HOUSE OF FINANCE

WORKING PAPER SERIES

Sven Thorsten Jakusch

On the Applicability of Maximum Likelihood Methods: From Experimental to Financial Data

SAFE Working Paper No. 148

SAFE | Sustainable Architecture for Finance in Europe
A cooperation of the Center for Financial Studies and Goethe University Frankfurt

House of Finance | Goethe University
Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main

Tel. +49 69 798 34006 | Fax +49 69 798 33910
info@safe-frankfurt.de | www.safe-frankfurt.de

Non-Technical Summary

Microeconomic modeling of investors' financial decision making and its results crucially depend on the mathematical shape of the underlying preference function as well as its parameterization. The majority of such models rely on expected utility, first established in the seminal work of John von Neumann and Oskar Morgenstern in 1947. Due to the shortcomings of the expected utility theory in conciliating empirical evidence with theoretical predictions, researchers proposed alternative and generalized utility concepts, such as Rank-dependent, Prospect Theory and Cumulative Prospect Theory as a refinement of the former to improve descriptive accuracy of these models. This raises the question of which assumption about the proposed utility specification is valid to properly characterize risk preferences in financial markets. Solutions to the question of how to identify the best fitting utility model in controlled experimental environments have been proposed and conducted by e.g. John Hey and Christ Orme, however, to the best of our knowledge, no study had yet embarked on a comparable endeavor for behavior in financial markets.

This paper addresses whether and to what extent econometric methods used in experimental studies can be adapted and applied to financial data to detect the best-fitting preference model. To address the research question, we implement a frequently used nonlinear probit model and base our analysis on a tailor-made simulation study. In detail, we simulate trading sequences for a set of utility models and try to identify the underlying utility model and its parameterization used to generate these sequences by maximum likelihood.

We find that for a very broad classification of utility models, this method provides acceptable outcomes. Yet, a closer look at the preference parameters reveals several caveats that come along with typical issues attached to financial data and that partly seem to drive our results. In particular, deviations are attributable to effects stemming from multicollinearity and coherent under-identification problems, where some of these detrimental effects can be remedied up to a certain degree by adjusting the error term specification and thus the assumption of the likelihood function.

Furthermore, additional uncertainty stemming from changing market parameter estimates affects the precision of our estimates for risk preferences and cannot be simply remedied by using a higher standard deviation of the error term or a different assumption regarding its stochastic process. Particularly, if the variance of the error term becomes large, we detect a tendency to identify Prospect Theory as utility model providing the best fit to simulated trading sequences regardless of the utility function used to generate these trading sequences. We also find that a frequent issue, namely serial correlation of the residuals, does not seem to be significant.

However, we detect a tendency to prefer nesting models over nested utility models, which is particularly prevalent if Rank-dependent Utility and Exponential Power-Utility models are estimated along with CRRA utility models.

ON THE APPLICABILITY OF MAXIMUM LIKELIHOOD METHODS: FROM EXPERIMENTAL TO FINANCIAL DATA

SVEN THORSTEN JAKUSCH

Abstract. This paper addresses whether and to what extent econometric methods used in experimental studies can be adapted and applied to financial data to detect the best-fitting preference model. To address the research question, we implement a frequently used nonlinear probit model in the style of Hey and Orme (1994) and base our analysis on a simulation study. In detail, we simulate trading sequences for a set of utility models and try to identify the underlying utility model and its parameterization used to generate these sequences by maximum likelihood. We find that for a very broad classification of utility models, this method provides acceptable outcomes. Yet, a closer look at the preference parameters reveals several caveats that come along with typical issues attached to financial data, and that some of these issues seems to drive our results. In particular, deviations are attributable to effects stemming from multicollinearity and coherent under-identification problems, where some of these detrimental effects can be captured up to a certain degree by adjusting the error term specification. Furthermore, additional uncertainty stemming from changing market parameter estimates affects the precision of our estimates for risk preferences and cannot be simply remedied by using a higher standard deviation of the error term or a different assumption regarding its stochastic process. Particularly, if the variance of the error term becomes large, we detect a tendency to identify prospect theory as utility model providing the best fit to simulated trading sequences. We also find that a frequent issue, namely serial correlation of the residuals, does not seem to be an important issue. However, we detected a tendency to prefer nesting models over nested utility models, which is particularly prevalent if rank-dependent utility and exponential power utility models are estimated along with expected utility with constant relative risk aversion utility models.

Date: 24 December, 2013.

Key words and phrases. Utility Functions, Model Selection, Parameter Elicitation.

House of Finance, Goethe University Frankfurt, Grueneburgplatz 1, D-60323 Frankfurt am Main, Germany. Sven Jakusch is a doctoral student at House of Finance, Goethe University Frankfurt and Senior Quantitative Consultant at Ernst & Young GmbH. We are grateful for comments by Andreas Hackethal, Glenn Harrison, John Hey, Steffen Meyer and Chris Orme. We gratefully acknowledge research support from the Research Center SAFE, funded by the state of Hessen initiative for research LOEWE. The corresponding authors can be reached by sven-jakusch@yahoo.de. Please note that parts of this paper were written when Sven Jakusch was working at Ernst & Young Wirtschaftspruefungsgesellschaft GmbH, however, any views, statements or opinions expressed in this paper are solely those of the authors and not related to Ernst & Young.

1. Introduction

Microeconomic modeling of investor financial decision making and its results depend on the mathematical shape of the underlying preference function as well as its parameterization. The majority of such models rely on expected utility, first established in the seminal von Neumann and Morgenstern (1947). Due to the shortcomings of the expected utility theory in conciliating empirical evidence and theoretical predictions, researchers have set forth alternative and generalized utility concepts, such as rank-dependent utility (Quiggin (1982)), prospect theory as conceptualized by Kahneman and Tversky (1979) and cumulative prospect theory as a refinement of the former (posed in Tversky and Kahneman (1992)) to improve the descriptive accuracy of these models. This raises the question: Which assumption about the proposed utility specification is valid to properly characterize risk preferences in financial markets? Solutions for how to identify the best-fitting utility model in controlled experimental environments are proposed by, for example, Hey and Orme (1994) and Laury and Holt (2005). However, to the best of our knowledge, no study considers a comparable endeavor for behavior in financial markets. Financial data, in contrast to experimental data, have certain characteristics, but also come with significant disadvantages, which might require some modifications of the methods adopted from experimental economics. In particular, such data comprise revealed rather than stated preferences (Train (2009)), offer a considerable sample size but might suffer from unobservable factors beyond the control of the researcher, first-order autocorrelation in the time series as well as between unobservable factors, under-identification problems regarding the utility models and a certain degree of multicollinearity (Campbell et al. (1997)), and introduction of additional uncertainty stemming from the way investors tend to extrapolate past returns into the future (Andreassen and Kraus (1990)) and carry over accumulated gains and losses over time.

We assess the compatibility of both econometric concepts from the experimental literature, namely, customized maximum likelihood methods, based on modeling an additional error term on top of the individuals decision rule as introduced by Hey and Orme (1994), and the effects of selected features of financial data, particularly the trading behavior from individual investors. We then identify and analyze potential problems that arise when using trade data such as multicollinearity, the effects of additional uncertainty regarding the stochastic properties of the likelihood function, autocorrelation, and the identifiability of the true but unknown functional shape of an investor's utility function. These problems can be generated via the way investors obtain estimates to approximate uncertain financial outcomes (Kahneman and Tversky (1973), Andreassen and Kraus (1990)) and by carrying forward accrued returns over time.

This paper proceeds as follows. We outline the research question and provide an overview of the current relevant experimental literature to frame the topic in the second section. In the third section, we sketch a frequently applied likelihood approach for utility model identification and present the inherent statistical properties that can be expected to hold in experimental data. This allows us to identify and highlight several weak points in the widely applied likelihood approach with regard to financial data and to show from which factors potential problems may arise. In Section 4, we focus on the problems identified in a preliminary likelihood analysis of the results from application of the maximum likelihood method from a simulation study. We then present and discuss each of these problems in detail, particularly how our results can be affected from a model selection procedure. In this section, we analyze in more detail the factors suspected to alter the surface of the likelihood function, to yield unreliable estimators for risk preference parameters, and to cause breakdowns in certain numerical search algorithms, thus affecting model selection results. We elaborate to what extent the above-mentioned effects arise and alter the results of the utility model identification strategy.

Our findings confirm that the additional uncertainty that yields a modification of the stochastics of the likelihood function, introduced by the unknown stochastic process stemming from carrying forward intermediate gains and losses, multicollinearity, and under-identification interferes with the precision of our estimators and thus the identification of the underlying utility model. Autocorrelation, on the other hand, appears to predominantly affect selection of the error term specification. However, if the error term specification interacts with model selection in a way that, by coincidence, captures part of the effects of the additional uncertainty and of multicollinearity, estimation of the variance of the error term, an additional nuisance parameter, can supplement and enhance identification of the correct utility model specification.

2. Preferences in financial markets and experiments

Empirical and theoretical research in finance places great emphasis on the mathematical specification of the utility function of an individual decision maker, usually concluding that most individuals are indeed risk averse, despite the fact that notable exceptions exist.[1] Findings from experimental and empirical studies paint a multifaceted picture in this matter, as they are based on various methods and data sets, and thus they are difficult to compare directly and yield results that are virtually impossible to reconcile. Despite apparent consensus about the general relevance of risk aversion in the theory of financial decision making, the exact characterization of an investor's utility function is a highly disputed topic. For example, regarding the classical expected utility paradigm, such studies as Friend and Blume (1975), Blume and Friend (1975), Schlarbaum et al. (1975), Morin and Suarez (1983) and Landskroner (1988) find evidence for decreasing absolute risk aversion (DARA) and constant relative risk aversion (CRRA), and estimate the coefficient for relative risk aversion to be approximately 2 or higher; this is incompatible with the assertions of logarithmic utility (Latane (1959), Hakansson (1971) and Markowitz (1976)), a conjecture supported by recent empirical studies based on household survey data.[2]

In contrast to this strand of the literature, experimental evidence outside the field of finance is vast and finds somewhat lower values, with ambiguous results. For example, Gordon et al. (1972), Kroll et al. (1988) and Levy (1994), who analyze study participant decision making in a portfolio choice context, show evidence for DARA and moderate support for either increasing relative risk aversion (IRRA) or CRRA, although other authors object that IRRA might be an artifact of the inherent positive-or-zero gain feature of such experiments (Levy (1994)), and that absolute risk aversion cannot be unambiguously recovered from actual choices (Wolf and Pohlman (1983)).

[1] Friedman and Savage (1948), Markowitz (1952), Tversky and Kahneman (1992) and Kahneman and Tversky (1979).

[2] Although Brunnermeier and Nagel (2008) conclude that logarithmic utility may provide an appropriate description of financial market risk aversion, Guiso and Paiella (2008) and Chiappori and Paiella (2011) detect signs of DARA and CRRA with (highly dispersed) coefficients of relative risk aversion above 2 for a significant proportion of households.

Although it appears that DARA and CRRA utility dominates financial markets, classical utility theories have been questioned due to their incompatibility with empirical phenomena, such as the observed equity premium[3], and on the observed violation of their inherent axiomatic properties (Allais (1953), see also Edwards (1996), Barberis and Thaler (2003), Glaser et al. (2004), Shefrin (2008), Wang (2006), Broihanne et al. (2008), the recent Wakker (2010) and Barberis (2013)). To address the latter issue, early experimental studies such as Preston and Baratta (1948), Edwards (1953), and Edwards (1954) reveal that decision makers systematically violate the independence axiom of von Neumann and Morgenstern (1947), thus concluding that subjects decide in discord with physical probabilities and seem to apply decision weights to making choices. These findings prompted development of generalized expected utility theories, such as Dual Theory (Yaari (1987)), and creation of rank-dependent utility (RDU) as advocated by Edwards (1962), Karmarkar (1978), Karmarkar (1979), Quiggin (1982) and Wakker (1994), whereas another strand of the literature proposes modifications of the utility function itself (e.g., Friedman and Savage (1948), Markowitz (1952), Kahneman and Tversky (1979), Wakker and Tversky (1993)). According to these generalized expected utility models, risk aversion is now not only determined by the curvature of the utility function, but is also dependent on the shape of the decision weight attached to the alternatives of the choice set. Despite the fact that increasing marginal utility causes a degree of discomfort for economists (Yaari (1965)), recent empirical and experimental evidence provides further support for these alternative utility models (e.g., Hakansson (1970), Hershey and Schoemaker (1980), Tversky and Kahneman (1991), Tversky and Kahneman (1992), Rabin (2000), Rabin and Thaler (2001), Levy and Post (2005), Wakker (2010)), although these concepts are not beyond criticism (Levy and Levy (2002b), see also Wakker (2003)).

Empirical studies that address whether alternative utility theories, particularly both versions of prospect theory (in which risk aversion is now captured by three different parameters) are effective and present in financial markets predominantly focus on selected features, such as the effects of reference points and loss aversion on financial decision making—features thought to be connected to individual investor trading behavior. For purposes of illustration, the pioneering Shefrin and Statman (1985) establishes a possible link between prospect theory and observed financial decisions, namely, the disposition effect, by basing its reasoning on these aforementioned characteristics, although this link has been questioned recently (e.g., Barberis and Xiong (2009), Linnainmaa (2010), Vlcek and Hens (2011) and Barberis (2012)).[4] Other studies focus on the interdependence between prospect theory and options exercise behavior (Heath et al. (1999), Poteshman and Serbin (2003)),

[3]For a critique on empirical findings for risk aversion based on their inconsistency with the equity premium, see Mehra and Prescott (1985), Mankiw and Zeldes (1991), Benartzi and Thaler (1995), Blake (1996), Kocherlakota (1996), Goetzman and Ibbotson (2005) and Mehra (2008)

[4]If individual preferences follow the predictions of prospect theory, phenomena such as the disposition effect should be observable in other environments. In fact, evidence for the disposition effect has been found among individual investors in the stock market (Schlarbaum et al. (1978), Ferris et al. (1988), Odean (1998), Weber and Camerer (1998), Odean (1998), Odean (1999), Garvey and Murphy (2004), Jordan and Diltz (2004), Lehenkari and Perttunen (2004), Frazzini (2006), Dhar and Zhu (2006)) and other environments, such as in the financial advice of stock brokers (Shapira and Venezia (2001)), the behavior of futures trades (Heisler (1994), Frino et al. (2004), Coval and Shumway (2005), as well as Locke and Mann (2005)), IPO trading volume (Kaustia (2004)), real estate markets (Genesove and Mayer (2001)), insurance contracts (i.e.Schoemaker and Kunreuther (1979), Camerer and Kunreuther (1989)), and observed risk behavior in laboratory environments for stocks (Weber and Camerer (1998), Oehler et al. (2003), Lee et al. (2008)) and monetary endowments (see Chui (2001)).

behavior of futures traders in real markets (Locke and Mann (2000), Locke and Mann (2005) and Coval and Shumway (2005)), and experiments (Haigh and List (2005); see also Harrison and Rutstrom (2009), who argue that these effects may also be consistent with constant absolute risk aversion (CARA) under variable risk aversion), as well as observed behavior in real estate markets (e.g., Genesove and Mayer (2001)).

In light of such preponderant, partly contradictory empirical evidence regarding the mathematical nature of individual investor preferences, comparison of utility functions is an ongoing topic in the experimental literature, as documented by Lattimore et al. (1992), Hey and Orme (1994) and Abdellaoui (2000). A major breakthrough in the quest to find underlying preferences is Hey and Orme (1994), who assess various parametric utility functions at the level of individual subjects using specific customized maximum likelihood procedures (Orme (1995)).[5] Maximum likelihood methods in general capture the idea that individuals err in their decision making, a presumption that is not explicitly recognized in previous methods of preference estimation (e.g., as in McCord and DeNeufville (1986), Currim and Sarin (1989) and others). These maximum likelihood methods have been enhanced and widely applied, generating a large number of studies addressing the best-fitting utility model, which predominantly find evidence favoring prospect theory models.[6] However, these results are criticized for their artificial setting concerning the payoff structure (Laury and Holt (2005)), as well as the way in which relevant information is presented (Kahneman and Tversky (1973)). Naturally, the question arises as to whether these well-established methods can be adapted and applied to financial data, assuming the decision process and data structure are similar to that obtained in the laboratory setting. In the next section, we present the frequently applied maximum likelihood approach for utility model selection in detail and elaborate on the econometric peculiarities that accompany financial data.

3. An econometric model of financial decision-making

Characterizing risk preferences of individual investors in financial markets typically involves extensive individual-level analysis, if one refrains from applying mixed models to capture the heterogeneity in preferences in the manner considered to be

[5]In contrast to the abundance of experimental evidence, studies directly addressing the best-fitting utility function in financial markets are surprisingly scarce. A notable exception is Blackburn and Ukhov (2006), who applies a modified pricing kernel approach from Jackwerth (2000) at the level of individual stocks to draw conclusions on the underlying utility function from the direction of the sign of the estimated pricing kernels. Blackburn and Ukhov find evidence for utility functions similar to those proposed by Friedman and Savage (1948), Markowitz (1952) and Kahneman and Tversky (1979), where classical utility functions, such as CRRA, prevail only in 3 out of 41 cases.

[6]For instance, Blondel (2002) fits linear, power, and exponential forms of utility to experimental data. This author finds strong evidence in favor of the power and the exponential function. In the same vein, Stott (2006) also find a best fit for power and exponential functions, while quadratic and linear specifications perform poorly. A further comparison between the first two functional forms shows, in line with Blondel (2002), that power specifications fit even better to the experimental data. Results from further experiments, in which maximum likelihood methods are applied, are mostly consistent with an inverse S-shaped probability weighting function (Wu and Gonzalez (1996), Wu and Gonzalez (1999), Abdellaoui (2000), Bleichrodt and Pinto (2000), Abdellaoui et al. (2005)); moreover, they are consistent with a concave value function in the domain of gains, corresponding to prospect theory, which is also backed in recent studies that address the best-fitting functional form (see, e.g., Wakker (2008)). The properties of diminishing sensitivity toward variations in areas of gains are further confirmed in Wakker and Deneffe (1996), Fox and Tversky (1998) and Fennema and van Assen (1999), whereas evidence of risk-seeking in the realm of losses is shown by Fishburn and Kochenberger (1979).

effective to unravel the respective utility functions and their parameterizations (e.g., Hey and Orme (1994), Laury and Holt (2005)). For this reason, the emphasis of empirical and experimental studies, particularly in the context of finance and asset markets, has shifted away from aggregated views that describe investors as a unified whole, usually modeled as a representative investor (Duffie (2001), Back (2012) or Munk (2013)), to disaggregated models in which the analysis is conducted on an individual level (for an early reference see Hensher and Johnson (1981)). In comparison to aggregate data, information on individual decisions is usually characterized by infrequent and discrete observations, by greater variation in each factor and less covariation among these factors (due to aggregation procedures), through summing the individual observations (Train (1986)).[7] A variety of econometric methods have been developed to address apparent discreteness, such as discrete choice models (Amemiya (1975), McFadden (1980), Amemiya (1981), Amemiya (1985)), popularized by Ben-Akiva and Lerman (1985), Train (1986), Train (2009) and advanced for estimation of nonlinear arguments, such as utility functions in a customized maximum likelihood model (Harrison and Rutstrom (2008) and de Palma et al. (2008)) as proposed by Hey and Orme (1994).

To test the effectiveness of such a likelihood-based model selection procedure, it is necessary to identify common and distinctive features in the design of laboratory and financial market environments. Experimental studies on utility functions, which apply such discrete choice models for preference measurement by customizing an underlying likelihood function, usually are similar in multiple aspects regarding underlying assumptions on the decision process. Narrow framing is one central assumption, as it allows definition of a finite and exhaustive set of alternatives via two mutually exclusive options, satisfying the requirements for a discrete choice set (Train (1986), Train (2009)) such that discrete choice models are applicable.[8] To define the set of choice options, it is commonly assumed that, for each investment decision of an individual investor, the respective choice set is spanned by a risky asset and an investment in an assumed riskless money market account in place of a representative riskless asset, yielding a gross return of $R_{f,t}$. Given the usual lottery-type design, the price of the risky asset, essentially any stock traded by the investor over the respective period, is assumed to be subject to a stochastic binomial process (Cox et al. (1979) and Rendleman and Bartter (1979), also Hull and White (1988)) in which two distinct states S of the world can be identified, yielding a gross return of $R_{S,t}$.[9]

[7] This is important, because the precision of estimation generally increases with sample size and variance of the variables entering the model, and decreases with its covariance. Further, standard econometric tools, such as regression analysis, implicitly assume a set of continuous variables– an assumption appropriate for aggregated models but that seems to be problematic when underlying factors on an individual investor level are the focus of interest.

[8] Empirical finance studies indicate that investors allocate different streams of income, such as dividends and cash flows resulting from corporate actions and other stocks (Shefrin and Statman (1984), Baker and Wurgler (2004)), to different mental accounts (Thaler (1985)). Further, the tendency to evaluate risky lotteries separately, known as narrow framing (Barberis and Huang (2001), Barberis et al. (2001), Barberis et al. (2001), Berkelaar et al. (2004), Gomes (2005), Barberis and Huang (2009)) is in line with Shefrin and Statman (1985), complementing recent studies on individual investors that examine trading decisions for each stock separately (see Odean (1998), Odean (1999), Barber and Odean (2000), Barberis and Huang (2001), Grinblatt and Keloharju (2001a), Grinblatt and Keloharju (2001b), Barber and Odean (2002), Dhar and Kumar (2002), Hong and Kumar (2002), Zhu (2002), Grinblatt and Han (2005), Lim (2006), Frazzini (2006), Barber and Odean (2008)).

[9] In the upside state U, associated with some unknown physical probability $p_t > 0$, indicated by an index t for time, the stock price rises and yields an upside return $R_{U,t} > 1$, whereas in the

The expected payoffs and accrued returns of the risky asset constitute a central, distinctive feature between, on the one hand, financial markets, in which parameterization of the underlying return distribution is unknown and intermediate (paper) gains and losses are followed up, and, on the other hand, experiments, in which gains and losses are not carried forward to the next lottery task to avoid strategic hedging behavior and where payoffs of the lotteries are clearly presented to the study participant. There is ample evidence showing that investors in financial markets experience difficulty in recognizing and learning the true but unknown market parameterizations, especially if they vary over time (Ehm et al. (2012)). To approximate financial payoffs, investors form their expectations on the outcomes of the risky asset by applying several mental shortcuts. DeBondt (1993) mentions that investors may consider recent past returns to be representative in formulating their expectations about the future to approximate financial payoffs (Kahneman and Tversky (1973), for evidence from stock markets see Andreassen (1987), Andreassen (1988) and Andreassen and Kraus (1990)). Considering that this mental pattern transforms the payoffs $R_{S,t}$ into expected profits extrapolated from the past over some lookback period, it generates an implicit correlation of these values as time proceeds and beliefs on market parameters are updated in each t.[10]

Studies on empirical dynamic programming suggest that individual investors face computational difficulties in determining the optimal trading strategy (see Eckstein and Wolpin (1989), Rust (1994) and Adda and Cooper (2003) for surveys). According to these studies, investor behavior is more likely reconcilable with a discrete decision process (Rust (1992)), which is found to be reflected in the stock market (Thaler et al. (1997) and Gneezy and Potters (1997), see also Normandin and St-Amour (2008)).[11] To formalize the decision process, the utility an individual investor obtains from a money market account is denoted as $V_k(W_t, R_{f,t}|\boldsymbol{\theta}_k)$ for utility model k, whereas the utility resulting from the risky asset is given the expression $V_k(W_t, R_{S,t}|\boldsymbol{\theta}_k)$ with $S \in \{U; D\}$. This allows introduction of a parameter set $\boldsymbol{\theta}_k$ to represent the utility-specific parameters of utility model-type k, which in turn, due to the discretionary (myopic) decision process, leads to the behavioral assumption that the investor invests a positive amount in the risky asset if

$$V_k(W_t, R_{S,t}|\boldsymbol{\theta}_k) \geq V_k(W_t, R_{f,t}|\boldsymbol{\theta}_k) \tag{3.1}$$

holds (see for this approach, e.g., Currim and Sarin (1989)). Discrete choice models as in Train (1986), Rust (1994) and Train (2009) are constructed around the assumption that only a minority of attributes that drive purchase and selling decisions in financial markets are observable. Consequently, the utility function has

downside state D with corresponding probability $1 - p_t$, the stock declines, generating a downside return $0 \leq R_{D,t} < 1$.

[10]In doing so, we distinguish from representativeness bias, in which investors base their judgments on stereotypes and seek out patterns in returns or prices (Weber and Camerer (1998), Shefrin (2008)). The intuition here is that, due to extrapolation bias with short horizons, investors may buy stocks whose prices have recently increased, especially when following a myopic trading strategy, which contradicts mean reversion expectation (Zuchel (2001)). This is backed in empirical studies e.g., Grinblatt and Keloharju (2000) and Kaustia (2010). These authors find that Finnish investors bought past winners and sold past losers, thus revealing a trend-following trading strategy, which is not consistent with an expectation of mean-reverting stock prices (see also Kaniel et al. (2008)). Dhar and Kumar (2002) investigate price trends of stocks bought by more than 62,000 households using discount brokerages, and conclude that investors prefer to buy stocks that have recently enjoyed an abnormal positive return.

[11]Further, under generalized utility concepts, dynamic programming with nonlinear decision weights can generate suboptimal and time-inconsistent results, as shown by Machina (1989) for non-expected utility in general, Nielssen and Jaffray (2004) for rank-dependent utility RDU and Barberis (2012) and Ebert and Strack (2012) for cumulative prospect theory CPT, respectively.

the additively separable decomposition

$$V_k(W_t, R_{S,t}|\boldsymbol{\theta}_k) = U_k(W_t, R_{S,t}|\boldsymbol{\theta}_k) + \epsilon, \tag{3.2}$$

which, combined with (3.1), implies that the investor holds the stock if the difference in utilities, abbreviated as $\Delta_t(U_k|\boldsymbol{\theta_k})$, is positive. There are two main purposes of the stochastic component ϵ: first the error term should fully capture hidden factors that affect the observed variations in the attractiveness of the respective stock without the necessity to explicitly model other (potentially unobservable) variables or data imperfections (Cramer (1986), also Rust (1994)). Second, without adding the error term ϵ to $\Delta_t(U_k|\boldsymbol{\theta_k})$, the lack of error in behavior yields imprecise estimates for $\boldsymbol{\theta_k}$. To illustrate the latter argument, consider the case of an investor whose decision is based exclusively on $\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}})$. While for each set of $\boldsymbol{\theta_k}$ there exists a unique optimal decision for the investor, the converse is not true. For any set of optimal decisions reflected in the trading sequence of this investor, there exists a set of parameters consistent with those decisions. In terms of the likelihood function, if no error term is added to $\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}})$, then for any parameter set $\boldsymbol{\theta_k}$, the observations are either consistent with the utility model or not. If they are not consistent with the utility model under consideration, the likelihood is zero and $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ is (theoretically) $-\infty$. However, if the observations are in accord with the utility model, then the likelihood is 1 and $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ is zero. As a consequence, $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ oscillates between zero and $-\infty$. Therefore, modeling and estimation of an individuals trading behavior should contain an additional element ϵ. Note that, based on the predictability of $R_{f,t}$ and the fact that the utility of the risk-free investment carries no uncertainty per se, ϵ stemming from the risk-free asset is assumed to be zero.[12]

To obtain the maximum likelihood function, conditional choice probabilities are derived given the stochastic properties of the error term, which is frequently assumed to be normally distributed (Hey and Orme (1994) and Carbone and Hey (2000)) as $\epsilon \sim N(0, \sigma_\epsilon^2)$ with density according to $\phi(\epsilon) = (2\pi\sigma_\epsilon^2)^{-\frac{1}{2}} e^{-\frac{1}{2}(\epsilon/\sigma_\epsilon)^2}$. Derivation of the conditional choice probabilities requires defining an index $I_{k,t} := I[\Delta_t(U_k|\boldsymbol{\theta_k}) + \epsilon \geq 0]$, taking the value 1 if the condition in the brackets is met and zero otherwise.[13] The resulting choice probabilities are usually denoted as $\Phi\left(\Delta_t(U_k|\boldsymbol{\theta_k})/\sigma_\epsilon\right)$, where by the finiteness of the choice set, the probability of investing in the riskless asset is $1 - \Phi\left(\Delta_t(U_k|\boldsymbol{\theta_k})/\sigma_\epsilon\right) = \Phi\left(-\Delta_t(U_k|\boldsymbol{\theta_k})/\sigma_\epsilon\right)$. In this case, Φ denotes the cumulative normal density function, but it may be substituted by any other distribution, such as the lognormal (as in Booij et al. (2009)) or logistic distribution (Harrison and Rutstrom (2008), Train (2009)). The term σ_ϵ^2 denotes the (heteroscedastic) variance of the error term on a daily basis, estimated as a nuisance parameter (Pawitan (2001)) along with $\boldsymbol{\theta}_k$ to absorb potential effects resulting from the lagged structure of the estimated market parameters within

---

[12]This is a minor technicality, as it avoids the necessity to evaluate all elements of the covariance matrix of errors (see Train (2009)). Further note that by assuming hedonic framing, the covariance in errors among stocks traded in a portfolio can be ignored.

[13]The probability (Rust (1994)) of buying or holding the risky asset is thus given as

$$\begin{aligned}
p(\Delta_t(U_k|\boldsymbol{\theta_k}) \geq 0) &= p(I[\Delta_t(U_k|\boldsymbol{\theta_k}) + \epsilon \geq 0] = 1) \\
&= \int_{-\infty}^{\infty} I[\Delta_t(U_k|\boldsymbol{\theta_k}) + \epsilon \geq 0]\phi(\epsilon)d\epsilon \\
&= \int_{-\infty}^{\frac{\Delta_t(U_k|\boldsymbol{\theta_k})}{\sigma_\epsilon}} \phi(\epsilon)d\epsilon,
\end{aligned}$$

satisfying the conditions if $\Delta_t(U_k|\boldsymbol{\theta_k}) \to \infty$, the choice probability converges to unity and approaches zero if $\Delta_t(U_k|\boldsymbol{\theta_k}) \to -\infty$.

$\Delta_t(U_k|\boldsymbol{\theta_k})$ (Dhrymes (1971), Cramer (1986)) and of its error terms.[14] Combining the normal distribution of the error term with the binary choice feature of the discrete choice setting leads to a likelihood function $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ similar to a nonlinear in arguments probit model (see Thurstone (1927), case V, for an early reference from the field of psychometrics and Marschak (1960) for a transition in terms of utility).[15] Experimental and empirical studies, which commonly apply likelihood methods to identify utility functions, implicitly use the convenient properties of maximum likelihood estimators (such as the frequently cited Jullien and Salanie (2000)), whereupon maximizing $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ for each of the K_k elements of $\boldsymbol{\theta_k}$ and σ_ϵ provides estimators $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}|\boldsymbol{n},\boldsymbol{t}}$ for a given sample size, indicated by the number of stocks n and trading days t, which are consistent, asymptotically efficient, and moreover, asymptotically normally distributed.[16]

To identify the best-fitting underlying utility function of type k, insights from likelihood theory provide the key to selecting the utility model that best explains the observed data. According to Fisher (1922) and Kullback (1968), the maximized likelihood function $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ contains information for each utility model k on the relative fit of this model to observed data. To distinguish the k utility models from one another, classical likelihood theory suggests that the utility model with the highest maximized likelihood $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ best fits the observed data (Kullback (1968), Akaike (1973), Schwarz (1978), Amemiya (1980), Pawitan (2001) and Burnham and Anderson (2004)). However, note that $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ does

---

[14]In addition, the economic intention of the heteroscedasticity feature of the error term ϵ allows us to disintegrate factors with varying impact, dependent on whether the decision at hand is a purchase or a sale, thus agreeing with Odean (1999), Glaser and Weber (2007) and Statman et al. (2006), whereupon purchase decisions may be motivated by factors other than sell decisions. For example, overconfident investors may suffer from biased beliefs about the anticipated returns they expect to generate by trading stocks even if these investors had average performance in the past (Odean (1999), Barber and Odean (1999) and Glaser and Weber (2007)), as such investors are induced to buy stocks more readily.

[15]The overall likelihood function of an investor of utility type k can accordingly be expressed as $\log L(\Delta_t(U_k|\boldsymbol{\theta_k})) = \sum_{t\in T}\sum_{I\in I_{k,t}} I_{k,t}\log p_{I_{k,t}}(\Delta_t(U_k|\boldsymbol{\theta_k}))$ in which $p_{I_{k,t}}(\Delta_t(U_k|\boldsymbol{\theta_k}))$ denotes the respective conditional probabilities as defined above. Given the binary choice assumption, the log-likelihood function can be explicitly written as $\log L(\Delta_t(U_k|\boldsymbol{\theta_k})) = \sum_{t\in T}\log\left(\left[\Phi\left(\frac{\Delta_t(U_k|\boldsymbol{\theta_k})}{\boldsymbol{\sigma_\epsilon}}\right)\right]^{I_{k,t}}\left[\Phi\left(-\frac{\Delta_t(U_k|\boldsymbol{\theta_k})}{\boldsymbol{\sigma_\epsilon}}\right)\right]^{1-I_{k,t}}\right)$ in which we omit constant combinatorial terms, since they add no further information about $\boldsymbol{\theta}_k$. An alternative expression, which explicitly recognizes the dichotomy of $I_{k,t}$, is obtained if the log-likelihood function is decomposed as $\log L(\Delta_t(U_k|\boldsymbol{\theta_k})) = \sum_1 \log\Phi\left(\frac{\Delta_t(U_k|\boldsymbol{\theta_k})}{\boldsymbol{\sigma_\epsilon}}\right) + \sum_0 \log\Phi\left(-\frac{\Delta_t(U_k|\boldsymbol{\theta_k})}{\boldsymbol{\sigma_\epsilon}}\right)$, which immediately carries over to the notation of the score vector and the Hessian matrix. Expressing the log-likelihood function this way is helpful in organizing the computations but cumbersome if one wishes to derive the information matrix such that the notation used in this paper appears to be more convenient, although it seems complex at first sight.

[16]In the appendix, we provide further details concerning these properties of the maximum likelihood methodology in the context of utility models for the interested reader. Concerning these estimators, given that certain regularity conditions of $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ hold, Lehmann (1983) elaborates that $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}|\boldsymbol{n},\boldsymbol{t}}$ has certain convenient properties, that is, its estimators cannot be located on a boundary without violating the regularity of the likelihood function. Further, the information matrix is bounded and positive, and thus satisfies the characteristics of a variance measure given that a second-order Taylor expansion of $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ is sufficient and valid. Experimental design can ensure that the presupposition according to which the single likelihood functions are indeed independent (e.g., Post et al. (2008) and others), is maintained, and consequently the asymptotic features of $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}|\boldsymbol{n},\boldsymbol{t}}$ are preserved. For instance, current studies on preferences in game shows allow for carrying forward gains and losses (Post et al. (2008)) in a way that preserves the likelihood properties but that may require simulation methods to establish the empirical distribution function upon which the likelihood approach is constructed.

not necessarily have to be exactly zero for the perfectly fitting model: Consider the case where the true model is $y \sim N(0, 1)$ and we fit the model $y \sim N(\mu, \sigma)$ by estimating μ and σ by maximum likelihood. Further assume that data are generated by sampling from a standard normal distribution using simulation. Even if maximizing the likelihood identifies zero and 1 as parameter estimates, the $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ of each observation will not be zero. It will rather be the standard normal density evaluated at y, so the total $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ will not be zero in general.[17]

Ranking all utility models according to $\log L(\Delta_t(U_k|\boldsymbol{\theta}_{\boldsymbol{k}}))$ and choosing the model with the highest likelihood as the model selection criterion is usually not recommended for utility model selection, since the maximized likelihood function is subject to overfitting, tendentially favoring multiparameter utility models such as Prospect Theory (3 parameters) through to expected utility (one parameter) (Carbone and Hey (1994), Hey and Orme (1994), Carbone and Hey (1995) and Stott (2006)). Rather, the literature on model selection suggests sorting utility models according to the *Akaike Information Criterion* (*AIC*), which controls explicitly for a varying number of parameters instead of using the maximized $\log L(\Delta_t(U_k|\boldsymbol{\theta}_{\boldsymbol{k}}))$ (Akaike (1973), Akaike (1974), Bozdogan (2000), Pawitan (2001) and Burnham and Anderson (2004)). The *AIC* is commonly expressed as

$$AIC = -\frac{2\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))}{nt} + \frac{2K_k}{nt}, \tag{3.3}$$

according to Akaike (1974) in the representation of Amemiya (1980), where dividing by nt, the number of observations in terms of trading days t and traded stocks n, corrects for the different number of observations, and where K_k denotes the rank of $\boldsymbol{\theta}_{\boldsymbol{k}}$, representing the number of parameters to be estimated in utility model k. Due to the general finiteness of our data set, we apply the *corrected Akaike Information Criterion* (*AICC*), defined by

$$AICC = -\frac{2\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))}{nt} + \frac{2K_k}{nt} + \frac{2K_k(K_k+1)}{nt(nt-K_k-1)}, \tag{3.4}$$

as first proposed by Sugiura (1978) for Ordinary Least Square (OLS) regressions (for a discussion of the original version of the Akaike Information Criterion *AIC* and *AICC* as model selection criteria, see Burnham and Anderson (2004)), which replaces the penalty term of *AIC* by its exact term for bias adjustment, resulting in a greater penalty for models with additional parameters in comparison to the original *AIC*. We provide a more formal outline in the appendix.

4. A Simulation Study of Likelihood-based Utility Model Selection

Thus far, we have sketched the elements of the likelihood theory as applied in experimental studies addressing individual preferences and argued that applying maximum likelihood estimation provides asymptotically efficient and unbiased estimators. Moreover, making use of the characteristics of the likelihood function $\log(L(\Delta_t(U_k|\boldsymbol{\theta}_{\boldsymbol{k}})))$, particularly the structure of its surface—its elevation and its steepness, allows filtering for the best-fitting model. As pointed out, experimental literature often assumes that each single likelihood function of $\log(L(\Delta_t(U_k|\boldsymbol{\theta}_{\boldsymbol{k}})))$

[17]Note that even if we remove randomness from our data (i.e., the variation of the error term is zero), the likelihood of each of these identical observations will be the normal density evaluated at zero, which is not zero but a positive number (approx. 0.399). Thus, the overall logarithm of the likelihood is $nt \cdot \ln(0.399) > 0$ even though we specified the correct model. Note that in our case, the observations are not equal to their means due to $\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}})$. I am grateful to John Hey, who made me aware of this.

is independent. Given the tendency to extrapolate past moments into the future and to keep track of paper gains and losses, the interlacing of the various functionals of the likelihood function introduces an implicit dependency within and across $\log(L(\Delta_t(U_k|\boldsymbol{\theta_k})))$, resulting in deficiencies in the surface of the likelihood function. In this case, the extent and direction of these effects are unclear and requires investigation into how these adverse effects transmit to estimators for $\hat{\boldsymbol{\theta}}_{\boldsymbol{k|n,t}}$ and by how much these imprecisions negatively affect identification of the underlying utility mode.

To investigate the reliability of likelihood-based model selection procedures given financial data and the identification of factors detrimental to this purpose, we conduct a conceptually simple simulation study that consists of four steps. First, to control effects stemming from the financial times series, we simulate a series of prices and their returns with known stochastic characteristics and market parameters. In the second step, we change perspective and take the position of an individual investor with known utility functions and risk parameters, who faces this set of hypothetical stocks, represented by the time series of returns. We estimate the unknown market parameters needed for the third step, in which the investor decides whether to buy, hold, or sell the hypothetical stocks according to inequality (3.1). From this step, we obtain a set of trading sequences (one for each hypothetical stock), which we use in the final step to identify the underlying utility function and infer its parameterization used to generate the trading sequence.

STEP 1: Simulation of a Set of Time Series for Given Market Parameters:

We simulate realizations of a sequence of returns based on a pre-specified stochastic Markovian process to avoid inherent autocorrelation of our time series up front. In total, we generate a set of 100 hypothetical stocks by simulating a series of identically and independently distributed (logarithmic) returns spanning 312 days each, denoted as

$$\{\left(R_{t,t+1}|\mu\Delta(t), \Phi^{-1}(0,\sigma)\sqrt{\Delta(t)}\right) : t = 1, 2, \ldots, 312\}.^{18}$$

The time series of returns is modeled as discrete time Geometric Brownian motion by the inverse of the cumulative standard normal distribution $\Phi^{-1}(0,\sigma)$ plus a trend μ, where the time step is set to $\Delta(t) = 1$. To calibrate the trend variable μ and standard deviation σ, we are guided by the results of Dimson et al. (2000), Dimson et al. (2003) as well as publicly available data for the German stock market, and we set the daily mean $\mu = 1.0004$ and volatility $\sigma = 0.0247$.[19] With regard to the risk-free rate, we opt for a fixed daily net return of 0.0001, corresponding to a return of 3.67% assuming annual compounding. For the inverse of the cumulative standard normal distribution of each simulated return sequence, we define the seed values of the random number generator in Stata Version 10.1 (Cameron and Trivedi (2005),

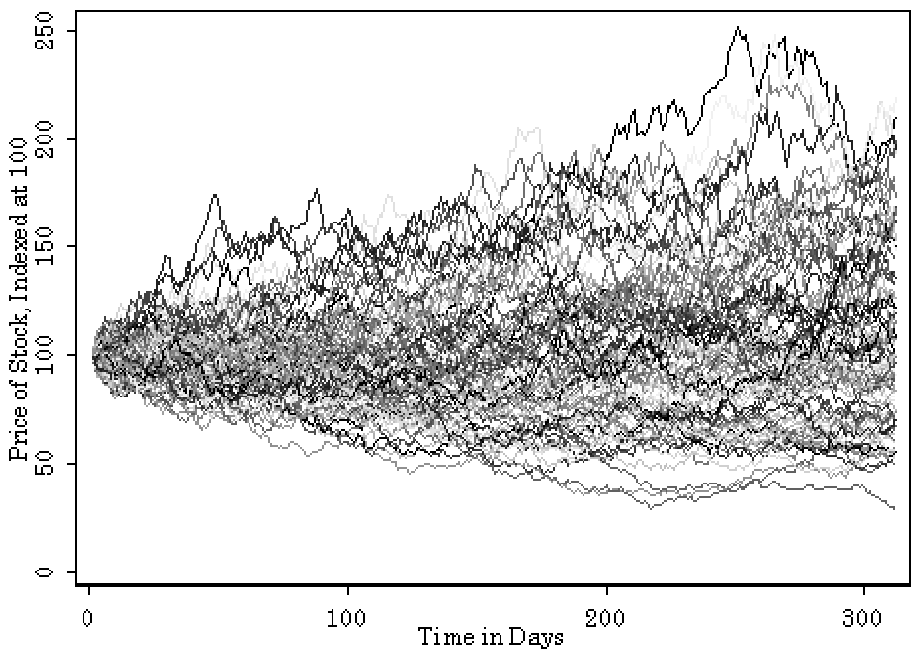
[18] Due to the rolling-window procedure of Step 2, where the unknown market parameters are estimated, we extended the time series of each stock by 60 days such that (technically) the length is 312; the final time series after Step 2 thus spans 252 days, which is a common approximation for the number of trading days in one year.

[19] Gross returns and risk premia (see Fama and French (1993)) are obtained from *http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#International* as accessed on 02.12.2012 to obtain values for μ, which we use in the second step to match σ with historical gross appreciation rates to make simulated returns correspond more closely to their empirical equivalents. However, Bruckner et al. (2015) note that for non-US markets, Fama–French factors cannot easily be exported, as these differ among the relevant data providers. This is a minor issue in our work here, as using the factors provided by the database mentioned in Bruckner et al. (2015) provides virtually similar market parameters.

Baum (2006)) by a Halton sequence (see Halton (1960)) based on the prime 11 to ensure the full spectrum of random numbers for the stochastic process is covered.[20] Since the draws from a Halton sequence tend to be negatively correlated with previous draws, this feature is beneficial to reduce simulation error, consequently less simulations are required to obtain statistically reliable results (Train (2009)). For example, Bath (2001) find that a sequence of 100 Halton draws provides more precise results for a mixed logit estimation than 1000 random draws. Table (1) illustrates the simulations conducted in Step 1.

Figure 1. Simulated Time Series of Prices

The figure below illustrates the development of simulated returns for 100 draws used in this simulation study. Market parameters are fixed at $\mu = 1.0004$ and $\sigma = 0.0247$, respectively. For this table, prices are calculated on a daily basis using the simulated returns generated in Step 1. The starting price of each stock is set to 100.



STEP 2: Estimation of Market Parameters from Simulated Time Series:

In the second step, we take the position of an individual investor and estimate the market parameters needed to fill the respective utility functionals. For this purpose, estimators for the mean $\hat{\mu}_t$ and volatility $\hat{\sigma}_t$ were obtained using a rolling-window procedure with a lookback horizon of $l = 60$ days. Values of upside and downside returns $\hat{R}_{D,t}$ and $\hat{R}_{U,t}$ were then derived from $\hat{\mu}_t$ and $\hat{\sigma}_t$ for each t, where

[20] A Halton sequence based on prime h is defined as $s_{t+1} = \{s_t, s_t + \frac{1}{h^t}, s_t + \frac{2}{h^t}, \ldots, s_t + \frac{h-1}{h^t}\}$, where s_{t+1} denotes the sequence at iteration $t+1$ of length h. The application of a Halton sequence constitutes a well-defined draw spanning the standard uniform density, as it systematically fills in the unit interval.

$\hat{R}_{U,t} = e^{\frac{\hat{\mu}_t}{l}+\sqrt{\frac{1-\hat{p}_t}{\hat{p}_t}\frac{\hat{\sigma}_t^2}{l}}}$ and $\hat{R}_{D,t} = e^{\frac{\hat{\mu}_t}{l}-\sqrt{\frac{\hat{p}_t}{1-\hat{p}_t}\frac{\hat{\sigma}_t^2}{l}}}$, respectively.[21] Corresponding upside probabilities $\hat{p}_t$ are derived by averaging observed up- and down-ticks, given a change in prices occurs, since the true probability p of the underlying binomial process is unknown to the individual investor (similar Weber and Camerer (1998)).[22]

STEP 3: Calculation of Trading Sequences for each Utility Function:

In the third step, given a vector $\boldsymbol{\theta_k}$ of predefined risk preference parameters for the respective utility function of type k of the simulated investors (one investor for each utility model k), we generate a variety of artificial trading histories in terms of roundtrip sequences (Shapira and Venezia (2001)), denoted accordingly as

$$\{(\hat{R}_{U,t}, \hat{R}_{D,t}, R_{f,t}, \hat{p}_t|\boldsymbol{\theta_k}) : I_t \in [0;1]\ \forall t = 1, 2, \ldots, 252\}.$$

The mathematical details of the utility functions used in this simulation study are given in the Appendix.[23] To calculate accrued returns, we set the gross realized return to unity if the index is zero and add the realized logarithmic returns to W_t if the index is 1 until the last day before it switches back to zero, which we take as realized return being invested in the risk-free asset.[24]

To substantiate the set of utility functions that specifies $\Delta_t(U_k|\boldsymbol{\theta_k})$, we generate the trading patterns for an investor characterized as a *expected utility theory (EUT)*-type investor of the CRRA form (A.3) with the following settings:

$$\begin{aligned}&\{U_{EUT}(W_t, R_{U,t}, R_{D,t}, R_{f,t}, p_t|\boldsymbol{\theta}_{EUT} \in \{\delta\})\},\\ &\text{where } \boldsymbol{\theta}_{EUT} : \{\delta \in \{2\}\},\end{aligned}$$

where we deliberately set W_t to zero at $t = 1$ and equal to the accrued return $\hat{R}_{S,t}$ from the risky investment for $t \in \{2, 3, \ldots, T\}$ for simplicity. Specifications of the utility functionals are the commonly applied *CRRA*-form and exponential power utility (*EXPO*) as proposed in Saha (1993). We contrast this simulated trading behavior with those generated by generalized expected utility and generate roundtrip sequences for an investor of *rank-dependent utility (RDU)*-type with the

[21] This is a standard procedure (Ingersoll (1987)) and is widely applied, as similar expressions can be found in Johnson et al. (1997), Barberis and Xiong (2009), Ebert and Strack (2009) and Johnson et al. (2012).

[22] Note that this behavior can imply rational behavior, since $\hat{p}_t$ also serves as a maximum likelihood estimator for the underlying true but unobservable probability p given a binomial distribution $\hat{p}_t = \binom{t+1}{j}p^{t-j+1}(1-p)^{j-1}$. The solution for the estimator $\hat{p}_t = \frac{t-j+1}{t}$ is derived by taking the logarithm and differentiating with respect to p.

[23] In detail, to generate a sequence of trades that resembles the lottery task features as commonly framed in experimental studies, we define an index $I[\Delta_t(U_k|\boldsymbol{\theta_k}) + \epsilon \geq 0]$ similar to the index used to obtain the choice probabilities in the previous chapter, yielding a series of zeros and ones dependent on whether the investor holds the stock on a particular day. These sequences allow us to define so-called roundtrips, defined as rows of ones similar to the definition of Shapira and Venezia (2001), since the various trade-inventory rules Odean (1998) coincide for binary choice situations. The effects of various accounting principles on the results is outside of the scope of this paper. Our results are robust to different parameter settings, as we also simulate trading sequences for different parameter values for all utility models k, and find virtually identical results.

[24] According to our definition, the roundtrip sequence comprises only the row of ones under the fiction that a new mental account is opened if the investor sells and buys at a later point in time. This avoids dilution effects due to accruing risk-free rates in the estimation and having different starting values for W_t, with the result that risk preferences would be measured at different locations of the utility functional, which we suspect yields different estimates if W_t is high.

Table 1. Utility Models and Parameters $\boldsymbol{\theta_k}$ used for Trading Sequences

This table describes the parameter set $\boldsymbol{\theta_k}$, its settings used for generating the trading sequences, and the key reference. Expected utility models are denoted as EUT, and Rank-dependent Utility is denoted as RDU. For Simple Prospect Theory (Tversky and Kahneman (1992)), we use the notation SPT, whereas Cumulative Prospect Theory (Tversky and Kahneman (1992)) is denoted as CPT. Further, for EUT and RDU, we use the notation $CRRA$ for utility functionals with constant relative risk aversion and $EXPO$ to denote exponential power utility functions (Saha (1993)). For SPT and CPT, we use the notation $POWR$ to indicate models with kinked power-functionals and $CRRA$ if a power functional is used to specify the prospect value functional (Gomes (2005)); in addition, $DHG0$ denotes value functionals as defined in DeGiorgi and Hens (2006).

		Set $\boldsymbol{\theta_k}$		Interpretation	Key Reference
EUT	**CRRA**	δ	$= 2$	Risk Aversion	Gollier (2001)
	EXPO	δ	$= 2$	Risk Aversion	Saha (1993)
		ρ	$= 1$	Scaling Parameter	Saha et al. (1994)
RDU	**CRRA**	δ	$= 2$	Risk Aversion	Quiggin (1982), Quiggin (1993)
		γ	$= 0.65$	Weighting Parameter	Quiggin (1982), Tversky and Kahneman (1992)
	EXPO	δ	$= 2$	Risk Aversion	Saha (1993)
		ρ	$= 1$	Scaling Parameter	Saha et al. (1994)
		γ	$= 0.65$	Weighting Parameter	Quiggin (1982)
SPT	**CRRA**	α	$= 0.88$	Risk Sensitivity	Gomes (2005)
		λ	$= -2.25$	Loss Aversion	Tversky and Kahneman (1991)
		γ	$= 0.65$	Weighting Parameter	Quiggin (1982), Tversky and Kahneman (1992)
	POWR	α	$= 0.88$	Risk Sensitivity	Kahneman and Tversky (1979)
		λ	$= -2.25$	Loss Aversion	Tversky and Kahneman (1991)
		γ	$= 0.65$	Weighting Parameter	Quiggin (1982), Tversky and Kahneman (1992)
	DGH0	$\alpha^{\pm}$	$= 0.20$	Scaling Parameter	DeGiorgi and Hens (2006)
		λ^{+}	$= 6.52$	Scaling Parameter	DeGiorgi and Hens (2006)
		λ^{-}	$= 14.7$	Scaling Parameter	DeGiorgi and Hens (2006)
		γ	$= 0.65$	Weighting Parameter	Quiggin (1982), Tversky and Kahneman (1992)
CPT		α	$= 0.88$	Risk Sensitivity	Gomes (2005)
	CRRA	λ	$= -2.25$	Loss Aversion	Tversky and Kahneman (1991)
		γ	$= 0.65$	Weighting Parameter	Quiggin (1982), Tversky and Kahneman (1992)
	POWR	α	$= 0.88$	Risk Sensitivity	Tversky and Kahneman (1992)
		λ	$= -2.25$	Loss Aversion	Tversky and Kahneman (1991)
		γ	$= 0.65$	Weighting Parameter	Quiggin (1982), Tversky and Kahneman (1992)
	DGH0	$\alpha^{\pm}$	$= 0.20$	Scaling Parameter	DeGiorgi and Hens (2006)
		λ^{+}	$= 6.52$	Scaling Parameter	DeGiorgi and Hens (2006)
		λ^{-}	$= 14.7$	Scaling Parameter	DeGiorgi and Hens (2006)
		γ	$= 0.65$	Weighting Parameter	Quiggin (1982), Tversky and Kahneman (1992)

following settings

$$\begin{aligned}&\{U_{RDU}(W_t, R_{U,t}, R_{D,t}, R_{f,t}, p_t|\boldsymbol{\theta}_{RDU} \in \{\delta, \gamma\})\},\\ &\text{where } \boldsymbol{\theta}_{RDU} : \{\delta \in \{2\}; \{\gamma \in \{0.65\}\}\}\end{aligned}$$

with decision-weightings according to Quiggin (1982) (*QU82*) and Tversky and Kahneman (1992) (*KT92*), regarding the utility functional, we use $CRRA$ and *EXPO*. To test the sensitivity of the likelihood-based model selection with respect to non-standard utility functions, we also consider non-expected utility investors of *simple prospect theory (SPT)*-type (Kahneman and Tversky (1979)) of the form

$$\begin{aligned}&\{U_{SPT}(W_t, W_{RP}, R_{U,t}, R_{D,t}, R_{f,t}, p_t|\boldsymbol{\theta}_{SPT} \in \{\alpha, \gamma\,\lambda)\},\\ &\text{where } \boldsymbol{\theta}_{SPT} : \{\alpha \in \{0.88\}; \{\gamma \in \{0.65\}; \{\lambda \in \{-2.25\}\},\end{aligned}$$

where we consider W_{RP} to be located at the purchase price of the risky asset without inherent dynamics.

We accompany the results from the generalized expected utility models by the trading sequences of a *cumulative prospect theory (CPT)*-type investor with the following settings

$$\{U_{CPT}(W_t, W_{RP}, R_{U,t}, R_{D,t}, R_{f,t}, p_t|\boldsymbol{\theta}_{CPT} \in \{\alpha, \gamma\, \lambda))\},$$

which covers approximately the upper and lower bounds of estimated parameterizations from experimental studies. As for the *SPT* case, we use a decision weight according to Quiggin (1982) and Tversky and Kahneman (1992) for the *CPT* function. For *SPT* and *CPT*, we specify the functional form of the value functional according to a power function (Kahneman and Tversky (1979)) and contrast the results with an *CRRA* functional (as in Barberis et al. (2001) and Gomes (2005)), exponential power functional (*EXPO*) and a piecewise negative exponential power functional (*DGH0*) as defined in DeGiorgi and Hens (2006). Table (1) provides an overview of the utility models used and their parameter settings in this simulation study.

As *SPT* and *CPT* type investors are expected to be sensitive with respect to the prospect horizon, such that we calculate the respective utilities according to a predetermined forecast period, denoted as τ. Due to the fact that *CPT* and *SPT* coincide under $\tau = 1$ and are thus indistinguishable, we set $\tau = 20$. To avoid side relations among the elements of $\boldsymbol{\theta_k}$ and to model decision errors as, e.g., Carbone and Hey (1994) and Carbone and Hey (1995), we add a normally distributed error term ϵ to the difference of the respective utility of the risky stock and the riskless money market account, for which we deliberately set the standard deviation of the error equal to 0.01. On each day, the hypothetical investor invests in the risky stock whenever $\Delta_t(U_k|\boldsymbol{\theta_k}) + \epsilon \geq 0$, which seems justified given the evidence regarding the day trading in stock markets (e.g., Jordan and Diltz (2004) and Linnainmaa (2005)). If this condition is satisfied, we set the indicator $I[\Delta_t(U_k|\boldsymbol{\theta_k}) + \epsilon \geq 0]$ to 1, zero otherwise. Repeating this for all 252 trading days of each time series of Step 1 yields the required trading sequences for the maximum likelihood estimation in Step 4.

STEP 4: Estimation and Selection of Utility Function from Trading Sequences:

In the final step, we take the trading sequences from Step 3, evaluate the likelihood function $\log(L(\Delta_t(U_k|\boldsymbol{\theta_k})))$ for each utility type k and perform the model selection. In detail, for each of the 100 trading sequences of each investor of utility type k, we loop through all conceivable utility functions, for which we estimate the associated risk preference vector $\boldsymbol{\theta_k}$ and the standard deviation of the error term σ_i. Regarding the latter, we transform σ_ϵ in the likelihood estimation by an exponential function (as described by Rabe-Hersketh and Everitt (2004), Chapter 13) to ensure that the ascertained estimator is strictly positive. We recover the estimator for σ_ϵ and the associated standard errors using the `nlcom` command in Stata version 10.1 (for details on the maximum likelihood estimation see Gould et al. (2006)). For each trading sequence of an investor of utility type k, every time the likelihood function is evaluated for each conceivable utility function, we rank the utility models by sorting the corrected Akaike criterion (*AICC*) before proceeding with the next trading sequence for this investor.

For linear-in-parameter logit models, McFadden (1974) shows that there exists a unique and global maximum of the likelihood function; however, due to utility models such as *RDU*, *SPT*, and *CPT*, and therefore the nonlinear structure of $\Delta_t(U_k|\boldsymbol{\theta_k})$, it cannot be expected that the likelihood function we face is well-behaved and can likewise be characterized by a unique global maximum. This has several consequences for the results of a model selection procedure: If it cannot be ruled out that $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}|\boldsymbol{n},\boldsymbol{t}}$ is the result of a stopped numerical search due to a local maximum in the likelihood function or a sufficiently flat region of $L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ (e.g., McCullough and Vinod (2003)), then estimates of $\boldsymbol{\theta_k}$ are potentially located far from the true values. If nesting models are tested against nested models, the imprecision in the estimation of $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}|\boldsymbol{n},\boldsymbol{t}}$ thus may favor the former, since the respective parameter constraints cannot be ascertained properly.

We address potential problems in the numerical search algorithm stemming from deficiencies in the surface of the likelihood function $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ in two ways: First, in accordance with suggestions from the literature (Judge et al. (1985), their Appendix B, Ruud (2000) and Gould et al. (2006)) we modify the numerical search algorithm every five steps. Thus, for the numerical search algorithm required to evaluate $\log(L(\Delta_t(U_k|\boldsymbol{\theta_k})))$, we run a Newton-Ralphson procedure for the first five steps. If no solution is obtained or the algorithm fails to converge, we switch to the Davidon-Fletcher-Powell algorithm (Fletcher (1980)) for the next five iterations to push the estimates outside the critical section of the likelihood function and then return to the former technique.[25] With regard to the number of iteration steps, we follow Cramer (1986) and implement a maximum of 30 iterations. Second, another frequent suggestion to address the local maximum problem is to repeatedly use different starting values for the numerical algorithm (Liu and Mahmassani (2000)) and to check whether the same solution is obtained. We adopt this idea and systematically change the vector of starting values within the boundaries of our parameter set $\boldsymbol{\theta}_k$ for the numerical algorithm by a Halton sequence based on the prime 7. Every time Stata reports successful convergence, we store the estimates and repeat this procedure using a new starting vector. The evaluation of $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ is repeated 11 times and the estimates as well as the value of the likelihood function with the highest absolute value for $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ are chosen.[26]

5. Preliminary Analysis of the Results

In this and the next section, we present the results from our simulation study. To begin with, we generate 100 hypothetical stocks by generating their time series of returns in Step 1 and estimate the market parameters in Step 2. Similar to Hey and Orme (1994), we estimate the utility functionals described in Table (1) based on the trading sequences generated in Step 3. The resulting number of investors, for which we evaluate the likelihood function $L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ is thus determined by the number of simulated stocks and the number of utility models considered. Accordingly, a total of 18 utility models have to be estimated, since, for each stock simulated in Step 1, one trading sequence per time series of returns for

[25] A trial-and-error search in terms of number of iterations and computational time shows that among the available numerical search techniques, the Berndt-Hall-Hall-Hausman algorithm (Berndt et al. (1974)) performs worst, which leaves the Newton-Ralphson and Davidon-Fletcher-Powell algorithm (Fletcher (1980)), a result that is in line with the results found by Griffiths et al. (1987). There is no clear winner between the latter two methods; thus, we compromise and use a mixed iteration procedure as described in the text. Note that if a quadratic approximation of $L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ provides a good description the log-likelihood function, then only a few steps are needed to find the maximum (if the fit is perfect, then the maximum can be found by one iteration only). We find that an acceptable quadratic approximation of $L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ prevails only for CRRA utility models (Train (2009)).

[26] Note that there are plenty more ways to deal with local maxima: For example, a stochastic optimization algorithm that has the potential to overcome the local maximum trap was developed by Kirkpatrick et al. (1983), and is known as *simulated annealing*. Other techniques include gradient-sensitive hill climbing and random restart methods (see for details Rich and Knight (1991)).

each utility-type investor is generated, comprising 53 preference parameters each plus 18 nuisance parameters for estimation of the variance of the error term, totaling 71 parameters. As the evaluation process loops 11 times through all possible likelihood functions (we repeat each estimation using different starting values as described in Step 4 (see Chapter 4)) of all 18 utility models for each investor, the number of utility models to be evaluated sums to $19,800$ utility models, requiring estimation of $1,405,800$ preference and nuisance parameters that need to be found numerically. Given the simulated time series of returns (252 trading days), this theoretically sums to $4,989,600$ single likelihood functions.

Before we begin to elaborate our results, we need to identify potential problems that might interfere with and compromise the quality of utility model selection. In particular, note that the quality and reliability of utility model selection procedures depend on the accuracy of the numerical evaluation of the respective utility models integrated in the likelihood function. Our estimations (see Table (2)) show that from $19,800$ utility models, we are able to evaluate $14,738$ utility models (approx. 74.43%) successfully; however, we observe some variations in these numbers: for EUT investors, we are able to evaluate $1,836$ of all $2,200$ models, whereas these figures are lower for RDU ($3,318$ out of $4,400$ models), and the lowest success rate is for SPT ($4,349$ out of $6,600$ models) and CPT investors ($4,364$ out of $6,600$ models). Recall that each of the 11 loops we perform to evaluate the likelihood function is characterized by a different parameter vector of starting values, which is an indication of problems in the surface of the likelihood function. Indeed, inspections of our results from evaluation of $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ reveal for $3,340$ models several problems associated with evaluation of the likelihood function. In particular, we detect missing values for $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ associated with a stopped numerical search algorithm.

Furthermore, regarding $1,722$ utility models, we identify 954 utility models where values for $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ and estimators for $\boldsymbol{\theta_k}$ are provided and numerical search reported convergence but standard errors were set to missing. Finally, for 768 models, we find large standard errors after evaluation of $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$. We discuss each of these problems in detail, as earlier simulation studies on utility model selection such as Carbone and Hey (1994) also explicitly report similar difficulties in the evaluation of the likelihood function but do not discuss their implications on their results in detail. Due to the fact that utility model selection based on $AICC$ strongly depends on the surface of $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$, problems such as insufficient steepness of the likelihood function or convex segments can negatively affect the ranking of utility models, and thus the accuracy of model selection (e.g., if the true utility model cannot be evaluated and is therefore not included in the model ranking because its $AICC$ cannot be determined).

As mentioned above, while running the evaluation of the utility models k in Step 4, we noted that from $5,062$ utility models, $3,340$ models cannot be evaluated, notably because the iteration gets stuck ($2,619$ utility models) or exceeds the maximum iteration steps (721 models) such that the program provides missing values for $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$. As part of a more detailed investigation, we check whether the true underlying utility model was assessed successfully in each evaluation and discard those cases where the true underlying utility model cannot be assessed, as the corresponding likelihood function cannot be evaluated with respect to $\boldsymbol{\theta_k}$.

Table 2. Frequency of Appearance for each Utility Model

This table captures the proportion of evaluated utility models to the total number of utility models. The left column reports the share of utility model k across all utility-type investors' trading sequences for which the corresponding likelihood function can be evaluated, denoted as % *calc.*. The right column captures the share of true utility model k for which the corresponding likelihood function cannot be evaluated (i.e., where the numerical search algorithm was terminated) given the k-type investors' trading sequence, denoted as % $\neg$ *calc.* k. Expected utility models are denoted as EUT, Rank-dependent Utility is denoted as RDU. Simple Prospect Theory (Kahneman and Tversky (1979)) uses the notation SPT, whereas Cumulative Prospect Theory (Tversky and Kahneman (1992)) is denoted as CPT. Decision weights in accord with Quiggin (1982) are denoted as $QU82$ and as $KT92$ for decision weights, as in Tversky and Kahneman (1992). If no decision weights are applicable, we use the abbreviation $None$. Further, we use the notation $CRRA$ for CRRA utility functionals and $EXPO$ to denote utility functions as in Saha (1993). For SPT and CPT, we use the notation $POWR$ for models with kinked power-functionals as in Kahneman and Tversky (1979) and $DHG0$ to denote value functionals as defined in DeGiorgi and Hens (2006). Note that the shares listed do not sum to 100%.

		EUT		**RDU**		**SPT**		**CPT**	
		% calc.	**% ¬ calc.k**	**% calc.**	**% ¬ calc.k**	**% calc.**	**% ¬ calc.k**	**% calc.**	**% ¬ calc.k**
CRRA	**None**	84.80%	0.60%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	QU82	0.00%	0.00%	73.40%	0.68%	81.60%	0.99%	69.50%	0.60%
	KT92	0.00%	0.00%	67.10%	0.55%	61.30%	0.23%	86.70%	1.37%
EXPO	**None**	82.10%	1.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	QU82	0.00%	0.00%	87.10%	0.70%	0.00%	0.00%	0.00%	0.00%
	KT92	0.00%	0.00%	74.00%	0.35%	0.00%	0.00%	0.00%	0.00%
POWR	**None**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	QU82	0.00%	0.00%	0.00%	0.00%	62.30%	0.24%	70.30%	0.31%
	KT92	0.00%	0.00%	0.00%	0.00%	64.70%	0.52%	68.80%	0.58%
DGH0	**None**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	QU82	0.00%	0.00%	0.00%	0.00%	86.00%	0.65%	75.20%	0.73%
	KT92	0.00%	0.00%	0.00%	0.00%	79.00%	0.43%	65.90%	0.27%

According to Table (2), the probability that the true underlying utility model is among those models that cannot be evaluated is 0.61%, which is smaller than the probability that this happens by coincidence (5.56%). To illustrate this for EUT, we find that among those 167 out of $1,100$ EUT-models where the likelihood function was set to missing, there is 1 case (approx. 0.598%), where the unevaluated utility model is a EUT model. If this were purely coincidental, we would expect that we could find approximately 9 cases, where an EUT-utility model is among those models that cannot be evaluated. These cases, where the true utility model cannot be evaluated, are distributed evenly across all utility types, thus we do not expect that this biases our findings systematically when sorting the utility models according to their attained Akaike criterion. However, note that we did not discard those cases where Stata reports successful convergence and provides values for the likelihood function and $\hat{\boldsymbol{\theta}}_{\boldsymbol{k|n,t}}$, but associated standard errors were set to missing ($1,722$ utility models were affected by this issue).

Since we run the calculations under a `lf` specification of $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ in Stata, under which $\boldsymbol{H}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ and $\boldsymbol{S}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ are approximated numerically, we change the method and essential parts of the program to obtain the elements of $\boldsymbol{H}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ for a detailed analysis. In detail, we select those models where Stata reported missing values for standard errors or where the search algorithm failed to proceed after a number of iterations, and rewrite the maximum likelihood program as a `d2`-evaluator to gain further insight into the characteristics of the Hessian and the information matrices. We find that for plausible values of $\boldsymbol{\theta_k}$, the determinant of the Hessian matrix $\mathbf{det}\,\boldsymbol{H}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ is indeed fairly close to zero.[27] This has several consequences for some of the numerical methods, such as for the Newton-Ralphson method, which runs into problems because the stepsize, governed by $-\boldsymbol{H}(\Delta_t(U_k|\boldsymbol{\theta_k}))^{-1}$, cannot be determined if the Hessian matrix is degenerate. Indeed, we find terminations of the search algorithm predominantly in those iterations where the Newton-Ralphson method governs. Note that the Davidon-Flechter-Powell algorithm does not require an evaluation of $-\boldsymbol{H}(\Delta_t(U_k|\boldsymbol{\theta_k}))^{-1}$. To circumvent the impossibility of evaluating the inverse of $\boldsymbol{H}(\Delta_t(U_k|\boldsymbol{\theta_k}))$, we identify those models for which the Newton-Ralphson algorithm failed to work, rerun the estimation with the Berndt-Hall-Hall-Hausman algorithm (Berndt et al. (1974)), and additionally invoke the `difficult` option at the `ml model` command. If the Hessian is singular for some values of $\boldsymbol{\theta_k}$, then the Berndt-Hall-Hall-Hausman algorithm should overcome this by construction. In detail, the Berndt-Hall-Hall-Hausman algorithm replaces $\boldsymbol{H}(\Delta_t(U_k|\boldsymbol{\theta_k}))$ by the outer product of the scores, which is an approximation of the covariance matrix of the score vectors if the average scores were zero (Cramer (1986)). Note that this is asymptotically equivalent to $\boldsymbol{H}(\Delta_t(U_k|\boldsymbol{\theta_k}))$ (Griffiths et al. (1987)) and thus independent from the assumption that $\boldsymbol{H}(\Delta_t(U_k|\boldsymbol{\theta_k}))$ has full rank.[28]

Re-estimation of these utility models using the Berndt-Hall-Hall-Hausman algorithm shows that we can cause the majority of them to converge, although a considerable sum of utility models (1, 730 utility models, corresponding to approximately 91.26%) still cannot be evaluated accurately (see Table (3)). However, we are able to reduce the number of cases where the true utility model cannot be evaluated to approximately 0.11%. This achievement comes with an additional computational burden, as more steps are required to provide a solution for $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$, and estimators $\hat{\boldsymbol{\theta}}_{\boldsymbol{k|n,t}}$ differ significantly from the true parameters. Note that a large number of iterations is susceptible to yielding dubious results (Cramer (1986)), because of likelihood functions with convex segments due to the incorporated decision-weighting function, as in RDU, or due to the convex structure of some value functions as in SPT. This re-assessment clearly requires more than 30 iterations; however, it nevertheless yields ambiguous outcomes.[29] The large number of iterations is consistent with the results of Griffiths et al. (1987), who find in their Monte Carlo simulation study that the Newton-Ralphson algorithm performs best under multicollinearity for a probit model, whereas the method of Berndt-Hall-Hall-Hausman is found to be least efficient. This is plausible, since the Berndt-Hall-Hall-Hausman algorithm is expected to work best if the likelihood function can be sufficiently approximated by a second-order Taylor approximation (Train (2009)), but provides poor results in the face of highly nonlinear likelihood functions (Pawitan (2001)). Consequently, literature from the field of experimental economics usually does not recommend the Berndt-Hall-Hall-Hausman method for estimation of utility functions, as pointed

[27]Technically, exact singularity of $\boldsymbol{H}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ is difficult to obtain because the determinant is a real-valued variable and the precision of any statistical program is limited; thus, in this paper, we refer to near-singularity when we mention the singularity of $\boldsymbol{H}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$. For completeness, one way to detect the extent of near-singularity is outlined in Belsley et al. (2004).

[28]Another possibility is to add a constant element to the diagonal elements of $\boldsymbol{H}(\Delta_t(U_k|\boldsymbol{\theta_k}))^{-1}$ until this expression becomes invertible. This is known as the Marquardt algorithm (Marquardt (1963)); it represents the analogous case for a ridge regression (Cramer (1986)) that is usually recommended if multicollinearity is present.

[29]Application of the Berndt-Hall-Hall-Hausman algorithm provides several nuances with respect to its effectiveness; in particular, for $CRRA$, the algorithm yields results accompanied by an acceptable number of iteration steps, being not significantly different from the baseline procedure as outlined in the text, whereas the quality of $\hat{\boldsymbol{\theta}}_{\boldsymbol{k|n,t}}$ as well as the number of steps necessary deteriorates with RDU, SPT, and CPT, as well as with inclusion of heterogeneous error terms.

Table 3. Frequency of Appearance for each Utility Model for BHHH

This table captures the proportion of evaluated utility models to the total number of utility models if the Berndt-Hall-Hall-Hausman algorithm is used. The left column reports the share of utility model k across all utility-type investors' trading sequences for which the corresponding likelihood function can be evaluated, denoted as % *calc.*. The right column captures the share of true utility model k for which the corresponding likelihood function cannot be evaluated (i.e., where the numerical search algorithm was terminated) given the k-type investors' trading sequence, denoted as % ¬ *calc. k*. Expected utility models are denoted as EUT, Rank-dependent Utility is denoted as RDU. Simple Prospect Theory (Kahneman and Tversky (1979)) uses the notation SPT, whereas Cumulative Prospect Theory (Tversky and Kahneman (1992)) is denoted as CPT. Decision weights in accord with Quiggin (1982) are denoted as $QU82$ and as $KT92$ for decision weights as in Tversky and Kahneman (1992). If no decision weights are applicable, we use the abbreviation $None$. Further, we use the notation $CRRA$ for CRRA utility functionals and $EXPO$ to denote utility functions as in Saha (1993). For SPT and CPT, we use the notation $POWR$ for models with kinked power-functionals as in Kahneman and Tversky (1979) and $DHG0$ to denote value functionals as defined in DeGiorgi and Hens (2006).

		EUT		**RDU**		**SPT**		**CPT**	
		% calc.	**% ¬ calc.k**	**% calc.**	**% ¬ calc.k**	**% calc.**	**% ¬ calc.k**	**% calc.**	**% ¬ calc.k**
CRRA	**None**	93.90%	0.18%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	QU82	0.00%	0.00%	90.60%	0.18%	91.60%	0.18%	87.90%	0.09%
	KT92	0.00%	0.00%	88.50%	0.09%	91.20%	0.09%	98.60%	0.09%
EXPO	**None**	91.40%	0.09%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	QU82	0.00%	0.00%	93.10%	0.09%	0.00%	0.00%	0.00%	0.00%
	KT92	0.00%	0.00%	93.50%	0.00%	0.00%	0.00%	0.00%	0.00%
POWR	**None**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	QU82	0.00%	0.00%	0.00%	0.00%	82.20%	0.09%	89.10%	0.18%
	KT92	0.00%	0.00%	0.00%	0.00%	85.70%	0.09%	88.40%	0.09%
DGH0	**None**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	QU82	0.00%	0.00%	0.00%	0.00%	98.60%	0.18%	98.30%	0.09%
	KT92	0.00%	0.00%	0.00%	0.00%	90.20%	0.09%	89.90%	0.09%

out by Harrison and Rutstrom (2008) and Harrison (2008).

Concerning the 954 models with missing standard errors, the search algorithm indicates convergence and a solution for $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$ is found, although it cannot be considered reliable (Gould et al. (2006)). We find this problem especially prevalent for RDU (158 models), SPT (165 models), and CPT (131 models) under decision weights according to Tversky and Kahneman (1992). Closer inspection of the iteration shows that Stata noted that the likelihood function is not concave in the last iteration step, indicating that the negative of the inverse of the Hessian matrix, which determines the step size of the numerical search algorithm, is not positive definite. Recall that if the likelihood function is concave, the Hessian matrix $\boldsymbol{H}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$, the second derivative of the likelihood function, is negative definite for the full spectrum of $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$, tantamount to a declining slope of $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ with respect to one element of $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$, ceteris paribus all other elements fixed. If $\boldsymbol{H}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ is negative definite, so is its inverse $\boldsymbol{H}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))^{-1}$ and the negative of the inverse $-\boldsymbol{H}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))^{-1}$, determining the stepsize in the Newton-Ralphson method, is therefore positive definite. However, if $-\boldsymbol{H}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))^{-1}$ is negative definite, the Newton-Ralphson algorithm moves down the slope of $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ and thus away from the maximum. Re-running the evaluation of the likelihood function using the Berndt-Hall-Hall-Hausman algorithm, we find that for the majority of our utility models, standard errors are now provided. Note that the Berndt-Hall-Hall-Hausman algorithm moderates the problems of $-\boldsymbol{H}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))^{-1}$ as the direction of the search is determined by the outer product of the scores, which is necessarily positive definite. This therefore guarantees an increase of $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ in each iteration, even if the log-likelihood function displays convex segments.

Furthermore, we also detect for 768 utility models for which Stata reported successful convergence and where estimators for $\boldsymbol{\theta_k}$ are provided, the associated standard errors, according to the Cramér-Rao Lower Bound (Rao (1945) and Cramer (1946)) derived from the inverse of $\boldsymbol{I}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$, are large, because the inverse of the information matrix is small. Deficiencies in the likelihood function that result in inflated standard errors for $\hat{\boldsymbol{\theta}}_{k|n,t}$ could indicate flat sections (such as plateaus and saddle points) in the surface of the likelihood function, probably stemming from $\Delta_t(U_k|\boldsymbol{\theta_k})$. Recall that large standard errors reflect the instability of $\hat{\boldsymbol{\theta}}_{k|n,t}$, since a flat segment of $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ contains theoretically infinitely many solutions of $\boldsymbol{\theta}_k$. In particular, a flat surface of $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ could point to a certain degree of multicollinearity in $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ (Griffiths et al. (1987)) or potential under-identification problems (Judge et al. (1985), Keele and Park (2006) and Greene (2008), for utility models see Carbone and Hey (1994)) might be present.[30] We discuss each of these possible reasons below.

The effects of multicollinearity in linear models are well established (see, for example, Judge et al. (1985), Lesaffre and Marx (1993), Belsley et al. (2004), Greene (2008) and Wooldridge (2010)),[31] For nonlinear models such as the probit or logit, however, the consequences are less certain, although some solutions have been suggested to transform the problem into a known linear one (e.g., Schaefer (1986)),[32] This, in turn, suggests that the asymptotic properties of $\hat{\boldsymbol{\theta}}_{\boldsymbol{k|n,t}}$ of the probit model under multicollinearity may also hold (McLeish (1974)) with consequences similar to the linear model. For a nonlinear likelihood model, multicollinearity can lead to dependencies between and within the score vectors $\boldsymbol{S}(\Delta_t(U_k|\boldsymbol{\theta_k}))$, thus invalidating the non-singularity assumption of the information matrix $\boldsymbol{I}(\Delta_t(U_k|\boldsymbol{\theta_k}))$ (Cramer (1986)). The likelihood function $\log(L(\Delta_t(U_k|\boldsymbol{\theta_k})))$ then displays a ridge instead of a sharp peak, yielding inflated standard errors, instability in parameter estimates (as infinitely many solutions for $\hat{\boldsymbol{\theta}}_{k|n,t}$ exist), and shortcomings in the numerical search algorithms (Cramer (1986)), additionally compromising the precision of the estimates of $\hat{\boldsymbol{\theta}}_{\boldsymbol{k|n,t}}$. Yet, note that multicollinearity, per se, does not automatically invalidate the maximum likelihood properties, as simulation studies provide some evidence that the normal distribution property of the resulting distribution may still be intact (Griffiths et al. (1987)), especially if financial data contains a larger number of observations such that it can be expected that these particular asymptotic properties are likely to hold.

[30] Under moderate multicollinearity, the step size of a search algorithm is reduced if entering flat segments of $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ as a flattening of the likelihood function might indicate that the maximum is close (Train (2009)). If $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ is characterized by a flat surface over a large range of plausible $\boldsymbol{\theta_k}$ due to a sufficient degree of multicollinearity, the application of such an algorithm results in an increased number of iteration steps or a termination of the search procedure given a maximum number of iteration steps, such that the respective utility model is not evaluated adequately.

[31] For instance, in linear regression models, perfect multicollinearity leads to difficulties in inverting the vector product of the predictor variables (Belsley et al. (2004)).

[32] In particular, Fomby et al. (1978) show for the linear probit model that by applying a principal component transformation, the information matrix $\boldsymbol{I}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ can be restated in a known form, particularly as the inverse of the covariance matrix for weighted least squares (Griffiths et al. (1987)).

Ridges and plateaus in the likelihood function $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$, both corresponding to singularity (or near-singularity) of the Hessian matrix $\boldsymbol{H}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$, could also point toward under-identification problems (Cramer (1986), Wooldridge (2010)), a problem also detected in simulation studies similar to ours (Carbone and Hey (1994) and Carbone and Hey (1995)).[33] The remedy for under-identification is to avoid side relations among the elements of $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$ similar to those reported by Gonzalez and Wu (1999). Note that if such an interdependence exists, then the variations of $\boldsymbol{\theta}_{\boldsymbol{k}}$ and the estimated parameters $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$ are restricted to a subset of $\boldsymbol{H}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ of less than the dimension of $\boldsymbol{\theta}_{\boldsymbol{k}}$, denoted as K_k. As a consequence, the rank of $\boldsymbol{H}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ is less than K_k and some of the elements of $\boldsymbol{\theta}_{\boldsymbol{k}}$ cannot be identified (Cramer (1986)). Consequently, the numerical search algorithm endlessly cycles over the parameter space of $\boldsymbol{\theta}_{\boldsymbol{k}}$, as no unique maximand $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$ of the likelihood function exists. Similar to multicollinearity, the fact that different admissible parameters of $\boldsymbol{\theta}_{\boldsymbol{k}}$ can define the same probability density, suggests that there exist infinitely many parameters that maximize $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$. Side relations among the parameters of $\boldsymbol{\theta}_{\boldsymbol{k}}$ are partly due to the features of the different utility models used as, for example, for RDU, Yaari (1987) shows that even under risk neutrality, risk-averse behavior can be introduced by the nonlinearity in the decision weights. Re-running our simulation and eliminating the error term ϵ in Step 3, we find that for a given trading sequence, we can determine several parameter combinations within the spectrum of $\boldsymbol{\theta}_{\boldsymbol{k}}$ that can explain the observed trading sequence equally successfully, as conjectured in Chapter 3, in which we discuss the role of the error term. Note that this observation is in line with such experimental studies as Gonzalez and Wu (1999), who report signs of significant correlation among their preference estimators for the CPT case, and Carbone and Hey (1995), who mention that for EUT, they detect a large number of admissible parameter values that fit their data. For example, Keele and Park (2006) report that the heteroscedastic linear probit model is quite prone to fragile identification (see also Judge et al. (1985) and Greene (2008)). These authors suggests this weakness to be evident if the likelihood function is smaller for the heteroscedastic likelihood than for the likelihood estimation where σ_ϵ is homogeneous and constrained to unity. We conclude that the magnitude of σ_ϵ as well as its stochastic properties are subject to discussion, as perhaps a higher level of σ_ϵ or a different assumption of the distribution of the error term might capture more accurately the inherent problems of the likelihood function and the embedded utility functions therein. We discuss these points in the next chapter.[34]

6. Utility Model Selection: Analysis and Discussion

A first inspection of the results of model selection, depicted in Table (4), indicates that throughout the simulated trading sequences, the true utility models obtained the highest rank in more than 50% of all cases. In detail, for EUT as the true model, the correct utility model obtained, on average, in 55% of all cases

[33] Multicollinearity and under-identification are difficult to disentangle and thus not widely discussed. We find this surprising, since, despite their prevalence in studies on utility model selection and parameter estimation, under-identification issues are usually not discussed in detail in terms of analysis, although a distinction is necessary to choose the remedies. If multicollinearity causes the breakdown in the likelihood estimation, collecting new data free from the defects can help. If a lack of identification is important to model assumptions, a modification of the model and a revision of extraneous information may be inevitable (Cramer (1986)).

[34] I am grateful to John Hey, who pointed out the importance of running simulated trading sequences by adding an error term ϵ to dampen the problem of under-identification.

the first rank, where for other utility models, the outcomes are better. For RDU, the true utility model obtained, on average, in 57.25% the first rank, for SPT on average in 66.45% and for CPT, on average, in 67.50% of all cases the first rank. If a broader classification is accepted, the ranking of the respective utility model improves as 92.70% of all utility functions can be classified correctly. This is reflected in the overall likelihood values across all utility models, which are close to zero for EUT and RDU and somewhat higher for SPT and CPT.[35] This is surprising, since despite the obscure stochastic properties of the additional uncertainty stemming from $\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}})$, nuisance parameter $\hat{\sigma}_\epsilon$ seems to capture the stochastics from the additional uncertainty reasonably well. Although we noted that the estimated distribution of the error term $\hat{\sigma_\epsilon} = 0.031$ is, on average, higher than the distribution of the error term used in the simulation to generate the trading sequences, which we set to $\sigma_\epsilon = 0.01$, a t-test indicates that the estimators for $\sigma_\epsilon = 0.01$ are significantly distinct from the true value only at a 10%-significance level (p-value 0.088). Inspecting the estimators $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$ of true utility models shows that the quality of our estimators varies among the various utility functions. We identify some cases where the estimators are close to the true values (predominantly for SPT-type (83%) and CPT-type investors (78%)); in other cases the estimators are significant and distinct even at the 1%-significance level (for more than 37% of all EUT-type investors and for 43% of RDU-type investors). We discuss the consequences of this bluntness of $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$ in a subsequent analysis. Henceforth, these results serve as a benchmark for our elaborations on the influence of modifications.

We pointed out earlier that our simulation is conducted using a normally distributed error term ϵ, arbitrarily setting its standard deviation to $\sigma_\epsilon = 0.01$. However, in light of the significant deviations in $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$, particularly for EUR and RDU, the magnitude of σ_ϵ might be too low. Recall that with regard to the significance of our parameter estimates of each utility model k, we find that the introduction of such an error term can help to resolve some under-identification problems. To elaborate the effects of ϵ, consider a situation in which the market parameters μ and σ are fixed and known to the investor; consequently, upside and downside returns R_U and R_D and the associated probabilities are fixed. If no error term is added to the difference in utility $\Delta_t(U_k|\boldsymbol{\theta}_{\boldsymbol{k}})$, the investor is invested in stocks always or never, dependent on the set of parameters $\boldsymbol{\theta}_{\boldsymbol{k}}$. This is not surprising; as the investor faces the same time-independent decision problem, the investor will decide always to do the same thing, as the investor is assumed to make no mistakes. Adding some ϵ to $\Delta_t(U_k|\boldsymbol{\theta}_{\boldsymbol{k}})$ generates trading sequences with more variations, and the application of maximum likelihood estimation of $\boldsymbol{\theta}_{\boldsymbol{k}}$ yields $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ to be close to zero with estimators for $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$ and $\hat{\sigma}_\epsilon$ closer to the true values.

However, if μ and σ are not fixed and/or known to the investor such that market parameters must be estimated from the time series of returns, additional fluctuations with an unclear stochastic pattern arise from $\Delta_t(U_k|\boldsymbol{\theta}_{\boldsymbol{k}})$, particularly the dynamics of market parameters (due to the rolling-window estimation) and the evolution of intermediate gains and losses. To locate the source of this additional uncertainty, we rerun simulations with a stationary stochastic process using an *Ornstein-Uhlenbeck-Process* (see Ornstein and Uhlenbeck (1930), Box et al. (2015)), generating almost time-invariant estimators for $R_{U,t}$, $R_{D,t}$ and for which we fixed

[35]The average of the likelihoods of each of $\log L(\Delta_t(U_{EUT}|\hat{\boldsymbol{\theta}}_{\boldsymbol{EUT}}))$ is -6.42, the average of $\log L(\Delta_t(U_{RDU}|\hat{\boldsymbol{\theta}}_{\boldsymbol{RDU}}))$ is -8.22, the average of $\log L(\Delta_t(U_{SPT}|\hat{\boldsymbol{\theta}}_{\boldsymbol{SPT}}))$ is -22.87 and for $\log L(\Delta_t(U_{CPT}|\hat{\boldsymbol{\theta}}_{\boldsymbol{CPT}}))$ is -21.09.

Table 4. Base Case: Ranking of each Utility Model

This table captures the median and average ranking (below), denoted as *Rank* of each utility model k for trading sequences where utility model k is the true utility model used to generate these trading sequences. We also report the type of the first-ranked ($1.^{st}$) and of the second-ranked ($2.^{nd}$) utility model (in brackets). Expected utility models are denoted as EUT, Rank-dependent Utility is denoted as RDU. Simple Prospect Theory (Kahneman and Tversky (1979)) uses the notation SPT, whereas Cumulative Prospect Theory (Tversky and Kahneman (1992)) is denoted as CPT. Decision weights in accordance with Quiggin (1982) are denoted as $QU82$ (Q) and as $KT92$ (K) for decision weights as in Tversky and Kahneman (1992). If no decision weights are applicable, we use the abbreviation *None*. Further, we use the notation $CRRA$ (C) for CRRA utility functionals and $EXPO$ (E) to denote utility functions as in Saha (1993). For SPT and CPT, we use the notation $POWR$ (P) for models with kinked power-functionals as in Kahneman and Tversky (1979) and $DHG0$ (D) to denote value functionals as defined in DeGiorgi and Hens (2006). Significance levels are calculated according to Wilcoxon signed-rank tests comparing the average ranking between the first-ranked and second-ranked utility models. We use $***$, $**$ and $*$ for significance at the 1%, 5% and 10% levels, respectively.

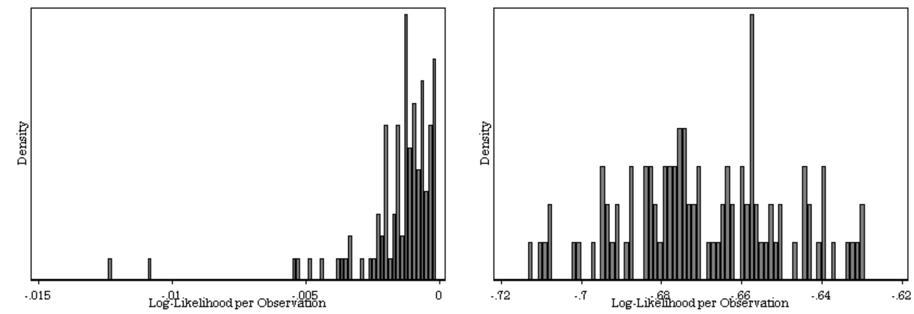
		EUT		**RDU**		**SPT**		**CPT**	
		Rank	$1.^{st}$ ($2.^{nd}$)	**Rank**	$1.^{st}$ ($2.^{nd}$)	**Rank**	$1.^{st}$ ($2.^{nd}$)	**Rank**	$1.^{st}$ ($2.^{nd}$)
CRRA	**None**	1^{***}	EUT^C	0	0	0	0	0	0
		1.560	$\left(EUT^E\right)$	-	-	-	-	-	-
	QU82	0	0	1^{***}	RDU^C_Q	1^{***}	SPT^C_Q	1^{***}	CPT^C_Q
		-	-	1.550	$\left(RDU^C_P\right)$	1.420	$\left(SPT^P_Q\right)$	1.480	$\left(CPT^P_Q\right)$
	KT92	0	0	1^{***}	RDU^C_K	1^{**}	SPT^C_K	1^{***}	CPT^C_K
		-	-	1.500	$\left(RDU^E_P\right)$	1.480	$\left(SPT^C_Q\right)$	1.530	$\left(CPT^D_Q\right)$
EXPO	**None**	1^{***}	EUT^E	0	0	0	0	0	0
		1.510	$\left(EUT^C\right)$	-	-	-	-	-	-
	QU82	0	0	1^{***}	RDU^E_Q	0	0	0	0
		-	-	1.510	$\left(RDU^C_P\right)$	-	-	-	-
	KT92	0	0	1^{***}	RDU^E_P	0	0	0	0
		-	-	1.490	$\left(RDU^E_Q\right)$	-	-	-	-
POWR	**QU82**	0	0	0	0	1^{***}	SPT^P_Q	1^{***}	CPT^P_Q
		-	-	-	-	1.480	$\left(SPT^P_K\right)$	1.420	$\left(CPT^D_K\right)$
	KT92	0	0	0	0	1^{***}	SPT^P_K	1^{***}	CPT^P_K
		-	-	-	-	1.510	$\left(SPT^P_Q\right)$	1.380	$\left(SPT^P_Q\right)$
DGH0	**QU82**	0	0	0	0	1^{***}	SPT^D_Q	1^{***}	CPT^D_Q
		-	-	-	-	1.530	$\left(SPT^P_Q\right)$	1.360	$\left(CPT^P_Q\right)$
	KT92	0	0	0	0	1^{**}	SPT^D_K	1^{**}	CPT^D_K
		-	-	-	-	1.420	$\left(SPT^P_Q\right)$	1.480	$\left(CPT^P_Q\right)$

$p_t = 50\%$. If additional disturbance arises from the estimation of market parameters, then we expect a notable improvement in $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$. Note that since a modification of the stochastic process affects all utility models, if additional uncertainty is caused by market parameter estimation, all likelihoods are expected to improve significantly. Inspections of $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ for the different utility models shows that for EUT, the maximized likelihood is close to zero, whereas for SPT and CPT models, the overall likelihood remains large, although still ranging in lower double figures, not significantly distinct from the baseline likelihood values on a 10%-significance level according to a likelihood-ratio test. The ranking of all utility models compared to the baseline case in Table (4) shows no significant changes according to a Wilcoxon signed-rank test.[36] We detect only small improvement in $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ for all utility models according to likelihood ratio tests for the usual significance levels. We conclude that the source of additional disturbance stems from accrued gains and losses within $\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}})$; as in the case of EUT and RDU, the accrued return can be truncated from $\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}})$.

Tracing back the additional uncertainty to the dynamics of the accrued return, it is not clear whether the estimation of the nuisance parameter σ_ϵ captures this additional uncertainty stemming from $\Delta_t(U_k|\boldsymbol{\theta}_{\boldsymbol{k}})$, particularly if we try to overlay the defects from $\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}})$ by increasing σ_ϵ. We observe that an increase in σ_ϵ alters the trading sequences, and thus the magnitude of the accrued return within the likelihood function. However, introducing higher disturbance of the error term could lead to deterioration of the estimators, altering the trading sequences, and thus affecting utility model selection negatively. To investigate the sensitivity of our utility model selection with respect to the magnitude of the error term, we rerun our simulation for various values of the standard deviation σ_ϵ. According to Table (5), a more pronounced standard deviation of the error term σ_ϵ is detrimental to utility model selection. In empirical data however, it cannot be expected that the error term plays such a modest role as in our simulation. Thus, we wish to eliminate the possibility that the likelihood approach shows a tendency to favor a particular utility model if the error term dominates (i.e., other trading factors matter and preference considerations are negligible).

Figure 2. Distribution of $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ for Different Investors

The figures on the left illustrate the distribution of the log-likelihood values $\log L(\Delta_t(U_{EUT,CRRA}|\hat{\boldsymbol{\theta}}_{\boldsymbol{EUT,CRRA}}))$, divided by the number of simulated draws for the trading sequence of a EUT-type investor with $CRRA$ utility functional. The figure on the right shows the distribution of the log-likelihood values $\log L(\Delta_t(U_{EUT,CRRA}|\hat{\boldsymbol{\theta}}_{\boldsymbol{EUT,CRRA}}))$ divided by the number of simulated draws of the EUT model with $CRRA$ utility functional for a trading sequence of a random trader where $\epsilon \sim N(0,1)$.



To investigate the sensitivity of our utility model selection with regard to a further increase in σ_ϵ, we extend our simulation and introduce the concept of a

[36]The significance of the differences in the rank scores between the baseline setting and the results from a re-estimation under modified settings are based on a Wilcoxon signed-rank test (Wilcoxon (1945); for the two-sample test see Mann and Whitney (1947)). This non-parametric test seems appropriate for this purpose, as it can be used for outcomes that are coded ordinally and that require no explicit distribution of the matched samples; yet, the test is comparable to a t-test (Siegel (1956)).

Table 5. Variations of Error Term: Ranking of each Utility Model

This table captures the median and average ranking (below) of each utility model k for trading sequences where utility model k is the true utility model used to generate these trading sequences given various values of σ_ϵ. Expected utility models are denoted as EUT, Rank-dependent Utility is denoted as RDU. Simple Prospect Theory (Kahneman and Tversky (1979)) uses the notation SPT, whereas Cumulative Prospect Theory (Tversky and Kahneman (1992)) is denoted as CPT. Decision weights in accordance with Quiggin (1982) are denoted as $QU82$ and as $KT92$ (K) for decision weights as in Tversky and Kahneman (1992). If no decision weights are applicable, we use the abbreviation $None$. Further, we use the notation $CRRA$ for CRRA utility functionals and $EXPO$ to denote utility functions as in Saha (1993). For SPT and CPT, we use the notation $POWR$ for models with kinked power-functionals as in Kahneman and Tversky (1979) and $DHG0$ to denote value functionals as defined in DeGiorgi and Hens (2006). Significance levels are calculated according to Wilcoxon signed-rank tests comparing the average ranking between the first-ranked and second-ranked utility models if the true utility model obtains first rank, and to the first-ranked model if the true model obtains second rank. For the respective utility models of first and second rank, see Table (7). We use $***$, $**$ and $*$ for significance at the 1%, 5% and 10% levels, respectively.

		EUT			RDU			SPT			CPT		
		$\sigma = 0.05$	$\sigma = 0.10$	$\sigma = 0.20$	$\sigma = 0.05$	$\sigma = 0.10$	$\sigma = 0.20$	$\sigma = 0.05$	$\sigma = 0.10$	$\sigma = 0.20$	$\sigma = 0.05$	$\sigma = 0.10$	$\sigma = 0.20$
CRRA	**None**	1^{***}	2^{**}	2^{**}	0	0	0	0	0	0	0	0	0
		1.630	1.710	1.790	-	-	-	-	-	-	-	-	-
	QU82	0	0	0	1^{**}	2^{*}	2^{**}	1^{***}	1^{***}	1^{**}	1^{***}	1^{**}	2
		-	-	-	1.550	1.550	1.550	1.500	1.620	1.720	1.600	1.720	1.830
	KT92	0	0	0	1^{***}	2	2^{*}	1^{*}	1^{*}	1^{*}	1^{***}	2^{*}	2^{*}
		-	-	-	1.570	1.680	1.750	1.580	1.700	1.770	1.620	1.700	1.820
EXPO	**None**	1^{***}	2^{***}	2^{***}	0	0	0	0	0	0	0	0	0
		1.610	1.770	1.750	-	-	-	-	-	-	-	-	-
	QU82	0	0	0	1^{*}	2	2	0	0	0	0	0	0
		-	-	-	1.600	1.700	1.770	-	-	-	-	-	-
	KT92	0	0	0	1^{**}	2^{*}	2^{**}	0	0	0	0	0	0
		-	-	-	1.550	1.650	1.720	-	-	-	-	-	-
POWR	**QU82**	0	0	0	0	0	0	1^{***}	2^{*}	2^{*}	1^{***}	1^{*}	2^{*}
		-	-	-	-	-	-	1.610	1.770	1.870	1.530	1.650	1.740
	KT92	0	0	0	0	0	0	1^{***}	2^{*}	2^{*}	1^{***}	1^{*}	2^{**}
		-	-	-	-	-	-	1.640	1.800	1.900	1.440	1.590	1.700
DGH0	**QU82**	0	0	0	0	0	0	1^{***}	2^{*}	2^{*}	1^{***}	1^{*}	2
		-	-	-	-	-	-	1.680	1.800	1.870	1.420	1.570	1.680
	KT92	0	0	0	0	0	0	1^{**}	1^{**}	2^{*}	1^{**}	1^{**}	2^{*}
		-	-	-	-	-	-	1.530	1.650	1.730	1.630	1.740	1.840

random trader, defined as an investor who trades based on economically irrelevant criteria that are purely independent from preferences (Kyle (1985) and Black (1986)). For our simulation of the trading sequences, we define a *Random Trader* as an investor who trades on other criteria that are unsystematic, standardized in their variance, independent from utility considerations, and thus approximately normally distributed with $\epsilon \sim N(0,1)$. We make the conjecture that the respective log-likelihood values $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ of each of the k utility models should be close to the baseline log-likelihood $\log L(\Delta_t(\epsilon)) = -174.67$ for a random traders' trading sequence, as none of the models contributes further information to observed trading data. In particular, to construct a random trader, we generate trading signals by replacing the difference in utilities $\Delta_t(U_k|\boldsymbol{\theta}_{\boldsymbol{k}})$ with a stochastic element ϵ, characterized by a standard normal density with zero mean and standardized variance, as mentioned above. Accordingly, the random trader has positive exposure to the risky stock if argument ϵ yields a cumulative density $\Phi(\epsilon)$ above 50% and the investor prefers to hold the riskless investment otherwise. In all cases, the obtained values of $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ are not significantly different from the baseline $\log L(\Delta_t(\epsilon))$ according to likelihood-ratio tests conducted for each utility model k.[37] The differences in the likelihood $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ for the *EUT*-type investor trading sequence and the likelihood of a Random Trader trading sequence is also illustrated in Figure 2. The log-likelihood of the simulated *EUT*-type investor is close to its theoretical maximum, whereas, in contrast, the log-likelihood values of the same utility model given a trading sequence of a Random Trader are distributed around ln(0.5) and display higher dispersion.

Given the trading sequence of a Random Trader, sorting the conceivable utility models according to *AICC* reveals a pronounced tendency to identify an *SPT* model specification as the best fit, regardless of the combination of the various return moment characterizations and stochastic processes (Wilcoxon signed-rank test p-value 0.054). Note also that this effect persists if we simulate the trading behavior with a certain degree of dispersion by increasing the magnitude of σ_ϵ by using $\epsilon \sim N(0, \sigma_\epsilon^2)$. In particular, we consider two cases in which we trigger a hold signal whenever the cumulative density $\Phi(\Delta_t(U_k|\boldsymbol{\theta}_{\boldsymbol{k}})/\sigma(\Delta_t(U_k|\boldsymbol{\theta}_{\boldsymbol{k}})))$ exceeds 50%, where the spread of the cumulative normal distribution depends on the difference in utility, as is frequently assumed in experiments (Moffatt and Peters (2001) and Loomes et al. (2002)). Note that this form of the error term corresponds to cases where an investor makes more mistakes the larger the difference in utility. While we obtain acceptable quality of model selection results for the endogenous case, we detect a bias toward *SPT* in the latter case, notably for particularly high values of $\sigma(\epsilon)$, as the resulting trading behavior resembles that of the random investor.[38]

[37]The corresponding p-values range from p-value 0.289 for *SPT* given *POWR* value functional and a decision weight according to Quiggin (1982), up to p-value 0.382 for *EUT* given *EXOP* utility functional.

[38]Note that the length of trading sequences from the Random Trader is sensitive to the mean of ϵ and insensitive with respect to σ_ϵ. The resulting trading sequences are short if the mean of ϵ is set to zero, comparable to a trading pattern from a day trader. If longer trading sequences are required, the mean of ϵ should be close to 0.013 to match the roundtrip length of an *EUT*-type investor in our case. Due to the insensitivity regarding σ_ϵ, the corresponding mean for ϵ can be approximated by matching the upper bound of the integral of the cumulative distribution of the Random Trader with the investment ratio of an *EUT*-type investor (number of observations where the indicator is 1 divided by trading days). We estimate a proportional hazard model as proposed by Cox (1972) to verify the conjecture that the generated trading sequences of the random investor are akin to the sequences generated by introducing the disturbance $\sigma(\Delta_t(U_k|\boldsymbol{\theta}_{\boldsymbol{k}}))$. A log-rank test (Mantel (1966)), as implemented in the `sts test` command in Stata, reveals no significant difference between the two hazard rates.

This may indicate some problems in identifying the correct utility model specification if differences in utility represent only a minor aspect for investor decision making in stock markets, as Grinblatt and Keloharju (2001c) and Levitt and List (2007) point out.

As a central component of the decision model (3.2), the role of the stochastic error term is crucial to determination of the likelihood function, as we show; however, the correct formal specification of ϵ is the subject of ongoing debate (see Harless and Camerer (1994), Hey and Orme (1994), Loomes and Sugden (1995) and Ballinger and Wilcox (1997) for a discussion of Cauchy and Laplace distributed errors), since the distributional assumption may have consequences on the determination of the best-fitting model. A frequently applied form is to define a utility ratio index, discussed in Harrison and Rutstrom (2008), related to a logit formulation as shown by McFadden (1974). To test the sensitivity of our results regarding modification of the distribution of the error term, we rerun our simulations under an independently, identically extreme value (type 1) distributed error term ϵ with density $\phi(\epsilon) \sim \frac{1}{\sigma_\epsilon} e^{-\epsilon/\sigma_\epsilon} e^{-e^{-\epsilon/\sigma_\epsilon}}$. The cumulative distribution of the error term used for the likelihood function is therefore $\Phi(\epsilon) = e^{-e^{-\epsilon/\sigma_\epsilon}}$ and the difference in utility plus the error term is described by a logistic function (Greene (2008), Train (2009) and Hosmer et al. (2013); for utility model selection see Carbone and Hey (1995), Loomes and Sugden (1995), Loomes et al. (2002) and Harrison (2008)).[39] According to Table (6), our results are virtually unchanged. A Wilcoxon signed-rank test indicates that the changes in the average rank between the baseline group and the average ranks, where we use a logit specification of the error term instead, are insignificant.[40] Although the logistic distribution is characterized by fatter tails in comparison to the standard normal distribution, the differences between these distributions are usually insignificant (Hosmer et al. (2013)), which explains why we find virtually the same results.[41] Despite logit and probit, various other error specifications and their relation to utility specifications are suggested and reviewed in Hey (1995), Loomes and Sugden (1995), Ballinger and Wilcox (1997), Hey (2002) and Loomes et al. (2002). Note that, under certain circumstances, using extreme value distributed error terms can bias the results: Wilcox (2008) shows that in a logit model according to Luce (1959), the finding that subjects behave according to $IARA$ may be biased by the fact that ϵ follows an extreme value distribution.[42]

With regard to the specification and estimation of σ_ϵ, we mention that the nuisance parameter may also be able to cope with other issues arising with financial data, such as the correlation structure within σ_ϵ, i.e., error in decision making of an investor carries over from previous periods, creating correlation among ϵ.

[39] Note that the mean of the extreme value distribution is not zero as for the standard normal distribution (the usual interpretation of a zero mean is that investors do not do errors on average, see Carbone and Hey (1995)); however, the mean appears to be immaterial.

[40] For example, for EUT, the corresponding p-value is 0.504, for RDU, its p-value is 0.373, and for SPT and CPT the corresponding p-values are 0.324 and 0.425, respectively.

[41] However, despite this result, we suggest not to use logit, for several reasons: First, we cannot exclude correlation effects in the error term, as Train (2009) points out. If these correlations exist, they need to be modeled and estimated explicitly (Train (1986) and Train (2009)). Second, for utility models where the value of the utility can be negative, such as SPT and CPT, this concept violates certain axioms of rationality, since a logit specification is justified only under a positive measure scale, such as, e.g., EUT (see Luce (1959)).

[42] In the experimental literature, the heterogeneous error likelihood specification is also referred to as *Fechnerian Error* or *white noise* (Fechner (1966), a choice model has been developed by Becker and Marschak (1963) and popularized by Hey and Orme (1994)), consistent with a probit specification, as shown above.

Table 6. Logit Error Specification: Ranking of each Utility Model

This table captures the median and average ranking (below), denoted as *Rank* of each utility model k for trading sequences where utility model k is the true utility model used to generate these trading sequences given a logit specification of the error term. We also report the type of the first-ranked ($1.^{st}$) and of the second-ranked ($2.^{nd}$) utility model (in brackets). Expected utility models are denoted as EUT, and Rank-dependent Utility is denoted as RDU. Simple Prospect Theory (Kahneman and Tversky (1979)) uses the notation SPT, whereas Cumulative Prospect Theory (Tversky and Kahneman (1992)) is denoted as CPT. Decision weights in accordance with Quiggin (1982) are denoted as $QU82$ (Q) and as $KT92$ (K) for decision weights as in Tversky and Kahneman (1992). If no decision weights are applicable, we use the abbreviation *None*. Further, we use the notation $CRRA$ (C) for CRRA utility functionals and $EXPO$ (E) to denote utility functions as in Saha (1993). For SPT and CPT, we use the notation $POWR$ (P) for models with kinked power-functionals as in Kahneman and Tversky (1979) and use $DHG0$ (D) to denote value functionals as defined in DeGiorgi and Hens (2006). Significance levels are calculated according to Wilcoxon signed-rank tests comparing the average ranking between the first-ranked and second-ranked utility models. We use $***$, $**$ and $*$ for significance at the 1%, 5% and 10% levels, respectively.

		EUT		**RDU**		**SPT**		**CPT**	
		Rank	$1.^{\text{st}}$ ($2.^{\text{nd}}$)	**Rank**	$1.^{\text{st}}$ ($2.^{\text{nd}}$)	**Rank**	$1.^{\text{st}}$ ($2.^{\text{nd}}$)	**Rank**	$1.^{\text{st}}$ ($2.^{\text{nd}}$)
CRRA	**None**	1^{***}	EUT^C	0	0	0	0	0	0
		1.560	(EUT^E)	-	-	-	-	-	-
	QU82	0	0	1^{***}	RDU_Q^C	1^{***}	SPT_Q^C	1^{***}	CPT_Q^C
		-	-	1.550	(RDU_P^C)	1.420	(SPT_Q^P)	1.480	(CPT_Q^P)
	KT92	0	0	1^{***}	RDU_K^C	1^{**}	SPT_K^C	1^{***}	CPT_K^C
		-	-	1.500	(RDU_K^E)	1.480	(SPT_Q^C)	1.530	(CPT_Q^D)
EXPO	**None**	1^{***}	EUT^E	0	0	0	0	0	0
		1.510	(EUT^C)	-	-	-	-	-	-
	QU82	0	0	1^{***}	RDU_Q^E	0	0	0	0
		-	-	1.510	(RDU_K^C)	-	-	-	-
	KT92	0	0	1^{***}	RDU_P^E	0	0	0	0
		-	-	1.490	(RDU_Q^E)	-	-	-	-
POWR	**QU82**	0	0	0	0	1^{***}	SPT_Q^P	1^{***}	CPT_Q^P
		-	-	-	-	1.480	(SPT_K^P)	1.420	(CPT_K^P)
	KT92	0	0	0	0	1^{***}	SPT_K^P	1^{***}	CPT_K^P
		-	-	-	-	1.510	(SPT_K^P)	1.380	(SPT_K^C)
DGH0	**QU82**	0	0	0	0	1^{***}	SPT_Q^D	1^{***}	CPT_Q^D
		-	-	-	-	1.530	(SPT_K^P)	1.360	(CPT_K^P)
	KT92	0	0	0	0	1^{*}	SPT_K^D	1^{**}	CPT_K^D
		-	-	-	-	1.420	(SPT_Q^P)	1.480	(CPT_Q^D)

Autocorrelation within the structure of the error term may arise if other factors affecting trading decisions, which are assumed to be independent from individual risk preferences, correlate across time, which is suspected to alter the dispersion of $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$, and thus should be captured by σ_ϵ. Since autocorrelation does not affect the functional form of $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ per se, despite increasing its spread via its dispersion (Pawitan (2001)), estimating σ_ϵ should have no consequences for the selection of the utility model, but it can indirectly interact with model identification via the precision of estimates $\hat{\boldsymbol{\theta}}_{\boldsymbol{k|n,t}}$.[43] To investigate the impact of autocorrelation on σ_ϵ, we consider the possibility that the error term ϵ may

[43] Note that the estimation of parameter σ_ϵ introduces an element that lowers the concavity of $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ and increases the number of iteration steps of our numerical search algorithm.

be potentially autocorrelated within each investor's trading sequence in reality, and that such correlation may drive the high standard errors we expect to find in this case (Wooldridge (2010)). Accordingly, in an earlier pretest of our likelihood approach, we modeled a lagged error term with lag 1 to analyze its effect on our utility model selection results. We follow Roger (1994) and use the `cluster` option in Stata within the `ml model` command, which invokes evaluation of the likelihood function (see Harrison and Rutstrom (2008) and Harrison (2008)) to control for the autocorrelation structure within each investor type (Harrison and Rutstrom (2009)). We find that this action leads to virtually no consequences for both our results and the precision of the estimated risk parameters, such that we conclude that $\hat{\sigma}_\epsilon$ is also able to capture possible autocorrelation sufficiently.

Other issues regarding the quality of utility model selection concern the likelihood function $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ and its susceptibility to overfitting issues. Although we control for the different number of parameters by using $AICC$ instead of $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$, we cannot be absolutely sure whether the penalization term for additional parameters is sufficient. The inclusion of a piecewise negative exponential value functional as proposed by DeGiorgi and Hens (2006), which contains four different risk sensitivity parameters, allows us to test for the susceptibility of the likelihood-based model selection procedure for possible overfitting issues.[44] The proposed value function added to our analysis is, accordingly, defined as a piecewise negative exponential value function with parameters set equal to the values presented in DeGiorgi and Hens (2006) to match with parameter estimates of Tversky and Kahneman (1992). If the likelihood function is prone to overfitting, then models containing a DeGiorgi and Hens (2006) functional should end up in higher ranks compared to models incorporating Kahneman and Tversky (1979) versions of the value function. We find no signs of an amassment of *DGH0* models in higher ranks across all utility models; thus, an overfitting problem does not appear to be significant.[45]

Since some utility models, for which we generate trading sequences in Step 3, are nested in more general formulations, we expect that nesting utility models should yield estimates for $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$ statistically indistinguishably close to those constraining values under which the nested utility model coincides with a nesting one. We find that for an increase in σ_ϵ, the trading sequences shorten and our parameter estimates for $\boldsymbol{\theta}_{\boldsymbol{k}}$ deteriorate; thus, due to the high standard errors and imprecision in the preference parameter estimates, nesting models are preferred to nested models even in those cases where nested models represent the true underlying utility model (see Table (7)). We elaborate above that our simulations reveal signs of shortcomings in the functionality of the applied search algorithm and the reliability of $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}|\boldsymbol{n},\boldsymbol{t}}$. These drawbacks compromise not only the obtained likelihood maxima and the information criteria, based upon which the ranking of models takes place, but they

[44] These authors mention that this functional form should offer a higher descriptive power with respect to the behavior of investors in financial markets than other functions, since this form is explicitly designed to capture the features of decreasing marginal utility if financial outcomes reach the edges of the return distributions. They argue that marginal utility stemming from the value function is still decreasing at the bounds of the return distribution, whereas the usually applied form of Kahneman and Tversky (1979) is virtually linear in the realm of higher stakes.

[45] In those cases where the $DGH0$-type is not the true model, the average rank for $DGH0$ utility model ranges between 12.4 (median rank: 11) for EUT and 7.2 (median rank: 7) for SPT. P-values from the Wilcoxon signed-rank tests range from 0.024 for EUT up to 0.051 for SPT; thus, the average ranks of the two models are at least significantly distinct on a 10%-significance level.

Table 7. Variations of Error Term: Tendency toward Nesting Utility Models

This table reports the type of the first-ranked ($1.^{st}$) and of the second-ranked ($2.^{nd}$) utility model (in brackets) for various values of σ_ϵ. Expected utility models are denoted as EUT, Rank-dependent Utility is denoted as RDU. Simple Prospect Theory (Kahneman and Tversky (1979)) uses the notation SPT, whereas Cumulative Prospect Theory (Tversky and Kahneman (1992)) is denoted as CPT. Decision weights in accordance with Quiggin (1982) are denoted as $QU82$ (Q) and as $KT92$ (K) for decision weights as in Tversky and Kahneman (1992). If no decision weights are applicable, we use the abbreviation *None*. Further, we use the notation $CRRA$ (C) for CRRA utility functionals and $EXPO$ (E) to denote utility functions as in Saha (1993). For SPT and CPT, we use the notation $POWR$ (P) for models with kinked power-functionals as in Kahneman and Tversky (1979) and use $DHG0$ (D) to denote value functionals as defined in DeGiorgi and Hens (2006).

		EUT			**RDU**			**SPT**			**CPT**		
		$\sigma = 0.05$	$\sigma = 0.10$	$\sigma = 0.20$	$\sigma = 0.05$	$\sigma = 0.10$	$\sigma = 0.20$	$\sigma = 0.05$	$\sigma = 0.10$	$\sigma = 0.20$	$\sigma = 0.05$	$\sigma = 0.10$	$\sigma = 0.20$
CRRA	**None**	EUT^C $\left(EUT^E\right)$	EUT^E $\left(RDU^E_Q\right)$	RDU^E_Q $\left(RDU^E_K\right)$	0 -	0 -	0 -	0 -	0 -	0 -	0 -	0 -	0 -
	QU82	0 -	0 -	0 -	RDU^C_Q $\left(RDU^E_K\right)$	RDU^E_K $\left(SPT^C_K\right)$	RDU^E_K $\left(SPT^C_K\right)$	SPT^C_Q $\left(SPT^P_Q\right)$	SPT^C_Q $\left(SPT^C_K\right)$	SPT^C_Q $\left(SPT^P_K\right)$	CPT^C_Q $\left(CPT^C_K\right)$	CPT^C_Q $\left(CPT^P_Q\right)$	SPT^C_Q $\left(SPT^C_K\right)$
	KT92	0 -	0 -	0 -	RDU^C_K $\left(RDU^E_Q\right)$	RDU^C_Q $\left(SPT^C_K\right)$	SPT^C_Q $\left(SPT^C_K\right)$	SPT^C_K $\left(SPT^P_Q\right)$	SPT^C_K $\left(CPT^P_Q\right)$	SPT^C_K $\left(SPT^P_K\right)$	CPT^C_K $\left(CPT^P_K\right)$	SPT^C_K $\left(CPT^P_K\right)$	SPT^C_Q $\left(SPT^P_K\right)$
EXPO	**None**	EUT^E $\left(RDU^E_Q\right)$	EUT^C $\left(RDU^E_Q\right)$	SPT^C_Q $\left(RDU^E_K\right)$	0 -	0 -	0 -	0 -	0 -	0 -	0 -	0 -	0 -
	QU82	0 -	0 -	0 -	RDU^E_Q $\left(RDU^E_K\right)$	RDU^C_K $\left(SPT^C_K\right)$	RDU^C_K $\left(SPT^C_K\right)$	0 -	0 -	0 -	0 -	0 -	0 -
	KT92	0 -	0 -	0 -	RDU^E_K $\left(RDU^E_Q\right)$	RDU^C_Q $\left(SPT^P_K\right)$	SPT^C_Q $\left(SPT^P_K\right)$	0 -	0 -	0 -	0 -	0 -	0 -
POWR	**QU82**	0 -	0 -	0 -	0 -	0 -	0 -	SPT^P_Q $\left(SPT^D_Q\right)$	SPT^C_K $\left(SPT^P_K\right)$	SPT^C_K $\left(CPT^C_K\right)$	CPT^P_Q $\left(SPT^P_K\right)$	CPT^P_Q $\left(SPT^P_K\right)$	SPT^P_Q $\left(SPT^C_K\right)$
	KT92	0 -	0 -	0 -	0 -	0 -	0 -	SPT^P_K $\left(SPT^C_Q\right)$	CPT^P_Q $\left(SPT^C_Q\right)$	SPT^C_Q $\left(CPT^P_K\right)$	CPT^P_K $\left(CPT^C_Q\right)$	CPT^P_K $\left(SPT^C_Q\right)$	SPT^P_K $\left(SPT^P_Q\right)$
DGH0	**QU82**	0 -	0 -	0 -	0 -	0 -	0 -	SPT^D_Q $\left(SPT^P_Q\right)$	SPT^C_Q $\left(SPT^P_K\right)$	SPT^C_Q $\left(SPT^P_Q\right)$	CPT^D_Q $\left(CPT^C_Q\right)$	CPT^D_Q $\left(CPT^C_Q\right)$	SPT^D_Q $\left(SPT^C_Q\right)$
	KT92	0 -	0 -	0 -	0 -	0 -	0 -	SPT^D_K $\left(SPT^C_Q\right)$	SPT^D_K $\left(CPT^P_K\right)$	SPT^C_Q $\left(CPT^D_K\right)$	CPT^D_K $\left(CPT^P_Q\right)$	CPT^D_K $\left(SPT^P_K\right)$	SPT^P_K $\left(CPT^P_K\right)$

also affect the precision of our estimators. If we expect some deviations from the implicitly inherent constraining parameter constellations in terms of $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$, then we obtain a ranking of utility models where nesting models should prevail in the upper ranks.

A closer inspection of Table (7) reveals, for an increase in the dispersion of the error term (increase in σ_ϵ), a tendency toward RDU, where EUT is the correct underlying model. Likewise, we detect a tendency to select exponential power functionals over power-functionals, which implies that the scaling parameter ρ in the $EXPO$ utility functional is significantly distinct from zero—the case where $EXPO$ coincides with $CRRA$.[46] We find that in those cases where the expo-power function obtains the top rank, the scaling parameter ρ is significantly different from zero according to a t-test (p-value 0.031), an indication that the expo-power function does not coincide with a $CRRA$ utility, although we would expect ρ not to be significantly distinct from zero.[47] The tendency to rank nesting utility models to nested models if the true underlying roundtrip sequence is determined by a nested utility model also carries over to the case of non-expected utility investor trading decisions. For instance, if the decision process is determined by RDU preferences with a $CRRA$ value function and decision weights according to (A.6), we detect significant differences of $\hat{\boldsymbol{\theta}}_{\boldsymbol{RDU}}$ from the true parameterization $\boldsymbol{\theta}_{\boldsymbol{RDU}}$. For example, for EUT, we expect γ to be close to 1, as EUT and RDU coincide if $\gamma = 1$ and for RDU, parameter γ should range near 0.65; however, in cases where EUT is the true utility model but RDU obtains the highest rank, γ is near 1 (mean value for γ is 0.92), but the hypothesis that γ is 1 is rejected at a 5%-significance level (p-value 0.049 according to an independent t-test). The decision weight $\omega(\hat{p}_{j,t}|\hat{\boldsymbol{\theta}}_{\boldsymbol{RDU}})$ seems to capture parts of the risk aversion of the EUT model as estimates for δ are close to risk neutrality (p-value 0.122). Consequently, $\log(L(\Delta_t(U_{RDU}|\hat{\boldsymbol{\theta}}_{\boldsymbol{RDU}})))$ obtains higher values than the corresponding likelihood function of EUT, although EUT is the true underlying utility model. In conclusion, the likelihood measure seems to identify models that are close to the true model, but are disturbed by the inherent imprecision of the parameter estimates of the expo-power function being significantly different from zero.

7. Conclusion

Focusing on the preferences of private investors in stock markets is fundamental to any research in finance. In particular, we would like to understand why individuals act in the way they do and we would like to assist them by providing normative guidelines toward better (or optimal) behavior. Financial theory implicitly assumes utility maximization to obtain pricing kernels, from which standard tools, such as stochastic discount factors, can be deployed to better understand how securities are priced and which class of investor assets are considered most valuable. To trace

[46] Note that the correct specification is a $CRRA$ utility function such that $\log L(\Delta_t(U_{EUT}|\hat{\boldsymbol{\theta}}_{\boldsymbol{EUT}}))$ is theoretically close to zero. We model the expected utility case explicitly using $CRRA$ as a ranking tendency toward expo-power utility, which may strengthen our conjecture that nesting models are systematically preferred. Inspecting the log-likelihood values and the information criterion of our simulations, the best-fitting model reveals an $AICC$ indeed fairly close to zero (a small differential is due to sample size and parameter correction), a result that is independent from the period for which we calculate returns, which is a result of the inherent horizon independence of $CRRA$ utility models (Merton (1969)).

[47] In this case, the solution would be acceptable if this estimator were not significantly different from zero, since it would imply that $CRRA$ is the true model.

an arc from the literature on investor behavior in financial markets to methods developed and applied in experimental economics, we first present a short summary on what we know about preferences and utility functions in financial markets and from where potential future streams of literature may spring. Then, we present and adopt a popular and frequently applied econometric method from experimental economics, namely, likelihood estimation and an application for model selection purposes. We check, based on a simulation study, whether a naive implementation of this method provides reliable results to address the research question, that is, to what extent it allows us to identify the correct utility model.

We find that for a very broad classification of utility models, this method provides acceptable outcomes. Yet, a closer look at the preference parameters reveals several caveats that come along with typical issues attached to financial data, and these issues may drive our results. In particular, deviations are attributable to effects stemming from multicollinearity and its concurrent parameter identification problems, where some of these detrimental effects can be captured up to a certain degree by adjusting the error term specification. Furthermore, additional uncertainty stemming from changing market parameter estimates affects the precision of our estimates for risk preferences and cannot simply be remedied by using a higher standard deviation of the error term or a different assumption regarding the stochastics of the error term. In particular, if the variance of the error term becomes large, we detect a tendency to identify SPT as the utility model providing the best fit to data. We also find that a frequent issue, namely, serial correlation of the residuals, does not seem to be significant. However, we detect a tendency to prefer nesting models over nested utility models, which is particularly prevalent if RDU and $EXPO$ utility models are estimated along with EUT and $CRRA$ utility models.

Appendix A. Utility Functions used in the Simulation Study

To substantiate the set of utility functions that specifies $\Delta(U_k|\boldsymbol{\theta_k})$, we consider several utility functions frequently used in the literature. For an *expected utility*-type investor (EUT), the preference over the risky outcomes of the stock are modeled as

$$U_{EUT}(W_t|\boldsymbol{\theta}_{EUT}) = \sum_{j=1}^{t+1} \hat{p}_{j,t} u_{EUT}(W_t|\boldsymbol{\theta}_{EUT}), \tag{A.1}$$

where $\hat{p}_{j,t}$ denotes the respective probabilities associated with the respective state.[48] We denote the utility functional as u_{EUT} given the expo-power specification proposed by Saha (1993)[49]

$$u_{EUT}(W_t|\boldsymbol{\theta}_{EUR}) = 1 - e^{-\rho(W_t \hat{R}_{U,t}^{t-j+1} \hat{R}_{D,t}^{j-1})^{1-\delta}} \rho^{-1}, \tag{A.2}$$

where ρ governs relative and δ governs absolute risk aversion such that the term $u_{EUT}(W_t|\boldsymbol{\theta}_{EUT})$ exhibits the properties of DARA for $1-\delta$ below unity and captures the behavior of CRRA if $1-\delta = 1$ and IARA for $1-\delta$ above 1. Regarding parameter ρ, functional (A.2) displays features of DRRA for $\rho < 0$ and IRRA if $\rho > 1$ (see also Saha et al. (1994), footnote 2). It is well established that $u_{EUT}(W_t|\boldsymbol{\theta}_{EUT})$ converges to CRRA utility if ρ approaches zero. To benchmark this case and to test for the hypothesis that the imprecision in estimating risk parameters is due to multicollinearity and its favoring of nesting utility, we explicitly model CRRA utility, where we specify the utility functional as

$$u_{EUT}(W_t|\boldsymbol{\theta}_{EUT}) = (W_t \hat{R}_{U,t}^{t-j+1} \hat{R}_{D,t}^{j-1})^{1-\delta}(1-\delta)^{-1}. \tag{A.3}$$

Note that for $\delta = -1$, expression (A.3) covers mean–variance preferences (see for a proof Back (2012); for other characteristics regarding δ see Gollier (2001)).

To capture the possible existence of generalized expected utility theories, we also consider an investor with *rank-dependent utility* (RDU) according to Quiggin (1993) and Wakker (1994), where the utility obtained from the risky asset is explicated as

$$U_{RDU}(W_t|\boldsymbol{\theta}_{RDU}) = \sum_{j=1}^{t+1} \pi_{j,t}(\Delta\omega(\hat{p}_{j,t}|\boldsymbol{\theta}_{RDU}))u_{RDU}(W_t|\boldsymbol{\theta}_{RDU}). \tag{A.4}$$

To return from this generalized version to expected utility as presented above, we use the utility functionals as presented in equations (A.2) and (A.3) and define the decision weights $\pi_{j,t}(\omega(\hat{p}_{j,t}|\boldsymbol{\theta}_{RDU}))$ as decumulative probability transformation functions according to Abdellaoui (2000) in the specification of the econometric model, to keep our results comparable to experimental evidence.[50] The probability

[48] Technically, we specify the state probabilities as $\hat{p}_{j,t} = \binom{t+1}{j}\hat{p}_t^{t-j+1}(1-\hat{p}_t)^{j-1}$. In an earlier version of our program, we calculated the respective values of $\hat{p}_{j,t}$ using Feller's famous *Reflection Principle* (Feller (1968)) as $\hat{p}_{j,t} = \left[\binom{t}{j} - \binom{t}{j-1}\right]\hat{p}_t^{t-j+1}(1-\hat{p}_t)^{j-1}$. In testing our program, we find virtually no differences in the generated results between both specifications, such that we opted for the simpler binomial version of $\hat{p}_{j,t}$.

[49] In the original paper, Saha (1993) suggests an exponential-power utility functional $u_{EUT}(W_t|\boldsymbol{\theta}_{EUT}) = c - e^{-\rho(W_t \hat{R}_{U,t}^{t-j+1} \hat{R}_{D,t}^{j-1})^{\delta}}$ where δ and ρ are the respective parameters of this functional and c denotes a constant. As Saha remarks (p. 906), setting $c = 1$ does not play a role in the characterization of risk attitudes or choices. In current research on asset pricing, this constant is usually set equal to zero and thus ignored.

[50] More precisely, the decision weights are specified as $\pi_{j,t}(\Delta\omega(\hat{p}_{j,t}|\boldsymbol{\theta}_{RDU})) = \omega\left(\sum_{k=j}^{t+1}\hat{p}_{k,t}\right) - \omega\left(\sum_{k=j+1}^{t+1}\hat{p}_{k,t}\right)$ for $j \leq t$ and for the highest node $\pi_{j,t}(\Delta\omega(\hat{p}_{j,t}|\boldsymbol{\theta}_{RDU})) = \omega(\hat{p}_{t+1,t})$ for $j = t+1$. In an earlier version of our program, we also implemented the logically equivalent representation with a cumulative probability transformation function $\pi_{j,t}(\Delta\omega(\hat{p}_{j,t}|\boldsymbol{\theta}_{RDU})) =$

transformation function $\omega(\hat{p}_{j,t}|\boldsymbol{\theta}_{RDU})$, for which we identify two dominantly used versions from the experimental literature represents a central ingredient of RDU. Following the original article of Quiggin (1982), the decision weights are adapted from Karmarkar (1978) and Karmarkar (1979) and defined as

$$\omega(\hat{p}_{j,t}|\boldsymbol{\theta}_{RDU}) = \hat{p}_{j,t}^{\gamma}(\hat{p}_{j,t}^{\gamma} + (1-\hat{p}_{j,t})^{\gamma})^{-1}. \tag{A.5}$$

It is noteworthy that the decision weights sum up to 1 and depend on the ranking of the payoffs of the risky asset, even for $\gamma \neq 1$. Further, if γ is equal to 1, RDU converges to EUT and the usual characterizations apply (Levy and Levy (2002a)). Studies on the impact of probability weighting, such as Barberis and Huang (2008) and Barberis (2012), instead impose a nonlinear weighting scheme, which can be reflected by the specification of $\pi_{j,t}(\omega(\hat{p}_{j,t}|\boldsymbol{\theta}_{RDU}))$ as in Kahneman and Tversky (1979), explicitly stated in Tversky and Kahneman (1992), and used in Wu and Gonzalez (1996) whereas

$$\omega(\hat{p}_{j,t}|\boldsymbol{\theta}_{RDU}) = \hat{p}_{j,t}^{\gamma}(\hat{p}_{j,t}^{\gamma} + (1-\hat{p}_{j,t})^{\gamma})^{-\frac{1}{\gamma}}, \tag{A.6}$$

in which the decision weights do not sum to unity. Note that the utility functional $u_{EUT}(W_t|\boldsymbol{\theta}_{RDU})$ remains the same as defined in EUT, since RDU reduces to expected utility if there is no probability weighting (e.g., if $\gamma = 1$ such that $\pi_{j,t}(\hat{p}_{j,t}|\boldsymbol{\theta}_{RDU}) = \hat{p}_{j,t} \; \forall \; \hat{p}_{j,t} \in (0,1)$).

As mentioned, some empirical studies on financial decision making suggest that there is some evidence that Prospect Theory may be at work in financial markets such that it seems advisable to model an investor's preferences toward financial outcomes according to the original formulation of prospect theory, namely *simple prospect theory* (SPT) as

$$U_{SPT}(W_t, W_{RP}|\boldsymbol{\theta}_{SPT}) = \sum_{j=1}^{t+1} \pi_{j,t}(\omega(\hat{p}_{j,t}|\boldsymbol{\theta}_{SPT}))u_{SPT}(W_t, W_{RP}|\boldsymbol{\theta}_{SPT}), \tag{A.7}$$

where W_{RP} marks a reference point (Kahneman and Tversky (1979)), assuming that preferences of the individual investor are based on changes of the initially invested wealth W_0 (for deviating reference points in stock markets such as historical extrema in prices Grinblatt and Keloharju (2001b), Garvey and Murphy (2004) or expectations Meng (2010)) and in which the decision weights $\pi_{j,t}(\omega(\hat{p}_{j,t}|\boldsymbol{\theta}_{SPT}))$ are defined for each possible state (Kahneman and Tversky (1979)).[51] In the original formulation of Kahneman and Tversky (1979) and as adapted in studies on various issues in finance (Berkelaar et al. (2004) with curvature parameter equal to 1, Berkelaar and Kouwenberg (2009), Kliger and Levy (2009) and others) the value function is captured by a power functional $u_{SPT}(W_t|\boldsymbol{\theta}_{SPT})$ of the form

$$u_{SPT}(W_t, W_{RP}|\boldsymbol{\theta}_{SPT}) = \lambda^{I[\Delta W_t < 0]}(|W_t \hat{R}_{U,t}^{t-j+1} \hat{R}_{D,t}^{j-1} - W_{RP}|)^{\alpha}, \tag{A.8}$$

where $I[\Delta W_t < 0]$ represents an indicator taking the value of 1 if the change in wealth, measured as difference from the reference point $\Delta W_t = W_t \hat{R}_{U,t}^{t-j+1} \hat{R}_{D,t}^{j-1} - W_{RP}$, is negative and zero otherwise, indicating states where losses, weighted with loss aversion parameter λ, occur.

$\omega\left(\sum_{k=j}^{t+1} \hat{p}_{k,t}\right) - \omega\left(\sum_{k=j}^{t} \hat{p}_{k,t}\right)$ for $j > 1$ and $\pi_{j,t}(\Delta\omega(\hat{p}_{j,t}|\boldsymbol{\theta}_{RDU})) = \omega(\hat{p}_{1,t})$ for $j = 1$ as in the original formulation of RDU according to Quiggin (1982) and Quiggin (1993). A test of our program reveals that the respective formulation of the decision weights appears to have no impact on our results, such that we refrain from an explicit distinction between these two possible cases.

[51] In detail, we establish a logical connection to RDU and define the decision weights as $\pi_{j,t} = \omega\left(\left(\sum_{k=j}^{t+1} \hat{p}_{k,t}\right) - \left(\sum_{k=j+1}^{t+1} \hat{p}_{k,t}\right)\right) \forall j$.

Some financial studies, such as Barberis et al. (2001), Gomes (2005) and Barberis and Huang (2008) model the demand of SPT-type investors according to a different form of the value functional and apply a mathematical construct similar to CRRA utility, specified as

$$u_{SPT}(W_t, W_{RP}|\boldsymbol{\theta}_{SPT}) = \lambda^{I[\Delta W_t<0]}(|W_t\hat{R}_{U,t}^{t-j+1}\hat{R}_{D,t}^{j-1} - W_{RP}|)^{1-\delta}(1-\delta)^{-1}. \quad \text{(A.9)}$$

It should be noted that, as in Barberis et al. (2001) and Barberis and Huang (2008), the individual's consumption enters as arguments, although other variables and fundamentals may be considered (e.g., Barberis and Xiong (2009)). For *cumulative prospect theory* (CPT), we calculate the utility of the financial prospects according to Tversky and Kahneman (1992) as

$$U_{CPT}(W_t, W_{RP}|\boldsymbol{\theta}_{CPT}) = \sum_{j=1}^{t+1} \pi_{j,t}(\Delta\omega(\hat{p}_{j,t}|\boldsymbol{\theta}_{CPT}))u_{CPT}(W_t, W_{RP}|\boldsymbol{\theta}_{CPT}), \quad \text{(A.10)}$$

with value functionals as defined in (A.8) and (A.9). Characteristic for CPT and distinct from SPT, the difference of the probability weighting functions $\omega(\hat{p}_{j,t})$ constitute rank-dependent decision weights as a decumulative function of the state-specific decision weights in the domain of losses and as a cumulative function of the state-specific decision weights if the investor's position in the risky asset generates positive returns (Fennema and Wakker (1997)).[52] It is noteworthy that for SPT and CPT, these decision weights sum to 1 if specified according to (A.5) and are usually subadditive under formulation (A.6) for $\gamma < 1$. Concerning $u_{CPT}(W_t, W_{RP}|\boldsymbol{\theta}_{CPT})$, we use the same specification as for the original version of Prospect Theory.

Appendix B. Remarks on the Maximum Likelihood Approach

As elaborated, experimental studies maximize the overall likelihood of an investor or decision maker, given the assumption of stochastically independent error terms yielding the likelihood function for a utility model of type k, expressed as

$$\log L(\Delta_t(U_k|\boldsymbol{\theta_k})) = \sum_{t\in T}\sum_{I\in I_{k,t}} I_{k,t}\log p_{I_{k,t}}(\Delta_t(U_k|\boldsymbol{\theta_k})),$$

in which it is required that $\Delta_t(U_k|\boldsymbol{\theta_k})$ is a one-to-one relationship connecting the functional values to particular values of $\boldsymbol{\theta}_k$ and where $p_{I_{k,t}}(\Delta_t(U_k|\boldsymbol{\theta_k}))$ denotes the respective conditional probabilities as defined in (3.3). To clarify notation and provided there exists a unique solution to the maximizing problem within the possible range of $\boldsymbol{\theta}_k$, maximizing the likelihood function (B.1) for a given sample and time periods $t \in \{1, \ldots, T\}$ returns a maximum likelihood estimate $\hat{\boldsymbol{\theta}}_{k|n,t}$, depending on the sample size, of the true but unknown parameter $\hat{\boldsymbol{\theta}}_k$, briefly denoted as

$$\hat{\boldsymbol{\theta}}_{\boldsymbol{k|n,t}} = \arg\max_{\theta_k\in\boldsymbol{\theta_k}} \log L(\Delta_t(U_k|\boldsymbol{\theta_k})). \quad \text{(B.1)}$$

Accordingly, the obtained estimator $\hat{\boldsymbol{\theta}}_{\boldsymbol{k|n,t}}$ is characterized by the usual standard conditions concerning the score vector $\boldsymbol{S}(\Delta_t(U_k|\boldsymbol{\theta_k}))$, which should be equal to a zero vector, and the Hessian matrix $\boldsymbol{H}(\Delta_t(U_k|\boldsymbol{\theta_k}))$, consequently being positive

[52] To be more precise, for CPT the decision weights are formulated as decumulative function $\pi_{j,t} = \omega\left(\sum_{k=1}^{\bar{j}}\hat{p}_{k,t}\right) - \omega\left(\sum_{k=1}^{\bar{j}-1}\hat{p}_{k,t}\right) \forall j < \lfloor\bar{j}\rfloor$ and $\pi_{j,t} = \omega(\hat{p}_{1,t})$ if $j = 1$ for positive returns and specified as cumulative function $\pi_{j,t} = \omega\left(\sum_{k=\bar{j}}^{t+1}\hat{p}_{k,t}\right) - \omega\left(\sum_{k=\bar{j}+1}^{t+1}\hat{p}_{k,t}\right) \forall j > \lfloor\bar{j}\rfloor$ and $\pi_{j,t} = \omega(\hat{p}_{t+1,t})$ if $j = t+1$ in the domain of losses where $\bar{j}$ denotes the break-even node that classifies a state to be only in the realm of negative returns.

definite. Ignoring σ_ϵ for a moment and following Edwards (1992), the score vector $\boldsymbol{S}(\Delta_t(U_k|\boldsymbol{\theta_k}))$ is

$$\boldsymbol{S}(\Delta_t(U_k|\boldsymbol{\theta_k})) = \sum_{I \in I_{k,t}} \boldsymbol{\delta}(U_k|\boldsymbol{\theta_k})\boldsymbol{S}(\Delta_t(\boldsymbol{\theta}_k)) \tag{B.2}$$

where we use the abbreviation $\boldsymbol{\delta_t}(U_k|\boldsymbol{\theta_k})$ to denote the square matrix of first derivatives of $\Delta_t(U_k|\boldsymbol{\theta_k})$ with respect to each of its parameters and denote the $(K_k \times 1)$ vector of outer derivatives of the likelihood function as $\boldsymbol{S}(\Delta_t(\boldsymbol{\theta}_k))$, being the product of a diagonal matrix $\boldsymbol{I}$ with elements $I_{k,t}/p_{I_{k,t}}$ and the diagonal matrix $\boldsymbol{P_I}$ containing the outer derivatives of $p_{I_{k,t}}$. Following this notation, the Hessian matrix $\boldsymbol{H}(\Delta_t(U_k|\boldsymbol{\theta_k}))$ consists of two terms, namely a matrix containing partial derivatives of the elements of $\boldsymbol{\delta}(U_k|\boldsymbol{\theta_k})$ and a matrix collecting the second derivatives of $\Delta_t(U_k|\boldsymbol{\theta_k})$ with respect to its parameters (see Edwards (1992) for details).[53]

To obtain the Information matrix $\boldsymbol{I}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}})$, the sign of the Hessian needs to be reversed and taken by its expectations, where we can use the fact that $E(I_{k,t}) = p_{I_{k,t}}$. Since the sum of the choice probabilities equals 1 $\sum_{I \in I_{k,t}} p_{I_{k,t}} = 1$, the last term of the Hessian vanishes if evaluated at $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$ such that the last term can be greatly simplified (Fisher (1956), Edwards (1992), Theorem 7.2.2) to

$$\boldsymbol{I}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_k)) = \sum_{I \in I_{k,t}} \boldsymbol{\delta}(U_k|\boldsymbol{\theta_k})\boldsymbol{I}(\Delta_t(\boldsymbol{\theta_k}))\boldsymbol{\delta}(\boldsymbol{U_k}|\boldsymbol{\theta_k})'. \tag{B.3}$$

Here, $\boldsymbol{\delta_t}(U_k|\boldsymbol{\theta_k})$ denotes the square matrix of first derivatives of $\Delta_t(U_k|\boldsymbol{\theta_k})$ with respect to each of its parameters and $\boldsymbol{I}(\Delta_t(\boldsymbol{\theta_k})) = \boldsymbol{P_I}\boldsymbol{P_I'}\boldsymbol{I}$ being the product of a diagonal matrix $\boldsymbol{I}$ with elements $I_{k,t}/p_{I_{k,t}}$ and the diagonal matrix $\boldsymbol{P_I}$ containing the outer derivatives of $p_{I_{k,t}}$. It is evident from this structure that for each $I_{k,t}$th term, the Hessian is a positive semi-definite matrix since $\boldsymbol{I}(\Delta_t(\boldsymbol{\theta_k})) = \boldsymbol{P_I}\boldsymbol{P_I'}\boldsymbol{I}$ is symmetrical. Disregarding the possibility that $\boldsymbol{H}(\Delta_t(U_k|\boldsymbol{\theta_k}))$ is singular, the Hessian is in fact positive definite. This implies that $\boldsymbol{I}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_k))$ is also a positive definite matrix over reasonable values of $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$.

Appendix C. The Akaike Information Criterion and Model Tests used

According to the likelihood approach described above, it can be shown that a connection to the Akaike Information Criterion (Carbone and Hey (1994), Hey and Orme (1994), Carbone and Hey (1995) and Stott (2006)) can be established and used to identify the best-fitting utility function specification. To sketch this idea, recall that $\log L(\Delta_t(U_k|\boldsymbol{\theta_k}))$ is continuous in $\boldsymbol{\theta_k}$ and twice differentiable.[54] A second-order Taylor-expansion of the log-likelihood (B.1) around $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$ yields

$$\begin{aligned}\log L(\Delta_t(U_k|\boldsymbol{\theta_k})) \approx & \log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}})) + \boldsymbol{S}_{k|n,t}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))+ \\ & + \boldsymbol{H}_{k|n,t}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k|n,t}}) - \Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}})).\end{aligned} \tag{C.1}$$

[53]Note that due to the independence assumption, each element of the score vector and the Hessian matrix consist of a series of sums. This is not surprising since, according to the independence assumption across time and choice sets, the log-likelihood function inherits the regularity property in the sense that differentiation and summation are interchangeable (e.g., Cramer (1986)), which in turn carries over to the entire sample if it holds for any single observation.

[54]As for the derivatives of, for example, SPT towards some elements of $\boldsymbol{\theta_k}$, the derivative is not defined for some combinations of $\hat{R}_{S,t}$. Under a nonlinear probit model such as ours, the normal distribution is continuous, and given that summation and differentiation are interchangeable, the probability that those critical combinations appear converges to zero and thus can be ignored.

If this expression is evaluated at $\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}$, the score vector $\boldsymbol{S}_{k|n,t}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ equals zero and the Hessian can be rewritten as

$$\log L(\Delta_t(U_k|\boldsymbol{\theta_k})) \approx \log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))-$$
$$-\frac{1}{2}nt(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k|n,t}}) - \Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))^{'}\boldsymbol{I}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k|n,t}}) - \Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}})).$$

In this step, we make use of the fact that

$$E\{nt(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k|n,t}}) - \Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))^{'}\boldsymbol{I}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{k|n,t}))(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k|n,t}}) - \Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))\}$$
$$\approx \boldsymbol{tr}(\boldsymbol{J}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{k|n,t}))\boldsymbol{I}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{k|n,t}))^{-1}), \tag{C.2}$$

as shown elsewhere (Bozdogan (2000), Pawitan (2001)). $\boldsymbol{J}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{k|n,t}))$ contains the product of the score vectors (Cramer (1986)) and can be written as $\boldsymbol{J}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{k|n,t})) = E(\boldsymbol{S}_{k|n,t}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))\boldsymbol{S}_{k|n,t}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))^{'})$ and where $\boldsymbol{tr}$ denotes the trace of the product of the matrices within the brackets. Taking expectations, by (C.2), expression (C.1) can be rewritten as

$$nt\bar{L}(\Delta_t(U_k|\boldsymbol{\theta_k})) \approx E(\log L(\Delta_t(U_k|\boldsymbol{\theta_k})))-\frac{1}{2}\boldsymbol{tr}\left(\boldsymbol{J}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{k|n,t}))\boldsymbol{I}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{k|n,t}))^{-1}\right).$$

If the number of observations or traded stocks increases beyond all bounds of the number of days t grows, according to Cramer (1986), $\boldsymbol{J}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{k|n,t})) \approx \boldsymbol{I}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{k|n,t}))$, such that $\boldsymbol{tr}\left(\boldsymbol{I}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{k|n,t}))\boldsymbol{I}(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{k|n,t}))^{-1}\right)$, being the dimension of $\boldsymbol{\theta_k}$, which is approximately equal to the number of parameters K_k of the respective utility model under consideration. Stated differently,

$$\bar{L}(\Delta_t(U_k|\boldsymbol{\theta_k})) \approx -\frac{2\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))}{nt} + \frac{2K_k}{nt}$$

which is the information criterion according to Akaike (1974) in the representation of Amemiya (1980) as stated above, where we correct for the different number of observations by nt. To contrast the results of the AIC by a finite correction version of Sugiura (1978) and Hurvich and Tsai (1989), we also invoke the corrected AIC, abbreviated as AICC, as

$$AICC = -\frac{2\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))}{nt} + \frac{2K_k}{nt} + \frac{2K_k(K_k+1)}{nt(nt-K_k-1)}. \tag{C.3}$$

This form is usually recommended if the number of observations does not outweigh the number of parameters by more than a factor of 40 (Burnham and Anderson (2004)).

One serious drawback of the AIC, AICC, or basically any information criterion, is the fact that it, per se, cannot provide significance levels or statistical statements about how good the discrimination between the two competing models actually is.[55] To obtain the usual significance levels and to retrieve further information

[55] Further, if two utility specifications, for instance model-k and competing model-m, share the same number of risk preference parameters, then sorting them according to AICC leads to a selection of one model against the other. However, we find that, in some cases, where the highest AICC is not close to zero, the order in which these models are sorted can be due to a sufficiently small difference in the obtained log-likelihood, where the correction for the number of parameters strongly affects the ranking and overcompensates the difference in the log-likelihood. We check the ranking of utility models and use the Schwartz Information Criterion (also known as Bayes Information Criterion (Schwarz (1978))) as well as the original AIC. Since the penalty of a higher number of parameters according to the Schwartz Information Criterion is higher, we find that this effect is aggravated, though it affects the results only in exactly those cases pointed out above. In light of such deficiencies, to validate the obtained position, we supplement the results

about the likelihood, the ranking according to the AICC is supplemented by the usual statistical tools made available by likelihood theory. For nested models where model k nests m, the usual likelihood ratio test can be applied (Kent (1982)). According to the null hypothesis H_0, both models are equally good in fitting the observed data such that the unconstrained maximum of $\log L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))$ should be close to the constraint maximum $\log L(\Delta_t(U_m|\hat{\boldsymbol{\theta}}_{\boldsymbol{m}}))$. The likelihood ratio test is calculated as

$$LR_{nt}(\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}, \hat{\boldsymbol{\theta}}_{\boldsymbol{m}}) = -2\ln\left[\frac{L(\Delta_t(U_i|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))}{L(\Delta_t(U_i|\hat{\boldsymbol{\theta}}_{\boldsymbol{m}}))}\right] \text{ where under } H_0 : LR_{nt}(\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}, \hat{\boldsymbol{\theta}}_{\boldsymbol{m}}) \xrightarrow{L} \chi^2_{df_{k,m}} \tag{C.4}$$

for which it is well established that this ratio is non-negative and under H_0 asymptotically chi-square distributed with degrees of freedom $df_{k,m}$ equal to the number of parameters of the unconstrained model minus the number of parameters of the constrained model (for a proof see, e.g., Rao (1973)). To derive p-values for contrasting non-nested models such as CRRA and CPT given $W_0(RP)$, we apply a non-nested likelihood ratio test according to Vuong (1989), their Theorem 5.1., where we denote the maximized likelihood values of competing models k and m, respectively. Vuong's contribution is to show that under general conditions and given that the null hypothesis holds, the expectations of the log-ratio of the two maximized likelihoods $(L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}})))$ and $(L(\Delta_t(U_m|\hat{\boldsymbol{\theta}}_{\boldsymbol{m}})))$ for two competing models k and m should be zero. The expectations can be consistently estimated by the average of the likelihood ratio statistic over nt observations such that given the null hypothesis that the log of the likelihood ratio has an expectation of zero

$$\begin{gathered}\frac{LR_{nt}(\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}, \hat{\boldsymbol{\theta}}_{\boldsymbol{m}})}{(\sqrt{nt})\hat{w}_{nt}} \xrightarrow{d} N(0,1) \text{ with } \frac{1}{nt} LR_{nt}(\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}, \hat{\boldsymbol{\theta}}_{\boldsymbol{m}}) \xrightarrow{L} E_0\left[\ln\frac{L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))}{L(\Delta_t(U_m|\hat{\boldsymbol{\theta}}_{\boldsymbol{m}}))}\right]. \\ \text{where } \hat{w}^2_{nt} \equiv \frac{1}{n}\sum_{i=1} nt\left[\ln\frac{L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))}{L(\Delta_t(U_m|\hat{\boldsymbol{\theta}}_{\boldsymbol{m}}))}\right]^2 - \left[\frac{1}{nt}\sum_{i=1}^{nt}\ln\frac{L(\Delta_t(U_k|\hat{\boldsymbol{\theta}}_{\boldsymbol{k}}))}{L(\Delta_t(U_m|\hat{\boldsymbol{\theta}}_{\boldsymbol{m}}))}\right]^2.\end{gathered} \tag{C.5}$$

for which it is shown that the resulting likelihood ratio statistic is asymptotically normally distributed.[56] If the time series is long enough, the asymptotic properties might hold on the individual level as well.[57]

References

References

Abdellaoui, M. (2000). Parameter-free elicitation of utilities and probability weighting functions. *Management Science 46*, 1497–1512.

Abdellaoui, M., F. Vossmann, and M. Weber (2005). Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. *Management Science 51*, 1384–1399.

by appropriate significance tests between the first and second rank utility models to identify those models that are found to be hardly satisfactorily distinguished according to the AICC.

[56]Technically, the Vuong test specifies that, under the null hypothesis, the expectation of the log ratio is symmetrically distributed around zero. In cases where the log of the likelihood ratio is not close to a normal distribution, alternative non-nested model tests have been proposed (e.g., Clarke (2003) and Clarke (2007)).

[57]The fact that we use time-series observations does not invalidate the possibility of applying the AIC, since Akaike (1978) provides a time-series version of the information criterion.

Adda, J. and R. Cooper (2003). *Dynamic Economics Quantitative Methods and Applications*. MIT Press.

Akaike, H. (1973). *Information Theory and an extension of the maximum likelihood principle, in: B.N. Petrox and F. Caski (Eds.) Second International Symposium on Information Theory*. Akademiai Kiado.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control AC-19*, 716–723.

Akaike, H. (1978). On the likelihood of a time series model. *The Statistician 27(4)*, 217–235.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulates et axioms de l'ecole americaine. *Econometrica 21*, 503–546.

Amemiya, T. (1975). Qualitative response models. *Annals of Economic and Social Measurement 4(3)*, 363–372.

Amemiya, T. (1980). Selection of regressors. *International Economic Review 21(2)*, 331–354.

Amemiya, T. (1981). Qualitative response models: A survey. *Journal of Econometric Literature 19*, 1483–1536.

Amemiya, T. (1985). *Advanced Econometrics*. Harvard University Press.

Andreassen, P. B. (1987). On the social psychology of the stock market: Aggregate attributional effects and the regressiveness of prediction. *Journal of Personality and Social Psychology 53*, 490–496.

Andreassen, P. B. (1988). Explaining the price-volume relationship: The difference between price changes and changing prices. *Organizational Behavior and Human Decision Processes 41*, 371–389.

Andreassen, P. B. and S. J. Kraus (1990). Judgmental extrapolation and the salience of change. *Journal of Forecasting 9(4)*, 347–372.

Back, K. E. (2012). *Asset Pricing and Portfolio Choice Theory*. Oxford University Press.

Baker, M. and J. Wurgler (2004). A catering theory of dividends. *Journal of Finance 59(3)*, 1125–1165.

Ballinger, T. P. and N. T. Wilcox (1997). Decision, error and heterogeneity. *Economic Journal 107*, 1090–1105.

Barber, B. and T. Odean (2000). Trading is hazardous to your wealth: The common stock performance of individual investors. *Journal of Finance 55(2)*, 773–806.

Barber, B. and T. Odean (2002). Online investors: Do the slow die first? *Review of Financial Studies 15(2)*, 455–487.

Barber, B. and T. Odean (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies 21(2)*, 785–818.

Barber, B. M. and T. Odean (1999). The courage of misguided convictions. *Financial Analysts Journal 55*, 41–55.

Barberis, N. C. (2012). A model of casino gambling. *Management Science 58(1)*, 35–51.

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *Working paper, Yale University*.

Barberis, N. C. and M. Huang (2001). Mental accounting, loss aversion, and individual stock returns. *Journal of Finance 56*, 1247–1292.

Barberis, N. C. and M. Huang (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review 98(5)*, 2066–2100.

Barberis, N. C. and M. Huang (2009). Preferences with frames: A new utility specification that allows for the framing of risks. *Journal of Economic Dynamics and Control 33(8)*, 1555–1576.

Barberis, N. C., M. Huang, and T. Santos (2001). Prospect theory and asset prices. *Quarterly Journal of Economics 141*, 1–53.

Barberis, N. C., M. Huang, and R. H. Thaler (2001). How distance, language, and culture influence stockholdings and trades. *Journal of Finance 56(3)*, 1053–1073.

Barberis, N. C. and R. Thaler (2003). *A survey of behavioral finance, in: Constantinides, M. Harris, and R. Stulz (editors) Handbook of the Economics of Finance*. Butterworth Heinemann.

Barberis, N. C. and W. Xiong (2009). What drives the disposition effect? an analysis of a long-standing preference-based explanation. *Journal of Finance 64*(2), 751–785.

Bath, C. (2001). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research B 35*, 677–693.

Baum, C. F. (2006). *An Introduction to Modern Econometrics Using Stata*. Stata-Press.

Becker, G. DeGroot, M. and J. Marschak (1963). Stochastic models of choice behavior. *Behavioral Science 8*, 41–55.

Belsley, D. A., E. Kuh, and R. E. Welsch (2004). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: John Wiley and Sons.

Ben-Akiva, M. and S. R. Lerman (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press.

Benartzi, S. and R. H. Thaler (1995). Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics 110*, 75–92.

Berkelaar, A. B. and R. Kouwenberg (2009). From boom 'til bust: How loss aversion affects asset prices. *Journal of Banking and Finance 33*, 1005–1013.

Berkelaar, A. B., R. Kouwenberg, and T. Post (2004). Optimal portfolio choice under loss aversion. *Review of Economics and Statistics 86*, 973–987.

Berndt, E. R., B. H. Hall, R. E. Hall, and J. A. Hausman (1974). Estimation and inference in non-linear structural models. *Annals of Economic and Social Measurement 3*, 653–665.

Black, F. (1986). Noise. *Journal of Finance 41*, 529–543.

Blackburn, D. W. and A. D. Ukhov (2006). Estimating preferences towards risk: evidence from dow jones. *Working Paper, University of Indiana*.

Blake, D. (1996). Efficiency, risk aversion and portfolio insurance: An analysis of financial asset portfolios held by investors in the united kingdom. *The Economic Journal 62(1)*, 107–118.

Bleichrodt, H. and J. L. Pinto (2000). A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management Science 46*, 1485–1496.

Blondel, S. (2002). Testing theories of choice under risk: Estimation of individual functionals. *Journal of Risk and Uncertainty 24(3)*, 251–265.

Blume, M. E. and I. Friend (1975). The asset structure of individual portfolios and some implications for utility functions. *Journal of Finance 30(2)*.

Booij, A. S., B. M. S. van Praag, and G. van de Kuilen (2009). A parametric analysis of prospect theory's functionals for the general population. *Decision Theory 68*, 115–148.

Box, G. E. P., G. M. Jenkins, G. C. rReinsel, and G. M. Ljung (2015). *Time Series Analysis: Forecasting and Control, 5th edition*. Hoboken, New Jersey:

John Wiley and Sohn.

Bozdogan, H. (2000). Akaike's information criterium and recent developments in information complexity. *Journal of Mathematical Psychology 44*, 62–91.

Broihanne, M.-H., M. Merli, and P. Roger (2008). Solving some financial puzzles with prospect theory and mental accounting: A survey. *Revue d'economie politique 118*, 475–512.

Bruckner, R., P. Lehmann, M. H. Schmidt, and R. Stehle (2015). Non-u.s. multi-factor data sets schould be used with caution. *Working Paper, Humbold University Berlin*.

Brunnermeier, M. and S. Nagel (2008). Do wealth fluctuations generate time-varying risk aversion? micro-evidence on individuals' asset allocation. *American Economic Review 98(3)*, 713–736.

Burnham, K. P. and D. R. Anderson (2004). Multimodel inference-understanding aic and bic in model selection. *Sociological Methods and Research 33(2)*, 261–304.

Camerer, C. and H. C. Kunreuther (1989). Subjectively weighted utility: A descriptive extension of the expected utility model. *Journal of Risk and Insurance 2*, 265–300.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997). *The Econometrics of Financial Markets*. Princeton University Press.

Carbone, E. and J. D. Hey (1994). Discriminating between preference functionals. *Journal of Risk and Uncertainty 8*, 223–242.

Carbone, E. and J. D. Hey (1995). A comparison of the estimates of eu and non-eu preference functionals using data from pairwise choice and complete ranking experiments. *Geneva Papers on Risk and Insurance Theory 20*, 111–133.

Carbone, E. and J. D. Hey (2000). Which error story is best. *Journal of Risk and Uncertainty 20*, 161–176.

Chiappori, P.-A. and M. Paiella (2011). Relative risk aversion is constant: evidence from panel data. *Journal of the European Economic Association 9(6)*, 1021–1052.

Chui, M. W. (2001). An experimental study of the disposition effect: Evidence from macao. *Journal of Psychology and Financial Markets 2(4)*, 215–221.

Clarke, K. A. (2003). Nonparametric model discrimination in international relations. *Journal of Conflict Resolution 47*, 72–93.

Clarke, K. A. (2007). A simple distribution-free test for non-nested model selection. *Political Analysis 15*, 347–363.

Coval, J. D. and T. Shumway (2005). Do behavioral biases affect prices? *Journal of Finance 60(1)*, 1–34.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society Series B (Methodological) 34*, 187–220.

Cox, J. C., S. A. Ross, and M. Rubinstein (1979). Option pricing: A simplified approach. *Journal of Financial Economics 7*, 229–264.

Cramer, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press.

Cramer, J. S. (1986). *Econometric Applications of Maximum Likelihood Methods*. Cambridge University Press.

Currim, I. and R. K. Sarin (1989). Prospect versus utility. *Management Science 35*, 22–41.

de Palma, A., M. Ben-Akiva, D. Brownstone, C. Holt, T. Magnac, D. McFadden, P. Moffatt, N. Picard, K. Train, P. Wakker, and J. Walker (2008). Risk, uncertainty and discrete choice models. *Marketing Letters 19*, 269–285.

DeBondt, W. (1993). Betting on trends: Intuitive forecasts of financial return and risk. *International Journal of Forecasting 9*, 355–371.

DeGiorgi, E. and T. Hens (2006). Making prospect theory fit for finance. *Journal of Financial Markets and Portfolio Management 20(3)*, 339–360.

Dhar, R. and A. Kumar (2002). A non-random walk down the main street: Impact of price trends on trading decisions of individual investors. *Working Paper, Yale University, New Heaven, CT*.

Dhar, R. and M. Zhu (2006). Up close and personal: An individual level analysis of the disposition effect. *Management Science 52(5)*, 726–740.

Dhrymes, P. J. (1971). *Distributed Lags- Problems of estimation and Formulation.* Holden Day.

Dimson, E., P. Marsh, and M. Staunton (2000). Risk and return in the 20th and 21st centuries. *Business Strategy Review 11(2)*, 1–18.

Dimson, E., P. Marsh, and M. Staunton (2003). Global evidence on the equity risk premium. *Journal of Applied Corporate Finance 15(4)*, 26–38.

Duffie, D. (2001). *Dynamic asset pricing theory, 3rd. ed.* Princeton University Press.

Ebert, S. and P. Strack (2009). An experimental methodology testing for prudence and third-order preferences. *Working paper, Bonn University*.

Ebert, S. and P. Strack (2012). Until the bitter end: On prospect theory in the dynamic context. *Working paper, Bonn University*.

Eckstein, Z. and K. I. Wolpin (1989). The specification and estimation of dynamic stochastic discrete choice models. *Journal of Human Resources 24(4)*, 562–598.

Edwards, A. W. F. (1992). *Likelihood, expanded edition.* Johns Hopkins University Press.

Edwards, K. D. (1996). Prospect theory-a literature overview. *International Review of Financial Analysis 5*, 18–38.

Edwards, W. (1953). Probability-preferences in gambling. *American Journal of Psychology 66(3)*, 349–364.

Edwards, W. (1954). Probability-preferences among bets with differing expected values. *American Journal of Psychology 67(1)*, 56–67.

Edwards, W. (1962). Subjective probabilities inferred from decisions. *Psychological Review 69(2)*, 109–135.

Ehm, C., C. Kaufmann, and M. Weber (2012). Investors care about risk, but can't cope with volatility. *Working Paper, Mannheim University*.

Fama, E. and K. R. French (1993). Common risk factors in the returns on stock and bonds. *Journal of Financial Economics 33(1)*, 3–56.

Fechner, G. (1966). *Elements of Psychophysics.* Holt, Rinehart and Winston.

Feller, W. (1968). *An introduction to probability theory and its applications, vol. 1, 3rd. ed.* New York, Wiley.

Fennema, H. and M. van Assen (1999). Measuring the utility of losses by means of the trade-off method. *Journal of Risk and Uncertainty 17(3)*, 277–295.

Fennema, H. and P. P. Wakker (1997). Original and cumulative prospect theory: A discussion of empirical differences. *Journal of Behavioral Decision Making 10*, 53–64.

Ferris, S. P., R. A. Haugen, and A. L. Makhija (1988). Predicting contemporary value at differential price levels: Evidence supporting the disposition effect. *Journal of Finance 43*, 677–697.

Fishburn, P. C. and G. Kochenberger (1979). Two-piece von neumann-morgenstern utility functions. *Decision Sciences 10*, 503–518.

Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London A222*, 309–368.

Fisher, R. A. (1956). *Statistical Methods and Scientific Inference.* Oliver and Boyd.

Fletcher, R. (1980). *Practical Methods of Optimization.* Wiley.

Fomby, T. B., R. Hill, and S. R. Johnson (1978). An optimality property of principal component regression. *Journal of the American Statistical Association 73*, 191–193.

Fox, C. and A. Tversky (1998). A belief-based account of decision under uncertainty. *Management Science 44(7)*, 879–895.

Frazzini, A. (2006). The disposition effect and underreaction to news. *Journal of Finance 61*, 2017–2046.

Friedman, M. and L. J. Savage (1948). The utility analysis of choices involving risk. *Journal of Political Economy 56(4)*, 279–304.

Friend, I. and M. E. Blume (1975). The demand for risky assets. *American Economic Review 65(5).*

Frino, A., D. Johnstone, and H. Zheng (2004). The propensity of local traders in futures markets to ride losses: Evidence of irrational or rational behavior? *Journal of Banking and Finance 28*, 353–372.

Garvey, R. and A. Murphy (2004). Are professional traders too slow to realize their losses? *Financial Analysts Journal 60*, 35–43.

Genesove, D. and C. Mayer (2001). Loss aversion and seller behavior: Evidence from the housing market. *Quarterly Journal of Economics 116*, 1233–1260.

Glaser, M., M. Noeth, and M. Weber (2004). *Behavioral Finance, in: D.J. Koehler, and N.Harvey (ed.), Blackwell Handbook of Judgment and Decision Making.* Blackwell.

Glaser, M. and M. Weber (2007). Overconfidence and trading volume. *Geneva Risk and Insurance Review 32(1)*, 1–36.

Gneezy, U. and J. Potters (1997). An experiment on risk taking and evaluation periods. *Quarterly Journal of Economics 112(2)*, 631–645.

Goetzman, W. N. and R. G. Ibbotson (2005). *The Equity Risk Premium: Essays and Explorations.* Oxford University Press.

Gollier, C. (2001). *The Economics of Risk and Time.* MIT Press.

Gomes, F. J. (2005). Portfolio choice and trading volume with loss-averse investors. *Journal of Business 78*, 675–706.

Gonzalez, R. and G. Wu (1999). On the shape of the probability weighting function. *Cognitive Psychology 38(1)*, 129–166.

Gordon, M. J., G. Paradis, and C. H. Rorke (1972). Experimental evidence on alternative portfolio decision rules. *American Economic Review 62(1)*, 107–118.

Gould, W., J. Pitblado, and W. Sribney (2006). *Maximum Likelihood Estimation with Stata.* Stata Press.

Greene, W. H. (2008). *Econometric Analysis.* Pearson Upper Saddle River.

Griffiths, W. E., R. C. Hill, and P. J. Pope (1987). Small sample properties of probit model estimators. *Journal of the American Statistical Association 82(399)*, 929–937.

Grinblatt, M. and B. Han (2005). Prospect theory, mental accounting, and the disposition effect. *Journal of Financial Economics 56*, 589–616.

Grinblatt, M. and M. Keloharju (2000). The investment behavior and performance of various investor types: A study of finlands unique data set. *Journal of Financial Economics 55*, 43–67.

Grinblatt, M. and M. Keloharju (2001a). How distance, language, and culture influence stockholdings and trades. *Journal of Finance 56(3)*, 1053–1073.

Grinblatt, M. and M. Keloharju (2001b). What makes investors trade? *Journal of Finance 56(2)*, 589–616.

Grinblatt, M. and M. Keloharju (2001c). What makes investors trade? *Journal of Finance 56*, 589–616.

Guiso, L. and M. Paiella (2008). Risk aversion, wealth and background risk. *Journal of the European Economic Association 6(6)*, 1109–1150.

Haigh, M. S. and J. A. List (2005). Do professional traders exhibit myopic loss aversion? *Journal of Finance 60(1)*, 523–534.

Hakansson, N. H. (1970). Friedman-Savage utility functions consistent with risk aversion. *Quarterly Journal of Economics 84(3)*, 472–487.

Hakansson, N. H. (1971). Capital growth and the mean-variance approach to portfolio selection. *Journal of Financial and Quantitative Analysis 6(1)*, 517–557.

Halton, J. (1960). On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik 2*, 84–90.

Harless, D. W. and C. Camerer (1994). The predictive utility of generalized expected utility theories. *Econometrica 62(2)*, 1251–1289.

Harrison, G. W. (2008). Maximum likelihood estimation of utility functions using stata. *Working paper, University of central Florida*.

Harrison, G. W. and E. E. Rutstrom (2008). Risk aversion in the laboratory, in:. *Research in Experimental Economics 12*, 41–196.

Harrison, G. W. and E. E. Rutstrom (2009). Expected utility theory and prospect theory: one wedding and a decent funeral. *Experimental Economics 12*, 133–158.

Heath, C., S. Huddart, and M. Lang (1999). Psychological factors and stock option exercise. *Quarterly Journal of Economics 114(2)*, 601–627.

Heisler, J. (1994). Loss aversion in a futures market: An empirical test. *Review of Futures Markets 13*, 793–822.

Hensher, D. A. and L. W. Johnson (1981). *Applied Discrete Choice Modelling.* Wiley and Sons.

Hershey, J. C. and P. J. H. Schoemaker (1980). Risk taking and problem context in the domain of losses: An expected utility analysis. *Journal of Risk and Insurance 47(1)*, 111–132.

Hey, J. D. (1995). Experimental investigations of errors in decision making under risk. *European Economic Review 39*, 633–640.

Hey, J. D. (2002). Experimental economics and the theory of decision making under uncertainty. *Geneva Papers on Risk and Insurance Theory 27(1)*, 5–21.

Hey, J. D. and C. Orme (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica 62*, 1296–1326.

Hong, D. and A. Kumar (2002). What induces noise trading around public announcement events? *Working Paper, Cornell University*.

Hosmer, D. W., S. Lemenshow, and R. X. Sturdivant (2013). *Applied Logistic Regression.* Wiley Series in Probability and Statistics.

Hull, J. and A. White (1988). The pricing of options on assets with stochastic volatilities. *Journal of Financial and Quantitative Analysis 23*, 237–251.

Hurvich, C. M. and C. L. Tsai (1989). Regression and time series model selection in small samples. *Biometrica 76*, 297–307.

Ingersoll, J. E. (1987). *Theory of Financial Decision Making.* Rowman and Littlefield.

Jackwerth, J. (2000). Recovering risk aversion from option prices and realized returns. *Review of Financial Studies 13(2)*, 433–451.

Johnson, R. S., J. E. Pawlukiewicz, and J. M. Metha (1997). Binomial option pricing with skewed asset returns. *Review of Quantitative Finance and Accounting 9*, 89–101.

Johnson, R. S., A. Sen, and B. Balyeat (2012). A skewness-adjusted binomial model for pricing futures options-the importance of the mean and carrying-cost parameters. *Journal of Mathematical Finance 2*, 105–120.

Jordan, D. and J. D. Diltz (2004). Day traders and the disposition effect. *Journal of Behavioral Finance 5(4)*, 192–200.

Judge, G. G., W. E. Griffiths, R. C. Hill, H. Luetkepohl, and T. C. Lee (1985). *The Theory and Practice of Econometrics*. North Holland.

Jullien, B. and B. Salanie (2000). Estimating preferences under risk: The case of racetrack bettors. *Journal of Political Economy 108(3)*, 503–530.

Kahneman, D. and A. Tversky (1973). On the psychology of prediction. *Psychological Review 80*, 237–251.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*, 263–291.

Kaniel, R., G. Saar, and S. Titman (2008). Individual investor trading and stock returns. *Journal of Finance 63(1)*, 273–310.

Karmarkar, U. S. (1978). Subjectively weighted utility: A descriptive extension of the expected utility model. *Organizational Behavior and Human Performance 21(1)*, 61–72.

Karmarkar, U. S. (1979). Subjectively weighted utility and the allais paradox. *Organizational Behavior and Human Performance 24(1)*, 67–72.

Kaustia, M. (2004). Market-wide impact of the disposition effect: Evidence form the ipo trading volume. *Journal of Financial Markets 7*, 207–235.

Kaustia, M. (2010). Prospect theory and the disposition effect. *Journal of Financial and Quantitative Analysis 47*, 263–291.

Keele, L. and D. Park (2006). Ambivalent about ambivalence: A re-examination of heteroskedastic probit models. *Working Paper, Ohio State University*.

Kent, J. T. (1982). Robust properties of likelihood ratio tests. *Biometrika 69*, 19–27.

Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi (1983). Optimization by simulated annealing. *Science 220*, 671–680.

Kliger, D. and O. Levy (2009). Theories of choice under risk: Insights from financial markets. *Journal of Economic Behavior and Organization 71*, 330–346.

Kocherlakota, N. R. (1996). The equity premium: Its still a puzzle. *Journal of Economic Literature 34(1)*, 42–71.

Kroll, Y., H. Levy, and A. Rapoport (1988). Experimental tests of the separation theorem and the capital asset pricing model. *American Economic Review 78(3)*, 500–519.

Kullback, S. (1968). *Information Theory and Statistics*. Dover Publications.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica 47*, 1315–1336.

Landskroner, Y. (1988). Risk aversion in securities markets. *Journal of Banking and Finance 6(1)*, 1247–1292.

Latane, H. A. (1959). Criteria for choices among risky ventures. *Journal of Political Economy 67(2)*, 144–155.

Lattimore, P. M., J. R. Baker, and A. D. Witte (1992). The influence of probability on risky choice. *Journal of Economic Behavior and Organization 17*, 377–400.

Laury, S. and C. A. Holt (2005). Further reflections on prospect theory. *Working Paper, Georgia State University*.

Lee, H., J. Park, L. J., and R. Wyer (2008). Disposition effects and underlying mechanism in e-trading of stocks. *Journal of Marketing Research 45*, 362–378.

Lehenkari, M. and J. Perttunen (2004). Holding on to losers: Finnish evidence. *Journal of Behavioral Finance 5(2)*, 116–126.

Lehmann, E. L. (1983). *Theory of Point Estimation*. Wiley.

Lesaffre, E. and B. D. Marx (1993). Collinearity in generalized linear regression. *Communications in Statistics - Theory and Methods 22(7)*, 1933–1952.

Levitt, S. and J. A. List (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives 21*, 153–174.

Levy, H. (1994). Absolute and relative risk aversion: an experimental study. *Journal of Risk and Uncertainty 8*, 289–307.

Levy, H. and T. Post (2005). Does risk seeking drive stock prices? a stochastic dominance analysis of aggregate investor preferences and beliefs. *Review of Financial Studies 18(3)*, 925–953.

Levy, M. and H. Levy (2002a). Arrow-pratt risk aversion, risk premium and decision weights. *Journal of Risk and Uncertainty 25(3)*, 265–290.

Levy, M. and H. Levy (2002b). Prospect theory: Much ado about nothing? *Management Science 48(10)*, 1334–1349.

Lim, S. S. (2006). Do investors integrate losses and segregate gains? mental accounting and investor trading decisions. *Journal of Business 79*, 2539–2573.

Linnainmaa, J. (2005). The individual day trader. *Working Paper, University of Chicago*.

Linnainmaa, J. T. (2010). Do limit orders alter inferences about investor performance and behavior? *Journal of Finance 65(4)*, 1473–1506.

Liu, Y. and H. Mahmassani (2000). Global maximum likelihood estimation procedures for multinomial probit (mnd) model parameters. *Transportation Research B 34*, 419–444.

Locke, P. R. and S. C. Mann (2000). Do professional traders exhibit loss realization aversion? *Working Paper, Washington University*.

Locke, P. R. and S. C. Mann (2005). Professional trader discipline and trade disposition. *Journal of Financial Economics 76*, 401–444.

Loomes, G., P. G. Moffatt, and R. Sugden (2002). A microeconomic test of alternative stochastic theories of risky choices. *Journal of Risk and Uncertainty 24*, 103–130.

Loomes, G. and R. Sugden (1995). Incorporating a stochastic element into decision theories. *European Economic Review 39*, 103–130.

Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis*. Dover Publication.

Machina, M. J. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature 27*, 1622–1668.

Mankiw, N. G. and S. Zeldes (1991). The consumption of stockholders and non-stockholders. *Journal of Financial Economics 29*, 97–112.

Mann, H. and D. Whitney (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics 18*, 50–60.

Mantel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports 50(3)*, 163–170.

Markowitz, H. M. (1952). The utility of wealth. *Journal of Political Economy 60*, 151–156.

Markowitz, H. M. (1976). Investment for the long run: New evidence for an old rule. *Journal of Finance 31(5)*, 1273–1286.

Marquardt, D. W. (1963). An algorithm for least squares estimation of nonlinear parameters. *Journal of the Society of Industrial Applied Mathematics 11*, 431–441.

Marschak, J. (1960). *Binary choice constraints on random utility indications, in K. Arrow, ed. Stanford Symposium on Mathematical Methods in the Social Sciences*. Stanford University Press.

McCord, M. and R. DeNeufville (1986). Lottery equivalents: Reduction of the certainty effect problem in utility assessment. *Management Science 32*, 56–60.

McCullough, B. D. and H. D. Vinod (2003). Verifying the solution from a nonlinear dolver: A case study. *American Economic Review 93(3)*, 873–892.

McFadden, D. (1974). *Conditional logit analysis of qualitative choice behavior, in:*.

McFadden, D. (1980). Econometric models for probabilistic choice among products. *Journal of Business 55(3)*, S13–S29.

McLeish, D. L. (1974). Dependent central limit theorems and invariance principles. *Annals of Probability 2(4)*, 620–628.

Mehra, R. (2008). *The Handbook of the Equity Risk Premium*. Elsevier.

Mehra, R. and E. Prescott (1985). The equity premium: A puzzle. *Journal of Monetary Economics 15*, 145–161.

Meng, J. (2010). The disposition effect and expectations as reference point. *Working Paper, Guanghua Peking University*.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics 51*, 247–257.

Moffatt, P. G. and S. A. Peters (2001). Testing for the presence of a tremble in economic experiments. *Experimental Economics 4*, 221–228.

Morin, R. A. and A. F. Suarez (1983). Risk aversion revisited. *Journal of Finance 106(438)*, 1175–1192.

Munk, C. (2013). *Financial asset pricing theory*. Oxford University Press.

Nielssen, T. D. and J. Jaffray (2004). Dynamic decision making without expected utility: An operational approach. *Working paper, Universite Paris 6*.

Normandin, M. and P. St-Amour (2008). An empirical analysis of aggregate household portfolios. *Journal of Banking and Finance 32*, 1583–1597.

Odean, T. (1998). Are investors reluctant to realize their losses? *Journal of Finance 53*, 1775–1798.

Odean, T. (1999). Do investors trade too much? *American Economic Review 89(5)*, 1279–1298.

Oehler, A., K. Heilmann, V. Laeger, and M. Oberlaender (2003). Coexistence of disposition investors and momentum traders in stock markets: Experimental evidence. *Journal of International Financial Markets, Institutions and Money 13*, 503–524.

Orme, C. (1995). On the use of artificial regressions in certain microeconometric models. *Econometric Theory 11(2)*, 290–305.

Ornstein, L. S. and G. E. Uhlenbeck (1930). On the theory of brownian motion. *Physical Review 36*, 823–841.

Pawitan, Y. (2001). *In all Likelihood-Statistical Modelling and Inference using likelihood*. Clarendon Press.

Post, T., M. J. van den Assem, G. Baltussen, and R. H. Thaler (2008). Deal or no deal? decision making under risk in a large-payoff game show. *American*

Economic Review 98(1), 38–71.

Poteshman, A. M. and V. Serbin (2003). Clearly irrational financial market behavior: Evidence from the early exercise of exchange traded stock options. *Journal of Finance 58(1)*, 37–70.

Preston, M. G. and P. Baratta (1948). An experimental study of the auction-value of an uncertain outcome. *American Journal of Psychology 61(2)*, 183–193.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization 3*, 323–343.

Quiggin, J. (1993). *Generalized Expected Utility Theory: The Rank Dependent Model*. Kluwer Academic Publishers.

Rabe-Hersketh, S. and B. Everitt (2004). *A Handbook of Statistical Analysis using Stata*. Chapman and Hall Press.

Rabin, M. (2000). Loss aversion and expected utility theory: A calibration theorem. *Econometrica 68*, 1281–1292.

Rabin, M. and R. H. Thaler (2001). Anomalies: Loss aversion. *Journal of Economic Perspectives 15*, 219–232.

Rao, C. R. (1945). Information and the accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society 37*, 81–89.

Rao, C. R. (1973). *Linear statistical inference and its applications, 2nd edition.* Wiley.

Rendleman, R. J. and B. J. Bartter (1979). Two-state option pricing. *Journal of Finance 34*, 1093–1110.

Rich, E. and K. Knight (1991). *Artificial Intelligence.* McGraw-Hill.

Roger, W. H. (1994). Regression standard errors in clustered samples. *Stata Technical Bulletin 13*, 19–23.

Rust, J. (1992). Do people behave according to bellman's principle of optimality? *Working paper, Stanford University*.

Rust, J. (1994). *Structural Estimation of Markov Decision Processes in R.F. Engle and D.L. McFadden (eds.)Handbook of Econometrics Volume IV*. Elsevir.

Ruud, P. (2000). *An Introduction to Classical Econometric Theory.* Oxford University Press.

Saha, A. (1993). Expo-power utility: A 'flexible' form of absolute and relative risk aversion. *American Journal of Agricultural Economics 75(4)*, 905–913.

Saha, A., R. C. Shumway, and H. Talpaz (1994). Joint estimation of risk preference structure and technology using expo-power utility. *American Journal of Agricultural Economics 76(5)*, 173–184.

Schaefer, R. L. (1986). Alternative estimators in logistic regression when data are collinear. *Journal of Statistical Computation and Simulation 25*, 75–91.

Schlarbaum, G., W. Lewellen, and R. Lease (1978). Realized returns on common stock investments: The experience of individual investors. *Journal of Business 51(2)*, 299–325.

Schlarbaum, G., W. Lewellen, R. Lease, and R. A. Cohn (1975). Individual investor risk aversion and investment portfolio composition. *Journal of Finance 30(2)*, 605–620.

Schoemaker, P. J. H. and H. C. Kunreuther (1979). An experimental study of insurance decisions. *Journal of Risk and Insurance 46(4)*, 603–618.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics 6*, 461–465.

Shapira, Z. and I. Venezia (2001). Patterns of behavior of professionally managed and independent investors. *Journal of Banking and Finance 25*, 1573–1587.

Shefrin, H. (2008). *A Behavioral Approach to Asset Pricing.* Elsevier Academic Press.

Shefrin, H. and M. Statman (1984). Explaining investors preference for cash dividends. *Journal of Financial Economics 13(2)*, 253–282.

Shefrin, H. and M. Statman (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance 40*, 777–790.

Siegel, S. (1956). *Nonparametric Statistics for Behavioral Science.* McGraw Hill.

Statman, M., S. Thorley, and K. Vorkink (2006). Investor overconfidence and trading volume. *Review of Financial Studies 19(4)*, 1531–1565.

Stott, H. P. (2006). Cumulative prospect theorys functional menagerie. *Journal of Risk and Uncertainty 32*, 101–130.

Sugiura, N. (1978). Further analysis of the data by akaike's information criterion and the finite correction. *Communications in Statistics - Theory and Methods 7(1)*, 13–26.

Thaler, R. H. (1985). Mental accounting and consumer choice. *Marketing Science 4(3)*, 199–214.

Thaler, R. H., A. Tversky, D. Kahneman, and A. Schwartz (1997). The effect of myopia and loss aversion on risk taking: An experimental test. *Quarterly Journal of Economics 112(2)*, 647–661.

Thurstone, L. L. (1927). A law of comparative judgement. *Psychological Review 34*, 273–286.

Train, K. E. (1986). *Qualitative Choice Analysis.* The MIT Press.

Train, K. E. (2009). *Discrete Choice Models with Simulation, 2nd. ed.* Cambridge University Press.

Tversky, A. and D. Kahneman (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics 106(4)*, 1039–1061.

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

Vlcek, M. and T. Hens (2011). Does prospect theory explain the disposition effect? *Journal of Behavioral Finance 12(3)*, 141–157.

von Neumann, J. and O. Morgenstern (1947). *Theory of Games and Economic Behavior.* Princeton University Press.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica 57(2)*, 307–333.

Wakker, P. P. (1994). Separating marginal utility and probabilistic risk aversion. *Theory and Decision 36*, 1–44.

Wakker, P. P. (2003). The data of levy and levy (2002) "prospect theory: Much ado about nothing?" actually support prospect theory. *Management Science 49(7)*, 979–981.

Wakker, P. P. (2008). Explaining the characteristics of the power (crra) utility function. *Health Economics 17*, 1329–1344.

Wakker, P. P. (2010). *Prospect Theory: For Risk and Ambiguity.* Cambridge University Press.

Wakker, P. P. and D. Deneffe (1996). Eliciting von neumann-morgenstern utilities when probabilities are distorted or unknown. *Management Science 42(8)*, 1131–1150.

Wakker, P. P. and A. Tversky (1993). An axiomatization of cumulative prospect theory. *Journal of Risk and Uncertainty 7(2)*, 147–176.

Wang, M. (2006). Prospect theory in behavioral finance. *Working Paper*.

Weber, M. and C. F. Camerer (1998). The disposition effect in securities trading: An experimental analysis. *Journal of Economic Behavior and Organization 33*, 167–184.

Wilcox, N. T. (2008). Stochastic models for binary discrete choice under risk, in:. *Research in Experimental Economics 12*, 197–292.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin 1*, 80–83.

Wolf, C. and L. Pohlman (1983). Notes and comments: The recovery of risk preferences from actual choices. *Econometrica 51(3)*, 843–850.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data, 2nd edition.* Cambridge: MIT-Press.

Wu, G. and R. Gonzalez (1996). Curvature of the probability weighting function. *Management Science 42*, 1676–1690.

Wu, G. and R. Gonzalez (1999). Nonlinear decision weights in choice under uncertainty. *Management Science 45*, 74–85.

Yaari, M. E. (1965). Convexity in the theory of choices under risk. *Quarterly Journal of Economics 79*, 278–290.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica 55(1)*, 95–115.

Zhu, M. (2002). The local bias of individual investors. *Working Paper, Yale University*.

Zuchel, H. (2001). What drives the disposition effect? *Working Paper, Mannheim University*.

Recent Issues

No. 147	Andreas Hackethal, Sven-Thorsten Jakusch, Steffen Meyer	**Taring all Investors with the same Brush? Evidence for Heterogeneity in Individual Preferences from a Maximum Likelihood Approach**
No. 146	Sven-Thorsten Jakusch, Steffen Meyer, Andreas Hackethal	**Taming Models of Prospect Theory in the Wild? Estimation of Vlcek and Hens (2011)**
No. 145	Christian Geppert, Alexander Ludwig, Raphael Abiry	**Secular Stagnation? Growth, Asset Returns and Welfare in the Next Decades: First Results**
No. 144	Mario Bellia, Loriana Pelizzon, Marti G. Subrahmanyam, Jun Uno, Darya Yuferova	**Low-Latency Trading and Price Discovery: Evidence from the Tokyo Stock Exchange in the Pre-Opening and Opening Periods**
No. 143	Peter Gomber, Satchit Sagade, Erik Theissen, Moritz Christian Weber, Christian Westheide	**Spoilt for Choice: Order Routing Decisions in Fragmented Equity Markets**
No. 142	Nathanael Vellekoop	**The Impact of Long-Run Macroeconomic Experiences on Personality**
No. 141	Brigitte Haar	**Freedom of Contract and Financial Stability through the lens of the Legal Theory of Finance**
No. 140	Reint Gropp, Rasa Karapandza, Julian Opferkuch	**The Forward-Looking Disclosures of Corporate Managers: Theory and Evidence**
No. 139	Holger Kraft, Claus Munk, Farina Weiss	**Predictors and Portfolios Over the Life Cycle**
No. 138	Mohammed Aldegwy, Matthias Thiemann	**How Economics Got it Wrong: Formalism, Equilibrium Modelling and Pseudo-Optimization in Banking Regulatory Studies**
No. 137	Elia Berdin, Cosimo Pancaro, Christoffer Kok	**A Stochastic Forward-Looking Model to Assess the Profitability and Solvency of European Insurers**
No. 136	Matthias Thiemann, Mohammed Aldegwy, Edin Ibrocevic	**Understanding the Shift from Micro to Macro-Prudential Thinking: A Discursive Network Analysis**
No. 135	Douglas Cumming, Jochen Werth, Yelin Zhang	**Governance in Entrepreneurial Ecosystems: Venture Capitalists vs. Technology Parks**