data Library	1
36	Rubber and Plastic Products	Ken French Data Library	1
37	Textiles	Ken French Data Library	1
38	Construction	Ken French Data Library	1
39	Construction Materials	Ken French Data Library	1
40	Steel	Ken French Data Library	1

41	Fabricated Products	Ken French Data Library	1
42	Machinery	Ken French Data Library	1
43	Electrical Equipment	Ken French Data Library	1
44	Automobiles and Trucks	Ken French Data Library	1
45	Aircraft	Ken French Data Library	1
46	Shipbuilding, Railroad Equipment	Ken French Data Library	1
47	Defense	Ken French Data Library	1
48	Precious Metals	Ken French Data Library	1
49	Mines	Ken French Data Library	1
50	Coal	Ken French Data Library	1
51	Petroleum and Natural Gas	Ken French Data Library	1
52	Utilities	Ken French Data Library	1
53	Telecommunication	Ken French Data Library	1
54	Personal Services	Ken French Data Library	1
55	Business Services	Ken French Data Library	1
56	Computer Hardware	Ken French Data Library	1
57	Computer Software	Ken French Data Library	1
58	Electronic Equipment	Ken French Data Library	1
59	Measuring and Control Equipment	Ken French Data Library	1
60	Business Supplies	Ken French Data Library	1
61	Shipping Containers	Ken French Data Library	1
62	Transportation	Ken French Data Library	1
63	Wholesale	Ken French Data Library	1
64	Retail	Ken French Data Library	1
65	Restaurants, Hotels, Motels	Ken French Data Library	1
66	Banking	Ken French Data Library	1
67	Insurance	Ken French Data Library	1
68	Real Estate	Ken French Data Library	1
69	Trading	Ken French Data Library	1
70	Other Industries	Ken French Data Library	1

Table OA.3: Textual Data series

No.	Description	Source	T-code
	Economic Policy Uncertainty (EPU) Index		
1	News Coverage about Policy-related Economic Uncertainty	Baker, Bloom & Davis (2016)	1
2	Tax Code Expiration Index	Baker, Bloom & Davis (2016)	1
3	CPI Forecast Disagreement	Baker, Bloom & Davis (2016)	1
4	Federal/State/Local Purchases Disagreement	Baker, Bloom & Davis (2016)	1
	U.S. Categorical EPU Data		
5	Monetary Policy	Baker, Bloom & Davis (2016)	1
6	Taxes	Baker, Bloom & Davis (2016)	1
7	Fiscal Policy	Baker, Bloom & Davis (2016)	1
8	Government Spending	Baker, Bloom & Davis (2016)	1
9	Health Care	Baker, Bloom & Davis (2016)	1
10	National Security	Baker, Bloom & Davis (2016)	1
11	Entitlement Programs	Baker, Bloom & Davis (2016)	1
12	Regulation	Baker, Bloom & Davis (2016)	1
13	Financial Regulation	Baker, Bloom & Davis (2016)	1
14	Trade Policy	Baker, Bloom & Davis (2016)	1
15	Sovereign Debt, Currency Crises	Baker, Bloom & Davis (2016)	1
	Equity Market Volatility (EMV) Trackers		
16	Policy-Related EMV	Baker, Bloom, Davis & Kost (2019)	1
17	Infectious Disease EMV	Baker, Bloom, Davis & Kost (2019)	1
18	Macroeconomic News and Outlook EMV	Baker, Bloom, Davis & Kost (2019)	1
19	Macro – Broad Quantity Indicators EMV	Baker, Bloom, Davis & Kost (2019)	1
20	Macro – Inflation EMV Indicator	Baker, Bloom, Davis & Kost (2019)	1
21	Macro – Interest Rates EMV	Baker, Bloom, Davis & Kost (2019)	1
22	Macro – Other Financial Indicators EMV	Baker, Bloom, Davis & Kost (2019)	1
23	Macro – Labor Markets EMV	Baker, Bloom, Davis & Kost (2019)	1
24	Macro – Real Estate Markets EMV	Baker, Bloom, Davis & Kost (2019)	1
25	Macro – Trade EMV	Baker, Bloom, Davis & Kost (2019)	1
26	Macro – Business Investment and Sentiment EMV	Baker, Bloom, Davis & Kost (2019)	1
27	Macro – Consumer Spending and Sentiment EMV	Baker, Bloom, Davis & Kost (2019)	1
28	Commodity Markets EMV	Baker, Bloom, Davis & Kost (2019)	1
29	Financial Crises EMV	Baker, Bloom, Davis & Kost (2019)	1
30	Exchange Rates EMV	Baker, Bloom, Davis & Kost (2019)	1
31	Healthcare Matters EMV	Baker, Bloom, Davis & Kost (2019)	1
32	Litigation Matters EMV	Baker, Bloom, Davis & Kost (2019)	1
33	Competition Matters EMV	Baker, Bloom, Davis & Kost (2019)	1
34	Labor Disputes EMV	Baker, Bloom, Davis & Kost (2019)	1
35	Intellectual Property Matters EMV	Baker, Bloom, Davis & Kost (2019)	1
36	Fiscal Policy EMV	Baker, Bloom, Davis & Kost (2019)	1
37	Taxes EMV	Baker, Bloom, Davis & Kost (2019)	1
38	Government Spending, Deficits, and Debt EMV	Baker, Bloom, Davis & Kost (2019)	1
39	Entitlement and Welfare Programs EMV	Baker, Bloom, Davis & Kost (2019)	1
40	Monetary Policy EMV	Baker, Bloom, Davis & Kost (2019)	1
41	Regulation EMV	Baker, Bloom, Davis & Kost (2019)	1
42	Financial Regulation EMV	Baker, Bloom, Davis & Kost (2019)	1
43	Competition Policy EMV	Baker, Bloom, Davis & Kost (2019)	1
44	Intellectual Property Policy EMV	Baker, Bloom, Davis & Kost (2019)	1
45	Labor Regulations EMV	Baker, Bloom, Davis & Kost (2019)	1
46	Immigration EMV	Baker, Bloom, Davis & Kost (2019)	1

47	Energy and Environmental Regulation EMV	Baker, Bloom, Davis & Kost (2019)	1
48	Lawsuit and Tort Reform, Supreme Court Decisions EMV	Baker, Bloom, Davis & Kost (2019)	1
49	Housing and Land Management EMV	Baker, Bloom, Davis & Kost (2019)	1
50	Other Regulation EMV	Baker, Bloom, Davis & Kost (2019)	1
51	National Security Policy EMV	Baker, Bloom, Davis & Kost (2019)	1
52	Government-Sponsored Enterprises EMV	Baker, Bloom, Davis & Kost (2019)	1
53	Trade Policy EMV	Baker, Bloom, Davis & Kost (2019)	1
54	Healthcare Policy EMV	Baker, Bloom, Davis & Kost (2019)	1
55	Food and Drug Policy EMV	Baker, Bloom, Davis & Kost (2019)	1
56	Transportation Infrastructure Utilities EMV	Baker, Bloom, Davis & Kost (2019)	1
57	Elections and Political Governance EMV	Baker, Bloom, Davis & Kost (2019)	1
58	Agricultural Policy EMV	Baker, Bloom, Davis & Kost (2019)	1
59	Petroleum Markets EMV	Baker, Bloom, Davis & Kost (2019)	1

Table OA.4: Portfolio Asset Return series

No.	Description	Source	T-code
1	S&P 500 Index Return	CRSP	1
2	Fama Bond Portfolio Return: 60month-120month	CRSP	1
3	Gold Price: 3:00 P.M London	FRED	5
	Fama-French 5 Industries Portfolio Return		
4	Consumer Goods	Ken French Data Library	1
5	Manufacturing	Ken French Data Library	1
6	High Technology	Ken French Data Library	1
7	Healthcare	Ken French Data Library	1
8	Other Industry	Ken French Data Library	1

Note: The column T-code denotes the data transformation applied to a time-series, which are: (1) Not transformed, (2) Δx_t, (3) $\Delta^2 x_t$, (4) $log(x_t)$, (5) $\Delta log(x_t)$. SPDR S&P 500 ETF Trust starts from 1993 January, so we use SP500 Index return as a proxy for S&P500 ETF return. iShares 7-10 Year Treasury Bond ETF starts from 2002 July, so we use Fama long-term bond portfolio return as a proxy for long-term treasury bond ETF return. SPDR Gold Shares ETF starts from 2004 November, so we use the gold commodity return as a proxy for the gold ETF return.

OA.2 Machine Learning Models and Tuning Parameter

Table OA.5: Machine Learning Models and Hyperparameters

Model Name	Hyperparameters
Linear Models:	
LASSO (Tibshirani (1996))	L1 penalty λ
RIDGE (Hoerl and Kennard (1970))	L2 penalty λ
Elastic Net (Zou and Hastie (2005))	penalty λ, and α
SCAD (Fan and Li (2001))	penalty λ, and α
MCP (Zhang (2010))	penalty λ, and α
Nonlinear Models:	
Random Forest (Breiman (2001))	depth, max# of features each split
Gradient Boosted Trees (Friedman (2001))	depth, learning rate, #Trees=500
Deep Learning (Goodfellow, Bengio, and Courville (2016))	L1 penalty λ #hidden Layers=1, #neurons=4 activation function: ReLU ensemble=10

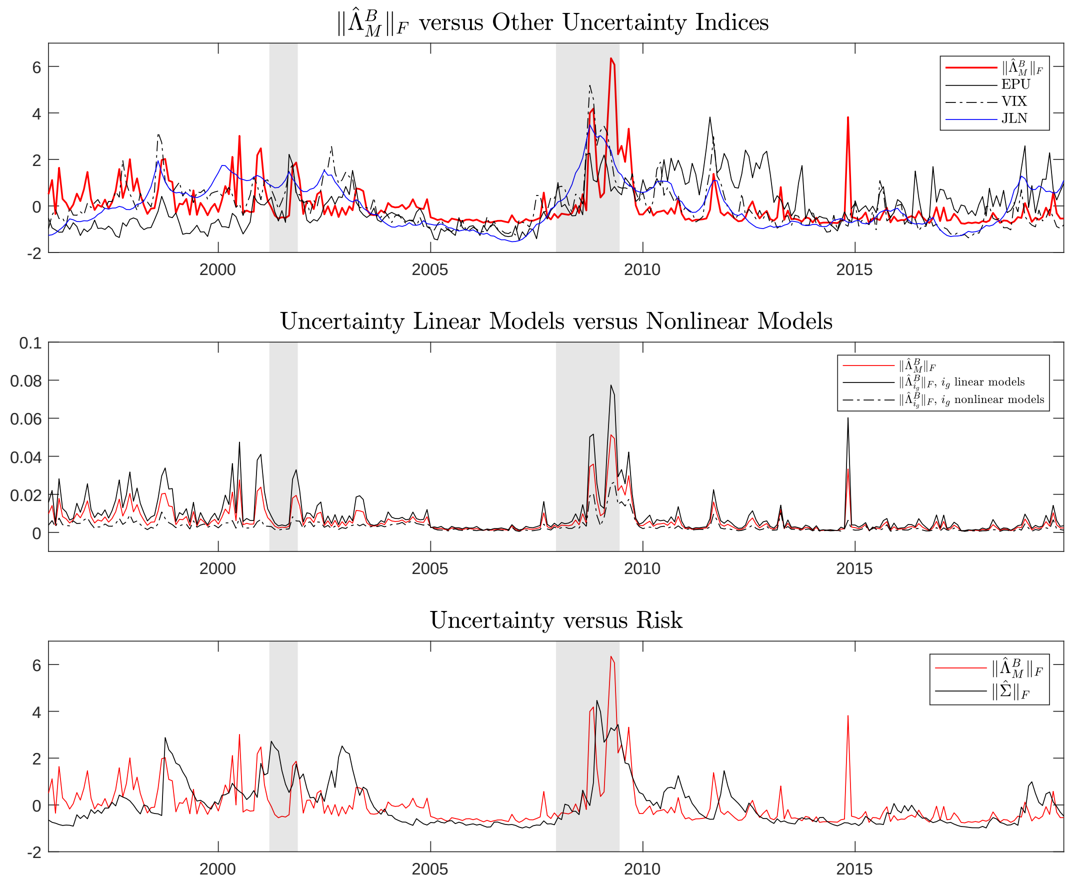

Fig. OA.1. **Alternative Knightian Uncertainty versus other Uncertainty Index and Volatility** The top and bottom panel of the figure plot the Frobenius norm $\|\hat{\Lambda}_{M,t}^B\|_F$ appearing in equation (10) against (a) the Economic Policy Uncertainty Index (EPU), (b) the Jurado et al. financial uncertainty index (JLN), (c) the CBOE Volatility Index (VIX), and finally (d) the Frobenius norm of the DCC(1,1) covariance matrix $\|\Sigma_t\|_F$ following the estimation procedure of Engle (2002). The series in the top and bottom panel are normalized. The middle panel plots the Frobenius norm $\|\hat{\Lambda}_{M,t}^B\|_F$ appearing in equation (10) against the Frobenius norm $\|\hat{\Lambda}_{i_g,t}^B\|_F$ appearing in equation (9) for i_g representing the group of all regularized linear models or the group of all nonlinear machine learning models. The sample period is from January 1996 to December 2019. The alternative uncertainty estimator is computed via equation OA.1.

OA.3 Alternative Uncertainty Measure: Properties and Portfolio Performance

In the main text of the paper, we calculate the bootstrap-based composite uncertainty estimator by treating all machine learning models symmetrically. The expected return vector is also estimated by averaging the model return forecasts. However, given that the number of linear machine learning models we consider is larger than the number of nonlinear models, the expected return and conditional uncertainty estimate may be skewed towards the linear model predictions.

In this section, we show that our main findings continue to hold when the group of linear models and the group of nonlinear models contribute equally to the expected return and composite uncertainty estimates. Specifically, we first compute $\hat{\Lambda}^B_{L,t}$ following equation 19 only using the group of linear machine learning models, and $\hat{\Lambda}^B_{NL,t}$ only using the group of nonlinear models, and then calculate the alternative composite uncertainty estimator via

$$\hat{\Lambda}^B_{L,t} = \frac{1}{2}\hat{\Lambda}^B_{L,t} + \frac{1}{2}\hat{\Lambda}^B_{NL,t} \tag{OA.1}$$

To be consistent, we compute the expected return $\hat{\mu}_{L,t}$ by first averaging the linear model predictions on the actual data and $\hat{\mu}_{NL,t}$ by averaging the nonlinear model predictions on the actual data, and then estimate the alternative expected return following

$$\hat{\mu}_t = \frac{1}{2}\hat{\mu}_{L,t} + \frac{1}{2}\hat{\mu}_{NL,t} \tag{OA.2}$$

The time series plot of the alternative composite uncertainty estimator versus other uncertainty indices and volatility is reported in Figure OA.1. The out-of-sample utility ratio between the Markowitz with uncertainty portfolio and the "MV" and "EW" portfolio is presented in Tables OA.6 and OA.7. Note that although the linear and nonlinear uncertainty estimator is computed in the same way as in the main text, the utility ratio between the "Linear UNC" and "Non-linear UNC" portfolio and the benchmark portfolios is not identical to the results in Tables 7 and 8, because the portfolio optimization involves a new expected return estimator.

Table OA.6. Out-of-Sample Utility: UNC Portfolio versus MV Portfolio, Alternative Uncertainty and Expected Return Estimator

	$\delta = 0$	$\delta = 2$	$\delta = 4$	$\delta = 6$	$\delta = 8$
Panel A: UNC with $\hat{\Lambda}^A_{M,t}$ / MV					
$\gamma = 2$	1.00	1.23***	1.68***	2.68***	5.97***
$\gamma = 4$	1.00	1.23***	1.60***	2.24***	3.53***
$\gamma = 6$	1.00	1.16***	1.40***	1.71***	2.19***
$\gamma = 8$	1.00	1.13***	1.30***	1.50***	1.77***
$\gamma = 10$	1.00	1.12**	1.25***	1.40***	1.60***
Panel B: UNC with $\hat{\Lambda}^A_{i_g,t}$, i_g linear models / MV					
$\gamma = 2$	1.00	1.23***	1.46***	1.99***	3.52***
$\gamma = 4$	1.00	1.12***	1.33***	1.73***	2.61***
$\gamma = 6$	1.00	1.11***	1.29***	1.60***	2.19***
$\gamma = 8$	1.00	1.08***	1.21***	1.41***	1.72***
$\gamma = 10$	1.00	1.07***	1.18***	1.33***	1.54***
Panel C: UNC with $\hat{\Lambda}^A_{i_g,t}$, i_g nonlinear models / MV					
$\gamma = 2$	1.00	1.01***	1.06***	1.14***	1.24***
$\gamma = 4$	1.00	1.05***	1.18***	1.39***	1.66***
$\gamma = 6$	1.00	1.14**	1.30***	1.50***	1.74***
$\gamma = 8$	1.00	1.10**	1.20***	1.32***	1.45***
$\gamma = 10$	1.00	1.06***	1.13***	1.21***	1.29***

Notes - The table compares the average out-of-sample utility for an uncertainty/risk averse versus risk averse portfolio under different relative risk aversion parameters γ and uncertainty aversion parameters δ, where the investor uses different machine learning models and associated uncertainty measures $\hat{\Lambda}^B_{M,t}$ and $\hat{\Lambda}^B_{i_g,t}$ for i_g respectively linear and nonlinear models. The alternative uncertainty estimator is computed via equation OA.1 and the alternative expected return estimator is computed via equation OA.2. Out-of-sample returns are at monthly frequency. One, two, and three stars denote statistical significance of Diebold and Mariano (2002) test at the 1%, 5%, and 10% level respectively.

TABLE OA.7. Out-of-Sample Utility: UNC Portfolio versus EW Portfolio, Alternative Uncertainty and Expected Return Estimator

	$\delta = 0$	$\delta = 2$	$\delta = 4$	$\delta = 6$	$\delta = 8$
			Panel A: UNC with $\hat{\Lambda}^A_{M,t}$ / MV		
$\gamma = 2$	0.82***	0.87***	0.96***	1.10***	1.25**
$\gamma = 4$	0.94***	1.06***	1.21*	1.39	1.60
$\gamma = 6$	1.19**	1.34	1.56	1.80	2.14**
$\gamma = 8$	1.48	1.73	2.07**	2.53***	3.27***
$\gamma = 10$	1.93*	2.39***	3.08***	4.29***	7.21***
			Panel B: UNC with $\hat{\Lambda}^A_{i_g,t}$, i_g linear models / EW		
$\gamma = 2$	1.30***	1.47***	1.55	1.72	2.04
$\gamma = 4$	1.44***	1.56	1.76	2.11	2.74*
$\gamma = 6$	1.61	1.83	2.20**	2.89***	4.43***
$\gamma = 8$	1.91	2.31***	3.06***	4.84***	12.98***
$\gamma = 10$	2.43***	3.26***	5.34***	17.32***	–
			Panel C: UNC with $\hat{\Lambda}^A_{i_g,t}$, i_g nonlinear models / EW		
$\gamma = 2$	0.86***	0.82***	0.80***	0.80***	0.80***
$\gamma = 4$	0.74***	0.73***	0.76***	0.81***	0.87***
$\gamma = 6$	0.83***	0.90***	0.98***	1.05**	1.13
$\gamma = 8$	1.10	1.20	1.30	1.41	1.53
$\gamma = 10$	1.49	1.63	1.79*	2.00**	2.25***

Notes - The table compares the average out-of-sample utility for an uncertainty/risk averse versus equally weighted portfolio under different relative risk aversion parameters γ and uncertainty aversion parameters δ, where the investor uses different machine learning models and associated uncertainty measures $\hat{\Lambda}^B_{M,t}$ and $\hat{\Lambda}^B_{i_g,t}$ for i_g respectively linear and nonlinear models. The alternative uncertainty estimator is computed via equation OA.1 and the alternative expected return estimator is computed via equation OA.2. Out-of-sample returns are at monthly frequency. One, two, and three stars denote statistical significance of Diebold and Mariano (2002) test at the 1%, 5%, and 10% level respectively. In Panel B, the EW portfolio has negative utility value when $\gamma = 10$ and $\delta = 8$. We omit the negative utility ratio from the table under this (γ, δ) specification.

OA.4 Uncertainty Estimator using only Ordinary Least Squares Model

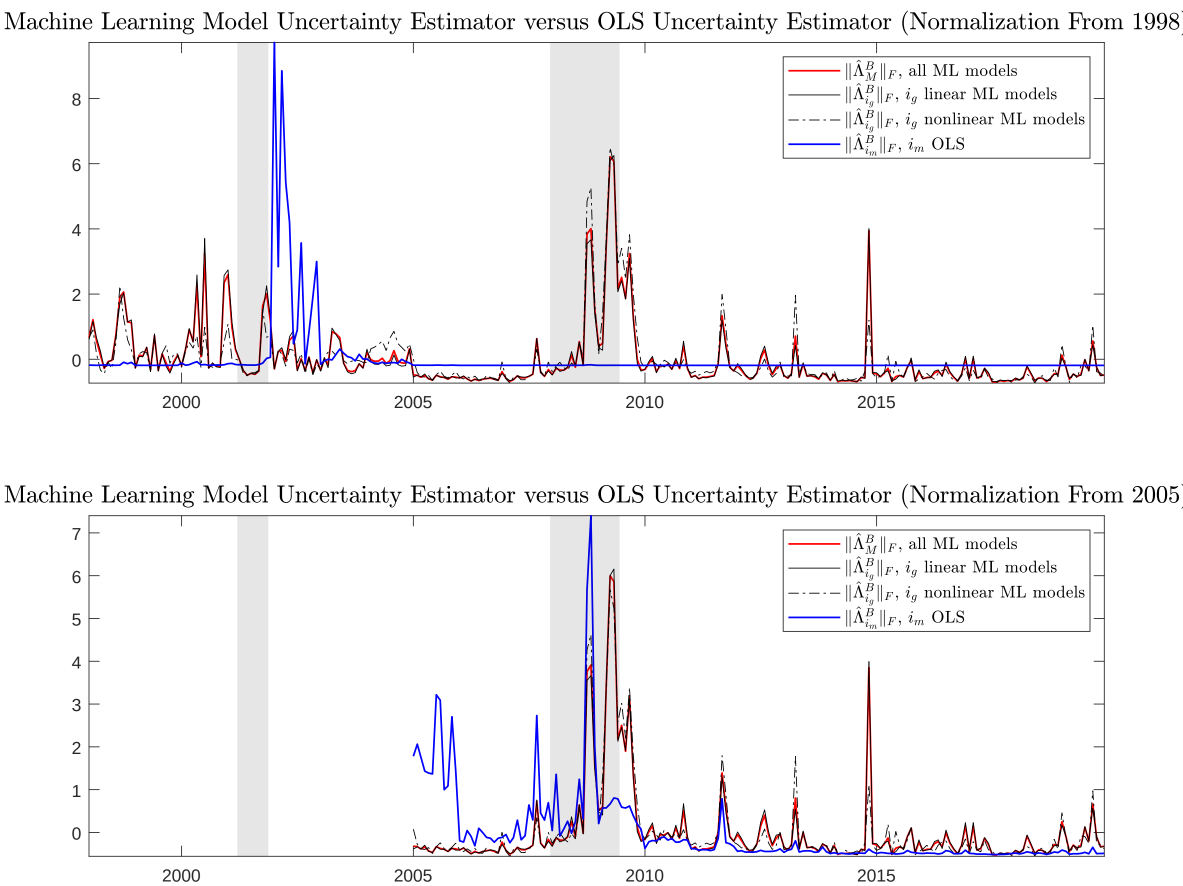

Fig. OA.2. **Machine Learning Model Uncertainty Estimator versus OLS Model Uncertainty Estimator** The figure plots the Frobenius norm $\|\hat{\Lambda}^B_{M,t}\|_F$ appearing in equation (10) and $\|\hat{\Lambda}^B_{i_g,t}\|_F$ appearing in equation (9) for i_g representing the group of all regularized linear predictions or the group of all nonlinear machine learning predictions, against the Frobenius norm $\|\hat{\Lambda}^B_{i_m,t}\|_F$ appearing in equation (8) for i_m representing the ordinary least squares model. The series in the top are normalized from 1998 and the series in the bottom are normalized from 2005.