

Semiparametric Estimation of Probability Weighting Functions Implicit in Option Prices*

H. Peter Boswijk

Amsterdam School of Economics, University of Amsterdam, and Tinbergen Institute

Jeroen Dalderop

Department of Economics, University of Notre Dame

Roger J. A. Laeven

Amsterdam School of Economics, University of Amsterdam, EURANDOM, and CentER

and

Niels Marijnen

Amsterdam School of Economics, University of Amsterdam, and Tinbergen Institute

September 23, 2025

Abstract

This paper develops a semiparametric estimation method that jointly identifies the probability weighting and utility functions implicit in option prices. Our econometric method obtains objective return distributions by transforming the options' implied risk-neutral distributions according to the posited rank-dependent utility model. We nonparametrically estimate the probability weighting function using the kernel density of suitable utility-adjusted probability integral transforms. The parameters of the utility function are estimated by maximizing the resulting profile likelihood. We establish the asymptotic properties of our estimation procedure, and demonstrate its good finite sample performance in Monte Carlo simulations. Empirical results based on S&P 500 index option prices and returns over the period 1996–2023 reveal the relevance of probability weighting, in particular at the monthly horizon where the weighting function is inverse S-shaped. These results are robust to various specifications of the utility function.

Keywords: Rank-Dependent Utility; Stochastic Discount Factor; Profile Likelihood; Kernel Estimation.

*We are very grateful to Yacine Aït-Sahalia, Carsten Chong, Glenn Harrison, Oliver Linton, Olivier Scaillet, as well as seminar and conference participants at the Financial Econometrics Conference at Cambridge University, the 2024 Netherlands Econometric Study Group, the 2024 QFFE conference at Aix-Marseille Université, the 2024 Bernoulli-IMS World Congress in Probability and Statistics at Ruhr University Bochum, the 2024 EEA-ESEM at Erasmus University Rotterdam, the 2024 NBER Time Series conference at the University of Pennsylvania, the 2024 ES Meeting at Palma de Mallorca, the 2025 World Congress of the Econometric Society in Seoul, the Tinbergen Institute, and the University of Amsterdam for comments and suggestions. This research was funded in part by the Netherlands Organization for Scientific Research under grant NWO Vici 2020–2027 (Laeven). *Corresponding author:* Jeroen Dalderop, jdalderop@nd.edu, Department of Economics, University of Notre Dame, Notre Dame, USA.

1 Introduction

Probability weighting is at the heart of the leading non-expected utility models for describing and understanding risky choices. By distinguishing attitudes toward wealth, as captured by the utility function, and attitudes toward probabilities, as represented by the probability weighting function, the decision-theoretic models developed by [Quiggin \(1982\)](#) and [Tversky and Kahneman \(1992\)](#) can explain a wide variety of traits in economics and finance. Key examples include consumption-savings decisions, portfolio choices, gambling behavior and insurance decisions ([Eeckhoudt et al., 2005](#); [Barberis, 2013](#)).

Existing evidence of probability weighting comes primarily from experimental studies, in which subjects choose from, predominantly fictitious, sets of lotteries designed and controlled by the researcher. Meanwhile, the scarcer real-world evidence of probability weighting is often based on market prices of contingent claims, such as betting odds ([Jullien and Salanié, 2000](#)) or financial option prices ([Kliger and Levy, 2009](#); [Polkovnichenko and Zhao, 2013](#)). Option markets, in particular, provide detailed information on investors' attitudes toward risk by trading large numbers of contracts for a wide range of payoff thresholds. However, unlike in the laboratory, their objective, or physical, payoff probabilities at any point in time are not known by the researcher. Market-based measures of probability weighting therefore require a model for the physical conditional distributions of the underlying asset price. Their estimation based on historical returns necessitates imposing a distributional form and/or listing the set of conditioning variables available to investors. The latter is particularly challenging, as forward-looking investors may gather relevant information beyond the historical data available to the econometrician.¹

This paper develops an estimation procedure that obtains physical return distributions from option-implied risk-neutral distributions by transformations that follow from the

¹In models without probability weighting, missing out on conditioning variables has been attributed to cause the so-called “pricing kernel puzzle” ([Chabi-Yo et al., 2008](#); [Song and Xiu, 2016](#); [Linn et al., 2018](#)).

asserted rank-dependent utility model. These transformations specify the conditional pricing kernel or stochastic discount factor used by investors, but do not restrict the physical conditional distributions given their information set, delineating only a relevant subset of econometric relations. The marginal utility and probability weighting functions are then estimated by maximizing the conditional likelihood of returns. As economic theory puts few restrictions on the probability weighting function, we estimate it nonparametrically. In particular, we consider a kernel density estimator based on the probability integral transforms (PITs) of the risk-neutral distributions after adjusting for a given utility function. With parametric models of the utility function, we obtain a tractable semiparametric profile likelihood estimator that only requires low-dimensional optimization. Our empirical analysis studies the sensitivity of the probability weighting function estimator to various parametrizations of the utility function.

Econometric theory to disentangle attitudes toward wealth and probabilities using option prices is limited. This paper fills this gap by formally establishing the identification and asymptotic properties of our profile likelihood estimator. First, we show that the utility parameters are identified separately from the probability weighting function by considering the utility-adjusted PITs. At the true parameter vector, the latter ought to be independent of conditioning information. However, for any other parameter value, the conditional quantiles of the PITs are time-varying under mild regularity conditions. Since the weighting component for a given probability is fixed, this time variation can be attributed to an incorrect utility parameter vector, which strictly lowers the profile likelihood criterion. The probability weighting function is then identified from the distribution of the PITs evaluated at the identified utility parameters.

Second, we provide conditions under which the nonparametric estimation of the probability weighting function does not affect the asymptotic distribution of the utility parameter estimator. We verify these conditions for the kernel density estimator based on the utility-

adjusted PITs. A key challenge is that the densities of popular probability weighting functions diverge for probabilities near zero and one, jeopardizing the required stochastic equicontinuity condition. We ensure the latter using a trimming rule that can be implemented prior to estimation, and by using a censored variant of the likelihood criterion.

Monte Carlo simulations demonstrate the good finite sample performance of our semiparametric estimator in realistic settings. When inverse S-shaped probability weighting of the form proposed by [Tversky and Kahneman \(1992\)](#) is present, our profile likelihood estimator of the risk aversion parameter of the utility function yields more than five times lower mean squared error (MSE) than the (then incorrectly specified) expected utility-based maximum likelihood estimator. Meanwhile, in the absence of probability weighting, the profile likelihood estimator typically has less than double the MSE of the (then correctly specified) maximum likelihood estimator. Furthermore, a nonparametric bootstrap procedure corresponding to our semiparametric estimator yields confidence intervals with accurate coverage rates. Finally, we propose a test for probability weighting based on the test for distributional specification by [Bai \(2003\)](#), and show it has good size and power properties.

An empirical analysis of a panel of S&P 500 index options and returns from 1996 to 2023 reveals the importance of probability weighting, especially at the monthly horizon, where the [Bai \(2003\)](#) test rejects the benchmark expected utility model. At the monthly frequency, the probability weighting function takes an inverse S-shape, with larger concave (i.e., locally risk averse) than convex (i.e., locally risk seeking) regions. In a shorter sample at the weekly frequency, we find a globally concave probability weighting function.² Furthermore, allowing for probability weighting leads to lower estimates of the constant relative risk aversion parameter than when probability weighting is ignored; bootstrap confidence intervals suggest that linear utility cannot be rejected. Moreover, the shape of the probability weighting function is robust to the specification of the marginal utility function. Using a flexible

²Experimental evidence documents both inverse S-shaped and globally concave probability weighting functions; see the literature referenced below.

exponential-polynomial model that allows for convex and concave regions, the estimated pricing kernel without probability weighting is U-shaped at the monthly horizon and non-decreasing in the right tail at the weekly horizon. However, when probability weighting is allowed, the estimated marginal utility functions decrease monotonically.

A large number of experimental studies report evidence of probability weighting. [Gonzalez and Wu \(1999\)](#), [Abdellaoui \(2000\)](#), and [Bruhin et al. \(2010\)](#), to name a few, find inverse S-shaped probability weighting functions that display diminishing sensitivity when probabilities move away from zero and one. Other experimental papers find a globally concave weighting function, consistent with probability-driven risk aversion ([Van de Kuilen and Wakker, 2011](#); [Qiu and Steiger, 2011](#)). Our paper contributes to a growing literature on measuring probability weighting in real-world settings. Several studies have used individual-level field data, such as insurance choices ([Cicchetti and Dubin, 1994](#); [Barseghyan et al., 2013](#)), to estimate rank-dependent utility models using micro-econometric methods. Like [Barseghyan et al. \(2013\)](#), we avoid parametric restrictions on the probability weighting function. By contrast, our method only requires observing the market prices of risky payoffs, rather than individual holdings. The reliance on market prices is also common in studies that use betting odds ([Jullien and Salanié, 2000](#); [Snowberg and Wolfers, 2010](#)) and transform them into predicted payoff probabilities by modeling risk preferences. However, betting odds are for discrete payoffs, whereas we consider financial assets with continuous payoffs.

Option prices are particularly informative about risk preferences, as their variation across strike prices reveals entire risk-neutral probability densities for the underlying return. Upon dividing these densities by estimated objective densities, [Rosenberg and Engle \(2002\)](#) and [Aït-Sahalia and Lo \(2000\)](#) estimate the pricing kernel, or the marginal utility as a function of the aggregate wealth return. Other pricing kernel estimators avoid, like ours, directly modeling the objective distribution, by using maximum likelihood ([Bliss and Panigirtzoglou, 2004](#); [Liu et al., 2007](#); [Schreindorfer and Sichert, 2025](#)), conditional density integration ([Linn](#)

et al., 2018), or inverse density weighting (Dalderop and Linton, 2025). However, none of the above studies allow for a probability weighting component in the pricing kernel. Conversely, studies that measure probability weighting using option prices, such as Kliger and Levy (2009), Polkovnichenko and Zhao (2013), Dierkes (2013), and Chen et al. (2024), all rely on “plug-in” methods for the objective return distributions. An exception is the GMM estimator of Baele et al. (2019), who specify a fully parametric model. To our best knowledge, our econometric method yields the first nonparametric estimator of the probability weighting function that foregoes direct specification of the objective conditional return distributions.

The asymptotic analysis of our profile likelihood estimator builds on general semiparametric estimation theory developed by Andrews (1994) and Newey (1994). Their key stochastic equicontinuity property has been established for other kernel-based profile likelihood estimators, such as for the single-index model in Ai (1997), the transformation model of Linton et al. (2008), and the independent component analysis in Hafner et al. (2024). Relative to these studies, our theoretical results are obtained under primitive conditions on structural objects, namely utility and probability weighting functions, that can be verified for specific economic models. Moreover, our theory allows for general time-series dependence, subject to stationarity and mixing conditions.

This paper is organized as follows. Section 2 introduces the model framework. Section 3 develops our semiparametric estimation theory. Section 4 describes the Monte Carlo simulations. Empirical results are in Section 5. Conclusions are in Section 6. Appendices provide the proofs, uniform convergence of the employed kernel estimator, verification of the required technical conditions for popular models, and additional simulation results.

2 Model Framework

Probability weighting is a pivotal ingredient of the rank-dependent utility (RDU) model of [Quiggin \(1982\)](#). The RDU model encompasses expected utility (EU) and the dual theory of [Yaari \(1987\)](#) as special cases, and is the main building block in (cumulative) prospect theory of [Tversky and Kahneman \(1992\)](#). These extensions of EU have been developed to address its descriptive failures for decision under risk, notably [Allais \(1953\)](#)-type behavior.³

Consider the one-period stochastic (gross) return on wealth, R_1 , defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Denote by $F(r) := \mathbb{P}(R_1 \leq r)$ its cumulative distribution function and by $f(r) := F'(r)$ its probability density, assumed to exist. Consider a representative investor with utility function $u(\cdot; \gamma)$, for some parameter vector γ , henceforth assumed to be strictly increasing and twice differentiable, with initial wealth $w_0 > 0$ and next-period wealth $W_1 = R_1 w_0$. Under RDU, the investor maximizes

$$V := \int_0^\infty u(rw_0; \gamma) dZ(F(r)), \quad (2.1)$$

with $Z : [0, 1] \rightarrow [0, 1]$ a non-decreasing function satisfying $Z(0) = 0$ and $Z(1) = 1$, referred to as the probability weighting function,⁴ and henceforth assumed to be strictly increasing and differentiable.⁵

Any nonlinearity in Z implies that outcomes are no longer weighted linearly in probabilities. Effectively, this transforms objective probabilities into subjective decision weights, which helps describe empirical and experimental behavior at odds with the EU model. Experimental studies into the shape of the probability weighting function indeed find

³Based on experimental evidence, [Harrison and Swarthout \(2023\)](#) argue that RDU emerges as the most promising non-EU model for descriptive purposes.

⁴In the microeconomic literature following [Yaari \(1987\)](#), the convention is often to distort decumulative probabilities instead of cumulative probabilities, using the decumulative probability weighting function $\bar{Z}(P) := 1 - Z(1 - P)$. Clearly, convexity of \bar{Z} is equivalent to concavity of Z , and \bar{Z} is inverse S-shaped if and only if Z is.

⁵Non-decreasing utility and probability weighting functions ensure that RDU is compatible with first-order stochastic dominance ([Quiggin, 1982](#); [Yaari, 1987](#)).

evidence of non-linear weighting, often with an inverse S-shaped function⁶ that is concave in the lower part of the domain and convex in the upper part of the domain (Gonzalez and Wu, 1999; Abdellaoui, 2000; Bruhin et al., 2010) or a globally concave function corresponding to probability-driven risk aversion (Van de Kuilen and Wakker (2011); Bruggen et al. (2024)). Prelec (1998) has axiomatized a popular functional form of the probability weighting function that is capable of rationalizing inverse S-shaped probability attitudes.

Suppose the investor allocates a fraction α_0 of initial wealth to a risk-free asset with return R^0 , and a fraction α_i to risky asset i with return R^i , for $i = 1, \dots, n$. Hence, under full allocation, $R_1 = \sum_{i=0}^n \alpha_i R^i$ with $\alpha_0 = 1 - \sum_{i=1}^n \alpha_i$. As both the utility and probability weighting functions are differentiable, and provided the R^i have absolutely continuous distributions, the first-order conditions (FOCs) for optimal allocation are (e.g., Ai, 2005)

$$\frac{\partial}{\partial \alpha_i} V = \mathbb{E} \left(u'(W_1; \gamma) Z'(F(R_1))(R^i - R^0) \right) = 0, \quad i = 1, \dots, n. \quad (2.2)$$

The FOCs can be represented as $\mathbb{E}(m(R^i - R^0)) = 0$ in terms of the pricing kernel, or stochastic discount factor, $m := u'(W_1; \gamma) Z'(F(R_1))$. The pricing kernel is positive and induces a linear pricing rule, and hence is arbitrage-free. It defines an equivalent risk-neutral probability measure \mathbb{Q} with probability density q given by

$$q(r) = c^{-1} u'(w; \gamma) Z'(F(r)) f(r), \quad (2.3)$$

where $c = \mathbb{E}(u'(W_1; \gamma) Z'(F(R_1)))$. For example, CRRA utility, i.e., $u(w; \gamma) = w^{1-\gamma}/(1-\gamma)$ for $\gamma \neq 1$, implies $q(r) \propto r^{-\gamma} Z'(F(r)) f(r)$.

The physical distribution has a closed-form expression in terms of the risk-neutral density. Specifically, re-arranging (2.3) yields $Z'(F(r)) f(r) = cq(r)/u'(w; \gamma)$. This relation

⁶The inverse S-shape does not conform to second-order stochastic dominance: the RDU maximizer is (globally strongly) risk averse if and only if the utility and probability weighting functions are concave (Chew et al., 1987; Roëll, 1987; Eeckhoudt and Laeven, 2022).

establishes the following string of identities:

$$Z(F(r)) = \int_0^{F(r)} Z'(F) dF = \int_0^r Z'(F(s)) f(s) ds = c \int_0^r \frac{q(s)}{u'(sw_0; \gamma)} ds =: U(r; q, \gamma). \quad (2.4)$$

The function $U(r; q, \gamma)$ is itself a (*utility-adjusted*) distribution function, so that the normalization constant equals $c = \left(\int_0^\infty \frac{q(r)}{u'(rw_0; \gamma)} dr \right)^{-1}$. Inverting (2.4) and next differentiating with respect to r yields the physical distribution and density functions as

$$F(r) = Z^{-1}(U(r; q, \gamma)), \quad f(r) = c \frac{q(r)}{u'(rw_0; \gamma)} Z^{-1'}(U(r; q, \gamma)). \quad (2.5)$$

3 Estimation Theory

Observing financial markets over time results in a sample $\{R_{t+1}, q_t\}_{t=1}^T$ of stock return realizations and risk-neutral densities from option prices.⁷ Using a dynamic version of the model-based physical density of R_{t+1} given q_t derived in Section 2, this section develops a profile likelihood estimator to identify the utility parameters and the probability weighting function. First, we propose a nonparametric estimator of the probability weighting density based on the PITs of the utility-adjusted risk-neutral distributions (Sections 3.1–3.2). Next, we construct the profile likelihood function by substituting the density estimator for each utility parameter value into the semiparametric likelihood (Section 3.3). We then establish the joint identification of the utility and probability weighting functions (Sections 3.4–3.5), and derive relevant asymptotic properties (Sections 3.6–3.7). Throughout we allow for general time-series dependence in the observations, subject to stationarity. Before each result, we specify the required assumptions. All proofs are in Appendix A. Useful convergence

⁷In practice, q_t is estimated, for each $t = 1, \dots, T$, from a cross-section of option prices at n different strikes. Under in-fill asymptotics as $n \rightarrow \infty$, the resulting estimator is consistent, see e.g., Dalderop (2020). The asymptotic results in this section disregard estimation uncertainty in $\{q_t\}_{t=1}^T$; a formal justification for this requires a double asymptotic scheme, assuming n to grow sufficiently fast relative to T .

results for the kernel estimator are established in Appendix B.

3.1 Dynamic conditional density specification

The static RDU model (2.1) naturally extends to a dynamic setting, in which investors re-balance their positions each period based on currently available information. Let $F_t(r) := \mathbb{P}(R_{t+1} \leq r | \mathcal{F}_t)$ and $f_t(r)$ be the conditional cumulative distribution function (CDF) and probability density function (PDF), respectively, of the return on wealth given the natural filtration $\{\mathcal{F}_t\}$. Similarly, denote the risk-neutral conditional CDF by Q_t and the corresponding PDF by q_t . Moreover, we re-define the utility u for the dynamic investor to be a function of the return on wealth, to avoid relying on the non-stationary wealth level. This holds automatically under CRRA utility,⁸ or under the multiplicative habit formation model of Abel (1990), with habit equal to current wealth. Additionally, it is consistent with the general pricing kernel specifications in Rosenberg and Engle (2002).

For any value of γ , define the utility-adjusted conditional risk-neutral CDF as follows:

$$U_t(r; \gamma) := c_t(\gamma) \int_0^r \frac{q_t(s)}{u'(s; \gamma)} ds, \quad (3.1)$$

where $c_t(\gamma) = \left(\int_0^\infty \frac{q_t(r)}{u'(r; \gamma)} dr \right)^{-1}$. The conditional physical CDF and PDF of the return are then modeled as counterparts of (2.5) by

$$F_t(r; \gamma, Z) = Z^{-1}(U_t(r; \gamma)), \quad f_t(r; \gamma, Z) = c_t(\gamma) \frac{q_t(r)}{u'(r; \gamma)} Z^{-1'}(U_t(r; \gamma)). \quad (3.2)$$

3.2 Estimating the probability weighting function via PITs

Define the utility-adjusted probability integral transform (PIT) as a function of γ , $U_{t+1}(\gamma) := U_t(R_{t+1}; \gamma) \in (0, 1)$. These PITs play a key role in measuring probability weighting. At γ_0 , the

⁸In fact, this holds somewhat more broadly under homothetic preferences with a homogeneous utility function, for which the previous period's wealth drops out from the conditional CDF.

true parameter value, $U_{t+1}(\gamma_0) = Z(F_t(R_{t+1}))$ by (3.2). Since $F_t(R_{t+1})$ are i.i.d. standard uniform, $U_{t+1}(\gamma_0)$ are i.i.d. with distribution function equal to the inverse probability weighting function Z^{-1} .⁹

For given γ , define the true CDF of $U_{t+1}(\gamma)$ as $G(v; \gamma) := \mathbb{P}(U_{t+1}(\gamma) \leq v)$, or G_γ for short, and note that $G(v; \gamma_0) = Z^{-1}(v)$. A natural estimator for $G_\gamma(v)$ is the empirical CDF

$$\hat{G}(v; \gamma) = \frac{1}{T} \sum_{t=1}^T 1(U_{t+1}(\gamma) \leq v). \quad (3.3)$$

Define the probability density $g(v; \gamma) := \frac{\partial}{\partial v} G(v; \gamma)$ and its shorthand g_γ . A similarly natural estimator for $g_\gamma(v)$ is the kernel estimator

$$\hat{g}(v; \gamma) = \frac{1}{T} \sum_{t=1}^T K_h(U_{t+1}(\gamma) - v), \quad (3.4)$$

where $K_h(\cdot) = \frac{1}{h} K\left(\frac{\cdot}{h}\right)$ for some kernel K and bandwidth h . As this estimator is generally consistent for any γ when $T \rightarrow \infty$ and $h \rightarrow 0$, the challenge is estimating γ itself.¹⁰

The utility-adjusted PITs can be computed by applying integration by parts to (3.1)

$$\int_0^{R_{t+1}} \frac{q_t(r)}{u'(r; \gamma)} dr = \frac{Q_t(R_{t+1})}{u'(R_{t+1}; \gamma)} + \int_0^{R_{t+1}} Q_t(r) \frac{u''(r; \gamma)}{u'(r; \gamma)^2} dr, \quad (3.5)$$

provided $\lim_{r \rightarrow 0} \frac{Q_t(r)}{u'(r; \gamma)} = 0$. The latter expression relies on the risk-neutral CDF, which can be estimated more efficiently from cross-sections of option prices than its density, as also exploited by Polkovnichenko and Zhao (2013).

The normalization constants $c_t(\gamma)$ cannot be directly computed using the integration-by-parts formula (3.5) in the realistic case that $\lim_{r \rightarrow \infty} \frac{1}{u'(r; \gamma)} = \infty$. However, we can split

⁹Henceforth, we often use the abbreviation ‘‘PITs’’ to refer to *utility-adjusted* probability integral transforms, recognizing that these are not standard uniform in the presence of probability weighting.

¹⁰In Appendix B, we establish consistency of \hat{g}_γ , uniformly over γ and v , allowing for general time-series dependence.

the domain into two parts for some threshold κ :

$$\int_0^\infty \frac{q_t(r)}{u'(r; \gamma)} dr = \int_0^\kappa \frac{q_t(r)}{u'(r; \gamma)} dr + \int_\kappa^\infty \frac{q_t(r)}{u'(r; \gamma)} dr. \quad (3.6)$$

The first term can be computed using (3.5), and, if $\lim_{r \rightarrow \infty} \frac{1-Q_t(r)}{u'(r; \gamma)} = 0$, the second using

$$\int_\kappa^\infty \frac{q_t(r)}{u'(r; \gamma)} dr = \frac{1 - Q_t(\kappa)}{u'(\kappa; \gamma)} - \int_\kappa^\infty (1 - Q_t(r)) \frac{u''(r; \gamma)}{(u'(r; \gamma))^2} dr. \quad (3.7)$$

3.3 Profile likelihood estimator

Given the conditional density in (3.2), the conditional log likelihood of the sample $\{R_{t+1} \mid q_t\}_{t=1}^T$ for any γ in some parameter space $\Theta \subseteq \mathbb{R}^k$ and density function g on $(0, 1)$ equals

$$\begin{aligned} \ell_T(\gamma, g) &:= \frac{1}{T} \sum_{t=1}^T \log f_t(R_{t+1}; \gamma, g) - \log q_t(R_{t+1}) \\ &= \frac{1}{T} \sum_{t=1}^T \log c_t(\gamma) - \log u'(R_{t+1}; \gamma) + \log g(U_{t+1}(\gamma)), \end{aligned}$$

subtracting the parameter-independent $\log q_t(R_{t+1})$ for convenience. Rather than optimizing this criterion over the infinite-dimensional joint parameter space, we propose the estimator $\hat{\gamma} := \operatorname{argmax}_{\gamma \in \Theta} \ell_T(\gamma)$, maximizing the profile likelihood (PL) defined by substituting in the kernel density estimator (3.4):¹¹

$$\ell_T(\gamma) := \ell_T(\gamma, \hat{g}_\gamma). \quad (3.8)$$

In a general semiparametric estimation framework, Newey (1994) shows that the asymptotic variance of $\hat{\gamma}$ does not depend on the estimation error in \hat{g}_γ provided its probability limit g_γ maximizes the expected log-likelihood $\ell(\gamma, g) := \mathbb{E} \ell_T(\gamma, g)$ for any γ . The following result

¹¹While we focus on the kernel estimator, our estimation theory allows for other consistent nonparametric estimators for g_γ .

confirms the latter property.

Lemma E. *For any γ , g_γ uniquely maximizes $\ell(\gamma, g)$ in the space of positive probability density functions on $(0, 1)$.*

As a result, the infeasible PL estimator based on g_γ attains the semiparametric efficiency bound. Section 3.7 establishes that our feasible PL estimator $\hat{\gamma}$ asymptotically behaves as if g_γ is known, confirming its asymptotic efficiency.

3.4 Identification of utility parameters

The identification of the parameter vector γ depends on whether the profiled log-likelihood population criterion $\ell(\gamma) := \ell(\gamma, g_\gamma)$ is uniquely maximized at the true parameter γ_0 . Jensen's inequality implies that $\ell(\gamma) \leq \ell(\gamma_0)$ for any γ , with equality holding if and only if

$$f_t(R_{t+1}; \gamma, g_\gamma) = f_t(R_{t+1}; \gamma_0, g_{\gamma_0}) \text{ a.s.} \quad (3.9)$$

Thus, identification requires that no parameter value $\gamma \neq \gamma_0$ always yields the true conditional density $f_t(r) = f_t(r; \gamma_0, g_{\gamma_0})$. Let $ARA(r; \gamma) := -\frac{u''(r; \gamma)}{u'(r; \gamma)}$ denote the absolute risk aversion function. The following assumption rules out (3.9) for any $\gamma \neq \gamma_0$.

Assumption I.

- (i) $Z'(P)$ is positive on $(0, 1)$, $u'(r; \gamma)$ is positive and differentiable in r on $\mathbb{R}_{++} \times \Theta$, and $c_t(\gamma)$ is finite a.s. for all $\gamma \in \Theta$;
- (ii) For any $\gamma \in \Theta \setminus \{\gamma_0\}$, there exists some $\bar{r} > 0$ such that $ARA(\bar{r}; \gamma) \neq ARA(\bar{r}; \gamma_0)$ and $u''(r; \gamma)$ is continuous at \bar{r} ;
- (iii) $f_t : \mathbb{R}_{++} \rightarrow \mathbb{R}_{++}$ is continuous a.s., and for some pair $(v_1, v_2) \in \text{supp}(F_t(\bar{r}))$ the conditional quantiles $F_t^{-1}(v_1)$ and $F_t^{-1}(v_2)$ are not one-to-one.

Lemma I. *Under Assumption I, $\ell(\gamma)$ is uniquely maximized at γ_0 .*

The intuition behind Lemma I is as follows. Condition I(i) ensures that the pricing kernel is positive and finite under model (3.2), and hence satisfies the fundamental theorem of asset pricing. This ensures that G_γ is strictly increasing for all $\gamma \in \Theta$. Therefore, (3.9) becomes equivalent to $U_{t+1}(\gamma) = U_{t+1}(\gamma_0)$ *a.s.*, which implies that $U_{t+1}(\gamma)$ must be independent of any information in \mathcal{F}_t . However, we prove that for some (v_1, v_2) , the conditional quantiles of $U_{t+1}(\gamma)$ vary over time unless either $ARA(r; \gamma) \equiv ARA(r; \gamma_0)$ or there is a particular one-factor structure among the corresponding physical quantiles. The former case is directly ruled out by condition I(ii). The latter case is ruled out by condition I(iii), which ensures that the physical densities are positive, continuous, and non-deterministic in a way that is satisfied for dynamic models with at least two state variables. Thus, $U_{t+1}(\gamma)$ must depend on time- t information for any $\gamma \neq \gamma_0$, so that resulting differences in the conditional densities of R_{t+1} result in a strictly lower value of the profile likelihood.

3.5 Trimming

The presence of probability weighting results in terms $\frac{\partial}{\partial \gamma} \log g_\gamma(U_{t+1}(\gamma))$ in the scores of the likelihood function. For several popular probability weighting functions (see Appendix C), the density $Z'(P)$ diverges to infinity for probabilities near zero and one. As a result, the density g_γ converges to zero at the boundaries, causing a small denominator issue.¹² Furthermore, $\frac{\partial}{\partial \gamma} g_\gamma$ may diverge at the boundaries for common utility functions, which prohibits the uniform convergence of its kernel estimator.

To overcome these issues, we propose the following trimming rule that bounds $U_{t+1}(\gamma)$ away from zero and one, uniformly over Θ . Let $R_l^*(q, v) := \max_{\gamma \in \Theta} U^{-1}(v; q, \gamma)$ and $R_u^*(q, v) := \min_{\gamma \in \Theta} U^{-1}(v; q, \gamma)$, and for some small $v^* > 0$ let $R_{t,l}^* = R_l^*(q_t, v^*)$ and $R_{t,u}^* = R_u^*(q_t, 1 - v^*)$. By design, $R_{t,l}^* \leq R_{t+1} \leq R_{t,u}^*$ ensures that $v^* \leq U_{t+1}(\gamma) \leq 1 - v^*$ for

¹²A similar issue occurs in semiparametric copula-based models (Chen and Fan, 2006), due to asymptotes in the copula densities.

all $\gamma \in \Theta$. We assume that v^* is chosen small enough to guarantee that $R_{t,l}^* < R_{t,u}^*$, or else the time- t observation is removed. Since the thresholds $(R_{t,l}^*, R_{t,u}^*)$ do not depend on the parameter, they only need to be computed once prior to estimation.

Define the censored profile log-likelihood function $\ell_T^*(\gamma) := \ell_T^*(\gamma, \hat{g}_\gamma)$, where

$$\ell_T^*(\gamma, g) := \frac{1}{T} \sum_{t=1}^T \left[1_{t+1}^l \log G(U_t^l(\gamma)) + 1_{t+1}^m \log \left(\frac{c_t(\gamma)g(U_{t+1}(\gamma))}{u'(R_{t+1}; \gamma)} \right) + 1_{t+1}^u \log (1 - G(U_t^u(\gamma))) \right], \quad (3.10)$$

with $U_t^i(\gamma) := U_t(R_{t,i}^*; \gamma)$ for $i \in \{l, u\}$ and $1_{t+1}^l := 1(R_{t+1} \leq R_{t,l}^*)$, $1_{t+1}^m := 1(R_{t,l}^* < R_{t+1} \leq R_{t,u}^*)$, and $1_{t+1}^u := 1(R_{t,u}^* < R_{t+1})$ indicators for the lower, middle, and upper parts of the distribution, respectively. This does not exclude the trimmed returns from the estimation of γ , but reduces them to binary tail events. The density estimator \hat{g}_γ still uses all observations.

The following lemma confirms that the censored profile likelihood population criterion $\ell^*(\gamma) := \mathbb{E}\ell_T^*(\gamma, g_\gamma)$ still identifies γ_0 , as long as the non-censored intervals are ‘wide’ enough.

Lemma I*. *Under Assumption I, with $\mathbb{P}(R_{t,l}^* \leq \bar{r} \leq R_{t,u}^*) > 0$ and $(v_1, v_2) \in \text{supp}(F_t(\bar{r}) | R_{t,l}^* \leq \bar{r} \leq R_{t,u}^*)$, $\ell^*(\gamma)$ is uniquely maximized at γ_0 .*

3.6 Consistency

Besides identification, consistency of $\hat{\gamma}$ requires the uniform convergence of $\ell_T^*(\gamma)$ to $\ell^*(\gamma)$ over the parameter space Θ . If the profile density g_γ were known, this would follow from the uniform convergence of $\ell_T^*(\gamma, g_\gamma)$ to $\ell^*(\gamma, g_\gamma)$. As we estimate g_γ nonparametrically, we additionally require the uniform convergence of \hat{g}_γ . Let $\|g\|_{\infty, v^*} = \sup_{v^* \leq v \leq 1-v^*} |g(v)|$ be the sup-norm over the trimmed support. We then establish the consistency of the profile-likelihood estimator under the following assumptions.

Assumption C.

- (i) $Z'(P)$ and $u'(r; \gamma)$ are continuous on $(0, 1)$ and $\mathbb{R}_{++} \times \Theta$, respectively, with Θ compact;
- (ii) $\mathbb{E} \left(\sup_{\gamma \in \Theta} |\log u'(R_{t+1}; \gamma)| \right) < \infty$ and $\mathbb{E} \left(\sup_{\gamma \in \Theta} |\log c_t(\gamma)| \right) < \infty$;
- (iii) $\sup_{\gamma \in \Theta} \|\hat{g}_\gamma - g_\gamma\|_{\infty, v^*} \xrightarrow{P} 0$ and $\sup_{\gamma \in \Theta} |\hat{G}_\gamma(v^*) - G_\gamma(v^*)| \xrightarrow{P} 0$;
- (iv) $f_t(r) = f(r|X_t)$ is a positive, measurable function of the stationary ergodic process (R_{t+1}, X_t) , where X_t is some state vector.

Proposition C. *Suppose Assumptions I and C hold. Then $\hat{\gamma} := \arg \max_{\gamma \in \Theta} \ell_T^*(\gamma) \xrightarrow{P} \gamma_0$ when $T \rightarrow \infty$.*

Assumption C contains sufficient conditions for those in Newey (1994, Lemma 5.2). Continuity of the probability weighting density and marginal utility function in C(i) is required for the continuity of the criterion function. Crucially, $Z'(P)$ only needs to be continuous on the open interval $(0, 1)$, as under inverse S-shaped probability weighting $Z'(P)$ can have asymptotes at the boundaries, as shown in Appendix C. The moment conditions in C(ii) ensure the criterion function is dominated by an integrable function. The first moment condition is readily verified for common utility functions. For example, for power utility it amounts to the existence of $\mathbb{E}|\log R_{t+1}|$.

The uniform convergence conditions C(iii) allow for a general nonparametric estimator, not necessarily a kernel estimator. This condition can in turn be implied from primitive conditions on specific estimators. Lemma K-1 in Appendix B provides sufficient conditions for the kernel density estimator (3.4), based on results in Kristensen (2009). Similarly, Lemma K-1* provides sufficient conditions for the kernel CDF estimator

$$\hat{G}_h(v; \gamma) = \frac{1}{T} \sum_{t=1}^T F_K \left(\frac{v - U_{t+1}(\gamma)}{h} \right) = \int_{-\infty}^v \hat{g}(u; \gamma) du, \quad (3.11)$$

with $F_K(x) = \int_{-\infty}^x K(z) dz$, where the second equality presumes K is symmetric around zero. Importantly, the uniform convergence of the density estimator is only required on the compact subset $[v^*, 1 - v^*]$, as the kernel estimator may not be consistent on $[0, 1]$ when g

or g' have asymptotes near the boundaries. For the CDF estimator this is not problematic, as its summands are uniformly bounded. Finally, condition $\mathbf{C}(iv)$ is used to establish a pointwise LLN for the criterion function, while the compact parameter space is used to extend this to a uniform law.

3.7 Asymptotic normality

The score of the profile likelihood function for each non-trimmed observation is given by

$$s(R_{t+1}, q_t, \gamma, g) := \frac{\partial}{\partial \gamma} \log f_t(R_{t+1}; \gamma, g_\gamma) = \frac{c'_t(\gamma)}{c_t(\gamma)} - \frac{\partial u'(R_{t+1}; \gamma) / \partial \gamma}{u'(R_{t+1}; \gamma)} + \frac{\frac{d}{d\gamma} g_\gamma(U_{t+1}(\gamma))}{g_\gamma(U_{t+1}(\gamma))},$$

for any family of density functions $g(v; \gamma)$. The scores for left and right censored observations equal $\frac{\partial}{\partial \gamma} \log G_\gamma(U_t(R_{t,l}^*; \gamma))$ and $\frac{\partial}{\partial \gamma} \log(1 - G_\gamma(U_t(R_{t,u}^*; \gamma)))$, respectively. The FOC for maximizing (3.8) then equals

$$\frac{\partial}{\partial \gamma} \ell_T^*(\gamma) = \frac{1}{T} \sum_{t=1}^T s(R_{t+1}, q_t, \gamma, \hat{g}) = 0,$$

where \hat{g} is given in (3.4).¹³ Hereafter we denote the true PDF and CDF of $U_{t+1}(\gamma)$ as $g_{0,\gamma}$ and $G_{0,\gamma}$ when distinguishing them from arbitrary families of distributions. The results in this section establish that when \hat{g} converges uniformly to its population counterpart g_0 , the nonparametric estimation error is asymptotically irrelevant for the parameter estimator $\hat{\gamma}$.

First, note that $g(v; \gamma)$ only enters the profile likelihood score at γ_0 via the functional

$$S(r, q, g) := \frac{\frac{d}{d\gamma} g_\gamma(U(r; q, \gamma)) \Big|_{\gamma=\gamma_0}}{g_{\gamma_0}(U_0(r, q))} = \frac{\dot{g}_{\gamma_0}(U_0(r, q)) + g'_{\gamma_0}(U_0(r, q)) \dot{U}_0(r, q)}{g_{\gamma_0}(U_0(r, q))},$$

where $\dot{f}_\gamma(v) := \frac{\partial}{\partial \gamma} f_\gamma(v)$ and $f'_\gamma(v) := \frac{\partial}{\partial v} f_\gamma(v)$ for any function f , and $U_0(r, q) := U(r; q, \gamma_0)$.

¹³Trimming requires adding \hat{G} as an argument to s , unless the CDF estimator equals the integrated PDF estimator. The latter holds for the kernel CDF estimator in (3.11).

The functional S only depends on the $(k + 1)$ -dimensional function $g(u; \gamma)$ through the univariate functions g_{γ_0} , g'_{γ_0} , and \dot{g}_{γ_0} . Therefore, define the following norm for functions $g(v; \gamma)$: $\|g\| := \max\{\|g_{\gamma_0}\|_{\infty, v^*}, \|g'_{\gamma_0}\|_{\infty, v^*}, \|\dot{g}_{\gamma_0}\|_{\infty, v^*}\}$. The following conditions, based on [Newey \(1994, Assumptions 5.1–5.3\)](#), ensure that the nonparametric estimation of g does not affect the limiting distribution of $\frac{\partial}{\partial \gamma} \ell_T^*(\gamma_0)$.

Assumption A.

- (i) $Z'(P)$, $u'(r; \gamma)$, and $c_t(\gamma)$ are positive and continuously differentiable on $(0, 1)$, \mathbb{R}_{++} at γ_0 , and at γ_0 a.s., respectively, and $\mathbb{E}\|\dot{U}_0\| < \infty$;
- (ii) $\|\hat{g} - g_0\| = o_p(T^{-\frac{1}{4}})$, and $(\hat{G}_{\gamma_0}, \hat{\dot{G}}_{\gamma_0})(v^*) = (G_{\gamma_0}, \dot{G}_{\gamma_0})(v^*) + o_p(T^{-\frac{1}{4}})$;
- (iii) $\frac{1}{\sqrt{T}} \sum_{t=1}^T D_0(R_{t+1}, q_t, \hat{g} - g_0) \xrightarrow{p} 0$, where $D_0(r, q, \bar{g})$ is the pathwise derivative of S at g_0 in direction \bar{g} .

Lemma A. Under Assumptions [A](#) and [C\(iv\)](#), when $T \rightarrow \infty$,

$$\sqrt{T} \frac{\partial}{\partial \gamma} \ell_T^*(\gamma_0) = \frac{1}{\sqrt{T}} \sum_{t=1}^T s(R_{t+1}, q_t, \gamma_0, g_0) + o_p(1).$$

Condition [A\(i\)](#) ensures the score functional is asymptotically linear in the unknown density g at the true parameter. Differentiability of $u'(r; \gamma_0)$ and $c_t(\gamma)$ ensures that \dot{U}_0 exists. The trimming avoids a small denominator problem, as the continuous and positive density g_{γ_0} is bounded away from zero on $[v^*, 1 - v^*]$. [A\(ii\)](#) requires the estimators \hat{g}_γ , $\hat{\dot{g}}_\gamma$ and \hat{g}'_γ to uniformly converge at the rate $T^{-\frac{1}{4}}$ or faster. For our suggested kernel estimator [\(3.4\)](#), [Lemma K-2](#) provides sufficient conditions for this rate. In particular, using at least fourth order kernels ensures that the kernel density derivative estimator converges fast enough. Similarly, [K-2*](#) provides conditions for the sufficiently fast convergence of the kernel CDF estimator [\(3.11\)](#) at the trimming point v^* . Assumption [A\(iii\)](#) is a stochastic equicontinuity condition, and is established in [Lemma SE](#) in [Appendix B](#) for the kernel estimator [\(3.4\)](#) under certain mixing and moment conditions.

The asymptotic distribution of $\hat{\gamma}$ follows from the mean-value expansion of the FOC around γ_0 ,

$$\sqrt{T}(\hat{\gamma} - \gamma_0) = \left(\frac{\partial^2}{\partial \gamma \partial \gamma^\top} \ell_T^*(\tilde{\gamma}) \right)^{-1} \sqrt{T} \frac{\partial}{\partial \gamma} \ell_T^*(\gamma_0), \quad (3.12)$$

for some $\tilde{\gamma}$ in between $\hat{\gamma}$ and γ_0 . The following additional assumptions therefore yield the \sqrt{T} -convergence rate and asymptotic normality of our profile likelihood estimator.

Assumption N.

- (i) $\mathbb{E} \left\| \frac{\partial}{\partial \gamma} \log c_t(\gamma) \Big|_{\gamma=\gamma_0} \right\|^2$ and $\mathbb{E} \left\| \frac{\partial}{\partial \gamma} \log u'(R_{t+1}; \gamma) \Big|_{\gamma=\gamma_0} \right\|^2$ are finite;
- (ii) $Z'(P)$, $u'(r, \gamma)$, $c_t(\gamma)$, and $F_t(r)$ are twice continuously differentiable a.s. on $(0, 1)$, $\mathbb{R}_+ \times \mathcal{N}_0$, \mathcal{N}_0 , and \mathbb{R}_+ , respectively, where \mathcal{N}_0 is a neighborhood of γ_0 , γ_0 is interior to Θ , and $M := \partial \mathbb{E}(s(R_{t+1}, q_t, \gamma, g_0)) / \partial \gamma \Big|_{\gamma=\gamma_0}$ is non-singular;
- (iii) $\sup_{\gamma \in \mathcal{N}_0} \left\| \frac{\partial^{i+|j|}}{\partial v^i \partial \gamma^j} (\hat{g}_\gamma - g_\gamma) \right\|_{\infty, v^*} \xrightarrow{P} 0$ for any non-negative integer i and multi-index j with $i + |j| \leq 2$, $\sup_{\gamma \in \mathcal{N}_0} \left| \frac{\partial^{|j|}}{\partial \gamma^j} (\hat{G}_\gamma(v^*) - G_\gamma(v^*)) \right| \xrightarrow{P} 0$ for $|j| \leq 2$;
- (iv) $\mathbb{E} \left(\sup_{\gamma \in \mathcal{N}_0} \|W_t(\gamma)\| \right) < \infty$ for $W_t \in \left\{ \frac{\partial^2}{\partial \gamma \partial \gamma^\top} \log c_t(\gamma), \frac{\partial^2}{\partial \gamma \partial \gamma^\top} \log u'(R_{t+1}; \gamma), \dot{U}_{t+1}, \ddot{U}_{t+1} \right\}$.

Proposition N. Let $V := \text{var}(s(R_{t+1}, q_t, \gamma_0, g_0))$. Under Assumptions I, C, A, and N, when $T \rightarrow \infty$,

$$\sqrt{T}(\hat{\gamma} - \gamma_0) \xrightarrow{d} N(0, M^{-1}VM^{-1}). \quad (3.13)$$

Assumption N(i) is required for the covariance matrix V of the score of the observations to exist. Meanwhile, the conditions on continuity in N(ii), uniform convergence in N(iii), and integrability in N(iv), ensure the convergence in probability of the Hessian term in (3.12) to a non-singular matrix uniformly in a neighborhood of γ_0 . Lemmas K-3 and K-3* in Appendix B provide sufficient conditions for the uniform convergence of the derivatives in N(iii) for the kernel density and CDF estimators (3.4) and (3.11), respectively.

Conducting inference requires estimators of the variance. However, plug-in estimation of V and M in (3.13) requires nonparametric estimation of higher order derivatives, which

is unstable in small samples. We propose instead a simple bootstrap algorithm: draw, with replacement, T pairs $(\tilde{R}_{t+1}, \tilde{q}_t)$ from $\{R_{t+1}, q_t\}_{t=1}^T$, and repeat estimation with the bootstrap sample to find $\tilde{\gamma}$. Doing so a large number K times results in a sample $\{\tilde{\gamma}_k - \hat{\gamma}\}_{k=1}^K$ whose distribution approximates that of $\hat{\gamma} - \gamma_0$. These bootstrap estimates can be used to construct quantile- or standard deviation-based confidence intervals. The algorithm described above is a standard (nonparametric) percentile bootstrap, see e.g. [Efron and Tibshirani \(1993\)](#), developed for independent and identically distributed data. The independence assumption is violated here for $\{R_{t+1}, q_t\}_{t=1}^T$. However, because the transformation $s(R_{t+1}, q_t, \gamma_0, g_0)$ is free of serial correlation, the dependence in the original data does not affect the limiting distribution in [Proposition N](#). This implies that a dependent bootstrap method (such as the block bootstrap) is not needed here: the proposed bootstrap algorithm replicates the dependence structure in the original data insofar as it appears in the asymptotic distribution of $\hat{\gamma}$. A similar result for autoregressive models was obtained by [Gonçalves and Kilian \(2004\)](#), who refer to this as the pairwise bootstrap.

4 Monte Carlo Simulations

In this section, we describe the results of a Monte Carlo simulation study that analyzes the finite sample properties of our estimation procedure. We consider three economic data generating processes (DGPs), which differ in the shape of the probability weighting function. Additional simulation results under a different sample size, different data frequency, larger risk aversion parameter, different \mathbb{P} -dynamics, and utility misspecification, can be found in [Appendix D](#).

We simulate $N = 1000$ replications of 25 years of monthly data ($T = 300$). We model

the price process of a futures contract, F_t , under the measure \mathbb{P} by the following dynamics:

$$\begin{aligned} d \log F_t &= \left(-\frac{1}{2}V_t - \mu_J \lambda_t\right) dt + \sqrt{V_t} dW_{1,t} + \sqrt{H_t} dW_{2,t} + J_t dN_t, \\ dV_t &= 12(0.015 - V_t)dt + 0.5\sqrt{V_t} \left(-0.9dW_{1,t} + \sqrt{1 - 0.9^2}dW_{3,t}\right) + J_t^V 1_{\{J_t < 0\}} dN_t, \\ dH_t &= (0.01 - H_t)dt + 0.125\sqrt{H_t} \left(-0.5dW_{2,t} + \sqrt{1 - 0.5^2}dW_{4,t}\right), \end{aligned}$$

where W_t is a four-dimensional standard Brownian motion, and N_t is a Poisson jump process with intensity $\lambda_t = 60V_t + 30H_t$. The presence of at least two volatility factors, which together drive the time-varying jump intensity under the measure \mathbb{P} , is consistent with findings in [Andersen et al. \(2015\)](#). The jumps J_t in the price process follow a double-exponential distribution: at a jump event, there is a probability $p^- = 0.7$ of a negative exponential jump with a mean of $\eta^- = 0.05$, and a probability $1 - p^-$ of a positive exponential jump with a mean of $\eta^+ = 0.02$. In case of a negative jump in the price process, volatility co-jumps: J_t^V is exponentially distributed with an average jump size of 0.01. The parameter μ_J ensures the jumps are compensated, and equals the expected relative jump size in returns:

$$\mu_J := \mathbb{E}^{\mathbb{P}} \left(e^{J_t} - 1 \right) = \frac{1 - p^-}{1 - \eta^+} + \frac{p^-}{1 + \eta^-} - 1,$$

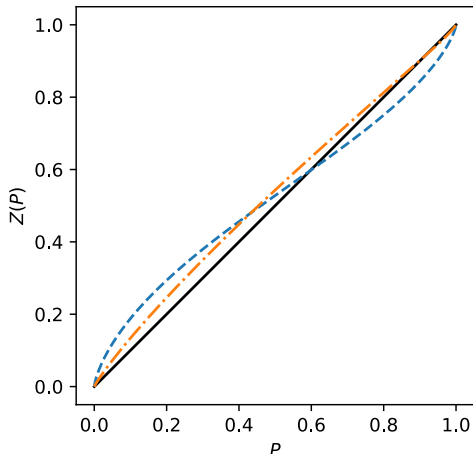
see also [Boswijk et al. \(2024\)](#). Sample paths are constructed by Euler discretization, with initial volatility drawn from its stationary distribution.¹⁴ Given the volatility at the start of each month, we compute the conditional characteristic function of the physical return distribution using [Duffie et al. \(2000\)](#). The physical density follows by Fourier inversion.

Constructing the risk-neutral densities requires modeling the conditional pricing kernel (cf. (2.2))

$$m_t(r) = u'(r; \gamma_0) Z'(F_t(r)). \tag{4.1}$$

¹⁴In practice, we simulate a single path of $100 + (T + 5)N$ years, of which the first 100 years serve as a burn-in period. The remainder is then cut into N blocks of $T + 5$ years, of which we drop the last 5 to reduce the dependence between the replications.

Figure 1: The probability weighting functions



Note: This figure displays the probability weighting functions Z_{TK} (blue, dashed), Z_P (orange, dash-dotted), Z_0 (black, solid).

We initially specify the utility function u as CRRA with $\gamma_0 = 2$. For the probability weighting function Z , we consider the one-parameter model by [Tversky and Kahneman \(1992\)](#) and the two-parameter model by [Prelec \(1998\)](#), which are defined, respectively, as

$$Z_{TK}(P) = \frac{P^\delta}{(P^\delta + (1 - P)^\delta)^{1/\delta}}, \quad Z_P(P) = 1 - \exp(-(-\beta \log(1 - P))^\alpha),$$

for some positive scalars δ , α , and β . [Appendix C](#) verifies that these functions satisfy the assumptions in [Section 3](#). For the [Tversky and Kahneman \(1992\)](#) model we set $\delta = 0.75$, which creates the typical inverse S-shape observed in experiments. For the [Prelec \(1998\)](#) model we set $(\alpha, \beta) = (0.9, 1.1)$, which describes nearly globally concave weighting. We also consider the linear weighting $Z_0(P) = P$, which describes the standard expected utility model. The resulting weighting functions are displayed in [Figure 1](#).¹⁵ We calculate the risk-neutral density q_t by multiplying the conditional physical density and pricing kernel, and normalizing to integrate to one. Our simulations thus yield N draws of the sample path $\{R_{t+1}, q_t\}_{t=1}^T$.

¹⁵The weighting densities Z'_{TK} and Z'_P diverge at 0 and 1. For numerical stability we therefore truncate the physical CDF F_t at 0.0001 and 0.9999 when calculating the pricing kernel [\(4.1\)](#). As this occurs far in the tail, minimal bias is introduced.

We simulate the profile likelihood estimator developed in Section 3 using the kernel density (3.4) to estimate the unknown weighting function g_γ . We use a fourth-order Gaussian kernel $K(u) = \frac{1}{2}(3 - u^2)\phi(u)$, with ϕ the standard normal density, and consider a variety of bandwidths. We consider the untrimmed ($v^* = 0$), and the trimmed estimator for the parameter space $\Theta = [-5, 10]$ and a range of v^* . The trimmed estimators depend on the kernel CDF estimator (3.11), for which we consider a smaller bandwidth $h^* = h/2$ than used for the density estimation.¹⁶

Table 1 summarizes the performance of the profile likelihood estimator (PLE). The bottom row of each panel also shows results for the parametric maximum likelihood estimator (MLE), which fixes the weighting function as the identity map. This MLE performs well under the expected utility model Z_0 , in which case it is correctly specified. However, its misspecification under the probability weighting models Z_{TK} and Z_P results in a positively biased risk aversion estimator, which is used to partially offset the effect of probabilistic risk aversion. Under probability weighting, the PLE outperforms the misspecified MLE for all considered bandwidths and trimming levels. The upward bias in the MLE can lift its MSE to five times that of the PLE, while without probability weighting, the PLE has less than double the MSE of the MLE for reasonably large bandwidths and trimming values. The PLE performs particularly well under the inverse S-shape weighting function Z_{TK} : its MSE values are the lowest under this DGP, and even lower than those of the MLE under Z_0 . Under the nearly globally concave Z_P parametrization, probability weighting and utility have a similar effect on the pricing kernel, which complicates their separate identification. This leads to a higher PL bias and variance under Z_P than under Z_{TK} . The integrated mean squared error (IMSE) of the nonparametric probability weighting function estimator, estimated as the inverse of (3.3), is fairly stable across tuning parameters and positively

¹⁶The factor half is motivated by the faster convergence of the kernel CDF estimator. With fourth-order kernels, the optimal bandwidth rates for the kernel PDF and CDF estimator are $O(T^{-1/9})$ and $O(T^{-1/6})$, respectively. For $T = 300$ and equal constants, the CDF bandwidth should scale by $T^{1/9-1/6} = 0.43$.

Table 1: Profile likelihood performance.

h	$v^* = 0$				$v^* = 0.001$				$v^* = 0.01$				$v^* = 0.05$			
	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE
0.15	-0.24	1.22	1.55	0.88	0.01	1.28	1.65	0.96	0.20	1.47	2.21	1.20	0.74	1.40	2.50	1.41
0.20	-0.09	0.97	0.95	0.55	0.14	1.02	1.06	0.62	0.39	1.18	1.55	0.84	0.86	1.24	2.28	1.23
0.25	0.09	0.85	0.74	0.41	0.31	0.89	0.88	0.51	0.63	1.03	1.46	0.77	1.13	1.16	2.63	1.35
0.30	0.30	0.82	0.76	0.40	0.52	0.84	0.97	0.52	0.90	0.96	1.73	0.88	1.48	1.07	3.34	1.65
MLE	1.92	0.76	4.26	1.88	1.92	0.76	4.26	1.88	1.90	0.77	4.21	1.85	1.81	0.80	3.93	1.71

(a) Z_{TK}

h	$v^* = 0$				$v^* = 0.001$				$v^* = 0.01$				$v^* = 0.05$			
	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE
0.15	-0.13	1.47	2.18	0.72	0.52	1.54	2.63	0.85	0.85	1.76	3.82	1.13	1.13	1.45	3.38	1.18
0.20	0.02	1.33	1.76	0.54	0.61	1.37	2.24	0.68	1.05	1.55	3.50	0.99	1.23	1.47	3.68	1.13
0.25	0.14	1.22	1.51	0.42	0.68	1.25	2.02	0.57	1.18	1.40	3.36	0.91	1.39	1.54	4.29	1.20
0.30	0.26	1.19	1.48	0.37	0.77	1.20	2.05	0.54	1.31	1.35	3.52	0.93	1.62	1.55	5.02	1.34
MLE	2.04	1.12	5.41	1.26	2.05	1.11	5.44	1.26	2.06	1.12	5.52	1.27	2.08	1.15	5.68	1.28

(b) Z_P

h	$v^* = 0$				$v^* = 0.001$				$v^* = 0.01$				$v^* = 0.05$			
	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE
0.15	-1.91	1.54	6.02	1.52	-0.97	1.60	3.52	1.00	-0.44	1.83	3.53	0.99	0.68	1.41	2.44	0.89
0.20	-1.90	1.40	5.58	1.36	-1.04	1.45	3.20	0.87	-0.32	1.64	2.79	0.77	0.41	1.48	2.35	0.77
0.25	-1.89	1.29	5.22	1.22	-1.09	1.32	2.94	0.76	-0.27	1.48	2.25	0.62	0.34	1.57	2.59	0.74
0.30	-1.85	1.24	4.99	1.13	-1.10	1.27	2.82	0.69	-0.21	1.41	2.04	0.56	0.47	1.63	2.86	0.75
MLE	0.07	1.18	1.40	0.28	0.08	1.18	1.39	0.28	0.07	1.19	1.42	0.28	0.06	1.22	1.48	0.26

(c) Z_0

Note: Subtables display the bias, standard deviation, and mean squared error (MSE) of the profile likelihood estimator $\hat{\gamma}$ over $N = 1000$ replications of samples of length $T = 300$ for various levels of trimming v^* , with bandwidth $h^* = h/2$ for the CDF estimator. Each subpanel represents a different true probability weighting function. Columns labeled IMSE display the integrated mean squared error of the nonparametric estimator \hat{Z} , multiplied by 1,000. The bottom rows represent the maximum likelihood estimator of the expected utility model, which fixes the weighting function as the identity map.

correlated with the MSE. The latter is not surprising as the nonparametric estimator directly depends on the parametric risk aversion estimator.

Our asymptotic theory establishes the consistency of the PLE assuming a trimming level $v^* > 0$. Indeed, without trimming, the simulated PLE under Z_0 displays a clear downward bias. This can be explained by the downward boundary bias of the kernel density estimator, which incentivizes a lower γ to increase values of $U_{t+1}(\gamma)$ that are close to zero. Trimming effectively counters this bias, as the trimmed likelihood only evaluates the kernel density away from the boundaries.¹⁷ Still, under Z_{TK} and Z_P , the PLE actually performs

¹⁷Alternatively, a simple first-order boundary bias correction can be applied to the kernel density

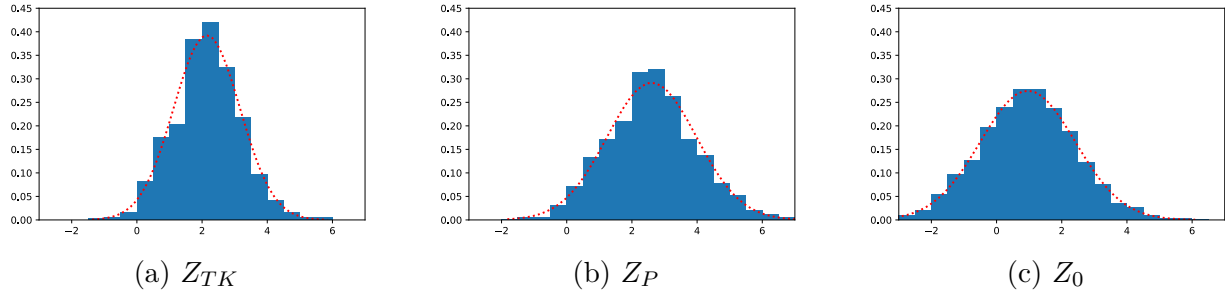
best without trimming. Both forms of probability weighting imply that $g(v)$ vanishes at the boundaries, which removes the leading boundary bias in the kernel density estimator. Meanwhile, trimming slightly increases the variance of the PLE as it does not use all information in the data. Overall, since not trimming is more harmful under Z_0 than it is beneficial under Z_{TK} and Z_P , a small amount of trimming appears prudent when it is unknown whether probability weighting is present.

Figure 2 displays histograms of the PLE $\hat{\gamma}$ at $h = 0.2$ and $v^* = 0.001$ for the three models considered. The histograms confirm that the PLE is approximately Gaussian for all three models at the realistic sample size of $T = 300$. The choice of tuning mainly affects the location and scale of the sampling distributions. The simulated nonparametric estimates of \hat{Z} are displayed in Figure 3 for the same bandwidth and trimming level. When either form of probability weighting is present, the estimated functions follow the shape of the true weighting function, and their pointwise means almost completely overlap with the true functions. In the expected utility model, the PL estimate is biased under the chosen tuning, which carries over to the nonparametric probability weighting function estimator. Nonetheless, the simulated pointwise quantiles uniformly contain the true weighting function.

In Appendix D, we discuss the performance of our feasible inference procedures. The bootstrap algorithm described in Section 3 leads to good coverage rates for the utility parameter in all settings. Furthermore, we consider a test for expected utility, or linear probability weighting, based on the method of Bai (2003). The empirical size of this test is close to, and generally below, its nominal value for all choices of tuning. Meanwhile, the power of the test against the inverse S-shaped alternative Z_{TK} is excellent, and more moderate for the nearly globally concave alternative Z_P , which is harder to distinguish from the marginal utility function, as discussed above.

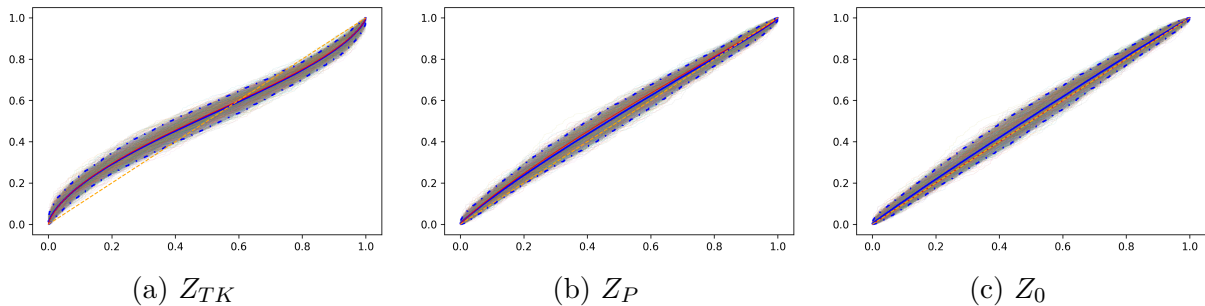
estimator. We found that this correction reduces the bias under Z_0 , but also substantially increases the estimation variance for all three DGPs. As this results in higher MSE levels, we do not consider the boundary correction further.

Figure 2: Simulation estimates of the utility parameter.



Note: The subpanels display histograms of the Monte Carlo estimates $\hat{\gamma}$ for $h = 0.2$ and $v^* = 0.001$, for three different true probability weighting functions and $\gamma_0 = 2$. The dotted red curve is a normal density with mean and standard deviation equal to that of $\hat{\gamma}$ over the $N = 1000$ replications.

Figure 3: Simulation estimates of the probability weighting function.



Note: The subpanels display the nonparametric estimates \hat{Z} of the probability weighting function, for three different true probability weighting functions. Each of the $N = 1000$ lines represents a single simulation. The mean (blue, solid), lower and upper 2.5% percentiles (blue, dash-dotted), and the true PW function (red, solid), are also displayed, along with the 45-degree line (orange, dashed).

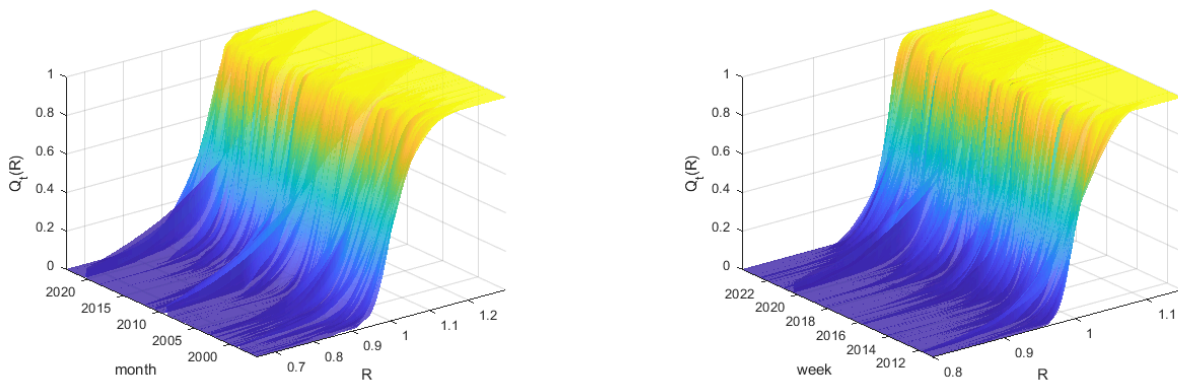
5 Empirical Results

In this section, we illustrate our econometric method using European-style option prices on the S&P 500 index (SPX) obtained from OptionMetrics. We consider monthly option data from January 1996 to January 2023 and weekly data from January 2011 to August 2023.

In particular, we consider monthly option contracts expiring on the third Friday of each month, and record their mid-quote prices on the last trading day with at least $\tau = 29$ days to maturity. The shorter weekly dataset contains the latest recorded mid prices for options expiring the next Friday with at least $\tau = 7$ days to maturity. The relevance of weekly options is discussed in detail in, e.g., [Andersen et al. \(2017\)](#). We remove options violating

static arbitrage bounds, with negative bid-ask spreads, or with unavailable implied volatility. We apply the constrained least squares method of [Aït-Sahalia and Duarte \(2003\)](#) to ensure each period’s cross section of option prices is monotone and convex in the strike price. Finally, the risk-neutral distributions Q_t of the futures return are extracted from the resulting option cross-sections using the local cubic kernel estimator of [Dalderop \(2020\)](#), with a plug-in pilot bandwidth based on fitting the single-factor stochastic volatility jump-diffusion model in [Bates \(1996\)](#).¹⁸ Figure 4 shows that this produces valid and smooth distributions throughout the sample. We complete the data set with futures returns computed from the SPX forward prices for matching expiry dates in OptionMetrics.

Figure 4: Nonparametrically estimated risk-neutral cumulative distributions for S&P 500 index returns.



(a) Monthly risk-neutral CDFs.

(b) Weekly risk-neutral CDFs.

Note: These figures display nonparametric kernel estimators of the risk-neutral CDFs estimated from monthly and weekly option prices.

With the sample $\{R_{t+1}, Q_t\}_{t=1}^T$ of returns and risk-neutral CDFs, we consider estimation under expected utility and under rank-dependent utility. We specify marginal utility as a

¹⁸Since nonparametric methods become unstable in the tails of the distribution due to sparse trading of deep OTM options, we ‘paste’ the tails of the [Bates \(1996\)](#) model to match at the lower and upper moneyness thresholds that leave 10 observed option prices in either tail. On average, this leaves about 1% of the risk-neutral probability mass in either tail, with the highest fractions in the early years of the sample.

flexible exponential-polynomial function used also by [Rosenberg and Engle \(2002\)](#):

$$u'(r; \gamma) = \exp\left(-\sum_{l=1}^L \gamma_l (\log r)^l\right). \quad (5.1)$$

To highlight the effect of probability weighting in a familiar setting, we start our discussion with the special case $L = 1$, which amounts to power utility. In particular, we define the MLE by maximizing (3.10) as a function of the CRRA parameter γ , while fixing the probability weighting function as the identity map. We use a trimming level $v^* = 0.05$ for the MLE, for which the [Bai \(2003\)](#) test has the highest power and lowest size in simulations for confidence level $\alpha = 0.05$, see Appendix D. We search for values of γ in the interval $[-5, 10]$, and thus allow for both risk-averse and risk-loving preferences. The normalization constants are computed using the integration-by-parts formulae (3.6) and (3.7). At the monthly frequency, we find a risk aversion parameter estimate of $\hat{\gamma}^{\text{ML}} = 2.34$ (0.62, 4.06), with the 95% bootstrap confidence interval in brackets. For the weekly sample, we find $\hat{\gamma}^{\text{ML}} = 4.39$ (1.49, 7.28). We test for the null of expected utility against the general alternative of non-expected utility, by computing the test statistic of [Bai \(2003\)](#) for the estimated utility-adjusted PITs $U_{t+1}(\hat{\gamma}^{\text{ML}})$. For the monthly data this results in p -values of 0.031 and 0.089 for the primal (left half) and dual (right half), respectively. For the weekly data the corresponding p -values are 0.170 and 0.581. This can be interpreted as significant evidence of non-expected utility in the monthly data, observable in the left-tail of the distribution.

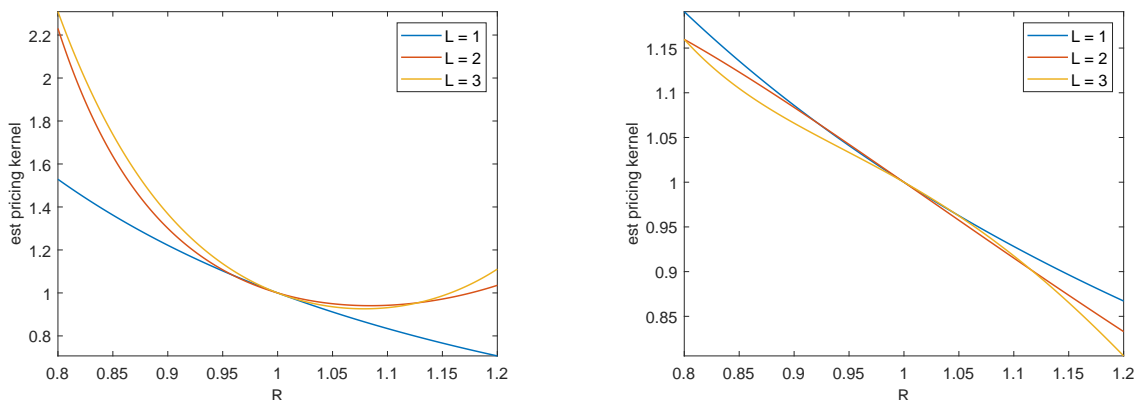
Next, we estimate the RDU model using our PL estimation procedure. Thus, we maximize the likelihood (3.9) as a function of the CRRA parameter γ by profiling out the probability weighting function based on the kernel density estimator (3.4) of the utility-adjusted PITs computed using (3.5). This results in estimated values of $\hat{\gamma}^{\text{PL}} = 0.59$ (-1.24, 2.43) and $\hat{\gamma}^{\text{PL}} = 2.87$ (-0.26, 6.00) for the monthly and weekly sample, respectively. The probability weighting function is then estimated based on the PITs at the profile likelihood maximizer

$\hat{\gamma}^{\text{PL}}$, see $L = 1$ in Figure 6.

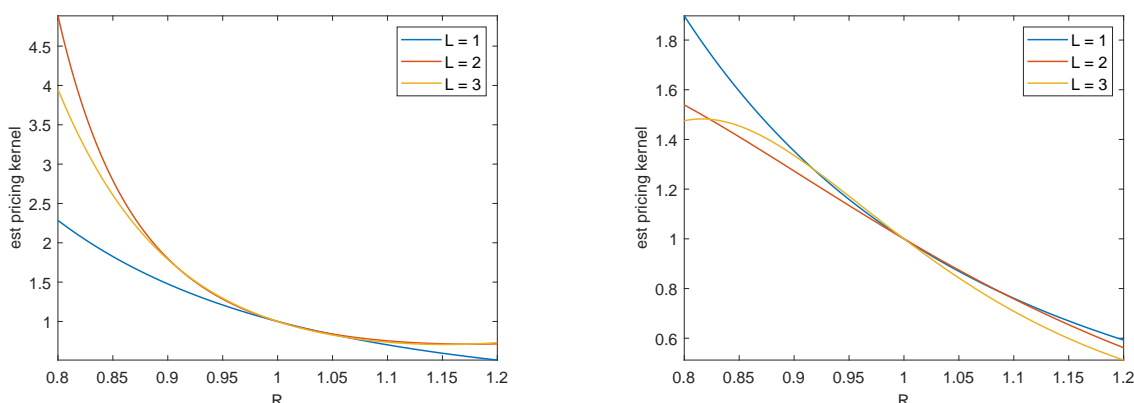
The bootstrap confidence intervals of the PL estimates contain zero at both frequencies, so that we cannot reject linear utility. This contrasts with the significantly positive risk aversion parameter estimate in the expected utility model. The lower risk aversion in utility is partially picked up by probabilistic risk aversion, with estimated probability weighting functions visibly distinct from the 45-degree line, see Figure 6. The magnitude of the respective divergences from linearity are in line with the p -values found in the Bai (2003) test. For monthly data, the probability weighting function exhibits an inverse S-shape, with larger concave than convex regions. The probability weighting function based on weekly data is closer to being linear, but appears to be globally concave.

The above analysis suggests the presence of pronounced probability weighting, particularly at the monthly frequency. However, these estimates are subject to the assumption of CRRA utility. An alternative explanation for the findings is that this marginal utility function is misspecified. For robustness, we consider more general marginal utility by taking $L > 1$ in (5.1). Figure 5 shows the resulting estimated marginal utility functions with and without probability weighting. The parameter space is chosen such that marginal utilities at returns $\pm 50\%$ are within a range of reasonable CRRA utility values. Without probability weighting, marginal utility displays clear non-monotonic U-shapes for monthly data when $L \geq 2$; for $L = 1$, marginal utility is monotone by construction. Increasing marginal utility for positive returns is in line with the literature on the “pricing kernel puzzle” (see Cuesdeanu and Jackwerth (2018) for a survey). For weekly data, the higher order marginal utility estimates are not strictly monotonically decreasing either, as they flatten for positive returns, but to a much lesser extent than for monthly data. However, when probability weighting is allowed for, the marginal utility estimates decrease monotonically on the interval $(0.8, 1.2)$ for all orders of L and both frequencies considered. Thus, the presence of probability weighting allows replacing puzzling U-shaped marginal utilities by more

Figure 5: Estimated exponential-polynomial marginal utility functions for varying order L .



(a) Without probability weighting, monthly data. (b) With probability weighting, monthly data.



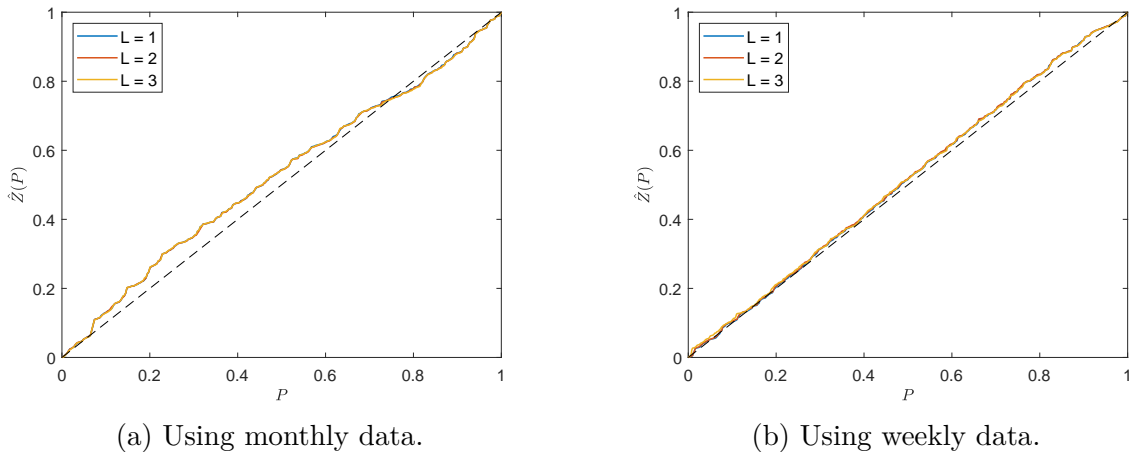
(c) Without probability weighting, weekly data. (d) With probability weighting, weekly data.

Note: Plots display (profile) maximum likelihood-estimated exponential-polynomial marginal utility functions for order $L = 1, 2, 3$, based on the local linear risk-neutral CDF estimator, and a Gaussian fourth-order kernel with bandwidth $h = 0.25$ for monthly and $h = 0.20$ for weekly data and trimming constant $v^* = 0.001$ for the probability weighting density estimator. For $L \geq 2$, trimming is based on the parameter set $\Theta_L = \{\theta \in \mathbb{R}^L : R_l^{\gamma_u} \leq u'_L(R_l; \theta) \leq R_l^{-\gamma_l} \text{ and } R_u^{\gamma_l} \leq u'_L(R_u; \theta) \leq R_u^{-\gamma_u} \text{ for } (R_l, R_u) = (0.5, 1.5), (\gamma_l, \gamma_u) = (-5, 10)\}$.

plausible monotonically decreasing functions that closely resemble CRRA utility. Moreover, probability weighting substantially reduces the amount of curvature in the marginal utility functions required to match option prices and returns.

Figure 6 shows the resulting estimates of the probability weighting function as the inverse of the empirical CDF (3.3). The weighting functions are nearly indistinguishable for different orders L of the exponential-polynomial utility model, confirming the robustness of our findings to the utility specification. Using monthly data yields particularly pronounced inverse S-shapes, while using weekly data yields weighting functions that are closer to linear.

Figure 6: Nonparametric estimates of the probability weighting function under exponential-polynomial utility.



Note: Plots display the estimates of the probability weighting function as the inverse of the empirical CDF (3.3) at the profile ML estimates $\hat{\gamma}^{\text{PL}}$ for varying order L of exponential-polynomial utility models from Figure 5. The weighting functions are nearly indistinguishable for different orders L . The dashed, blue curve is the 45-degrees line.

The plots also reveal a slight asymmetry, with stronger nonlinear weighting for the left than the right tail, implying the probabilities of large negative returns are most overweighted.

6 Conclusion

The intertwined nature of attitudes toward wealth and toward probabilities challenges their joint identification based on financial market data. Using the time series of risk-neutral distributions implicit in option contracts and realized returns, this paper develops a semiparametric profile likelihood estimator of the rank-dependent utility model. Our estimation procedure jointly identifies the probability weighting and utility functions, is easy to compute, and avoids directly specifying physical return dynamics. We establish its asymptotic properties, and report favorable performance in realistic Monte Carlo simulations. Our empirical analysis of large samples of monthly and weekly S&P 500 index option prices and returns unveils the importance of probability weighting. In particular, at the monthly horizon the weighting function implicit in option prices shows a pronounced inverse S-shape.

This finding appears robust to the parametric specification of the utility function. Our results, and the nonlinearities in probabilities of investors' risk evaluation that they entail, contribute to our understanding of risk preferences, and have relevant implications for option pricing, hedging, and risk management.

7 Disclosure statement

The authors report there are no competing interests to declare.

SUPPLEMENTARY MATERIAL

Appendix: Containing: (A) all proofs, (B) technical results on the kernel estimator, (C) verification of technical conditions, and (D) additional simulation results. (.pdf file)

References

- ABDELLAOUI, M. (2000): "Parameter-free elicitation of utility and probability weighting functions," *Management Science*, 46, 1497–1512.
- ABEL, A. B. (1990): "Asset prices under habit formation and catching up with the Joneses," *American Economic Review*, 80, 38–42.
- AI, C. (1997): "A semiparametric maximum likelihood estimator," *Econometrica*, 65, 933–963.
- AI, H. (2005): "Smooth nonexpected utility without state independence," *Working paper, Federal Reserve Bank of Minneapolis*.
- AÏT-SAHALIA, Y. AND J. DUARTE (2003): "Nonparametric option pricing under shape restrictions," *Journal of Econometrics*, 116, 9–47.
- AÏT-SAHALIA, Y. AND A. W. LO (2000): "Nonparametric risk management and implied risk aversion," *Journal of Econometrics*, 94, 9–51.
- ALLAIS, M. (1953): "Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine," *Econometrica*, 21, 503–546.
- ANDERSEN, T. G., N. FUSARI, AND V. TODOROV (2015): "The risk premia embedded in index options," *Journal of Financial Economics*, 117, 558–584.
- (2017): "Short-term market risks implied by weekly options," *The Journal of Finance*, 72, 1335–1386.

- ANDREWS, D. W. (1994): “Asymptotics for semiparametric econometric models via stochastic equicontinuity,” *Econometrica*, 62, 43–72.
- BAELE, L., J. DRIESSEN, S. EBERT, J. M. LONDONO, AND O. G. SPALT (2019): “Cumulative prospect theory, option returns, and the variance premium,” *The Review of Financial Studies*, 32, 3667–3723.
- BAI, J. (2003): “Testing parametric conditional distributions of dynamic models,” *Review of Economics and Statistics*, 85, 531–549.
- BARBERIS, N. C. (2013): “Thirty years of prospect theory in economics: A review and assessment,” *Journal of Economic Perspectives*, 27, 173–196.
- BARSEGHYAN, L., F. MOLINARI, T. O’DONOGHUE, AND J. C. TEITELBAUM (2013): “The nature of risk preferences: Evidence from insurance choices,” *American Economic Review*, 103, 2499–2529.
- BATES, D. S. (1996): “Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options,” *The Review of Financial Studies*, 9, 69–107.
- BLISS, R. R. AND N. PANIGIRTZOGLU (2004): “Option-implied risk aversion estimates,” *The Journal of Finance*, 59, 407–446.
- BOSWIJK, H. P., R. J. A. LAEVEN, AND E. VLADIMIROV (2024): “Estimating option pricing models using a characteristic function-based linear state space representation,” *Journal of Econometrics*, 244, 105864.
- BRUGGEN, P. V., R. J. A. LAEVEN, AND G. VAN DE KUILEN (2024): “Higher-order risk attitudes for non-expected utility,” Working paper, Center for Economic Research, centER Discussion Paper Nr. 2024-019, <https://research.tilburguniversity.edu/files/101140130/2024-019.pdf>.
- BRUHIN, A., H. FEHR-DUDA, AND T. EPPER (2010): “Risk and rationality: Uncovering heterogeneity in probability distortion,” *Econometrica*, 78, 1375–1412.
- CHABI-YO, F., R. GARCIA, AND E. RENAULT (2008): “State dependence can explain the risk aversion puzzle,” *The Review of Financial Studies*, 21, 973–1011.
- CHEN, T.-Y., Y.-L. LIN, AND L. Y. TZENG (2024): “Estimating Probability Weighting Functions through Option Pricing Bounds,” *The Review of Asset Pricing Studies*, 14, 513–543.
- CHEN, X. AND Y. FAN (2006): “Estimation of copula-based semiparametric time series models,” *Journal of Econometrics*, 130, 307–335.
- CHEW, S., E. KARNI, AND Z. SAFRA (1987): “Risk aversion in the theory of expected utility with rank dependent probabilities,” *Journal of Economic Theory*, 42, 370–381.
- CICCHETTI, C. J. AND J. A. DUBIN (1994): “A microeconomic analysis of risk aversion and the decision to self-insure,” *Journal of Political Economy*, 102, 169–186.

- CUESDEANU, H. AND J. C. JACKWERTH (2018): “The pricing kernel puzzle: Survey and outlook,” *Annals of Finance*, 14, 289–329.
- DALDEROP, J. (2020): “Nonparametric filtering of conditional state-price densities,” *Journal of Econometrics*, 214, 295–325.
- DALDEROP, J. AND O. B. LINTON (2025): “Estimating a conditional density ratio model for asset returns and option demand,” *Available at SSRN 4779268*.
- DIERKES, M. (2013): “Probability weighting and asset prices,” *Available at SSRN 2253817*.
- DUFFIE, D., J. PAN, AND K. SINGLETON (2000): “Transform analysis and asset pricing for affine jump-diffusions,” *Econometrica*, 68, 1343–1376.
- EECKHOUDT, L. R., C. GOLLIER, AND H. SCHLESINGER (2005): *Economic and Financial Decisions under Risk*, Princeton: Princeton University Press.
- EECKHOUDT, L. R. AND R. J. A. LAEVEN (2022): “Dual moments and risk attitudes,” *Operations Research*, 70, 1330–1341.
- EFRON, B. AND R. J. TIBSHIRANI (1993): *An Introduction to the Bootstrap*, Chapman & Hall, New York.
- GONZALEZ, R. AND G. WU (1999): “On the shape of the probability weighting function,” *Cognitive Psychology*, 38, 129–166.
- GONÇALVES, S. AND L. KILIAN (2004): “Bootstrapping autoregressions with conditional heteroskedasticity of unknown form,” *Journal of Econometrics*, 123, 89–120.
- HAFNER, C. M., H. HERWARTZ, AND S. WANG (2024): “Statistical identification of independent shocks with kernel-based maximum likelihood estimation and an application to the global crude oil market,” *Journal of Business & Economic Statistics*, 1–16.
- HARRISON, G. W. AND J. T. SWARTHOUT (2023): “Cumulative prospect theory in the laboratory: A reconsideration,” in *Models of Risk Preferences: Descriptive and Normative Challenges*, ed. by G. W. Harrison and D. Ross, Leeds: Emerald Group Publishing Limited, vol. 22, 107–192.
- JULLIEN, B. AND B. SALANIÉ (2000): “Estimating preferences under risk: The case of racetrack bettors,” *Journal of Political Economy*, 108, 503–530.
- KLIGER, D. AND O. LEVY (2009): “Theories of choice under risk: Insights from financial markets,” *Journal of Economic Behavior & Organization*, 71, 330–346.
- KRISTENSEN, D. (2009): “Uniform convergence rates of kernel estimators with heterogeneous dependent data,” *Econometric Theory*, 25, 1433–1445.
- LINN, M., S. SHIVE, AND T. SHUMWAY (2018): “Pricing kernel monotonicity and conditional information,” *The Review of Financial Studies*, 31, 491–531.
- LINTON, O., S. SPERLICH, AND I. VAN KEILEGOM (2008): “Estimation of a semiparametric transformation model,” *The Annals of Statistics*, 36, 686–718.

- LIU, X., M. B. SHACKLETON, S. J. TAYLOR, AND X. XU (2007): “Closed-form transformations from risk-neutral to real-world distributions,” *Journal of Banking & Finance*, 31, 1501–1520.
- NEWKEY, W. K. (1994): “The asymptotic variance of semiparametric estimators,” *Econometrica*, 62, 1349–1382.
- POLKOVNICHENKO, V. AND F. ZHAO (2013): “Probability weighting functions implied in options prices,” *Journal of Financial Economics*, 107, 580–609.
- PRELEC, D. (1998): “The probability weighting function,” *Econometrica*, 66, 497–527.
- QIU, J. AND E.-M. STEIGER (2011): “Understanding the two components of risk attitudes: An experimental analysis,” *Management Science*, 57, 193–199.
- QUIGGIN, J. (1982): “A theory of anticipated utility,” *Journal of Economic Behavior & Organization*, 3, 323–343.
- ROËLL, A. (1987): “Risk aversion in Quiggin and Yaari’s rank-order model of choice under uncertainty,” *The Economic Journal*, 97, 143–159.
- ROSENBERG, J. V. AND R. F. ENGLE (2002): “Empirical pricing kernels,” *Journal of Financial Economics*, 64, 341–372.
- SCHREINDORFER, D. AND T. SICHERT (2025): “Conditional risk and the pricing kernel,” *Journal of Financial Economics*, 171, 104106.
- SNOWBERG, E. AND J. WOLFERS (2010): “Explaining the favorite–long shot bias: Is it risk-love or misperceptions?” *Journal of Political Economy*, 118, 723–746.
- SONG, Z. AND D. XIU (2016): “A tale of two option markets: Pricing kernels and volatility risk,” *Journal of Econometrics*, 190, 176–196.
- TVERSKY, A. AND D. KAHNEMAN (1992): “Advances in prospect theory: Cumulative representation of uncertainty,” *Journal of Risk and Uncertainty*, 5, 297–323.
- VAN DE KUILEN, G. AND P. P. WAKKER (2011): “The midweight method to measure attitudes toward risk and ambiguity,” *Management Science*, 57, 582–598.
- YAARI, M. (1987): “The dual theory of choice under risk,” *Econometrica*, 55, 95–115.