CFR Working Paper NO. 23-01

Deep Parametric Portfolio Policies

F. Simon ● S. Weibels
● T. Zimmermann

centre for financial research
cologne

Deep Parametric Portfolio Policies[*]

Frederik Simon[†]   Sebastian Weibels[‡]   Tom Zimmermann[§]

First Version: July 2022; This Version: February 2025

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 (DPPP) generate 43-102 basis points higher monthly certainty equivalent returns compared to linear policies. Looking beyond expected returns, non-linear 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.

JEL classification: G11, G12, C58, C45

Keywords: Portfolio Choice, Machine Learning, Expected Utility

---

[†]LIQID Asset Management, frederik.simon@liqid.de

[‡]University of Cologne, Institute for Econometrics and Statistics and Centre for Financial Research, weibels@wiso.uni-koeln.de

[§]University of Cologne, Institute for Econometrics and Statistics and Centre for Financial Research, tom.zimmermann@uni-koeln.de

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 non-linear 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 non-linear and interactive effects, thereby offering a more flexible approach to portfolio optimization. Third, our empirical results provide compelling evidence that the DPPP outperforms traditional linear models across a variety of settings and utility specifications, delivering economically meaningful gains in investor utility. 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.

2

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 non-linearities 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 dataset 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 function. We explore portfolio policies with and without transaction costs, short-selling constraints, and for different utility functions such as mean–variance and loss-aversion preferences. Overall, complex portfolio policies can be beneficial in all these scenarios, and 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 (DeMiguel et al., 2020; Jensen et al., 2022) that examines the importance of firm characteristics under economic constraints. Additionally, we also find important differences for different investor utility functions. For example, mean-variance investors maintain exposure to risk-related characteristics with increasing levels of risk aversion, consistent with their explicit focus on the mean-variance trade-off. In contrast, loss-averse investors increasingly emphasize reversal patterns when loss aversion increases, aligning with their asymmetric treatment of gains and losses and preference for positive skewness.

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.

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; Gu et al., 2020; Bianchi 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, we optimize a utility function and model portfolio weights directly as a function of firm characteristics.

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 non-linear 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 non-linearity 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. (2021), Chevalier et al. (2022), Jensen et al. (2022), Guijarro-Ordonez et al. (2022), Coulombe and Goebel (2024), and Feng et al. (2024). Each of these differs from ours in one or more aspects. Cong et al. (2021) 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 non-linear 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. Similar to us, Guijarro-Ordonez et al. (2022) utilize neural networks for portfolio optimization in a framework based on the idea of statistical arbitrage. In contrast, we directly map portfolio weights to stock-specific signals. Coulombe and Goebel (2024) propose a machine learning framework for directly optimizing portfolio weights with non-linear 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. (2024) 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. (2024) employ feed-forward neural networks to estimate portfolio weights by modeling a deep factor, i.e. a long-short factor based on a non-linear 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. (2023) develop a random forest-based approach to discover pricing factors and build a corresponding SDF. In contrast, our paper sidesteps the explicit factor-identification step required for SDF construction, instead directly learning the portfolio-weight function in a one-step setting under real-world constraints such as transaction costs, leverage limits, and different utility specifications.

In addition, our work relates to the literature that deals with estimation risk arising from

6

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 non-linearities 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.[1] 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 $r_{p,t+1}$:

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

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

---

[1]In the classical mean–variance setting with a risk-free asset the tangency portfolio remains constant across investors with different levels of risk aversion. However, for investors with CRRA preferences —especially when constraints or non-linearities (as in our DPPP) are present— the optimal risky asset allocation can differ not only in scale (i.e. leverage) but also in composition. Even in a mean-variance world with a risk-free asset available, the presence of non-linearities and binding constraints could result in optimal risky portfolios that differ across investors with different risk aversion levels.

approach), we follow Brandt et al. (2009) and parameterize the weights as a function of firm characteristics $x_{i,t}$, i.e.

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

where θ is the coefficient vector to be estimated.

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

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

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

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

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.[2]

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 non-linear, 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

---

[2]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.

(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 to 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 Ostdiek, 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. 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 the 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 follows:

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

where $r_{b,t+1}$ is the benchmark return and $r_{c,t+1}$ contains characteristic portfolio returns. Following 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 $S_t = A(\Omega X_t)$ represents non-linear transformations of the original characteristics. Here, Ω is a $P \times K$ matrix of random weights that creates linear combinations of the K original characteristics, $A(\cdot)$ is a non-linear activation function applied elementwise, and Ξ is a matrix of weights that maps the P transformed features to the original coefficients. While this transformation introduces high-dimensional intermediate representations through S_t, the final coefficients $\theta^T \Xi^T$ maintain the linear structure of the optimization problem. This interpretation allows us to extend our theoretical results from the linear characteristic-sorted portfolios to a more flexible deep portfolio policy, where the mapping between characteristics and characteristic-sorted portfolio coefficients is enriched through non-linear transformations implemented by a neural network, while preserving the fundamental structure of the optimization over these portfolio coefficients.

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

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

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

The optimal deviations from the benchmark portfolio converge as follows:

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

In particular, if the benchmark is uncorrelated with the active positions (i.e., $\hat{\sigma}_{bc} = 0$), then the active positions converge to zero.

The proof is in Appendix A.1.

High risk aversion forces both PPP and DPPP to prioritize risk minimization, leading to convergence for the mean absolute weight differences between the models. We summarize this in

the following proposition

Proposition 2. *Compare two models of different complexity, with corresponding portfolio weights $w_{PPP}(\gamma)$ and $w_{DPPP}(\gamma)$. The portfolio weight difference can be decomposed as*

$$w_{PPP}(\gamma) - w_{DPPP}(\gamma) = \underbrace{\frac{1}{\gamma}\Delta_{RP}}_{\text{Risk premium difference}} + \underbrace{\Delta_{RM}}_{\text{Risk minimization difference}}, \tag{8}$$

where Δ_{RP} and Δ_{RM} are deterministic quantities independent of γ that depend only on the problem structure (X_t, b_t, and r_{t+1}). In the high risk–aversion limit, the risk premium differences goes to zero, so that

$$\lim_{\gamma \to \infty} \|w_{PPP}(\gamma) - w_{DPPP}(\gamma)\| = \|\Delta_{RM}\| = c. \tag{9}$$

In particular, if the benchmark is uncorrelated with the active positions (i.e., $\hat{\sigma}_{bc,PPP} = 0$ and $\hat{\sigma}_{bc,DPPP} = 0$), then the difference between the PPP and DPPP weights converge to zero.

The proof is in Appendix A.2.

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

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:

$$\text{EDF} = \text{tr}(G^{-1}V)/T, \tag{10}$$

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 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*

$$\lim_{\gamma \to \infty} EDF = \lim_{\gamma \to \infty} \frac{p}{\gamma} = 0. \qquad (11)$$

The proof is in Appendix A.3.

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 PPP and DPPP established in the previous section. As $\gamma \to \infty$, the EDF approaches zero for both models, meaning they effectively become less complex regardless of their nominal parameter count. For PPP, the number of effective parameters p/γ goes to zero. Similarly for DPPP, despite having a richer non-linear structure through $S_t = A(\Omega X_t)$, 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.

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.

Following our theoretical framework where $S_t = A(\Omega X_t)$, we expand the feature space with random Fourier features. Specifically, we draw random vectors $w^j \stackrel{\text{iid}}{\sim} N(0, \eta^2 I)$ for $j = 1, ..., p/2$. For each j, we create a pair of new features using sine and cosine transformations. The complete random Fourier feature vector for firm i at time t is:

$$RF(x_{i,t}) = \frac{1}{\sqrt{p}} [\sin(\langle w^1, x_{i,t} \rangle), \cos(\langle w^1, x_{i,t} \rangle), ..., \sin(\langle w^{p/2}, x_{i,t} \rangle), \cos(\langle w^{p/2}, x_{i,t} \rangle)]', \qquad (12)$$

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.

Returns are generated with predictability based on the base characteristics $r_{i,t+1} = x_{i,t}^\top \beta + \eta_{i,t+1}$, 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.

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 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.
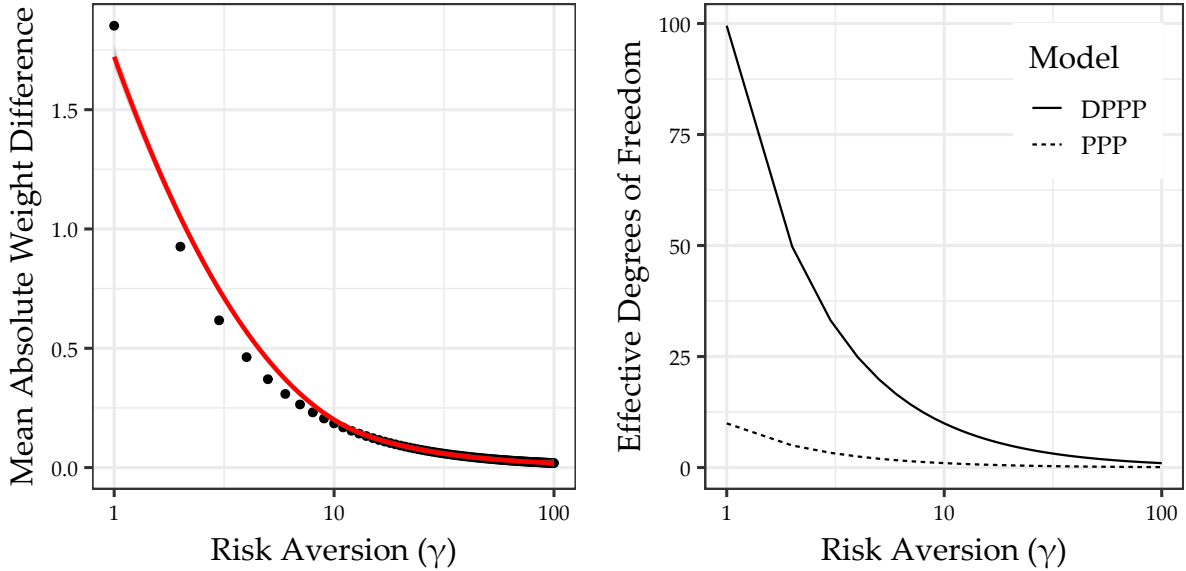
Figure 1: Risk aversion as economic regularization
This figure presents simulation evidence demonstrating that risk aversion acts as an economic regularization mechanism. We compare two nested parametric portfolio policies of different complexity: one using 10 characteristics (PPP) and a second one using 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.

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 (with or without networks) tries to minimize some distance metric (e.g., mean squared error) between e.g., observed stock returns 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 θ is the constant relative risk aversion (CRRA) utility:

$$\mathcal{L}(\theta) = -\frac{1}{T} \sum_{t=1}^{T} \left(\frac{(1 + r_{p,t+1}(\theta))^{1-\gamma}}{1 - \gamma} \right),$$ (13)

where γ is the relative risk aversion parameter. Note that minimizing Equation (13) 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 non-linearities. In other words, unlike time series models (such as RNNs or LSTMs) 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.
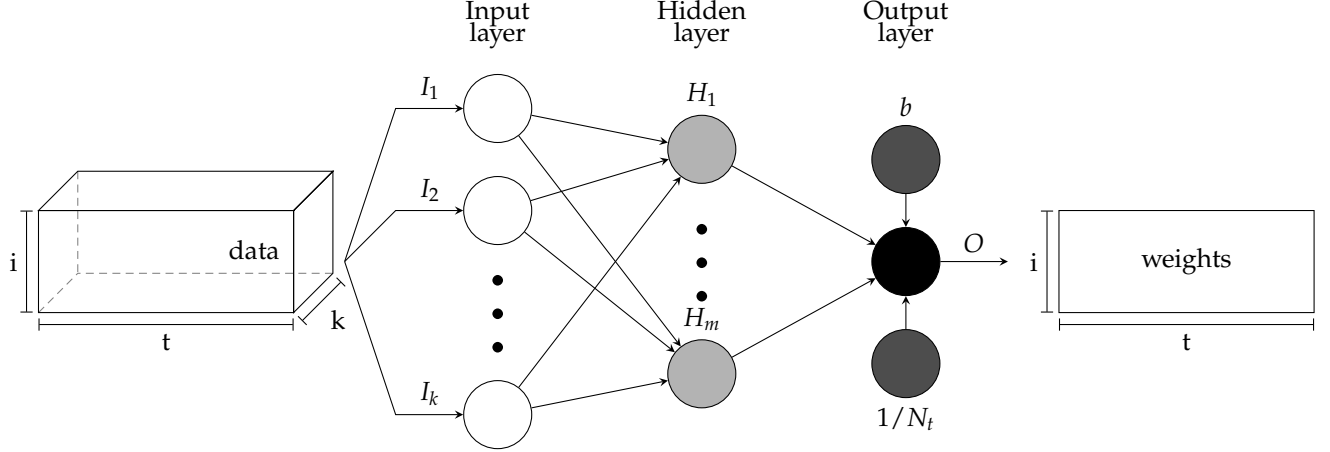
Figure 2: Neural network structure
This figure presents the core structure of our neural networks. White circles denote the input layer, grey circles denote the hidden layer and black circles denote the output layer. The data cube on the left depicts the structure of our data, i.e., we have k variables across i cross-sections in t periods. The rectangle on the right depicts our output, i.e., 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.

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 is 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 time period into a portfolio return that is the input of the loss function in Equation (1).

Constructing a neural network requires many design choices, including e.g. the depth (number of layers) and width (units per layer) of the model, or the activation function for different units and layers,[3] and selecting the optimal network architecture is a challenging task. We simplify

---

[3]The activation function introduces non-linearity 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.

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 $|3\%|$, i.e.

$$|w_{i,t}| \leq 0.03, \tag{14}$$

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.[4] This constraint is formulated for every period t as

$$\sum_{i=1}^{N_t} w_i I(w_i < 0) \geq -1, \tag{15}$$

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 non-linear 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

---

[4]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%.

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 Supplementary Appendix 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).[5] 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 dataset of Chen and Zimmermann (2022). The dataset contains monthly US stock-level data on 205 cross-sectional stock return predictors, covering the period from January 1925 to December 2020.

We focus on the period from January 1971 to December 2020, since comprehensive accounting data is only sparsely available in the years prior to that. In addition, we only keep common stocks, i.e. 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.[6]

Finally, we follow Gu et al. (2020) and replace missing values with the cross-sectional median at each month for each stock, respectively.[7] 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

---

[5]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.

[6]All characteristics are calculated at a monthly frequency. For variables that are updated at a lower frequency, the monthly value is simply the last observed value. We assume the standard lag of six months for annual accounting data availability and a lag of one quarter for quarterly accounting data availability. For IBES, we assume that earnings estimates are available by the end date of the statistical period. For other data, we follow the respective original research in regards to availability.

[7]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.

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 dataset contains 157 predictors for a total of 5,154 firms. Each month, the dataset 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.5.1 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.[8] 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.

The economic significance of the deep learning approach is most evident in the certainty equivalent returns (CE).[9] At a risk aversion level of $\gamma = 2$, the DPPP achieves a CE of 2.97% compared to the PPP's 1.95%, representing a significant enhancement of 102 basis points (p-value = 0.0003).[10] This substantial improvement suggests that capturing non-linear 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

---

[8]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.

[9]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.

[10]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.

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
$\sum |w_i|/N_t * 100$	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
$\sum I(w_i < 0)/N_t$	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

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 Dataset spanning January 1971 to December 2020. Results are shown for Constant Relative Risk Aversion (CRRA) investors with relative risk aversion coefficients (γ) 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 Deep Parametric Portfolio Policy (DPPP) and 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 DPPP and 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.

(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. 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 (e.g. average absolute weights of 0.19% versus 0.17% for the PPP at $\gamma = 2$) and higher turnover than the linear portfolio policy.[11] 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 similar maximum drawdowns (e.g. 49.79% for the DPPP versus 63.02% for the PPP at $\gamma = 2$), maximum one-month losses (22.64% versus 21.40%) or Conditional Value at Risk (CVaR) (11.07% versus 10.44%). 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 1.52 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 (solid lines)

---

[11]See Section 3.6 for a formal definition of turnover.
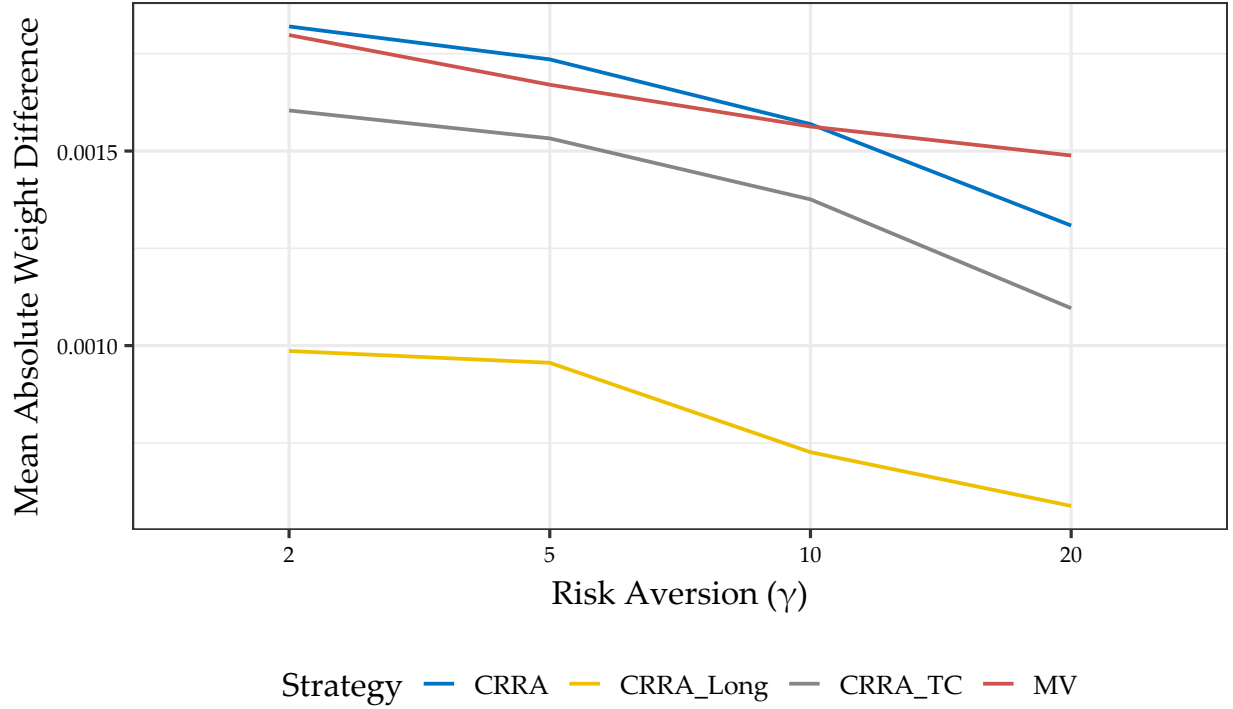
Figure 3: Mean absolute portfolio weight differences by risk aversion
The plot shows mean absolute portfolio weight differences between the PPP and DPPP models. Different lines refer to different model specifications, with CRRA (the blue line) the models in Table 1, CRRA_TC (grey line) the models in Table 3, CRRA_Long (yellow line) the models in Table 3 and MV (red line) the models in Table 4.

consistently outperforms its linear counterpart (dotted lines) across all risk aversion levels, with the outperformance becoming more pronounced after the 2008 financial crisis. Second, lower risk aversion portfolios ($\gamma = 2$, blue line) 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 to 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 non-linear modeling are particularly valuable in capturing upside potential while still providing some downside protection.
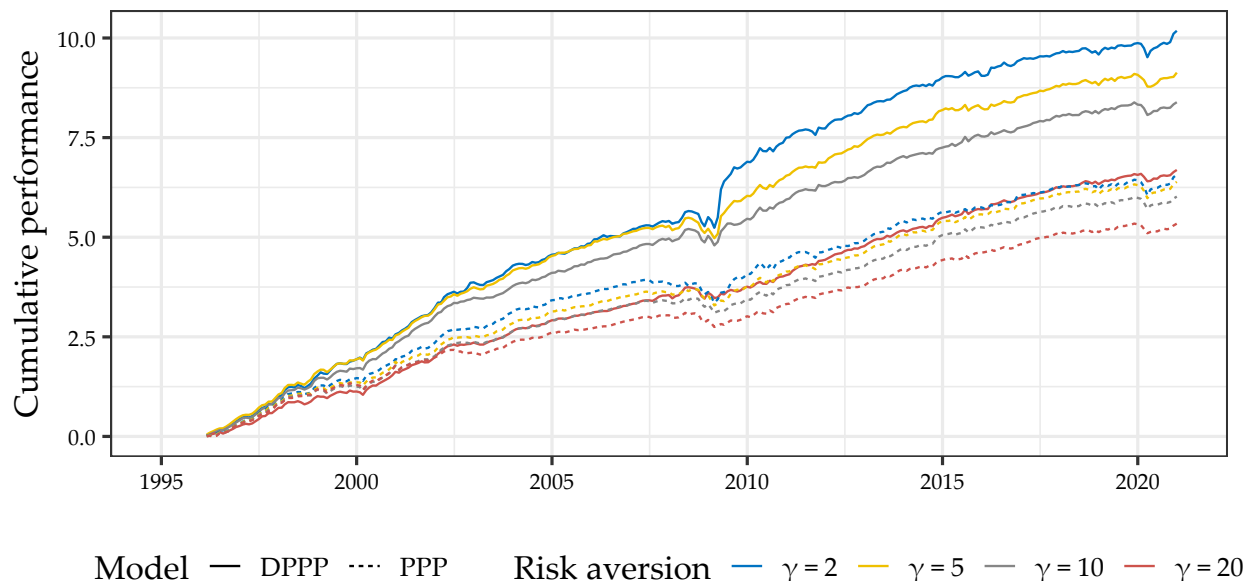
Figure 4: Cumulative performance over time for CRRA preferences
The plot shows the cumulative sum of portfolio returns for the DPPP and PPP. We show the results for each of the degrees of relative risk aversion considered and across all out-of-sample periods.

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 expected 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 to 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 to the PPP. This suggests that the non-linear 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 balanced importance across characteristics, particularly in the DPPP, consistent with the results of DeMiguel et al. (2020).

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 non-linear 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 non-linear 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 non-linearity plays a role, we fit linear surrogate models to explain the portfolio weights of the DPPP policy. The findings indicate that 30–60% of the
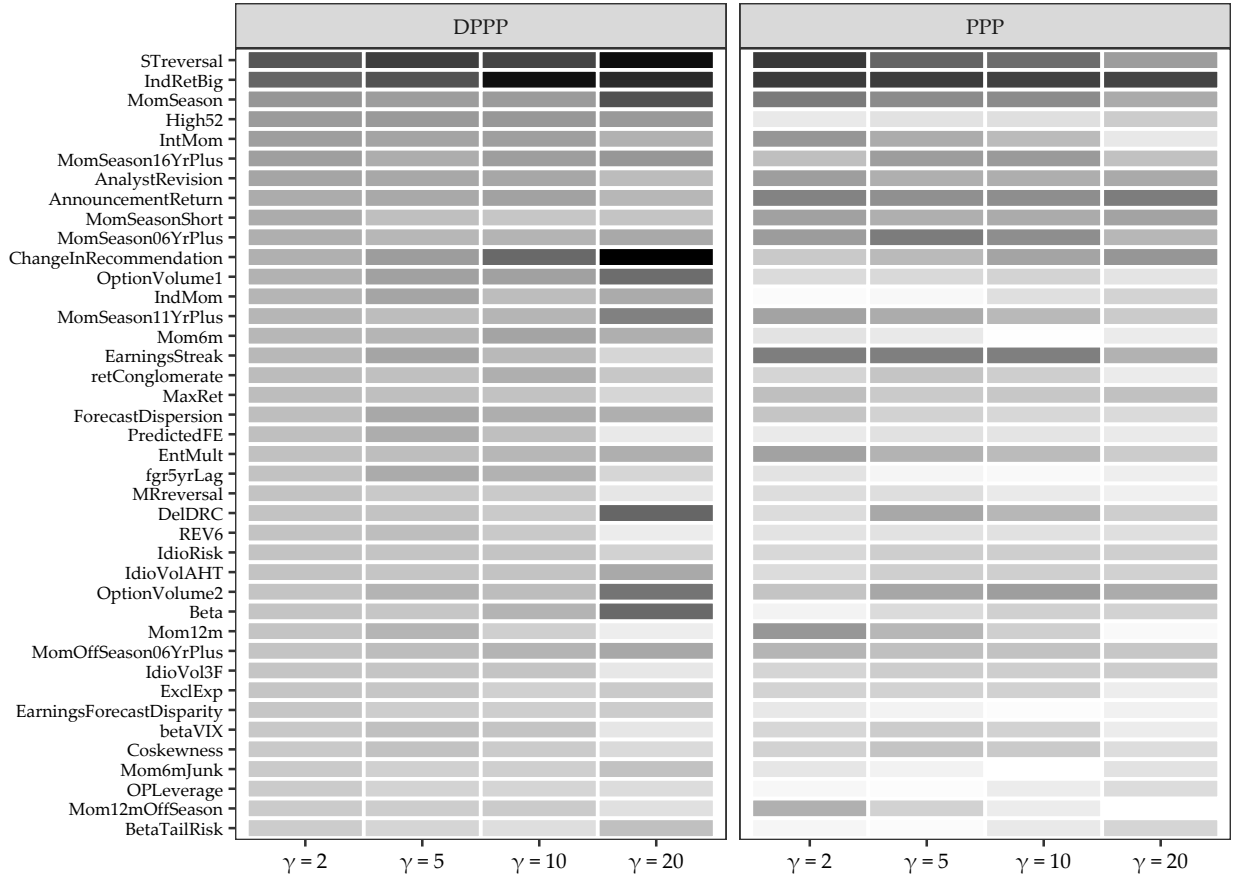
Figure 5: Variable importance for the CRRA preference for the DPPP and the PPP
Variable importance for the 40 most influential variables in the PPP and DPPP across model specifications and risk aversions, respectively. Variable importance is computed as the average absolute gradient over all training 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 $\gamma = 2$.

characteristic–weight relationship is linear, an additional 20–30% is explained by interactions, and the remaining 10–50% is due to higher-order non-linearities. Moreover, the economic significance of these non-linear components—measured via certainty equivalent differences—is most pronounced for lower risk aversion levels, underscoring that the flexibility of non-linear 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 to several benchmark models

What do investors gain by optimizing their utility function directly rather than optimizing standard benchmarks such as e.g. the Sharpe ratio? Table 2 presents a comprehensive cross-evaluation of the DPPP against several important benchmarks, including machine learning-based return forecasting, Sharpe ratio optimization, a factor portfolio, and traditional passive strategies. In other words, for each utility function, Table 2 shows the resulting investor utility had they employed a. a portfolio policy optimized for another risk aversion or b. portfolio optimization based on one the considered benchmarks.

Results show that the portfolio policy optimized for a specific risk aversion level consistently outperforms alternative portfolio policies optimized for other risk aversion levels. For instance, at $\gamma = 2$, the DPPP strategy optimized for this risk aversion achieves the highest utility, outperforming portfolio policies optimized for $\gamma = 5, 10$ or 20 by a small margin. The outperformance is more pronounced compared to classical allocation strategies (e.g. Sharpe Ratio optimization) and machine learning-based approaches. This outperformance holds across different risk aversion levels. The relative advantage becomes increasingly pronounced as risk aversion increases. This is particularly evident when compared to 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.

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, with the magnitude of outperformance being most pronounced at higher levels of risk aversion.

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.

Table 2: Cross-Evaluation of portfolio strategies against different utility preferences

Strategy	CRRA			
	$\gamma = 2$	$\gamma = 5$	$\gamma = 10$	$\gamma = 20$
DPPP ($\gamma = 2$)	**1.0000**	0.9998	0.9663	0.5374
DPPP ($\gamma = 5$)	0.9981	**1.0000**	0.9980	0.7937
DPPP ($\gamma = 10$)	0.9963	0.9956	**1.0000**	0.8207
DPPP ($\gamma = 20$)	0.9915	0.9819	0.9944	**1.0000**
SR (PPP)	0.9896	0.9752	0.9829	0.9991
SR (DPPP)	0.9887	0.9715	0.9750	0.9831
ML	0.9899	0.9595	0.8725	0.4247
Factor	0.9739	0.9139	0.8386	0.5471
EW	0.9784	0.9173	0.7887	0.3067
VW	0.9784	0.9197	0.8057	0.3707

This table presents the out-of-sample utility for various investment strategies evaluated for a CRRA investor across different risk aversion levels. The utility of each strategy is normalized by the maximum utility within each risk aversion category, so that the best performing strategy is set to one (100%), and all other values are expressed as a fraction of that optimum. Bold values indicate the best performing strategy for each preference. 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 strategy of a simple Fama-French five-factor model plus momentum. EW and VW are passive equal-weighted and value-weighted strategies.

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. We next 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 non-linear 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:

$$TC = \frac{1}{T} \sum_{t=1}^{T} \left[\sum_{i=1}^{N_t} |\kappa_{i,t}(w_{i,t} - w_{i,t-1}^+)| \right], \tag{16}$$

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,

$$w_{i,t-1}^+ = \frac{w_{i,t-1}(1 + r_{i,t})}{1 + \sum\limits_{j=1}^{N_t} w_{j,t-1} r_{j,t}}. \tag{17}$$

Our transaction cost estimates come from Chen and Velikov (2023).[12] 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 multi-period 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 (16). 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

---

[12]We thank the authors for making an updated version of the data available.

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), i.e., we restrict portfolios weights to be non-negative 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:

$$w_{i,t}^* = \frac{max[0, w_{i,t}]}{\sum\limits_{j=1}^{N_t} max[0, w_{j,t}]}. \tag{18}$$

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 to Table 1, but provide similar tables with all results in Table S.5.2 for transaction cost and Table S.5.3 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 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 78–84% 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

Table 3: Long-only & transaction costs constrained deep portfolio policy for CRRA investors with different degrees of risk aversion

Transaction costs	$\gamma = 2$		$\gamma = 5$		$\gamma = 10$		$\gamma = 20$	
	PPP	DPPP	PPP	DPPP	PPP	DPPP	PPP	DPPP
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
Long-only								
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

This table presents out-of-sample performance estimates for deep portfolio policies with the transaction costs penalty from Equation (16) 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 Dataset spanning January 1971 to December 2020. Results are shown for Constant Relative Risk Aversion (CRRA) investors with relative risk aversion coefficients (γ) 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 Deep Parametric Portfolio Policy (DPPP) and 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 DPPP and PPP.

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 76.56-135.08% versus 32.74-58.83% for the PPP. Additionally, 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.5.4 in the Supplementary Appendix).

4 Other 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

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

(19)

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

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

(20)

where $\overline{W}$ 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 $\overline{W}$ 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 to Table 1, but provide similar tables with all results in Table S.5.5 for mean-variance preference and Table S.5.6 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, i.e., 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 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 to 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 in the Supplementary Appendix, increasing loss aversion l does indeed penalize negative outcomes more severely.

Finally, it is instructive to compare portfolio return moments across utility functions and portfolio policies. Under mean-variance utility, increasing risk aversion is associated with lower portfolio return variance for both the linear (PPP) and complex (DPPP) models. There are also no differences in higher-order moments because mean-variance investors are indifferent to these characteristics. In contrast, with loss aversion utility, higher loss aversion leads to lower skewness

31

Table 4: Deep portfolio policy for mean-variance and loss-averse investors with different degrees of risk aversion (γ) and loss aversion (l)

Mean-variance preference	$\gamma = 2$		$\gamma = 5$		$\gamma = 10$		$\gamma = 20$	
	PPP	DPPP	PPP	DPPP	PPP	DPPP	PPP	DPPP
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
$\sum \lvert w_{i,t} - w_{i,t-1}^+ \rvert$	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
$\sum \lvert w_{i,t} - w_{i,t-1}^+ \rvert$	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

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 Dataset spanning January 1971 to December 2020. Results are shown for mean-variance investors with relative risk aversion coefficients (γ) 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 Deep Parametric Portfolio Policy (DPPP) and 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 DPPP and PPP.

in absolute terms, yet the portfolios diverge significantly: the PPP model produces slightly left-skewed returns, whereas the DPPP model yields right-skewed returns that better align with the preferences of loss-averse investors. This distinction highlights the relevance of accurately capturing optimally skewed portfolio return distributions when investor utility is sensitive to skewness.

Figure S.4.1 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.

4.2 Comparison of portfolio weights and variable importance

To analyze the economic differences between CRRA, mean-variance and loss-averse portfolio policies, we examine their exposure to characteristics. For each topical cluster of characteristics k and portfolio p, we calculate the portfolio's exposure as

$$E_{p,k,t} = \sum_{i=1}^{N_t} w_{i,t}^p X_{i,k,t},$$ (21)

where $w_{i,t}^p$ represents the portfolio weight of stock i at time t in portfolio p, and $X_{i,k,t}$ is the standardized value of characteristic k for stock i at time t. Since we allow for short-selling, $w_{i,t}^p$ can be negative, implying that positive characteristic exposures can arise from either long positions in stocks with positive characteristic values or short positions in stocks with negative characteristic values. Conversely, negative exposures result from long positions in stocks with negative characteristic values or short positions in stocks with positive characteristic values. We assess the economic significance of these net exposures by examining the time-series average

$$\overline{E_{p,k}} = \frac{1}{T} \sum_{t=1}^{T} E_{p,k,t}.$$ (22)

Figure 6 shows the time-series averages of net exposures $\overline{E_{p,k}}$ for a set of eight clusters. Across the panels, distinct patterns emerge in the exposure of portfolios to firm characteristics, with notable differences between utility functions and risk (or loss) aversions. The short-term reversal and size clusters exhibit a declining trend in net exposure as aversion increases, suggesting that
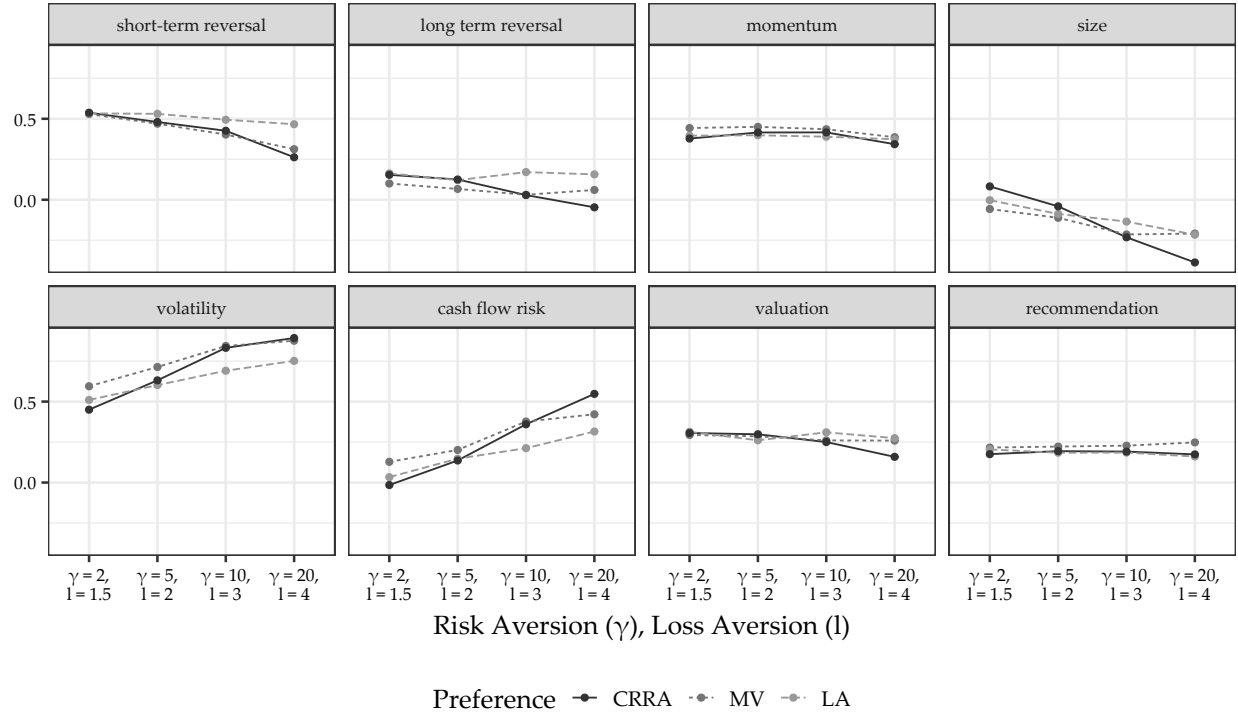
Figure 6: Net exposure to clusters of firm characteristics for different preferences
This figure shows the net expose to clusters for CRRA, mean-variance (MV) and loss-averse (LA) portfolio policies across different levels of risk and loss aversion. We group the characteristics into clusters according to the economic category specified in the Open Source Asset Pricing data set by Chen and Zimmermann (2022). Each panel presents time-series averages of net exposures to a given cluster for a specific risk aversion level γ and corresponding loss aversion parameter l.

more risk-averse investors allocate less weight to stocks associated with these characteristics. However, the decline is least pronounced for loss-averse portfolios, which exhibit systematically higher exposures to especially short-term reversal. This suggests 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.

In contrast, the volatility and cash flow risk clusters show increasing exposure with higher aversions, particularly for CRRA and mean-variance investors, who appear willing to allocate more weight to firms with higher risk in pursuit of potential returns. However, loss-averse portfolios consistently exhibit lower exposure to these clusters, indicating that risk-related characteristics are less important, likely since volatility is not a specific part of the loss function.

For the long term reversal, momentum, valuation and recommendations clusters, exposures remain relatively stable across aversion levels for all three preferences, implying that these charac-

teristics play a more neutral role in portfolio construction. Nevertheless, the long-term reversal cluster displays slight variations across investor types, with loss-averse portfolios consistently maintaining higher exposures than CRRA and mean-variance portfolios, especially for higher aversion levels. This suggests that loss-averse investors are more willing to take contrarian positions on previously underperforming stocks, whereas CRRA and mean-variance investors are less open to these opportunities.

5 Conclusion

Building on the seminal work of Brandt et al. (2009) and the extensive literature on asset allocation 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 not only maps a large set of firm characteristics to optimal portfolio weights in a non-linear and interactive manner, but it also directly incorporates market friction constraints as well as investor-specific utility functions, whether 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 non-linear 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. This result offers an economically intuitive mechanism by which investor preferences regulate model complexity.

Our empirical investigation reinforces the theoretical insights. We document substantial improvements in investor utility when adopting the DPPP relative to standard linear models. Certainty equivalent gains range between 43 and 102 basis points per month, with the magnitude of these improvements decreasing systematically with increasing risk aversion. Furthermore, the DPPP delivers robust performance across various settings, including transaction costs, short-selling constraints as well as different utility functions. 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.

Highlighting the growing role of machine learning and non-linear models in finance, our approach thus resembles 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. Moreover, the built-in economic regularization mechanism provides practitioners with a practical tool to select the optimal level of model complexity based on their individual risk-return preferences.

References

Ammann, M., G. Coqueret, and J.-P. Schade (2016). Characteristics-based portfolio choice with leverage constraints. *Journal of Banking & Finance 70*, 23–37.

Ang, A., S. Gorovyy, and G. B. van Inwegen (2011). Hedge fund leverage. *Journal of Financial Economics 102*(1), 102–126.

Bianchi, D., M. Büchner, and A. Tamoni (2020). Bond Risk Premiums with Machine Learning. *The Review of Financial Studies 34*(2), 1046–1089.

Brandt, M. W., P. Santa-Clara, and R. Valkanov (2009). Parametric Portfolio Policies: Exploiting Characteristics in the Cross-Section of Equity Returns. *The Review of Financial Studies 22*(9), 3411–3447.

Bryzgalova, S., M. Pelger, and J. Zhu (2023). Forest through the trees: Building cross-sections of stock returns. Working paper.

Chen, A. Y. and J. McCoy (2024). Missing values handling for machine learning portfolios. *Journal of Financial Economics 155*, 103815.

Chen, A. Y. and M. Velikov (2023). Zeroing in on the expected returns of anomalies. *Journal of Financial and Quantitative Analysis 58*(3), 968–1004.

Chen, A. Y. and T. Zimmermann (2022). Open source cross-sectional asset pricing. *Critical Finance Review 27*(2), 207–264.

Chen, L., M. Pelger, and J. Zhu (2024). Deep learning in asset pricing. *Management Science 70*(2), 714–750.

Chevalier, G., G. Coqueret, and T. Raffinot (2022). Supervised portfolios. *Quantitative Finance 22*(12), 2275–2295.

Cong, L., K. Tang, J. Wang, and Y. Zhan (2021). Alphaportfolio: Direct construction through deep reinforcement learning and interpretable ai. Working paper.

Coulombe, P. G. and M. Goebel (2024). Maximally machine-learnable portfolios. Working paper.

DeMiguel, V., L. Garlappi, and R. Uppal (2009). Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *The Review of Financial Studies 22*(5), 1915–1953.

DeMiguel, V., A. Martin-Utrera, and R. Uppal (2024). A multifactor perspective on volatility-managed portfolios. *Journal of Finance 79*(6), 3859–3891.

DeMiguel, V., A. Martín-Utrera, F. J. Nogales, and R. Uppal (2020). A Transaction-Cost Perspective on the Multitude of Firm Characteristics. *The Review of Financial Studies 33*(5), 2180–2222.

Detzel, A., R. Novy-Marx, and M. Velikov (2023). Model comparison with transaction costs. *Journal of Finance 78*(3), 1743–1775.

Didisheim, A., S. B. Ke, B. T. Kelly, and S. Malamud (2023). Complexity in factor pricing models. Working paper, National Bureau of Economic Research.

Feng, G., L. Jiang, J. Li, and Y. Song (2024). Deep tangency portfolio. Working paper.

Freyberger, J., A. Neuhierl, and M. Weber (2020). Dissecting characteristics nonparametrically. *The Review of Financial Studies 33*(5), 2326–2377.

Gu, S., B. Kelly, and D. Xiu (2020). Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies 33*(5), 2223–2273.

Guijarro-Ordonez, J., M. Pelger, and G. Zanotti (2022). Deep learning statistical arbitrage. Working paper.

Hautsch, N. and S. Voigt (2019). Large-scale portfolio allocation under transaction costs and model uncertainty. *Journal of Econometrics 212*(1), 221–240.

Heaton, J. B., N. G. Polson, and J. H. Witte (2017). Deep learning for finance: deep portfolios. *Applied Stochastic Models in Business and Industry 33*(1), 3–12.

Hjalmarsson, E. and P. Manchev (2012). Characteristic-based mean-variance portfolio choice. *Journal of Banking & Finance 36*(5), 1392–1401.

Ioffe, S. and C. Szegedy (2015, 07–09 Jul). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Proceedings of the 32nd International Conference on Machine Learning 37*, 448–456.

Jagannathan, R. and T. Ma (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance 58*(4), 1651–1683.

Jensen, T. I., B. T. Kelly, S. Malamud, and L. H. Pedersen (2022). Machine learning and the implementable efficient frontier. Research paper no. 22-63, Swiss Finance Institute.

Kelly, B., S. Malamud, and K. Zhou (2024). The virtue of complexity in return prediction. *The Journal of Finance 79*(1), 459–503.

Kingma, D. P. and J. Ba (2014). Adam: A method for stochastic optimization. Working paper, arXiv.

Kirby, C. and B. Ostdiek (2012a). It's all in the timing: simple active portfolio strategies that outperform naive diversification. *Journal of financial and quantitative analysis 47*(2), 437–467.

Kirby, C. and B. Ostdiek (2012b). Optimizing the performance of sample mean-variance efficient portfolios. AFA 2013 San Diego Meetings Paper.

Kozak, S., S. Nagel, and S. Santosh (2020). Shrinking the cross-section. *Journal of Financial Economics 135*(2), 271–292.

Lassance, N., A. Martín-Utrera, and M. Simaan (2024). The risk of expected utility under parameter uncertainty. *Management Science 70*(11), 7644–7663.

Ledoit, O. and M. Wolf (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance 15*(5), 850–859.

Liu, Y., G. Zhou, and Y. Zhu (2024). Maximizing the sharpe ratio: A genetic programming approach. Working paper.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance 7*(1), 77–91.

Masters, T. (1993). *Practical Neural Network Recipes in C++.* Academic Press Professional, Inc.

Moritz, B. and T. Zimmermann (2016). Tree-based conditional portfolio sorts: The relation between past and future stock returns. Working paper.

Murata, N., S. Yoshizawa, and S. Amari (1994). Network information criterion-determining the number of hidden units for an artificial neural network model. *IEEE Transactions on Neural Networks 5*(6), 865–872.

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

Skouras, S. (2007). Decisionmetrics: A decision-based approach to econometric modelling. *Journal of Econometrics 137*, 414–440.

Srivastava, N., G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research 15*(56), 1929–1958.

Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty 5*(4), 297–323.

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

Appendix A   Proofs

A.1   Proof of Proposition 1

For this proof consider a second-order Taylor expansion of the CRRA utiltiy function around the expected portfolio return:

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

This approximation is well-known and shows that the CRRA utility framework naturally reduces to mean-variance preferences, where γ 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:

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

with first-order condition:

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

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

A.2   Proof of Proposition 2

We aim to show that the difference between the portfolio weights of the PPP and DPPP decreases with risk aversion.

Using Equation (6), the portfolio weight difference between the PPP and DPPP models is given by Equation (8). Thus, in the high risk aversion limit, the difference simplifies to:

$$\lim_{\gamma \to \infty} \|w_{PPP}(\gamma) - w_{DPPP}(\gamma)\| = \underbrace{\lim_{\gamma \to \infty} \frac{1}{\gamma} \|\Delta_{\text{RP}}\|}_{=0} + \underbrace{\lim_{\gamma \to \infty} \|\Delta_{\text{RM}}\|}_{=c} = c, \tag{A.4}$$

where c is a constant that depends on the structure of the risk minimization term. This establishes the stated convergence rate and Proposition 2.

A.3   Proof of Proposition 3

For this proof consider the definition of the EDF in Equation (10). We define the two key matrices as:

$$G = \mathbb{E} \left[\frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'} \right], \tag{A.5}$$

and

$$V = \mathbb{E} \left[\frac{\partial L(\theta)}{\partial \theta} \frac{\partial L(\theta)}{\partial \theta'} \right]. \tag{A.6}$$

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

$$G = \frac{1}{T}\gamma\hat{\Sigma}_c, \tag{A.7}$$

$$\frac{\partial L(\theta)}{\partial \theta} = r_{c,t+1} - \gamma\hat{\Sigma}_c \left(\frac{1}{\gamma}\hat{\Sigma}_c^{-1}\hat{\mu}_c - \hat{\Sigma}_c^{-1}\hat{\sigma}_{bc}\right) - \gamma\hat{\sigma}_{bc} = r_{c,t+1} - \hat{\mu}_c, \tag{A.8}$$

and

$$V = \frac{1}{T}\left(r_{c,t+1} - \hat{\mu}_c\right)^T \left(r_{c,t+1} - \hat{\mu}_c\right) = \hat{\Sigma}_c. \tag{A.9}$$

Therefore, our EDF measure simplifies to

$$\text{EDF} = \text{tr}(G^{-1}V)/T = \frac{1}{\gamma}\text{tr}(\hat{\Sigma}_c^{-1}\hat{\Sigma}_c) = \frac{p}{\gamma}, \tag{A.10}$$

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

Supplementary Appendix for "Deep Parametric Portfolio Policies"

- Appendix S.1: Neural Network Configuration

- Appendix S.2: Loss Aversion as Economic Regularization

- Appendix S.3: Additional results and extensions

  - S.3.1: Partial dependence and surrogate models

  - S.3.2: Rolling window estimation

  - S.3.3: Adding macroeconomic variables

- Appendix S.4: Supplementary figures

- Appendix S.5: Supplementary tables

Appendix S.1 Neural network configuration

Our benchmark model consists of an input layer, three to five hidden layers and an output layer. We apply the geometric pyramid rule (Masters, 1993), i.e., the first hidden layer consists of 32 nodes, and for each subsequent layer, the number of units is halved.

At each node of the network, a linear transformation of the preceding outputs is fed into an activation function. We choose to use the leaky rectified linear unit (leaky ReLU) activation function at every node:

$$R(z) = \begin{cases} z & \text{if } z > 0 \\ \alpha z & \text{otherwise} \end{cases}, \tag{S.1.1}$$

where z denotes the input and α denotes some small non-zero constant, in our case 0.01. ReLU is the most popular activation function because it is cheap to compute, converges fast and is sparsely activated. The disadvantage of transforming all negative values to zero is a problem called "dying ReLU". A ReLU neuron is "dead" if it is stuck in the negative range and always outputs zero. Since the slope of ReLU in the negative range is also zero, it is unlikely that a neuron will recover once it goes negative. Such neurons play no role in discriminating inputs and are essentially useless. Over time, a large part of the network may do nothing. Leaky ReLU fixes this problem because it has small slope for negative values instead of a flat slope. Moreover, we shift the activation function at every node in every hidden layer by adding a constant. This is commonly referred to as bias in the machine learning literature.

Our benchmark network is estimated by minimizing the loss function (utility function) given in Equation (13). To do so, we apply the commonly used ADAM stochastic gradient descent optimization technique developed by Kingma and Ba (2014). One technical detail in our implementation involves handling missing firm-date observations in our three-dimensional input tensor (see Figure 2). Since not every firm is observed at every point in time, some entries in the tensor are missing. To maintain a consistent tensor shape for computational purposes, we fill these missing entries with zeros. However, because these zeros do not represent actual data, we add a masking layer to our network. This layer ensures that any firm missing in a particular month is excluded from the utility calculation, so that only real observations contribute to the model's performance.

To control for the non-linearity and heavy parametrization of the model, we employ different regularization techniques to prevent overfitting: first, as mentioned in Section 3.1, we impose constraints on portfolio weights, i.e., the absolute portfolio weight of a single stock cannot exceed |3%| and the portfolio leverage in any period cannot exceed 100%. The weight constraint is imposed via an additional penalty term in the objective function. Specifically, we augment the utility function with a term of the form

$$\lambda \sum max(0, |w_i| - .03), \tag{S.1.2}$$

where λ is a large positive constant. This penalty becomes positive when any weight w_i exceeds ± 0.03. Effectively, this forces the network to keep $|w_i|$ below 0.03 in order to avoid incurring a large penalty in the objective function.

In practice, we set λ to be sufficiently large so that the violation of the constraint becomes very costly, while still allowing the optimizer enough flexibility to search for the best feasible solution. Hence, although our implementation is not a "hard" constraint in the sense of truncating outputs directly, in our estimations the model learns to remain strictly within $|w_i| < 0.03$. We verified empirically that none of the weights exceed this bound during optimization or at convergence. We implement the constraint on leverage analogously.

Second, we add a lasso (l_1) penalty term to the loss function to be minimized. Adding the penalty implies a potential shrinkage of coefficients towards zero. This in turn reduces the variance of the prediction, i.e., prevents overfit of the model.

Third, we employ early stopping on the validation data. Early stopping refers to a very general regularization technique. At each new iteration, predictions are estimated for the validation sample, and the loss (utility) is constructed. The optimization is terminated when the validation sample loss starts to increase by some small specified number (tolerance) over a specified number of iterations (patience). Typically, the termination occurs before the loss is minimized in the training sample. Early stopping is a popular regularization tool because it reduces the computational cost.

Fourth, we implement a dropout layer before the first hidden layer (Srivastava et al., 2014). The basic idea of dropout is to randomly remove units (and their connections) from the neural network during training. This prevents the units from becoming too similar. During training,

3

samples are taken from an exponential number of different thinned networks. At test time, it is easy to approximate the effect of averaging the predictions of all these thinned networks by simply using a single, unthinned network with smaller weights. The combination of a dropout layer, l_1-regularization and early stopping tremendously helps to reduce overfitting and model complexity.

Finally, we adopt our own version of a batch normalization algorithm (Ioffe and Szegedy, 2015). In general, training deep neural networks is complicated by the fact that the distribution of inputs to each layer changes during training as the parameters of the previous layers change. This phenomenon is referred to as internal covariate shift and can be remedied by normalizing the layer inputs. The strength of this method is that normalization is part of the model architecture and is performed for each training mini-batch. Batch normalization allows much higher learning rates to be used and less care to be taken in initialization. Brandt et al. (2009) standardize characteristics cross-sectionally to have zero mean and unit standard deviation across all stocks at date t. Hence, the model predictions represent deviations from the benchmark portfolio. However, applying the aforementioned activation function destroys this structure. In our model each observation can be interpreted as a complete cross-section (e.g., a batch size of 12 refers to 12 complete cross-sections of data). However, the model of Brandt et al. (2009) requires normalization on a cross-sectional level instead of a batch level. Thus, we employ our own version of cross-sectional normalization after applying the activation function in each hidden layer, such that the output of each node in the hidden layer is standardized 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 also be interpreted as a deviation from the benchmark portfolio.

We provide a summary of the relevant hyperparameters in Table S.1.1. Models are estimated using TensorFlow via Keras. The estimation of a single DPPP model including hyperparameter tuning takes about six hours with a standard retail GPU.

For our estimation strategy, we follow Brandt et al. (2009) and Gu et al. (2020) and use an expanding window strategy to generate out-of-sample results. More specifically, we split our data into a training sample used to estimate the model, a validation sample used to tune the hyperparameters of the model and a test sample used to evaluate the out-of-sample performance of the model.

Table S.1.1: Hyperparameters

	PPP	DPPP
L1 penalty	$\lambda \in \{0, 10^{-5}, 10^{-3}\}$	$\lambda \in \{0, 10^{-5}, 10^{-3}\}$
Learning Rate	0.001	0.001
Dropout	0	$D \in \{0, 0.2, 0.4\}$
Batch Size	60	60
Epochs	200	200
Patience	30	30
Hidden Layers	−	$H \in \{3, 4, 5\}$
Leaky ReLU	−	0.01

This table gives the hyperparameters that we tune. The first column shows the hyperparameters for the linear parametric portfolio policy (PPP). The second column shows the hyperparameters for the deep parametric portfolio policy (DPPP). For the DPPP, we start with 32 units in the first hidden layer, and for each subsequent layer, the number of units is halved.

We initially train the model on the first 20 years of the dataset, validate it on the following five years and evaluate its out of-sample-performance on the 12 months following the validation window. We then recursively increase the training sample by one year. Each time the training sample is increased, we refit the entire model while holding the size of the validation and test window fixed. The result is a sequence of out-of-sample periods corresponding to each expanding window, in our case 25 in total. This corresponds to a total out-of-sample period of 300 months. Note that this approach ensures that the temporal ordering of the data is maintained. The testing strategy is depicted graphically in Figure S.1.1.
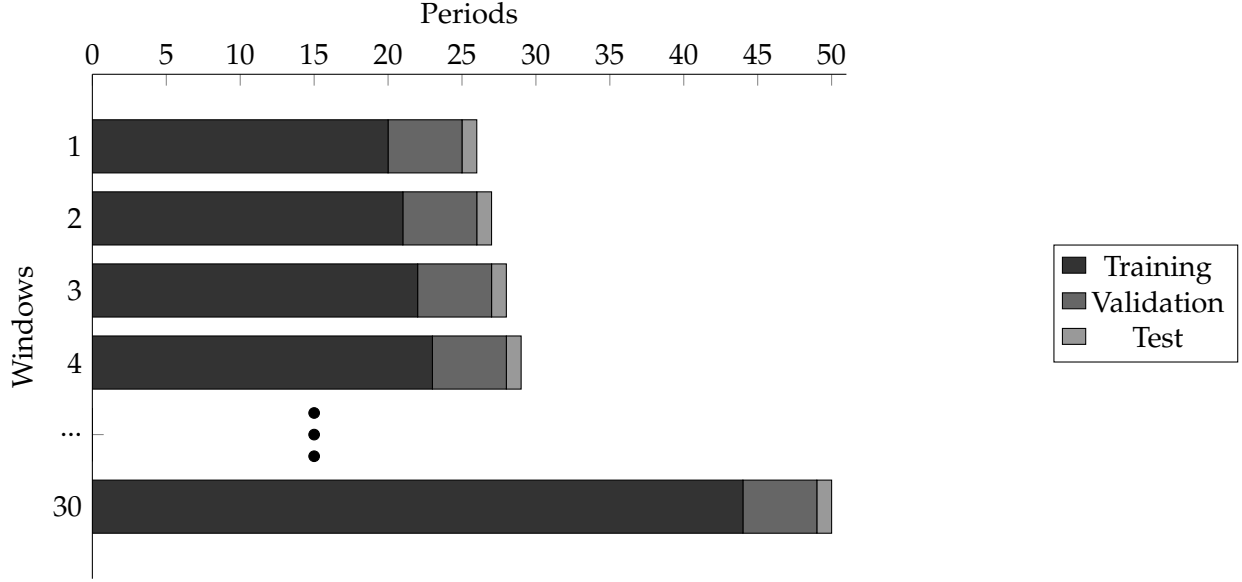
Figure S.1.1: Out-of-sample testing strategy
This figure presents our out-of-sample testing strategy. We recursively increase our training window, presented by the black portion of each bar, while holding the validation and the test window constant, presented by the grey portions of each bar.

Appendix S.2 Loss aversion as economic regularization

Building on Tversky and Kahneman (1992), consider a loss-averse investor with the piecewise-linear utility function as in section 4:

$$
u(r_{p,t+1}) = \begin{cases} -l(\overline{W} - r_{p,t+1})^b & \text{if } r_{p,t+1} < \overline{W} \\ (r_{p,t+1} - \overline{W})^b & \text{otherwise} \end{cases}, \tag{S.2.1}
$$

where l measures the strength of loss aversion relative to gains, and $\overline{W}$ is a reference return (here $\overline{W} = 0$) and the parameter b captures the degree of risk seeking over losses and risk aversion over gains (here $b = 1$).

When l is large, negative deviations $1 + r_{p,t+1} < \overline{W}$ are penalized heavily. This parallels the way high risk aversion γ suppresses active exposures in mean-variance and CRRA frameworks: high l shrinks portfolios toward safer strategies that minimize downside realizations. Hence, loss aversion naturally acts as an economic regularizer against overfitting predictive signals, just as high γ penalizes variance in mean-variance or CRRA utility.

6

Consider an investor's optimization problem

$$\max_{\theta} \; \mathbb{E}\big[u_{(r_{p,t+1}(\theta))}\big], \tag{S.2.2}$$

where $r_{p,t+1}(\theta)$ denotes the portfolio return. We decompose the expected utility:

$$\mathbb{E}[u_l(r_{p,t+1}(\theta))] = \mathbb{E}\big[(r_{p,t+1} - \overline{W})\,\mathbf{1}\{r_{p,t+1} \geq \overline{W}\}\big] \; + \; l\,\mathbb{E}\big[(r_{p,t+1} - \overline{W})\,\mathbf{1}\{r_{p,t+1} < \overline{W}\}\big]. \tag{S.2.3}$$

As $l \to \infty$, the second term dominates unless the probability of shortfalls $\{r_{p,t+1} < \overline{W}\}$ is driven toward zero. Consequently, the optimal portfolio θ^* converges to a solution that minimizes downside risk.

If a near riskless benchmark b_t with return $r_{b,t+1} \approx \overline{W}$ exists, active deviations eventually go to zero, mirroring the high-γ limit in Section 2.2. Formally,

$$\lim_{l\to\infty} \theta^* \; \to \; \arg\min_{\theta} \; \mathbb{E}\Big[(r_{p,t+1}(\theta) - \overline{W})\,\mathbf{1}\{r_{p,t+1}(\theta) < \overline{W}\}\Big]. \tag{S.2.4}$$

Hence, infinite loss aversion collapses the portfolio to a baseline strategy that nearly eliminates downside deviations, thereby reducing active complexity in the limit.

If $r_{b,t+1}$ is a benchmark with limited downside, then the optimal deviation from b_t converges to zero in the loss-aversion limit:

$$\lim_{l\to\infty} \|\theta^*\| = 0, \tag{S.2.5}$$

and the investor asymptotically holds the benchmark.

The above limit behavior implies an effective shrinkage of parameters in parametric portfolio policies or more flexible deep portfolio policies. In analogy to the CRRA and mean-variance case (see Section 2.2), one can interpret l as scaling the penalty on negative outcomes. Large l forces the model to reduce the probability of downside, thereby reducing the susceptibility to overfitting in high-dimensional or non-linear representations.

Although the piecewise function is not globally differentiable (making a simple closed-form Effective Degrees of Freedom (EDF) derivation more involved), the economic intuition is clear: as $l \to \infty$, any parameter that increases downside risk is curtailed. This downside penalty

channels into the Hessian of the loss, compressing active exposures in a manner analogous to L2 regularization, but founded on investor preferences rather than a purely statistical criterion.

A simple simulation parallel to Section 2.2 confirms that as l increases, we see a convergence in portfolio weights. The results are depicted in Figure S.2.1. Thus, loss aversion plays a role akin to a built-in regularization term that trims complexity to mitigate estimation risk.

Our findings complement the results in Section 2.2: while risk aversion γ penalizes variance, loss aversion l penalizes negative deviations, yet both yield parallel shrinkage outcomes in the respective limits $\gamma \to \infty$ or $l \to \infty$. Thus, behavioral preferences such as loss aversion can be viewed as alternative economic mechanisms that align well with the notion of regularization against overfitting.
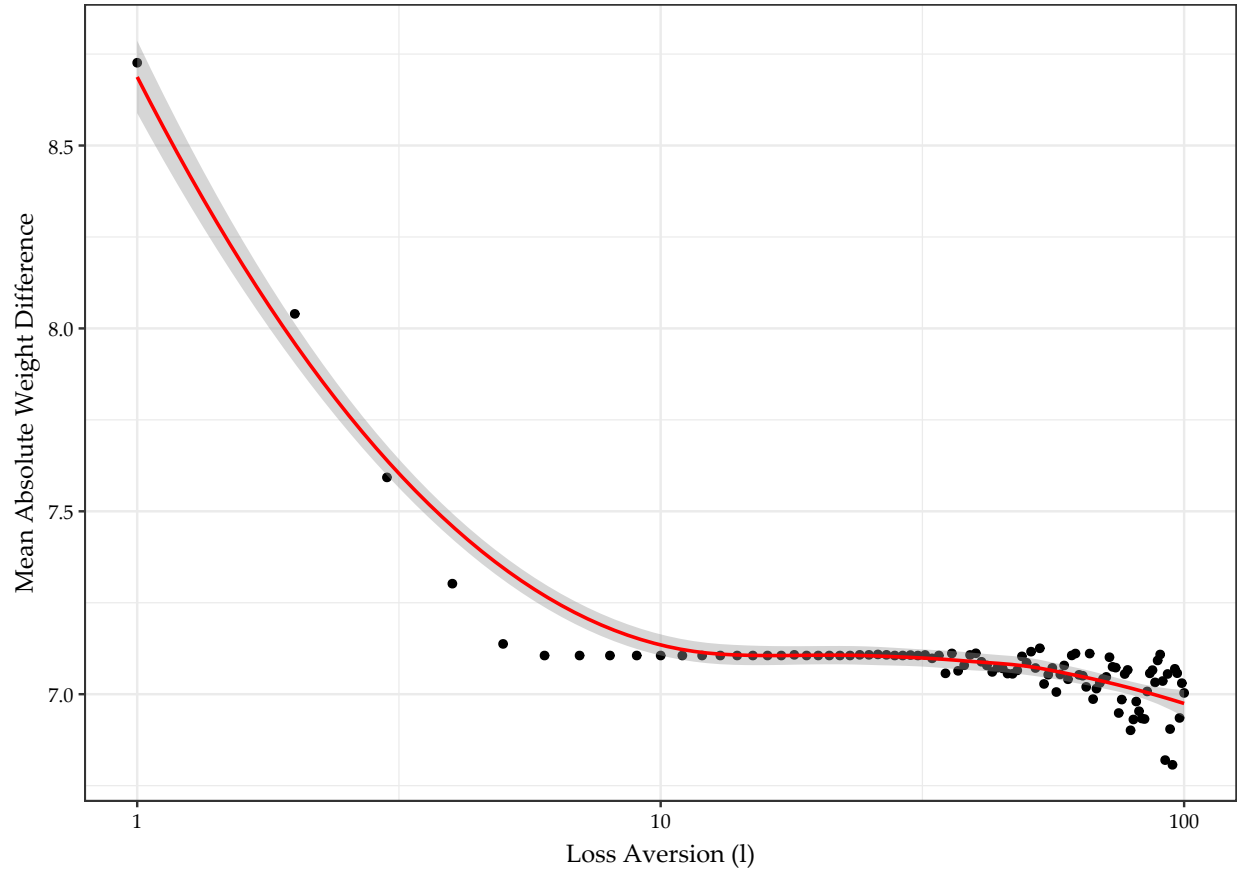
Figure S.2.1: Loss aversion as economic regularization
This figure presents simulation evidence demonstrating that risk aversion acts as an economic regularization mechanism. The simulation compares two nested parametric portfolio policies of different complexity: one using 10 characteristics (PPP) and another using 100 characteristics (DPPP) constructed through random Fourier transformations of the base characteristics. The figure shows the mean absolute difference in portfolio weights between models across loss aversion levels. All panels use a logarithmic scale (base 10) for loss aversion.

Appendix S.3 Additional results and extensions

S.3.1 Partial dependence and surrogate models

The importance of non-linear modeling of portfolio weights becomes evident when considering an investor who trades off mean return against return volatility. The investor uses standard one-dimensional portfolio sorting techniques as pictured in Figure S.3.1. Decile portfolios formed on short-term reversal or sales-to-price display monotonically increasing mean return.[13] At the same time, the standard deviations of decile portfolios are non-linear in deciles, with top and bottom decile portfolios having high standard deviations. This leads to extreme portfolios having comparatively low Sharpe ratios relative to decile portfolios in the middle of the distribution. A (long-only) investor would therefore potentially be indifferent between investing in any portfolio in the upper half of the short-term reversal distribution, and she would prefer to invest in portfolios in the middle of the sales-to-price distribution rather than investing in the extreme portfolios. Non-linear portfolio policies are able to capture these kinds of relationships.

In our application, understanding the estimated relation between input (firm characteristics) and output (estimated portfolio weights) is essential in order to shed light on the relation between firm characteristics and utility. Moreover, such an understanding allows us to compare our results to the existing literature. We provide two additional ways of interpreting the non-linearity in our models.

---

[13]We picked these two variables for illustrative purposes as these variables are the most important return- and fundamental-based variables in Gu et al. (2020).

(a) Short-Term Reversal
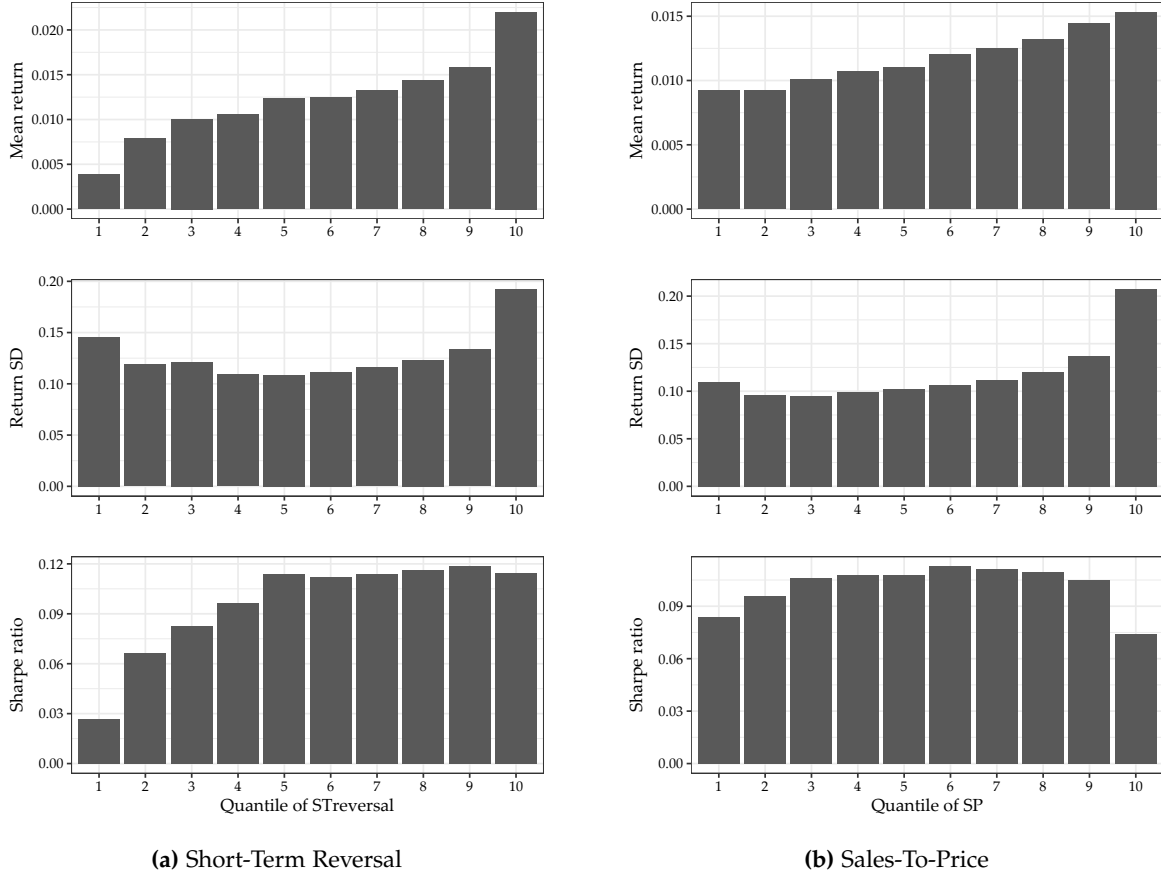
(b) Sales-To-Price

Figure S.3.1: Mean returns, standard deviations and Sharpe ratios of one-dimensional portfolio sorts
Mean returns, standard deviations and Sharpe ratios of decile portfolios sorted on short-term reversal (left panel) and sales-to-price ratio (right panel). Data is from Chen and Zimmermann (2022) and spans from 1925 to 2021.

Partial dependence

We evaluate the sensitivity of the model output to each variable. Typically, partial dependence plots provide an assessment of the variables of interest over a range of values. At each value of the variable, the model is evaluated while the remaining variables remain unchanged, and the results are then averaged across the cross-section. However, since the sum of all weights in each cross-section is equal to one and thus the mean weight prediction is always the same, applying this method to parametric portfolio policies does not yield reasonable results. To address this, we apply our own algorithm: when assessing the sensitivity with respect to variable k, we compute two sets of predictions - one with all features and another where feature k is set to zero (equivalent to its mean in our standardized setting). The difference between these predictions represents the marginal contribution of variable k to the portfolio weights. We then plot these marginal

11

contributions against the actual values of feature k to understand how the feature's impact varies across its range. We interpret this difference in predicted weights conditional on values of k as the marginal sensitivity of weights (i.e., its partial dependence) with respect to k, which allows us to assess both the magnitude and direction of each variable's influence on the portfolio allocation decision.

Figure S.3.2 depicts how different characteristics contribute to the DPPP's portfolio allocation decisions. We assess this by computing the marginal contribution of each characteristic, defined as the difference between predictions using all features and predictions where the respective feature is set to zero (its mean in our standardized setting). We examine the sensitivity with respect to three fundamental variables, namely the book-to-market ratio (BM), liquid assets (cash), and quarterly return on assets (roaq), as well as an analyst variable, namely earnings forecast revisions per share (AnalystRevision), and four past return-based variables, namely 12-month momentum (Mom12m), short-term reversal (STreversal), seasonal momentum (MomSeason), and intermediate momentum (IntMom). Recall that each predictor is signed, so that a larger value implies a higher expected return. To assess whether the marginal association of the deep model is more in line with the actual risk and return associated with each characteristic than a linear model, we include the overall Sharpe ratio for each decile portfolio sorted on each of the characteristics.

The results reveal distinct patterns in how the DPPP utilizes different characteristics. In line with the findings in regards to importance, short-term reversal exhibits the most pronounced marginal effect, as indicated by the steepness of the depicted relationship, with its impact on portfolio weights varying substantially across its range and risk aversion levels. This effect is particularly strong for low risk aversion ($\gamma = 2$), where extreme values of STreversal trigger the largest portfolio weight adjustments. The strong response aligns with the monotonically increasing Sharpe ratios across STreversal deciles, suggesting the model effectively captures this signal's risk-return profile.

Most characteristics exhibit non-linear marginal contributions, though their magnitude and patterns differ notably. BM, for instance, shows minimal impact in lower deciles but increasingly affects portfolio weights in higher deciles, particularly under lower risk aversion. This pattern partially reflects the underlying Sharpe ratio profile of BM-sorted portfolios. In contrast, momentum variables (Mom12m, MomSeason, IntMom) display more modest marginal effects, suggesting

12

they serve a more supplementary role in the portfolio allocation decision.

Risk aversion systematically influences how the DPPP incorporates characteristic information. Higher risk aversion levels generally lead to more muted marginal effects, as evidenced by the flatter curves for $\gamma = 10$ and $\gamma = 20$. This finding suggests that as risk aversion increases, the model takes a more conservative approach to characteristic-based tilts, consistent with theoretical expectations about risk-return tradeoffs. It also confirms the reasoning that increasing risk aversion leads to a decrease in model complexity.

The relationship between Sharpe ratios of characteristic-sorted portfolios and marginal contribution patterns varies across characteristics. While some characteristics like STreversal show strong alignment between Sharpe ratios and marginal contributions, others exhibit more complex relationships. This suggests that the DPPP captures both direct risk-return relationships and potentially more sophisticated interactions between characteristics in its portfolio allocation decisions.

Surrogate model

We evaluate the extent to which non-linearity contributes to the estimated DPPP. Put differently, we assess the extent to which different forms of non-linearity play a role when optimizing portfolios conditional on firm characteristics. To do so, we estimate a sequence of increasingly complex surrogate models. First, we regress the out-of-sample weight predictions from the DPPP on all firm characteristics in a linear model. This allows us to assess the extent to which simple linear relationships explain the predicted weights. In a next step, we estimate a second surrogate model that includes all possible two-way interactions between variables. The incremental explanatory power of this model captures the importance of variable interactions in the DPPP's portfolio allocation decisions. We attribute the remaining unexplained portion of predicted DPPP weights to higher-order non-linearities in the functional form. To assess the economic significance of these non-linearities, we additionally examine the certainty equivalent differences of the ex-post fitted surrogate models compared to the actual model during the out-of-sample periods.

Figure S.3.3 shows both the adjusted R^2s of linear surrogate models for the out-of-sample predicted weights and the resulting differences in certainty equivalents across different levels of risk aversion. For each risk aversion level, we estimate two surrogate models: a simple linear
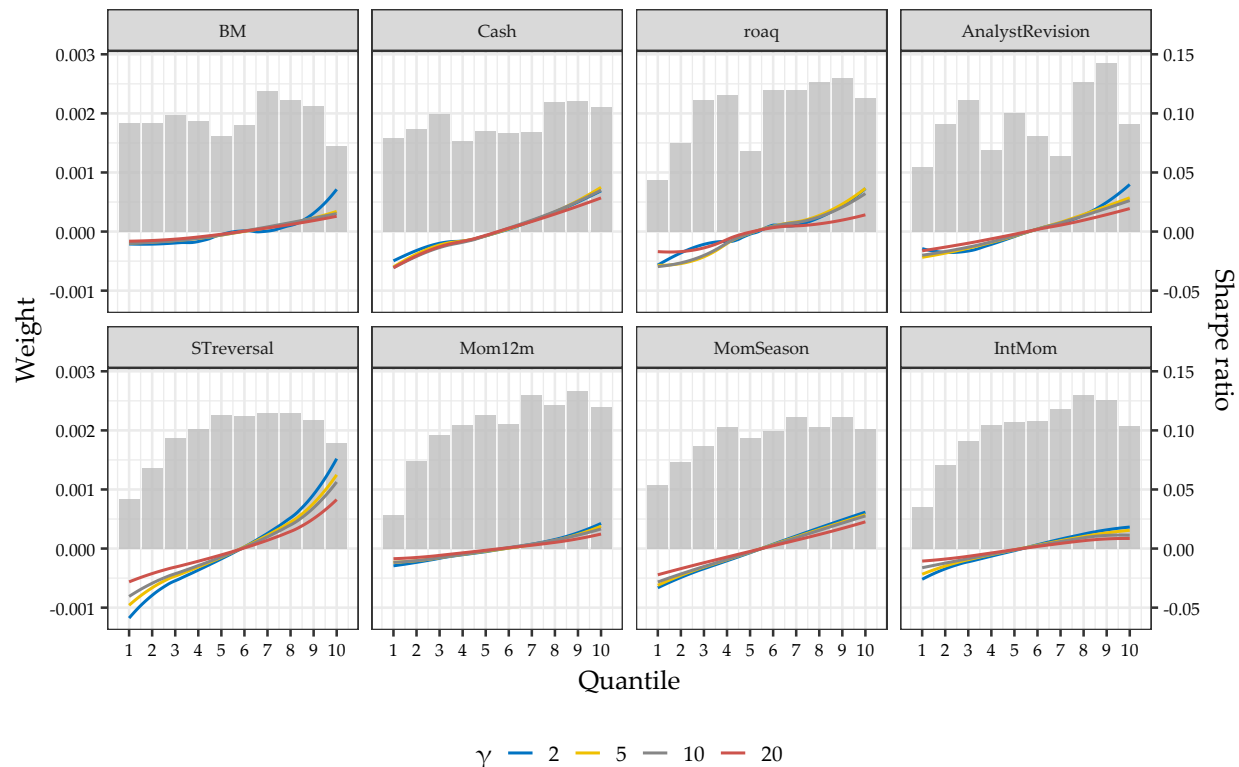
Figure S.3.2: Marginal contribution of characteristics to portfolio weights in the DPPP
This figure shows the sensitivity of predicted weights (left vertical axis) with respect to values of the respective variable (horizontal axis) across different risk aversions γ. The marginal contribution is computed as the difference between predictions with all features and predictions where the respective characteristic is set to zero. The aforementioned relationship is depicted by curves, fitted via local polynomial regressions. The figure also includes bars, depicting the Sharpe ratio (right vertical axis), per variable decile (horizontal axis).

model and an extended version that includes all possible two-way interactions between the 50 most important characteristics.

The results strongly support our theoretical findings that higher risk aversion reduces model complexity and acts as an economic regularization parameter, i.e., that the importance of non-linearity varies with the degree of risk aversion. The simple linear surrogate model explains about 30-40% of the variation in predicted portfolio weights for $\gamma = 2$, while the R^2 ranges between 40-60% for higher degrees of risk aversion. This underscores that risk aversion acts as an economic regularization parameter, reducing model complexity. Adding interactions substantially improves the model fit. The R^2 increases by approximately 20-30 percentage points across all degrees of risk aversion when including two-way interactions. Moreover, we observe that models with interactions show more stable explanatory power over time, as evidenced by less fluctuation in R^2

across periods.

The economic significance of these non-linearities is assessed through the certainty equivalent differences between the DPPP and surrogate models. The right panel of Figure S.3.3 reveals that the economic impact of non-linearities is most pronounced for lower levels of risk aversion. For $\gamma = 2$ and $\gamma = 5$, we observe certainty equivalent differences of up to 200-300 basis points in some periods, indicating substantial economic value from the DPPP's non-linear portfolio allocation decisions. This effect diminishes with increasing risk aversion, as shown by the smaller and more stable certainty equivalent differences for $\gamma = 10$ and $\gamma = 20$. Adding interactions to the surrogate models generally reduces these certainty equivalent differences, suggesting that a significant portion of the DPPP's economic value comes from capturing interaction effects between characteristics.

Based on these findings, we can decompose the DPPP's portfolio allocation decisions as follows: approximately 30-60% of the underlying characteristic-weight relationship is linear in nature, depending on the degree of risk aversion. An additional 20-30% can be captured by two-way interactions, while the remaining 10-50% can be attributed to higher-order non-linearities in the DPPP model. The economic significance of these non-linear components, as measured by certainty equivalent differences, is most pronounced for lower risk aversion levels and during periods of market stress, suggesting that the DPPP's flexibility in capturing complex relationships becomes particularly valuable under these conditions. This empirical evidence strongly supports our theoretical framework showing how risk aversion serves as an economic regularization mechanism that naturally constrains model complexity.

S.3.2 Rolling window estimation

For additional robustness, we consider a rolling-window estimation procedure of fixed length 20 years (240 months). At each date t, we estimate our model parameters using the most recent 240 months of data, keeping five years (60 months) of validation data and then form one-period-ahead forecasts of portfolio weights. Compared to the baseline expanding-window approach, this rolling scheme is designed to adapt more readily to potential structural changes in the data by discarding older observations.

15
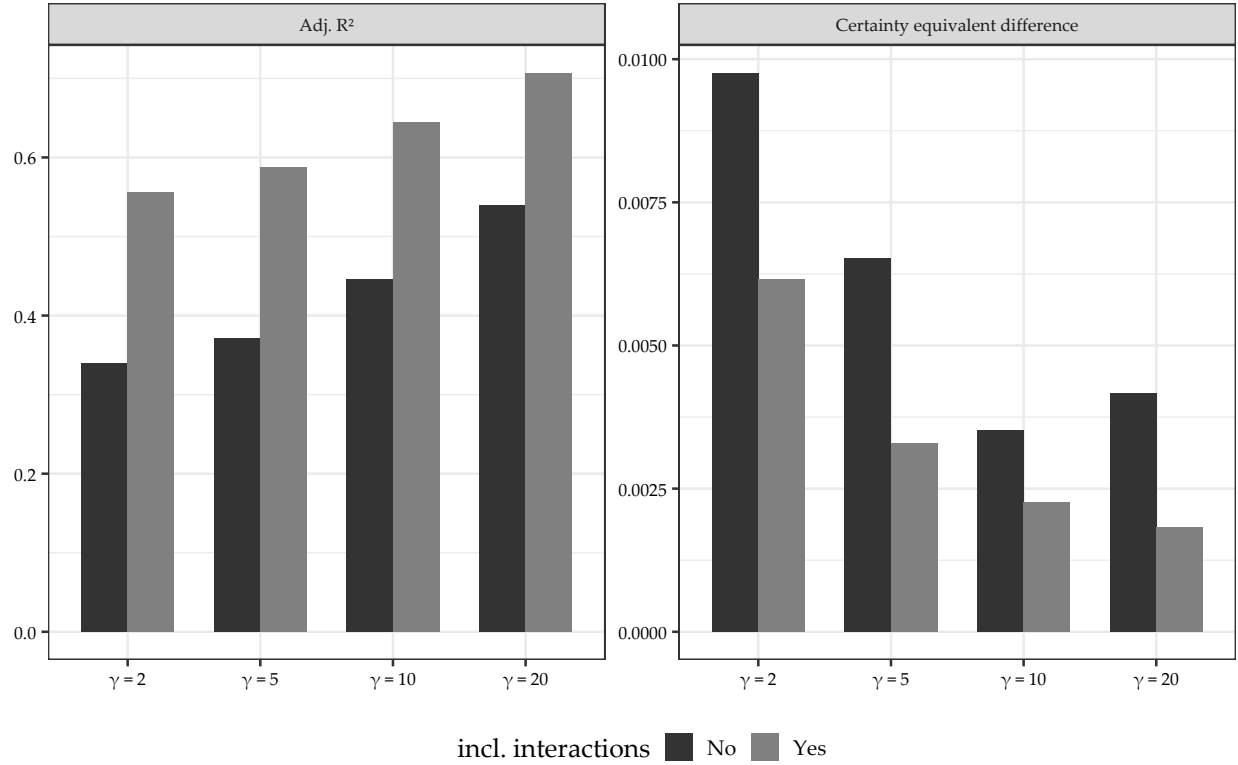
Figure S.3.3: Surrogate R^2 and difference in certainty equivalent in the DPPP
This figure depicts two key metrics evaluating the complexity and economic significance of the DPPP's portfolio allocation decisions. The left panel shows the adjusted R^2 of linear surrogate models fitted to the DPPP's weight predictions and the right panel displays the certainty equivalent differences between the DPPP and the surrogate models. More specifically, the bars show the R^2 and certainty equivalent differences for a linear surrogate model of the estimated weights by the deep models on the 50 most important variables in each model for all out-of-sample periods and across different risk aversion levels. Interactions include all possible two-way interactions between the variables.

Table S.3.1 presents out-of-sample results for the Deep Parametric Portfolio Policy (DPPP) under both rolling- and expanding-window estimation, using the same 157 firm-level characteristics. The table reports certainty equivalent returns for investors with CRRA preference coefficients $\gamma \in \{2, 5, 10, 20\}$. While the rolling-window approach can be more responsive to recent market conditions, our results indicate that it does not substantially outperform the expanding-window method in terms of certainty equivalent return. Moreover, for the most risk-averse investors ($\gamma = 20$), the baseline expanding-window approach achieves statistically higher certainty equivalent return compared to rolling-window estimation. We hypothesize that this pattern arises because the rolling procedure reduces the effective sample size, thereby increasing estimation uncertainty precisely in those regimes where the portfolio is most sensitive to parameter estimates.

Other performance metrics exhibit similar behavior. For instance, Sharpe ratios under rolling estimation are generally on par with those from the baseline method, showing little evidence of consistent improvement. Maximum drawdown and conditional value at risk measures also do not favor one approach conclusively, reinforcing the notion that the benefits of a rolling scheme can be offset by the increased estimation variability.

In short, while a rolling-window procedure may be appealing in settings prone to regime shifts, our empirical findings suggest that its advantages over the expanding-window baseline are limited in this particular application. 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.

S.3.3 Adding macroeconomic variables

To investigate how our portfolio policies interact with the state of the economy, we augment our models with macroeconomic variables. Specifically, we construct eight macro predictors based on Welch and Goyal (2008): the dividend-price ratio (dp), earnings-price ratio (ep), book-to-market ratio (bm), net equity expansion (ntis), the Treasury-bill rate (tbl), the term spread (tms), the default spread (dfy), and stock variance (svar).

Following Gu et al. (2020), we introduce a transformation layer in our network that generates a new set of interaction variables by multiplying stock-level characteristics with these macro predictors. Formally, for each stock i at time t,

$$z_{i,t} = x_{i,t} \otimes c_t, \tag{S.3.1}$$

where $x_{i,t}$ is a $P_c \times 1$ matrix of characteristics for each stock i, and c_t is a $P_x \times 1$ vector of macroeconomic predictors (including a constant). Hence, $z_{i,t}$ is a $P \times 1$ ($P = P_c P_x$) that captures the interaction between stock-level characteristics and macro-level factors. This yields a total of $157 \times (8 + 1) = 1,413$ covariates in our model.

Table S.3.2 compares the results of our baseline DPPP model with those of the augmented model that incorporates the macro variables. The difference in certainty equivalent returns is statistically insignificant for common levels of risk aversion ($\gamma \in \{2, 5, 10\}$). However, for

Table S.3.1: Deep portfolio policy for CRRA investors with different degrees of risk aversion with expanding vs. rolling window estimation

	$\gamma = 2$		$\gamma = 5$		$\gamma = 10$		$\gamma = 20$	
	Expanding	Rolling	Expanding	Rolling	Expanding	Rolling	Expanding	Rolling
CE	0.0297	0.0285	0.0232	0.0209	0.0152	0.0163	0.0040	0.0002
p-value($CE_{Expanding} - CE_{Rolling}$)		0.2759		0.1405		0.2230		0.0071
$\sum |w_i| / N_t * 100$	0.1907	0.1914	0.1938	0.1930	0.1933	0.1907	0.1729	0.1420
$max\ w_i * 100$	1.1483	1.0082	0.9843	0.7808	0.8305	0.6637	0.4582	0.3675
$min\ w_i * 100$	-1.2824	-1.1808	-1.2053	-1.1702	-0.9743	-0.8930	-0.7224	-0.4995
$\sum w_i I(w_i < 0)$	-0.8748	-0.8795	-0.8974	-0.8912	-0.8932	-0.8751	-0.7464	-0.5234
$\sum I(w_i < 0)/N_t$	0.3400	0.3247	0.3368	0.3160	0.3319	0.3340	0.3202	0.3095
$\sum |w_{i,t} - w_{i,t-1}^+|$	2.6342	2.4667	2.6022	2.3367	2.3813	2.1576	1.7516	1.2103
Mean	0.0341	0.0327	0.0305	0.0275	0.0281	0.0268	0.0224	0.0181
StdDev	0.0710	0.0655	0.0550	0.0507	0.0475	0.0432	0.0378	0.0350
Skew	2.6646	0.3897	0.8411	-0.2837	-0.2470	-0.5179	-0.5201	-0.8981
Kurt	26.4755	5.5252	10.9695	2.7665	4.0705	2.3334	1.9954	2.6755
Max DD	0.4979	0.5745	0.5601	0.5388	0.4662	0.5192	0.3027	0.3812
Max 1M loss	0.2264	0.2753	0.1789	0.2087	0.1838	0.1779	0.1446	0.1623
CVaR (95%)	0.1107	0.1157	0.0978	0.0911	0.0882	0.0815	0.0713	0.0717
SR	1.6607	1.7277	1.9230	1.8819	2.0446	2.1501	2.0491	1.7856
p-value($SR_{Expanding} - SR_{Rolling}$)		0.2777		0.4351		0.1309		0.0018
$FF5 + Mom\ \alpha$	0.0232	0.0199	0.0205	0.0159	0.0182	0.0165	0.0130	0.0089
$StdErr(\alpha)$	0.0029	0.0024	0.0024	0.0021	0.0020	0.0017	0.0016	0.0012

This table presents out-of-sample performance estimates for deep portfolio policies using 157 firm characteristics, as specified in Equation 1. The results show the DPPP for two different forecasting methods, namely expanding window (baseline) and rolling window. The analysis employs a feed-forward neural network model and data from the Open Source Asset Pricing Dataset spanning January 1971 to December 2020. Results are shown for Constant Relative Risk Aversion (CRRA) investors with relative risk aversion coefficients (γ) 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 Deep Parametric Portfolio Policy (DPPP) and 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 DPPP and 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 highest risk aversion ($\gamma = 20$), the baseline model outperforms the augmented model at a statistically significant level. We attribute this result to overfitting: the large number of parameters introduced by the macro interaction terms can degrade performance in high-risk-aversion scenarios, where a simpler model better aligns with our economic regularization theory. These findings are corroborated by the Sharpe ratio comparisons.

All other performance metrics exhibit a similar pattern, suggesting that macro variables do not substantially enhance our model's predictive power. Nevertheless, we further examine the mean absolute gradient of each macro predictor in Figure S.3.4. We observe that the importance of net equity expansion (ntis) increases monotonically with higher levels of risk aversion, while the other macro variables become comparatively less influential. This is notably different from the behavior of firm-level characteristics, whose relative importance tends to be more evenly distributed as risk aversion increases.
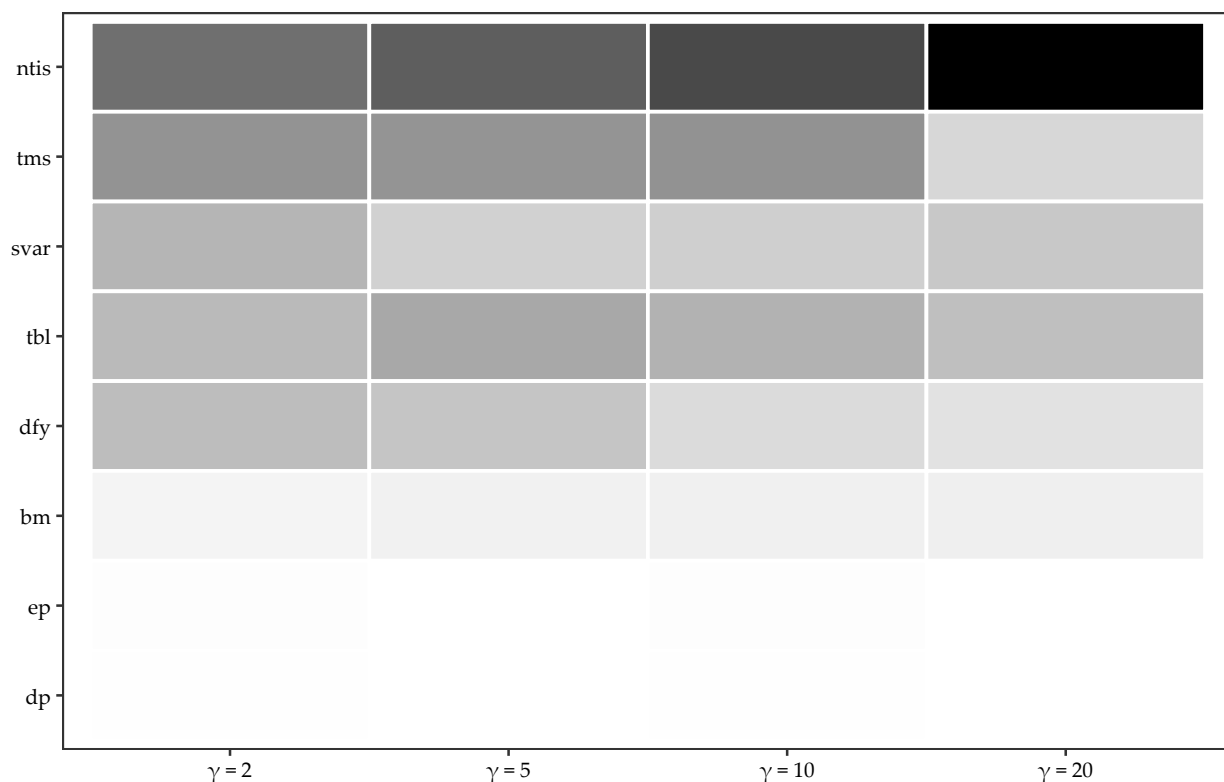
Figure S.3.4: Variable importance for the CRRA including macro variables for the DPPP
Variable importance for the eight macro variables in the DPPP across risk aversions. Variable importance is computed as the average absolute gradient over all training samples and normalized to sum to one within each 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 $\gamma = 2$.

Table S.3.2: Deep portfolio policy for CRRA investors with different degrees of risk aversion including macro variables

	$\gamma = 2$		$\gamma = 5$		$\gamma = 10$		$\gamma = 20$	
	Base	+Macro	Base	+Macro	Base	+Macro	Base	+Macro
CE	0.0297	0.0273	0.0232	0.0210	0.0152	0.0156	0.0040	-0.0022
p-value($CE_{Base} - CE_{+Macro}$)		0.1058		0.1022		0.3727		0.0001
$\sum |w_i| / N_t * 100$	0.1907	0.1880	0.1938	0.1864	0.1933	0.1863	0.1729	0.1500
$max\ w_i * 100$	1.1483	1.2595	0.9843	1.0513	0.8305	0.8609	0.4582	0.4140
$min\ w_i * 100$	-1.2824	-1.2296	-1.2053	-1.0027	-0.9743	-0.9916	-0.7224	-0.6849
$\sum w_i I(w_i < 0)$	-0.8748	-0.8555	-0.8974	-0.8436	-0.8932	-0.8433	-0.7464	-0.5814
$\sum I(w_i < 0) / N_t$	0.3400	0.3447	0.3368	0.3464	0.3319	0.3304	0.3202	0.2854
$\sum |w_{i,t} - w_{i,t-1}^+|$	2.6342	2.6562	2.6022	2.4281	2.3813	2.3192	1.7516	1.5513
Mean	0.0341	0.0308	0.0305	0.0290	0.0281	0.0271	0.0224	0.0198
StdDev	0.0710	0.0612	0.0550	0.0567	0.0475	0.0464	0.0378	0.0401
Skew	2.6646	1.0584	0.8411	0.8840	-0.2470	-0.0614	-0.5201	-0.4935
Kurt	26.4755	8.7724	10.9695	13.8198	4.0705	2.9006	1.9954	2.7715
Max DD	0.4979	0.5091	0.5601	0.6141	0.4662	0.4275	0.3027	0.4104
Max 1M loss	0.2264	0.2101	0.1789	0.2460	0.1838	0.1686	0.1446	0.1577
CVaR (95%)	0.1107	0.1045	0.0978	0.1052	0.0882	0.0826	0.0713	0.0814
SR	1.6607	1.7409	1.9230	1.7712	2.0446	2.0208	2.0491	1.7116
p-value($SR_{Base} - SR_{+Macro}$)		0.1901		0.0801		0.4413		0.0016
$FF5 + Mom\ \alpha$	0.0232	0.0221	0.0205	0.0178	0.0182	0.0166	0.0130	0.0102
$StdErr(\alpha)$	0.0029	0.0025	0.0024	0.0022	0.0020	0.0018	0.0016	0.0014

This table presents out-of-sample performance estimates for deep portfolio policies using 157 firm characteristics and eight macro variables, as well as their interactions, as specified in Equation 1. The analysis employs a feed-forward neural network model and data from the Open Source Asset Pricing Dataset spanning January 1971 to December 2020. Results are shown for Constant Relative Risk Aversion (CRRA) investors with relative risk aversion coefficients (γ) 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 Deep Parametric Portfolio Policy (DPPP) and 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 DPPP and 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.

Welch and Goyal (2008) show that net equity expansion acts as a market-timing signal that benefits investors with moderate risk aversion ($\gamma \approx 3$) but results in negative certainty equivalent returns for those with higher risk aversion. In our portfolio optimization framework, the increased importance of net equity expansion for higher risk aversion can be interpreted as the model relying more on net equity expansion to adjust portfolio weights in order to mitigate the risks and trading costs associated with excessive leverage.

Appendix S.4   Supplementary figures
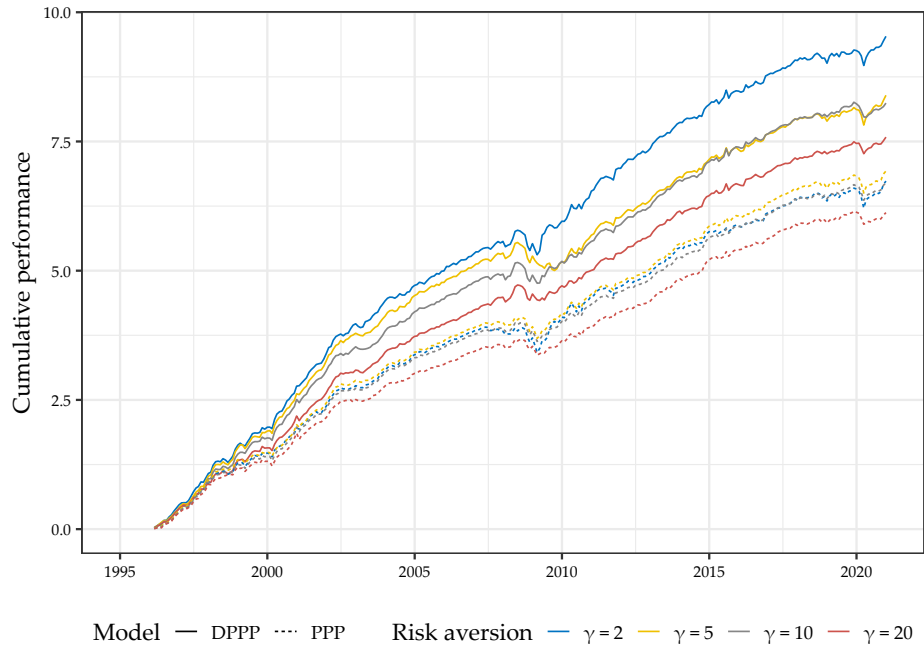
(a) Mean-variance utility
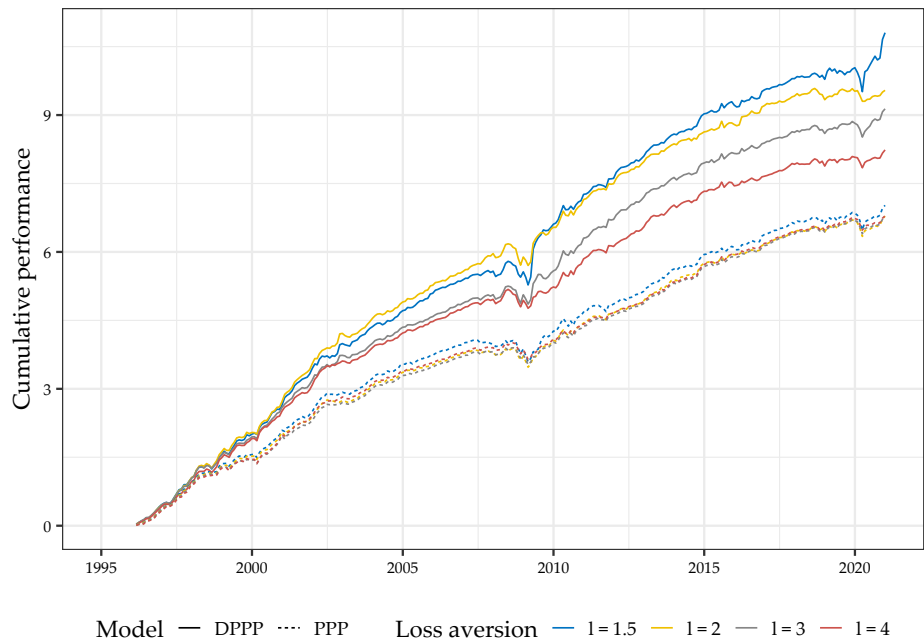
(b) Loss aversion

Figure S.4.1: Cumulative performance over time for MV and LA preferences
Panel (a) shows the cumulative sum of portfolio returns for the DPPP and PPP of investors with mean-variance preferences. Panel (b) shows the cumulative sum of portfolio returns net of the DPPP and PPP of investors with loss-aversion preferences. We show the results for each of the degrees of absolute risk aversion and loss aversion considered and across all out-of-sample periods.

Appendix S.5 Supplementary tables

Table S.5.1: Characteristic documentation

Acronym	Long Description	Author(s)	Year, Journal	Frequency	Cat.Data	Cat.Economic
ChInvIA	Change in capital inv (ind adj)	Abarbanell and Bushee	1998, AR	yearly	Accounting	investment growth
GrSaleToGrInv	Sales growth over inventory growth	Abarbanell and Bushee	1998, AR	yearly	Accounting	sales growth
GrSaleToGrOverhead	Sales growth over overhead growth	Abarbanell and Bushee	1998, AR	yearly	Accounting	sales growth
IdioVolAHT	Idiosyncratic risk (AHT)	Ali, Hwang, and Trombley	2003, JFE	monthly	Price	volatility
EarningsConsistency	Earnings consistency	Alwathainani	2009, BAR	yearly	Accounting	earnings
Illiquidity	Amihud's illiquidity	Amihud	2002, JFM	monthly	Trading	liquidity
BidAskSpread	Bid-ask spread	Amihud and Mendelsohn	1986, JFE	monthly	Trading	liquidity
grcapx	Change in capex (two years)	Anderson and Garcia-Feijoo	2006, JF	yearly	Accounting	investment growth
grcapx3y	Change in capex (three years)	Anderson and Garcia-Feijoo	2006, JF	yearly	Accounting	investment growth
betaVIX	Systematic volatility	Ang et al.	2006, JF	monthly	Price	volatility
IdioRisk	Idiosyncratic risk	Ang et al.	2006, JF	monthly	Price	volatility
IdioVol3F	Idiosyncratic risk (3 factor)	Ang et al.	2006, JF	monthly	Price	volatility
CoskewACX	Coskewness using daily returns	Ang, Chen and Xing	2006, RFS	monthly	Price	risk
Mom6mJunk	Junk Stock Momentum	Avramov et al	2007, JF	monthly	Price	momentum
OrderBacklogChg	Change in order backlog	Baik and Ahn	2007, Other	yearly	Accounting	accruals
roaq	Return on assets (qtrly)	Balakrishnan, Bartov and Faurel	2010, JAE	quarterly	Accounting	profitability
MaxRet	Maximum return over month	Bali, Cakici, and Whitelaw	2010, JF	monthly	Price	volatility
ReturnSkew	Return skewness	Bali, Engle and Murray	2015, Book	monthly	Price	risk
ReturnSkew3F	Idiosyncratic skewness (3F model)	Bali, Engle and Murray	2015, Book	monthly	Price	risk
CBOperProf	Cash-based operating profitability	Ball et al.	2016, JFE	yearly	Accounting	profitability
OperProfRD	Operating profitability R&D adjusted	Ball et al.	2016, JFE	yearly	Accounting	profitability
Size	Size	Banz	1981, JFE	monthly	Price	size
SP	Sales-to-price	Barbee, Mukherji and Raines	1996, FAJ	yearly	Accounting	valuation

Continued on next page

Table S.5.1: Characteristic documentation

Acronym	Long Description	Author(s)	Year, Journal	Frequency	Cat.Data	Cat.Economic
EP	Earnings-to-Price Ratio	Basu	1977, JF	monthly	Price	valuation
InvGrowth	Inventory Growth	Belo and Lin	2012, RFS	yearly	Accounting	profitability
BrandInvest	Brand capital investment	Belo, Lin and Vitorino	2014, RED	yearly	Accounting	investment
Leverage	Market leverage	Bhandari	1988, JFE	monthly	Price	leverage
ResidualMomentum	Momentum based on FF3 residuals	Blitz, Huij and Martens	2011, JEmpFin	monthly	Price	momentum
Price	Price	Blume and Husic	1972, JF	monthly	Price	other
NetPayoutYield	Net Payout Yield	Boudoukh et al.	2007, JF	monthly	Price	valuation
PayoutYield	Payout Yield	Boudoukh et al.	2007, JF	monthly	Price	valuation
NetDebtFinance	Net debt financing	Bradshaw, Richardson, Sloan	2006, JAE	yearly	Accounting	external financing
NetEquityFinance	Net equity financing	Bradshaw, Richardson, Sloan	2006, JAE	yearly	Accounting	external financing
XFIN	Net external financing	Bradshaw, Richardson, Sloan	2006, JAE	yearly	Accounting	external financing
DolVol	Past trading volume	Brennan, Chordia, Subra	1998, JFE	monthly	Trading	volume
FEPS	Analyst earnings per share	Cen, Wei, and Zhang	2006, WP	monthly	Analyst	profitability
AnnouncementReturn	Earnings announcement return	Chan, Jegadeesh and Lakonishok	1996, JF	monthly	Price	earnings
REV6	Earnings forecast revisions	Chan, Jegadeesh and Lakonishok	1996, JF	monthly	Analyst	earnings
AdExp	Advertising Expense	Chan, Lakonishok and Sougiannis	2001, JF	monthly	Accounting	R&D
RD	R&D over market cap	Chan, Lakonishok and Sougiannis	2001, JF	monthly	Accounting	R&D
CashProd	Cash Productivity	Chandrashekar and Rao	2009, WP	yearly	Accounting	profitability
std_turn	Share turnover volatility	Chordia, Subra, Anshuman	2001, JFE	monthly	Trading	liquidity
VolSD	Volume Variance	Chordia, Subra, Anshuman	2001, JFE	monthly	Trading	liquidity
retConglomerate	Conglomerate return	Cohen and Lou	2012, JFE	monthly	Price	delayed processing
RDAbility	R&D ability	Cohen, Diether and Malloy	2013, RFS	yearly	Accounting	other
AssetGrowth	Asset growth	Cooper, Gulen and Schill	2008, JF	yearly	Accounting	investment

Table S.5.1: Characteristic documentation

Acronym	Long Description	Author(s)	Year, Journal	Frequency	Cat.Data	Cat.Economic
EarningsForecastDisparity	Long-vs-short EPS forecasts	Da and Warachka	2011, JFE	monthly	Analyst	earnings
CompEquIss	Composite equity issuance	Daniel and Titman	2006, JF	monthly	Accounting	external financing
IntanBM	Intangible return using BM	Daniel and Titman	2006, JF	yearly	Accounting	long term reversal
IntanCFP	Intangible return using CFtoP	Daniel and Titman	2006, JF	yearly	Accounting	long term reversal
IntanEP	Intangible return using EP	Daniel and Titman	2006, JF	yearly	Accounting	long term reversal
IntanSP	Intangible return using Sale2P	Daniel and Titman	2006, JF	yearly	Accounting	long term reversal
ShareIss5Y	Share issuance (5 year)	Daniel and Titman	2006, JF	monthly	Accounting	external financing
LRreversal	Long-run reversal	De Bondt and Thaler	1985, JF	monthly	Price	long term reversal
MRreversal	Medium-run reversal	De Bondt and Thaler	1985, JF	monthly	Price	long term reversal
EquityDuration	Equity Duration	Dechow, Sloan and Soliman	2004, RAS	yearly	Price	valuation
cfp	Operating Cash flows to price	Desai, Rajgopal, Venkatachalam	2004, AR	yearly	Accounting	valuation
ForecastDispersion	EPS Forecast Dispersion	Diether, Malloy and Scherbina	2002, JF	monthly	Analyst	volatility
ExclExp	Excluded Expenses	Doyle, Lundholm and Soliman	2003, RAS	quarterly	Analyst	composite accounting
ProbInformedTrading	Probability of Informed Trading	Easley, Hvidkjaer and O'Hara	2002, JF	yearly	Trading	liquidity
OrgCap	Organizational capital	Eisfeldt and Papanikolaou	2013, JF	yearly	Accounting	R&D
sfe	Earnings Forecast to price	Elgers, Lo and Pfeiffer	2001, AR	monthly	Analyst	valuation
GrLTNOA	Growth in long term operating assets	Fairfield, Whisenant and Yohn	2003, AR	yearly	Accounting	investment
AM	Total assets to market	Fama and French	1992, JF	yearly	Accounting	valuation
BMdec	Book to market using December ME	Fama and French	1992, JPM	yearly	Accounting	valuation
BookLeverage	Book leverage (annual)	Fama and French	1992, JF	yearly	Accounting	leverage
OperProf	operating profits / book equity	Fama and French	2006, JFE	yearly	Accounting	profitability
Beta	CAPM beta	Fama and MacBeth	1973, JPE	monthly	Price	risk
EarningsSurprise	Earnings Surprise	Foster, Olsen and Shevlin	1984, AR	quarterly	Analyst	earnings

Continued on next page

Table S.5.1: Characteristic documentation

Acronym	Long Description	Author(s)	Year, Journal	Frequency	Cat.Data	Cat.Economic
AnalystValue	Analyst Value	Frankel and Lee	1998, JAE	monthly	Analyst	valuation
AOP	Analyst Optimism	Frankel and Lee	1998, JAE	monthly	Analyst	other
PredictedFE	Predicted Analyst forecast error	Frankel and Lee	1998, JAE	monthly	Accounting	earnings
FR	Pension Funding Status	Franzoni and Marin	2006, JF	monthly	Accounting	composite accounting
BetaFP	Frazzini-Pedersen Beta	Frazzini and Pedersen	2014, JFE	monthly	Price	other
High52	52 week high	George and Hwang	2004, JF	monthly	Price	momentum
IndMom	Industry Momentum	Grinblatt and Moskowitz	1999, JFE	monthly	Price	momentum
PctAcc	Percent Operating Accruals	Hafzalla, Lundholm, Van Winkle	2011, AR	yearly	Accounting	accruals
PctTotAcc	Percent Total Accruals	Hafzalla, Lundholm, Van Winkle	2011, AR	yearly	Accounting	accruals
tang	Tangibility	Hahn and Lee	2009, JF	yearly	Accounting	asset composition
Coskewness	Coskewness	Harvey and Siddique	2000, JF	monthly	Price	risk
RoE	net income / book equity	Haugen and Baker	1996, JFE	yearly	Accounting	profitability
VarCF	Cash-flow to price variance	Haugen and Baker	1996, JFE	monthly	Accounting	cash flow risk
VolMkt	Volume to market equity	Haugen and Baker	1996, JFE	monthly	Trading	volume
VolumeTrend	Volume Trend	Haugen and Baker	1996, JFE	monthly	Trading	volume
AnalystRevision	EPS forecast revision	Hawkins, Chamberlin, Daniel	1984, FAJ	monthly	Analyst	earnings
Mom12mOffSeason	Momentum without the seasonal part	Heston and Sadka	2008, JFE	monthly	Price	momentum
MomOffSeason	Off season long-term reversal	Heston and Sadka	2008, JFE	monthly	Price	momentum
MomOffSeason06YrPlus	Off season reversal years 6 to 10	Heston and Sadka	2008, JFE	monthly	Price	momentum
MomOffSeason11YrPlus	Off season reversal years 11 to 15	Heston and Sadka	2008, JFE	monthly	Price	momentum
MomOffSeason16YrPlus	Off season reversal years 16 to 20	Heston and Sadka	2008, JFE	monthly	Price	momentum
MomSeason	Return seasonality years 2 to 5	Heston and Sadka	2008, JFE	monthly	Price	momentum
MomSeason06YrPlus	Return seasonality years 6 to 10	Heston and Sadka	2008, JFE	monthly	Price	momentum

Continued on next page

Table S.5.1: Characteristic documentation

Acronym	Long Description	Author(s)	Year, Journal	Frequency	Cat.Data	Cat.Economic
MomSeason11YrPlus	Return seasonality years 11 to 15	Heston and Sadka	2008, JFE	monthly	Price	momentum
MomSeason16YrPlus	Return seasonality years 16 to 20	Heston and Sadka	2008, JFE	monthly	Price	momentum
MomSeasonShort	Return seasonality last year	Heston and Sadka	2008, JFE	monthly	Price	momentum
NOA	Net Operating Assets	Hirshleifer et al.	2004, JAE	yearly	Accounting	asset composition
dNoa	change in net operating assets	Hirshleifer, Hou, Teoh, Zhang	2004, JAE	yearly	Accounting	investment
EarnSupBig	Earnings surprise of big firms	Hou	2007, RFS	quarterly	Accounting	delayed processing
IndRetBig	Industry return of big firms	Hou	2007, RFS	monthly	Price	delayed processing
PriceDelayRsq	Price delay r square	Hou and Moskowitz	2005, RFS	monthly	Price	delayed processing
PriceDelaySlope	Price delay coeff	Hou and Moskowitz	2005, RFS	monthly	Price	delayed processing
PriceDelayTstat	Price delay SE adjusted	Hou and Moskowitz	2005, RFS	monthly	Price	delayed processing
STreversal	Short term reversal	Jegadeesh	1989, JF	monthly	Price	short-term reversal
RevenueSurprise	Revenue Surprise	Jegadeesh and Livnat	2006, JFE	quarterly	Accounting	sales growth
Mom12m	Momentum (12 month)	Jegadeesh and Titman	1993, JF	monthly	Price	momentum
Mom6m	Momentum (6 month)	Jegadeesh and Titman	1993, JF	monthly	Price	momentum
ChangeInRecommendation	Change in recommendation	Jegadeesh et al.	2004, JF	monthly	Analyst	recommendation
OptionVolume1	Option to stock volume	Johnson and So	2012, JFE	monthly	Trading	volume
OptionVolume2	Option volume to average	Johnson and So	2012, JFE	monthly	Trading	volume
BetaTailRisk	Tail risk beta	Kelly and Jiang	2014, RFS	monthly	Price	risk
fgr5yrLag	Long-term EPS forecast	La Porta	1996, JF	monthly	Analyst	earnings
CF	Cash flow to market	Lakonishok, Shleifer, Vishny	1994, JF	monthly	Accounting	valuation
MeanRankRevGrowth	Revenue Growth Rank	Lakonishok, Shleifer, Vishny	1994, JF	yearly	Accounting	sales growth
RDS	Real dirty surplus	Landsman et al.	2011, AR	yearly	Accounting	composite accounting
Tax	Taxable income to income	Lev and Nissim	2004, AR	yearly	Accounting	tax

Continued on next page

Table S.5.1: Characteristic documentation

Acronym	Long Description	Author(s)	Year, Journal	Frequency	Cat.Data	Cat.Economic
RDcap	R&D capital-to-assets	Li	2011, RFS	yearly	Accounting	asset composition
zerotrade	Days with zero trades	Liu	2006, JFE	monthly	Trading	liquidity
zerotradeAlt1	Days with zero trades	Liu	2006, JFE	monthly	Trading	liquidity
zerotradeAlt12	Days with zero trades	Liu	2006, JFE	monthly	Trading	liquidity
ChEQ	Growth in book equity	Lockwood and Prombutr	2010, JFR	yearly	Accounting	investment
EarningsStreak	Earnings surprise streak	Loh and Warachka	2012, MS	monthly	Accounting	earnings
NumEarnIncrease	Earnings streak length	Loh and Warachka	2012, MS	quarterly	Accounting	earnings
GrAdExp	Growth in advertising expenses	Lou	2014, RFS	yearly	Accounting	investment
EntMult	Enterprise Multiple	Loughran and Wellman	2011, JFQA	monthly	Accounting	valuation
CompositeDebtIssuance	Composite debt issuance	Lyandres, Sun and Zhang	2008, RFS	yearly	Accounting	external financing
InvestPPEInv	change in ppe and inv/assets	Lyandres, Sun and Zhang	2008, RFS	yearly	Accounting	investment
Frontier	Efficient frontier index	Nguyen and Swanson	2009, JFQA	yearly	Accounting	valuation
GP	gross profits / total assets	Novy-Marx	2013, JFE	yearly	Accounting	profitability
IntMom	Intermediate Momentum	Novy-Marx	2012, JFE	monthly	Price	momentum
OPLeverage	Operating leverage	Novy-Marx	2010, ROF	yearly	Accounting	other
Cash	Cash to assets	Palazzo	2012, JFE	quarterly	Accounting	asset composition
BetaLiquidityPS	Pastor-Stambaugh liquidity beta	Pastor and Stambaugh	2003, JPE	monthly	Price	liquidity
BPEBM	Leverage component of BM	Penman, Richardson and Tuna	2007, JAR	monthly	Accounting	leverage
EBM	Enterprise component of BM	Penman, Richardson and Tuna	2007, JAR	monthly	Accounting	valuation
NetDebtPrice	Net debt to price	Penman, Richardson and Tuna	2007, JAR	monthly	Accounting	leverage
PS	Piotroski F-score	Piotroski	2000, AR	yearly	Accounting	composite accounting
ShareIss1Y	Share issuance (1 year)	Pontiff and Woodgate	2008, JF	monthly	Accounting	external financing
DelDRC	Deferred Revenue	Prakash and Sinha	2012, CAR	yearly	Accounting	investment

Continued on next page

Table S.5.1: Characteristic documentation

Acronym	Long Description	Author(s)	Year, Journal	Frequency	Cat.Data	Cat.Economic
OrderBacklog	Order backlog	Rajgopal, Shevlin, Venkatachalam	2003, RAS	yearly	Accounting	sales growth
DelCOA	Change in current operating assets	Richardson et al.	2005, JAE	yearly	Accounting	investment
DelCOL	Change in current operating liabilities	Richardson et al.	2005, JAE	yearly	Accounting	external financing
DelEqu	Change in equity to assets	Richardson et al.	2005, JAE	yearly	Accounting	investment
DelFINL	Change in financial liabilities	Richardson et al.	2005, JAE	yearly	Accounting	external financing
DelLTI	Change in long-term investment	Richardson et al.	2005, JAE	yearly	Accounting	investment
DelNetFin	Change in net financial assets	Richardson et al.	2005, JAE	yearly	Accounting	investment
TotalAccruals	Total accruals	Richardson et al.	2005, JAE	yearly	Accounting	investment
BM	Book to market using most recent ME	Rosenberg, Reid, and Lanstein	1985, JF	monthly	Accounting	valuation
Accruals	Accruals	Sloan	1996, AR	yearly	Accounting	accruals
ChAssetTurnover	Change in Asset Turnover	Soliman	2008, AR	yearly	Accounting	sales growth
ChNNCOA	Change in Net Noncurrent Op Assets	Soliman	2008, AR	yearly	Accounting	investment
ChNWC	Change in Net Working Capital	Soliman	2008, AR	yearly	Accounting	investment
ChInv	Inventory Growth	Thomas and Zhang	2002, RAS	yearly	Accounting	investment
ChTax	Change in Taxes	Thomas and Zhang	2011, JAR	quarterly	Accounting	tax
Investment	Investment to revenue	Titman, Wei and Xie	2004, JFQA	yearly	Accounting	investment
realestate	Real estate holdings	Tuzel	2010, RFS	yearly	Accounting	asset composition
AbnormalAccruals	Abnormal Accruals	Xie	2001, AR	yearly	Accounting	accruals
FirmAgeMom	Firm Age - Momentum	Zhang	2004, JF	monthly	Price	momentum

The table shows all available characteristics used, the author(s), the year and the journal of publication. In addition, this table shows the update frequency, the data category as well as the economic category.

Table S.5.2: Transaction cost 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
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
$\sum |w_i|/N_t * 100$	0.1684	0.1925	0.1750	0.1938	0.1753	0.1914	0.1607	0.1631
$max\ w_i * 100$	0.6553	0.7983	0.6721	0.6514	0.6595	0.4899	0.5866	0.3813
$min\ w_i * 100$	-0.6305	-1.0150	-0.6774	-0.9952	-0.6812	-1.0398	-0.5735	-0.6183
$\sum w_i I(w_i < 0)$	-0.7139	-0.8877	-0.7612	-0.8973	-0.7638	-0.8798	-0.6588	-0.6756
$\sum I(w_i < 0)/N_t$	0.3267	0.3296	0.3417	0.3267	0.3440	0.3185	0.3319	0.3367
$\sum |w_{i,t} - w_{i,t-1}^+|$	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
Skew	-0.6222	-0.1700	-0.8459	0.0763	-0.9009	-0.5885	-0.8134	-0.7482
Kurt	3.1259	4.9340	2.4793	7.0330	2.4114	1.9342	1.8024	2.0122
Max DD	0.6062	0.7288	0.4937	0.5344	0.4224	0.5714	0.4020	0.3975
Max 1M loss	0.2228	0.2280	0.1812	0.2015	0.1559	0.1546	0.1303	0.1513
CVaR (95%)	0.1037	0.1164	0.0937	0.0981	0.0891	0.0873	0.0794	0.0727
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
$FF5 + Mom\ \alpha$	0.0065	0.0112	0.0071	0.0100	0.0069	0.0082	0.0056	0.0070
$StdErr(\alpha)$	0.0013	0.0021	0.0013	0.0019	0.0013	0.0017	0.0014	0.0014

This table presents out-of-sample performance estimates for deep portfolio policies with the transaction costs penalty from Equation (16) 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 Dataset spanning January 1971 to December 2020. Results are shown for Constant Relative Risk Aversion (CRRA) investors with relative risk aversion coefficients (γ) of 2, 5, 10, and 20. All results are reported net of transaction costs. 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 Deep Parametric Portfolio Policy (DPPP) and 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 DPPP and 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.

Table S.5.3: Long-only 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	0.0118	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
$\sum |w_i| / N_t * 100$	0.0694	0.0694	0.0694	0.0694	0.0694	0.0694	0.0694	0.0694
$max\ w_i * 100$	0.3543	1.9711	0.3761	1.9329	0.3588	1.2718	0.3608	0.7288
$min\ w_i * 100$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\sum w_i I(w_i < 0)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\sum I(w_i < 0) / N_t$	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
Skew	-0.4486	0.1148	-0.6717	0.3396	-0.5582	-0.6934	-0.9488	-0.9849
Kurt	3.3286	3.9959	3.0863	7.3953	3.6061	3.9808	3.0547	2.8460
Max DD	0.7942	0.7985	0.7257	0.8282	0.6778	0.6520	0.6011	0.5064
Max 1M loss	0.2483	0.2603	0.2171	0.2667	0.1968	0.2260	0.1832	0.1780
CVaR (95%)	0.1266	0.1472	0.1168	0.1295	0.1037	0.1093	0.0964	0.0898
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
$FF5 + Mom\ \alpha$	0.0040	0.0105	0.0045	0.0114	0.0042	0.0076	0.0023	0.0047
$StdErr(\alpha)$	0.0007	0.0017	0.0008	0.0016	0.0008	0.0010	0.0008	0.0009

This table presents out-of-sample performance estimates for deep portfolio policies including a long-only constraint 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 Dataset spanning January 1971 to December 2020. Results are shown for Constant Relative Risk Aversion (CRRA) investors with relative risk aversion coefficients (γ) 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 Deep Parametric Portfolio Policy (DPPP) and 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 DPPP and 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.

Table S.5.4: Long-only & transaction cost 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
CE	0.0101	0.0128	0.0063	0.0067	0.0003	0.0015	-0.0157	-0.0087
p-value($CE_{DPPP} - CE_{PPP}$)		0.0079		0.4253		0.0245		0.0001
$\sum |w_i| / N_t * 100$	0.0694	0.0694	0.0694	0.0694	0.0694	0.0694	0.0694	0.0694
$max\ w_i * 100$	0.2776	1.7957	0.3111	1.6695	0.3280	1.1258	0.3575	0.5347
$min\ w_i * 100$	0.0013	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\sum w_i I(w_i < 0)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\sum I(w_i < 0) / N_t$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\sum |w_{i,t} - w_{i,t-1}^+|$	0.3104	1.1833	0.3388	1.1537	0.2910	0.8621	0.2509	0.4959
Mean	0.0130	0.0172	0.0127	0.0153	0.0120	0.0137	0.0106	0.0113
StdDev	0.0536	0.0658	0.0481	0.0559	0.0439	0.0457	0.0401	0.0363
Skew	-0.5117	0.1447	-0.6868	-0.5080	-0.6450	-0.3657	-0.9713	-1.0731
Kurt	3.7273	5.0898	3.6745	4.2885	3.7214	4.1127	3.0701	2.8108
Max DD	0.7754	0.8689	0.7044	0.8739	0.6672	0.6905	0.6061	0.5223
Max 1M loss	0.2419	0.2587	0.2192	0.2513	0.1899	0.1843	0.1808	0.1617
CVaR (95%)	0.1229	0.1383	0.1119	0.1255	0.1019	0.1020	0.0960	0.0858
SR	0.8413	0.9042	0.9130	0.9496	0.9432	1.0342	0.9152	1.0760
p-value($SR_{DPPP} - SR_{PPP}$)		0.1453		0.2827		0.0238		0.0001
$FF5 + Mom\ \alpha$	0.0019	0.0059	0.0023	0.0045	0.0022	0.0041	0.0014	0.0024
$StdErr(\alpha)$	0.0007	0.0014	0.0007	0.0012	0.0007	0.0010	0.0008	0.0009

This table presents out-of-sample performance estimates for deep portfolio policies with the transaction costs penalty from Equation (16) and including a long-only constraint 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 Dataset spanning January 1971 to December 2020. Results are shown for Constant Relative Risk Aversion (CRRA) investors with relative risk aversion coefficients (γ) of 2, 5, 10, and 20. All results are reported net of transaction costs. 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 Deep Parametric Portfolio Policy (DPPP) and 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 DPPP and 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.

Table S.5.5: Deep portfolio policy for mean-variance investors with different degrees of risk aversion

	$\gamma = 2$		$\gamma = 5$		$\gamma = 10$		$\gamma = 20$	
	PPP	DPPP	PPP	DPPP	PPP	DPPP	PPP	DPPP
CE	0.0201	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
$\sum |w_i|/N_t * 100$	0.1748	0.1926	0.1811	0.1952	0.1816	0.1928	0.1786	0.1932
$max\ w_i * 100$	0.7115	1.0449	0.7675	0.9073	0.7761	0.7741	0.7586	0.6787
$min\ w_i * 100$	-0.6847	-1.3109	-0.7234	-1.2246	-0.7227	-1.1856	-0.6987	-1.0437
$\sum w_i I(w_i < 0)$	-0.7602	-0.8882	-0.8059	-0.9070	-0.8093	-0.8899	-0.7879	-0.8925
$\sum I(w_i < 0)/N_t$	0.3488	0.3250	0.3564	0.3193	0.3560	0.3227	0.3517	0.3381
$\sum |w_{i,t} - w_{i,t-1}^+|$	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
Kurt	2.8505	3.3104	2.5837	2.3437	2.1502	2.1743	1.9302	2.2886
Max DD	0.5478	0.5284	0.4404	0.5957	0.3857	0.4392	0.3219	0.3290
Max 1M loss	0.2151	0.2134	0.1867	0.1980	0.1551	0.1837	0.1235	0.1620
CVaR (95%)	0.1039	0.1035	0.0921	0.0979	0.0871	0.0833	0.0801	0.0761
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
$FF5 + Mom\ \alpha$	0.0108	0.0198	0.0122	0.0158	0.0121	0.0171	0.0108	0.0151
$StdErr(\alpha)$	0.0013	0.0022	0.0014	0.0020	0.0014	0.0020	0.0014	0.0017

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 Dataset spanning January 1971 to December 2020. Results are shown for mean-variance investors with relative risk aversion coefficients (γ) 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 Deep Parametric Portfolio Policy (DPPP) and 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 DPPP and 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.

Table S.5.6: Deep portfolio policy for loss-averse investors with different degrees of loss aversion

	$l = 1.5$		$l = 2$		$l = 3$		$l = 4$	
	PPP	DPPP	PPP	DPPP	PPP	DPPP	PPP	DPPP
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
$\sum |w_i| / N_t * 100$	0.1793	0.1931	0.1801	0.1919	0.1816	0.1917	0.1815	0.1901
$max\ w_i * 100$	0.7510	1.0462	0.7490	1.1024	0.7652	1.0074	0.7625	0.9322
$min\ w_i * 100$	-0.7062	-1.3205	-0.7093	-1.2292	-0.7215	-1.1298	-0.7172	-1.1071
$\sum w_i I(w_i < 0)$	-0.7929	-0.8918	-0.7980	-0.8833	-0.8090	-0.8823	-0.8083	-0.8702
$\sum I(w_i < 0) / N_t$	0.3537	0.3219	0.3522	0.3471	0.3559	0.3353	0.3566	0.3326
$\sum |w_{i,t} - w_{i,t-1}^+|$	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
Kurt	2.7196	15.8393	2.6054	6.7495	2.1086	4.8422	1.9761	1.6728
Max DD	0.5395	0.5974	0.4395	0.5395	0.4019	0.4461	0.3955	0.4560
Max 1M loss	0.2115	0.2903	0.1901	0.2144	0.1588	0.1786	0.1358	0.1604
CVaR (95%)	0.1032	0.1183	0.0920	0.1039	0.0874	0.0971	0.0828	0.0926
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
$FF5 + Mom\ \alpha$	0.0116	0.0235	0.0117	0.0223	0.0122	0.0198	0.0131	0.0167
$StdErr(\alpha)$	0.0014	0.0029	0.0014	0.0026	0.0014	0.0023	0.0014	0.0019

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 Dataset spanning January 1971 to December 2020. Results are shown for loss-averse investors with loss aversions coefficients (l) of 1.5, 2, 3, and 4. 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 Deep Parametric Portfolio Policy (DPPP) and 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 DPPP and 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.

centre for financial research
cfr/university of cologne
Albertus-Magnus-Platz
D-50923 cologne
fon +49(0)221-470-6995
fax +49(0)221-470-3992
kempf@cfr-cologne.de
www.cfr-cologne.de