



## Management Science

Publication details, including instructions for authors and subscription information:  
<http://pubsonline.informs.org>

### Deep Parametric Portfolio Policies

Frederik Simon, Sebastian Weibels, Tom Zimmermann

To cite this article:

Frederik Simon, Sebastian Weibels, Tom Zimmermann (2026) Deep Parametric Portfolio Policies. Management Science

Published online in Articles in Advance 03 Jun 2026

. <https://doi.org/10.1287/mnsc.2025.00721>

This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as “*Management Science*. Copyright © 2026 The Author(s). <https://doi.org/10.1287/mnsc.2025.00721>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by/4.0/>.”

Copyright © 2026 The Author(s)

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

# Deep Parametric Portfolio Policies

Frederik Simon,<sup>a</sup> Sebastian Weibels,<sup>b,c</sup> Tom Zimmermann<sup>b,c,\*</sup>

<sup>a</sup>LIQID Asset Management, 10707 Berlin, Germany; <sup>b</sup>Department of Econometrics and Statistics, University of Cologne, 50931 Cologne, Germany; <sup>c</sup>Centre for Financial Research, 50923 Cologne, Germany

\*Corresponding author

Contact: [frederikahsimon@gmail.com](mailto:frederikahsimon@gmail.com), <https://orcid.org/0009-0008-5102-3313> (FS); [weibels@wiso.uni-koeln.de](mailto:weibels@wiso.uni-koeln.de),

<https://orcid.org/0000-0001-8711-5393> (SW); [tom.zimmermann@uni-koeln.de](mailto:tom.zimmermann@uni-koeln.de), <https://orcid.org/0000-0002-0368-5303> (TZ)

Received: February 21, 2025

Revised: October 31, 2025

Accepted: December 30, 2025

Published Online in Articles in Advance:  
June 3, 2026

<https://doi.org/10.1287/mnsc.2025.00721>

Copyright: © 2026 The Author(s)

**Abstract.** We consider parametric portfolio policies of any complexity using deep neural networks to optimize investor utility. Risk aversion acts as an economic regularization mechanism, with higher risk aversion constraining model complexity. Empirically, Deep Parametric Portfolio Policies generate 43–102 basis points higher monthly certainty equivalent returns compared with linear policies. Looking beyond expected returns, nonlinear portfolio policies better capture the complex relationship between investor preferences and firm characteristics but the benefits of using complex models vary with investor preferences. Results hold across different utility functions and remain robust to transaction costs and short-selling restrictions. Overall, economic regularization constrains model complexity much like statistical regularization but emerges endogenously from investor preferences.

**History:** Accepted by Kay Giesecke, finance.

**Open Access Statement:** This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as “Management Science. Copyright © 2026 The Author(s). <https://doi.org/10.1287/mnsc.2025.00721>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by/4.0/>.”

**Supplemental Material:** The online appendix and data files are available at <https://doi.org/10.1287/mnsc.2025.00721>.

**Keywords:** portfolio choice • machine learning • expected utility

## 1. Introduction

Consider the formidable problem of an investor who wants to choose an optimal asset allocation within her equity portfolio. The literature provides her with a few options: she can opt for a traditional Markowitz approach (Markowitz 1952) which requires estimating expected returns, variances and covariances, with the number of moments to estimate increasing rapidly in the number of assets. At the other end of the spectrum, she might estimate a low-dimensional parametric portfolio policy (PPP) (Brandt et al. 2009) but a linear model might not provide sufficient flexibility. She can also consult a large literature that relates characteristics to expected returns but even studies that consider a multitude of firm-level characteristics (e.g., Gu et al. 2020) only investigate expected returns and do not speak to risk as perceived by different investors’ objective functions.

We provide a general solution to the portfolio optimization challenge. In short, we combine the parametric portfolio policy approach that can estimate portfolio weights for any utility function with the flexibility of feed-forward networks from the machine learning literature. The resulting approach that we label *Deep Parametric*

*Portfolio Policy* (DPPP) is well-suited to accommodate flexible nonlinear and interactive relationships between portfolio weights and stock characteristics, to integrate different utility functions, to deal with leverage or portfolio weight constraints, and to incorporate transaction costs. Importantly, the model also allows us to study the relationship between model complexity and investor preferences.

The contributions of our paper are fourfold. First, we advance the theoretical literature by formally linking investor risk preferences with effective model complexity through the mechanism of economic regularization. Second, we extend the parametric portfolio policy framework by integrating deep neural networks to capture nonlinear and interactive effects, thereby offering a more flexible approach to portfolio optimization. Third, our empirical results show that the DPPP delivers economically meaningful utility gains over traditional linear models across a variety of settings, objective functions and benchmark models. Notably, investor preferences shape the optimal portfolio allocation in a manner similar to statistical regularization, but they do so with respect to the objectives investors actually care about. Fourth, our analysis of variable importance contributes to a deeper

understanding of how different types of firm characteristics influence portfolio construction, particularly under varying degrees of risk aversion and preferences.

To the best of our knowledge, our study is the first to systematically explore how the benefits of a complex and flexible model vary for investors with different levels of risk aversion or different utility functions. A natural concern with parameter-rich models is their potential to overfit historical data. Overfitting leads to less reliable out-of-sample estimates and higher prediction variance. Since our portfolio policy approach maximizes the investor's objective function directly (as opposed to minimizing a statistical objective such as the squared distance between realized and predicted returns (Moritz and Zimmermann 2016, Gu et al. 2020)), volatility of predictions becomes a systematic part of the economic objective. As risk aversion increases, the variance of portfolio returns becomes more important and leans against overfitting and thus model complexity. We refer to this mechanism as *economic regularization* (in contrast to purely statistically motivated regularization techniques), and present theoretical and simulation-based findings which demonstrate that models with different degrees of complexity converge as risk aversion increases.

Our empirical work represents a significant conceptual departure from linear parametric portfolio policies in two ways: first, by replacing the linear specification with a neural network, we allow for nonlinearities and interactions in the relationship between firm characteristics and portfolio weights. Research on using machine learning for return prediction shows that such flexibility is relevant to model the relationship between firm characteristics and future returns and can lead to substantial improvements over less flexible specifications (Moritz and Zimmermann 2016, Freyberger et al. 2020, Gu et al. 2020). It is conceivable that such flexibility will also help to model the relation between portfolio weights and firm characteristics. Second, this flexibility comes at the cost of having to estimate a model with a high-dimensional parameter vector. This is a deviation from the original motivation of the parametric portfolio policy literature which aims to reduce portfolio optimization to a low-dimensional problem with only a small number of coefficients that need to be estimated. Kelly et al. (2024) argue that model complexity is a virtue for return prediction, and our approach can be viewed as an exploration of that point in the context of parametric portfolio policies.

Our empirical investigation further underscores the value of the DPPP approach. Utilizing a comprehensive data set of firm-level characteristics, we document substantial improvements in investor utility when using the DPPP relative to a standard linear model. Empirically, our complex model significantly improves over a standard linear parametric portfolio policy, with certainty equivalent gains ranging from about 43 basis

points to 102 basis points. Further, in line with our theoretical results, we find that the benefit of model complexity decreases when an investor's risk aversion increases.

While our benchmark investor is a classical constant relative risk aversion (CRRA) optimizer, our setup easily accommodates other utility functions. We explore portfolio policies with and without transaction costs and short-selling constraints, as well as for different utility functions, such as mean–variance and loss aversion preferences, and a statistically motivated utility function. We find that economic regularization via risk aversion has similar dampening effects on higher-order portfolio moments as statistical regularization but economic regularization targets precisely the investor's objective function and thereby the risks most relevant to their investment decisions. Overall, complex portfolio policies can be beneficial in all these scenarios, but utility gains are higher for lower risk (or loss) aversion.

Beyond the aggregate performance improvements, our study offers novel insights into the relative importance of different types of firm characteristics. Past return-based stock characteristics turn out to be more important to the portfolio policy than accounting-based characteristics. However, while prior research has highlighted the dominance of return-based signals in asset pricing, our results indicate that the inclusion of a large set of predictors with both return-based and accounting-based measures leads to a more balanced importance profile as risk aversion increases, extending the existing literature that examines the importance of firm characteristics under economic constraints (DeMiguel et al. 2020, Jensen et al. 2022).

Overall, our work bridges the gap between traditional portfolio optimization and modern machine learning techniques. By directly mapping firm characteristics to portfolio weights through neural networks, we offer a flexible, robust, and economically intuitive framework that adapts to the complexities of real-world investment challenges. This novel approach not only improves upon classical methods in terms of performance but also advances our theoretical understanding of how investor preferences can naturally regulate model complexity in high-dimensional settings.

### 1.1. Related Literature

Our work relates to several different streams of the literature. First, we add to a growing literature that explores the potential of machine learning algorithms in finance (e.g., Heaton et al. 2017, Bianchi et al. 2020, Gu et al. 2020, Kelly et al. 2024). Studies in this literature typically consider a prediction task (e.g., predicting stock returns), and minimize a statistical loss function such as the mean squared error (or a related distance metric) between the actual and predicted values. Predicted values are used to construct portfolio weights (e.g., Gu et al. 2020). In contrast, our methodology uses machine learning to provide

a direct, one-step mapping from firm characteristics to portfolio weights that explicitly targets the investor's objective rather than a statistical loss function.

Second, our paper serves as a natural extension of the parametric portfolio approach by Brandt et al. (2009). While Brandt et al. (2009) argue that it may be worthwhile to consider nonlinear functions and interactions in weight modeling, subsequent papers that have implemented and extended parametric portfolio policies parameterize portfolio weights as a linear function of firm characteristics (e.g., Hjalmarsson and Manchev 2012, Ammann et al. 2016). DeMiguel et al. (2020) incorporate transaction costs, a larger set of firm characteristics, and statistical regularization but stay within the linear framework. Our DPPP replaces the linear model with a feed-forward neural network that accounts for both nonlinearity and possible interactions of firm characteristics. In addition, we use a larger set of firm characteristics than previous studies and explore different utility functions, constraints, and degrees of risk aversion.

Further, we contribute to the literature that employs alternative methods to direct portfolio optimization via machine learning. Particularly relevant approaches in our context include Cong et al. (2020), Chevalier et al. (2022), Jensen et al. (2022), Guijarro-Ordóñez et al. (2025), Coulombe and Goebel (2023), and Feng et al. (2022). Each of these differs from ours in one or more aspects. Cong et al. (2020) propose a reinforcement learning-based approach (as opposed to our feed-forward framework) and connect to a related literature in computer sciences that puts additional emphasis on more technical parts of the model implementation. Our study naturally connects to the preceding finance literature, and generalizes the approach of Brandt et al. (2009) to explicitly analyze differences between a linear and nonlinear specification for different utility functions, constraints, and levels of risk aversion, and we derive theoretical results for the convergence of these specifications under economic regularization. Chevalier et al. (2022) derive optimal in-sample weights based on investor preferences and subsequently predict these weights conditional on covariates. This is conceptually different from our approach, primarily because we do not require the preprocessing step of computing the optimal in-sample weights. Jensen et al. (2022) take a different approach. They specifically address the issue of integrating transaction costs into mean-variance portfolio optimization with machine learning. While their focus is the derivation of an efficient frontier including transaction costs, we explicitly analyze how different types of investor preferences and constraints affect the benefit of complexity in portfolio optimization. Guijarro-Ordóñez et al. (2025) also employ neural networks for portfolio optimization. Their framework, however, is grounded in

statistical arbitrage, whereas we directly map portfolio weights to stock-specific signals. Coulombe and Goebel (2023) propose a machine learning framework for directly optimizing portfolio weights with nonlinear algorithms, building on Lo and MacKinlay's (1997) maximally predictable portfolio approach. Their method aligns with mean-variance utility maximization. In contrast, our framework supports any utility function, offering broader flexibility beyond mean-variance preferences. Liu et al. (2021) propose an alternative one-step optimization approach, mapping predictors to optimal portfolio weights through genetic programming. Our methodology leverages feed-forward neural networks while incorporating a substantially larger set of stock characteristics and practical constraints such as transaction costs and leverage limits. Feng et al. (2022) employ feed-forward neural networks to estimate portfolio weights by modeling a deep factor, that is, a long-short factor based on a nonlinear combination of characteristics. They apply their method to bond data.

Our work also connects to the growing literature on machine learning approaches for estimating stochastic discount factors (SDFs). In this stream, Kozak et al. (2020) propose shrinking the cross-section of returns into a parsimonious set of factors that price all assets. Chen et al. (2024) use deep neural networks to construct a flexible, high-dimensional SDF, showing improved explanatory power for cross-sectional returns. Similarly, Bryzgalova et al. (2025) introduce a decision-tree-based method for constructing managed portfolios that span the SDF. Their approach focuses on identifying a parsimonious and interpretable set of characteristic-based portfolios that capture complex nonlinear interactions while remaining tractable and economically meaningful. In contrast, our paper sidesteps the explicit construction of test assets for SDF estimation and instead directly parameterizes and estimates the portfolio-weight function in a one-step framework.

In addition, our work relates to the literature that deals with estimation risk arising from parameter uncertainty (Kirby and Ostdiek 2012a, b; Lassance et al. 2024) and the literature that explores regularization in this regard through economic mechanisms (Jagannathan and Ma 2003, Skouras 2007, Hautsch and Voigt 2019). Adding to the literature, we explore how risk aversion affects model complexity and uncertainty, as well as how this affects the difference between models of different complexity.

Finally, our paper relates to research that explicitly analyzes how transaction costs and other forms of optimization constraints impact portfolio choice (DeMiguel et al. 2020, Jensen et al. 2022, Detzel et al. 2023). Complementing the literature, we study how nonlinearities contribute to the portfolio optimization, and how risk aversion regularizes optimization on top of and beyond the effects of transaction costs.

## 2. Theory

### 2.1. Expected Utility Framework and Parametric Portfolio Policies

The starting point of our framework is the parametric portfolio policy model in Brandt et al. (2009). Consider a universe of  $N_t$  stocks that an investor can invest in at each month  $t \in T$ . Following Brandt et al. (2009) and to focus on the rich dynamics of risky asset allocations, we do not include a risk-free asset.<sup>1</sup> Each stock  $i$  is associated with a vector of firm characteristics  $x_{i,t}$  and a return  $r_{i,t+1}$  from date  $t$  to  $t+1$ . The investor maximizes the conditional expected utility of future portfolio returns  $r_{p,t+1}$ :

$$\max_{\{w_{i,t}\}_{i=1}^{N_t}} E_t [u(r_{p,t+1})] = E_t \left[ u \left( \sum_{i=1}^{N_t} w_{i,t} r_{i,t+1} \right) \right], \quad (1)$$

where  $w_{i,t}$  is the weight of stock  $i$  in the portfolio at date  $t$  and  $u(\cdot)$  denotes the respective utility function.

Instead of directly deriving the weights  $w_{i,t}$  (as e.g., following the traditional Markowitz approach), we follow Brandt et al. (2009) and parameterize the weights as a function of firm characteristics  $x_{i,t}$ , that is,

$$w_{i,t} = f(x_{i,t}; \theta), \quad (2)$$

where  $\theta$  is the coefficient vector to be estimated.

The parameter vector  $\theta$  remains constant across assets  $i$  and periods  $t$ , that is, it maximizes the conditional expected utility at every period  $t$ . This necessarily implies that  $\theta$  also maximizes the unconditional expected utility. Hence, one can estimate  $\theta$  by maximizing the unconditional expected utility via the return distribution's sample analogues:

$$\max_{\theta} \frac{1}{T} \sum_{t=1}^T u(r_{p,t+1}(\theta)) = \frac{1}{T} \sum_{t=1}^T u \left( \sum_{i=1}^{N_t} f(x_{i,t}; \theta) r_{i,t+1} \right). \quad (3)$$

The idea behind parametric portfolio policies is that one may exploit firm characteristics in order to tilt some benchmark portfolio toward stocks that increase an investor's utility, so that  $f(\cdot)$  can be expressed as

$$w_{i,t} = b_{i,t} + \frac{1}{N_t} g(x_{i,t}; \theta), \quad (4)$$

where  $b_{i,t}$  denotes benchmark portfolio weights such as the equally weighted or value weighted portfolio and  $x_{i,t}$  denotes the characteristics of stock  $i$ , standardized cross-sectionally to have zero mean and unit standard deviation in each cross section  $t$ .<sup>2</sup>

In essence, our model can be interpreted as a generalization of the linear parametric portfolio policy approach, as we allow  $x_{i,t}$  to enter the model flexibly. Brandt et al. (2009) and the subsequent literature (e.g., DeMiguel et al. 2020) restrict firm characteristics to affect the portfolio in a linear, additive manner. In contrast, we model  $g(\cdot)$  in Equation (4) as a feed-forward neural network, arguably one of the most flexible forms. As discussed in the

introduction, this represents a significant conceptual deviation from the literature in at least two respects: first, by replacing the linear specification with a neural network, we allow the relationship between firm characteristics and weights to be nonlinear, and we account for potential interactions of firm characteristics, in line with the recent literature that finds that such flexibility can be important to predict returns (Moritz and Zimmermann 2016, Freyberger et al. 2020, Gu et al. 2020). Here, our approach explores whether such flexibility also helps to model the relationship between *portfolio weights* and firm characteristics. Second, this flexibility comes at the cost of having to estimate a model with a high-dimensional parameter vector. Thus, it departs from the original motivation of the parametric portfolio policy literature, which aimed to reduce portfolio optimization to a low-dimensional problem where only a small number of coefficients need to be estimated. In fact, our benchmark model below has about 5,700 to 5,900 parameters compared with the three parameters that need to be estimated when following Brandt et al. (2009).

### 2.2. Risk Aversion as Economic Regularization

This section establishes the theoretical underpinnings for how risk aversion serves as an economic regularization mechanism in our setting. Our key insight is that risk aversion naturally constrains model complexity when estimation risk is a concern. Intuitively, estimation risk arises from the uncertainty about the parameters of the data generating process. This leads to errors in the estimation of portfolio weights which increases portfolio risk (Kirby and Ost diek 2012b, Lassance et al. 2024). As risk aversion increases, the investor places a greater penalty on portfolio return variance, leading to more conservative portfolios with simpler investment strategies.

Formally, when portfolio returns are evaluated under the predictive distribution, the law of total variance implies that estimation uncertainty about model parameters  $\theta$  enters expected utility as an *additional* source of variance. Risk aversion penalizes not only intrinsic return risk but also the excess variance induced by parameter estimation. For example, in the mean-variance case with one risky asset, Brandt (2010) shows that the certainty equivalent loss due to estimation uncertainty is proportional to the risk aversion coefficient.

This stands in contrast to previous approaches to regularization in portfolio choice. While Hautsch and Voigt (2019) and Jagannathan and Ma (2003) show that transaction costs and short-selling constraints can serve as economically motivated penalties, our framework demonstrates how investor preferences themselves create a natural regularization mechanism. In the following two subsections, we formalize this idea through two approaches: first, from an economic perspective, and then using results from statistical learning theory.

**2.2.1. Economic Intuition.** To establish economic intuition, consider a CRRA investor maximizing expected utility over portfolio returns as in Equation (1). Following Brandt et al. (2009), we express portfolio returns as:

$$r_{p,t+1}(\theta) = b_t^T r_{t+1} + \theta^T X_t^T r_{t+1}/N_t = r_{b,t+1} + \theta^T r_{c,t+1}, \quad (5)$$

where  $r_{b,t+1}$  is the benchmark return and  $r_{c,t+1}$  contains characteristic portfolio returns. As outlined in Didisheim et al. (2023), we can interpret this linear portfolio framework through a neural network lens by replacing our characteristic-sorted portfolios  $X_t^T r_{t+1}/N_t$  with transformed portfolios  $S_t^T r_{t+1}/N_t$ , where  $S_t$  represents nonlinear transformations of the original characteristics.

The following proposition shows how the active deviations from the benchmark defined in Equation (4) critically depend on risk aversion:

**Proposition 1.** *Define the optimal active deviations from the benchmarks:*

$$\frac{1}{N_t} \theta^{*T} S_t = \underbrace{\frac{1}{\gamma} \hat{\Sigma}_c^{-1} \hat{\mu}_c^T S_t}_{\text{Risk premium term}} - \underbrace{\hat{\Sigma}_c^{-1} \hat{\sigma}_{bc}^T S_t}_{\text{Risk minimization term}}. \quad (6)$$

*The optimal deviations from the benchmark portfolio converge as follows:*

$$\lim_{\gamma \rightarrow \infty} \frac{1}{N_t} \theta^T S_t = -\hat{\Sigma}_c^{-1} \hat{\sigma}_{bc}^T S_t. \quad (7)$$

The proof is in Appendix A.1.

In particular, if the benchmark is uncorrelated with the active positions (i.e.,  $\hat{\sigma}_{bc} = 0$ ), then the active positions converge to zero. If the benchmark portfolio is correlated with active positions, that is,  $\hat{\sigma}_{bc} \neq 0$ , then as  $\gamma \rightarrow \infty$  the optimal active tilt converges to the benchmark-hedge  $-\hat{\Sigma}_c^{-1} \hat{\sigma}_{bc}$  rather than zero. Economically, higher  $\gamma$  shuts down premium-seeking tilts but preserves positions that hedge benchmark risk. Only under the special case  $\hat{\sigma}_{bc} = 0$  do all active tilts vanish in the limit.

High risk aversion forces both the PPP and the DPPP to prioritize risk minimization, leading to convergence for the mean absolute weight differences between the models. We summarize this in the following proposition, which follows directly and trivially from Proposition 1:

**Proposition 2.** *Compare two models of different complexity, with corresponding portfolio weights  $w_{PPP}(\gamma)$  and  $w_{DPPP}(\gamma)$ . Let  $\Delta_{RP}$  and  $\Delta_{RM}$  be the differences in the risk premium and risk minimization terms, respectively, from Equation (7). The absolute portfolio weight difference can be decomposed as*

$$\lim_{\gamma \rightarrow \infty} \|w_{PPP}(\gamma) - w_{DPPP}(\gamma)\| = \lim_{\gamma \rightarrow \infty} \underbrace{\frac{1}{\gamma} \|\Delta_{RP}\|}_{=0} + \lim_{\gamma \rightarrow \infty} \|\Delta_{RM}\|, \quad (8)$$

where  $\Delta_{RP}$  and  $\Delta_{RM}$  are deterministic quantities independent of  $\gamma$  that depend only on the problem structure ( $X_t$ ,  $b_t$ , and  $r_{t+1}$ ).

*If both models build active portfolios orthogonal to the benchmark, then the gap between the PPP and DPPP weights converges to zero, otherwise a constant hedge difference can remain.*

In Section S.2 in the Supplementary Appendix we provide a similar intuition for loss aversion preference.

This convergence in weights is consistent with the classical separation theorem. In our mean–variance setting, all investors share the same risky tangency portfolio and risk aversion only scales overall exposure. As risk aversion increases, the premium-seeking component vanishes and the DPPP weights contract toward those of the PPP, leading to the convergence established in Proposition 2. In this sense, the convergence in weights corresponds to investors moving toward a common region of the efficient set where model flexibility has little incremental effect.

**2.2.2. Complexity Interpretation.** The concept of risk aversion as a regularization mechanism is also grounded in statistical learning theory. Following Skouras (2007), estimation uncertainty directly affects economic decisions through the utility function. This leads to a natural complexity measure known as the *effective degrees of freedom* (EDF), originally developed by Murata et al. (1994) which is defined as:

$$\text{EDF} = \text{tr}(G^{-1}V)/T, \quad (9)$$

where  $G$  is the hessian of the model with respect to the parameters and  $V$  is the outer product of the gradients of the model with respect to the parameters.

This measure has an intuitive interpretation: for linear models like the PPP, the trace reduces to the parameter count  $p$ , making EDF interpretable as the number of “effective” parameters in more complex models. This leads to our key result about model complexity.

**Proposition 3.** *As risk aversion increases, the effective model complexity converges to zero*

$$\lim_{\gamma \rightarrow \infty} \text{EDF} = \lim_{\gamma \rightarrow \infty} \frac{p}{\gamma} = 0. \quad (10)$$

The proof is in Appendix A.2.

This result shows that increasing risk aversion effectively reduces model complexity, providing a theoretical link between risk preferences and model complexity in our portfolio framework. The complexity interpretation provides additional insight into the convergence between the PPP and the DPPP established in the previous section. As  $\gamma \rightarrow \infty$ , the EDF approaches zero for both models, meaning they effectively become less complex regardless of their nominal parameter count. For the PPP, the number of effective parameters  $p/\gamma$  goes to zero. Similarly for the DPPP, despite having a richer nonlinear structure through  $S_t$  its effective complexity also converges to zero as risk aversion increases.

In essence, high risk aversion forces both models to prioritize risk minimization over exploiting their different parametric structures, leading to their convergence.

Naturally, one might wonder how these theoretical results relate to the trade-off between bias (less complex models might be misspecified) and variance (less complex models might have lower estimation variance) in the traditional statistical sense. While traditional bias-variance analysis focuses on minimizing prediction error, we focus on maximizing expected utility from decisions, that is, models are evaluated based on their utility for decision-making rather than statistical criteria. Skouras (2007) shows that when models are potentially misspecified, the standard bias-variance tradeoff is not the most relevant consideration for decision-makers. Instead, models should balance complexity against performance. Expected utility reflects both traditional statistical channels: higher risk aversion amplifies the penalty on estimation variance and thus constrains functional forms. Conversely, low risk aversion allows the model to exploit richer functional forms. Preferences thus endogenously determine the optimal degree of complexity from a decision-theoretic rather than purely statistical perspective. The fact that regularization/shrinkage can lead to improvements in nonstatistical objective functions is not new to academic finance (see e.g., Jorion (1986)). In Section 4.2, we compare model regularization via risk aversion to statistical model regularization, and find that both have dampening effects on portfolio return volatility.

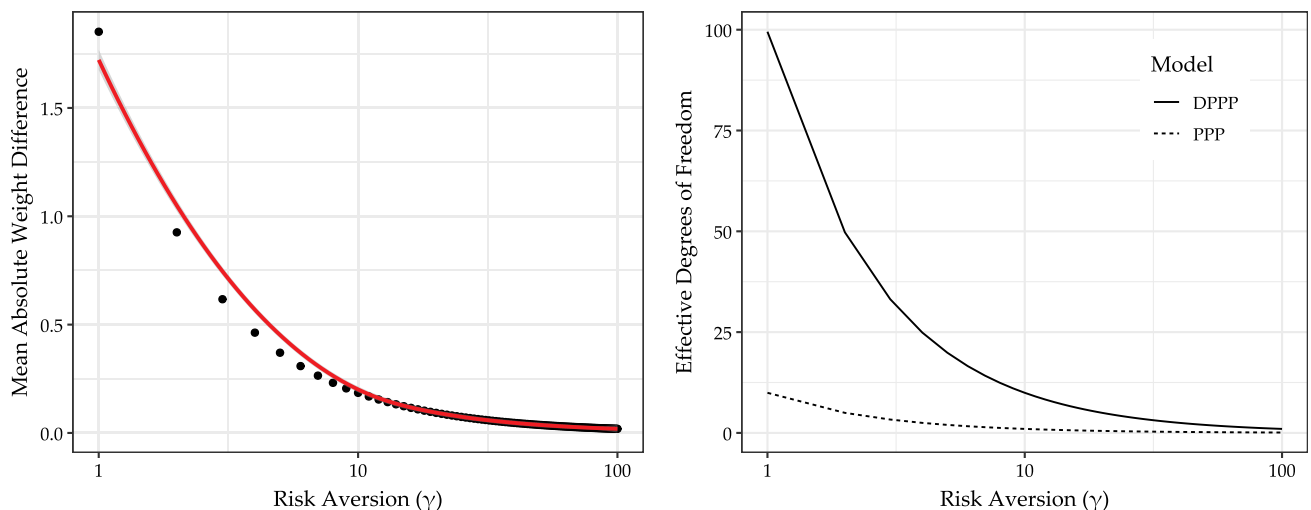
**2.2.3. Simulation Evidence.** We illustrate our theoretical results through a simulation study featuring two

nested parametric portfolio policies that share the same base information set but differ in complexity. We generate a panel of  $N = 100$  firms over  $T = 200$  months with  $K = 10$  base firm characteristics that follow persistent AR(1) processes  $x_{i,k,t} = \rho x_{i,k,t-1} + \epsilon_{i,k,t}$ , where  $\rho = 0.8$  captures the empirically observed persistence in firm characteristics, and  $\epsilon_{i,k,t} \sim N(0, 1 - \rho^2)$ . All characteristics are standardized cross-sectionally. Returns are generated with predictability based on the base characteristics  $r_{i,t+1} = x_{i,t}^T \beta + \eta_{i,t+1}$ , where  $\beta \sim N(0, 0.1^2 I_K)$  and  $\eta_{i,t+1} \sim N(0, 0.15^2)$ . Lastly, both models use an equally-weighted portfolio as a benchmark.

Following our theoretical framework with  $S_t$  representing nonlinear transformations of the original characteristics, we expand the feature space with random Fourier features. Specifically, we draw random vectors  $w^j \stackrel{iid}{\sim} N(0, \eta^2 I)$  for  $j = 1, \dots, p/2$ . For each  $j$ , we create a pair of new features using sine and cosine transformations, where we generate  $p = 10$  features for the simple model (PPP) and  $p = 100$  features for the complex model (DPPP). This approach ensures both models operate on transformations of the same underlying information while differing substantially in their parametric complexity.<sup>3</sup>

Figure 1 presents the key findings from estimating both models across a grid of risk aversion values  $\gamma \in [1, 100]$ . We examine two metrics that directly correspond to our theoretical results: (i) the mean absolute difference in portfolio weights, and (ii) the effective degrees of freedom of each model. Consistent with our theoretical predictions, we find that the weight difference between models decreases and the EDF of both

**Figure 1.** (Color online) Risk Aversion as Economic Regularization



*Notes.* This figure presents simulation evidence demonstrating that risk aversion acts as an economic regularization mechanism. We compare two nested parametric portfolio policies of different complexity: one using 10 characteristics (PPP) and a second one using 100 characteristics (DPPP) constructed through random Fourier transformations of the base characteristics. The left panel shows the mean absolute difference in portfolio weights between models across risk aversion levels. The right panel plots the effective degrees of freedom (EDF) for both models, demonstrating how increasing risk aversion reduces model complexity. All panels use a logarithmic scale (base 10) for risk aversion.

models converge as risk aversion increases, effectively constraining model complexity through the investor's utility function rather than via statistical penalties. Supplementary Appendix S.2 shows similar results for investors with loss aversion utility.

### 3. Estimation and Results

#### 3.1. Network Architecture

We model function  $g(\cdot)$  in Equation (4) as a feed-forward network. Conceptually, our feed-forward networks are structured to estimate optimal portfolio weights and as such differ from networks used in pure prediction contexts in two important ways.

First, the objective of our estimation is to maximize expected utility. Standard use of predictive modeling for stock returns (with or without networks) tries to minimize some distance metric (e.g., the mean squared error) between observed and predicted stock returns. For example, Gu et al. (2020) use neural networks to predict stock returns using a penalized mean squared error as the statistical loss function. In contrast, we follow Brandt et al. (2009) and directly estimate portfolio weights. More specifically, we predict portfolio weights by maximizing the unconditional sample analogue of a utility function as given in Equation (3). For example, in our base case, the loss function  $\mathcal{L}$  that we aim to minimize with respect to  $\theta$  is the constant relative risk aversion (CRRA) utility:

$$\mathcal{L}(\theta) = -\frac{1}{T} \sum_{t=1}^T \left( \frac{(1 + r_{p,t+1}(\theta))^{1-\gamma}}{1-\gamma} \right), \quad (11)$$

where  $\gamma$  is the relative risk aversion parameter. Note that minimizing Equation (11) is equivalent to maximizing CRRA utility.

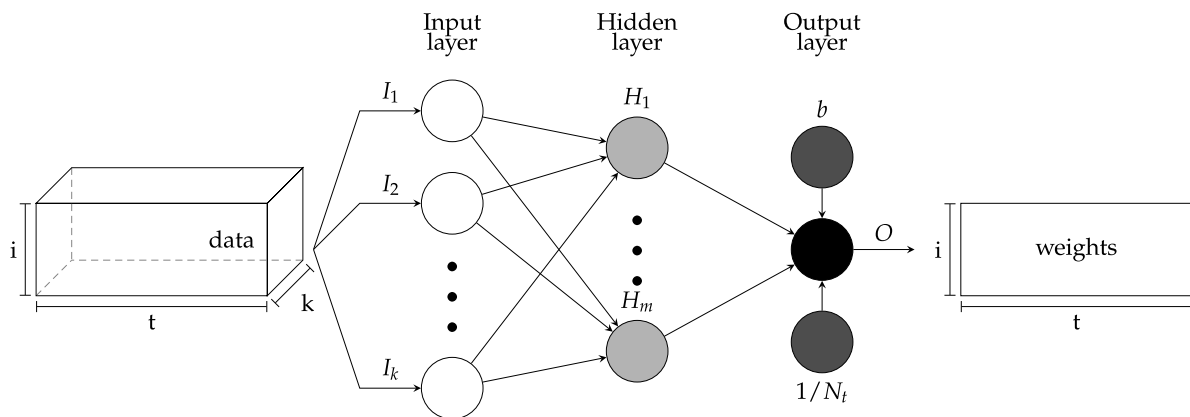
Second, unlike applications that predict stock-level returns using neural nets, our estimated stock-level

portfolio weights are only intermediate outputs of the neural network in that the loss function is based on the *portfolio* return. Hence, we need to aggregate intermediate network outputs (stock-level weights in period  $t$ ; that are a function of stock-level characteristics in period  $t$ ) and stock-level returns in period  $t + 1$  cross-sectionally (see Equations (2)–(4)).

To operationalize this, we maintain the three-dimensional structure of our data (time / stocks / characteristics) where the three-dimensional input tensor reflects the panel structure of the data. Still, portfolio weights at time  $t$  are determined by that period's stock characteristics, maintaining the original spirit of Brandt et al.'s (2009) approach while leveraging the additional flexibility of neural networks to capture cross-sectional nonlinearities. In other words, unlike time series models (such as Recurrent Neural Networks or Long Short-term Memory Networks) that explicitly model sequential dependencies, our network makes independent decisions at each time step based on the concurrent cross-sectional relationships between characteristics and expected returns. This is by design, as our goal is to identify robust cross-sectional patterns rather than temporal dependencies.

Conceptually, our models can be depicted as shown in Figure 2. The input data on the left form a cube (or 3D tensor, the three-dimensional structure described above) with dimensions time  $t$ , stocks  $i$  and input variables  $k$ . Input data are fed into networks with different numbers of hidden layers. In line with Equation (4), the output of the neural network is then normalized by  $1/N_t$  and added to the benchmark portfolio  $b$ . The output of the model  $O$  is a two-dimensional matrix with dimensions  $t \times i$  of portfolio weights for each stock and time period that is then aggregated (as a weighted sum of period  $t + 1$  stock returns) across all stocks in each

Figure 2. Neural Network Structure



Notes. This figure presents the core structure of our neural networks. White circles denote the input layer, grey circles denote the hidden layer and black circles denote the output layer. The data cube on the left depicts the structure of our data, that is, we have  $k$  variables across  $i$  cross-sections in  $t$  periods. The rectangle on the right depicts our output, that is, weights across  $i$  cross-sections in  $t$  periods. The output of the neural network is normalized by  $1/N_t$  and added to the benchmark portfolio  $b$ . The final output is labeled  $O$ .

time period into a portfolio return that is the input of the loss function in Equation (11).

Constructing a neural network requires many design choices, including, for example, the depth (number of layers) and width (units per layer) of the model, or the activation function for different units and layers,<sup>4</sup> and selecting the optimal network architecture is a challenging task. We simplify the process by tuning the number of hidden layers only, evaluating configurations with three, four, and five layers. In every configuration, the first hidden layer starts with 32 nodes, and each subsequent hidden layer contains half as many nodes as the preceding layer.

As discussed in Section 2.1, the network's output needs to be normalized and can then be interpreted as the deviation from a benchmark portfolio. In our application, the benchmark portfolio is the equal-weighted portfolio in all models. A common alternative would be a value-weighted benchmark portfolio where weights are determined by a stock's market capitalization. We stick to the equal-weighted benchmark because of empirical evidence that it outperforms other benchmarks for longer periods (DeMiguel et al. 2009).

Lastly, we impose two constraints to ensure that the model's performance stems from diversified positions with economically reasonable leverage levels rather than from concentrated bets or excessive leverage. First, we impose an ex-ante upper bound on an individual stock's absolute portfolio weight of  $|3\%|$ , that is,

$$|w_{i,t}| \leq 0.03, \quad (12)$$

where  $w_{i,t}$  represents the portfolio weight of stock  $i$  at time  $t$ . Second, we limit leverage to 100% of the portfolio value in any single period during model training.<sup>5</sup> This constraint is formulated for every period  $t$  as

$$\sum_{i=1}^{N_t} w_i I(w_i < 0) \geq -1, \quad (13)$$

where  $I(w_i < 0)$  is an indicator function that equals one if the corresponding portfolio weight is negative and zero otherwise.

Additionally, we maintain the full investment constraint. Due to the nonlinear nature of the model, it is not obvious that the full investment constraint holds (unlike in Brandt et al. (2009)). To make sure that the full investment constraint is satisfied, we standardize the outputs of each unit in the hidden layers cross-sectionally to have zero mean and unit standard deviation across all stocks at date  $t$ . Hence, the output of each node in each hidden layer can be interpreted as a deviation from the benchmark portfolio (see Supplementary Appendix S.1 for details).

We also employ a range of different additional regularization techniques that are standard in the deep learning literature. We give an outline of these techniques

and a more detailed description of the structure of the model including its hyperparameters in S.1.

To estimate our model, we use an expanding-window strategy with a rolling 12-month out-of-sample period (details in Supplementary Appendix S.1).<sup>6</sup> Specifically, we train on the first 20 years, validate on the next 5 years, and then test on the next 12 months. We then roll forward one year at a time, continually re-estimating.

### 3.2. Data

We use the Open Source Asset Pricing data set of Chen and Zimmermann (2022) for the period from January 1971 to December 2020, as comprehensive accounting data are only sparsely available in prior years. In addition, we only keep common stocks, that is, stocks with share codes 10 and 11, and stocks that are traded on the NYSE (exchange code equal to 1) to ensure that results are not driven by small or illiquid stocks. We match the data with monthly stock return data from the Center for Research in Security Prices (CRSP). We drop any observation with missing return, size and/or a return of less than  $-100\%$ . We include continuous firm characteristics from Chen and Zimmermann (2022)'s categories *Price*, *Trading*, *Accounting* and *Analyst*, respectively.

Finally, we follow Gu et al. (2020) and replace missing values with the cross-sectional median at each month for each stock, respectively.<sup>7</sup> Additionally, similar to Gu et al. (2020) we rank all stock characteristics cross-sectionally. As in Brandt et al. (2009) and DeMiguel et al. (2020), each predictor is then standardized to have a cross-sectional mean of zero and standard deviation of one. Note that each predictor is signed so that a larger value implies a higher expected return in the original in-sample period.

Our final data set contains 157 predictors for a total of 5,154 firms. Each month, the data set contains a minimum of 1,213, a maximum of 1,855 and an average of 1,422 firms. These numbers are consistent with a sample of established, liquid companies rather than the broader universe including small and illiquid stocks. Table S.6.2 in the Supplementary Appendix lists the included predictors by original paper. The three columns in the table describe the update frequency of each predictor, the predictor category and the economic category, both taken from Chen and Zimmermann (2022).

### 3.3. Performance Results for CRRA Investors

Table 1 compares the results of the optimization process for CRRA investors with different degrees of risk aversion for our DPPP with its linear counterpart.<sup>8</sup> Results reveal substantial economic gains from employing deep learning in portfolio optimization across different levels of risk aversion. The DPPP consistently outperforms the PPP, with the magnitude and statistical significance of improvements varying systematically with investor risk preferences.

**Table 1.** Deep Portfolio Policy for CRRA Investors with Different Degrees of Risk Aversion

	$\gamma = 2$		$\gamma = 5$		$\gamma = 10$		$\gamma = 20$	
	PPP	DPPP	PPP	DPPP	PPP	DPPP	PPP	DPPP
CE	0.0195	0.0297	0.0163	0.0232	0.0109	0.0152	-0.0006	0.0040
$p$ -value( $CE_{DPPP} - CE_{PPP}$ )		0.0003		0.0002		0.0338		0.0278
$\sum  w_i  / N_i * 100$	0.1696	0.1907	0.1770	0.1938	0.1769	0.1933	0.1690	0.1729
$\max w_i * 100$	0.6815	1.1483	0.7221	0.9843	0.7087	0.8305	0.6710	0.4582
$\min w_i * 100$	-0.6581	-1.2824	-0.6953	-1.2053	-0.6981	-0.9743	-0.6322	-0.7224
$\sum w_i I(w_i < 0)$	-0.7228	-0.8748	-0.7762	-0.8974	-0.7754	-0.8932	-0.7180	-0.7464
$\sum I(w_i < 0) / N_i$	0.3426	0.3400	0.3498	0.3368	0.3490	0.3319	0.3426	0.3202
$\sum  w_{i,t} - w_{i,t-1}^+ $	1.3426	2.6342	1.4571	2.6022	1.4224	2.3813	1.2204	1.7516
Mean	0.0220	0.0341	0.0214	0.0305	0.0201	0.0281	0.0179	0.0224
StdDev	0.0492	0.0710	0.0435	0.0550	0.0401	0.0475	0.0372	0.0378
Skew	-0.5991	2.6646	-0.8212	0.8411	-0.8161	-0.2470	-0.7878	-0.5201
Kurt	2.8950	26.4755	2.5283	10.9695	2.1622	4.0705	1.9090	1.9954
Max DD	0.6302	0.4979	0.4467	0.5601	0.3953	0.4662	0.3803	0.3027
Max 1M loss	0.2140	0.2264	0.1855	0.1789	0.1489	0.1838	0.1369	0.1446
CVaR (95%)	0.1044	0.1107	0.0938	0.0978	0.0881	0.0882	0.0803	0.0713
SR	1.5465	1.6607	1.7007	1.9230	1.7404	2.0446	1.6635	2.0491
$p$ -value( $SR_{DPPP} - SR_{PPP}$ )		0.3709		0.0985		0.0445		0.0042
$FF5 + Mom \alpha$	0.0104	0.0232	0.0103	0.0205	0.0097	0.0182	0.0085	0.0130
$StdErr(\alpha)$	0.0012	0.0029	0.0013	0.0024	0.0014	0.0020	0.0014	0.0016

*Notes.* This table presents out-of-sample performance estimates for deep portfolio policies using 157 firm characteristics, as specified in Equation (1). The analysis employs a feed-forward neural network model and data from the Open Source Asset Pricing Data set spanning January 1971 to December 2020. Results are shown for constant relative risk aversion (CRRA) investors with relative risk aversion coefficients ( $\gamma$ ) of 2, 5, 10, and 20. The first set of rows reports the certainty equivalent for each investor type, along with bootstrapped one-sided  $p$ -values comparing the certainty equivalents between the Deep Parametric Portfolio Policy (DPPP) and the Parametric Portfolio Policy (PPP). The second set of rows presents time-averaged portfolio weight statistics, including absolute weights, maximum and minimum weights, negative weight metrics (sum and proportion), and portfolio turnover. The third set of rows displays the return distribution characteristics: the first four moments, risk metrics (maximum drawdown, maximum monthly loss, and Conditional Value at Risk), annualized Sharpe ratios, and bootstrapped one-sided  $p$ -values comparing Sharpe ratios between the DPPP and the PPP. The bottom set of rows reports the alphas and their standard errors relative to the Fama-French five-factor model augmented with the momentum factor.

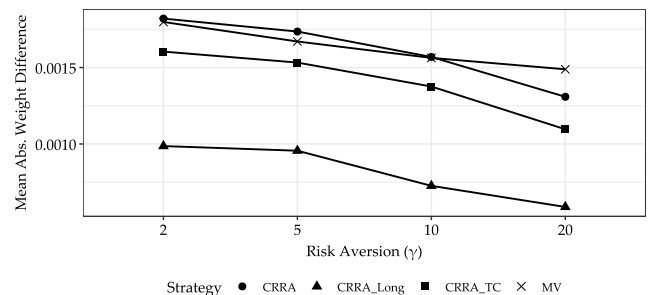
The economic significance of the deep learning approach is most evident in the certainty equivalent returns (CE).<sup>9</sup> At a risk aversion level of  $\gamma = 2$ , the DPPP achieves a CE of 2.97% compared with the PPP’s 1.95%, representing a significant enhancement of 102 basis points ( $p$ -value = 0.0003).<sup>10</sup> This substantial improvement suggests that capturing nonlinear relationships and interactions between predictors creates meaningful economic value for investors.

Notably, the performance differential between the DPPP and the PPP exhibits a decline with increasing risk aversion. The difference in monthly certainty equivalent narrows to 43-69 basis points at higher levels of risk aversion, with statistical significance declining correspondingly ( $p$ -values increase from 0.0003 at  $\gamma = 2$  to 0.0278 at  $\gamma = 20$ ). Intriguingly, we also find that the mean absolute weight differences between the PPP and DPPP models decrease with risk aversion, consistent with our analytical results and simulations in Section 2.2. Empirically, different from the simulation, constraints and modeling choices influence the composition of risky assets across risk aversion. Yet Figure 3 mirrors our simulation results in Figure 1 and shows a steady decline of weight differences for the models of Table 1 but also for other model specifications considered in the next sections,

suggesting that risk aversion serves as an effective economic regularization mechanism.

The portfolio characteristics in Table 1 provide further insights into the sources of outperformance. The DPPP exhibits more aggressive position-taking, as evidenced by the maximum portfolio weights (1.15% versus 0.68% at  $\gamma = 2$ ) and minimum weights (-1.28% versus -0.66% at  $\gamma = 2$ ), more concentrated portfolios (average

**Figure 3.** Mean Absolute Portfolio Weight Differences by Risk Aversion



*Notes.* The plot shows mean absolute portfolio weight differences between the PPP and DPPP models. Different lines refer to different model specifications, with CRRA the models in Table 1, CRRA\_TC the models in Table 3, CRRA\_Long the models in Table 3 and MV the models in Table 4.

absolute weights of 0.19% versus 0.17% at  $\gamma = 2$ ) and higher turnover than the linear portfolio policy (263% versus 134% at  $\gamma = 2$ ).<sup>11</sup> All these differences decline with risk aversion, reflecting the regularizing effect of risk preferences.

Risk metrics offer insights into downside protection, a crucial consideration for many investors that may not be fully accounted for by variance-based measures or standard utility functions. Despite its more aggressive positioning, the DPPP achieves risk control comparable to the PPP with lower maximum drawdowns (49.79% versus 63.02% at  $\gamma = 2$ ), similar maximum one-month losses (22.64% versus 21.40% at  $\gamma = 2$ ) or Conditional Value at Risk (CVaR) (11.07% versus 10.44% at  $\gamma = 2$ ).<sup>12</sup> Qualitatively similar results hold for all values of risk aversion, demonstrating that the DPPP approach can yield utility benefits without sacrificing practically relevant performance dimensions, even when those dimensions are not explicitly targeted in the optimization.

The improvements in risk-adjusted performance are substantial. Sharpe ratios are consistently higher for the DPPP across all risk aversion levels (1.66 versus 1.55 at  $\gamma = 2$ ), with the differences becoming statistically significant at higher risk aversion levels ( $p$ -value = 0.0042 at  $\gamma = 20$ ). The outperformance is robust to controlling for standard risk factors, as evidenced by significant monthly alphas against the Fama-French five-factor model augmented with momentum (2.32% versus 1.04% at  $\gamma = 2$ , with standard errors of 0.29% and 0.12%, respectively).

The main results from Table 1 are visually summarized in Figure 4, which shows the cumulative performance of portfolio returns over time for both the PPP and DPPP across different degrees of risk aversion. The figure demonstrates several key patterns. First, the DPPP consistently outperforms its linear counterpart across all risk aversion levels, with the outperformance becoming more pronounced after the 2008 financial crisis. Second, lower risk aversion portfolios ( $\gamma = 2$ , solid lines) achieve higher

cumulative returns but exhibit more volatility during periods of market stress. For instance, during the dot-com bubble burst (2000–2002), the global financial crisis (2008–2009), and the COVID-19 market crash (2020), the  $\gamma = 2$  portfolio experiences larger drawdowns compared with higher risk aversion portfolios.

Notably, the outperformance of the DPPP over the PPP persists across market environments, though the magnitude varies. The gap between the DPPP and the PPP tends to widen during strong market periods (e.g., 2003–2007 and 2009–2020) and narrows during market stress, suggesting that the benefits of nonlinear modeling are particularly valuable in capturing upside potential while still providing some downside protection.

### 3.4. Supporting Results

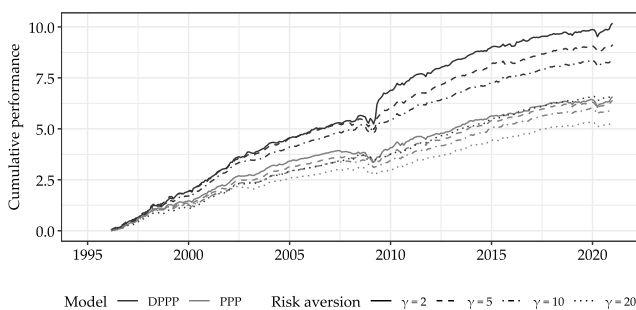
**3.4.1. Variable Importance.** We calculate the importance of the variables in the model as the mean absolute gradient of the model with respect to the input features. That is, for each period, we calculate the gradient of the investor's utility with respect to an input feature, take the absolute value of each value, and then take the average over all values. We repeat this for each feature in every out-of-sample period and take the average across all models. For the sake of comparability, we scale the average utility losses across all variables for each model so that they add up to one. As a result, we are able to rank the variables according to the average absolute gradient.

Figure 5 displays the relative importance of the 40 most influential characteristics across different risk aversion levels for both models, measured using absolute average gradients. The variables are ordered according to the importance of the DPPP model optimized for  $\gamma = 2$ .

Several key patterns emerge from this analysis. First, past return-based characteristics dominate the importance rankings across all specifications, with short-term reversal (STreversal), industry returns of big firms (IndRetBig), and momentum seasonality (MomSeason) consistently appearing among the most influential features, mirroring the findings in Moritz and Zimmermann (2016), Gu et al. (2020), and Chen et al. (2024) for prediction of returns rather than utility maximization. This finding holds for both the DPPP and the PPP, though the relative magnitudes differ substantially.

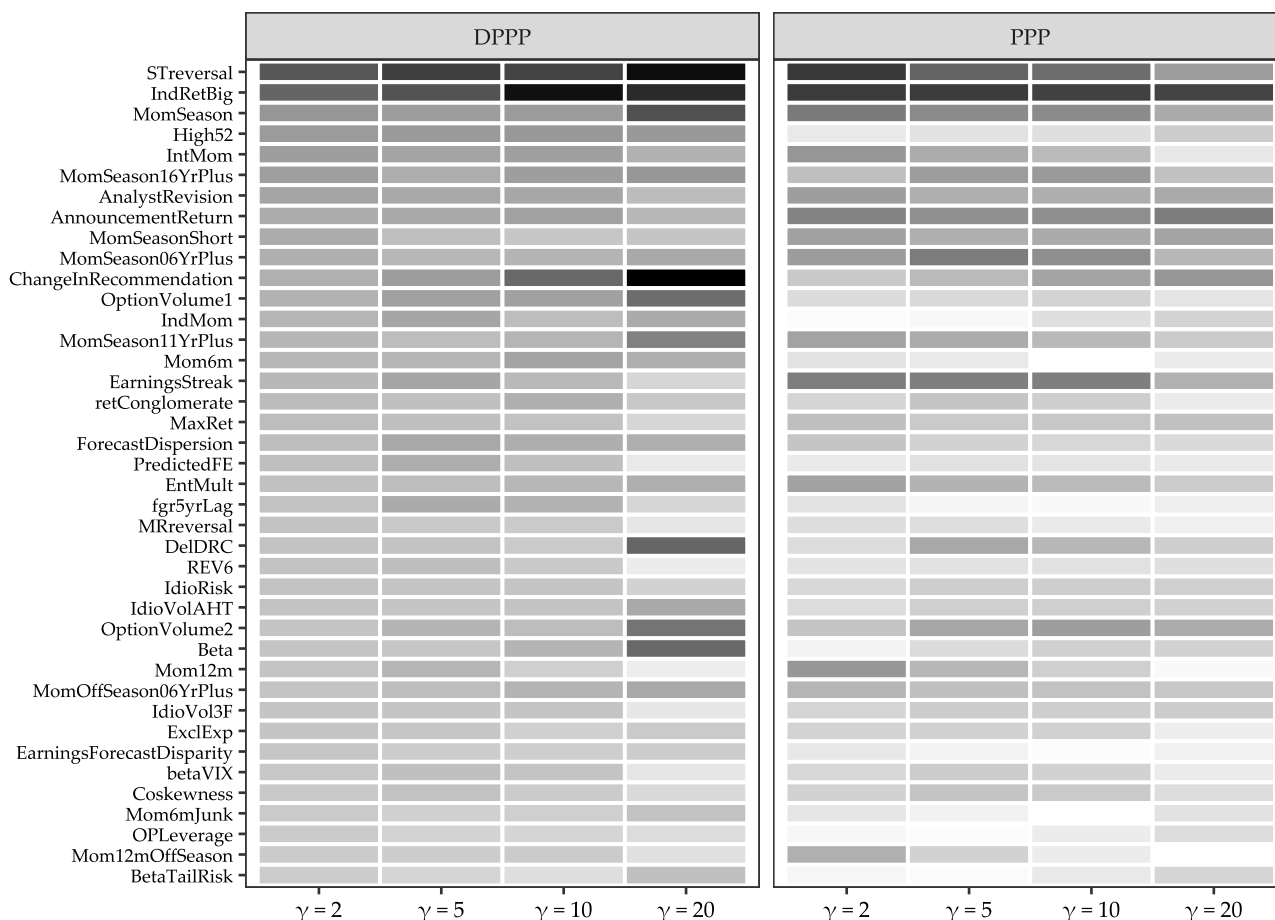
Second, the DPPP shows more pronounced differentiation in feature importance compared with the PPP, particularly at lower risk aversion levels. For instance, with  $\gamma = 2$ , short-term reversal exhibits approximately twice the importance in the DPPP compared with the PPP. This suggests that the nonlinear model is better at capturing and exploiting the dynamic nature of return reversal patterns. Importantly, the pattern of feature importance varies systematically with risk aversion. As risk aversion increases, our analysis reveals a more

**Figure 4.** Cumulative Performance over Time for CRRA Preferences



*Notes.* The plot shows the cumulative sum of portfolio returns for the DPPP and PPP. We show the results for each of the degrees of relative risk aversion considered and across all out-of-sample periods.

**Figure 5.** Variable Importance for the CRRA Preference for the DPPP and the PPP



*Notes.* Variable importance for the 40 most influential variables in the PPP and DPPP across model specifications and risk aversions, respectively. Variable importance is computed as the average absolute gradient over all training samples and normalized to sum to one within each model. The darker the color gradient, the higher the respective importance. The variables are ordered according to the importance of the DPPP model optimized for  $\gamma = 2$ .

balanced importance across characteristics, particularly in the DPPP, consistent with the results of DeMiguel et al. (2020). Therefore, model flexibility matters precisely when investors' preferences allow them to bear estimation risk in pursuit of complex return premia.

Third, the analysis also reveals interesting differences in how the two models utilize similar information. While both models draw heavily on momentum-related signals (MomSeason, IntMom, High52), the DPPP appears to extract more nuanced information, as evidenced by the higher importance weights on various momentum components (seasonal, intermediate, and price-based momentum). Notably, characteristics related to fundamental firm information (earnings, analyst forecasts, and balance sheet measures) show relatively stable importance across risk aversion levels, particularly in the DPPP. This suggests that these features provide complementary information that remains valuable even as the portfolio becomes more conservative.

In Supplementary Appendix S.3.1, we examine the marginal contribution (partial dependence) of

characteristics to portfolio weights in the DPPP and we find that nonlinear modeling is beneficial for capturing the complex relationship between firm characteristics and portfolio allocations. Key variables such as short-term reversal, book-to-market, and various momentum measures exert nonlinear and risk-aversion-dependent effects on portfolio weights. For instance, short-term reversal shows a strong, varying impact across its range, particularly under low risk aversion, which aligns well with the risk-return profiles observed in decile portfolios. In contrast, other characteristics exhibit more subdued or context-specific influences, with higher risk aversion generally dampening these effects.

To analyze further the extent to which nonlinearity plays a role, we fit linear surrogate models to explain the portfolio weights of the DPPP. The findings indicate that 30%–60% of the characteristic–weight relationship is linear, an additional 20%–30% is explained by interactions, and the remaining 10%–50% is due to higher-order nonlinearities. Moreover, the economic significance of these nonlinear components, measured via certainty

equivalent differences, is most pronounced for lower risk aversion levels, underscoring that the flexibility of nonlinear models adds substantial value. It also provides empirical support for our understanding that risk aversion acts as an economic regularization parameter. See Supplementary Appendix S.3.1 for detailed results.

**3.4.2. Comparison with Benchmark Models.** What do investors gain by optimizing their utility function directly rather than optimizing standard benchmarks such as, for example, the Sharpe ratio? Table 2 comprehensively evaluates DPPP portfolios against several important benchmarks, including machine learning-based return forecasting, Sharpe ratio optimization, a factor portfolio, and traditional passive strategies.

Results indicate that involved methods such as directly optimizing the Sharpe ratio, or a two-step strategy that uses machine-learning to forecast returns and subsequently builds portfolios based on those forecasts, have similar overall portfolio characteristics in terms of leverage, turnover or maximal one-month losses to the DPPP strategies, especially those DPPP strategies optimized for higher levels of risk aversion. DPPP strategies generally deliver higher out-of-sample alphas against involved benchmarks and also passive benchmarks such as equal- or value-weighted index portfolios. Consequently, DPPP strategies also deliver higher cumulative returns than the benchmark strategies (see Figure S.5.1 in the Supplementary Appendix).

Interestingly, the DPPP strategies deliver higher out-of-sample Sharpe ratios than even the strategies that target maximization of the Sharpe ratios directly. In Table S.6.1 in the Supplementary Appendix, we check whether the reverse (i.e., whether benchmark strategies deliver higher out-of-sample utility than strategies that target utility maximization directly) is also true, and find that this is not the case: strategies optimized for a specific level of risk aversion deliver higher out-of-sample utility than benchmark models or strategies optimized for other levels of risk aversion. This outperformance holds across different risk aversion levels. The relative advantage becomes increasingly pronounced as risk aversion increases. This is particularly evident when compared with strategies that do not explicitly incorporate portfolio risk in the optimization objective, underscoring the critical importance of integrating risk considerations directly into the optimization process.

Figure S.5.2 in the Supplementary Appendix shows variable importance rankings for the involved benchmark models. We find that, similar to the DPPP strategies, Sharpe ratio-maximizing strategies and machine learning-based return forecasting put high weight on past return-based characteristics such as short-term reversal or industry returns of big firms (IndRetBig). PPP-based maximization of the Sharpe ratio leads to a more even distribution across variable importance scores than DPPP-based maximization which mirrors our findings for variable importance for utility maximization above.

**Table 2.** Comparing Deep Portfolio Policies to Benchmark Strategies

	$\gamma = 2$	$\gamma = 5$	$\gamma = 10$	$\gamma = 20$	SR <sub>PPP</sub>	SR <sub>DPPP</sub>	ML	Factor	EW	VW
$\sum  w_i  / N_t * 100$	0.1907	0.1938	0.1933	0.1729	0.1868	0.1951	0.1802	0.1128	0.0694	0.0694
$\max w_i * 100$	1.1483	0.9843	0.8305	0.4582	0.8462	0.6434	0.4230	0.3987	0.0704	0.1113
$\min w_i * 100$	-1.2824	-1.2053	-0.9743	-0.7224	-0.7137	-0.7960	-0.2801	-0.3262	0.0704	0.0410
$\sum w_i I(w_i < 0)$	-0.8748	-0.8974	-0.8932	-0.7464	-0.8466	-0.9068	-0.7987	-0.3132	0.0000	0.0000
$\sum I(w_i < 0) / N_t$	0.3400	0.3368	0.3319	0.3202	0.3654	0.3566	0.2009	0.2558	0.0000	0.0000
$\sum  w_{i,t} - w_{i,t-1}^+ $	2.5891	2.5616	2.3473	1.7311	1.5163	1.8824	1.8640	0.4587	0.0829	0.0659
Mean	0.0341	0.0305	0.0281	0.0224	0.0202	0.0192	0.0193	0.0115	0.0110	0.0105
StdDev	0.0710	0.0550	0.0475	0.0378	0.0354	0.0354	0.0667	0.0416	0.0587	0.0552
Skew	2.6646	0.8411	-0.2470	-0.5201	-0.7604	-0.6342	-0.0544	-1.0598	-0.3716	-0.5039
Kurt	26.4755	10.9695	4.0705	1.9954	1.8370	1.8973	7.2504	3.2232	3.6591	3.3455
Max DD	0.4979	0.5601	0.4662	0.3027	0.2607	0.4356	1.0025	0.7836	0.9026	0.8585
Max 1M loss	0.2264	0.1789	0.1838	0.1446	0.1304	0.1336	0.3531	0.1924	0.2556	0.2398
CVaR (95%)	0.1107	0.0978	0.0882	0.0713	0.0732	0.0693	0.1466	0.1037	0.1351	0.1286
SR	1.6607	1.9230	2.0446	2.0491	1.9783	1.8815	1.0042	0.9590	0.6461	0.6609
FF5 + Mom $\alpha$	0.0232	0.0205	0.0182	0.0130	0.0110	0.0102	0.0048	0.0010	-0.0002	-0.0003
StdErr( $\alpha$ )	0.0029	0.0024	0.0020	0.0016	0.0015	0.0016	0.0019	0.0009	0.0007	0.0006

*Notes.* This table presents out-of-sample performance estimates for various portfolio strategies using data from the Open Source Asset Pricing Data set spanning January 1971 to December 2020. Strategies include CRRA-based Deep Parametric Portfolio Policies for different risk aversions, Sharpe ratio optimization approaches (SR) for linear and Deep Parametric Portfolio Policies, machine learning portfolios (ML), factor-based CRRA portfolio (Factor), and benchmark strategies (EW, VW). The DPPP strategy represents our baseline strategy for different risk aversions. The SR strategy is a PPP and DPPP optimized for Sharpe ratio preference. ML is the portfolio of a machine learning model trained to predict expected returns. The factor strategy is the linear PPP of a simple Fama-French five-factor model plus momentum. EW and VW are passive equal-weighted and value-weighted strategies. The first set of rows presents time-averaged portfolio weight statistics, including absolute weights, maximum and minimum weights, negative weight metrics (sum and proportion) and portfolio turnover. The second set of rows displays the return distribution characteristics: the first four moments, risk metrics (maximum drawdown, maximum monthly loss, and Conditional Value at Risk), and annualized Sharpe ratios. The bottom set of rows reports the alphas and their standard errors relative to the Fama-French five-factor model augmented with the momentum factor.

These results collectively reinforce three key findings. First, the value of deep learning in portfolio optimization extends beyond simple return prediction to the direct optimization of investor utility. Second, the benefits of preference-aligned optimization are robust across different utility specifications. Finally, sophisticated modeling approaches consistently outperform traditional passive strategies.

### 3.5. Robustness

We examine a number of alternative and extended model specifications. For the sake of brevity, the results are presented in the Supplementary Appendix S.3 and we only discuss the main take-aways here.

Our main results are based on an expanding-window framework that uses successively more data for model estimation (see Section 3.1). Rolling-window estimation that uses a fixed number of months for training might be able to adapt more readily to potential structural changes in the data by discarding older observations. Results in Supplementary Appendix S.3.2 show that rolling-window estimation does not consistently lead to better (or worse) results than expanding-window estimation. In fact, for high levels of risk aversion, the stability provided by a longer estimation sample can be crucial in achieving robust portfolio outcomes, aligning with our broader argument that model simplicity and regularization often yield more reliable results.

An important question concerns the interaction of cross-sectional characteristics and the state of the macroeconomy in the portfolio weight function. To study the impact of macroeconomic variables, we expand our baseline model with 8 macroeconomic variables from Welch and Goyal (2008) as in Gu et al. (2020), and we interact each macroeconomic variable with each cross-sectional characteristic for a total of 1,413 covariates. Results in Supplementary Appendix S.3.3 show that models including macroeconomic variables do not lead to higher investor utility than models that do not include macroeconomic variables, for all levels of risk aversion.

### 3.6. Market Frictions

In the benchmark setting, average turnover and leverage are economically high for both the PPP and the DPPP. Next, we compare both approaches in a more economically feasible scenario that explicitly accounts for market frictions by imposing a transaction cost penalty and using a long-only constraint in the optimization task. Note that both frictions act as regularization mechanisms (Jagannathan and Ma 2003, Hautsch and Voigt 2019) on top of regularization via risk aversion. We therefore expect nonlinear and linear models to be closer for all levels of risk aversion in these scenarios, making it harder to isolate the risk aversion channel.

To account for transaction costs, we follow DeMiguel et al. (2020) and add the following penalty term to the

optimization problem:

$$TC = \frac{1}{T} \sum_{t=1}^T \left[ \sum_{i=1}^{N_t} |\kappa_{i,t}(w_{i,t} - w_{i,t-1}^+)| \right], \quad (14)$$

where  $\kappa_{i,t}$  are transaction costs for stock  $i$  at time  $t$  and  $w_{i,t-1}^+$  is the portfolio weight before rebalancing, that is,

$$w_{i,t-1}^+ = \frac{w_{i,t-1}(1 + r_{i,t})}{1 + \sum_{j=1}^{N_t} w_{j,t-1}r_{j,t}}. \quad (15)$$

Our transaction cost estimates come from Chen and Velikov (2023).<sup>13</sup> Thus, we define transaction costs  $\kappa_{i,t}$  as the effective half bid-ask spread.

An important consideration when incorporating transaction costs into portfolio optimization is the inherently dynamic nature of the problem. In a multiperiod setting, optimal portfolio weights at time  $t$  depend not only on current characteristics but also on expected future optimal positions and the associated trading costs. Jensen et al. (2022) formalize this intuition by deriving a closed-form solution in the mean-variance case that explicitly accounts for these dynamic effects. Our approach, following DeMiguel et al. (2020), instead incorporates transaction costs through a penalty term in the objective function (14). While this simplifies the dynamic aspect of the problem, it maintains tractability when dealing with a large cross-section of assets and characteristics while still capturing the first-order effects of trading frictions on portfolio choice. As our empirical results demonstrate, this formulation effectively constrains turnover and generates economically reasonable portfolios.

A large majority of equity portfolios face restrictions on short selling. We incorporate short-sale constraints as in Brandt et al. (2009), that is, we restrict portfolio weights to be nonnegative in the optimization problem of Equation (1) (and still keep the cap of 3% per stock). In particular, to make sure that portfolio weights still sum up to one, we add the following portfolio rebalancing term to our optimization process:

$$w_{i,t}^* = \frac{\max[0, w_{i,t}]}{\sum_{j=1}^{N_t} \max[0, w_{j,t}]}. \quad (16)$$

Table 3 shows separately the results of the optimization process with the transaction cost penalty and the long-only constraint for CRRA investors with different degrees of risk aversion. We show a selected set of results compared with Table 1, but provide similar tables with all results in Table S.6.3 for transaction cost and Table S.6.4 for long-only in the Supplementary Appendix.

The first panel of Table 3 shows that even with transaction costs, the DPPP outperforms the PPP across all risk aversion levels, with monthly certainty equivalent differences ranging from 10 to 39 basis points (certainty

**Table 3.** Long-Only & Transaction Costs Constrained Deep Portfolio Policy for CRRA Investors with Different Degrees of Risk Aversion

	$\gamma = 2$		$\gamma = 5$		$\gamma = 10$		$\gamma = 20$	
	PPP	DPPP	PPP	DPPP	PPP	DPPP	PPP	DPPP
<b>Transaction costs</b>								
CE	0.0155	0.0194	0.0129	0.0157	0.0077	0.0087	-0.0029	-0.0006
$p$ -value( $CE_{DPPP} - CE_{PPP}$ )		0.0118		0.0620		0.3195		0.1555
$\sum w_i I(w_i < 0)$	-0.7139	-0.8877	-0.7612	-0.8973	-0.7638	-0.8798	-0.6588	-0.6756
$\sum  w_{i,t} - w_{i,t-1}^+ $	0.8441	2.0257	0.8794	1.9002	0.8593	1.5947	0.7754	1.1407
Mean	0.0179	0.0225	0.0178	0.0221	0.0170	0.0182	0.0144	0.0157
StdDev	0.0482	0.0551	0.0427	0.0498	0.0397	0.0412	0.0360	0.0349
Max 1M loss	0.2228	0.2280	0.1812	0.2015	0.1559	0.1546	0.1303	0.1513
SR	1.2851	1.4123	1.4453	1.5370	1.4823	1.5296	1.3805	1.5552
$p$ -value( $SR_{DPPP} - SR_{PPP}$ )		0.2090		0.2962		0.3852		0.0447
<b>Long-only</b>								
CE	0.0118	0.0164	0.0076	0.0107	0.0011	0.0020	-0.0157	-0.0104
$p$ -value( $CE_{DPPP} - CE_{PPP}$ )		0.0001		0.0143		0.3308		0.0114
$\sum w_i I(w_i < 0)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\sum  w_{i,t} - w_{i,t-1}^+ $	0.5883	1.3508	0.6426	1.3417	0.5000	1.0914	0.3274	0.7656
Mean	0.0150	0.0215	0.0147	0.0216	0.0137	0.0174	0.0114	0.0135
StdDev	0.0566	0.0713	0.0510	0.0647	0.0459	0.0490	0.0406	0.0390
Max 1M loss	0.2483	0.2603	0.2171	0.2667	0.1968	0.2260	0.1832	0.1780
SR	0.9213	1.0418	0.9996	1.1580	1.0342	1.2262	0.9717	1.1974
$p$ -value( $SR_{DPPP} - SR_{PPP}$ )		0.0119		0.0077		0.0006		0.0001

Notes. This table presents out-of-sample performance estimates for deep portfolio policies with the transaction costs penalty from Equation (14) and including a long-only constraint using 157 firm characteristics separately, as specified in Equation (1). The analysis employs a feed-forward neural network model and data from the Open Source Asset Pricing Data set spanning January 1971 to December 2020. Results are shown for constant relative risk aversion (CRRA) investors with relative risk aversion coefficients ( $\gamma$ ) of 2, 5, 10, and 20. Results in the first panel are reported net of transaction costs. For each panel the first set of rows reports the certainty equivalent for each investor type, along with bootstrapped one-sided  $p$ -values comparing the certainty equivalents between the Deep Parametric Portfolio Policy (DPPP) and the Parametric Portfolio Policy (PPP). The second set of rows presents time-averaged portfolio weight statistics, including leverage and portfolio turnover. The third set of rows displays the return distribution characteristics: the first two moments, maximum monthly loss, annualized Sharpe ratios, and bootstrapped one-sided  $p$ -values comparing Sharpe ratios between the DPPP and the PPP.

equivalents are reported net of transaction costs). This suggests that, like risk aversion, the transaction cost penalty acts as an economic regularizer that reduces model complexity. Consequently, both models exhibit lower certainty equivalents and smaller, less significant differences, as supported by the reduced mean absolute weight differences in Figure 3. This is in line with the results of Hautsch and Voigt (2019), who show that a transaction cost penalty is analogous to a ridge penalty and thus acts as a natural economic regularization. As risk aversion increases, the significance of these differences declines with  $\gamma = 10$  showing no significant difference at the 5% level while the constraints reduce turnover to 78%–88% for the PPP and 114%–203% for the DPPP. Despite higher turnover, the DPPP delivers notably larger net returns and higher Sharpe ratios.

The second panel of Table 3 presents long-only portfolio optimization results for CRRA investors. Here, the DPPP again outperforms the PPP (monthly certainty equivalent differences from 9 to 53 basis points), although the benefits of model complexity diminish more rapidly as risk aversion increases. The long-only

constraint, like risk aversion, acts as an economic regularizer that reduces complexity, as evidenced by lower certainty equivalents and minimal weight differences in Figure 3. This is consistent with the results of Jagannathan and Ma (2003), who show that short-selling restrictions can also be interpreted as a form of regularization that implicitly shrinks the set of possible weights and prevents extreme allocations. Therefore, it also leads to more concentrated positions, with DPPP turnover ranging from 77%–135% versus 33%–64% for the PPP. Additionally, the DPPP achieves significantly higher Sharpe ratios at the 5% level across all risk aversion levels.

Similar patterns are observed when both constraints are applied jointly (see Table S.6.5 in the Supplementary Appendix).

## 4. Alternative Investor Utility Functions

### 4.1. Mean-Variance and Loss Aversion

We explore results for different investor types by changing the utility function that we use to optimize the models. In particular, we consider linear and deep portfolio

policies for an investor with mean-variance utility defined as

$$u(r_{p,t+1}) = r_{p,t+1} - \frac{\gamma}{2} \left( r_{p,t+1} - \frac{1}{T} \sum_{t=1}^T r_{p,t+1} \right)^2, \quad (17)$$

where  $\gamma$  is the absolute risk aversion of the investor, and for a loss-averse investor (Tversky and Kahneman 1992) with utility defined as

$$u(r_{p,t+1}) = \begin{cases} -l(\bar{W} - (1 + r_{p,t+1}))^b & \text{if } (1 + r_{p,t+1}) < \bar{W} \\ ((1 + r_{p,t+1}) - \bar{W})^b & \text{otherwise} \end{cases}, \quad (18)$$

where  $\bar{W}$  is a reference wealth level determined in the editing stage, the parameter  $l$  measures the investor's loss aversion and the parameter  $b$  captures the degree of risk seeking over losses and risk aversion over gains. For simplicity, we fix the parameters  $\bar{W}$  and  $b$  at one and only change the loss aversion parameter  $l$ . We include the constraints specified in Section 3.1 in the optimization process for both preferences.

Table 4 shows separately the results of the optimization process for the mean-variance investors with different degrees of risk aversion and loss-averse investors

with different degrees of loss aversion. We show a selected set of results compared with Table 1, but provide similar tables with all results in Table S.6.6 for mean-variance preference and Table S.6.7 for loss aversion preference in the Supplementary Appendix.

The first panel of Table 4 shows that for a mean-variance investor, the deep portfolio policy yields higher certainty equivalent returns than the linear policy across all risk aversion levels. While the DPPP's results (certainty equivalents, Sharpe ratios, and weight characteristics) are similar to those for a CRRA investor, the linear model performs relatively better in the mean-variance setting, reducing the monthly certainty equivalent difference to 23–86 basis points. The mean-variance utility function perfectly illustrates that the degree of absolute risk aversion determines the strength of the penalty on the variance of portfolio returns, that is, the strength of regularization, since portfolio return variance is an explicit part of the utility function (see Section 2.2). Figure 3 illustrates the convergence of mean absolute weight differences between the two models with increasing risk aversion.

The second panel of Table 4 reports results for a loss-averse investor. Here, the DPPP outperforms the PPP at all levels of loss aversion, with improvements ranging

**Table 4.** Deep Portfolio Policy for Mean-Variance and Loss-Averse Investors with Different Degrees of Risk Aversion ( $\gamma$ ) and Loss Aversion ( $l$ )

	$\gamma = 2$		$\gamma = 5$		$\gamma = 10$		$\gamma = 20$	
	PPP	DPPP	PPP	DPPP	PPP	DPPP	PPP	DPPP
Mean-variance preference								
CE	0.0201	0.0287	0.0184	0.0217	0.0143	0.0170	0.0065	0.0088
$p$ -value( $CE_{DPPP} - CE_{PPP}$ )		0.0001		0.0292		0.0291		0.0849
$\sum w_i I(w_i < 0)$	-0.7602	-0.8882	-0.8059	-0.9070	-0.8093	-0.8899	-0.7879	-0.8925
$\sum  w_{i,t} - w_{i,t-1}^+ $	1.5185	2.6428	1.7406	2.5648	1.6789	2.4174	1.4693	2.2676
Mean	0.0225	0.0319	0.0232	0.0281	0.0224	0.0276	0.0205	0.0254
StdDev	0.0492	0.0566	0.0435	0.0505	0.0402	0.0459	0.0373	0.0407
Skew	-0.6239	-0.1348	-0.8530	-0.6631	-0.8516	-0.4331	-0.7727	-0.5940
SR	1.5843	1.9506	1.8438	1.9259	1.9317	2.0786	1.9007	2.1596
$p$ -value( $SR_{DPPP} - SR_{PPP}$ )		0.0019		0.2768		0.1185		0.0171
Loss aversion preference								
	$l = 1.5$		$l = 2$		$l = 3$		$l = 4$	
CE	0.0188	0.0311	0.0147	0.0235	0.0082	0.0137	0.0025	0.0036
$p$ -value( $CE_{DPPP} - CE_{PPP}$ )		0.0002		0.0015		0.0247		0.3014
$\sum w_i I(w_i < 0)$	-0.7929	-0.8918	-0.7980	-0.8833	-0.8090	-0.8823	-0.8083	-0.8702
$\sum  w_{i,t} - w_{i,t-1}^+ $	1.6336	2.6846	1.5951	2.6742	1.6887	2.5599	1.7273	2.4745
Mean	0.0235	0.0361	0.0227	0.0319	0.0226	0.0306	0.0227	0.0275
StdDev	0.0494	0.0751	0.0442	0.0580	0.0412	0.0548	0.0395	0.0485
Skew	-0.6194	1.9765	-0.7339	0.5481	-0.7651	0.3536	-0.7475	-0.2204
SR	1.6475	1.6666	1.7793	1.9049	1.9052	1.9331	1.9905	1.9677
$p$ -value( $SR_{DPPP} - SR_{PPP}$ )		0.4931		0.2424		0.4498		0.4400

*Notes.* This table presents out-of-sample performance estimates for deep portfolio policies using 157 firm characteristics, as specified in Equation (1). The analysis employs a feed-forward neural network model and data from the Open Source Asset Pricing Data set spanning January 1971 to December 2020. Results are shown for mean-variance investors with relative risk aversion coefficients ( $\gamma$ ) of 2, 5, 10, and 20 in the first panel and loss-averse investors with loss aversion ( $l$ ) of 1.5, 2, 3, and 4 in the second panel. The first set of rows reports the certainty equivalent for each investor type, along with bootstrapped one-sided  $p$ -values comparing the certainty equivalents between the Deep Parametric Portfolio Policy (DPPP) and the Parametric Portfolio Policy (PPP). The second set of rows presents time-averaged portfolio weight statistics, including leverage and portfolio turnover. The third set of rows displays the return distribution characteristics: the first three moments, annualized Sharpe ratios, and bootstrapped one-sided  $p$ -values comparing Sharpe ratios between the DPPP and the PPP.

from 11 to 123 basis points—differences significant at the 1% level for  $l = 1.5$  and  $l = 2$ , at 5% for  $l = 3$ , and insignificant for  $l = 4$ . Because a loss-averse investor values the tail behavior of returns more than the mean–variance trade-off, both models show higher skewness compared with mean–variance or CRRA optimizations. Notably, the DPPP produces significantly higher (right) skewness, which explains its higher certainty equivalent. In line with our theoretical results in Supplementary Appendix S.2, increasing loss aversion  $l$  does indeed penalize negative outcomes more severely.

Leverage of the PPP and DPPP strategies using different utility functions has about the same magnitude as in the CRRA benchmark results in Table 1. In Table S.6.8 in the Supplementary Appendix, we show that certainty-equivalent returns are lower for long-only than for leveraged strategies for mean-variance or loss aversion utility, mirroring our results for CRRA utility (Table 3). The PPP–DPPP return gap is of comparable magnitude to that with leverage and remains statistically significant, except under very high loss aversion, consistent with the theoretical framework and simulations. Figure S.5.3 in the Supplementary Appendix further illustrates that the DPPP consistently outperforms the PPP over time, with the degree of outperformance varying with the investor’s risk or loss aversion.

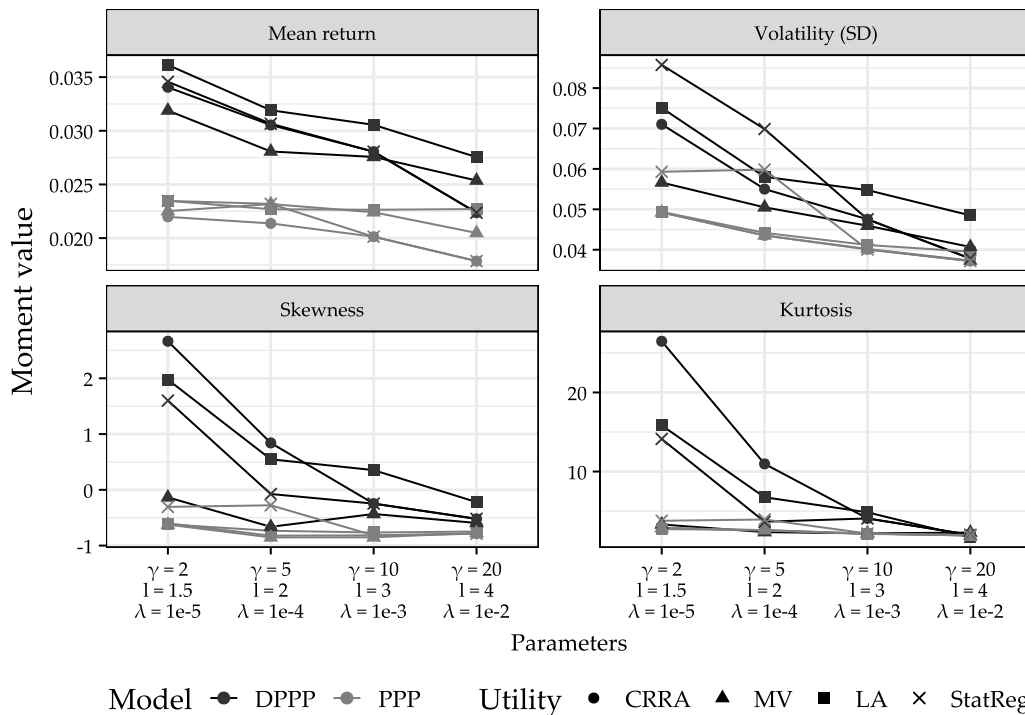
Finally, it is instructive to compare portfolio return moments across utility functions and portfolio policies in Table 4. Under mean-variance utility, increasing risk aversion is associated with lower portfolio return variance for both the linear (PPP) and complex (DPPP) models. In contrast, with loss aversion utility, higher loss aversion leads to lower skewness in absolute terms. This distinction highlights the relevance of capturing optimal portfolio return distributions when the investor’s objective function is sensitive to higher moments. We study this in more detail in the next section.

### 4.2. Portfolio Moments Across Utility Functions and Models

To highlight the economic implications of different utility functions, Figure 6 reports the first four moments of optimal portfolios across investor preferences, model classes, and regularization schemes. The figure compares CRRA, mean-variance, and loss-averse preferences, as well as a statistical objective function (StatReg) that maximizes mean return subject to an  $\ell_1$  penalty on model parameters with strength  $\lambda$ . The StatReg baseline provides a useful benchmark for understanding how purely statistical regularization compares to preference-induced regularization.

Three patterns are evident. First, increasing risk aversion or loss aversion systematically reduces volatility

Figure 6. Portfolio Moments Across Utility Functions



Notes. This figure shows portfolio moments (mean return, volatility, skewness, and kurtosis) across utility functions, preference parameters, and model classes. The utility functions are CRRA (dots), mean–variance (triangles), loss aversion (squares) and StatReg (x). Dark lines denote DPPP; lighter lines denote PPP. StatReg optimizes mean return with an  $\ell_1$  penalty (lasso).

and higher-order moment exposures. Both CRRA and loss-averse investors display a monotonic decline in variance and kurtosis as parameters  $\gamma$  and  $\lambda$  increase, confirming the theoretical mechanism that preferences penalize estimation variance in a manner analogous to statistical shrinkage. This comparison is reinforced by the StatReg benchmark: purely statistical lasso regularization produces similar dampening of higher-order moments, but the utility-based approach targets precisely those risks that the investor's preferences place weight on.

Second, the different utility functions emphasize distinct moment trade-offs. CRRA portfolios initially exhibit very high mean return and positive skewness, but these come with extreme kurtosis that vanishes as risk aversion rises. Loss-averse portfolios are particularly distinctive: as loss aversion strengthens, these portfolios generate the highest skewness of any utility specification, but this comes jointly with high variance. In other words, loss aversion leads to portfolios that sacrifice stability in exchange for both higher upside and protection against extreme downside states. In contrast, the StatReg portfolio with the highest regularization yields the lowest mean return and relatively muted higher moments. This reflects the effect of the lasso penalty, which shrinks exposures but does not respond to investor preferences.

Third, a comparison of different model classes shows that the DPPP generally delivers higher means and stronger moment tilts at low levels of risk tolerance, and it converges more strongly. In contrast, the PPP exhibits a less exposed profile across preferences. As risk aversion or loss aversion increases, the gap between the DPPP and the PPP closes, reflecting the theoretical result that preferences shrink the range of possible outcomes and dampen the role of nonlinear interactions.

In Supplementary Appendix S.4, we analyze the net exposure of different portfolios to different stock characteristics. While characteristics exposure is similar across utility functions and risk/loss aversion levels for most clusters of characteristics, some intuitive differences emerge: As risk or loss aversion increases, investors decrease net exposure to short-term reversal and firm size. Since predictors are signed in the Chen and Zimmermann (2022) database to have positive mean return in the original in-sample periods (and smaller firms are associated with higher returns), this implies for size that more risk or loss averse investors load *more* on larger firms, presumably because such firms are less volatile. For short-term reversal, the decline in exposure is least pronounced for loss-averse portfolios, suggesting that loss-averse investors, unlike CRRA and mean-variance investors, are more willing to invest in stocks that have recently underperformed even for higher degrees of loss aversion. See Supplementary Appendix S.4 for a more comprehensive discussion.

## 5. Conclusion

Building on the seminal work of Brandt et al. (2009) and the extensive literature on portfolio optimization and machine learning, we develop a novel Deep Parametric Portfolio Policy (DPPP) that integrates the structural advantages of traditional parametric portfolio policies with the flexibility of deep neural networks. Our approach maps a large set of firm characteristics to optimal portfolio weights in a nonlinear, interactive manner while directly incorporating market frictions and investor-specific utility functions such as CRRA, mean-variance, or loss aversion into the optimization process.

A key contribution of our work is the introduction of the concept of *economic regularization*. We provide a theoretical framework demonstrating how an investor's risk aversion naturally limits effective model complexity. As risk aversion increases, the incentive to exploit nonlinear relationships is tempered by the heightened penalty on return variance. Our simulations and analytical derivations show that, under higher risk aversion, the benefits of additional complexity diminish, leading the flexible DPPP to converge toward its linear counterpart. The results provide diagnostic guidance: if an investor is conservative or faces strong frictions, then a linear PPP is close to optimal; if the investor is aggressive and frictions are light, then the DPPP offers meaningful gains. In other words, one can match model complexity to investor preferences.

Our empirical investigation reinforces the theoretical insights. We document substantial improvements in investor utility when adopting the DPPP relative to standard linear models, across a variety of settings with different constraints or objective functions. Across all settings, certainty-equivalent gains over linear models decrease systematically with increasing risk aversion. Further, our analysis of variable importance reveals that while past return-based signals dominate for low risk aversions, a more balanced mix of return-based and accounting-based characteristics emerges as risk aversion increases.

Overall, our approach puts forward a comparably simple and flexible neural network-based model that enables practitioners and researchers alike to create reasonable portfolio allocations based on firm characteristics and preferences, highlighting the growing role of machine learning and nonlinear models in finance. Moreover, the built-in economic regularization mechanism selects the optimal level of model complexity based on investor preferences directly. As such, it provides an intuitive alternative to purely statistical model regularization.

## Acknowledgments

The authors thank Victor DeMiguel, Christian Fieberg, Bryan Kelly, Alexander Klos, Simon Rottke, Mark Salmon,

Fabricius Somogyi (discussant), Bastidon Cécile (discussant), Heiner Beckmeyer (discussant) and seminar participants at the Amsterdam Business School, the Research in Behavioral Finance Conference (RBF), the Cardiff Fintech Conference, the 2022 New Zealand Finance Meeting (NZFM), the Paris Financial Management Conference (PFMC), the Theory-based Empirical Asset Pricing Research (TBEAR) Network Workshop 2023, the University of Liechtenstein, University Pompeu Fabra, the CEQURA Conference 2023 on Advances in Financial and Insurance Risk Management, the BVI-CFR Event 2023, the 4th Frontiers of Factor Investing 2024 Conference, the Oxford-Man Institute of Quantitative Finance, and the International Monetary Fund (IMF) Brownbag Research Seminar for helpful comments and suggestions.

## Appendix A. Proofs

**A.1. Proof of Proposition 1** For this proof consider a second-order Taylor expansion of the CRRA utility function around the expected portfolio return:

$$\mathbb{E}[u(r_{p,t+1}(\theta))] \approx \mathbb{E}[r_{p,t+1}(\theta)] - \frac{\gamma}{2} \mathbb{E}[r_{p,t+1}(\theta)^2]. \quad (\text{A.1})$$

This approximation is well-known and shows that the CRRA utility framework naturally reduces to mean-variance preferences, where  $\gamma$  represents risk aversion and directly scales the penalty on portfolio return variance. We can express returns as specified in Equation (5). For mean-variance utility, this yields the optimization problem:

$$\max_{\theta} \theta^T \hat{\mu}_c - \frac{\gamma}{2} \theta^T \hat{\Sigma}_c \theta - \gamma \theta^T \hat{\sigma}_{bc}, \quad (\text{A.2})$$

with first-order condition:

$$\theta^* = \frac{1}{\gamma} \hat{\Sigma}_c^{-1} \hat{\mu}_c - \hat{\Sigma}_c^{-1} \hat{\sigma}_{bc}. \quad (\text{A.3})$$

Plugging in the optimal coefficients  $\theta^*$  into Equation (4) yields Proposition 1.

**A.2. Proof of Proposition 3** For this proof consider the definition of the EDF in Equation (9). We define the two key matrices as:

$$G = \mathbb{E} \left[ \frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'} \right], \quad (\text{A.4})$$

and

$$V = \mathbb{E} \left[ \frac{\partial L(\theta)}{\partial \theta} \frac{\partial L(\theta)}{\partial \theta'} \right]. \quad (\text{A.5})$$

Under mean-variance utility from Equation (A.2):

$$G = \frac{1}{T} \gamma \hat{\Sigma}_c, \quad (\text{A.6})$$

$$\frac{\partial L(\theta)}{\partial \theta} = r_{c,t+1} - \gamma \hat{\Sigma}_c \left( \frac{1}{\gamma} \hat{\Sigma}_c^{-1} \hat{\mu}_c - \hat{\Sigma}_c^{-1} \hat{\sigma}_{bc} \right) - \gamma \hat{\sigma}_{bc} = r_{c,t+1} - \hat{\mu}_c, \quad (\text{A.7})$$

and

$$V = \frac{1}{T} (r_{c,t+1} - \hat{\mu}_c)^T (r_{c,t+1} - \hat{\mu}_c) = \hat{\Sigma}_c. \quad (\text{A.8})$$

Therefore, our EDF measure simplifies to

$$\text{EDF} = \text{tr}(G^{-1}V)/T = \frac{1}{\gamma} \text{tr}(\hat{\Sigma}_c^{-1} \hat{\Sigma}_c) = \frac{p}{\gamma}, \quad (\text{A.9})$$

where  $p$  denotes the number of characteristics. This leads to our key result about model complexity in Proposition 3.

## Endnotes

<sup>1</sup> As noted by Brandt et al. (2009), the presence of a risk-free asset is not required for our setting, because it only adds a scalar leverage decision that is orthogonal to the cross-sectional allocation problem.

<sup>2</sup> The  $1/N_t$  term is a normalization that allows the portfolio weight function to be applied to a time-varying number of stocks. Without this normalization, an increase in the number of stocks with an otherwise unchanged cross-sectional distribution of characteristics leads to more radical allocations, although the investment opportunities are basically unchanged.

<sup>3</sup> While return prediction using random Fourier features has become popular following Kelly et al. (2024), Nagel (2025) shows that predictions from such (ridgeless) regressions become a weighted average of past returns in the training window, with weights being functions of simple Gaussian kernels. Our results are invariant to the transformation function and therefore also hold for using simple polynomial features.

<sup>4</sup> The activation function introduces nonlinearity into the model by applying a transformation that isn't simply a straight line. We use the leaky rectified linear unit (ReLU) as activation function throughout all layers to prevent the issue of "dying ReLU" (see Supplementary Appendix S.1). The leaky ReLU is a piecewise linear function: it behaves like a regular ReLU for positive inputs but applies a small, non-zero slope to negative inputs instead of completely zeroing them out. Because of this change in slope, the overall function is not purely linear, which lets the network capture more complex, non-linear relationships in the data.

<sup>5</sup> Ang et al. (2011) show that the average gross leverage of hedge fund companies amounts to 120% in the period after the financial crisis 2007–2008. We use a slightly more conservative number of a maximum leverage of 100%.

<sup>6</sup> We also experimented with a single split of the data into an estimation and a test period but results are significantly worse. This suggests that the relationship between stock weights and characteristics varies over time. Hence, more frequent coefficient updates (either via expanding- or rolling-window strategies) are crucial to achieve promising results.

<sup>7</sup> Chen and McCoy (2024) show that simple median imputation of missing values outperforms more sophisticated methods in the context of machine learning portfolio formation. In fact, they explicitly recommend applying simple median imputation in this context. They argue that there is little to be gained from other methods (and that such methods might even introduce estimation noise) that try to exploit the cross-sectional or time-series structure because a. missingness occurs in blocks and b. non-missing predictors display low cross-sectional correlations.

<sup>8</sup> To ensure comparability between the linear and the Deep Parametric Portfolio Policy we differ slightly from Brandt et al. (2009) in that the linear model includes  $l_1$ -regularization and early stopping, similar to the deep model. A more detailed description is given in Supplementary Appendix S.1.

<sup>9</sup> The certainty equivalent return is the guaranteed monthly return an investor would require to achieve the same expected utility as via following the corresponding estimated portfolio policy.

<sup>10</sup> We follow DeMiguel et al. (2024) and construct one-sided  $p$ -values from 10,000 bootstrap samples using the stationary bootstrap

method of Politis and Romano (1994) with an average block size of five and the procedure of Ledoit and Wolf (2008). This method is also used when assessing the statistical significance of Sharpe ratio differences between the deep and the linear parametric portfolio policy hereafter.

<sup>11</sup> See Section 3.6 for a formal definition of turnover.

<sup>12</sup> Following Gu et al. (2020), we calculate maximum drawdown based on cumulative log returns.

<sup>13</sup> We thank the authors for making an updated version of the data available.

## References

- Ammann M, Coqueret G, Schade J-P (2016) Characteristics-based portfolio choice with leverage constraints. *J. Banking Finance* 70:23–37.
- Ang A, Gorovyy S, van Inwegen GB (2011) Hedge fund leverage. *J. Financial Econom.* 102(1):102–126.
- Bianchi D, Büchner M, Tamoni A (2020) Bond risk premiums with machine learning. *Rev. Financial Stud.* 34(2):1046–1089.
- Brandt MW (2010) Portfolio choice problems. Ait-Sahalia Y, Hansen LP, eds. *Handbook of Financial Econometrics: Tools and Techniques*, Handbooks in Finance, vol. 1 (North-Holland, San Diego), 269–336.
- Brandt MW, Santa-Clara P, Valkanov R (2009) Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Rev. Financial Stud.* 22(9):3411–3447.
- Bryzgalova S, Pelger M, Zhu J (2025) Forest through the trees: Building cross-sections of stock returns. *J. Finance* 80:2447–2506.
- Chen AY, McCoy J (2024) Missing values handling for machine learning portfolios. *J. Financial Econom.* 155:103815.
- Chen AY, Velikov M (2023) Zeroing in on the expected returns of anomalies. *J. Financial Quant. Anal.* 58(3):968–1004.
- Chen AY, Zimmermann T (2022) Open source cross-sectional asset pricing. *Critical Finance Rev.* 27(2):207–264.
- Chen L, Pelger M, Zhu J (2024) Deep learning in asset pricing. *Management Sci.* 70(2):714–750.
- Chevalier G, Coqueret G, Raffinot T (2022) Supervised portfolios. *Quant.* 22(12):2275–2295.
- Cong LW, Tang K, Wang J (2020) AlphaPortfolio: Goal-oriented investment management through deep reinforcement learning. Preprint, submitted April 20, <http://dx.doi.org/10.2139/ssrn.3554486>.
- Coulombe PG, Goebel M (2023) Maximally machine-learnable portfolios. Preprint, submitted June 8, <https://arxiv.org/abs/2306.05568>.
- DeMiguel V, Garlappi L, Uppal R (2009) Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Rev. Financial Stud.* 22(5):1915–1953.
- DeMiguel V, Martín-Utrera A, Uppal R (2024) A multifactor perspective on volatility-managed portfolios. *J. Finance* 79(6):3859–3891.
- DeMiguel V, Martín-Utrera A, Nogales FJ, Uppal R (2020) A transaction-cost perspective on the multitude of firm characteristics. *Rev. Financial Stud.* 33(5):2180–2222.
- Detzel A, Novy-Marx R, Velikov M (2023) Model comparison with transaction costs. *J. Finance* 78(3):1743–1775.
- Didisheim A, Ke SB, Kelly BT, Malamud S (2023) Complexity in factor pricing models. NBER Working Paper No. 31689, National Bureau of Economic Research, Cambridge, MA.
- Feng G, Jiang L, Li J, Song Y, Wang Y (2022) Deep tangency portfolio. Preprint, submitted March 11, <http://dx.doi.org/10.2139/ssrn.3971274>.
- Freyberger J, Neuhierl A, Weber M (2020) Dissecting characteristics nonparametrically. *Rev. Financial Stud.* 33(5):2326–2377.
- Gu S, Kelly B, Xiu D (2020) Empirical asset pricing via machine learning. *Rev. Financial Stud.* 33(5):2223–2273.
- Guijarro-Ordóñez J, Pelger M, Zanotti G (2025) Deep learning statistical arbitrage. *Management Sci.*, ePub ahead of print December 4, <https://doi.org/10.1287/mnsc.2022.03132>.
- Hautsch N, Voigt S (2019) Large-scale portfolio allocation under transaction costs and model uncertainty. *J. Econom.* 212(1): 221–240.
- Heaton JB, Polson NG, Witte JH (2017) Deep learning for finance: Deep portfolios. *Appl. Stoch. Models Bus. Indust.* 33(1):3–12.
- Hjalmarsson E, Manchev P (2012) Characteristic-based mean-variance portfolio choice. *J. Banking Finance* 36(5):1392–1401.
- Jagannathan R, Ma T (2003) Risk reduction in large portfolios: Why imposing the wrong constraints helps. *J. Finance* 58(4):1651–1683.
- Jensen TI, Kelly BT, Malamud S, Pedersen LH (2022) Machine learning and the implementable efficient frontier. Research paper no. 22-63, Swiss Finance Institute.
- Jorion P (1986) Bayes-Stein estimation for portfolio analysis. *J. Financial Quant. Anal.* 21(3):279–292.
- Kelly B, Malamud S, Zhou K (2024) The virtue of complexity in return prediction. *J. Finance* 79(1):459–503.
- Kirby C, Ostdiek B (2012a) It's all in the timing: Simple active portfolio strategies that outperform naive diversification. *J. Financial Quant. Anal.* 47(2):437–467.
- Kirby C, Ostdiek B (2012b) Optimizing the performance of sample mean-variance efficient portfolios. <https://ssrn.com/abstract=1821284>. AFA 2013 San Diego Meetings Paper.
- Kozak S, Nagel S, Santosh S (2020) Shrinking the cross-section. *J. Financial Econom.* 135(2):271–292.
- Lassance N, Martín-Utrera A, Simaan M (2024) The risk of expected utility under parameter uncertainty. *Management Sci.* 70(11): 7644–7663.
- Ledoit O, Wolf M (2008) Robust performance hypothesis testing with the Sharpe ratio. *J. Empir. Finance* 15(5):850–859.
- Liu Y, Zhou G, Zhu Y (2021) Maximizing the sharpe ratio: A genetic programming approach. Preprint, submitted January 13, <http://dx.doi.org/10.2139/ssrn.3726609>.
- Lo AW, MacKinlay AC (1997) Maximizing predictability in the stock and bond markets. *Macroeconomic Dynam.* 1(1):102–134.
- Markowitz H (1952) Portfolio selection. *J. Finance* 7(1):77–91.
- Moritz B, Zimmermann T (2016) Tree-based conditional portfolio sorts: The relation between past and future stock returns. Working paper, Ludwig Maximilian University of Munich (LMU); FINVIA, Munich, Germany.
- Murata N, Yoshizawa S, Amari S (1994) Network information criterion-determining the number of hidden units for an artificial neural network model. *IEEE Trans. Neural Netw.* 5(6): 865–872.
- Nagel S (2025) Seemingly virtuous complexity in return prediction. NBER Working Paper No. 34104, National Bureau of Economic Research, Cambridge, MA.
- Politis DN, Romano JP (1994) The stationary bootstrap. *J. Amer. Statist. Assoc.* 89(428):1303–1313.
- Skouras S (2007) Decisionmetrics: A decision-based approach to econometric modelling. *J. Econom.* 137(2):414–440.
- Tversky A, Kahneman D (1992) Advances in prospect theory: Cumulative representation of uncertainty. *J. Risk Uncertainty* 5(4):297–323.
- Welch I, Goyal A (2008) A comprehensive look at the empirical performance of equity premium prediction. *Rev. Financial Stud.* 21(4):1455–1508.