Essays On Statistical Economics With Applications To Financial Market Instability, Limit Distribution Of Loss Aversion, And Harmonic Probability Weighting Functions

G. Charles-Cadogan

Thesis Presented for the Degree of
Doctor of Philosophy
in the School of Economics
University of Cape Town

Supervisor: Dr. Sure Mataramvura

Examiners:
Prof. Ganna Pogrebna, University of Warwick
Prof. Pablo Brañas Garza, Middlesex University–London
Prof. John Doe (Anonymous)

December 9, 2015

Abstract

This dissertation is comprised of four essays. It develops statistical models of decision making in the presence of risk with applications to economics and finance. The methodology draws upon economics, finance, psychology, mathematics and statistics. Each essay contributes to the literature by either introducing new theories and empirical predictions or extending old ones with novel approaches.

The first essay (Chapter II) includes, to the best of our knowledge, the first known limit distribution of the myopic loss aversion (MLA) index derived from micro-foundations of behavioural economics. That discovery predicts several new results. We prove that the MLA index is in the class of α-stable distributions. This striking prediction is upheld empirically with data from a published meta-study on loss aversion; published data on cross-country loss aversion indexes; and macroeconomic loss aversion index data for US and South Africa. The latter results provide contrast to Hofstede's cross-cultural uncertainty avoidance index for risk perception. We apply the theory to information based asset pricing and show how the MLA index mimics information flows in credit risk models. We embed the MLA index in the pricing kernel of a behavioural consumption based capital asset pricing model (B-CCAPM) and resolve the equity premium puzzle. Our theory predicts: (1) stochastic dominance of good states in the B-CCAPM Markov matrix induce excess volatility; and (2) a countercyclical fourfold pattern of risk attitudes.

The second essay (Chapter III) introduces a probability model of "irrational exuberance" and financial market instability implied by index option prices. It is based on a behavioural empirical local Lyapunov exponent (BELLE) process we construct from micro-foundations of behavioural finance. It characterizes stochastic stability of financial markets, with risk attitude factors, in fixed point neighbourhoods of the probability weighting functions implied by index option prices. It provides a robust early warning system for market crash across different credit risk sources. We show how the model would have predicted the Great Recession of 2008. The BELLE process characterizes Minskys financial instability hypothesis that financial markets transit from financial relations that make them stable to those that make them unstable.

The third essay (Chapter IV) introduces an outcome dependent harmonic probability weighting function (HPWF) based on an information theory of stochastic choice. We use the HPWF to resolve the preference reversal (PR) puzzle–which is observed in economics and psychology experiments when a decision makers (DMs) preferences over the same items change depending upon how she is subsequently asked to construct a preference. We use the principle of maximum entropy to synthesize information processing, probabilistic choice, and momentary fluctuation hypotheses proposed by various researchers to explain intransitivity implied by PR phenomenon. The HPWF theory is illustrated via simulation. Additionally, we show how the HPWF decomposes regret theory, and rank dependent utility (RDU), into core expected utility theory (EUT) plus functionally equivalent stochastic error addends. This theoretical prediction finds support in Hey and Orme (1994) seminal experiments on the difference between generalized EUT and core EUT models. We also prove that experimenter interference with the probability cycle of DMs HPWF causes them to observe preference reversal in stochastic choice experiments even though the true state is transitive and there is no violation of procedure invariance.

The fourth essay (Chapter V) shows that Bernoulli (1738) original utility function is alive and well. For example, several papers reexamine Bernoulli (1738) expected utility resolution of the St Petersburg Paradox in the context of cumulative prospect theory (CPT). We go a step further. We reexamine the geometry of Bernoulli's original sketch of his utility function. We prove that contrary to received literature, which alleges that Bernoulli's utility function is unable to generate a loss aversion index (ULA), the geometry of Bernoulli's original sketch accommodates a ULA index with smooth reference dependent utility functions. In fact, it provides a solution to the open problem of closed form global ULA index formula in prospect theory. Like in the first essay, the ULA index predicted by Bernoulli's utility function is α-stable. Under fairly mild assumptions, we show how it supports a Fisher z-transform statistical test for the loss aversion index and we show how the test can be applied.

Dedication

This dissertation is dedicated to my deceased parents from whom I received my work ethic and life philosophies.

Declaration

I declare that the instant thesis entitled “ Essays On Statistical Economics With Applications To Financial Market Instability, Limit Distribution Of Loss Aversion, And Harmonic Probability Weighting Functions” is my own work, that it has not been submitted for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged by complete references.

Acknowledgment

This dissertation was made possible by the contributions of several people. First and foremost I am eternally grateful to my principal adviser Prof. Sure Mataramvura who went beyond the call of duty to make the production of this document possible. His wise counsel, comments, suggestions, and careful attention to mathematical details greatly improved the manuscript. I thank the examiners for helpful comments and suggestions which improved the manuscript and stimulated ideas for future research. I thank Prof. Haim Abraham and Prof. Amos Peters for many useful discussions at various stages in the process. Prof. Toni Leiman has been an ambassador at large and braaimeister extraordinaire whose knowledge of the history of economic thought improved my understanding of the literature. Prof. Ojelanki Ngwenyama was instrumental in facilitating this dissertation from start to finish. I am eternally grateful for his prescience and encouragement throughout. Prof. Don Ross and Prof. Glenn W. Harrison were catalysts in the early stages of this dissertation. I thank Prof. Ingrid Woolard and Prof. Edwin Muchapondwa for facilitating the administrative process of this dissertation. I acknowledge funding received from Research Unit in Behavioural Economics and Neuroeconomics (RUBEN), National Research Foundation, and Institute for Innovation and Technology Management (IITM).

Many researchers have provided feedback throughout the development of the Chapters in this dissertation. They are acknowledged elsewhere in this document. Of course, any errors which may remain are my own.

I thank my family Audrey, Delisia (and Dre Dre), and Sam for their unwavering support throughout. Omarion, Beyonce, Alexisa and Jamela you guys are special. I have a lot of friends–too numerous to mention–to thank for encouraging and supporting me in various ways throughout my professional endeavours. However, Doreen Thompson, Oliver Martin, Shelley Walker, John A. Cole, Debby Mitchell, Don Mullings, and Mark Zanecki are more like family.

How can I forget the “Drop Squad”–the group of friends I made in the PhD program and who provided a sounding board for ideas, wise counsel. Elizabeth Nanziri Lwanga, Ramaele Moeshoshoe, Threza Mtenga, David Kaoya, Muna Shifa, Mahawiya Suleman, Grieva Chelwa, Marmelo Nchake, and Dambala Gelo I thank you all.

Statement On Published Material

Chapter II is based in part on Cadogan (2013), "Myopic Loss Aversion and Intolerance to Decline in Consumption" which appear in the *Proceedings of Foundations of Utility and Risk XVI 2014*; and "Asymptotic Theory Of Myopic Loss Aversion: Applications To Intolerance for Decline in Standard Of Living and Asset Pricing" in *Proceedings of Midwest Econometrics Group 2015, Federal Reserve Bank, St. Louis*. I thank Mark Machina, Peter Wakker, Mario Siniscalchi, Haim Abraham, Martine Visser, Dambala Gelo, Brian Munroe, J. Atsu Amegashie, Gregory N. Price, Hersh Shefrin, seminar participants at UCT School of Ecoomics, and participants in FUR XIV 2014 session B20 "Risk attitude - Loss Aversion: Theory" for their comments on that version of the paper.

Chapter III is an expansion with correction of Cadogan, G. (2014). "Noisy Chaos in a Large System of Decision Makers with Heterogeneous Beliefs with Application to Index Option Prices," *Systems Research and Behavioral Science*, Special Issue on Behavioral Risk 31(4), 487-501. I thank the editor Damon Dash Wu (editor) and two anonymous referees, Oliver Martin, Diane Wilcox, G. Di Nunno, John Mwamba (discussant), Co-Pierre Georg, Salma Kagee, Alex Zimper, Haim Abraham, Nico Katzke (discussant), and participants of the Southern Africa Mathematical Sciences Association (SAMSA) 2013 International Conference, Economic Research Southern Africa (ERSA) Financial Economics Workshop 2013, and ERSA Economic Theory Workshop 2014 for comments on different versions of this paper. This paper was accepted for presentation at 38th Conference on Stochastic Processes and their Applications 2015, Oxford Mann Institute of Quantitative Finance; and 5th International Conference of the Financial Engineering 2015, Audencia Nantes School of Management Campus.

Chapter IV is an extension of an article under revise and resubmit status at a peer reviewed journal. I thank the editor and two anonymous referees for their comments. This paper benefited from discussions with Glenn W. Harrison, Don Ross, Amos Peters, detailed comments from Peter Wakker, and comments from Vyacheslav Yukalov, Nat Wilcox, and seminar participants at the UCT School of Economics. I thank Paul Slovic, Michael J. Stutzer and Jerome R. Busemeyer for sharing their knowledge of the literature and directing me to references which greatly improved the presentation within. Any errors which remain are my own.

Chapter V is an extension of Cadogan, G, (2014), "Prospect Theory's Cognitive Error about Bernoulli's Utility Function," *AER Bulletin Peer Review Working Paper Series* available at https://drive.google.com/file/d/0B95uP0Flg3KJQkFOZkprYVZhRWM/view?usp=sharing. I thank Mark Taylor (editor) , three anonymous referees, Haim Abraham, Ramaele Moeshoeshoe, and Mario Siniscalchi for their comments, which improved the paper.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Great Recession of 2008 highlighted a failure of neoclassical economic models, predicated on rational economic theories, to anticipate the consequences of seemingly "irrational exuberance" associated with risk attitudes that precipitated the crisis. This led to a clarion call for new modelling paradigms and creation of organizations such as the Institute for New Economic Thinking (INET), and market monitoring research units in major central banks. For example, the International Monetary Fund (IMF) publishes the biannual Global Financial Stability Report which "provides an assessment of the global financial system and markets". The Financial Policy Committee, Bank of England; Financial Stability Committee, Federal Reserve Board, and Center for Financial Stability in the US are some of the newly formed entities tasked with monitoring financial stability.

One approach to new economic thinking that has traction, and the one taken in this dissertation, is the synthesis of neoclassical economics paradigms with behavioural assumptions (Shefrin, 2014, 2015a). At its core this dissertation develops statistical models of decision making in the presence of risk, and rigorous theoretical models that extend to several applications. Its highly interdisciplinary nature draws upon the fields of economics, finance, psychology, mathematics and statistics. It applies probability theory to microfoundations of behavioural economics and finance in an ambient topological space, and it derives consequential structural models. This approach contributes to the literature by extending the solution space of economics and finance with several new results. Many of which are important in their own right as indicated in the chapters that follow.

In Chapter II, we introduce an asymptotic theory of and statistical distribution for the myopic loss aversion (MLA) index. To the best of our knowledge this theory is new to the literature,[1]

[1]We note that Stott (2006) conducted a metastudy on functional forms of CPT value functions and probability weighting functions. Zeisberger et al. (2012) and references therein examined the stability of CPT parameter estimates over time. And Baucells and Heukamp (2012) studied time varying probability weighting functions. Booij et al. (2010) study presents heterogeneity and large disparities in loss aversion index estimates. However, none of those studies represented the MLA index as a random variable drawn from a statistical distribution. Our novelty claim rests on representation of the loss aversion index as a random variable

and it leads to a groundswell of new results. The loss aversion concept in economics was introduced in Kahneman and Tversky (1979). Whereas the loss aversion index was first introduced to behavioural economics by Tversky and Kahneman (1991) as a deterministic parameter under riskless choice, subsequently presented in the context of risky choice in Tversky and Kahneman (1992), and axiomatized by Wakker and Tversky (1993). The MLA concept was popularized by Benartzi and Thaler (1995) in their simulated solution of the equity premium puzzle.

First, we prove that the loss aversion index is an independent and identically distributed Cauchy random variable. This striking result implies that any positive value of the loss aversion index is admissible. So the oft reported statement "loss aversion index median value around 2.0" (e.g., Bowman et al. (1999); De Neve et al. (2015)) may be superfluous. We tested the theory by fitting the distribution to four different data sets of loss aversion indexes estimated in different contexts: a metastudy, loss aversion indexes estimated from hypothetical choices in survey based experiments conducted in 45 countries, and our own cross-country macroeconomic loss aversion index estimates for US and South Africa income and consumption data. The distribution theory was upheld in each case.

Second, we introduce a simple empirical strategy for computing a macroeconomic loss aversion index by a method of moments estimator which asymptotic distribution is in the Cauchy class of α-stable distributions. This permits cross-country risk attitude comparison. It provides a "hard data" alternative to Hofstede (1980, 1983) survey driven cross-cultural anticipative uncertainty avoidance index (UAI) of risk perception, popularized in the international business and cross-cultural psychology literatures. For example, whereas Hofstede's UAI posits that attitudes towards uncertainty in the US and South Africa are similar, our macroeconomic loss aversion index estimator shows that they are dissimilar: South Africans tend to be gain seeking relative to their US counterparts. South Africa's risk attitude is sensitive to political uncertainty, whereas US risk attitude is sensitive to natural disasters and financial market instability.

Third, we show how the Cauchy random variable prediction for the MLA index performs in asset pricing contexts. Specifically, we show that the MLA index mimics information flows in a Cauchy bridge process in the credit risk model employed by Hoyle (2010); Hoyle et al. (2011);

drawn from a specific family of statistical distributions.

Ikpe et al. (2014). Furthermore, we show that Mehra and Prescott (1985) neoclassical consumption based asset pricing model (CCAPM) is unable to explain the equity premium because risk attitudes in their model are misspecified. Their model depends on risk aversion and volatility of consumption growth to explain the equity premium puzzle. Köbberling and Wakker (2005, p. 128) presented stylized arguments which show how risk aversion is tied to loss aversion by and through a utility function they devise. Charles-Cadogan (2016b, Appendix A.2) provides an explicit relationship between the loss aversion index and Arrow-Pratt risk aversion indexes over gain and loss domains. We extend that result by deriving a cross-sectional specification for risk aversion as a function of loss aversion. The Cauchy random variable MLA index in that specification is a missing variable in the Mehra-Prescott model. It accommodates the impact of rare disasters on asset prices. And it admits large countercyclical swings in the equity premium that neoclassical models do not pick up. Furthermore, our behavioural asset pricing approach also predicts that investor response to stochastic dominance of preferred states induce excess volatility in equity prices.

Chapter III examines behavioural dynamics in fixed point neighbourhoods of probability weighting functions (pwfs) with applications to financial markets. Even though Prelec (1998) and subsequently Luce (2001) axiomatized a fixed point probability for pwfs, their models focused on the stable (concave-convex) shape. We show that pwfs come in stable and unstable (convex-concave) vintages that have the same fixed point probability. This implies that fixed point dynamics inform pwfs shapes and phase transitions in pwfs.

First, we derive a behavioural empirical local Lyapunov exponent (BELLE) process from neuronal noise in fixed point neighbourhoods. This introduces a behavioural stochastic differential equation that is important in its own right. We examine stochastic stability criteria for the BELLE process and verify how they control pwfs shapes in a simple Monte Carlo experiment.

Second, the theory is upheld *a fortiori* by the pwfs implied by index option prices in Polkovnichenko and Zhao (2013). We find that the shapes of pwfs reflect probabilistic risk attitudes towards credit risk sources. This result lends credence to the source function theory of pwfs in the Abdellaoui et al. (2011) study. We calibrate our BELLE theory with parameter estimates for the pwfs implied by index option prices in Polkovnichenko and Zhao (2013)). We show how that exercise supports an early warning system for financial market instability. In fact, we show

how our model would have anticipated the Great Recession of 2008. The shape of the pwf implied by index option prices, when the market crashed in 2008, mimics the shapes of the pwf predicted by our Monte Carlo experiment on market crash. Thus, we prove that the pwf source functions predicted by the BELLE theory are sufficient statistics for Hyman Minskys financial instability hypothesis: An economy has stable and unstable regimes, and it transits from financial relations that make it stable to those that make it unstable.

Chapter IV employs the principle of maximum entropy to derive a coherent harmonic probability weighting function (HPWF) that extends Hogarth and Einhorn (1990) descriptive venture theory probability weighting model. The HPWF resolves the preference reversal (PR) phenomenon first observed in experiments conducted by psychologists Lichtenstein and Slovic (1971) and Lindman (1971). The phenomenon is observed when the choice decision makers (DMs) make, between two alternatives, are inconsistent with the price they are willing to accept (or pay) for the given choice when it is presented later as part of the same two alternatives. Assuming preferences are invariant to elicitation method, e.g., procedure invariance, PR presents a dilemma for decision theory because it implies intransitive preferences–even though axiomatic decision theories in economics and psychology often feature transitive preferences as a building block. Preference reversal implies cyclic behaviour that leaves decision makers vulnerable to money pumps. It has policy implications for cost-benefit analysis and environment studies where the phenomenon has been reported. For example, imputed values for non-tradeable items may be elicited from surveys and policymakers may use the responses to inform resource allocation. If the preference ranking extrapolated from survey responses do not reflect the choices that respondents would actually make, then the survey elicited valuation is unreliable and policy is misinformed.

Seidl (2002) identified four determinants of PR in his literature review: (1) mode of elicitation of certainty equivalents, (2) intransitivity of preferences, (3) overpricing and/or underpricing of lotteries, and (4) nonlinear probabilities. We introduce an outcome dependent HPWF which resolves the PR puzzle in the context of (2) and (4). We impose procedure invariance and transitivity of choice on the model, and show how observers can still be misled into reporting PR when there is none. According to our theory, the HPWF is controlled by probability cycles that are disturbed or broken by observers who interpret the observation as PR. This is a manifestation of the uncertainty principle or observer effect explained at length in Von Neumann (1955, Ch. VI). However, the phenomenon vanishes when probability cycles are completed. In the neuroeconomics and psychophysics literature, Takahashi (2006) proposed a stylized outcome dependent probability weighting function based on the difference between perceived probability and Claude Shannon's entropy measure of probability uncertainty. Several other authors have proposed outcome dependent probability weighting functions, e.g., Pfanzagl (1967); Fellner (1961); Schneeweiss (1974); Hogarth and Einhorn (1990); Dillenberger et al. (2013) to name a few. To the best of our knowledge our model is the first to employ maximum entropy analysis and identify behavioural harmonics in outcome dependent probability weighting functions. Our model addresses the open issue of seemingly intransitive preferences identified by Seidl (2002, p. 637). We reiterate here that our model proves that PR can be reported by an observer even when procedure invariance is not violated, and true preference is transitive, if the probability weighting function is harmonic. Specifically, we identify an observer effect for PR. For application, we show how the HPWF decomposes regret theory and rank dependent utility (RDU) into core expected utility theory (EUT) plus functionally equivalent stochastic error addends. Thus, the HPWF provides a theoretical explanation for why Hey and Orme (1994) found no statistically significant difference between core EUT and generalized EUT models in their seminal economic experiments.

Chapter V re-examines Bernoulli (1738) original utility function, in the context of cumulative prospect theory (CPT), to evaluate the claim made by some proponents of prospect theory that Bernoulli's utility function cannot accommodate a loss aversion index. This exercise is reminiscent of Blavatskyy (2005); Rieger and Wang (2006); Pfiffelmann (2011) who reexamined Bernoulli (1738) "resolution" of the St Petersburg Paradox in the context of CPT. We go one step further by examining CPT loss aversion index in the context of Bernoulli (1738) sketch of his utility function. Cursory inspection of the geometry of Bernoulli's original utility function shows that it accommodates relative wealth or a reference point where it cuts the horizontal axis. So that points to the left of the reference point support "disutility of losses" and points to the right support "utility of gains". We show how a loss aversion index can be constructed from that observation. Furthermore, we show how one can develop a statistical test for the loss aversion index derived from that process. Thus, we refute claims alleging that Bernoulli's utility function is unable to accommodate a loss aversion index. It should be noted in passing that Charles-Cadogan (2016b) introduced an EUT based model that also generates a loss aversion index.

In conclusion, the instant dissertation provides several new results that impact decision theory, behavioural finance, and behavioural economics going forward. Perhaps the most promising areas for future research, motivated by the instant dissertation, lie in applications of (1) the family of statistical distributions supported by MLA indexes, (2) the BELLE process identified in fixed point neighbourhoods generated by probability weighting functions, and (3) the HPWF. For example, there is a nascent asset pricing literature in which loss aversion is synthesized with exotic preferences to (a) resolve the equity premium puzzle, and (b) account for excess volatility. The results in this thesis imply that the α-stable feature of the MLA index makes the latter a sufficient statistic that (1) resolves the equity premium puzzle, (2) explains excess volatility, and (3) explains rare disaster effects on asset prices in the context of standard preferences. Future research also includes axiomatization of financial structures identified in Minsky's financial instability hypothesis in the context of the BELLE process. The path properties of the BELLE process are also important in their own right. For example, one can analyze the escape time of the BELLE path from a set, and extend the BELLE process solution space with different background driving stochastic processes. Another avenue for further research contemplates a matrix operator theory of loss aversion motivated by the empirical gain-loss topology we introduce. That theory would tie a time varying Arrow-Pratt risk aversion index to a CRRA index and the loss aversion index generated by our gain-loss topological basis set. Finally, the HPWF decomposition of regret theory and RDEU into core EUT plus functionally equivalent addends, introduced in this thesis, shows promise as a methodology that can be axiomatized and applied to decompose other generalized EUT specifications into a core EUT theory plus stochastic error.

Chapter 2

Asymptotic Theory Of Myopic Loss Aversion with Applications To Asset Pricing

2.1 Introduction

This chapter contributes to the literature by introducing an asymptotic theory of the myopic loss aversion (MLA) index which show that it is drawn from a specific family of α-stable distributions. Thus, the MLA index is not constant or deterministic as is often portrayed in extant literature. One implication of our result is that several papers that assume constant or deterministic MLA index can now be reexamined to see if the results they reported are robust to α-stable MLA.

The loss aversion concept, first introduced in Kahneman and Tversky (1979) original version of prospect theory (OPT), is a pillar of behavioural economics and finance (Shefrin, 2009; Hirshleifer, 2015). It reflects investor psychology and risk perception (Hirschleifer, 2001). It posits that losses loom larger than gains when a decision maker (DM) chooses in a mixed lottery or gamble, i.e., one that is comprised of gains and losses. That is, a DM is more sensitive to losses than she is to gains of the same absolute size. That risk attitude concept was subsequently refined to include MLA which involves a decision maker's (DM's) response to and evaluation of losses incurred over short periods, together with a mental accounting process (Benartzi and Thaler, 1995).[1] Moreover, the MLA phenomenon has been affirmed in economic experiments, e.g., Gneezy and Potters (1997); Thaler et al. (1997); Benartzi and Thaler (1999); Haigh and List (2005).

Gneezy and Potters (1997) illustrated the MLA concept with the following example motivated by (Samuelson, 1963). Suppose a decision maker (DM) is faced with a simple gamble L where she stands to win \$200 with probability 1/2 or loose -\$100 with probability 1/2, i.e., $L = (200, 1/2; -100, 1/2)$. Suppose further that the DM is characterized by loss aversion and has a utility function $u(z) = z$ for $z \geq 0$ and $u(z) = 2.5z$ for $z < 0$, where z is the change in

[1]The mental accounting concept introduced by Thaler (1985, 1990) deals with the frequency of transactions evaluation, and how they are aggregated or segregated. However, it is not the subject matter of this paper.

wealth due to the gamble. Then, the *expected utility* of a single play of the gamble is negative: $1/2(200) + 1/2(-250) < 0$. Hence, the DM will reject the single play gamble, and also two single play gambles-if each is evaluated separately. However, the DM would accept two gambles $L \times L$ if (s)he evaluates them in combination: $1/4(400) + 1/2(100) + 1/4(-500) > 0$.[2] Thus, rejecting a single gamble while accepting two gambles is explained by the combined hypotheses of individuals being more sensitive to losses than to gains and evaluating the outcomes of the sequence of gambles in combination.

Benartzi and Thaler (1995, p. 75) used the MLA paradigm to resolve the equity premium puzzle. The latter is based on the notion that the historic equity premium in financial markets is too large to be explained by predictions of neoclassical models. Myopic investors concentrate on short-term price movements. They frame financial decision making such that loss aversion is the predominant risk attitude that motivate them to overweight short term price movements. Consequently, they tend to be more conservative investors who demand higher equity premia to hold stocks.

Thus far, the literature is silent on the statistical properties of the MLA index estimator derived from foundations of behavioural economics. This paper fills that gap in the literature and provides several independently important applications impacted by our novel derivation of the statistical distribution of the MLA index.

The modern foundations of behavioural economics specifications were established in Tversky and Kahneman (1992) who amended OPT with cumulative prospect theory (CPT) to address, inter alia, OPT's violation of stochastic dominance. "The key elements of [OPT] are 1) a value function that is concave for gains, convex for losses, and steeper for losses than for gains, and 2) a nonlinear transformation of the probability scale, which overweighs small probabilities and underweighs moderate and high probabilities," (Tversky and Kahneman, 1992, pp. 297-298). "In prospect theory, the carriers of value (meaning arguments of the utility function) are gains and

[2] Gneezy and Potters (1997) did not state which "mental accounting" rules were used. However, it appears that two wins (200 + 200) were assigned compound probability $1/2 \times 1/2 = 1/4$; two losses $(-2.5 \times 100 + -2.5 \times 100 = -500)$ assigned compound probability $1/2 \times 1/2 = 1/4$; and "a win and a loss" $(200 + -100)$ assigned compound probability $1/2 \times 1/2 = 1/4$ or "a loss and a win" $(-100 + 200)$ assigned compound probability $1/2 \times 1/2 = 1/4$ were aggregated to get $(200 + -100)$ assigned the probability 1/4+1/4=1/2. Evidently, the "loss factor" 2.5 was not consistently applied to the -100 loss in the aggregation in the "mental accounting".

losses relative to a pre-specified reference point." (Shefrin, 2009, p. 17). Key CPT amendments include (i) introduction of a loss aversion index, and (ii) the incorporation of rank dependent utility (RDU) (Quiggin, 1982) and the nonadditive feature of Choquet expected utility (CEU) theory (Schmeilder, 1989).[3]

In the psychology literature, a set of papers by Birnbaum and Navarrete (1998); Birnbaum et al. (1999); Birnbaum (2004, 2005, 2008) introduce competing models such as transfer attention exchange (TAX), among others, and produce evidence that falsify aspects of OPT and CPT. Wu (1994) conducted a study which revealed violation of ordinal independence, i.e., the notion that substitution of a common right tail should not affect decision makers rank ordered preferences (an important component of the "rank and sign dependent" (Luce, 2000) features of CPT).Wakker et al. (1994) introduced a study which showed that comonotonic independence was also violated. Studies by Wu and Markle (2008); Por and Budescu (2013) implicate the gain-loss separability theory of CPT which is crucial for construction of the loss aversion index in CPT (Tversky and Kahneman, 1992). However, no single model of decision making in the presence of risk is a panacea and CPT is no exception. Despite its shortcomings, a recent review paper by Barberis (2013, p. 173) asserts that "prospect theory is still widely viewed as the best available description of how people evaluate risk in experimental settings". Thus, our analysis in the sequel is motivated by CPT.

More important for the subject matter of this paper, Tversky and Kahneman (1992) introduced, specified, and estimated a utility based loss aversion index with data from controlled experiments.[4] They reported a median value of 2.25 for the loss aversion index estimated in their study. However, they were silent on the characteristics of the underlying statistical distribution that led to the median value 2.25. It is known that the median is a consistent estimator of central tendency for double exponential (Laplace) and Cauchy distributions. The mean and variance of such distributions may not exist because they are susceptible to extreme values and they do not adhere to the central limit theorem (Johnson et al., 1994; Kleiber and Kotz, 2003). Somewhat

[3] CEU is based on the non-additive "capacity measure" of a set introduced by Choquet (1954) as an alternative to probability measure of a set.

[4] Schmidt and Zank (2008) introduced the concept of probabilistic loss aversion based on curvature properties of the probability weighting function over gain domain, and the probability weighting function over loss domain. Quiggin (1993, p. 83) introduced a probabilistic risk aversion measure he called "probability weighting coefficient".

surprisingly, the literature on decision theory is also silent on the statistical distribution of the loss aversion index.[5]

This paper's contribution to the literature lies in its introduction of an asymptotic theory of the MLA index based on microfoundations of behavioural economics and finance. In particular, microfoundations of the loss aversion index were axiomatized in Wakker and Tversky (1993). We exploit the Euclidean topology (Dugundji, 1966, p. 63) induced by reference point(s) popularized by Kahneman and Tversky (1979); Tversky and Kahneman (1992); Kőszegi and Rabin (2006), and we identify a simple estimator for the MLA index from generating sets. Whereupon we derive statistical properties of the estimator. We prove that the statistical distribution of the MLA index is α-stable.[6] In particular, it admits a generalized Cauchy distribution characterized by extreme values. Thus, the narrow range of values for the loss aversion index induced by controlled laboratory experiments is misleading.

The Euclidean topology supports a new proof of the symbiotic relationship between utility functions and probability weighting functions (pwfs), as carriers of risk attitudes like loss aversion and risk aversion.[7] The proof is based on the utility transform of mollifiers[8] in a neighbourhood base of a reference point topology for ranked outcomes. This induces a skewed mixture distribution which characterizes low ranked and high ranked outcomes that support inverted S-shaped pwfs for risk aversion, and skewed S-shaped for risk seeking. This helps us to identify preference based foundations of observed phenomena like stocks as lotteries, long shot bias, and preference for skewness, and to interpret empirical results in this paper.

We tested the MLA index theory in a tournament by fitting and ranking a battery of dis-

[5] Lopes (1981) makes the case for the median value as a measure of central tendency in evaluating a gamble outside the context of loss aversion. But she did not specify a statistical distribution function that drives the estimator. In the econophysics literature Jensen et al. (2003) fitted statistical distributions for the first passage of returns over a given threshold for the Dow Jones Industrial Average that revealed "gain-loss asymmetry". However, they did not develop a micro-foundation theory that predicted their choice of fitted distribution as we do here.

[6] Stable distributions are characterized by an index of stability α wherein $0 < \alpha \leq 2$. They are "stable" because they retain their basic shape after certain transformations. For example, the well known normal distribution has an index $\alpha = 2$. The distribution of a normalized sum of normal random variables is also normal. So it retains the shape of a normal distribution. "When $\alpha < 2$, the tails of the distributions decay like a power function. This means that a stable random variable exhibits much more variability than a Gaussian one: it is much more likely to take values far away from the median. Stable distributions have been used to model such diverse phenomena as gravitational fields of stars, temperature distributions in nuclear reactors, stresses in crystalline lattices, stock market prices and annual rainfall." Samoradnitsky and Taqqu (1994, pp. 1-2). This paper adds myopic loss aversion to the pantheon of applications of stable distributions.

[7] The symbiotic relationship between utility and probability functions is known since at least Pfanzagl (1967) who axiomatized the relationship. See also Schneeweiss (1974) who provides a review of this issue.

[8] These are also known as bump functions or test functions. See e.g., Siddiqui (2004, §5.4) for further details on test functions.

tribution functions to MLA index estimates from the loss aversion index data in the Fishburn and Kochenberger (1979) metastudy. We tested it on the distribution of loss aversion indexes around the world reported in the Rieger et al. (2011) study. We also tested the theory with income and consumption time series data for the US and South Africa. In each case the theory is upheld. So it is robust across domains.[9]

A comparative analysis of MLA index estimates for intolerance to decline in standard of living, based on income and consumption data for a developed economy like the US, and an emerging economy like South Africa, show that the macroeconomic MLA index is leptokurtic. In South Africa it is explosive during periods of political uncertainty. Vizly, the Soweto uprising in 1976, P.W. Bothas hardliner Rubicon speech in 1985, and transition to democracy talks with the ANC in the early 1990s. However, in non-turbulent epochs the MLA index distribution for South Africa data exhibit mostly gain seeking behaviour with only a couple years where the index was not statistically different from the median value of 2.25 popularized by behavioural economics.

In contrast, for US data, the macroeconomic MLA index distribution has median value close to 2.25, and it is less responsive to political uncertainty. It is explosive during periods of financial market instability and natural disasters such as Hurricane Charlie and Ivan in 2004, and Katrina in 2005, and financial disasters such the Great Recession of 2008. To the best of our knowledge, the MLA index distributions for intolerance to decline in standard of living in the US and South Africa are new to the literature. Moreover, our measure of macroeconomic uncertainty implicates the Hofstede uncertainty avoidance index (UAI) which predicts similar risk attitudes in the US and South Africa.[10]

One important implication of the findings in this paper is the MLA index is stochastic. It constitutes independent and identically distributed jumps of a subordinate Lévy process. Therefore, it should be modelled accordingly. For example, a seminal paper by Benartzi and Thaler (1995) used a starting value of 2.25 in a simulation model to predict that a MLA index value of 2.7 resolves the equity premium puzzle upon convergence of their algorithm. An important paper by Bowman et al. (1999) assumed a constant loss aversion index value of 2.0 in their behavioural

[9] Details of the goodness of fit diagnostics for the statistical distribution are provided in Appendix 2.D.

[10] Technically, Hofstede's UAI is a riskless measure. It is based on questionnaire response and data reduction techniques like cluster analysis and there are no probabilities involved.

challenge to the permanent income hypothesis. More recent, Merkle (2015) used a ratio of slopes method to "infer loss aversion", and a Wald test to draw inference, in his study of investor subjective well-being relative to anticipated portfolio returns. The results in those papers are implicated by the α-stable MLA index finding in this paper. For example, since the index is a random variable, it is perhaps better specified as a random coefficient that accommodates large values, instead of extant specification as a nonstochastic parameter.

We provide an application which shows how our model contributes to emergent literature on information based asset pricing of defaultable bonds. Recent work by Hoyle (2010); Hoyle et al. (2011) model the filtration of information flow in binary bond pricing with a Lévy random bridge process. That model was recently extended by Ikpe et al. (2014) to include more abstract processes characterized by conditioning information. Our contribution extends the information process with embedded myopic loss aversion to default. In our model the credit risk index is based on a time changed Lévy process adapted to the natural filtration of a Cauchy bridge process to default. Throughout the life of the default path each jump in the credit rate index mimics myopic loss aversion to default. It signals market pessimism that is impounded in the price of the credit instrument. The cumulative effect of market loss aversion coincides with the terminal default date. We provide closed form expressions for the price of the defaultable credit instrument and examine its ability to explain the results in a Monte Carlo experiment in Hoyle (2010).

In an independently important application we resolve the equity premium puzzle with a novel behavioural asset pricing approach which produced several new results. The "puzzle" introduced by Mehra and Prescott (1985) stems from the unrealistically high risk aversion index required by their neoclassical model to explain the observed equity premium. We introduce a CCAPM which embeds the α-stable MLA index in the pricing kernel.[11] Grüne and Semmler (2008); Yogo (2008); Hung and Wang (2011); Curatola (2015) also introduced behavioural CCAPMs with loss aversion and consumption growth. An emerging class of models by Andries (2014); Easley and Yang (2014) and Guo and He (2015) combine exotic Epstein and Zin (1989) preferences with a MLA index in exposition of their asset pricing models. However, the loss aversion index in the prior models is either constant or deterministic or it is based on past gains or

[11] We provide an illustration in Appendix 2.A.1.

losses. Our behavioural pricing kernel with embedded α-stable MLA index does the following:

- predicts large equity premia;
- explains why the equity premium is counter-cyclical to business cycle peaks and troughs;
- identifies the source/cause of excess volatility in asset prices;
- identifies an endogenous fourfold ranking of asset prices;
- predicts large swings in the price of consumption risk.

According to our theory, the source of risk required to resolve the puzzle is not risk aversion but loss aversion. However, unlike Benartzi and Thaler (1995); Barberis et al. (2001) who used a constant or deterministic loss aversion index to resolve equity premium issues in their models, the MLA index in our model is an independent and identically distributed random variable. Since it is in the domain of attraction of an α-stable distribution, it admits MLA index values that induce a match with any observed equity premium. This feature of the model accommodates rare disaster effects on the equity premium. Cf. (Rietz, 1988; Barro, 2006, 2009; Wachter, 2013). Whereas the neoclassical pricing kernel is linear, our pricing kernel is nonlinear–a finding consistent with financial econometrics (Rosenberg and Engle, 2002), and behavioural finance (Shefrin, 2008). Furthermore, our model predicts large swings in the price of consumption risk–a result consistent with findings in Duffee (2005) study.

We prove that loss averse decision makers (DMs) in a current gain state in income, but who anticipate a loss state in income in the next period (GL states), place the highest value on assets. Evidently, they hope the asset would make up the anticipated short fall in income. Those in a current loss state in income, but who anticipate a gain in income in the next period (LG states), had the second highest evaluation. We prove that DMs are risk seeking in each of those GL and LG states.

The asset price valuation for lower ranked states LL (current loss, anticipated loss) and GG (current gain, anticipated gain) is more complex. We find that there is a critical level of consumption growth beyond which the equity premium predicted by the behavioural pricing kernel is higher than that predicted by the neoclassical model for risk averse DMs in GG states. It is lower otherwise. In LL states there is a small probability of risk aversion at which point the risk premium is negative. Otherwise DMs are risk seeking almost surely.

Our model predicts that DMs distort probabilities in the state transition probability matrix for the CCAPM. Investors in a current state of loss anticipate gains and shift probability mass from loss state to stochastically dominant gain states and this manifests itself as risk seeking over losses. This causes average returns and variances to be uniformly higher than that predicted by neoclassical models. In particular, this stochastic dominance feature of the model explains why the equity premium is high at the trough of a business cycle–an empirical regularity of equity premia (Campbell and Cochrane, 1999; Yogo, 2008). To the best of our knowledge, this stochastic dominance explanation of counter-cyclicality of the equity premium is new to the literature. Refer to Ludvigson (2013); Campbell (2015) for recent reviews of the literature.

Among the various pwf shapes implied by the permutation of ranked asset prices induced by our behavioural pricing kernel, is an endogenous fourfold ranking of *anticipated asset prices*: $\underbrace{p^D_{LL} \prec p^D_{GG} \prec p^D_{LG} \prec p^D_{GL}}_{\text{risk aversion}}$ or $\underbrace{p^D_{GG} \prec p^D_{LL} \prec p^D_{GL} \prec p^D_{LG}}_{\text{risk seeking}}$ where superscript D stands for behavioural asset pricing, and subscripts denote the transition states in which the asset was priced. The anticipated price is what DMs act on. It is a disposition effect. So for example, they sell winners in GG states too soon and hold on to losers in LG states too long. Risk aversion is supported by concave-convex pwfs, and risk seeking by convex-concave pwfs. We characterize the ranking in the context of Tversky and Kahneman (1992) fourfold pattern of risk attitudes. Perhaps more important, our model predicts the shape of the probability weighting function for relatively low income investors because it predicts switching in the shape of lower and upper tail ranks based on whether income growth crosses a certain threshold. This prediction of our theory finds support in Hartzmark (2015) study where "[t]he [rank] effect indicates that trades in a given stock depend on how it compares to other positions in an investors portfolio".

2.1.1 Positioning the paper in context of related literature

Subsequent to the influential paper on CPT by Tversky and Kahneman (1992), the loss aversion index is mostly estimated with data generated from controlled experiments in behavioural and experimental economics.[12] Refer to Abdellaoui et al. (2007); Booij et al. (2010) for a review. Moreover, loss aversion index estimates derived from economic experiments are used to calibrate influential models in behavioural economics and finance, e.g., Benartzi and Thaler (1995); Barberis et al. (2001); Barberis and Huang (2008). However, the conditions induced in controlled laboratory experiments, with student subjects, do not capture the full range of experiences a decision maker faces in the real economy.

To be sure, papers by Fox et al. (1996); Haigh and List (2005); Abdellaoui et al. (2013) have shown that the behaviour of financial professionals mirror that of student subjects in controlled laboratory experiments. However, the type of stress induced by extreme conditions of market failure were not replicated in those studies. For instance, losing a job or a home in an economic downturn should invoke a more visceral response than that observed from the relatively small stakes in economic experiments or hypothetical choices in survey instruments. Dunn and Mirzaie (2015) provide evidence that the stress levels in extreme conditions are more sensitive to uncollaterlized loans than collateralized loans. So one would expect the loss aversion index for calamitous events like the Great Recession of 2008 to be different from that found under controlled conditions.

Recently, De Neve et al. (2015) examined myopic subjective well being response to economic growth. They found that subjective well being is twice as sensitive to negative economic growth as it is to positive economic growth. Their results are based on an "experienced utility" approach wherein survey response (as opposed to economic time series data) to questions pertaining to life satisfaction is the dependent variable. However, our results show that estimates of the MLA index derived from survey data is α-stable.[13] So it is subject to large deviations which impact may be muted by aggregation bias. Thus, our results implicate the "twice as sensitive" finding in the De Neve et al. (2015) study.

More on point, our model is based on a novel interpretation of Duesenberry (1949) relative income hypothesis (RIH) consumption function with a skewed S-shaped value function for

[12] Notable exceptions include Hardie et al. (1993) who used supermarket scanner data to estimate a loss aversion index i a marketing context.

[13] Charles-Cadogan (2015b) introduced an econometric theory of loss aversion index recovery in cross sectional regression models of subjective well-being that also predict α-stable distribution for loss aversion index.

changes in consumption.[14] Duesenberry's value function is functionally equivalent to Kahneman and Tversky (1979) original prospect theory (OPT) value function and precedes it by at least three decades. The value function is related to the "surplus consumption" feature of Campbell and Cochrane (1999) habit formation model. Breeden et al. (2015) likened aspects of habit formation paradigm to the RIH. They posit that Constantinides (1990) "formulation is an extreme version of habit formation that implies a [Duesenberry (1949)] type racheting consumption demand that prevents consumption from falling below the exponentially weighted average of past consumption." For instance, Dusenberry's habitual consumption is a running maximum of past consumption over a given sliding window.[15] That is a robust estimator of the exponentially weighted average of past consumption. However, our model is distinguished from the class of habit formation models because we embed a MLA index in the consumption function. That specification provides a nexus for the joint distribution of the MLA index and consumption growth that drives the behavioural pricing kernel in our exposition of the CCAPM.

The rest of the paper proceeds as follows. In section 2.2 we provide the main results of our asymptotic theory of the MLA index from foundations of behavioural economics. In section 2.3 we embed the MLA index in Duesenberry (1949) RIH and identify its estimator. We provide empirical evidence that risk attitudes in the RIH induce probability distortions which support α-stable distributions. In section 2.4 we introduce a new proof of the nexus between probability weighting functions as outcome dependent transformations of utility functions. In section 2.5 we devise a strategy for estimating the MLA index from time series data. We apply and test the asymptotic theory in different contexts to establish robustness. In section 2.6 we show how our theory applies to information based asset pricing of defaultable binary bonds. In section 2.7 we show how our theory resolves the equity premium puzzle. We conclude in section 5.4 with some perspectives on avenues for further research.

[14] Our approach is distinguished from Dybvig (1995) and Riedel (2009) who derived Duesenberry's consumption ratcheting from a continuous time model of habit formation.

[15] The running maximum as reference point is also popular in the behavioural finance literature under rubric of "52-week high". See e.g., George and Hwang (2004); Huddart et al. (2009); Baker et al. (2012).

2.2 Foundations of the myopic loss aversion index estimator

In this section we provide some preliminaries, and we present the microfoundations of our model in a topological basis set in the Euclidean topology. Whereupon we derive an existence theorem for the statistical distribution of the MLA index. This theorem is tested (and upheld) empirically in subsequent sections.

Tversky and Kahneman (1992, p. 32) introduced a robust ratio of slopes procedure for estimating the loss aversion index. It amounts to a calibration exercise in which subjects were presented with two simple mixed lotteries $L_1 = \left(a, \frac{1}{2};\ b, \frac{1}{2}\right)$ and $L_2 = \left(c, \frac{1}{2};\ x, \frac{1}{2}\right)$, where a nd c are losses, b and x are gains, and $\frac{1}{2}$ is the corresponding probability of occurrence for each outcome. They reported the median value of x which subjects used to establish equivalence between the two lotteries , i.e., $L_1 \sim L_2$, for various a, b, c combinations. They employed the ratio

$$\lambda_\theta = \frac{x-b}{c-a} \tag{2.2.1}$$

as a robust estimator of the loss aversion index. Tversky and Kahneman (1992, p. 310) noted that "when the possible loss is increased by k the compensating gain must be increased by about $2k$". So λ_θ is a ratio of the slope of gains over the slope of losses in this "compensatory" framework. This sets the stage for a ratio type estimator for the loss aversion index.

2.2.1 The empirical myopic loss aversion (MLA) index estimator

We start with microfoundations of behavioural economics to motivate the theory behind our MLA index. Let v be CPT's value function bifurcated at a reference point x_r, and separated by sub-utility functions v_g and v_ℓ over gain and loss domains respectively. So that if x is change in income then we write the value function as

$$v(x) = v_g(x)\mathbb{I}_{\{x>x_r\}} - v_\ell(-x)\mathbb{I}_{\{x<x_r\}} \tag{2.2.2}$$

where $\mathbb{I}$ is an indicator function and points to the right of x_r are gains and points to the left are loses. Assume the existence of measurement error ε with mean 0 and variance σ_ε^2, i.e., $\varepsilon \sim (0, \sigma_\varepsilon^2)$, for change of income.

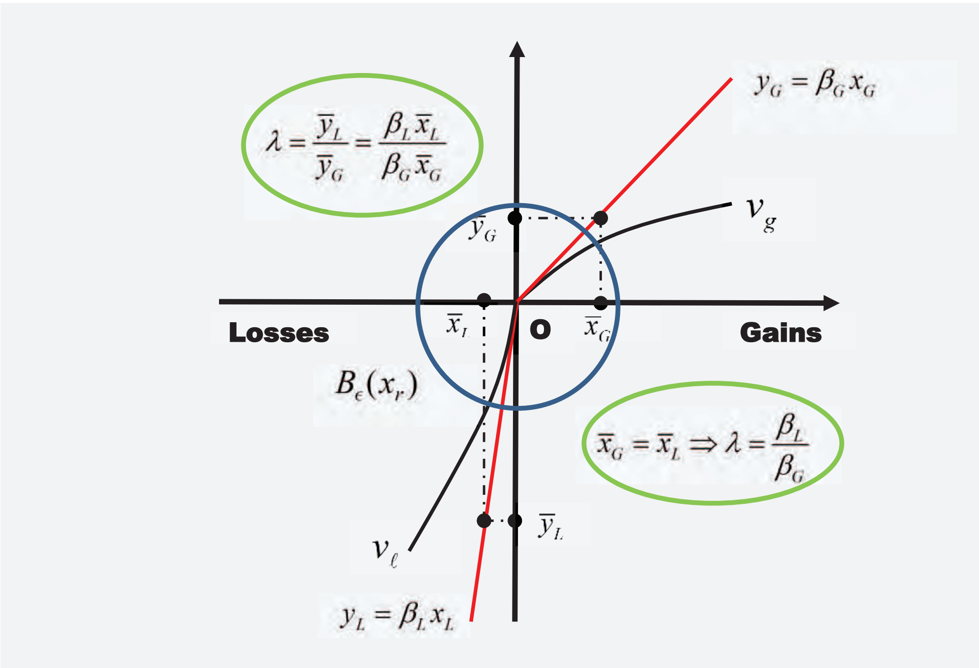
Definition 2.2.1 (Reference point topology)**.** Define a small open set centered at the reference point x_r with radius ε, as $B_\varepsilon(x_r) = \{x|\,|x-x_r| < \varepsilon\}$. The family $\mathscr{T}$ of such open sets forms a topology in $\mathbb{R}$. □

We make the following technical assumption which implies existence of convergent and approximate probability distributions in $B_\varepsilon(x_r)$. Refer to Gikhman and Skorokhod (1969, p. 441) for technical entails.

Assumption 2.2.1 (Existence of approximation)**.** *Assume that there exist a compact set* $K_\delta \subset B_\varepsilon(x_r)$ *such that* $Pr\{x \notin B_\varepsilon(x_r)\backslash K_\delta\} < \varepsilon_\delta$ *where* $\varepsilon_\delta \downarrow 0$.

This assumption accommodates the existence of "bump functions" or mollifiers (Karatzas and Shreve, 1991, p. 206) on the compact set K_δ that vanish on the set $B_\varepsilon(x_r)\backslash K_\delta$.

Figure 2.1: Geometry of MLA estimator



MLA estimator in an open ε disk $B_\varepsilon(x_r)$ for the bifurcated value function v introduced in Kahneman and Tversky (1979). It collapses to the Köbberling and Wakker (2005) ratio of slopes estimator when $\bar{x}_G = \bar{x}_L$ so that $\lambda = \dfrac{\beta_L}{\beta_G}$.

Definition 2.2.2 (Mollifier or bump function)**.** A mollifier or bump function is of type

$$\varphi(\varepsilon) = \begin{cases} C_0 \exp\left(-\dfrac{1}{a^2-\varepsilon^2}\right) & |\varepsilon| < a \\ 0 \quad \text{o.w} \end{cases} \tag{2.2.3}$$

where C_0 is a normalizing constant such that $C_0 \int_{-a}^{a} \varphi(\varepsilon) d\varepsilon = 1$. □

Remark 2.2.1. Siddiqui (2004, §5.4.1) defines the space of mollifiers or test unctions as "$C^\infty(u)$ equipped with the topology induced through the convergence is called the space of test functions and often denoted by $D(\Omega)$. In other words, a test function is an infinitely differentiable function on $\mathbb{R}^n$ identically zero outside of some compact set."

A first order Taylor expansion of the left hand side of (2.2.2) around a reference point x_r implies

$$v(x) = v(x_r) + v'(x_r)(x - x_r) + O_p(x^2) \tag{2.2.4}$$

where O_p is a function that is bounded in probability.[16] This first order expansion is applied separately to each subutility function on the right hand side of (2.2.2). The proof of Proposition 2 in Kőszegi and Rabin (2006, p. 1160) used a similar argument. The utility based loss aversion index measure for riskless choice (Tversky and Kahneman, 1991) implied by Tversky and Kahneman (1992) is $\lambda = \dfrac{v_\ell(-1)}{v_g(1)}$, and the measure proposed by Köbberling and Wakker (2005) is $\lambda = \dfrac{v'_\ell(0^-)}{v_g(0^+)}$.[17] Blavatskyy (2011) introduced a generalized loss aversion index formula given by

$$\lambda = -\left(\frac{u(x_-) - u(r)}{u(x_+) - u(r)}\right) \tag{2.2.5}$$

where r is a reference point and x_+ and x_- are gains and losses respectively. When $r = 0$ his formula collapses to the ones above. Refer to Wakker (2010) for a review of loss aversion index formulae. Thus, our task is to establish a correspondence between (2.2.4) and those loss aversion

[16] We note that nondifferentiability at the reference point x_r implies that $\lim_{x \downarrow x_r^+} v'(x) \neq \lim_{x \uparrow x_r^-} v'(x)$ if $v(x)$ is piecewise continuous at x_r, i.e., it has no jump discontinuity. If $\eta = O_p(x^2)$, then by virtue of Assumption 2.2.1, for some constant C_ε we have $\Pr\{|\eta| > C_\varepsilon x^2\} \leq 2\varepsilon$. See Chow and Teicher (1988, p. 255) for details.

[17] The related concept of probabilistic loss aversion is presented as a ratio of slopes of gain and loss domain dependent probability weighting functions in Schmidt and Zank (2005, 2008). However, that is outside the scope of the present paper.

index formulae.

Kahneman and Tversky (1979) assumed a reference point $x_r = 0$[18] and we do the same here. Let $y = v(x)$, $x_r = 0$, $\alpha = v(x_r) = 0$, $\beta = v'(x_r)$ and $\eta = O_p(x^2)$. The piecewise linear function in Figure 2.1 has a kink at $x_r = 0$. Since $x_r = 0$ is "known" we can treat α and β as parameters. Whereupon $\alpha = 0$ and $\beta_L = v'(0^-)$ over loss domain, and $\beta_G = v'(0^+)$ over gain domain. Thus we rewrite (2.2.4) as a simple linear model:

$$y = \alpha + \beta(x - x_r) + \eta \Longrightarrow \bar{y} = (\alpha - \beta\, x_r) + \beta\, \bar{x} + \bar{\eta} = \beta\, \bar{x} + \bar{\eta} \tag{2.2.6}$$

This is the equation of a straight line that passes through the point $(\bar{x}, \bar{y})$, with intercept $\alpha = 0$. Since $v(x) = v_g(x)$ over gain domain, and $v(x) = v_\ell(x)$ over loss domain, application of (2.2.4) and (2.2.6) gives us the MLA index corresponding to $\lambda = \dfrac{v_\ell(-1)}{v_g(1)}$. So that

$$\lambda = -\frac{\bar{y}_L}{\bar{y}_G} = -\frac{\beta_L\, \bar{x}_L + \bar{\eta}_L}{\beta_G\, \bar{x}_G + \bar{\eta}_G} \tag{2.2.7}$$

where the G and L subscripts pertain to gain and loss domains, and the *negative sign* is retained so that $-\bar{y}_L > 0$ and $-\bar{x}_L > 0$ since λ is positive. The line in (2.2.6) passes through the origin so $\bar{y} = \beta\, \bar{x}$. This procedure applies to gain and loss domains.[19] It is depicted in Figure 2.1 for $\beta > 0$. Under the identifying restriction $x_r = 0$ we have the following empirical measure of the loss aversion index which corresponds to $\lambda = \dfrac{v'_\ell(0^-)}{v_g(0^+)}$

$$\lambda = -\frac{\bar{y}_L}{\bar{y}_G} = -\frac{\beta_L \bar{x}_L}{\beta_G \bar{x}_G} \tag{2.2.8}$$

We summarize the foregoing in the following:

Theorem 2.2.2 (MLA estimator)**.** *The empirical MLA index estimator in* (2.2.7) *pertains to Tversky and Kahneman* (*1992*) *utility ratio formula. Whereas the empirical MLA index estimator in* (2.2.8) *pertains to Köbberling and Wakker* (*2005*) *ratio of marginal utility formula in a small* ε-

[18] This was subsequently extended by Kőszegi and Rabin (2006) to include a generalized reference point.

[19] In their reference dependent model, Kőszegi and Rabin (2006, p. 1146) write this relation as $\mu(x) = \eta\, x$, $x > 0$ and $\mu(x) = \lambda\, \eta\, x$, $x < 0$ where their $\eta > 0$ (our β) is the weight a subject attaches to "gain-loss utility".

neighbourhood of the origin. □

Remark 2.2.2. The linear approximation implies that the MLA index estimator should produce reasonably close estimates for the Tversky and Kahneman (1992) and Köbberling and Wakker (2005) estimators in (2.2.7) and (2.2.8), respectively. However, in economic experiments the values produced by those estimators differ (Abdellaoui et al., 2007, p. 1662).

2.2.2 The limit distribution of MLA index

Before we derive the distribution for λ we need the following preliminary definitions.

Preliminaries

Definition 2.2.3 (Generalized Cauchy)**.** A probability density function f is standard Cauchy, written $\mathscr{C}(0,1)$ if $f(x) = 1/\pi(1+x^2)$, $-\infty < x < \infty$. If Y follows a standard Cauchy law, then Z has a generalized Cauchy law $\mathscr{C}(a,b)$ if $Z = bY + a$, where b is a scale and a is a location parameter. □

Definition 2.2.4 (Stable distribution)**.** Samoradnitsky and Taqqu (1994). A random variable X is said to have a stable distribution if for any positive numbers A and B, there is a positive number C and a real number D such that

$$AX_1 + BX_2 \stackrel{d}{=} CX + D \tag{2.2.9}$$

where X_1 and X_2 are independent copies of X and $\stackrel{d}{=}$ denotes equality in distribution. □

Another popular definition of α-stable is if $X_1, \ldots, X_n$ are independent and identically distributions (iid) random variables, and there exist constants c_n, d_n such that $X_1 + \cdots + X_n \stackrel{dist}{=} c_n X + d_n$ where X has the same distribution as the $X_i'^s$, and $c_n = n^{\frac{1}{\alpha}}$, $0 < \alpha \leq 2$, then X is α-stable. An α-stable distribution (such as the normal distribution) retains its shape up to scale c and shift d after addition. Refer to Samoradnitsky and Taqqu (1994) for further details.

Theorem 2.2.3 (α-stable distribution)**.** *Samoradnitsky and Taqqu (1994). For any stable random* *variable X , there is a number $\alpha \in (0,2)$ such that the number C in (2.2.9) satisfies*

$$C^{\alpha} = A^{\alpha} + B^{\alpha} \tag{2.2.10}$$

The number α is called the index of stability or characteristic exponent. A stable random variable X with index α is called α-stable. □

Proof. See Feller (1970, § VI.1, pp. 170-171). □

Definition 2.2.5 (Spherically symmetric vector)**.** Arnold and Brockett (1992). A random vector $\boldsymbol{U}$ is said to be spherically symmetric if $\boldsymbol{\Gamma U}$ has the same distribution as $\boldsymbol{U}$ (i.e., $\boldsymbol{\Gamma U} \sim \boldsymbol{U}$) for all orthogonal matrices $\boldsymbol{\Gamma}$. □

Definition 2.2.6 (Elliptically symmetric)**.** Arnold and Brockett (1992). A random vector $\boldsymbol{X}$ is said to be elliptically symmetric if there exists an invertible matrix $\boldsymbol{A}$ such that $\boldsymbol{X} = \boldsymbol{AU}$ where $\boldsymbol{U}$ has a spherically symmetric distribution. □

Existence theorem for generalized Cauchy for MLA index

Assume that gains and losses are symmetric around the reference point x_r. Assumption 2.2.1 implies the existence of bump functions characterized by a symmetric elliptic distribution that vanishes outside of K_{δ}.[20] Let $\boldsymbol{X} = (X_1, \ldots, X_r, \ldots, X_n)^T$ be a $n \times 1$ vector of random variables for gains and losses with an elliptically symmetric distribution around a reference point X_r. We state the following theorems implied by the foregoing assessment, and provide proofs in the appendix.

Theorem 2.2.4 (Standard Cauchy distribution)**.** *If $\boldsymbol{U} = (U_1, \ldots, U_n)^T$, has a spherically symmetric distribution, then, for $i \neq k$ U_i/U_k has a standard Cauchy distribution.*

Proof. See Appendix 2.B.1. □

Theorem 2.2.5 (Generalized Cauchy distribution)**.** *If $\boldsymbol{X} = (X_1, \ldots, X_r, \ldots, X_n)^T$ has an elliptically symmetric distribution, then, for $i \neq k, X_i/X_k$ has a general Cauchy distribution.*

[20] Anderson (2003, p. 47) provides analysis for spherical and elliptically symmetric distributions. Owen and Rabinovitch (1983); Landsman and Valdez (2003) provide applications of symmetric elliptic distributions in economics and finance.

Proof. See Appendix 2.B.2. □

Theorem 2.2.6 (Existence of Generalized Cauchy distribution for MLA estimator)**.** *Let* $\bar{\boldsymbol{U}} = \left(\bar{X}^L \; X_r \; \bar{X}^G\right)^T$ *be a vector comprised of the sample mean of loss and gain values relative to a reference point* X_r*. Assume that* $\bar{\boldsymbol{U}}$ *is elliptic symmetric, and define* $\bar{\boldsymbol{X}} = \bar{\boldsymbol{A}}\bar{\boldsymbol{U}}$*. There exist a* 3×3 *invertible matrix* $\bar{\boldsymbol{A}}$ *such that* $\bar{X}^L / \bar{X}^G \sim \mathscr{C}(a,b)$*.*

Proof. See Appendix 2.B.3. □

2.3 The relative income hypothesis with myopic loss aversion

The main purpose of this section is to establish a nexus between loss aversion to decline in the standard of living and the relative income hypothesis (RIH). This provides a basis for the MLA index theory to be tested. We derive the consumption function under Duesenberry (1949) RIH, and prove that it is a piecewise linear version of Kahneman and Tversky (1979); Tversky and Kahneman (1992) value function over gain and loss of income.[21] According to Shea (1995, pp. 798-799) "Under myopia, consumption tracks current income. Thus, the failure of the [Life Cycle Hypothesis/Permanent Income Hypothesis] should be symmetric: consumption should respond equally to predictable income increases and decreases." We use that observation as a basis for the following

Axiom 1 (Myopia)**.** *Under myopia, consumption tracks income.*

Figure 2.37 in Appendix 2.C.2 illustrates "myopia" under the "consumption tracking income" postulate for US nondurable consumption and real disposable income series. The US and South Africa data used in the sequel were taken from publicly available data at the Federal Reserve Bank-St. Louis (FRED database) and the South African Reserve Bank (SARB database) websites.[22]

[21] Benartzi and Thaler (1995); Barberis et al. (2001) used related value function specification in their behavioural asset pricing models.

[22] Available at https://research.stlouisfed.org/fred2/ and http://wwwrs.resbank.co.za/webindicators/EconFinDataForSA.aspx.

2.3.1 Reference dependence and the relative income hypothesis

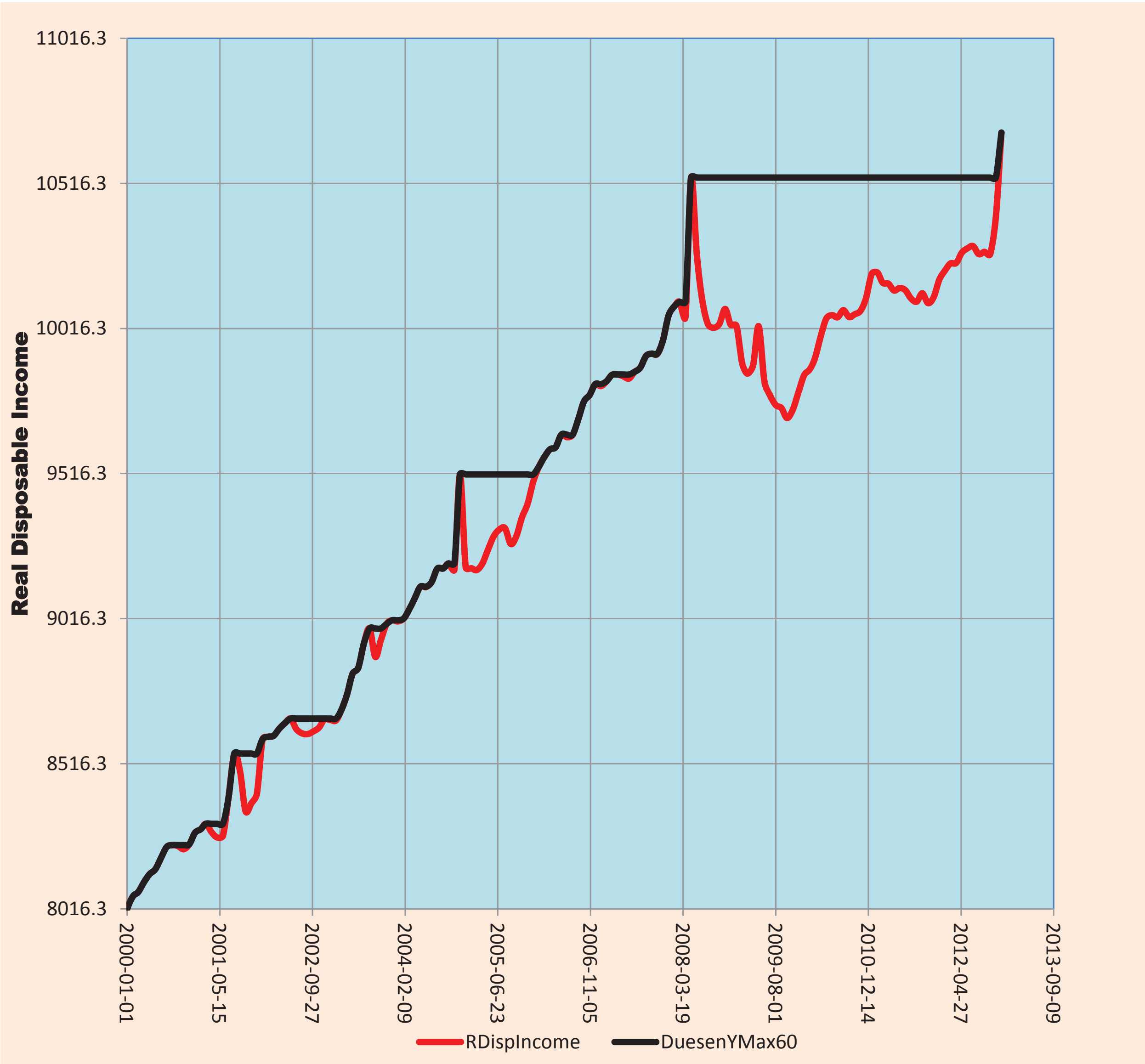
Let Y_t be disposable income, $t = 1, \ldots, T$, S_t be savings and C_t be consumption of an investor at time t so $C_t + S_t = Y_t$. Let M_t be the running maximum income measured at time t *for all periods before* t, i.e., $M_t = \max_{0<s<t} Y_s$.

The running maximum M_t is critical to our analysis. It is the variable that captures retrospective standard of living and induce nonseparability of consumption. Moreover, it is an independently important subordinate income process. As a practical matter, it is measured over a finite sliding window u so that $M_t(u) = \max_{t-u \leq s<t} Y_s$. For example, if the real income distribution of an individual over the last five years is $\{20,000; 25,000; 19,000; 27,000; 22,000\}$, then her maximum is $27,000$. As that 5-year window slides over time it retains the highest maxima attained in retrospect. To see this, suppose we considered a 10-years period obtained by concatenating the five years above with the following five years income distribution $\{18,000; 21,000; 26,000; 23,000; 26,500\}$. The highest income in the last set is $26,500$. However, the 5-year rolling window over the 10-years includes

$$\max\{25,000; 19,000; 27,000; 22,000; 18,000\} = 27,000;$$
$$\max\{19,000; 27,000; 22,000; 18,000; 21,000\} = 27,000;$$
$$\max\{27,000; 22,000; 18,000; 21,000; 26,000\} = 27,000;$$
$$\max\{22,000; 18,000; 21,000; 26,000; 23,000\} = 26,000;$$
$$\max\{18,000; 21,000; 26,000; 23,000; 26,500\} = 26,500$$

So the highest standard of living attained remained fairly stable at 27,000 until income systematically dropped in the last 5-years period. If income did not systematically drop over any 5-year period, it will have only upward jumps. Over the entire 10-years period $M_{10} = \max_{0<s<10} Y_s = 27,000$. The five years period was arbitrarily chosen. However, it is consistent with empirical literature on consumption and income, e.g., Parker and Julliard (2005); Guvenen et al. (2014); Bandi and Tamoni (2015), and evaluation of life satisfaction over a period of time popularized in the subjective well being literature spawned by Cantril (1965). For example,

Figure 2.2: US Relative Income Over 60-months Sliding Window



Duesenberry's monthly relative income $M_t = \max_{t-u<s<t}\{Y_s\}$ is a running maximum for real disposable income Y_t over a select window u taken here to be a 60-months or 5-years sliding window over the period 2000:10–2012:11. The first 5-years of data in the monthly time series between 1995 and 2000 is used to derive the first maximum value 8016.3 above.

Guvenen et al. (2014) studied the cyclical behaviour of 5-year average income of top income earners over the business cycle. Bandi and Tamoni (2015) studied aggregate consumption growth over heterogenous durations that subsume 5-year cycles to explain equity premia in their model. Figure 2.2 depicts a plot of the highest standard of living attained for US real disposable income over a rolling 5-year window. More will be said about the construction of that plot in the sequel.

Duesenberry (1949, p. 4) described his relative income model as follows:

> "If in periods of steadily rising income the savings ratio is constant while in depressions the ratio depends on current income and previous peak income, we can explain saving with the relation $S_t/Y_y = 0.25Y_t/Y_0 - 0.196$, where S_t, and Y_t, are current saving and disposable income respectively and Y_0 is highest previous disposable income. When fitted to the data, this relation yields a high correlation. Moreover, it accurately predicts the savings rates of 1947." (Duesenberry, 1949, p. 4).

We parameterize those statements as follows:

$$\frac{S_t}{Y_t} = \alpha_0 + \alpha_1 \frac{Y_t}{M_t}, \; M_t = \max_{0 \leq s \leq t} \{Y_s\}, \;\; \alpha_0 > 0, \alpha_1 > 0 \Rightarrow \frac{C_t}{Y_t} = 1 - \alpha_0 - \alpha_1 \frac{Y_t}{M_t} \tag{2.3.1}$$

$$Y_t > M_t \Rightarrow \text{ income gain } \quad \frac{Y_t}{M_t} = 1 + g_t^G \tag{2.3.2}$$

$$Y_t = M_t \Rightarrow \text{ reference income } \quad \frac{Y_t}{M_t} = 1 \tag{2.3.3}$$

$$Y_t < M_t \Rightarrow \text{ income loss } \quad \frac{Y_t}{M_t} = 1 - g_t^L \tag{2.3.4}$$

where $g_t^L > 0$ and $g_t^G > 0$ are relative-growth rates of income, and α_1 is a savings rate factor. Note that M_t does not include the current period in its evaluation. In continuous time the evaluation is over $0 < s < t^-$. In discrete time it is $0 < s \leq t-1$. We can rewrite (2.3.1)-(2.3.4) for *change in consumption* $\Delta^b C_t^D$ as follows:

Duesenberry's reference dependent change in consumption

Theorem 2.3.1 (RIH consumption function with embedded MLA index)**.**

$$C_t^D = \begin{cases} a(d))Y_t + \Delta^b C_t^D & \text{if gain in income} \\ a(d))Y_t & \text{if reference income} \\ a(d))Y_t - \lambda_t \Delta^b C_t^D,\ \lambda > 0 & \text{if loss of income} \end{cases} \tag{2.3.5}$$

$$\Delta^b C_t^D = -\alpha_1 g_t^G Y_t, \qquad \lambda_t = \frac{|g_t^L|}{g_t^G}, \qquad a(d) = 1 - \alpha_0 - \alpha_1 \tag{2.3.6}$$

Remark 2.3.1. Since λ_t is constrained to be positive by definition, we use the absolute value $|g_t^L|$ instead of $g_t^L < 0$.

Proof. See Appendix 2.B.4. □

We formalize the embedded myopic loss aversion index in the foregoing with the following

Theorem 2.3.2 (MLA to decline in standard of living)**.** *Myopic loss aversion to decline in standard of living induces asymmetric response to anticipated gains and losses in relative income.* □

In (2.3.5) a gain in relative income signifies an increase in savings and decline in consumption in (2.3.1). Whereas a loss in relative income induces decreased savings and asymmetric or "irreversible" increase in consumption in (2.3.1) (Duesenberry (e.g., 1949, p. 101); and Komlos (2014)). λ_t is a *reference dependent* loss aversion index, and $\Delta^b C_t^D$ is the *piecewise [linear] change in consumption* consistent with that in Figure 2.1. Benartzi and Thaler (1995, p. 83) and Barberis et al. (2001, p. 12) used a functionally equivalent piecewise linear value function specification for stock returns in their analyses. In fact, the variable referred to as "historical benchmark level Z_t" in Barberis and Huang (2001, p. 9) is our M_t in (2.3.1). In Campbell and Cochrane (1999) habit formation model a quantity like $\Delta^b C_t^D$ is referred to as a consumption surplus.

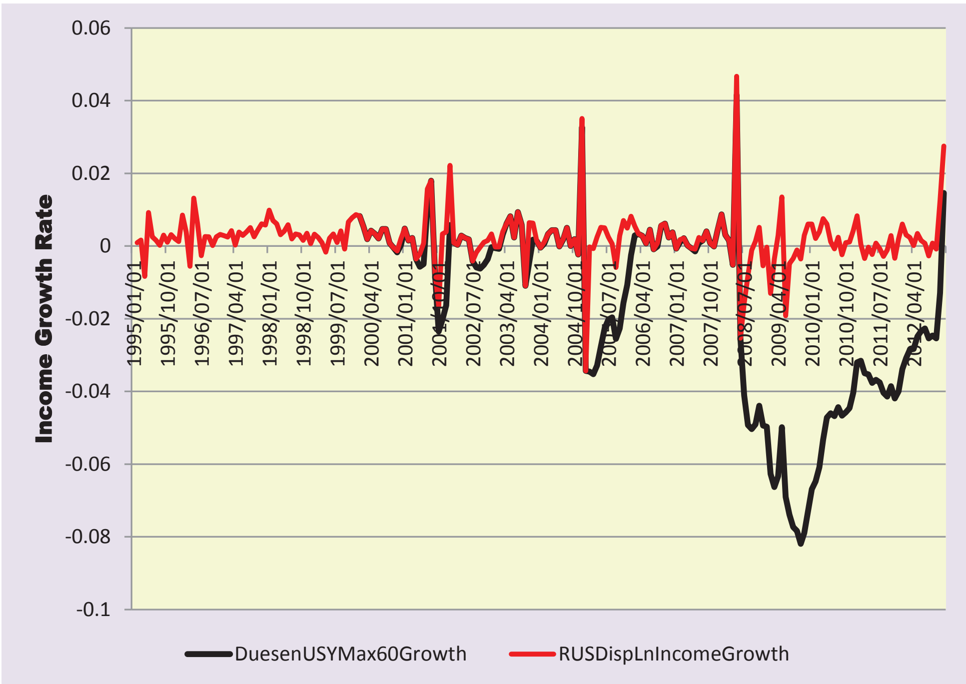
2.3.2 Relative income dynamics for US and South Africa

Figure 2.2 depicts a plot of the RIH for monthly real disposable income in the US. The standard of living $M_t(u)$ is measured over a 5-year or 60-months rolling window. It jumps only when there is an increase in real income or it stays flat otherwise. Hence $M_t(u)$ is a subordinate income process.

Cursory inspection shows that there was a persistent decline in the standard of living for at least 5-years after the onset of the Great Recession of 2008.

Figure 2.3 depicts US real income growth, and relative income growth. The "pain" or intolerance associated with a decline in standard of living is reflected by the exaggerated downward growth. According to (2.3.5) and (2.3.6) this reflects loss aversion to decline in consumption and intolerance to decline in standard of living. The negative growth implies that the level of happiness or well-being in the economy has declined.

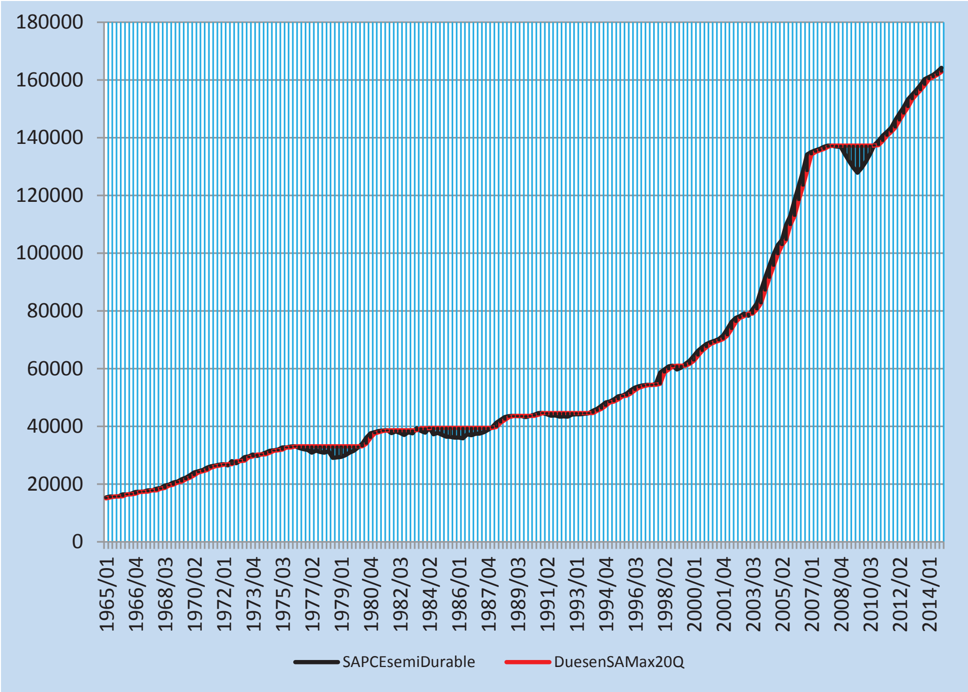
Figure 2.3: US Relative Income Growth Over 60-months Sliding Window



US monthly relative income growth is computed from $g_t = \ln(Y_t) - \ln(M_t)$ where $M_t = \underset{t-u<s<t}{max}\{Y_s\}$ is a running maximum for real disposable income Y_t over a select window u taken here to be a 60-months or 5-years sliding window over the period 2000:10–2012:11. The exaggerated negative growth reflects the psychological pain associated with loss of income and consequent intolerance for decline in standard of living.

Figure 2.4 is the South Africa analog of Figure 2.2. By virtue of Axiom 1 on myopic consumption tracking income, and without loss of generality, we used semi-durable personal consumption expenditure in 2010 prices as an instrument for income since a suitable income series was unavailable at the SARB website. Cursory inspection shows that with few exceptions, the "consumption ratchet" is persistent.

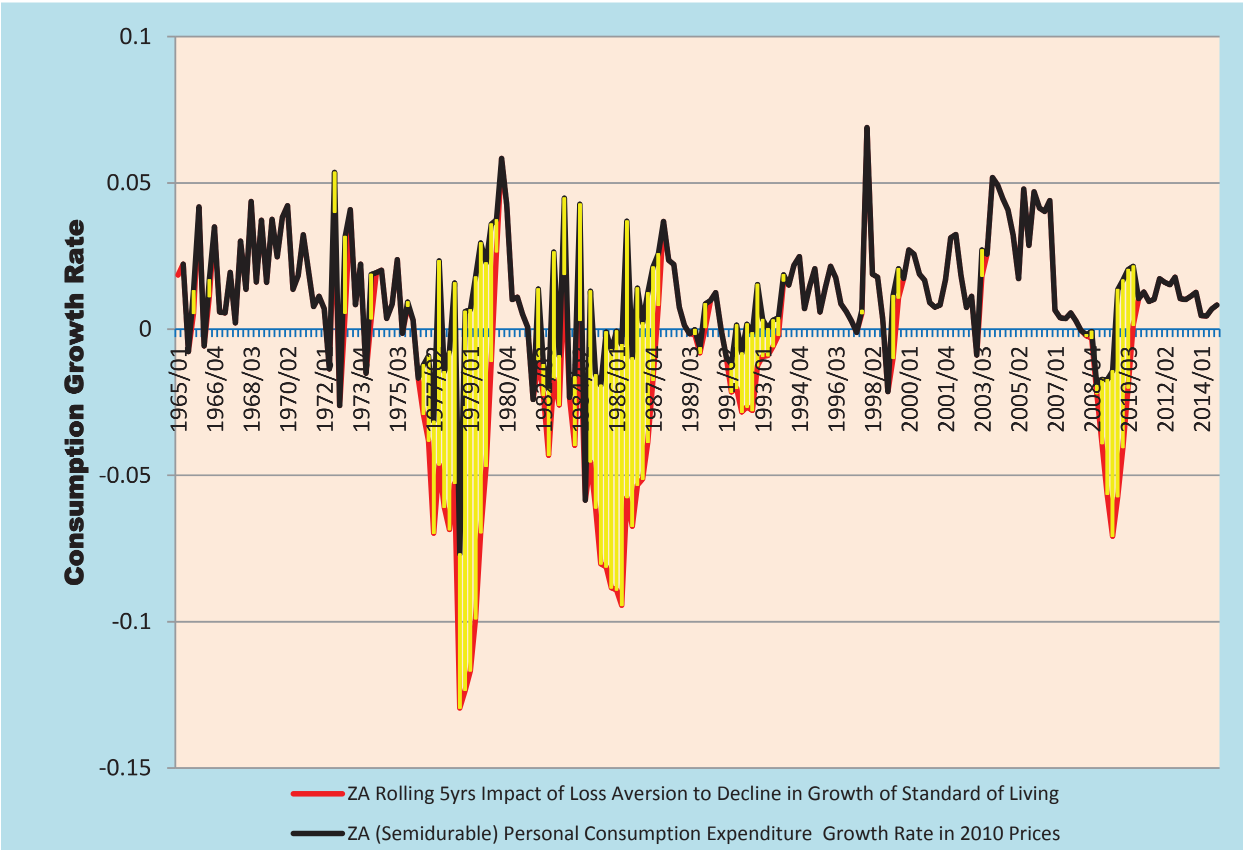
Figure 2.4: South Africa's Relative PCE Over 20Q Sliding Window



Semidurable personal consumption expenditure (PCE) in 2010 prices is used as an instrument for South Africa's quarterly relative income. $M_t = \max_{t-u<s<t} \{Y_s\}$ is a running maximum for the income instrument Y_t over a select window u taken here to be a 20-quarters or 5-years sliding window over the period 1960:1-2014:1. The first 20-quarters or 5-years of data in the quarterly time series is used to derive the first maximum value. Under the consumption tracking income myopia hypothesis, the exaggerated negative growth reflects the psychological pain associated with loss of income and consequent intolerance for decline in standard of living.

Figure 2.5 depicts the growth rates in quarterly personal consumption expenditure (ZA PCE) for 2010 base year for South Africa. The shaded regions coincides with the Soweto uprisings around 1976; P. W. Botha's hardliner Rubicon speech in 1984, prelude to democracy talks with the ANC in the early 1990s, and the Great Recession in 2008. Thus, political uncertainty drives the MLA index implied by Figure 2.5. Loss aversion accentuates the decline in growth rates.

Figure 2.5: South Africa's Relative PCE Growth Over 20Q Sliding Window



South Africa's quarterly relative PCE growth in 2010 prices is computed from $g_t = \ln(Y_t) - \ln(M_t)$ where $M_t = \max_{t-u<s<t}\{Y_s\}$ is a running maximum for the PCE instrument for income Y_t over a select window u taken here to be a 20-quarters or 5-years sliding window over the period 1960:1-2014:1.

2.4 Probability weighting functions induced by myopic loss aversion

In this section, we show how α-stable distributions are induced by utility based loss aversion, and risk aversion, towards outcomes by projection of mollifiers (test functions) in the base topology $B_\varepsilon(x_r)$ in Definition 2.2.1. One of the empirical regularities of prospect theory is risk seeking over losses, and risk aversion over gains (Kahneman and Tversky, 1979, p. 268). These are typically represented by convex and concave utility functions over loss and gain domains, respectively. Another empirical regularity is nonlinear transformation of probabilities P with a probability weighting functions $w(P)$. The latter carries probabilistic risk attitudes. Mosteller and Nogee (1951) conducted one of the earliest experiments to elicit utility functions from a mixed gamble in which probabilities were fixed and a certainty equivalent payoff was elicited. Preston and Baretta (1948) were the first to publish a plot of an inverted S-shape probability weighting functions (pwf) from experiment data. Mosteller and Nogee (1951) found that the utility functions of several subjects were convex over small sums of money and concave over comparatively larger sums. Furthermore, they found that subjects probabilistic risk attitudes coincided with the Preston-Baretta finding. This implies that probabilistic risk attitudes and utility risk attitudes are related to each other. This relationship is formalized in Pfanzagl (1967) and Schneeweiss (1974).

The inverted S-shape probability weighting function was popularized in the economics literature by Quiggin (1982) rank dependent utility (RDU) model, and Tversky and Kahneman (1992) CPT model. The literature often treats pwfs as carriers of probabilistic risk attitudes distinct from the risk attitudes characterized by utility functions. For example, Schmidt and Zank (2005, 2008) argue in favor of probabilistic loss aversion based on analysis of pwfs. In this section, we establish a theory based nexus between utility and probabilistic risk attitudes below. We also provide a new proof of the nexus and use it to explain a seemingly misinterpreted experiment result.

2.4.1 Skewed mixture distributions induced by reference point topology

Figure 2.6 and Figure 2.7 depict projection of the distribution of ranked outcomes that support a mollifier centered at reference outcome. CPT is a rank and sign dependent model (Luce, 2000) based on separate application of rank dependent utility (RDU) (Quiggin, 1993) over gain and loss domains. To the extent that RDU is a generalized expected utility model, is the extent to which concepts like Arrow-Pratt risk measure is applicable to the underlying subutility functions in CPT (Köbberling and Wakker, 2005, p. 128). In the economics literature a concave utility function like that in Figure 2.6 is often associated with risk aversion.[23] A risk averter transforms the underlying objective distribution of outcomes x relative to a reference outcome x_r. Points to the left of the reference point x_r are treated as “losses”, and points to the right are treated as ”gains”. This distorts the distribution and induces a left tail skew that corresponds to a fear of losses, i.e., loss aversion. Thus, Figure 2.6 illustrates the connection between risk aversion and loss aversion.

[23] Hansson (1988) argues against this notion and provides alternative definitions of risk aversion.

Figure 2.6: Risk averter utility function

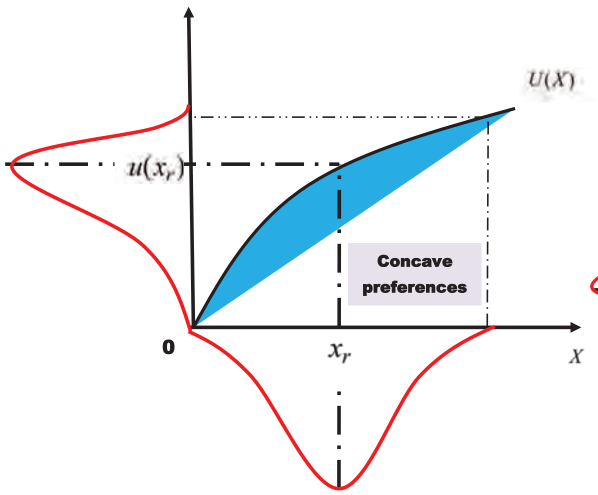
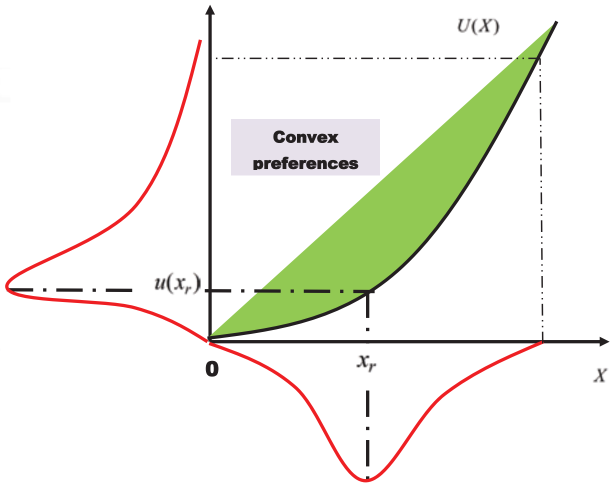


Figure 2.7: Risk seeker utility function



Risk attitudes over mollifiers centered at reference point x_r. Left tail skew distribution projected by risk averter in Figure 2.6 implies risk averter avoids losses from tail risk. Right tailed skew distribution projected by risk seeker in Figure 2.7 implies risk seeker attracted to gains from long shot bias.

Figure 2.7 depicts a risk seeker's transformation of outcomes relative to the same reference outcome x_r. In this case, the effect of losses is muted and the effect of gains are magnified. This is a manifestation of long shot bias (Ali, 1977; Golec and Tamarkin, 1998), and stocks as lotteries phenomenon (Barberis and Huang, 2008; Kumar, 2009). It shows how DMs transform probability distributions based on their preference for certain outcomes, see e.g., Lopes (1987); Lopes and Oden (1999) security/potential, aspiration (S/P A) theory. The risk seeker is overly optimistic about gains relative to losses. This gain seeking behaviour is reflected by right tail skew.

We present the following analytic framework in the topological basis set $B_\varepsilon(x_r)$ for the geometry above. Let $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. Consider a second order Taylor expansion of a utility function u around a reference point x_r with "measurement error" ε.

$$u(x_r + \varepsilon) \approx u(x_r) + \varepsilon u'(x_r) + \frac{\varepsilon^2}{2} u''(x_r) \tag{2.4.1}$$

$$\implies \Delta_r^\varepsilon u(x_r) = u(x_r + \varepsilon) - u(x_r) = \alpha_1 \varepsilon + \alpha_2 \varepsilon^2 \tag{2.4.2}$$

where $\alpha_0 = u(x_r)$, $\alpha_1 = u'(x_r)$, and $\alpha_2 = u''(x_r)/2$ are parameters. $\alpha_2 < 0$ for concave utility functions and $\alpha_2 > 0$ for convex utility functions. So the Arrow-Pratt risk measure at x_r is given by

$$r_A(x_r) = -\frac{u''(x_r)}{u'(x_r)} = -\frac{2\alpha_2}{\alpha_1} \tag{2.4.3}$$

It is known that if $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$, then $\varepsilon^2 \sim \sigma_\varepsilon^2 \chi_1^2$, where $E[\chi_1^2] = 1$ and variance$[\chi_1^2] = 2$ are the mean and variance of a chi-squared distribution with 1-degree of freedom (Johnson et al., 1994). Therefore, the utility transformation in (2.4.2) is a mixture normal and χ^2-distribution weighted by α_1 and α_2 respectively. The skewness and long tail features of the χ^2 distribution are evident in Figure 2.6 and Figure 2.7.[24] Furthermore, (2.4.2) is characterized by excess kurtosis inherited from the χ^2 component. Moreover, imposition of Inada conditions (Inada, 1963) on u imply that $\lim_{x \downarrow 0} u'(x) = \infty^+$. Thus, α_1 can become quite large and explosive under certain conditions. In which case the mixture distribution portended by the transformation is an α-stable distribution. Thus we prove the following

Proposition 2.4.1 (Existence of utility induced α-stable distribution of outcomes)**.** *The utility transformation of reference outcome with normally distributed measurement error in the base topology $B_\varepsilon(x_r)$ induces a skewed mixture distribution that admits an alpha-stable distribution.* □

Proposition 2.4.1 provides further proof of the existence of an α-stable distribution induced by risk attitudes around a reference point x_r. It differs from the proof in Dillenberger et al. (2013) stake dependent probability weighting function. Those authors assume monotonic transformation of preferences to facilitate their representation theory of presences based on a "state dependent" probability distribution theory.

[24]Skewness is measured by the formula $\gamma = \frac{\mu_3}{\mu_2^{3/2}}$ where μ_2 and μ_3 are second and third central moments, respectively. $\gamma_\varepsilon = 0$ for the normal distribution and $\gamma_{\varepsilon^2} = 2\sqrt{2/k}$ for the χ^2-distribution with k-degrees of freedom. For kurtosis $\kappa_\varepsilon = 3$ for the normal and $\kappa_{\varepsilon^2} = 3 + (12/k)$ for the χ^2-distribution. (Johnson et al., 1994). For excess kurtosis we subtract 3 so its 0 for the normal and $12/k$ for χ^2 distributions.

2.4.2 Optimism, pessimism, and probabilistic risk attitudes over ranked outcomes

Preliminaries–Normal distribution mollifier

Assume that the outcomes in Figure 2.6 and Figure 2.7 are ranked from worst to best on outcome space $X \subset \mathbb{R}$. Furthermore, Assumption 2.2.1 allows us to use a normal distribution as mollifier at reference point x_r. For purposes of exposition we assume that the mollifier vanishes for $x > |x_r|$. Thus, in an ε-neighbourhood $B_\varepsilon(x_r)$ centered ar x_r with $\varepsilon \sim \mathscr{N}(0, \sigma_\varepsilon^2)$, for $x \in B_\varepsilon(x_r)$, we have $x = x_r + \varepsilon$ where $E[x] = x_r$ and $V(x) = \sigma_\varepsilon^2$. This allows us to use a Z-transformation of outcome space X such that $Z = \dfrac{X - x_r}{\sigma_\varepsilon}$. So the reference point x_r is transformed to reference point 0 in Z transform space. If we choose $X = \{x | 0 \leq x \leq 2x_r\}$, then the mollifier condition is such that it has compact support on $-x_r/\sigma_\varepsilon \leq z \leq x_r/\sigma_\varepsilon$ and vanishes outside of that set. Thus, $u(x_r) \mapsto u(0)$ and $u(x) \mapsto u(z) = u((x - x_r)/\sigma_\varepsilon)$ where lower case z is an observed Z-score. Moreover, $u(-x_r/\sigma_\varepsilon) \leq u(z) \leq u(x_r/\sigma_\varepsilon)$. Note that $\Delta_r^\varepsilon u(x_r) = u(x_r + \varepsilon) - u(x_r) = \alpha_1 Z + \alpha_2 \chi$ is the change in utility measured on the vertical axis relative to reference utility $u(x_r)$ where χ is a χ^2 r.v. So the change in utility is a mixture distribution.

Endogenous outcome dependent probability weighting functions

Claim: If subjects overweight low ranked outcomes, then we expect that if the probability supported by $0 \leq x \leq x_r$ is p, then the probability supported by the projected utility interval $0 \leq u(x) \leq u(x_r)$ will be larger than p. We prove this claim as follows.

We consider the probabilistic relationships

$$\Pr\{0 \leq x \leq x_r\} = \Pr\{-\frac{x_r}{\sigma_\varepsilon} \leq z = \frac{x - x_r}{\sigma_\varepsilon} \leq 0\} = p \qquad (2.4.4)$$

$$\Pr\{x_r < x \leq \infty\} = \Pr\{0 < z \leq \infty\} = 1 - p \qquad (2.4.5)$$

Recall that the mollifier vanishes for $x > |x_r|$, i.e., it is zero over $x_r/\sigma_\varepsilon < z \leq \infty$ in (2.4.5). Hence the relationship in (2.4.5) supports the corresponding probability relationships on the vertical axis:

$$\Pr\{u(x_r+\varepsilon) > u(x_r)\} = \Pr\{u(x_r+\varepsilon) - u(x_r) > 0\} \tag{2.4.6}$$

$$= \Pr\{\Delta_r^{\varepsilon} u(x_r) > 0\} = \Pr\{\alpha_1 Z + \alpha_2 \chi > 0\} \tag{2.4.7}$$

$$= \Pr\{Z > -\frac{\alpha_2}{\alpha_1}\chi\} = 1 - \Pr\{Z \leq -\frac{\alpha_2}{\alpha_1}\chi\} \tag{2.4.8}$$

$$= 1 - \int_{\{0\leq x\leq\infty\}} \Pr\{Z \leq -\frac{\alpha_2}{\alpha_1}x | \chi = x\} f_k(x)dx \tag{2.4.9}$$

$$= 1 - \int_{\{0\leq x\leq\infty\}} \Phi\left(-\frac{\alpha_2}{\alpha_1}x\right) f_k(x)dx \tag{2.4.10}$$

where $f_k(x)$ is the probability density function for a χ^2-distribution with k degrees of freedom. In our case $k = 1$. For the sake of exposition we consider the case of a risk averter so that $\alpha_2 < 0$. In which case $-\frac{\alpha_2}{\alpha_1}x > 0$ and $\Phi\left(-\frac{\alpha_2}{\alpha_1}x\right) > p$. Recall that $x_r \mapsto 0$ in Z-space so the latter inequality is consistent with values to the right of the reference point x_r. After substitution of the Arrow-Pratt risk measure from (2.4.3), the inequality in (2.4.10) becomes

$$\Pr\{\Delta_r^{\varepsilon} u(x_r) > 0\} = 1 - \int_{\{0\leq x\leq\infty\}} \Phi\left(-\frac{\alpha_2}{\alpha_1}x\right) f_k(x)dx \tag{2.4.11}$$

$$1 - \int_{\{0\leq x\leq\infty\}} \Phi\left(\frac{r_A(x_r)}{2}x\right) f_k(x)dx < 1 - p\int_{\{0\leq x\leq\infty\}} f_k(x)dx = 1-p \tag{2.4.12}$$

We are interested in the probabilistic relationship $\Pr\{0 \leq u(x) \leq u(x_r)\}$. This is functionally equivalent to $\Pr\{\Delta_r^{-\varepsilon} u(x_r) \leq 0\} = \Pr\{0 \leq u(x-\varepsilon) \leq u(x_r)\}$ for $0 \leq x \leq x_r$, as ε runs through the interval $[0, x_r]$. By virtue of the symmetry of the mollifier around the reference point x_r, we have from (2.4.12)

$$\Pr\{\Delta_r^{-\varepsilon} u(x_r) \leq 0\} = 1 - \Pr\{\Delta_r^{\varepsilon} u(x_r) > 0\} \tag{2.4.13}$$

$$\Longrightarrow \Pr\{\Delta_r^{-\varepsilon} u(x_r) \leq 0\} > 1 - (1-p) = p \tag{2.4.14}$$

A similar argument to the ones above hold for risk seekers. In that case $\alpha_2 > 0$ and the direction of the inequalities above change. Since x_r was arbitrary, as x_r runs through X, and p runs through

Figure 2.8: Pessimist probability weighting function

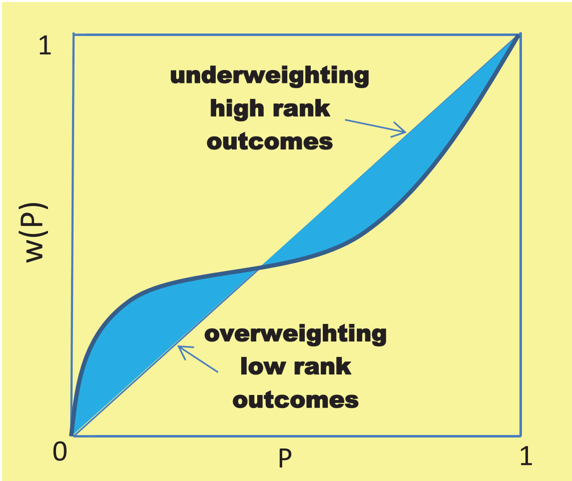


Figure 2.9: Optimist probability weighting function

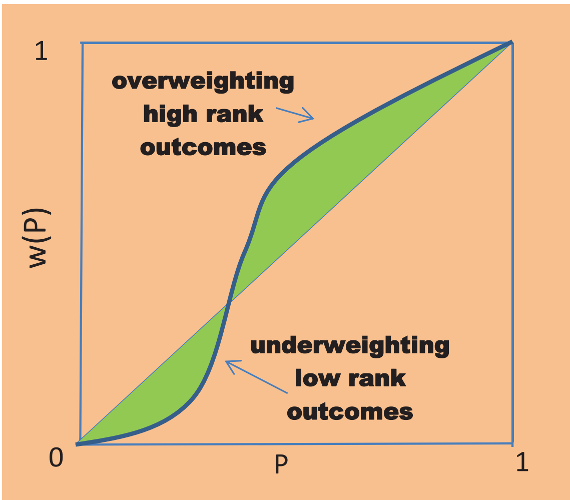


Figure 2.8 is a sketch of the probability weighting function induced by our risk averter in Figure 2.6. Figure 2.9 is a sketch of the probability weighting function induced by our risk seeker in Figure 2.7. Our ranking of outcomes is from worst to best. In Wakker (2010, p. 174) the ranking of outcomes is from best to worse so concavity implies optimism and convexity implies pessimism.

the interval $0,1]$, the corresponding pwf that overweighs in (2.4.12) and underweighs in (2.4.14) is traced.

So according to (2.4.14) low ranked outcomes less that x_r are underweighted and high ranked outcomes above x_r are overweighted. Thus, our claim is proved. Interestingly, the Arrow-Pratt risk measure for outcomes is embedded in the probabilistic statements above, e.g., (2.4.12). Thus establishing a nexus between utility risk aversion and probabilistic risk aversion. So in addition to probability distortion $w(p)$, pwfs are outcome dependent. Thus, we established:

Proposition 2.4.2 (Probability distortion)**.** *Risk averters overweight low ranked outcomes, and underweight high ranked outcomes. Risk seekers underweight low ranked outcomes, and overweight high ranked outcomes. Probability weighting functions are outcome dependent functions of type $w(p,x)$ where x is ranked outcome, and p is the cumulative probability that corresponds to the ranked outcome.* □

The predictions of Proposition 2.4.2 are depicted in Figure 2.8 and Figure 2.9. This probability weighting scenario lies at the heart of rank dependent utility theory (RDU) popularized by Quiggin (1982, 1993). To the best of our knowledge, the proof of Proposition 2.4.2 based on mollifiers centered at a reference point is new to the literature. Armed with Propositions 2.4.1 and 2.4.2 one can see that probability weighting functions are useful devices for analyzing tail risks. Wakker (2010) provides a review of the "likelihood insensitivity" induced by tail risk. That is, outcomes near rank extremities are overweighted or underweighted accordingly, while outcomes in the "middle" ranks are relatively flat compared to the linear probability associated with expected utility.

Explaining economic experiments results on skewness preferences

Kraus and Litsenberger (1976, p. 1086) argued for a class of utility functions that accommodates risk averse investors while at the same time allowing for third moments. Their reason for doing so is based on factors that support concave utility functions. Golec and Tamarkin (1998) proposed a cubic utility model in which they interpreted the coefficients of polynomial terms as risk attitude parameters. They tested that specification with data on horse race track betting and found that bettors were skewness loving but not risk loving.

Figure 2.10: Distribution of US Real Disposable Income Growth

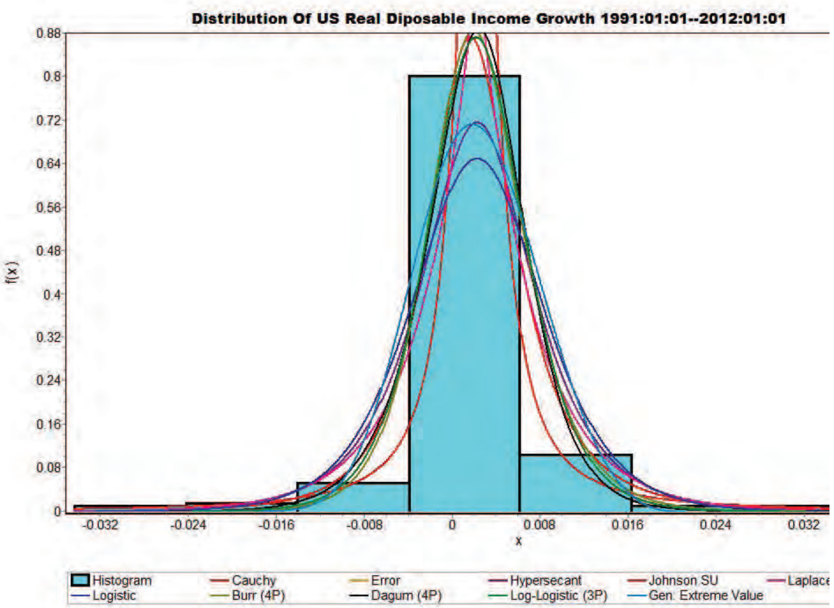


Figure 2.11: Distribution of US Real Disposable Relative Income Growth

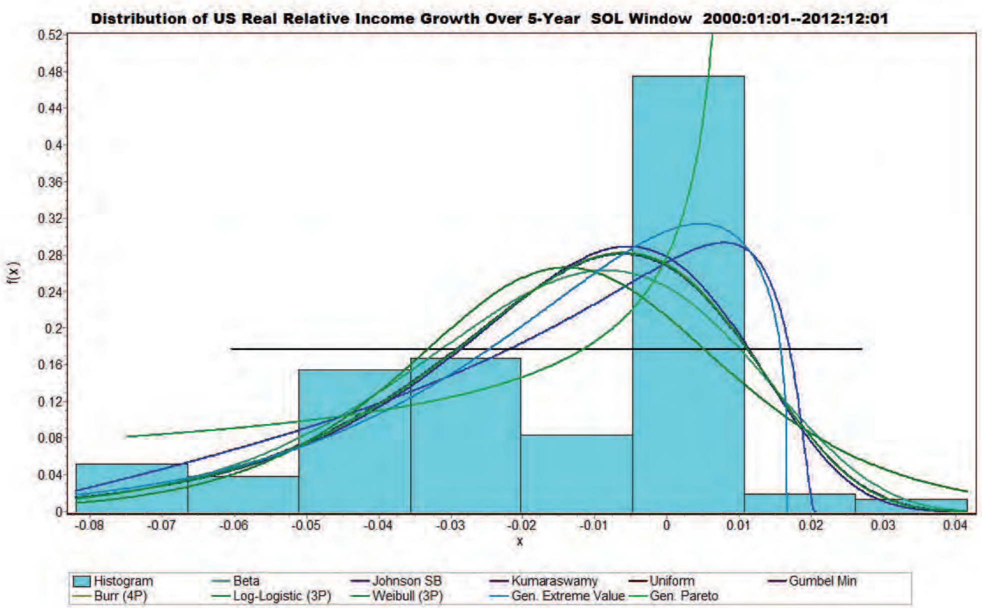


Figure 2.12: Distribution of South Africa Real semidurable PCE Growth

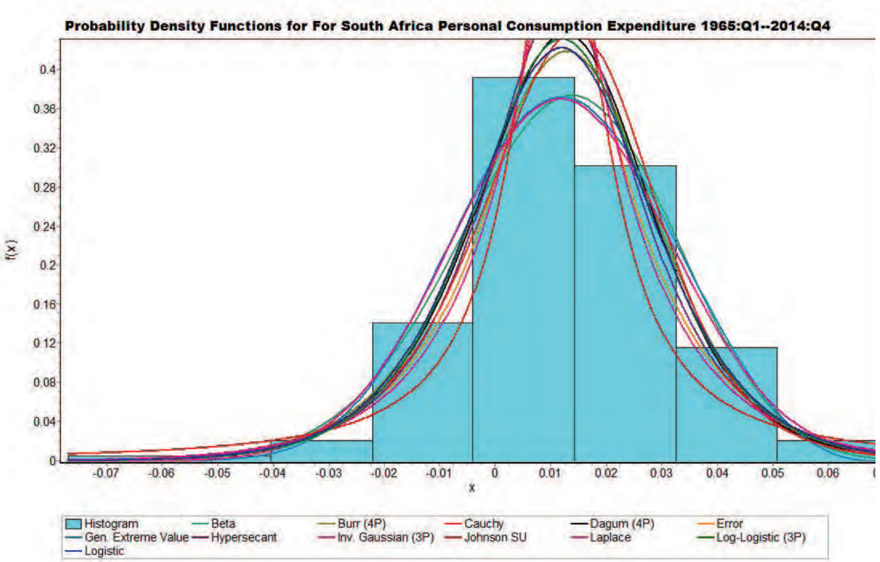
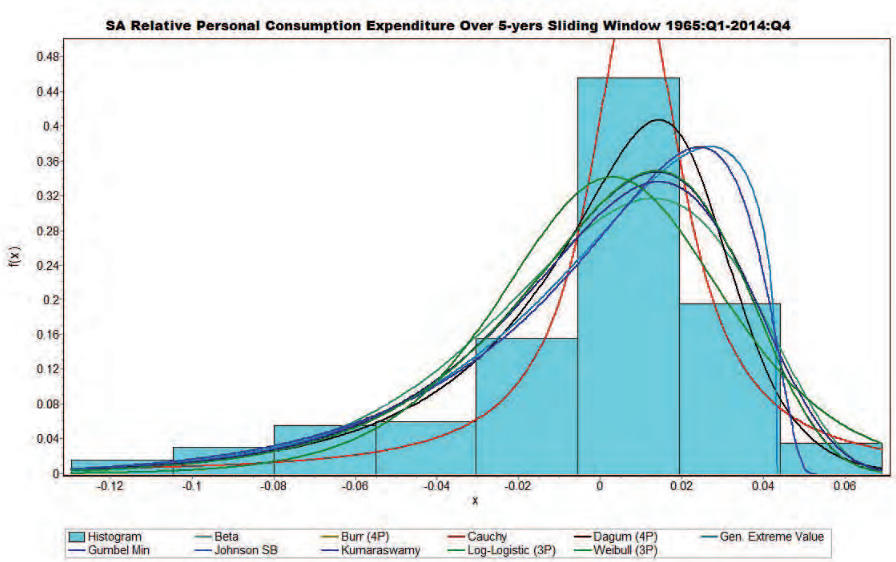


Figure 2.13: Distribution of South Africa semidurable Relative PCE Growth



A recent experimental study by Åstebro et al. (2015) used a structural model, based on constant relative risk aversion (CRRA)[25] and Prelec (1998) model of probability weighting, to test the skewness preference hypothesis. Prelec's two-factor pwf is $w(p) = \exp\left(-\beta(-\ln p)^{\alpha}\right)$ where β controls elevation, i.e., degree of optimism, and α controls curvature, i.e., degree of likelihood insensitivity. For $0 < \beta < 1$ the pwf captures optimism. For $\beta > 1$ pessimism. If $0 < \alpha < 1$, then the pwf is inverse S-shaped like in Figure 2.8. For $\alpha > 1$ the shape is skewed S-shaped like that in Figure 2.9. On the basis of those structures, Åstebro et al. (2015) found that, among other things, subjects made riskier choices when lotteries displayed positive skew. They identified probabilistic risk seeking, i.e., the shape of probability weighting functions is as in Figure 2.9, as the explanatory factor for this behaviour. However, they claimed that their study found no evidence of utility based risk seeking.

According to Proposition 2.4.1 one can introduce skewness and kurtosis in preference based models without resorting to third order expansion by virtue of inclusion of the χ^2-factor. This observation seems to have escaped analysts who use third order expansions. For example, the probability relationships in (2.4.12) show that likelihood insensitivity or tail risk is controlled by the presence of the Arrow-Pratt risk measure $r_A(x_r)$ in the probabilistic statement. That risk measure is negative for risk seekers and positive for risk averters. Thus, Åstebro et al. (2015) misinterpreted their results when they claimed to have found evidence of optimism (from probabilistic risk attitude factors) but not risk loving (from convex utility) among the subjects in their experiments. Figure 2.7 plainly shows that a convex utility function induces an optimistic pwf so the two are substitutes for each other. A finding for one is equivalent to a finding for the other. The more likely explanation is that the Åstebro et al. (2015) econometric specification may have crowded out the convex utility factor. This is all the more likely given their finding that "[t]here are significant rates of decision errors" (p. 203) among subjects in their experiments.

2.4.3 Probabilistic preference for skewness in US and South Africa income growth

This section provides empirical evidence that support our theory of skewed distributions induced by risk attitudes over mollifiers centered at reference points.

[25] Under CRRA $u(x) = x^{1-r_A}/(1-r_A)$ where r_A is Arrow-Pratt risk measure.

Figure 2.10 represents the distribution of US real income growth fitted to a battery of candidate probability density functions. Refer to Appendix 2.D.3 for computer diagnostics on fits. All the fitted distributions are leptokurtic–a characteristic of α-stable distributions, and they are candidate mollifiers. The Cauchy distribution emerges as one of the best fitted distributions. We applied the RIH procedure inhered in Theorem 2.3.2 to generate a distribution of income growth relative to a 5-years standard of living. Thus, M_t is a *de facto* reference point. We fitted a battery of probability distributions to the transformed distribution as shown in Figure 2.11. Refer to Appendix 2.D.4 for computer diagnostics on fits. Cursory inspection shows that the geometry of Figure 2.6 and predictions of Proposition 2.4.1 are bourne out. Most important, the α-stable prediction in Theorem 2.2.6 for the MLA index also has currency.

Figure 2.12 depicts the probability distribution functions fitted to South Africa PCE growth data. Again, the Cauchy distribution emerges as a fitted probability distribution function. Consistent with Figure 2.6 and Propositon 2.4.1 the underlying probability distribution is distorted by risk attitudes about outcomes.

It should be noted that the analyses above apply equally to Figure 2.7. In Appendix 2.C.3 and Appendix 2.C.4, Figure 2.38 and Figure 2.39 depict a "reflection effect" (Kahneman and Tversky, 1979, p. 268) for growth rates. That is, the signs of the growth rates are changed. Substitution of the reflection growth rates would reverse the direction of the skew in the analysis above to characterize risk seeking behaviour depicted in Figure 2.7.

2.5 Application to income and consumption growth

In this section we provide a simple apparatus for estimating the MLA index with time series data for income and consumption. We derive a "marked MLA index process"[26] and introduce some new results and econometric tests for MLA index behaviour.

Data on real income and consumption growth are characterized by runs. For example, in

[26]Data from a marked point process are represented by $(t_j, X(t_j))$ where the time-point t_j of observing the phenomenon is generated by a process independent of the process generating X. The idea behind the term "marked MLA index process" is that some independent event in the economy causes a run of losses and gains in income, relative to a reference income, from which an MLA index can be computed. This event is a "mark" that can be modelled as a random variable attached to t_j. Refer to Cox and Isham (1980, p. 131) for illustrate examples of marked point process. Marked point processes are outside the scope of this paper.

periods of increasing income we can expect to see runs of positive relative income growth g_t^G. For falling incomes, we would see runs of negative income growth g_t^L. These runs are manifestations of information based economic activity. Assume that there are a total of K non-zero runs each of length n_k generated by an arbitrary time series of length T, where $k = 1, \dots, K$ and $T = n_1 + \cdots + n_K$. According to (2.3.6) the loss aversion index is characterized by a gain-loss pair (g_t^G, g_t^L). De Neve et al. (2015, p. 19) used the phrase "macroeconomic loss aversion" to recommend that "future research should consider positive and negative economic growth rates separately in piecewise analyses in order to more accurately interpret the gradient for the general relationship between economic growth rates and subjective well-being." That statement is functionally equivalent to Theorem 2.2.2 which contemplates the estimator $\lambda = \bar{g}^L/\bar{g}^G$. Thus, we propose the following estimation strategy for the "macroeconomic loss aversion" index in our study.

2.5.1 An empirical strategy for constructing the MLA index process

We illustrate the procedure for implementing Theorem 2.2.2 in discrete time with the following example. Suppose that $\underbrace{g_{-3}^L, g_{-2}^L, g_{-1}^L,}_{\text{block of losses}} \underbrace{0}_{\text{local reference point}}, \underbrace{g_1^G, g_2^G}_{\text{block of gains}}$ is an observed sequence of losses $(g^L < 0)$, and gains $(g^G > 0)$ adjacent to, and separated from, each other by the local reference point 0, i.e., $g_{-1}^L < 0 < g_1^G$. The negative subscript is used to highlight the fact that the runs are to the left of 0. Each run of gains or losses constitutes a block. Here $\bar{g}^L = (g_{-3}^L + g_{-2}^L + g_{-1}^L)/3$ is the average over the run (or block) of losses; and $\bar{g}^G = (g_1^G + g_2^G)/2$ the average over the run (or block) of gains. According to Theorem 2.2.2, the MLA index estimator in correspondence with the joint block of gains and losses is $\widehat{\lambda} = -\bar{g}^L/\bar{g}^G$. We reiterate with a numerical example. Consider the sequence of losses and gains: $-0.04, -0.02, -0.03, 0, 0.01, 0.02$. Here $\bar{g}^L = -(0.04 + 0.02 + 0.02)/3 = -0.03$ and $\bar{g}^G = (0.01 + 0.02)/2 = 0.015$. Hence $\widehat{\lambda} = -(-0.03/0.015) = 2.0$. We abstract from this procedure below.

Axiom 2 (Blocks of growth)**.** *Every economic time series generates a block or run of negative (loss) growth rates followed by a block or run of positive (gain) growth rates.*

Let $g_{t_i^{2k-1}}^L$ be the observation at time t_i included in the $2k-1$ "marked block" of losses. For example, the block is marked by some random event in the economy which generated the run of losses. For notational convenience we write t_i^{2k-1} instead of $-t_i^{2k-1}$. It being understood that the observed block of negative growth ($g^L < 0$) is to the left of the local reference point 0. So $g^L_{t_i^{2k-1}}$ is a random variable. It is the "marked observation" at time t_i^{2k-1} where the latter should read time t_i in the "marked block" $2k-1$ for $i = 1,2,\ldots,n_{2k-1}$. Let $g^G_{t_j^{2k}}$ be the observation at time t_j included in the adjacent $2k$ block (separated from the $2k-1$ block by local reference point 0) of gains where $j = 1,2,\ldots,n_{2k}$. Here, the observed marked block of positive growth ($g^G > 0$) are to the right of the local reference point 0. Similarly, $g^G_{t_j^{2k}}$ is a random variable. We use even and odd subscripts for blocks to emphasize their non-overlapping feature.

Definition 2.5.1 (Topology of gains and losses)**.** Let $\mathscr{B}^L_{2k-1}$ and $\mathscr{B}^G_{2k}$ be adjacent open sets or marked blocks of growth rates for losses and gains, respectively, in a univariate time series of T growth rates. Let $\mathscr{N} = \{n_1,\ldots,n_{2K-1},n_{2K}\}$ be the set of K block lengths, and $T = n_1 + \cdots + n_{2K-1} + n_{2K}$. So $n_{2k-1} \in \mathscr{N}$ is the number of observations in $\mathscr{B}^L_{2k-1}$, and $n_{2k} \in \mathscr{N}$ is the number of observations in $\mathscr{B}^G_{2k}$. Thus, we have

$$\mathscr{B}^L_{2k-1} = \{\underbrace{g^L_{t_1^{2k-1}},\ldots,g^L_{t_{n_{2k-1}}^{2k-1}}}_{\text{block of losses}}\}, \quad \mathscr{B}^G_{2k} = \{\underbrace{g^G_{t_1^{2k}},\ldots,g^G_{t_{n_{2k}}^{2k}}}_{\text{block of gains}}\} \tag{2.5.1}$$

$$\mathscr{B}_{2k-1,2k} = \underbrace{\left\{\mathscr{B}^L_{2k-1} \oplus \{0\} \oplus \mathscr{B}^G_{2k}\right\}}_{\text{MLA index generator}}, \quad \mathscr{B}_T = \bigoplus\nolimits_k \mathscr{B}_{2k-1,2k} \tag{2.5.2}$$

$$\mathscr{B}^0 = \{g = 0 | g \in \mathscr{B}_T\} \subseteq \bigcap_{k=1}^{K} \mathscr{B}_{2k-1,2k} \tag{2.5.3}$$

where $\oplus$ is a concatenation operation defined so that

$$\mathscr{B}_{2k-1,2k} = \mathscr{B}^L_{2k-1} \oplus \{0\} \oplus \mathscr{B}^G_{2k} = \left\{g^L_{t_1^{2k-1}},\ldots,g^L_{t_{n_{2k-1}}^{2k-1}},0,g^G_{t_1^{2k}},\ldots,g^G_{t_{n_{2k}}^{2k}}\right\}$$

is the joint distribution of growth rates that generate the $(k-1)$th MLA index. And $\bigoplus$ is the global concatenation operation that reconstructs the joint distribution of growth rates for the entire time series $\mathscr{B}_T$ in (2.5.2). Here $\mathscr{B}^0$ is the (closed) zero set comprised of the local reference point 0 common to all blocks. □

Lemma 2.5.1 (MLA index operation)**.** *There exist an operator $\mathfrak{L}_{k-1}$ such that $\mathfrak{L}_{k-1} : \mathscr{B}_{2k-1,2k} \longrightarrow$*

λ_{k-1}. $\square$

Remark 2.5.1. After dropping the subscript $k-1$, the "existence" operation $\mathfrak{L}$ is decomposed into the operations $\mathfrak{L} = \mathfrak{L}^{div} \star \mathfrak{L}^{add}$. In matrix operator terms $\mathfrak{L}^{div}$ is a 1×2 matrix, and $\mathfrak{L}^{add}$ is a $2 \times n$ matrix where the elements of each matrix are suitably defined for a $n \times 1$ vector $\boldsymbol{v}$ comprised of gains and losses such that $\boldsymbol{v} \in \mathscr{B}_{2k-1,2k}$. For example, one row of $\mathfrak{L}^{add}$ contains 1's that pick up gains, and the other row contains -1 that pick up losses, with zeros used as filler for non-gains and non-losses in each row. Thus, if $\boldsymbol{v} = [g_1, g_2, g_3, -g_4, -g_5]^T$ is 5×1 then the 2×5 matrix for addition is $\mathfrak{L}^{add} = \left[\begin{smallmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 \end{smallmatrix}\right]$. The elements of the $\mathfrak{L}^{div} = \left[\begin{smallmatrix} 0, & \frac{3}{2}\sum_{i=1}^{3} g_i \end{smallmatrix}\right]$ matrix are suitably defined to compute the MLA index from the column(s) of $\mathfrak{L}^{add}\boldsymbol{v}$. So $\widehat{\lambda} = \mathfrak{L}\boldsymbol{v} = (\sum_{i=1}^{2} g_i/2)/(\sum_{i=1}^{3} g_i/3)$ $\square$

Implicit in Definition 2.5.1 is the presumably negligible event

$$\widetilde{\mathscr{B}}^0 = \{g^L_{t_i^{2k-1}} = 0 \text{ or } g^G_{t_j^{2k}} = 0\}, \; i = 1, 2, \ldots, n_{2k-1}, \; j = 1, 2, \ldots, n_{2k} \tag{2.5.4}$$

Here, $\widetilde{\mathscr{B}}^0$ contains the event of zero growth included in a block of losses or gains where such zero growth *is not a reference point*. That is, it does not separate gains and losses but is contained *within* a gain or loss block accordingly. The foregoing gain-loss topology set the stage for the following

Theorem 2.5.2 (Empirical MLA index)**.** *Let $\mathscr{B}_T$ be an observed sequence of growth rates (g) for income. Define the arithmetic mean growth for a block of gains (G) and losses (L) as follows:*

$$\bar{g}^L_{n_{2k-1}} = \frac{\sum_{i=1}^{n_{2k-1}} g^L_{t_i^{2k-1}}}{n_{2k-1}}, \quad g^L_{t_i^{2k-1}} \in \mathscr{B}^L_{2k-1} \tag{2.5.5}$$

$$\bar{g}^G_{n_{2k}} = \frac{\sum_{j=1}^{n_{2k}} g^G_{t_j^k}}{n_{2k}}, \qquad g^G_{t_j^{2k}} \in \mathscr{B}^G_{2k} \tag{2.5.6}$$

The empirical myopic loss aversion index estimator generated by the operation $\mathfrak{L}_{k-1}$ on $\mathscr{B}_{2k-1,2k}$, i.e., $\mathfrak{L}_{k-1} : \mathscr{B}_{2k-1,2k} \to \lambda_{k-1}$, is given by:

$$\widehat{\lambda}_{k-1} = \frac{\bar{g}^L_{n_{2k-1}}}{\bar{g}^G_{n_{2k}}}, \quad k = 1, \ldots, K, \; \lambda_0 = 0 \tag{2.5.7}$$

In particular, the sequence of ordered pairs $(\bar{g}^L_{n_1}, \bar{g}^G_{n_2}), \ldots, (\bar{g}^L_{n_{2k-1}}, \bar{g}^G_{n_{2k}})$ generates the distribution

$\{\widehat{\lambda}_1,\ldots,\widehat{\lambda}_{K-1}\}$ *of MLA index estimates from the operations* $\mathfrak{L}_1 : \mathscr{B}_{1,2} \to \lambda_1$,..., $\mathfrak{L}_{k-1} : \mathscr{B}_{2k-1,2k} \to \lambda_{k-1}$. ☐

Remark 2.5.2. In this setup, the sequence always starts with a block of losses followed by a block of gains–the two blocks being separated by 0 and so on. One weakness of the procedure is that it is sensitive to the starting block. So if we start with a block of gains followed by a block of losses, we will generate different MLA index estimates compared to if we start with a block of loss followed by a block of gains. So the estimates are not "ordinal independent".

By virtue of the above, we claim that $\widehat{\lambda}_{k-1}$ is iid.

Proposition 2.5.3 (MLA index independence)**.** *The distribution* $\{\widehat{\lambda}_1,\ldots,\widehat{\lambda}_{K-1}\}$ *of MLA index estimates, generated by the operations* $\mathfrak{L}_1 : \mathscr{B}_{1,2} \to \lambda_1$,..., $\mathfrak{L}_{k-1} : \mathscr{B}_{2k-1,2k} \to \lambda_{k-1}$, *is independent and identically distributed.*

Proof. See Appendix 2.B.5. ☐

Functional implications of lognormally distributed income

In keeping with the literature on income and consumption, we assume that income is lognormally distributed (see Battistin et al. (2009) for a recent review of the literature). Thus, we assume $Y \sim \ln\mathcal{N}(\mu_y,\sigma_y^2)$ with probability density function $f(y) = \left(y\sigma\sqrt{2\pi}\right)^{-1}\exp\left(-\{\frac{\ln y-\mu}{2\sigma_y^2}\}\right)$. See e.g., Borowiak and Shapiro (2014, p. 26) for details. Thus, if $Y = 1$ at the start of a given period, and g is income growth over the period, then at the end of the period $Y = 1+g$ and $\ln Y = \ln(1+g) \approx g \sim \mathcal{N}(\mu_y,\sigma_y^2)$. Accordingly, we make the following

Assumption 2.5.4 (Lognormal income)**.** *Income* Y_t *is lognormally distributed, and income growth* g_Y *is normally distributed.* ☐

It is known that if the numerator and denominator in (2.5.7) are normally distributed then $\widehat{\lambda}_k$ has a standard Cauchy distribution characterized by probability density function $f(\lambda) = [\pi(1+(\lambda-\mu)^2)]^{-1}$, where μ is a measure of location, and the sample median $\widehat{\pi}_{.5} \sim AN\left(\pi_{.5},\frac{\pi^2}{4n}\right)$ is an asymptotically normal consistent estimator for sample size n (Serfling, 1980, p. 85). However, the numerator and denominator are drawn from a truncated normal. Vizly, gains from the right half and losses from the left half. By symmetry, gains and losses are drawn from the same law (Çinlar, 2011, p. 331). So the ratio is actually a truncated or generalized Cauchy distribution which lies entirely in the positive quadrant. Thus, our estimator is consistent with the predictions of Theorem 2.2.6. We summarize this formally.

Proposition 2.5.5 (Statistical test for MLA index)**.** *Let $Y \in \mathscr{B}_T$ be a lognormally distributed realization of income with growth rate $g_Y \sim \mathscr{N}(\mu_y, \sigma_y^2)$ in the reference point basis set $B_\varepsilon(\mu_y)$ under assumption 2.2.1. Let μ_y be a reference point, and define $g^L = \mathbb{I}_{\{g_Y < \mu_y\}}$ and $g^G = \mathbb{I}_{\{g_Y > \mu_y\}}$. We claim that $\widehat{\lambda} \sim \mathscr{C}(a, b)$ with sample median $\widehat{\pi}_{.5} \sim AN\left(\pi_{.5}, \dfrac{\pi^2 b^2}{4n}\right)$.*

Proof. See Theorem 2.2.6 and Walck (2007, Ch. 7, p. 31). □

Remark 2.5.3. In practice b is replaced by a maximum likelihood estimate (MLE) $\overset{\triangle}{b}$. For robust tests $b = \frac{1}{2}(Q_3 - Q_1)$ where $Q_3 - Q_1$ is the interquartile range for $\widehat{\lambda}$. We set $b = 1$ for standard Cauchy $\widehat{\lambda} \sim \mathscr{C}(0, 1)$ so $\widehat{\pi}_{.5} \sim AN\left(\pi_{.5}, \dfrac{\pi^2}{4n}\right)$. □

Motivated by Theorem 2.2.6 we introduce the following processes.

Definition 2.5.2 (Subordinator)**.** (Çinlar, 2011, p. 279) Let $S = (S_t)_{t \in \mathbb{R}^+}$ be an increasing right-continuous stochastic process with state space $\mathbb{R}^+$ and $S_0 = 0$. It is said to be an increasing Lévy process (or subordinator) if

(a) the increments $S_{t_1} - S_{t_0}$, $S_{t_2} - S_{t_1}, \ldots, S_{t_n} - S_{t_{n-1}}$ are independent for $n \geq 2$ and $0 \leq t_0 < t_1 < \cdots < t_n$, and

(b) the distribution of the increment $S_{t+u} - S_t$ is the same as that of S_u for every t and u in $\mathbb{R}^+$.

The property (a) is called the independence of increments, and (b) the stationarity of the increments. □

Definition 2.5.3 (Cauchy process)**.** (Çinlar, 2011). Let $(\Omega, \mathscr{F}, P)$ be a probability space where Ω is a sample space, P is a probability measure on Ω, and $\mathscr{F}$ is the σ-field of Borel measurable subsets of Ω. Let T_a be the first crossing time of level a for the running maximum of a Brownian motion B_t, i.e., $M_t(\omega) = \max_{0\leq s\leq t} B_s(\omega)$ for some $\omega \in \Omega$. Thus, we have

$$T_a(\omega) = \inf\{t > 0;\, M_t(\omega) > a\} \tag{2.5.8}$$

$$M_t(\omega) = \inf\{a > 0;\, T_a(\omega) > t\} \tag{2.5.9}$$

If $W_t(\omega)$ is a Brownian motion independent of $B_t(\omega)$, then $C_t^u(\omega) = W_{T_a}(\omega)$ is a Cauchy process that depends on the subordinator T_a. □

The next proposition follows readily from Theorem 2.2.6, Proposition 2.5.3, and Proposition 2.5.5

Proposition 2.5.6 (MLA index as Cauchy r.v)**.** *The MLA index estimator* $\widehat{\lambda}$ *admits a Cauchy random variable (r.v.).*

Proof. See Appendix 2.B.6. □

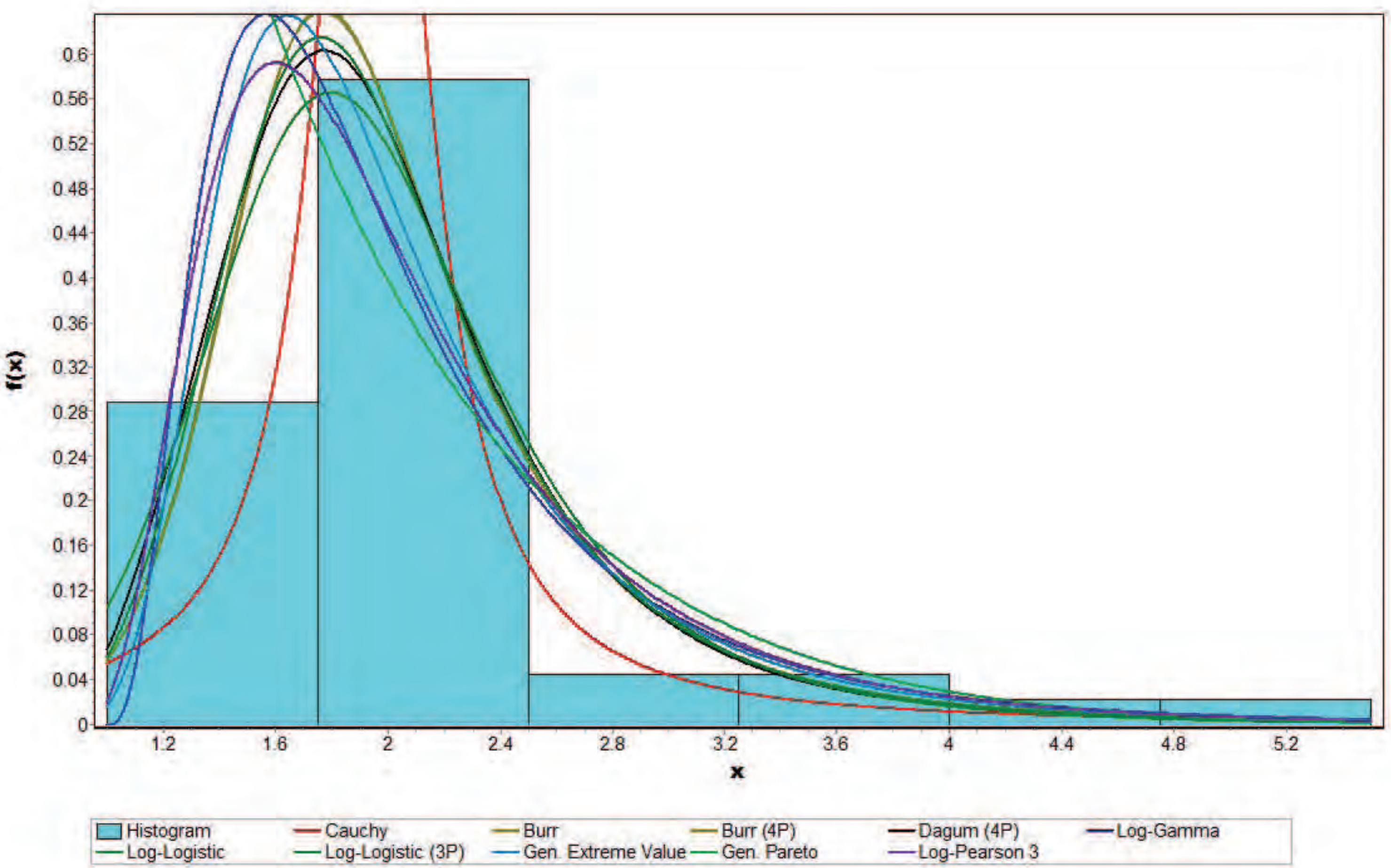
Assumption 2.5.4 implies that Y_t admits a Geometric Brownian motion (GBM). Thus, the Cauchy process induced by lognormally distributed income can be modelled as a subordinate or time changed GBM $C_t^u(\omega) = W_{T_a}^Y$ for some GBM $W_t^Y(\omega)$. Refer to Karatzas and Shreve (1991, p. 174) for details on time changed Brownian motion. Hence, Axiom 1 implies that intolerance to decline in standard of living can be modelled as the running maximum of a Cauchy process. We state that formally as

Theorem 2.5.7 (Intolerance to decline in standard of living)**.** *If* $C_t^Y(\omega)$ *is the Cauchy process induced by myopic loss aversion, then the subordinate process* $C_t^{Y\star}(\omega) = \max_{0<s<t} C_t^Y(\omega)$ *mimics a decision maker's intolerance to decline in standard of living.* □

2.5.2 Statistical tests of MLA index estimator theory

In this subsection we provide the results of statistical tests of our α-stable theory of the MLA index. The tests are applied to published data from surveys, a meta study, and our own estimates from applying Theorem 2.5.2 to income and consumption data from US and South Africa.

Figure 2.14: Probability distribution of MLA index around the world



The top of the Cauchy distribution was cut-off in the computer generated plot because it was so much higher that for other candidate distribution in the α-stable class.

Table 2.1: Diagnostics for MLA index around the world

Statistic	Value	Percentile	Value
Sample Size	45	Min	1
Range	4.5	5%	1.132
Mean	2.0731	10%	1.318
Variance	0.63346	25% (Q1)	1.635
Std. Deviation	0.7959	50% (Median)	2
Coef. of Variation	0.38392	75% (Q3)	2.06
Std. Error	0.11865	90%	3.132
Skewness	2.2889	95%	3.832
Excess Kurtosis	7.3994	Max	5.5

Fitting the distribution of MLA indexes around the world

Summary statistics for α-stability of MLA indexes around the world

To illustrate the robustness of our theory, we analyze the MLA index estimates for loss aversion around the world published in Rieger et al. (2011, Table 2, p. 7). The data is reproduced in Appendix 2.C.1. Those numbers were generated from hypothetical choices in a survey instrument administered to mostly university students in forty-five different countries. All the index values are greater than or equal to 1.0 with a maximum of 5.5 reported for Georgia.

The descriptive statistics in Table 2.1 indicate that the distribution of MLA index estimates is skewed (skewness coeff.=2.2889) and leptokurtic (excess kurtosis = 7.3994). Figure 2.14 shows the ten best distributions fitted to the world MLA index data. Refer to Appendix 2.D.6 for details on distribution ranks and fits. All the distributions are members of the α-stable class popularized in the actuarial science and economics literatures (Samoradnitsky and Taqqu, 1994; Kleiber and Kotz, 2003). In fact, the generalized Cauchy distribution with MLE scale $\overset{\triangle}{\sigma} = 0.21558$, and MLE location $\overset{\triangle}{\mu} = 1.9422$), is ranked as the best fitting distribution by the Kolmogorov-Smirnov and Chi-squared goodness of fit tests, and ranked as the fourth best fitted distribution by the Anderson-Darling[27] goodness of fit test.

Fitting the Fishburn-Kochenberger MLA index metatstudy data

[27] This is a nonparametric goodness of fit test that is sensitive to tail behaviour in distributions. Refer to Anderson and Darling (1954).

Kahneman and Tversky (1979, p. 280) referenced the Fishburn and Kochenberger (1979) metastudy as one of many pieces of evidence which supports their loss aversion theory. The MLA index estimates in that metastudy were generated from data points collected from eyeballing plots in published papers. Thus, the data is quite noisy. Nonetheless, they employed a procedure, related to Theorem 2.2.2 in this paper, to estimate local piecewise linear ("two-piece") utility functions. The majority of functions were concave over gains and convex over losses. The Fishburn-Kochenberger procedure provides contrast to the distribution of MLA indexes obtained from hypothetical choice data in survey instruments in the Rieger et al. (2011) study described above.

Table 2.2: Diagnostics for Fishurn-Kochenberger
MLA index metastudy data

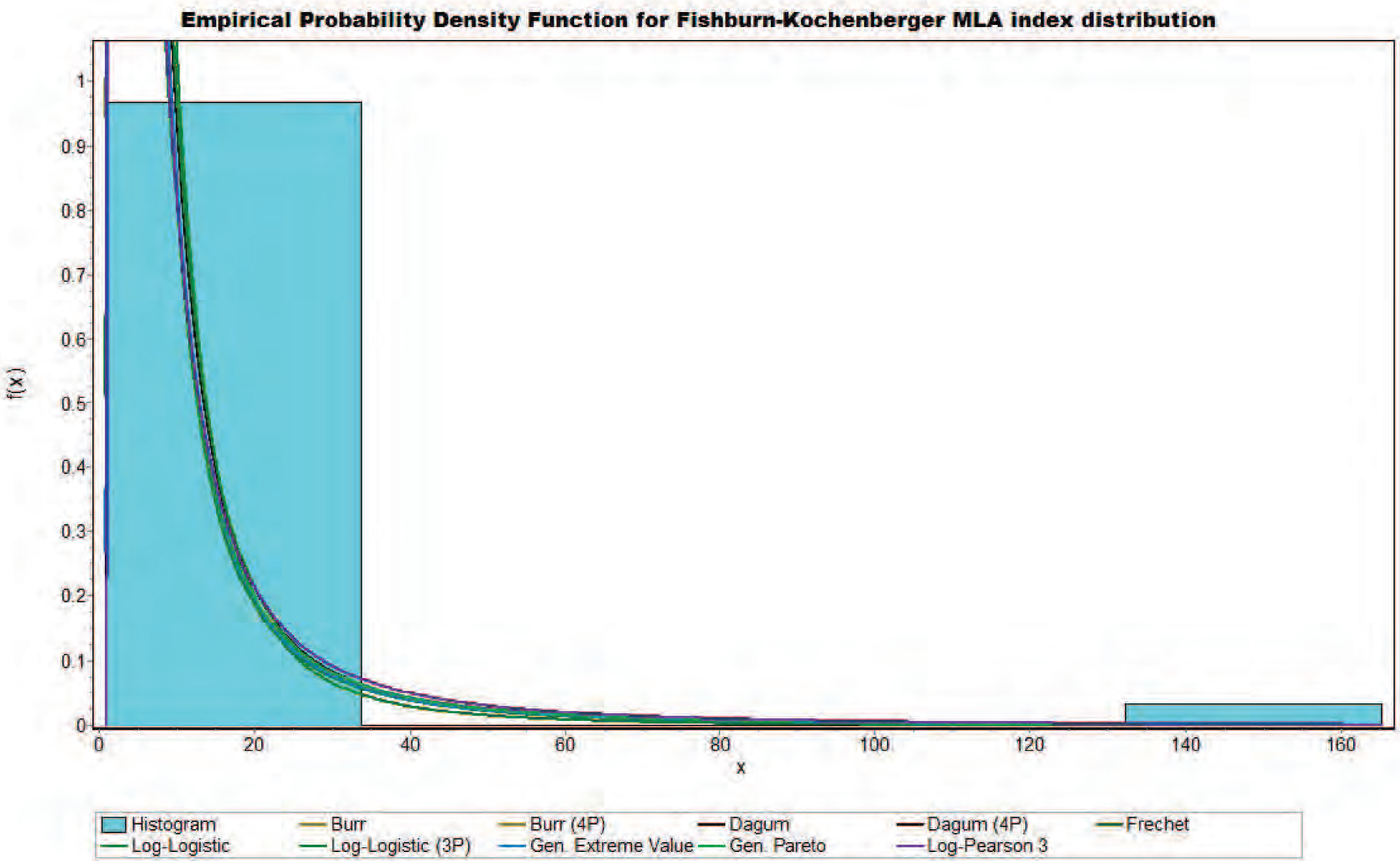
Statistic	Value	Percentile	Value
Sample Size	30	Min	0.8
Range	164.4	5%	1.295
Mean	12.34	10%	1.83
Variance	876.65	25% (Q1)	2.825
Std. Deviation	29.608	50% (Median)	4.85
Coef. of Variation	2.3994	75% (Q3)	7.725
Std. Error	5.4057	90%	22.96
Skewness	5.0684	95%	87.925
Excess Kurtosis	26.809	Max	165.2

Summary statistics for α-stability of MLA index metastudy data

The descriptive statistics in Table 5.3 are for the loss aversion index estimates reported in Fishburn and Kochenberger (1979, Tales 1A, 1B, pp. 508-509) for their two-piece linear $(L^- L^+)$ local utility function. Conceptually, the "two-piece $(L^- L^+)$" is functionally equivalent to the piecewise linearized utility function depicted in Figure 2.1. The statistics show that the index is right skewed and extremely leptokurtic. The Fishburn and Kochenberger (1979) data were best fitted to a Log Pearson Type III distribution ($p = 0.32514$ for Anderson-Darling goodness of fit statistic). However, the data admits a generalized Cauchy distribution ($\overset{\triangle}{\sigma} = 1.9938$, $\overset{\triangle}{\mu} = 4.1153$) which was not rejected at the $p = 0.20$ level for the Kolmogorov-Smirnov and Chi-squared test statistic. Refer to Appendix 2.D.5for details on function ranks and fits.[28]

[28] Fishburn and Kochenberger (1979) also reported extreme values for the loss aversion index, i.e., $\lambda = 3300$, $\lambda = \infty$ for a two-piece exponential $(E^- E^+)$ local utility function. However, that functional form is not part of the analysis depicted in Figure 2.1.

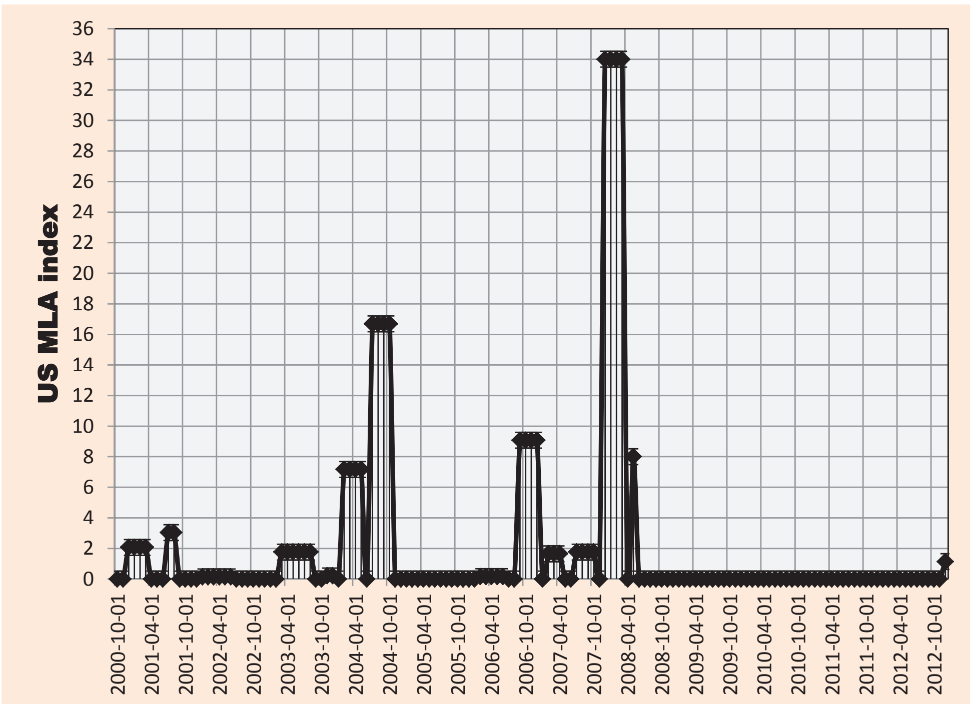
Figure 2.15: Fitted distributions for MLA index metastudy



US MLA index for intolerance to decline of standard of living

Figure 2.16 depicts a plot of the empirical MLA index for US monthly real income computed according to Theorem 2.5.2. The plot is based on a "marked time series" constructed by replacing the elements of each $\mathscr{B}^{L}_{2k-1}$ set with the loss aversion index $\widehat{\lambda}_{k-1}$ generated from $\mathscr{B}_{2k-1,2k}$. The observations in the $\mathscr{B}^{G}_{2k}$ component of that block were "zeroed out". In that way the MLA index is restricted to loss domain. That procedure gave the plot a tableau look over losses. The explosive values for the index, in or around 2004, coincides with Hurricanes Charlie and Ivan which resulted in many deaths. Hurricane Katrina was even more devastating in 2005. In 2008, the Great Recession 2008 with prolonged effect.

Figure 2.16: Pseudo time series plot of US MLA index



The MLA index in (2.3.5) is estimated over monthly real disposable income in US by dividing the average relative growth rate for loss of income by the average relative growth rate for gains in income for 60-months sliding windows between 2000:10 and 2012:11. The loss aversion index is time and state dependent. Its value is backfilled and displayed in the plot over periods with loss of income.

Summary statistics for α-stability of US MLA index

Table 2.3 provides descriptive statistics for the US MLA index estimated in accord with Theorem 2.5.2. The excess kurtosis of 6.3515 confirms that the MLA index is leptokurtic. Even though the median MLA index value of 1.9163 is consistent with that reported in the behavioural and experimental economics literature, the mean is over 3 times as large and the variance is quite high. These are characteristics of an α-stable distribution.

Table 2.3: Diagnostics for US MLA index

Statistic	Value	Percentile	Value
Sample Size	14	Min	0.13896
Range	33.862	5%	0.13896
Mean	6.2046	10%	0.14787
Variance	85.959	25% (Q1)	0.90275
Std. Deviation	9.2714	50% (Median)	1.9163
Coef. of Variation	1.4943	75% (Q3)	8.2694
Std. Error	2.4779	90%	25.349
Skewness	2.4209	95%	34.001
Excess Kurtosis	6.3515	Max	34.001

Figure 2.17 is an empirical plot of the US MLA index series adjusted for the median MLA index value of 2.25. So that $f(\widehat{\lambda}) = 1/[\pi(1+(\widehat{\lambda}-2.25)^2)]$. The plot provides a visual image for the descriptive statistics in Table 2.3. However, the MLE estimates for a generalized Cauchy distribution returned $\overset{\triangle}{\sigma} = 1.4716$ for the scale parameter, and $\overset{\triangle}{\mu} = 1.6182$ for measure of location. Three popular goodness of fit measures: Kolmogorov-Smirnov, Anderson-Darling, and Chi-squared tests, upheld the generalized Cauchy fit at the $p = 0.05$ level.

Statistical tests for US MLA index independence

Proposition 2.5.3 posits that the MLA index is independent. Consequently, we should not find autocorrelation between them. To test this hypothesis we ran a simple autoregression which produced

$$\widehat{\lambda}_{t_{2k-1,2k}} = \underset{(p=0.048)}{1.76} - \underset{(p=0.438)}{0.245}\ \widehat{\lambda}_{t_{2k-3,2k-2}} \tag{2.5.10}$$

where the subscripts for $\widehat{\lambda}$ emphasize that MLA indexes were estimated in accord with the set $\mathscr{B}_{2k-1,2k}$, $k = 1, \ldots, K$. The intercept term is statistically significant at $p = 0.05$ but the coefficient on the autoregressive term is not statistically different from 0 at the $p = 0.438$ level, i.e., we could not reject the null hypothesis $H_0 : \beta = 0$. Hence, the proposition is upheld at the $p = .05$ level. Figure 2.40 in Appendix 2.C.5 provides a plot of the relationship.

Figure 2.17: Empirical distribution of US MLA index

$$f(\widehat{\lambda}) = \frac{1}{\pi[1+(\widehat{\lambda}-2.25)^2]}$$

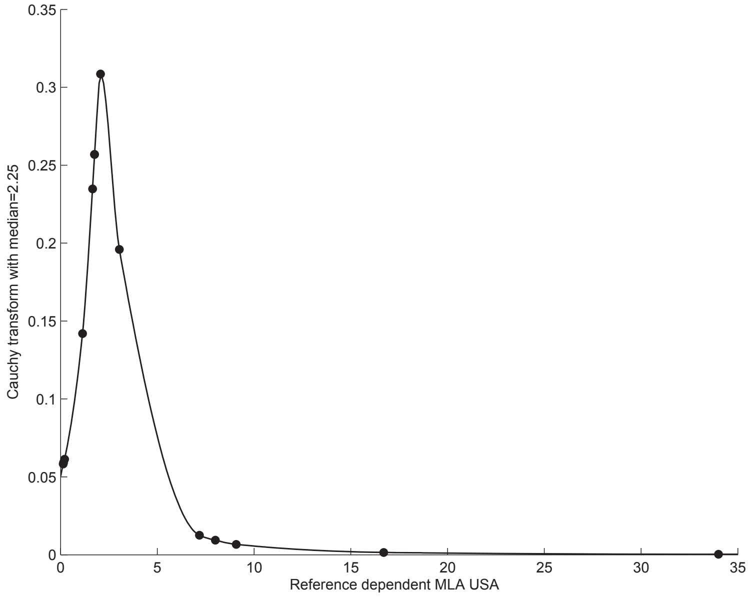


Table 2.4 provides further diagnostic tests for the MLA index. We compared test results for the median-value hypothesis $H_0 : \lambda = 2.25$ vs. $H_a : H_0$ not true, for three variations of the asymptotic distribution for the sample median of a Cauchy distribution: standard Cauchy, robust Cauchy, and generalized Cauchy. Keeping in mind that the tests statistics contemplate large samples while our sample size is small at $n = 14$. At the $p = 0.01$ level the standard Cauchy failed to reject H_0 in five out of fourteen or 26% of the MLE index estimates. In contrast, the robust Cauchy

Table 2.4: Cauchy test for US MLA index
$H_0 : \lambda = 2.25 \ \ vs. \ \ H_a : \lambda \neq 2.25$

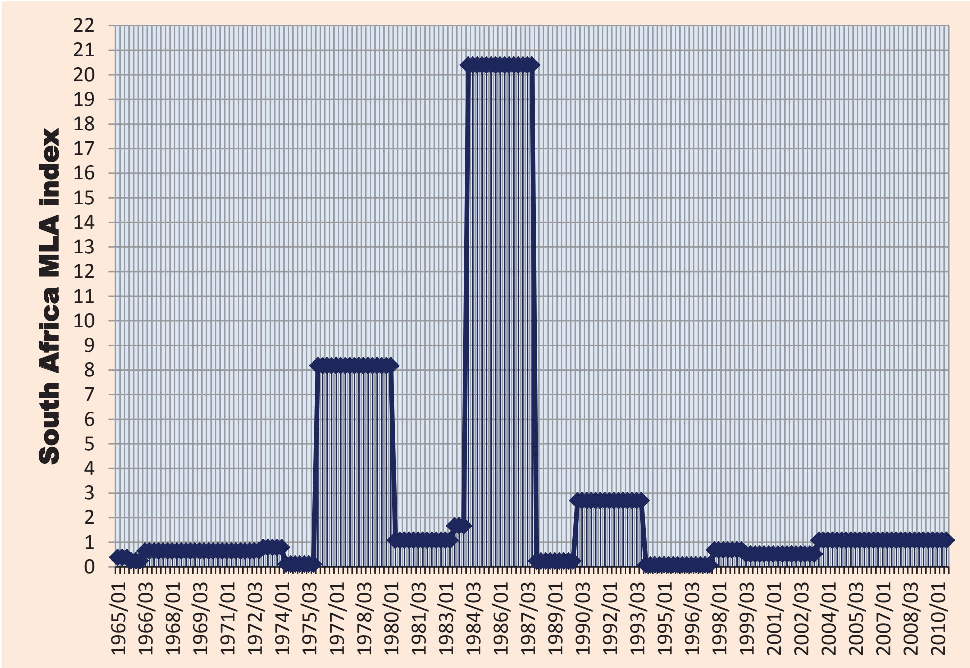
Loss Aversion Index US	Z-score[a]	P-value	Z-score[b]	P-value	Z-score[c]	P-value
2.071649094	-0.424834189	0.335478776***	-0.130139867	0.448227885***	-0.121195234	0.4517682***
3.040267584	1.882427716	0.029888983***	0.576645896	0.282089344***	0.537012492	0.295629493***
0.138960041	-5.028524782	2.47134E-07	-1.540392842	0.061732313***	-1.434520222	0.075711931***
1.760777186	-1.165335138	0.121941681***	-0.356978236	0.36055405***	-0.332442793	0.369777465***
0.203805358	-4.87406239	5.46632E-07	-1.493076229	0.067708617***	-1.390455723	0.082195268***
7.171284087	11.72256304	0	3.590984031	0.000164716	3.344172389	0.000412642
16.69620192	34.41104184	0	10.5411676	0	9.81666344	0
0.156774472	-4.986090575	3.08066E-07	-1.52739393	0.063331542***	-1.422414738	0.077452944***
9.072612087	16.25155119	0	4.978353335	3.20638E-07	4.636186523	1.77448E-06
1.653350981	-1.421225764	0.077625568***	-0.435365458	0.331648589***	-0.405442389	0.342576179***
1.761011665	-1.164776608	0.122054715***	-0.356807141	0.360618095***	-0.332283457	0.369837615***
34.00091144	75.63108615	0	23.1681435	0	21.57577564	0
8.001603231	13.70039409	0	4.196854924	1.35324E-05	3.908401217	4.64545E-05
1.13573267	-2.654199347	0.003974842	-0.813063443	0.208090842***	-0.757180844	0.224470763***

*** not significant at p=0.01 n=14

$Q1 = 1.265137248$ $Q2 = 1.916330379$ $Q3 = 7.794023445$ $Q4 = 34.00091144$

[a] $Z = \frac{\widehat{\lambda} - 2.25}{\sqrt{\frac{\pi^2}{4n}}}$ for standard Cauchy [b] $Z = \frac{\widehat{\lambda} - 2.25}{\sqrt{\frac{\pi^2\left(\frac{1}{2}(Q_3 - Q_1)\right)}{4n}}}$ for robust Cauchy [c] $Z = \frac{\widehat{\lambda} - 2.25}{\sqrt{\frac{\pi^2 \overset{\triangle}{\sigma}^2}{4n}}}$ for generalized Cauchy with MLE $\overset{\triangle}{\sigma}$ for scale measure

Figure 2.18: Pseudo time series plot of South Africa MLA index



South Africa's numbers are generated from personal consumption expenditure in 2010 prices.

and its MLE counterpart upheld nine out of fourteen or 64% of the estimated MLA index values.

In this case the standard Cauchy over rejected H_0. This may be because that test statistic fails to take the scale parameter into account. The scale parameter is related to the interquartile range for the robust Cauchy statistic, and it influences the MLE Cauchy statistic.

ZA MLA index for intolerance to decline of standard of living

Summary statistics for α-stability of South Africa MLA index

Figure 2.18 depicts a plot of the empirical MLA index for South Africa quarterly semidurable PCE data computed according to Theorem 2.5.2. [29] Cursory inspection of the plot shows that it is explosive around the time of the Soweto uprisings in 1976, P. W. Botha's hardliner Rubicon

[29] We restate here the description used in section 2.5.2 The plot is based on a pseudo time series which was constructed by replacing the elements of each $\mathscr{B}^{L}_{2k-1}$ with the loss aversion index $\widehat{\lambda}_{k-1}$ computed from concatenation block $\mathscr{B}_{2k-1,2k}$. The observations in the $\mathscr{B}^{G}_{2k}$ component of that block were "zeroed out". In that way the MLA index only corresponds to losses. That procedure gave the plot a tableau look over losses.

speech around 1985, and pre-democracy talks with the African National Congress (ANC) in the early 1990s. Thus, the index is sensitive to political uncertainty.

Table 2.5: Diagnostics for South Africa MLA index

Statistic	Value	Percentile	Value
Sample Size	16	Min	0.07807
Range	20.315	5%	0.07807
Mean	2.4611	10%	0.09636
Variance	26.705	25% (Q1)	0.27523
Std. Deviation	5.1677	50% (Median)	0.66892
Coef. of Variation	2.0997	75% (Q3)	1.5212
Std. Error	1.2919	90%	11.842
Skewness	3.2348	95%	20.393
Excess Kurtosis	10.911	Max	20.393

Statistical tests for South Africa MLA index independence

A weak form test of Proposition 2.5.3 for MLA index independence was conducted with a simple autoregression which produced

$$\widehat{\lambda}_{t_{2k-1,2k}} = \underset{(p=0.090)}{2.90} - \underset{(p=0.679)}{0.116}\,\widehat{\lambda}_{t_{2k-3,2k-2}} \tag{2.5.11}$$

where the subscripts for $\widehat{\lambda}$ emphasize that MLA indexes were estimated in accord with the set $\mathscr{B}_{2k-1,2k}$, $k = 1,\ldots,K$. The intercept term is statistically significant at the $p = 0.10$ level so $\widehat{\lambda}$ is independent. However, the coefficient on the autoregressive term is not statistically different from 0 at the $p = 0.679$ level. A plot of this relationship is presented in Appendix 2.C.6.

Figure 2.19 is an empirical plot of the South Africa MLA index series adjusted for the median MLA index value of 2.25. So that $f(\widehat{\lambda}) = 1/[\pi(1+(\widehat{\lambda}-2.25)^2)]$. The plot provides a visual image for the descriptive statistics in Table 2.5. The MLE estimates for a generalized Cauchy distribution for South Africa data are $\overset{\triangle}{\sigma} = 0.36797$ for the scale parameter, and $\overset{\triangle}{\mu} = 0.57702$ for measure of location. Three popular goodness of fit measures: Kolmogorov-Smirnov, Anderson-Darling, and Chi-squared tests, uniformly upheld the generalized Cauchy fit at the $p = 0.05$ level.

Figure 2.19: Empirical distribution of South Africa MLA index

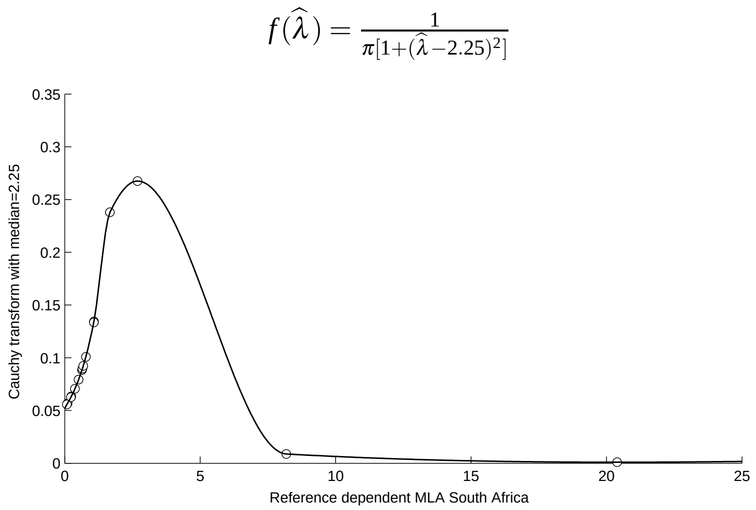


Table 2.6 displays the results of statistical tests for whether South Africa MLE index values are statistically equivalent to the median value of 2.25. The robust Cauchy test uniformly rejected the South Africa MLA index values as being statistically equivalent to 2.25. The standard Cauchy and MLE Cauchy upheld the null hypothesis in only two of the sixteen or 12.5% of the values at $p = 0.05$. Most of the MLA index values were less than 1. This implies that South Africans tend to be risk (gain) seeking (Wakker, 2010). That is, their utility function is convex over gain domain and steeper than that for loss domain. In the context of (2.3.5) this implies that the impact of loss of income on consumption is much lower than it would be for a similar loss in the US. Here, the standard Cauchy and MLE Cauchy upheld the same values at the $p = 0.05$ level. For US data, the two tests statistics did not agree on the same MLA index values.

Table 2.6: Cauchy test for South Africa MLA index
$H_0 : \lambda = 2.25 \; vs. \; H_a : \lambda \neq 2.25$

Loss Aversion Index ZA	Z-score[a]	P-value	Z-score[b]	P-value	Z-score[c]	P-value
0.375921603	-4.77230145	9.10663E-07	-10.79088648	1.90053E-27	-5.093019532	1.76203E-07
0.241664476	-5.114184416	1.5755E-07	-11.56393493	3.13826E-31	-5.457878424	2.40929E-08
0.637231784	-4.106880537	2.00519E-05	-9.286270392	7.98947E-21	-4.382879625	5.85604E-06
0.653345363	-4.065847647	2.39291E-05	-9.193488897	1.90151E-20	-4.339089158	7.15372E-06
0.782003075	-3.738223473	9.26626E-05	-8.452681697	1.42344E-17	-3.989447305	3.31137E-05
0.104196035	-5.464244927	2.32441E-08	-12.35547404	2.27533E-35	-5.831464426	2.74715E-09
8.177276227	15.09368497	0	34.12907645	0	16.10804203	0
1.07967304	-2.980213132	0.001440239	-6.738707083	7.99011E-12	-3.180495584	0.000735117
1.668433819	-1.480946119	0.069310466**	-3.348640404	0.000406046	-1.580471726	0.056999438**
20.39275185	46.2001382	0	104.465414	0	49.30497553	0
0.229222009	-5.145868898	1.33143E-07	-11.63557828	1.35871E-31	-5.491692232	1.9905E-08
2.685613422	1.109280471	0.133654604**	2.508248852	0.006066559	1.183828633	0.118240453**
0.078074308	-5.530763358	1.5942E-08	-12.50588215	3.4663E-36	-5.902453167	1.79068E-09
0.684503331	-3.986504533	3.35269E-05	-9.014082265	9.92628E-20	-4.254413864	1.04799E-05
0.51411166	-4.42040336	4.92584E-06	-9.995192328	7.99882E-24	-4.717472457	1.19396E-06
1.074078184	-2.994460316	0.001374654	-6.770922096	6.39821E-12	-3.195700238	0.00069746

** not significant at p=0.05 n=16

$Q1 = 0.342357321$ $Q2 = 0.668924347$ $Q3 = 1.226863235$ $Q4 = 20.39275185$

[a] $Z = \frac{\widehat{\lambda} - 2.25}{\sqrt{\frac{\pi^2}{4n}}}$ for standard Cauchy [b] $Z = \frac{\widehat{\lambda} - 2.25}{\sqrt{\frac{\pi^2 \left(\frac{1}{2}(Q_3 - Q_1)\right)}{4n}}}$ for robust Cauchy

[c] $Z = \frac{\widehat{\lambda} - 2.25}{\sqrt{\frac{\pi^2 \overset{\triangle}{\sigma}^2}{4n}}}$ for generalized Cauchy with MLE $\overset{\triangle}{\sigma}$ for scale measure

2.5.3 Macroeconomic loss aversion and cross-country uncertainty avoidance

The Hofstede uncertainty avoidance index (UAI is a popular measure of cross-cultural attitudes towards uncertainty in the literature on international business. It is computed from survey item response and data reduction techniques like cluster analysis (Hofstede, 1983).[30] The survey is typically given to employees in a sample of firms within a given country and the UAI score is assigned to the country. For example, the US UAI score is 46 while that for South Africa is 49 and the international business literature treats the two countries as if their attitudes towards uncertainty is similar. Hofstede (1983, p. 53) provides details on the theory and credits the 1963 edition of Cyert and March (1992) with the phrase "uncertainty avoidance".

Our cross-country analysis of loss aversion between the US and South Africa provides contrast to UAI type survey item response, and use of incentivized experiments in measuring risk attitudes. Anderson and Mellor (2009) found that risk preferences were not stable across survey response and incentivized experiments for the subjects in their experiments. Sahm (2012) finds that risk preference is stable at the level of the individual but very heterogeneous among individuals from the survey item responses in her study. However, Coppola (2014) finds that domain specific surveys are more effective in eliciting risk preferences in some situations. The empirical evidence in income and consumption data for US and South Africa in our study show that risk attitudes in the two countries are markedly different from that reported in surveys and experiments. So our empirical strategy provides a mechanism for measuring macroeconomic risk attitudes like loss aversion with hard economic data. This allows for a more internally consistent method for cross-country risk attitude comparison up to measurement error in macroeconomic data.

2.6 Application to information based asset pricing

In this section we show how the MLA index applies to information based asset pricing. It serves double duty as jumps in a subordinate Cauchy process for information flows in asset pricing of defaultable binary bonds, and as the source of loss aversion to default. Unless otherwise stated we

[30] Refer to Hofstede index http://www.clearlycultural.com/geert-hofstede-cultural-dimensions/uncertainty-avoidance-index/ The index is computed from a formula $UAI = 40 \times (\overline{Q18} - \overline{Q15}) + 25 \times (\overline{Q21} - \overline{Q24}) + C$ where C is a constant, and $\overline{Q\#}$ is the average item response for question $Q\#$ in the Values Survey 2013 Module available at http://www.geerthofstede.nl/vsm2013.

make the standard assumption that uncertainty in the model is characterized by a probability space $(\Omega, \mathscr{F}, \mathbb{F}, P)$ where Ω is the underlying sample space, $\mathscr{F}$ is a σ-field of Borel measurable subsets of Ω, $\mathbb{F}$ is a filtration of information $\mathscr{F}_s \subset \mathscr{F}_t$, $s < t$, and P is a probability measure.

2.6.1 MLA and Cauchy information flow for defaultable binary bonds

Preliminaries

To fix ideas we define a defaultable binary bond over a finite duration $[0, T]$ as one which has a terminal cash flow $H_T \in \{h_0, h_1\}$. The creditor recovers h_0 upon default with probability p_0 or she receives the contracted amount h_1 with probability p_1 where $p_0 + p_1 = 1$. So H_T is based on a binary outcome: default or no-default. Let P_{tT} be a discount factor, and B_{tT} be the bond price at time $t < T$. We assume that information flows with an "activity rate" c. The bond is thinly traded when c is small so the price moves only when there is new information that affects the beliefs of traders. Refer to Clark (1973) for an early review of the activity rate or trading times concept. More recent models by Carr and Wu (2004) use the phrase "business time" to distinguish the activity rate from "clock time" in their option price model based on subordinate Lévy processes. Cartea and Jaimungal (2013) used marked point processes to model activity rate in a market microstructure model of high frequency trading.

Defaultable binary bond pricing

The price of the bond is given by

$$B_{tT} = P_{tT} E^P [H_T | \mathscr{F}_t] \tag{2.6.1}$$

The goal of information based asset pricing is to characterize the information in $\mathscr{F}_t$ with a suitable stochastic process. For instance, Hoyle (2010, §7) model information flow with a Cauchy random bridge (CRB) among other candidate bridge processes. Hoyle et al. (2011) model the information flow as a Lévy random bridge (LRB) process. More recent, Ikpe et al. (2014) extended the model to more abstract processes based on a conditioning information theory. Our contribution extends the information process to myopic loss aversion to default that mimic jumps in a Cauchy information process.

Figure 2.20: A Cauchy process sample path

Source: https://almostsure.wordpress.com/2011/02/25/properties-of-levy-processes/#more-1052

Figure 2.20 depicts a simulated Cauchy process $C_t^Y(\omega)$. In Theorem 2.5.7, the process $C_t^{Y\star}(\omega)$ is comprised solely of the positive jumps of $C_t^Y(\omega)$ and remains flat otherwise.

Figure 2.21: US MLA Subordinate Credit Risk Index

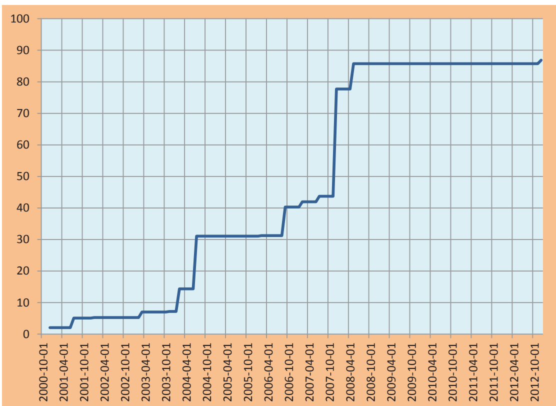
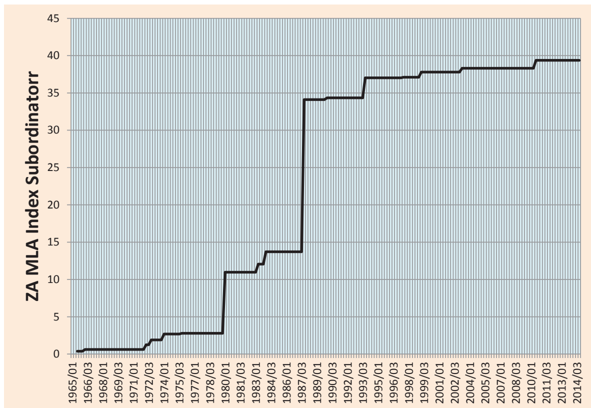


Figure 2.22: South Africa MLA Subordinate Credit Risk Index



The subordinate "credit risk index" process $C_t^{Y\star}(\omega)$ is comprised solely of the independent (positive) jumps of the Cauchy process $C_t^Y(\omega)$ and remains flat otherwise. The positive jumps are the MLA index estimates for US and South Africa constructed from the activity rate in the set $\mathscr{B}_{2k-1,k}$ defined in Definition 2.5.1, and rationalized in Propositions 2.5.3 and 2.5.6.

Figure 2.21 and Figure 2.22 depict the $C_t^{Y\star}(\omega)$ process for US and South Africa. In the context of information based asset pricing, $C_t^{Y\star}(\omega)$ is functionally equivalent to a macroeconomic credit risk index. Its jumps reflect fear of loss motivated by some "credit event" in the economy. By virtue of

Figure 2.23: US MLA synthetic Credit Risk Index and FNMA Asset Price Paths

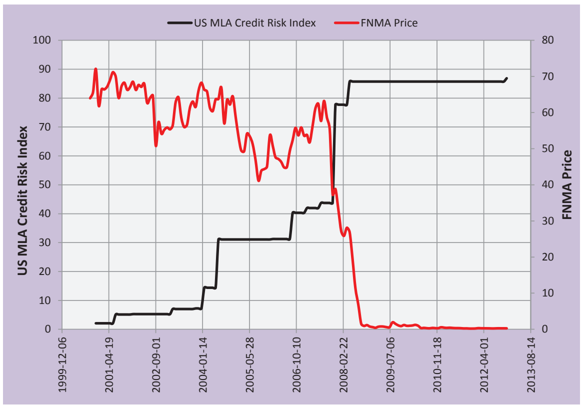
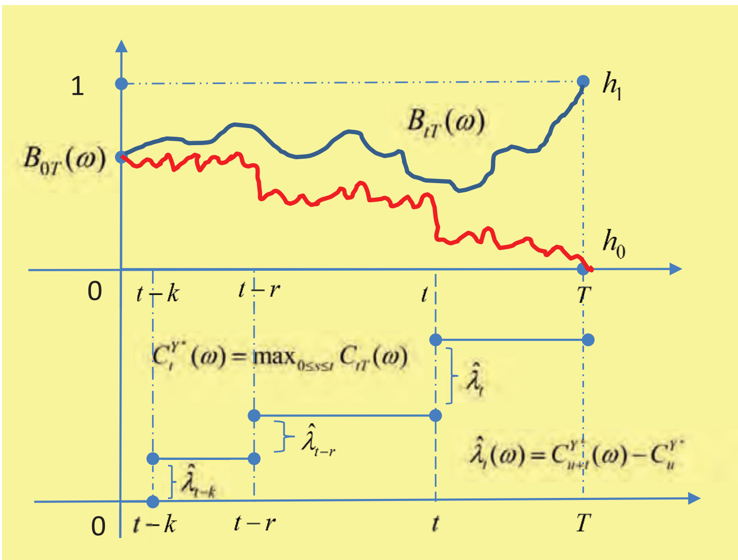


Figure 2.24: Lehman Brothers Credit Default Swap Prices and Asset Price Paths

[Figure 2.23](#) depicts a cross-plot of our synthetic credit risk index based on the subordinate Cauchy bridge process $C_t^{Y\star}(\omega)$. The Federal National Mortgage Association (FNMA) had a lot of mortgage default exposure leading up to the Great Recession of 2008. That information was impounded in the asset price. Jumps in $C_t^{Y\star}(\omega)$ coincide with credit events arising from natural disasters and financial market instability. [Figure 2.24](#) is a cross-plot of Lehman Brothers credit default swap (CDS) risk premia, and Lehman Brothers asset price in the months leading up to the financial crisis. Positive jumps in the CDS risk premium coincide with drops in the asset price.

Figure 2.25: Binary bond path process



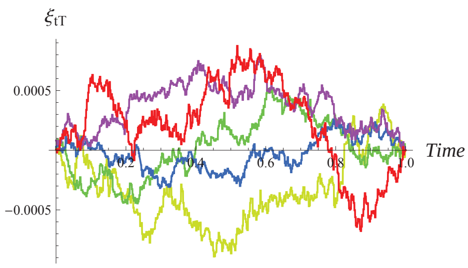
The MLA index estimator $\widehat{\lambda}$ mimics independent *positive jumps* of a Cauchy bridge $C_{tT}(\omega)$.

Propositions 2.5.3 and 2.5.6, the MLA index estimator $\widehat{\lambda}$ mimics independent *positive jumps* of the subordinate Cauchy process $C_t^{Y^\star}(\omega)$. So that

$$\widehat{\lambda}_t = C_{u+t}^{Y^\star}(\omega) - C_u^{Y^\star}(\omega) \tag{2.6.2}$$

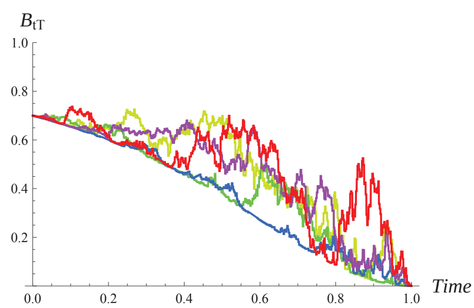
Figure 2.23 depict cross plots of the asset price for the Federal National Mortgage Association (FNMA) which had a large exposure to mortgage default leading up to the financial crisis in 2008. Figure 2.24 is a similar plot for Lehman Brothers in the months preceding their bankruptcy. In each case the credit risk index jumps up when there is a credit event, and the asst price drops. Thus, lending support to our claim that positive jumps in a suitable credit index process mimic loss aversion to default. We provide a sketch of the binary bond pricing process in Figure 2.25. Anticipated bad news increases trade activity. According to Definition 2.5.1, bad news causes a run of type $\mathscr{B}_{2k-1}^L$ and induces an MLA index estimate $\widehat{\lambda}$ from $\mathscr{B}_{2k-1,2k}$. The default path follows a "bridge" or "tied down" process from B_{0T} until default at h_0. The non-default path ends at h_1. Throughout the life of the default path, each jump in the information process signals investor pessimism and fear of loss.

Figure 2.26: Subordinate Cauchy information processes for $c = 10^{-3}$



Information process when activity rate for information arrival $c = 10^{-3}$

Figure 2.27: Defaultable bond price paths for $c = 10^{-3}$



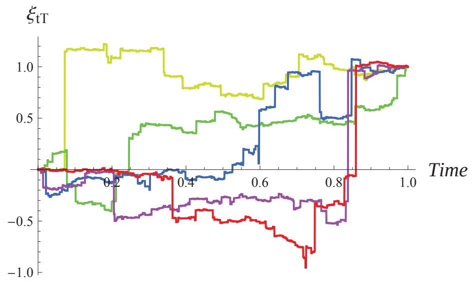
Bond price when activity rate for information arrival $c = 10^{-3}$

Source: Hoyle (2010, p. 112). Positive jumps in the Cauchy bridge process mimic the MLA index. As default date approaches for the bond, the information flow is pessimistic with persistent loss aversion.

For example, the jumps at $t-k$, $t-r$ and t are represented by the MLA index random variables

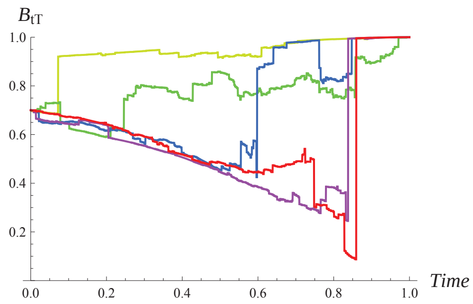
$\widehat{\lambda}_{t-k}$, $\widehat{\lambda}_{t-r}$ and $\widehat{\lambda}_t$, respectively. The cumulative effect of this pessimism, i.e., $(\widehat{\lambda}_{t-k} + \widehat{\lambda}_{t-r} + \widehat{\lambda}_t$, coincides with the default date at T.

Figure 2.28: Subordinate Cauchy information processes for $c = 1$



Information process when activity rate for information arrival $c = 1$

Figure 2.29: Non-defaultable bond price paths for $c = 1$



Bond price when activity rate for information arrival $c = 1$

Source: Hoyle (2010, p. 113). Positive jumps in the Cauchy bridge process mimic the MLA index. As terminal date approaches for the non-defaultable bond, the information flow is optimistic so loss aversion is overcome.However, bond prices were decreasing for the most part when the Cauchy bridge process jumped up.

We note in passing that inasmuch as the natural filtration $\mathscr{F}_t^{C^Y} \subset \mathscr{F}_t$ is given by $\mathscr{F}_t^{C^Y} \triangleq \sigma\left(C_s^Y; 0 \leq s \leq t\right)$, the information flow in the asset pricing paradigm above is equivalent to $\mathscr{F}_t^{C^Y} \subseteq \sigma(\mathscr{B}_{2s-1,s}; 0 \leq s \leq t)$, where $\mathscr{B}_{2s-1,s}$ is defied in Definition 2.5.1. The latter implies that, in principle, we should be able to price binary bonds with the information we have about income and consumption in $\mathscr{B}_{2s-1,s}$. In other words, there exist a consumption based capital asset pricing approach to binary bond pricing. Cf. Backus et al. (1989). Moreover, in (2.6.2) $\widehat{\lambda}_t$ is independent of the natural filtration $\mathscr{F}_u^{C^Y}$.

Behavioural explanation of Hoyle (2010) simulation model

The following example is adapted from Hoyle (2010, §7). It is effectively, a simulation of our model sketch. Hoyle (2010) constructed an example of an information-based defaultable binary bond for which the information flow follows a Cauchy random bridge (CRB) process ξ_{tT}. The binary bond price in Hoyle (2010, p. 108) is given by $B_{tT} = \Lambda(t, \xi_{tT})$ where

$$\Lambda(t, \xi_{tT}) = P_{tT}\frac{Y_1}{Y_2} \tag{2.6.3}$$

$$Y_1 = \alpha_0(t) + \alpha_1\xi_{tT} + \alpha_2\xi_{tT}^2 \tag{2.6.4}$$

$$Y_2 = \beta_0(t) + \beta_1\xi_{tT} + \beta_2\xi_{tT}^2 \tag{2.6.5}$$

where $\alpha_0(t)$ and $\beta_0(t)$ are functions of t and the other coefficients are relative constants that depend on terminal date T and h_0, h_1, p_0, p_1 and activity rate c. The ratio Y_1/Y_2 is a pricing kernel $M_t(\xi_{tT}|h_0, h_1, p_0, p_1, c)$ which can be written as

$$M_t(\xi_{tT}|\cdot) = \frac{Y_1}{Y_2} = 1 + \frac{(\alpha_0(t) - \beta_0(t) + (\alpha_1 - \beta_1\xi_{tT} + (\alpha_2 - \beta_2)\xi_{tT}^2)}{\beta_0(t) + \beta_1\xi_{tT} + \beta_2\xi_{tT}^2} \tag{2.6.6}$$

Cursory inspection show that (2.6.6) contains a stochastic trend driven by innovations in ξ_{tT} and ξ_{tT}^2. These innovations include the MLA index for loss aversion to credit default. In principle, given our knowledge about the behaviour of the MLA index we should be able to estimate the probability of a market crash. We leave that for another day.

For our purposes, ξ_{tT} is subordinate to the Cauchy process $C_t^Y(\omega)$, and its jumps reflect loss aversion to decline in the bond price. According to Theorem 2.5.7 Hoyle's Cauchy bridge process is related to our MLA index driven Cauchy process by

$$C_t^{Y\star} = \max_{\substack{0 \leq s \leq t \\ 0 < t \leq T}} \xi_{sT} \tag{2.6.7}$$

Thus, the positive jumps in ξ_{tT} mimic the MLA index $\widehat{\lambda}$ as in (2.6.2). Figure 2.26 depicts Hoyle's simulated model of information flows for Cauchy bridge processes with activity rate $c = 10^{-3}$ (represented by σ in our appendix (2.B.13)). Cursory inspection of Figure 2.27 shows that when there is an increase in information flow about the bond–presumably bad news, the CRB ξ_{tT} increases. However, the bond price B_{tT} tends to decrease. So investor loss aversion is being mimicked by the CRB.

The information flows and price patterns for non-defaultable bonds in Figure 2.28 and

Figure 2.29 show that as long as the Cauchy bridge process is positive the bond price remains fairly stable. Even though there are small jumps which may signal loss aversion. However, when information flow is negative, i.e., the bridge process in negative, bond prices continue to decline. It is only after the first crossing of the bridge into the positive quadrant does bond price paths jump on this good news to h_1.

2.7 Application to the equity premium puzzle and behavioural pricing kernel

The equity premium puzzle is one of the most actively researched areas in financial economics and macro-finance (Campbell, 2003; Ludvigson, 2013; Campbell, 2015). Therefore, this section is limited in scope to a theoretical characterization of the equity premium in the context of our relative income model with embedded MLA index. In particular, our goal is to establish our behavioural model's ability to do the following:

- predict large equity premia;
- explain why equity premium is counter-cyclical to business cycles;
- identify the source/cause of excess volatility.

We do not address the issue of price/dividend ratio or Hansen-Jagannathan bounds Sharpe ratio connection in the CCAPM (Hansen and Jagannathan, 1991). However, we note that Campbell and Cochrane (1999, pp. 220-221) showed empirically that the price/dividend ratio is linear in surplus consumption ratio. And their surplus consumption ratio $(C-X)/C$, relative to habit formation consumption X, is similar to our RIH consumption growth $(C_t^D - a(d))Y_t)/C_t^D$, relative to reference consumption $a(d))Y_t$ in (2.3.6). Therefore, in principle our theoretical model should be able to explain the behaviour of price/dividend ratio in a manner analogous to Campbell and Cochrane (1999) albeit with MLA index included in our growth formulation. Additionally, the α-stable feature of the the MLA index embedded in the behavioural pricing kernel makes it (the kernel) unmeasurable. So in order to apply the Hansen-Jagannathan bounds one needs a time change transformation of the α-stable process which is subordinate to a Gaussian distribution. Whereupon a modified Sharpe ratio can be applied. See e.g., Charles-Cadogan (2016a) for time changed Sharpe ratio analysis.

In this section, we introduce a behavioural pricing kernel that depends on the joint distribution of consumption growth rates and MLA index. So it synthesizes aspects of the Constantinides (1990) (habit formation) and Benartzi and Thaler (1995) (myopic loss aversion) models resolution of the equity premium puzzle. Yogo (2008); Hung and Wang (2011); Curatola (2015) also introduced models with loss aversion and consumption. However, in their models loss aversion is deterministic. Our model is distinguished because in it the MLA index is a random variable based on the α-stable theory we introduce in this paper. Furthermore, we embed the MLA index in the pricing kernel and show how its statistical distribution affects asset prices.

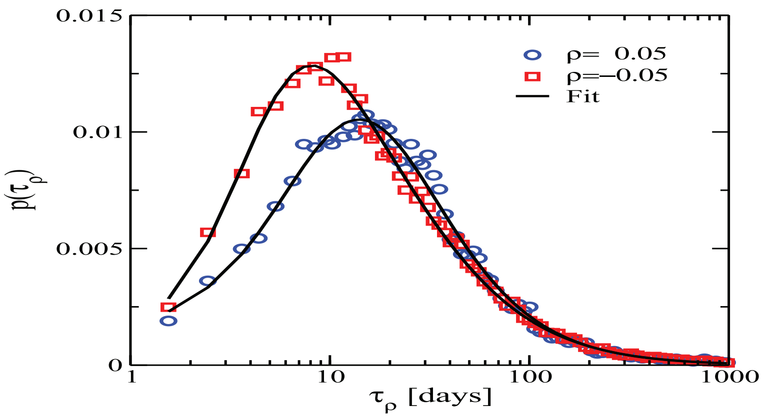
2.7.1 Myopic loss aversion and holding periods of stocks

The impact of myopic loss aversion is illustrated by the following example adapted from Jensen et al. (2003). Let $S(t)$ be the levels price of a stock and $r_{\Delta t}(t) = \ln S(t+\Delta t) - \ln S(t)$ be the short horizon return. We are interest in the first passage time $\tau_\rho(t)$ of returns to attain a level ρ, i.e., $\tau_\rho(t) = \inf\{t > 0 | \, |r_{\Delta t}(t)| > \rho\}$. Thus, an investor may implement a simple trading strategy based on short sell $r_{\Delta t}(t) < -\rho$ or hold for $r_{\Delta t}(t) > \rho$. They used wavelet analysis to filter out the long term trend in the Dow Jones Industrial Average (DJIA) between May 1896 and June 2001 so that $s(t) = \tilde{s}(t) + d(t)$, where $s(t) = \ln S(t)$, $d(t)$ is filtered trend and $\tilde{s}(t)$ is the gain or loss component or returns around trend. Jensen et al. (2003) fitted the following generalized Gamma probability distribution function separately over gains $\tilde{s}^+(t)$ and losses $\tilde{s}^-(t)$

$$p(t) = \frac{\nu}{\Gamma\left(\frac{\alpha}{\nu}\right)} \frac{|\beta|^{2\alpha}}{(t+t_0)^{\alpha+1}} \exp\left(-\left(\frac{\beta^2}{t+t_0}\right)^2\right) \tag{2.7.1}$$

A cross plot of the fits is depicted in Figure 2.30. The probability density function $p^-(t)$ for losses has a much higher elevation for losses over a given time period compared to that for gains $p^+(t)$ over a similar period. Thus, the probability of experiencing a loss or "draw down" is much higher over a given period. However, as the length of the trading intervals increase the probability distributions cross such that the probability of gains exceeds that of losses in an interval $[t_\ell, t_u] = \{t > 0 | p^+(t) > p^-(t),\ t_\ell \leq t \leq t_u\}$. Whereafter the window of profitability disappears and the two distributions coincide. The phenomenon of higher returns over less frequent evaluation periods was confirmed in an economic experiment by Gneezy and Potters (1997). We note in passing that the generalized Gamma distribution is an admissible α-stable probability distribution function for the MLA index estimated for US and South Africa.

Figure 2.30: Gain loss asymmetry over
investment horizons on DJIA

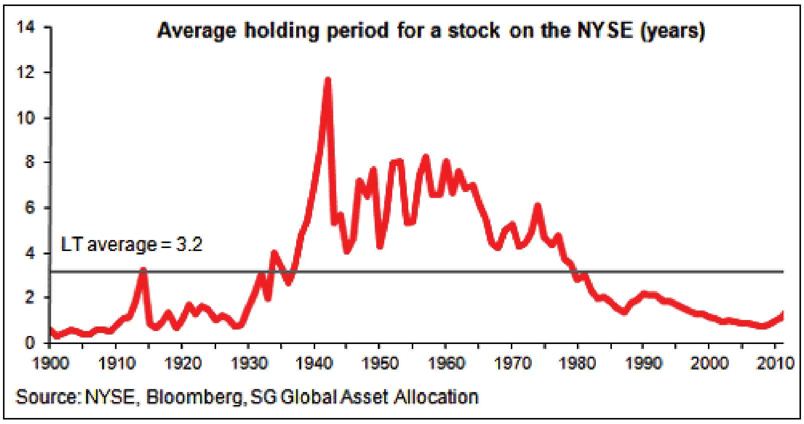


Source: Jensen et al. (2003)

According to Jensen et al. (2003) the fitting parameters used to obtain these fits are for $\rho^+ = 0.05$: $\alpha = 0.50, \beta = 4.5 days^{1/2}, \nu = 2.4$, and t_0 = 11.2 days; and for $\rho^- = 0.05$: $\alpha = 0.50, \beta = 5.0 days^{1/2}, \nu = 0.7$, and t_0 = 0.6 days

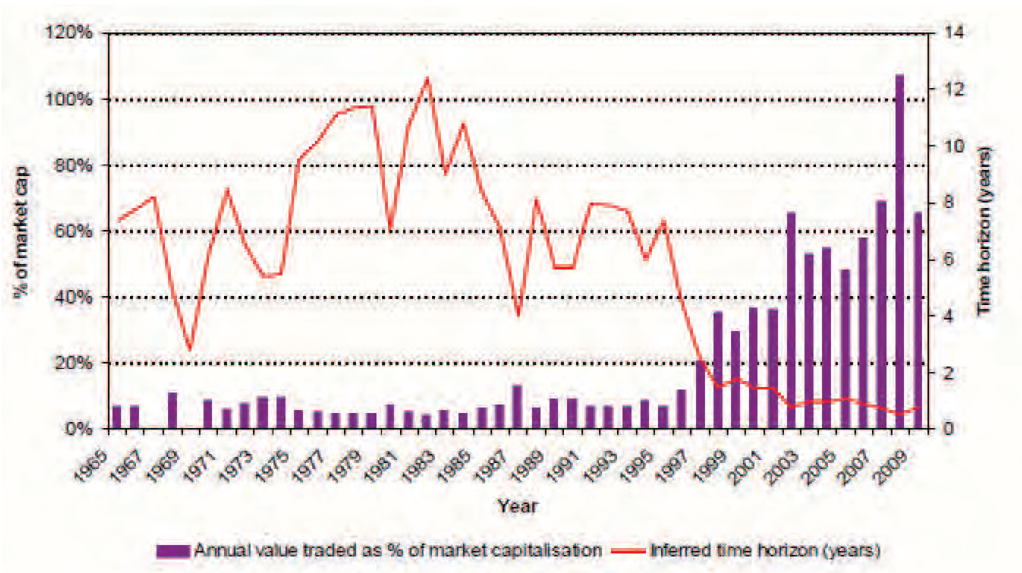
Figure 2.31 and Figure 2.32 depict the distribution of the holding period of a stock on the NYSE and JSE over the last 100 years and 50 years respectively. It shows that the holding period is in steady decline. According to Dumontier (2012), for the US it bottomed out at 9-months in 2008 and climbed back to 14-months in 2012. “In other words, investors are now roughly turning over their entire portfolio on an annual basis. When you consider that the two fundamental components of real equity returns are the dividend yield and real earnings growth, you quickly realize that the current level of trading activity is impossible to justify in terms of economic fundamentals.” Benartzi and Thaler (1995) argue that myopic loss aversion is the cause of such behaviour. Their simulation show that a MLA index of around 2.77 coincided with a holding period of around

Figure 2.31: Average holding period of a stock traded on US NYSE



Source: Dumontier (2012), "Best Practices for Long-Term Investors in a Microsecond Market", *The Motley Fool*

Figure 2.32: Average holding period of a stock traded on South Africa JSE



Source: Wessels (2010), "Dividends: The Major Source of Real Equity Returns", *DRW Investment Research Report*

The plots depict the average holding period of a stock on the New York Stock Exchange (NYSE) and Johannesburg Stock Exchange (JSE) over the last 100 years, and 50 years respectively. The holding period is in steady decline. An eyeball of the plots show that the most recent holding period is below 12-months in each exchange. The exponential growth in the volume of stocks traded on the JSE originated around 1998.

12-months. However, they used the "prospective utility" function

$$V(P) = \sum_{i=1}^{n} \pi_i v(x_i) \tag{2.7.2}$$

where π_i is the decision weight obtained from Quiggin (1982) transformation procedure for probabilities associated with a gamble P, and v is Tversky and Kahneman (1992) value function over gains and losses (in returns) relative to a reference point.

The equity premium puzzle introduced by Mehra and Prescott (1985) posits that the observed equity premium in financial markets is too high to be explained by neoclassical equilibrium models. In particular, an unrealistically high risk aversion index and covariance between consumption growth and asset returns are required to match the observed equity premium. Campbell (2003, pp. 803-803) list 14-different empirical regularities of stock market returns and consumption growth that motivate research in consumption based capital asset pricing. Two important models that predict large equity premia are Constantinides (1990) continuous time habit formation in consumption model, and Benartzi and Thaler (1995) simulation model with myopic loss aversion and mental accounting in equity portfolios. Campbell (2003) and Ludvigson (2013) survey the literature on alternative preference based theories that resolve aspects of the puzzle.

In this section we compare, and contrast, the asset pricing behaviour of investors under our modified RIH with myopic loss aversion, to fundamental asset pricing under the neoclassical model. As shown in Theorem 2.3.1 the RIH consumption function retains several aspects of the CPT value function v in (2.7.2) for linear utility. Furthermore, to the extent that $\sum_{i=1}^{n} \pi_i = 1$ in (2.7.2), the transformed probability weights π_i are functionally equivalent exchange of probability distributions that admit an expectations operator of the type used in (2.7.8) below.

2.7.2 Asymmetric distribution of behavioural pricing kernel for RIH

The ensuing analysis is based on the pricing kernel approach to the consumption asset pricing model (CCAPM) in Cochrane (2005, pp. 15-17). p_t is the ex-dividend price of an asset, C_t is consumption at time t; β is a subjective discount factor; x_{t+1} is a future payoff at time $t+1$; E_t is an expectations operator based on information at time t and U is a utility function; M_{t+1} is a *pricing kernel* related to the intertemporal rate of substitution obtained from an Euler equation.

The first order necessary conditions amount to:

$$M_{t+1} = \beta \frac{U'(C_{t+1})}{U'(C_t)}, \qquad p_t = E_t[M_{t+1} x_{t+1}] \tag{2.7.3}$$

Assuming a CRRA specification $U(C_t) = \dfrac{C^{1-\gamma}}{1-\gamma}$ with constant risk aversion parameter γ, after a loglinear approximation, the pricing kernel is reduced to

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \approx \beta\, e^{-\gamma g^c_{t+1}} \tag{2.7.4}$$

where g^c_{t+1} is consumption growth. Under our myopia Axiom 1, consumption tracks income so the growth rates of the two are approximately equal, i.e., $g^c_{t+1} \approx g^Y_{t+1}$. According to Wickens (2011, p. 303) the equity premium puzzle arises because (2.7.4) fails empirically. The observed asset price p_t in (2.7.3) is much higher than that predicted by (2.7.4). As will be shown below, our RIH approach to asset pricing predicts a much larger asset price (and hence equity premium) than the neoclassical model. So it is sufficient to resolve the equity premium puzzle.

According to Kahneman and Tversky (1979); Tversky and Kahneman (1992) DMs do not formulate decisions on the basis of terminal wealth. Rather they formulate it on the basis of changes in wealth relative to a reference point. Benartzi and Thaler (1995) and Barberis and Huang (2001) extended that concept to asset pricing in tandem with MLA in order to resolve the equity premium puzzle. MLA reflects investor psychology and risk perception wherein losses loom larger than gains. Myopic investors concentrate on short-term price movements. They frame financial decision making such that loss aversion is the predominant risk attitude that motivate them to overweight short term price movements. Consequently, they tend to be more conservative investors who demand higher equity premia to hold equities. Yogo (2008); Hung and Wang (2011); Curatola (2015) also extend the MLA concept to specification of preferences in a CCAPM with loss aversion. We extend that concept to the change in the RIH consumption function by substituting

$\Delta^b C_t^D$ in (2.3.5) for C_t in (2.7.4) to get the *behavioural pricing kernel*[31]

$$M_{t+1}^D = \beta \left(\frac{\Delta^b C_{t+1}^D}{\Delta^b C_t^D} \right)^{-\gamma} = \beta \left(1 + \varphi_{t+1}(\boldsymbol{g}, \boldsymbol{\lambda}) \right)^{-\gamma} \tag{2.7.5}$$

where the growth rate $\varphi_{t+1}(\boldsymbol{g}, \boldsymbol{\lambda})$ is a function of the *joint distribution of income and consumption growth rates and MLA indexes*, i.e., $\boldsymbol{g} = (g_t, g_{t+1}, g_{t+1}^Y)$ and $\boldsymbol{\lambda} = (\lambda_t, \lambda_{t+1})$ where $g_t = -\alpha_1 g_t^G < 0$. In Campbell and Cochrane (1999, pp. 208-209) habit formation model $S_t = C_t - X_t$ is referred to as "surplus" consumption, where X_t is the level of habitual consumption. In our model, the equivalent of their X_t is a reference consumption level $a(d)Y_t$, and loss aversion enters through intolerance to decline below reference consumption as shown in (2.3.6). The joint distributions above are motivated by (2.3.5) which induce four possible values for consumption growth rate as shown below.

The following state transitions play a prominent role in the rest of this section.

Definition 2.7.1 (State transitions)**.**

Let $\{G, L\}$ be the possible states of income where G is gain and L is loss. $\mathscr{I} = \{G, L\} \times \{G, L\} = \{GG, GL, LG, LL\}$ is the index set for the possible state transitions for current and anticipated states. For example, the consumption pair $\{\Delta^b C_t^D, \Delta^b C_{t+1}^D\}$ reflects current states at time t and anticipated states at time $t+1$. The possible current and anticipated gain (G) and loss (L) states in income are:

GG current gain state G and an anticipated gain state G for next period;

GL current gain state G and an anticipated loss state L for next period;

LL current loss state L and an anticipated loss state L for next period;

LG current loss state L and an anticipated gain state G for next period;

□

Assumption 2.7.1 (Time dependent loss aversion)**.** *Loss aversion is time dependent so that $\lambda \sim$ iid $\mathscr{C}$ is substituted with λ_t.*

This assumption is motivated by Barberis and Huang (2001); Barberis et al. (2001) who used time and state dependent loss aversion in their model. The RIH consumption function in (2.3.5)

[31](Shefrin, 2008) provides a review of the literature on this issue.

induce the following heterogenous consumption growth rates

$$\frac{\Delta^b C_{t+1}^D}{\Delta^b C_t^D} = 1 + \varphi_{t+1}(\boldsymbol{g}, \boldsymbol{\lambda}), \qquad \theta_{t+1}(\boldsymbol{g}) = \left(\frac{g_{t+1}}{g_t}\right) e^{g_{t+1}^Y} \tag{2.7.6}$$

$$\varphi_{t+1}(\boldsymbol{g}, \boldsymbol{\lambda}) = \begin{cases} \theta_{t+1}(\boldsymbol{g}) - 1 & \text{if GG state transitions} \\ -\lambda_{t+1}\theta_{t+1}(\boldsymbol{g}) - 1 & \text{if GL state transitions} \\ -\dfrac{1}{\lambda_t}\theta_{t+1}(\boldsymbol{g}) - 1 & \text{if LG state transitions} \\ \dfrac{\lambda_{t+1}}{\lambda_t}\theta_{t+1}(\boldsymbol{g}) - 1 & \text{if LL state transitions} \end{cases} \tag{2.7.7}$$

We use a log growth rate relationship for $Y_{t+1} = (1 + g_{t+1}^Y)Y_t$ to construct $\theta_{t+1}(\boldsymbol{g})$ in (2.7.6).[32] According to (2.7.7) GG is the only consumption state transition that is not affected by the loss aversion index.

The behavioural pricing kernel depends on the CRRA parameter γ and the joint distribution of growth rates and MLA indexes through $\varphi_{t+1}(\boldsymbol{g}, \boldsymbol{\lambda})$.[33] Log linear approximation of M_{t+1}^D in (2.7.5) implies the existence of γ such that we have a behavioural pricing kernel

$$M_{t+1}^D(\beta, \gamma, \boldsymbol{g}, \boldsymbol{\lambda}) \approx \beta e^{-\gamma\varphi_{t+1}(\boldsymbol{g}, \boldsymbol{\lambda})}, \qquad p_t^D = E_t[M_{t+1}^D(\beta, \gamma, \boldsymbol{g}, \boldsymbol{\lambda}) x_{t+1}] \tag{2.7.8}$$

where p_t^D is the price of the asset predicted by the RIH pricing kernel. We reiterate here that the expectations operator in (2.7.8) is admissible for M_{t+1}^D, defined in (2.7.5), by virtue of the exchangeability of probability measures implied by the "prospect utility" valuation in (2.7.2).

2.7.3 Resolution of equity premium puzzle with behavioral pricing kernel

The pricing equations (2.7.3) and (2.7.5) can be rewritten as

$$1 = E_t\left[M_{t+1}\frac{x_{t+1}}{p_t}\right], \qquad 1 = E_t\left[M_{t+1}^D \frac{x_{t+1}}{p_t^D}\right] \tag{2.7.9}$$

[32] The common RHS factor is derived from $\ln(\theta_{t+1}(\boldsymbol{g})) = \ln\left(\frac{g_{t+1}}{g_t}\frac{Y_{t+1}}{Y_t}\right) \approx \ln\left(\frac{g_{t+1}}{g_t}\right) + g^{Y_{t+1}}$.

[33] Choi et al. (2007, pp. 1931-32) also found that the risk premium in their model of attitudes towards risk in insurance microdata are driven jointly by loss aversion (or (Gul, 1991) disappointment aversion) and risk aversion indexes.

Let r_t^f be the risk free rate. The price of the risk free asset in (2.7.9) is such that the payoff 1-period ahead is $x_{t+1} = 1 + r_t^f$ so that

$$1 = E_t\left[M_{t+1}\frac{x_{t+1}}{p_t}\right] = E_t\left[M_{t+1}\frac{1+r_t^f}{p_t}\right] \Rightarrow E_t\left[M_{t+1}\right] = \frac{1}{1+r_t^f} \tag{2.7.10}$$

$$1 = E_t\left[M^D_{t+1}\frac{x_{t+1}}{p_t}\right] = E_t\left[M^D_{t+1}\frac{1+r_t^f}{p_t}\right] \Rightarrow E_t\left[M^D_{t+1}\right] = \frac{1}{1+r_t^f} \tag{2.7.11}$$

Let r_{t+1} be the return on the asset in the next period. Assuming that M_{t+1} and $(1+r_{t+1})$ are jointly lognormal Wickens (2011, pp. 284-285) shows that under the no arbitrage hypothesis

$$\begin{aligned} E_t\left[r_{t+1}\right] - r_t^f &= -\frac{1}{2}V_t(r_{t+1}) + \gamma_t(c)Cov_t\left(\Delta \ln C_{t+1}, r_{t+1}\right) \\ &= -\frac{1}{2}V_t(r_{t+1}) + \gamma_t(c)Cov_t\left(g^c_{t+1}, r_{t+1}\right) \end{aligned} \tag{2.7.12}$$

where $\gamma_t(c)$ is Arrow-Pratt measure of relative risk aversion, $V_t(r_{t+1})$ is *conditional volatility* of asset returns, and $Cov_t\left(g^c_{t+1}, r_{t+1}\right)$ is the *conditional covariance* between asset returns and consumption growth.[34] The equity premium puzzle arises empirically because the observed left hand side (LHS) of (2.7.12) is too large to be explained by asset price volatility $V_t(r_{t+1})$, risk aversion $\gamma_t(c)$ and the covariance $Cov_t\left(g^c_{t+1}, r_{t+1}\right)$ between consumption growth and asset returns.

In the context of our RIH specification (2.7.12) is rewritten as

$$E_t\left[r^D_{t+1}\right] - r_t^f = -\frac{1}{2}V_t(r^D_{t+1}) + \gamma_t^D(c)Cov_t\left(\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda}), r^D_{t+1}\right) \tag{2.7.13}$$

where the superscript D implies RIH. However, under Theorem 2.3.1 DMs evaluate their consumption stream relative to a reference consumption, and they have MLA and intolerance to decline in consumption. So that if $C_1, C_2, \ldots, C_{r-1}, C_r, C_{r+1}, \ldots, C_T$ is a T-period consumption stream with reference consumption C_r, then $\eta_t = C_t - C_r$ is the relative change in consumption. We assume that $\eta_t \sim (\mu_\eta, \sigma^2_\eta)$ where $E[\eta_t] = \mu_\eta$ and σ^2_η is the variance of η_t.[35] On the basis of those prerequisites

[34] The quantity $-\frac{1}{2}V_t(r_{t+1})$ is referred to as Jensen's inequality correction factor. For example, if $R_A = \sum_{t=1}^{T}(1+r_t)/T - 1$ is an arithmetic mean return and $R_G = \left(\Pi_{t=1}^{T}(1+r_t)\right) - 1$ is the corresponding geometric mean, then $1+R_G \approx e^{R_A} \approx 1 + R_A - R_A^2/2$. So that $R_G \approx R_A - R_A^2/2$. See Grinblatt and Linnainmaa (2011); Becker (2012) for further details on this issue.

[35] If $W \sim \mathcal{N}(\mu_w, \sigma^2_w)$ and $C = \ln W$, then consumption is lognormally distributed and $\mu_\eta = \exp(\mu_w + \frac{\sigma_w^2}{2})$, $\sigma^2_\eta = \mu^2_\eta \exp(\frac{\sigma_w^2}{2} - 1)$. So we treat μ_η and σ^2_η as parameters.

we claim

Proposition 2.7.2 (MLA index and risk aversion nexus)**.** *Under RIH the relationship between Arrow-Pratt measure of relative risk aversion and myopic loss aversion to decline in consumption is given by*

$$\gamma_t^D(\mu_\eta) = \mu_\eta \, sgn(\widetilde{\gamma}(\mu_\eta))\, \widetilde{\gamma}(\mu_\eta)\lambda_t(\mu_\eta) \tag{2.7.14}$$

where $\gamma_t^D(\mu_\eta)$ *is Arrow-Pratt measure of relative risk aversion,* $\lambda_t(\mu_\eta)$ *is a MLA index, and* $\widetilde{\gamma}(\mu_\eta) = -\dfrac{U''_+(\mu_\eta)}{U'_-(-\mu_\eta)}$ *is a pseudo Arrow-Pratt risk measure for sub-utility functions* U_+ *(gain) and* U_- *(loss) such that* $sgn(\widetilde{\gamma}(\mu_\eta)) = +1$ *for risk aversion and* $sgn(\widetilde{\gamma}(\mu_\eta)) = -1$ *for risk seeking.*

Remark 2.7.1. The right hand side of (2.7.14) collapses to Arrow-Pratt risk measure when there are no loss domains involved in the analysis. Since $\lambda_t(\mu_\eta)$ is undefined over gain domains we cannot say that $\mu_\eta \lambda_t(\mu_\eta) = 1$. This proposition is a manifestation of Wakker and Tversky (1993, Thm 9.1) and generalizes the example in Köbberling and Wakker (2005, p. 128) relating Arrow-Pratt risk measure to loss aversion.

Proof. See Appendix 2.B.8. □

This proposition shows that time varying risk aversion, in a context where DMs are faced with possible loss, is linked to the MLA index in our model. In particular, the α-stable properties of the MLA index implies that time varying risk aversion admits large values by and through the MLA index.

Procedure invariance and the price of consumption risk

Theoretically, the *observed risk premium* $\widehat{\pi}^E_{t+1}$, and observed asset returns r_{t+1} should be invariant to the RIH behavioural pricing kernel, and neoclassical pricing kernel, when the two models price the same asset. However, the covariance between excess returns and consumption growth is dispositive of the price of consumption risk (Breeden (1979, pp. 275-276), Duffee (2005, p. 1674)). Cursory inspection of (2.7.12) and (2.7.13) shows that the measure of covariance and risk aversion differ between the two asset pricing paradigms. So that is a source of pricing differential.

According to Proposition 2.7.2 under RIH behavioural asset pricing, risk aversion depends on the MLA index. Furthermore, we proved in Proposition 2.5.6 that the MLA index is a Cauchy random variable that admits large values. Thus, the RIH behavioural pricing kernel resolves the observed equity premium according to:

Proposition 2.7.3 (RIH resolution of equity premium puzzle)**.** *Let $\widehat{\pi}^E_{t+1}$ be the observed equity premium and $V(\widehat{r}_{t+1})$ be the variance of observed asset returns. $\lambda_t(\mu_\eta)$ is the MLA index Cauchy random variable under RIH for reference consumption level C_r with average consumption deviation from reference consumption μ_η. Then we have*

$$\widehat{\pi}^E_{t+1} = -\frac{1}{2}\widehat{V}(r_{t+1}) + c(\mu_\eta)\,\lambda_t(\mu_\eta)\,Cov_t\left(\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda}),\, r_{t+1}\right) \tag{2.7.15}$$

where $c(\mu_\eta) = \mu_\eta\, sgn(\widetilde{\gamma}(\mu_\eta))\,\widetilde{\gamma}(\mu_\eta)$ is a relative constant scale and $c(\mu_\eta)\,\lambda_t(\mu_\eta)$ is the price of consumption risk.

Proof. Substitute value of $\lambda_t(\mu_\eta)$ in Proposition 2.7.2 for $\lambda_t(c)$ in (2.7.12). □

According to (2.7.15) the empirical equity premium depends on a Cauchy random variable so it is susceptible to large swings. However, Charles-Cadogan (2016b) introduced a model which predicts that the loss aversion index is nonmeasurable–a prediction supported by the Cauchy random variable. In which case the empirical equity premium above may be nonmeasurable.

Anticipative risk attitudes and the covariance of equity premia

In the analysis that follow, we assume that the equity premium is positive, i.e., $\widehat{\pi}_{t+1} > 0$ so $r_{t+1} > 0$. Our focus is on the impact of risk attitudes on conditional covariance and the equity premium. In other words, we try to answer the question:

QUESTION:

What are the risk attitudes that support an investors' anticipation of a positive risk premium across states?

A more detailed analysis that accounts for statistical properties of λ_t (based on a comparison of the neoclassical and behavioural asset paradigms) is provided in subsection 2.7.4.

GG states

Here, our DMs is in a current "income gain state" and she anticipates income gain in the next period. This scenario plays out entirely in gain domains. According to (2.7.7) $\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda})$ depends on $\theta_{t+1}(\boldsymbol{g})$. Thus, $Cov_t\left(\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda}), r_{t+1}\right) > 0$. This implies that in Proposition 2.7.3 $\gamma_t^D(\mu_\eta) > 0$ and *our DM is risk averse*. This inference is consistent with prospect theory's prediction of risk aversion over gain domain almost surely.

GL states

In this case, our DM in current state gain, and she anticipates a loss in income in the next period. Now loss aversion kicks in and $\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda})$ depends on $-\lambda_{t+1}\,\theta_{t+1}(\boldsymbol{g})$. Our DM will demand a higher risk premium to buffer the shortfall in income and myopic consumption. If she expects $r_{t+1} > 0$, then $Cov_t\left(-\lambda_{t+1}\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda}), r_{t+1}\right) < 0$. In order for $\widehat{\pi}_{t+1} > 0$ we must have $\gamma_t^D(\mu_\eta) < 0$. That is, *DMs are risk seeking*. They decrease myopic consumption tracking income under Axiom 1 and invest as much as possible in the stock market hoping that it will rise (Curatola, 2015, p. 106).

LG states

Like the GL state, loss aversion kicks in in this state. Our DM is in current loss state and she anticipates a gain in income in the next period. Now $\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda})$ depends on $-\frac{1}{\lambda_t}\,\theta_{t+1}(\boldsymbol{g})$ and $Cov_t\left(-\frac{1}{\lambda_t}\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda}), r_{t+1}\right) < 0$. So *DMs are risk seeking* for the same reasons above.

LL states

This is perhaps the most complicated state of all for the following reason. Our DM is in current loss state, and she anticipates a loss in income in the next period. In this state $\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda})$ depends on $\frac{\lambda_{t+1}}{\lambda_t}\,\theta_{t+1}(\boldsymbol{g})$. However, the ratio of two Cauchy random variables has a probability distribution function, f, say, i.e., $\frac{\lambda_{t+1}}{\lambda_t} \sim f$. We show in section 2.7.4 below that f is approximated by a symmetric mollifier that admits negative values for the ratio relative to a central location. Nonetheless, our hypothesis is $Cov_t\left(\frac{\lambda_{t+1}}{\lambda_t}\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda}), r_{t+1}\right) > 0$ and $\gamma_t^D(\mu_\eta) > 0$ so our DM is risk averse. According to the fourfold pattern of risk attitudes in Tversky and Kahneman (1992) this is a low probability event. That is, even though DMs demand higher risk premia in LL states the probability that the covariance is positive is low. This implies that the more likely hypothesis is that the covariance is negative, and *DMs are risk seeking almost surely*. This would generate a positive counter-cyclical risk premium at the trough of the business cycle for LL states that characterize falling income.

2.7.4 Behavioural asset pricing vs. Neoclassical asset pricing over business cycle

One of the empirical regularities of the equity premium is its counter-cyclical. It is highest at the troughs and lower at the peaks of business cycles (Campbell and Cochrane, 1999; Campbell, 2003; Ludvigson, 2013). In this sub-section we show how the regularity is explained by our behavioural pricing kernel. In particular, we identify the role of the statistical distribution of the MLA index in affecting the process. We compare the asset prices predicted by the neoclassical pricing kernel to the asset prices predicted by our behavioural pricing kernel. In that way we identify differences between the two pricing paradigms.

Let $p_t^\star$ be the price of the asset predicted by the neoclassical pricing kernel in (2.7.3). Under RIH preferences DMs price the asset as p_t^D in (2.7.8). So their risk premium $p_t^D - b_t$ relative to a risk free asset b_t differs from the neoclassical risk premium $p_t^\star - b_t$. A first order approximation of the pricing kernel in (2.7.4) and (2.7.5) leads to the following price differential in equity premia

$$\Delta_t(p^\star, p^D) = p_t^\star - p_t^D \tag{2.7.16}$$

$$= E_t[\beta(1-\gamma g_{t+1}^c)x_{t+1}] - E_t[\beta(1-\gamma\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda}))x_{t+1}] \tag{2.7.17}$$

$$= -\beta\gamma E_t[(g_{t+1}^c - \varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda}))x_{t+1}] \tag{2.7.18}$$

The sign of the quantity inside the brackets determines whether the difference in predicted prices is positive or negative.

Asset prices under GL, LG income state transitions

The *growth rates* $\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda})$ *are negative for GL and LG state transitions* in (2.7.7). Hence $e^{-\gamma\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda})} = e^{\gamma|\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda})|} \gg e^{-\gamma g_{t+1}^c}$ and $M_{t+1}^D(\boldsymbol{g},\boldsymbol{\lambda}) \gg M_{t+1}$ where the symbol $\gg$ reads "much greater than". To see this, under myopia Axiom 1 consumption tracks income, we can write g^Y

instead of g^c. Suppose that anticipated state dependent growth rates $g_{GL}^Y < 0$ and $g_{LG}^Y > 0$. In GL states the Cauchy random variable component λ_{t+1} implies $\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda}) = -\lambda_{t+1}\theta_{t+1}(\boldsymbol{g}) - 1 \ll g_{GL}^Y$ which implies $\lambda_{t+1}\theta_{t+1}(\boldsymbol{g}) + 1 \gg g_{GL}^Y$ almost surely, since λ_{t+1} is restricted to the positive real axis and its variance very large. A similar argument holds for $-\frac{1}{\lambda_t}\theta_{t+1}(\boldsymbol{g}) - 1 \ll g_{LG}^Y$ which implies $\frac{1}{\lambda_t}\theta_{t+1}(\boldsymbol{g}) + 1 \gg -g_{LG}^Y$ in LG states. The following lemma guarantees the inequality $M_{t+1}^D(\boldsymbol{g},\boldsymbol{\lambda}) \gg M_{t+1}$.

Lemma 2.7.4 (Necessary RIH consumption growth condition)**.** *The consumption growth rate predicted by the neoclassical pricing kernel cannot exceed the growth rate predicted by the behavioural pricing kernel, i.e.,* $g_{t+1}^c \not\gg \varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda})$

Proof. See Appendix 2.B.7. □

Remark 2.7.2. According to Lemma 2.7.4 the volatility of RIH consumption growth is uniformly higher than that for neoclassical consumption growth. This observation is formally proven below in Proposition 2.7.10. It is a necessary condition for RIH to predict a larger risk premium than the neoclassical hypothesis. Cf. Wickens (2011, p. 203).

In GL and LG states $\varphi(\cdot) < 0$. So the quantity in (2.7.18) is negative, and investors with RIH preferences price assets much higher than their neoclassical counterparts, i.e., $p_t^D \gg p_t$, when either a gain in income is followed by an anticipated loss of income or a loss of income is followed by an anticipated gain in income. For a given benchmark asset b_t the equity premium is such that $p_t^D - b_t \gg p_t - b_t$.

Recall that Benartzi and Thaler take consumption as given in their model and focus solely on equity portfolios. Their assumption is equivalent to neutralizing consumption growth. In our model this can be achieved by setting $g_{t+1} = g_t = const$ and $g_{t+1}^Y = const$ so that M_{t+1}^D only depends on the *marginal distribution for* $\boldsymbol{\lambda}$. In which case the asset price p_t^D and equity premium $p_t^D - b_t$ only depend on the CRRA parameter γ and MLA vector $\boldsymbol{\lambda}$. Thus, our theory derives a Benartzi and Thaler (1995) type result, that under myopia the equity premium driven by MLA is much greater than that predicted by (Mehra and Prescott, 1985) neoclassical model. According to Propositions 2.5.3 and 2.5.6 the MLA index is an independent and identically distributed Cauchy random variable that takes values in the range $(0,\infty]$. This implies the following scenarios.

Scenario 1: $\boxed{\lambda_t > 1 \Longrightarrow \dfrac{1}{\lambda_{t+1}} < 1}$

Recall that the *iid* loss aversion index generated by the set $\mathscr{B}_{2k-1,2k}$ is used to replace and back fill observations in $\mathscr{B}^L_{2k-1}$ for analytic convenience. So for $\{\lambda_t, \lambda_{t+1}\} \in \mathscr{B}^L_{2k-1}$, we have $\lambda_t, = \lambda_{t+1}$ whereupon $\lambda_t > 1 \Longrightarrow \dfrac{1}{\lambda_{t+1}} < 1$. This implies

$$\varphi_{GL}(\cdot) \ll \varphi_{LG}(\cdot) \Longrightarrow p^D_{GL} \succ p^D_{LG} \tag{2.7.19}$$

where the subscript on behavioural growth rate φ denotes its value in that state transition, and $p^D_j,\ j \in \{GL, LG\}$ is the corresponding *anticipated price*. We interpret this scenario as one in which decision makers in GL states engage in risk seeking behaviour in anticipation of losses. So they require much larger equity premium to hold stocks. In which case they need stock prices to be much higher than that predicted by the neoclassical model. We believe that this is plausible because a recent experimental study by Doerrenberg et al. (2015) found that subjects who anticipated a loss of income held more risky assets.

Scenario 2: $\boxed{\lambda_t < 1 \Longrightarrow \dfrac{1}{\lambda_{t+1}} > 1}$

Similarly, $\lambda_t < 1 \Longrightarrow \dfrac{1}{\lambda_{t+1}} > 1$ implies

$$\varphi_{GL}(\cdot) \gg \varphi_{LG}(\cdot) \Longrightarrow p^D_{LG} \succ p^D_{GL} \tag{2.7.20}$$

We interpret this scenario as one in which decision makers in LG states engage in risk averse behaviour *relative* to GL states in anticipation of gains. They demand a higher risk premium to hold assets in order to compensate them for the recent lost they experienced.

Asset prices under GG income state transitions

Inspection of the GG state transitions in (2.7.5) shows that $\varphi_{t+1}(\boldsymbol{g}, \boldsymbol{\lambda})$ does not depend on the MLA index. Thus, for this relative income state with increasing income we have $Y_{t+1} > Y_t$ and $g^Y_{t+1} > 0$. Only the CRRA factor γ and growth rates for income affects the pricing kernel M^D_{t+1}. Substituting

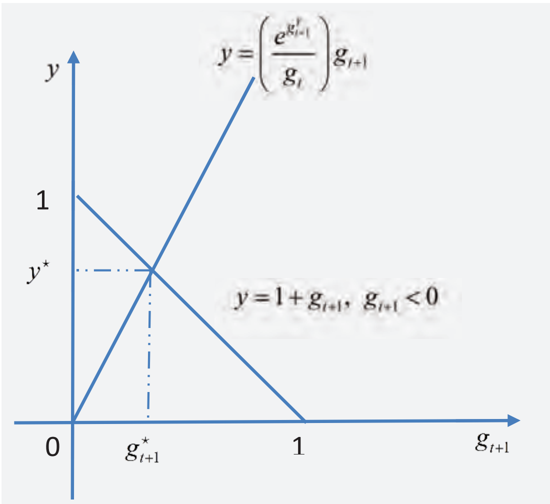
$\theta_{t+1}(\boldsymbol{g}) - 1$ for $\varphi_{t+1}(\boldsymbol{g}, \boldsymbol{\lambda})$ in (2.7.18) gives us the difference between neoclassical and RIH asset prices is $(p_t^\star - p_t^D)$ in GG state transitions:

$$\Delta_t(p^\star, p^D) = -\gamma\beta E_t[(g_{t+1} - \theta_{t+1}(\boldsymbol{g}) + 1)x_{t+1}] \tag{2.7.21}$$

$$= -\gamma\beta E\left[\left(-|g_{t+1}| - \frac{g_{t+1}}{g_t}e^{g^Y_{t+1}} + 1\right)x_{t+1}\right] \tag{2.7.22}$$

since $g_t < 0$, $g_{t+1} < 0$ by definition. In this case, the price differential is more nuanced.

Figure 2.33: Critical income growth for
GG state transitions



The critical growth $g_{t+1}^\star$ implies if $g_{t+1} > g_{t+1}^\star$, then $p_t^\star < p_t^D$. If $g_{t+1} < g_{t+1}^\star$, then $p_t^\star > p_t^D$.

Figure 2.33 shows that in income states GG, under RIH, the equity premium is greater than under neoclassical if income growth $g_{t+1} > g_{t+1}^\star$. It is lesser if $g_{t+1} < g_{t+1}^\star$. We characterize this result as follows.

Lemma 2.7.5 (Critical income growth rate)**.** *There exist a critical income growth rate $g_{t+1}^\star$ which, if exceeded, then the equity premium predicted by RIH is much greater than that predicted by the neoclassical model in states of rising income.* □

Asset prices under LL income state transitions

For this relative income group with *decreasing income*, we have $Y_{t+1} < Y_t \Rightarrow g^Y_{t+1} < 0$. The CRRA factor γ, growth rates $\boldsymbol{g}$ for consumption, and growth rate for loss aversion index affect the pricing kernel M^D_{t+1}. Substitute $\frac{\lambda_{t+1}}{\lambda_t}\theta_{t+1}(\boldsymbol{g}) - 1$ for $\varphi_{t+1}(\boldsymbol{g},\boldsymbol{\lambda})$ in (2.7.18) to get

$$\Delta_t(p^\star, p^D) = -\gamma\beta E_t[(g_{t+1} - \frac{\lambda_{t+1}}{\lambda_t}\theta_{t+1}(\boldsymbol{g}) + 1)x_{t+1}] \tag{2.7.23}$$

$$= -\gamma\beta E_t\Big[\left(\Big(1 - \frac{\lambda_{t+1}}{\lambda_t}\frac{e^{g^Y_{t+1}}}{g_t}\Big)g_{t+1} + 1\right)x_{t+1}\Big] \tag{2.7.24}$$

The quantity inside the bracket is nuanced (recall $g_t < 0$, $g_{t+1} < 0$) so the sign of $\Delta_t(p^\star, p^D)$ has to be analyzed. We posit two cases.

Scenario: $\boxed{\frac{\lambda_{t+1}}{\lambda_t} \mapsto \frac{\lambda_{t_{2k-1,2k}}}{\lambda_{t_{2k-3,2k-2}}}}$

The analysis here pertains to loss followed by anticipated loss. In the context of our loss aversion generator $\mathscr{B}_{2k-1,2k}$ the ratio $\frac{\lambda_{t+1}}{\lambda_t} = 1$ for LL states in $\mathscr{B}^L_{2k-1}$ and (2.7.7). However, $\lambda_t, \lambda_{t+1} \sim iid\ \mathscr{C}$. Thus, the ratio behaves as if it was constructed from adjacent blocks $\mathscr{B}_{2k-1,2k}$ and $\mathscr{B}_{2k-3,2k-2}$. In which case, we have $\frac{\lambda_{t+1}}{\lambda_t} \mapsto \frac{\lambda_{t_{2k-1,2k}}}{\lambda_{t_{2k-3,2k-2}}}$. Therefore, we will treat the ratio $\frac{\lambda_{t+1}}{\lambda_t} = 1$ as a given value for the ratio of two Cauchy random variables by virtue of Proposition 2.5.3. If λ has a Cauchy p.d.f $f(x) = \frac{a}{\pi(a^2+x^2)}$, then the ratio of two Cauchy random variables is known to have a distribution of type $f(y) = \frac{a^2}{\pi^2(y^2-a^4)}\ln\left(\frac{y^2}{a^4}\right)$, $-\infty < y < \infty$ (Springer, 1979, p. 158). According to Definition 2.2.2, $f(y)$ admits a candidate mollifier on $[-a^2, a^2]$ since $f(a^2) = f(-a^2) = 0$. We consider two cases: **Case A** where the neoclassical price is greater than the behavioural price, and *vice versa* for **Case B**.

Case A: $\boxed{p^\star > p^D}$

If $p_t^\star > p_t^D$, then we have

$$\left(1 - \frac{\lambda_{t2k-1,2k}}{\lambda_{t2k-3,2k-2}} \frac{e^{g_{t+1}^Y}}{g_t}\right) g_{t+1} + 1 < 0 \tag{2.7.25}$$

$$\Longrightarrow \frac{\lambda_{t2k-1,2k}}{\lambda_{t2k-3,2k-2}} > \left(1 + \frac{1}{g_{t+1}}\right) g_t . e^{-g_{t+1}^Y} = \left(\frac{|g_t|}{|g_{t+1}|} - |g_t|\right) e^{-g_{t+1}^Y} \tag{2.7.26}$$

where $g_t < 0$, $g_{t+1} < 0$, $|g_{t+1}| < 1$ and the absolute values are used to distinguish the signs of values.

Case B: $\boxed{p^\star < p^D}$

In this case the inequality in (2.7.26) changes so that

$$\frac{\lambda_{t2k-1,2k}}{\lambda_{t2k-3,2k-2}} < \left(\frac{|g_t|}{|g_{t+1}|} - |g_t|\right) e^{-g_{t+1}^Y} \tag{2.7.27}$$

Impact of mollifier for ratio of two Cauchy random variables

Let $\left(\frac{|g_t|}{|g_{t+1}|} - |g_t|\right) e^{-g_{t+1}^Y} = v$. The relationship in (2.7.26) and (2.7.27) implies that

$$\left|\frac{\lambda_{t2k-1,2k}}{\lambda_{t2k-3,2k-2}}\right| < v \tag{2.7.28}$$

We are only interested in the positive values of the ratio. So the foregoing relationships leads to the following

Lemma 2.7.6 (Impact of ratio of two MLA indexes on risk premium)**.** *Let* $f(y) = \frac{a^2}{\pi^2(y^2 - a^4)} \ln\left(\frac{y^2}{a^4}\right)$, $-\infty < y < \infty$ *and* $f(y) = 0$ *otherwise, be the pdf for the ratio* $y = \frac{\lambda_{t2k-1,2k}}{\lambda_{t2k-3,2k-2}}$ *of two independent MLA indexes, i.e., two independent Cauchy random variables. We consider the mollifier for the restricted range* $-a^2 \le y \le a^2$ *and* $f(y) = o$ *otherwise. If* $y \notin [v, a^2]$*, then* $p_t^\star < p_t^D$*. If* $y \in [v, a^2]$*, then* $p_t^\star > p_t^D$*. The latter condition fails empirically if* $v < 0$*. In which case* $p_t^\star < p_t^D$ *almost surely.* □

Suppose $\frac{\lambda_{t+1}}{\lambda_t} = 1$. According to Lemma 2.7.6 and the fourfold pattern of risk attitudes in Tversky and Kahneman (1992), if $\Pr\{1 \in [v, a^2]\}$ is small, then DMs are risk averse over losses. In this case $p_t^\star > p_t^D$ and neoclassical prices are higher than RIH prices. So risk aversion is very high under RIH in this case. On the other hand, $\Pr\{1 \notin [v, a^2]\}$ is large so DMs are risk seeking over losses under the fourfold pattern of risk attitudes, and $p_t^\star < p_t^D$. The small probability associated with risk aversion over losses implies that DMs are risk seeking over losses almost surely. So that $p_t^\star < p_t^D$ almost surely.

Figure 2.34: PDF bump function truncated at $a^2 = \pm 1$ for ratio of two MLA indexes

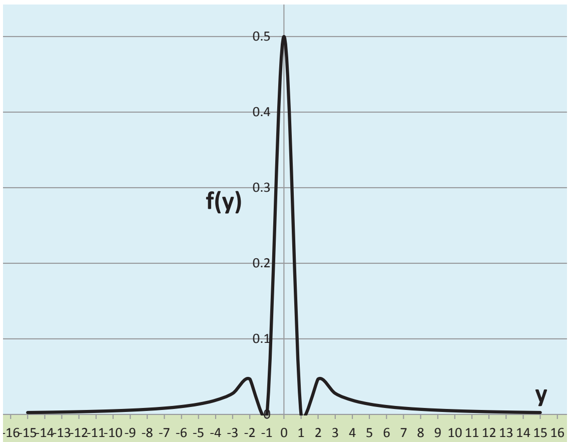
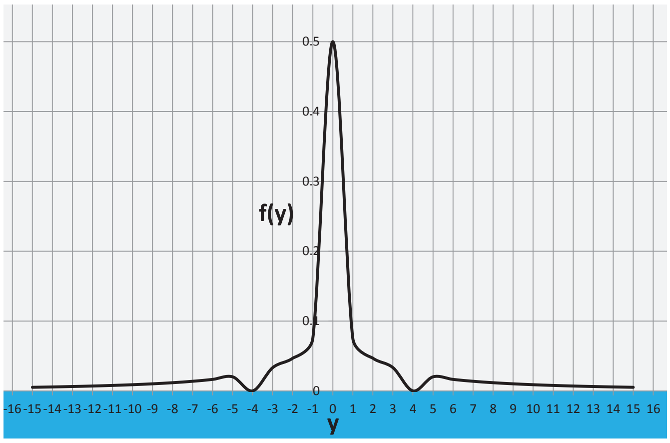


Figure 2.35: PDF bump function truncated at $a^2 = \pm 4$ for ratio of two MLA indexes



Plot depicts the probability distribution function for the ratio $y = \dfrac{\lambda_{t+1}}{\lambda_t}$ where

$$f(y) = \frac{a^2}{\pi^2(y^2 - a^4)} \ln\left(\frac{y^2}{a^4}\right), \quad -\infty < y < \infty$$

The curve has unbounded kurtosis, i.e., $f(y) \uparrow \infty$, $y^+ \downarrow 0$. The sharp peak was arbitrarily chosen at $f(0) = 0.5$ which is 500 times a probability unit of 0.001 in order to satisfy software requirements for the plot. We impose the mollifier condition $a = 1$ in Figure 2.34 and $a = 2$ in Figure 2.35. That removes the tail probabilities for $|y| \geq 1$ and $|y| \geq 4$ as indicated.

Figure 2.34 and Figure 2.35 depict a plot of the probability distribution for $f(y)$ in Lemma 2.7.6 when $a = 1$ and $a = 2$, respectively. When $a = 1$, $\Pr\{1 \in [v, a^2]\} = 0$ in Figure 2.34. However, in Figure 2.35, $\Pr\{0 < y < 1\}$ is much larger than $\Pr\{1 \leq y \leq 4\}$ by virtue of the asymptote at 0. So $p_t^\star < p_t^D$ uniformly under Figure 2.34 and DMs are risk seeking over losses. When $a = 2$, $p_t^\star < p_t^D$ almost surely under Figure 2.35 since the associated probability of that event is much larger than that for the event $p_t^\star > p_t^D$. Hence DMs are risk seeking almost surely in LL states.

2.7.5 Probabilistic risk attitudes towards the Markov state transition matrix

Let $q_j,\ j \in \mathscr{I}$ be the j-th transition probability such that $q_{LL} + q_{LG} = 1$ and $q_{GL} + q_{GG} = 1$. Recall that in Definition 2.7.1 $\mathscr{I} = \{LL, LG, GG, GL\}$ is an index set for the possible gain loss states $\{G, L\} \times \{G, L\}$. Furthermore, let $w^+(\cdot)$ be the probability weight assigned over gain domains and $w^-(\cdot)$ be the probability weighting function of loss domain.

We introduce two synthetic lotteries $L_g = (r_{GG}, q_{GG}; r_{GL}, q_{GL})$ and $L_\ell = (r_{LL}, q_{LL}; r_{LG}, q_{LG})$ where $r_j\ j \in \mathscr{I}$ corresponds to the j-th state dependent asset returns. L_g is comprised of *anticipated returns* when the current state is "gain", and L_ℓ is comprised of *anticipated returns* when the current state is "loss". Assume that utility is linear in asset returns so that risk attitudes are driven by probability weighting (Yaari, 1987). We assume that if a DM is in gain state G and she anticipates a gain state G in the next period, then she weights probability according to $w^+(q_{GG})$. If she anticipates a loss L, then she employs $w^-(q_{GL})$. Similarly, if she is in loss state, and anticipates a loss state then the assigned weight is $w^-(q_{LL})$. If she anticipates a gain state from a loss state, then the assigned weight is $w^+(q_{LG})$. These assumptions are consistent with probability weighting in two-outcomes lotteries (Quiggin, 1993, p. 57). The state transition probabilities are given by

$$q_{GG} = \Pr[\Delta^b C^D_{t+1} = g_{t+1} Y_{t+1} | \Delta^b C^D_t = g_t Y_t] \tag{2.7.29}$$

$$q_{LG} = \Pr[\Delta^b C^D_{t+1} = g_{t+1} Y_{t+1} | \Delta^b C^D_t = -\lambda_t g_t Y_t] \tag{2.7.30}$$

$$q_{GL} = \Pr[\Delta^b C^D_{t+1} = -\lambda_{t+1} g_{t+1} Y_{t+1} | \Delta^b C^D_t = g_t Y_t] \tag{2.7.31}$$

$$q_{LL} = \Pr[\Delta^b C^D_{t+1} = -\lambda_{t+1} g_{t+1} Y_{t+1} | \Delta^b C^D_t = -\lambda_t g_t Y_t] \tag{2.7.32}$$

These probabilities satisfy the requirements for a state transition probability matrix

$$Q = \begin{pmatrix} q_{GG} & q_{GL} \\ q_{LG} & q_{LL} \end{pmatrix} \tag{2.7.33}$$

Q is ergodic because every state can be attained from any other state. Mehra and Prescott (1985, pp. 150-151) assumed the existence of a symmetric Markov matrix, i.e., state transition probability matrix in their model.[36] Here we show how it is a natural artefact of RIH preferences. In the simple synthetic lottery setup, we have the 2-states probability weighting relationship (Quiggin, 1993, p. 57)

$$w^{+}(q_{GG})+w^{-}(q_{GL})=1 \tag{2.7.34}$$

$$w^{+}(q_{LG})+w^{-}(q_{LL})=1 \tag{2.7.35}$$

The 2×2 nature of Q implies the existence of a probability weighted transition matrix

$$w(Q)=\begin{pmatrix} w^{+}(q_{GG}) & w^{-}(q_{GL}) \\ w^{+}(q_{LG}) & w^{-}(q_{LL}) \end{pmatrix} \tag{2.7.36}$$

We make the following

Assumption 2.7.7 (Preference for good states)**.** *Decision makers prefer good states over bad states.*

Under the hypothesis that DMs prefer good states to bad, we have $L_g \succ L_\ell$. So that DMs evaluate the lotteries such that

$$w^{+}(q_{GG})r_{GG}+w^{-}(q_{GL})r_{GL} > w^{+}(q_{LG})r_{LG}+w^{-}(q_{LL})r_{LL} \tag{2.7.37}$$

Following Mehra and Prescott (1985), assume that Q in (2.7.33) is symmetric so that $q_{GL}=q_{LG}=1-\phi$ and $q_{GG}=q_{LL}=\phi$. The relationship in (2.7.37) is reduced to

$$w^{+}(\phi)r_{GG}+(1-w^{+}(\phi))r_{GL} > (1-w^{-}(\phi))r_{LG}+w^{-}(\phi)r_{LL} \tag{2.7.38}$$

$$\Longrightarrow r_{GL}+w^{+}(\phi)(r_{GG}-r_{GL}) > r_{LG}+w^{-}(\phi)(r_{LL}-r_{LG}) \tag{2.7.39}$$

$$\Longrightarrow \frac{w^{-}(\phi)}{w^{+}(\phi)} > \frac{r_{GL}-r_{GG}}{r_{LG}-r_{LL}} \tag{2.7.40}$$

Assuming that assets are priced under more likely (2.7.19) in Scenario 1, and more likely (2.7.27)

[36] Lucas (1978, p. 1431) also used a Markov transition function in his model.

for Case B, we get the asset price ranking $r_{GL} \succ r_{LG} \succ r_{LL} \succ r_{GG}$. In which case (2.7.40) implies

$$\frac{w^-(\phi)}{w^+(\phi)} > 1 \Longrightarrow w^-(\phi) = w^+(\phi) + \Delta_1, \qquad w^+(1-\phi) = w^-(1-\phi) + \Delta_2 \tag{2.7.41}$$

where Δ_1 and Δ_2 are equating values that remove the inequality. Substitution of these values in (2.7.36) leads to

$$w(Q) = \begin{pmatrix} w^+(\phi) & w^-(1-\phi) \\ w^+(1-\phi)+\Delta_2 & w^-(\phi)+\Delta_1 \end{pmatrix} \tag{2.7.42}$$

The identifying restriction in (2.7.36) and (2.7.42) are such that

$$w^+(1-\phi) + \Delta_2 + w^-(\phi) + \Delta_1 = 1 \Longrightarrow \Delta_2 + \Delta_1 = 0 \tag{2.7.43}$$

Source of excess volatility and counter-cyclical equity premia

There are two scenarios implied by (2.7.42) and (2.7.43):

$$w^+(q_{LG}) = w^+(1-\phi) + \Delta_2, \qquad w^-(q_{LL}) = w^-(\phi) - \Delta_2 \tag{Optimistic}$$

$$w^+(q_{LG}) = w^+(1-\phi) - \Delta_1, \qquad w^-(q_{LL}) = w^-(\phi) + \Delta_1 \tag{Pessimistic}$$

To help interpret the results above we introduce the following

Definition 2.7.2 (Stochastic dominance)**.** Montes et al. (2013). Given two cumulative distribution functions F and G, we say that F stochastically dominates G, and denote it $F \succeq_{FSD} G$, if $F(t) \leq G(t)$ for every $t \in [0,1]$, and given two random variables X, Y taking values on $[0,1]$, we say that X stochastically dominates Y , and denote it $X \succeq_{FSD} Y$ when its associated distribution function F_X stochastically dominates F_Y, where $F_X(t) = P(X \leq t)$ and $F_Y(t) = P(Y \leq t)$ $\forall t \in [0,1]$. □

In the optimistic scenario DMs shift probability mass from LL states to LG states. So r_{LG} stochastically dominates r_{LL} (Machina, 1987a, p. 236). They underweight the probability of LL states and overweight the probability of LG states. In the pessimistic scenario, they shift mass from LG state to LL states. So r_{LL} stochastically dominates r_{LG} in that case. Thus, we prove

Lemma 2.7.8 (Asymmetric $w(Q)$)**.** *The weighted transition probability matrix $W(Q)$ is asymmetric because DMs shift probability mass within good and bad states to reflect stochastic dominance of anticipated state dependent probability distributions.* □

The probabilistic risk attitude associated with pessimism (or risk aversion) is characterized by a concave (or concave-convex) probability weighting function. The probabilistic risk attitude associated with optimism (or risk seeking) is characterized by a convex (or convex-concave) probability weighting function. On the basis of our evaluations in (2.7.19), (2.7.20), and $w(Q)$ in Lemma 2.7.8 we posit the following

Proposition 2.7.9 (Endogenous rankings of anticipated asset prices)**.** *The behavioural pricing kernel induces an endogenous ranking of assets:* $\underbrace{p^D_{LL} \prec p^D_{GG} \prec p^D_{LG} \prec p^D_{GL}}_{\text{risk aversion}}$ *or* $\underbrace{p^D_{GG} \prec p^D_{LL} \prec p^D_{GL} \prec p^D_{LG}}_{\text{risk seeking}}$ *where the subscripts correspond to the* anticipated price *of the asset generated by the growth rate in* (2.7.7)*. In particular, the rankings accommodate the outcome dependent probability weighting functions sketched in Figure 2.8 and Figure 2.9.* □

Remark 2.7.3. The analyses above accommodate different permutations of prices and hence various shapes of the corresponding probability weighting functions (Stewart et al., 2015). So in a sense the proposition describes two of many possible sample function shapes. This proposition and other results presented above show that the behavioural pricing kernel is nonlinear (Shefrin, 2009, p. 86).

The implications of probability mass shifting in the asymmetric transition probability matrix $w(Q)$ can be seen in how DMs evaluate state dependent average returns and volatility. Average returns predicted by the behavioural pricing kernel are lower than that predicted by the neoclassical kernel in good states. In contrast, in bad states average returns (and hence average equity premium) predicted by the behavioural pricing kernel is higher than that predicted by the neoclassical pricing kernel. Furthermore, volatility is uniformly higher under the behavioural pricing paradigm. Thus, the behavioural model captures the counter-cyclical behaviour of equity premia. We summarise these results in the following.

Proposition 2.7.10 (Stochastic dominance, probabilistic risk aversion, and excess volatility)**.** *Assuming that DMs prefer good states over bad states, if the transition probability matrix Q is symmetric, then the probability weighted transition matrix $w(Q)$ is asymmetric. In bad states, the behavioural pricing kernel predicts large risk premia and excess volatility are caused by stochastic dominance of anticipated good states. In good states, the behavioural pricing kernel predicts lower average equity premium, and excess volatility due to probabilistic risk aversion. The equity premium is counter-cyclical over the business cycle peak and troughs for good and bad states.*

Proof. See Appendix 2.B.9. □

2.8 Conclusion

This paper fills a gap in the literature by introducing an asymptotic theory of prospect theory's myopic loss aversion (MLA) index. It produces several new results that are important in their own right. We prove that the MLA index is α-stable, and that it follows a generalized Cauchy law in most cases. Because a Cauchy process belongs to the class of Lévy processes, our theory expands the solution space for loss aversion to include its embedding in Lévy type processes.

We provide a simple procedure for estimating the MLA index from economic time series data. Whereupon we applied it to US and South Africa income and consumption data. A distribution of macroeconomic MLA indexes was computed for each country, and a battery of statistical tests upheld the α-stable law prediction of our theory. We checked for robustness of the theory by applying it to different domains such as MLA index data from around the world, and MLA index data from a meta study. In each case, the theory was upheld.

An independently important result of our theory is that the MLA index mimics positive jumps of a Cauchy process. We show how this can be applied to information based asset pricing when the information process is a Cauchy bridge. Jumps in the bridge process mimic bad news and loss aversion and the asset price falls. Thus, we contribute to the strand of information based asset pricing.

We employed a novel specification of the relative income hypothesis (RIH) to embed the MLA index in a behavioural pricing kernel in the consumption based asset pricing model

(CCAPM). We show how this resolves the equity premium puzzle. Theory predicts a much larger equity premium than that obtained from the neoclassical pricing kernel. Consistent with the empirical literature, our RIH behavioural pricing kernel is nonlinear. In what appears to be new to the literature, our model predicts that the price of consumption risk includes myopic loss aversion to decline in consumption. The Cauchy random variable attribute of the MLA index predicts wide swings in the price of risk observed in the empirical literature. Moreover, we prove that the behavioural pricing kernel induces an endogenous ranking of assets supported by outcome dependent probability weighting functions. In fact, we prove that DMs who prefer gain states to loss states shift probability mass from loss states to gain states. This stochastic dominance behaviour distorts the state transition probability matrix, and induces a uniformly higher risk premium and asset price volatility than that predicted by neoclassical models.

Further research includes identifying the small sample properties of the MLA index estimator to facilitate statistical inference in economic experiments. A natural extension of our results is to cross-country data in order to identify and compare cross-cultural risk attitudes. In related work, we show that cross sectional regressions of economic growth on subjective well being are misspecified by virtue of simultaneity bias induced by the MLA index embedded in the growth series. This result has implications for economic growth policies formulated on the basis of such models. For if the parameter to be estimated in drawn from an α-stable process, analysts are likely to get a "policy surprise" when observed values of the purported parameter are much larger than that predicted by their models. Preliminary results show that the model makes empirically testable predictions about leverage effects that depend on risk aversion, loss aversion, and income and consumption growth in cross-sectional asset pricing.

2.A APPENDIX–EMBEDDING MLA INDEX IN PRICING KERNEL

2.A.1 State dependent pricing kernels and loss aversion embedding

This appendix provides the basic idea behind how we embed the MLA index in our behavioural pricing kernel. We assume a single good pure exchange economy. In what follows "A state contingent consumption claim is a security that pays one unit of the consumption good when one particular state of the world occurs and nothing otherwise. A state contingent claim is an elementary claim. Existing assets, however, may be viewed as complex bundles of elementary claims." (Huang and Litzenberger, 1988, p. 119). For example, we consider a DM faced with a state contingent claim on consumption over two periods: the present 0, and future 1, with the following primitives:

Primitives

Ω	the set of all possible states of nature
ω	an element of the set Ω
c_0	present period 0 certain consumption
$c_1(\omega)$	next period 1 uncertain consumption
e_0	present period certain endowment
$e_1(\omega)$	next period's uncertain endowment
ϕ_ω	Arrow security state price of insurance per unit of consumption
β	subjective discount factor
$u(c_0, c_1(\omega); \beta)$	utility over consumption plan over periods 0, 1
π_ω	DM's subjective probability about the occurrence of state $\omega \in \Omega$

The contingent claim pays 1 if state ω occurs in period 1 and 0 otherwise. So that for $c_1(\omega)$ units of state contingent consumption, the total price is $\phi_\omega c_1(\omega)$. Similarly, the total price for state contingent endowment is $\phi_\omega e_1(\omega)$. For the sake of exposition in what follows we assume that utility is nonseparable. We relax that assumption later. We note that Epstein and Zin (1989) introduced a class of flexible nonseparable utility functions that are used to derive pricing kernels. Those functions admit separability as a special case with constant elasticity of substitution (CES). The expected utility of the consumption plan is $\sum_{\omega\in\Omega} \pi_\omega u(c_0, c_1(\omega); \beta)$. So our DM faces the optimization problem

$$\begin{aligned} &\max_{c_0, c_1(\omega)} \sum_{\omega\in\Omega} \pi_\omega u(c_0, c_1(\omega); \beta) \\ &\text{s. t. } c_0 + \sum_{\omega\in\Omega} \phi_\omega c_1(\omega) = e_0 + \sum_{\omega\in\Omega} \phi_\omega e_1(\omega) \end{aligned} \tag{2.A.1}$$

The Langrangian (L) for this problem is given by

$$L = \sum_{\omega\in\Omega} \pi_\omega u(c_0, c_1(\omega); \beta) - \mu\Big(c_0 + \sum_{\omega\in\Omega} \phi_\omega c_1(\omega) - e_0 - \sum_{\omega\in\Omega} \phi_\omega e_1(\omega)\Big) \tag{2.A.2}$$

$$= \sum_{\omega\in\Omega} \pi_\omega u(c_0, c_1(\omega); \beta) - \mu(c_0 - e_0) - \mu\Big(\sum_{\omega\in\Omega} \phi_\omega (c_1(\omega) - e_1(\omega))\Big) \tag{2.A.3}$$

The necessary first order conditions (FOCs) for optimality in (2.A.3) is applied to each period. In period 0 our DM does not know which ω state will occur so she takes the expected value over all possible states Ω and the FOC is applied there. In period 1, the state is revealed so there is no need for expected values and she takes the FOC there. That process is characterized by:

$$\frac{\partial L}{\partial c_0} = 0 \Longrightarrow \sum_{\omega\in\Omega} \pi_\omega \frac{\partial}{\partial c_0} u(c_0, c_1(\omega); \beta) = \mu \tag{2.A.4}$$

$$\frac{\partial L}{\partial c_1(\omega)} = 0 \Longrightarrow \pi_\omega \frac{\partial}{\partial c_1(\omega)} u(c_0, c_1(\omega); \beta) = \mu\phi_\omega \tag{2.A.5}$$

$$\Longrightarrow \phi_\omega = \frac{\pi_\omega \frac{\partial}{\partial c_1(\omega)} u(c_0, c_1(\omega); \beta)}{\sum_{\omega\in\Omega} \pi_\omega \frac{\partial}{\partial c_0} u(c_0, c_1(\omega); \beta)} \tag{2.A.6}$$

Since the contingent claim is priced ϕ_ω per unit of consumption, and it pays 1 only if the state ω occurs or 0 otherwise, the total price for consumption in a given state is a simple function given by

$$P(c) = \sum_{\omega\in\Omega} \phi_\omega c_1(\omega) \Longrightarrow \sum_{\omega\in\Omega} \pi_\omega \left(\frac{\phi_\omega}{\pi_\omega}\right) c_1(\omega) = \sum_{\omega\in\Omega} \pi_\omega M_\omega c_1(\omega) \tag{2.A.7}$$

$$= E[M_\omega c_1(\omega)], \quad M_\omega = \frac{\phi_\omega}{\pi_\omega} \tag{2.A.8}$$

where M_ω is the "pricing kernel" for state contingent consumption. Following Lucas (1978), an asset is a claim to all the output in our simple exchange economy. So $P(c)$ is a consumption based asset price. Dividing the left hand and right hand side of 2.A.7 by $P(c)$ we get

$$1 = \sum_{\omega\in\Omega} \pi_\omega M_\omega \frac{c_1(\omega)}{P(c)} = \sum_{\omega\in\Omega} \pi_\omega M_\omega x_1(\omega) \tag{2.A.9}$$

$$= E[M_\omega x_i(\omega)] \tag{2.A.10}$$

where $x_1(\omega)$ is the total return on assets, i.e., $x_1(\omega) = 1 + r_1(\omega)$. If the risk free rate is given by r_f, then (2.A.10) also prices the risk free asset. Substituting $1 + r_f$ for $x_1(\omega)$, and using the definition of covariance, gives us $E[M_\omega]$ and the equity premium relationship

$$E[r_1(\omega)] - r_f = (1 + r_f)Cov(M_\omega, x_1(\omega)) \tag{2.A.11}$$

Thus, the equity premium is driven by the covariance structure of the pricing kernel and $x_1(\omega)$ which serves double duty as a consumption growth relation. In (2.A.6) we proved that the Arrow state price of insurance ϕ_ω is related to the intertemporal marginal rate of substitution (IMRS) between c_0 and $c_1(\omega)$.

The problem is simplified by assuming that utility is separable with a subjective time discount factor β such that

$$u(c_0, c_1(\omega); \beta) = u(c_0) + \beta E[u(c_1(\omega))] \tag{2.A.12}$$

$$\frac{\partial}{\partial c_0} u(c_0, c_1(\omega); \beta) = u'(c_0), \qquad \frac{\partial}{\partial c_1(\omega)} u(c_0, c_1(\omega); \beta) = \beta E[u'(c_1(\omega))] \tag{2.A.13}$$

$$\phi_\omega = \frac{\pi_\omega \beta E[u'(c_1(\omega))]}{u'(c_0)}, \qquad M_\omega = \frac{\phi_\omega}{\pi_\omega} = \beta \frac{E[u'(c_1(\omega))]}{u'(c_0)} \tag{2.A.14}$$

Thus, the pricing kernel in (2.A.14) includes a time discount factor and DM's subjective expectation about future consumption.

Embedding myopic loss aversion

Under Tversky and Kahneman (1992) CPT the arguments of the utility function are gains and losses relative to a pre-specified reference point (Shefrin, 2009, p. 17). So DMs care about changes in consumption relative to a reference consumption level–not about terminal consumption. Moreover, they have "value functions" to distinguish this utility attitude from classic utility functions. Furthermore, the value function is bifurcated into sub-utility functions over gains and loss relative to a reference consumption level c_r and its parameterizations includes a loss aversion index λ. Thus, we replace the utility function u by a "carrier of value" $v(\cdot|\lambda)$ for consumption, we replace $c_1(\omega)$ with $c_1(\omega) - c_r$, and c_0 with $c_0 - c_r$. Note that in our model the *loss aversion index is inside the utility function* as opposed to extant literature where it lies outside the utility function. So that we get the "relative consumption" pricing kernel

$$M_{\omega}^{D}(\lambda) = \beta \frac{E[v'((c_1(\omega) - c_r)|\lambda)]}{v'((c_0 - c_r)|\lambda)} \tag{2.A.15}$$

If $(c_1(\omega) - c_r) < 0$, and $(c_0 - c_r) > 0$, then according to (2.2.7) $-(c_1(\omega) - c_r) = \lambda(\omega)(c_0 - c_r) > 0$. Substitution in (2.A.15) gives us

$$M_{\omega}^{D}(\lambda) = \beta E \left(\frac{\frac{\partial}{\partial c_1(\omega)}[v(-(c_1(\omega) - c_r))]}{\frac{\partial}{\partial c_0}[v(c_0 - c_r)]} \right) \tag{2.A.16}$$

$$= \beta \frac{E[v'(\lambda(\omega)(c_0 - c_r))]}{v'((c_0 - c_r))} \tag{2.A.17}$$

Thus, we embed the loss aversion index in a behavioural pricing kernel. $M_{\omega}^{D}(\lambda)$ inherits the α-stable property of the MLA index random variable $\lambda(\omega)$. So it is subject to large fluctuations. We note in passing that the pricing kernel also depends on subjective probability assessment π_{ω} which admits a probability weighting function (pwf) $w(\cdot)$. So that formation of expectations in (2.A.17) is amenable to pwf analysis. The result in (2.A.17) extends readily to Kőszegi and Rabin (2006) reference dependent preference model.

Anticipating Gain (G) and Loss (L) states

In order for loss aversion to affect our analysis, a DM must either be in a current loss state, i.e., $c_0 - c_r < 0$ or in an anticipating loss state, i.e., $c_1(\omega) - c_r < 0$ or both. To simplify the analysis, we will assume a *current* gain state G and an *anticipating* loss state L.

A probabilistic approach. Suppose that $\Omega \triangleq \{G, L\}$. In what follows an up-arrow $\uparrow$ signifies that the associated quantity goes up, a down-arrow $\downarrow$ signifies that it goes down. Let $w(\cdot)$ be our DMs probability weighting function. For a given price $P(c,\lambda) = E[M_{\omega}(\lambda)c_1(\omega)]$, if our DM anticipates state L so that $c_1(\omega)\downarrow$, then in order to maintain $P(c,\lambda)$ we must have $M_{\omega}(\lambda)\uparrow = \frac{\phi_{\omega}}{\pi_{\omega}}\uparrow$. The latter condition is satisfied if $\pi_{\omega}\downarrow$ or $\phi_{\omega}\uparrow$ or both. That is, the Arrow security price of insurance, i.e., risk premium, goes up in bad states or DMs revise their subjective probability downwards in bad states or both.

Suppose that ϕ_ω, the Arrow security price of insurance goes up in bad states. Then $P(c)$ goes up in (2.A.7). If our DM revised her subjective probability downwards, i.e., $w^-(\pi_\omega) < \pi_\omega$, she is more optimistic about L state, then the anticipated given price of the asset goes up higher than that predicted by ϕ_ω alone. We represent this with a double arrow $P(c) \uparrow\uparrow$. Thus, our DM is risk seeking in this state, and the risk premium associated with $P(c)$ is in excess of that predicted by the Arrow security price of insurance alone. The complementary event is if our DM anticipates $c_1(\omega) \uparrow$. Then, $M_\omega \downarrow$ and either $\phi_\omega \downarrow$ or $(1-\pi_\omega) \uparrow$ or both. Assuming both, our DM revises her subjective probability upwards and $w^+(1-\pi_\omega) > 1-\pi_\omega$. Thus, the risk premium associated with $P(c) \downarrow\downarrow$ is much lower than that predicted by ϕ_ω alone. In this two-state case $w^-(\pi_\omega) + w^+(1-\pi_\omega) = 1$ and our DM is optimistic because she revises her probability to underweight bad states, and overweight good states. This example is not exhaustive of the taxonomy of behaviours that is accommodated by combinations of ϕ_ω and π_ω in M_ω.

A utility based approach. In this case $M_\omega(\lambda) \uparrow$ implies $E[v'(\lambda(\omega)(c_0 - c_r))] \uparrow$. This implies that the argument inside the utility function is decreasing. Since $c_0 - c_r$ is fixed, it implies that $\lambda(\omega) \downarrow$. Our DM is becoming less loss averse as she anticipates $c_1(\omega) \downarrow$ in L states. This is the utility analog of risk seeking over losses. Since $\lambda(\omega)$ is α-stable, its value can drop dramatically. According to the fourfold pattern, if there is a high probability of a large drop in loss aversion when consumption drops, then DMs will be risk seeking. For $\lambda(\omega) < 1$ this behaviour is called gain-seeking (Wakker, 2010).

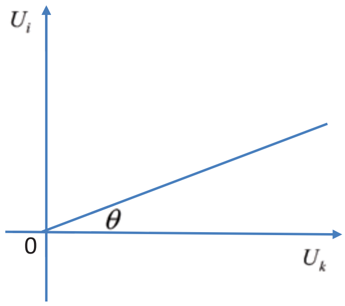
2.B APPENDIX OF PROOFS

2.B.1 Proof of Theorem 2.2.4 Standard Cauchy Spherically Symmetric

Proof. The following is a modified proof of Arnold and Brockett (1992, Thm. 1). Let $F(t) = \Pr\{U_i/U_k \leq t\}$. By symmetry, we need only consider the first quadrant $U_i > 0, U_k > 0$. Accordingly, $\Pr\{U_i/U_k \leq t | U_i > 0, U_k > 0\}$ is the area under the joint density of $U_i, U_k)^T$ in the region

$0 < U_i \leq tU_k$.

Figure 2.36: Geometry of distribution in (U_i, U_k) space



By spherical symmetry this area depends only on the angle, $\theta = \tan^{-1}(t)$, that the line $U_i = tU_k$, makes with the U_k axis as shown in Figure 2.36. Thus, $F(t) = F(\tan(\theta)) = h(\theta)$ and $\Pr\{U_i/U_k \leq t\}$, is considered as a function $h(\theta)$. For each point $\boldsymbol{U} = (U_i, U_k)^T$ in $\mathbb{R}^2$ there corresponds a homeomorphism (i.e., mapping) $H : \mathbb{R}^2 \times \Theta \to \mathbb{R}^2$ of $\mathbb{R}^2$ into itself such that $\widehat{\boldsymbol{U}} = H(\boldsymbol{U}, \theta_1) \in \mathbb{R}^2$. Thus, there exist a point

$$\widetilde{\boldsymbol{U}} = H(\widehat{\boldsymbol{U}}, \theta_2) = H(H(\boldsymbol{U}, \theta_1), \theta_2) \tag{2.B.1}$$

$$= H(\boldsymbol{U}, \theta_1 + \theta_2) \tag{2.B.2}$$

These operations are consistent with a transformation group with group operation addition on Θ. Refer to Guggenheimer (1977, p. 88) for further details. Since the erstwhile group maps into itself, there exist a group homomorphism $h(\theta_1 + \theta_2) = h(\theta_1) + h(\theta_2)$. The group isomorphism implies that this must satisfy the Cauchy functional equation $h(\theta_1 + \theta_2) = h(\theta_1) + h(\theta_2)$. Aczel (1966, pp. 31-32) proves that $h(\theta) = c\,\theta$ satisfies the Cauchy functional equation. Consequently, the distribution function $F(t) = \Pr\{U_i/U_k \leq t\} = h(\tan^{-1}(t)) = c\tan^{-1}(t)$. This is the distribution function for a standard Cauchy. Thus, we prove that the ratio U_i/U_k has a standard Cauchy distribution. □

2.B.2 Proof of Theorem 2.2.5 Generalized Cauchy Elliptically Symmetric

Proof. The proof is adapted from Arnold and Brockett (1992, Thm. 2) with slight modification to fill gaps. The general idea of the proof is related to transformation groups (Guggenheimer, 1977, pp. 88-90). By hypothesis, $\boldsymbol{X}$ being elliptically symmetric implies that it has a representation $\boldsymbol{X} = \boldsymbol{AU}$ where $\boldsymbol{A}$ is invertible, and $\boldsymbol{U}$ has a symmetrically symmetric distribution. Any two elements of $X_i, X_k,\ i \neq k$ of $\boldsymbol{X}$ can be written in matrix form as

$$\begin{bmatrix} X_i \\ X_k \end{bmatrix} = \boldsymbol{BU} \tag{2.B.3}$$

where B is a $2 \times n$ matrix. Under the Gram-Schmidt orthogonalization process, a set of independent vectors (that comprise a matrix) can be mapped into a set of mutually orthogonal and orthonormal vectors that constitute an orthonormal matrix (Gentle, 2007, p. 27). Under the $\boldsymbol{LU}$ matrix transformation method (Gentle, 2007, p. 186), where $\boldsymbol{L}$ is a lower triangular matrix and $\boldsymbol{U}$ is an upper triangular matrix, there exist a 2×2 matrix $\boldsymbol{C}$ and an orthonormal matrix $\boldsymbol{Q}$ such that $\boldsymbol{QQ}^T = \boldsymbol{I}$ and $\boldsymbol{B} = \boldsymbol{CQ}$. In which case, $\boldsymbol{BU} = (\boldsymbol{CQ})\boldsymbol{U} = \boldsymbol{C}(\boldsymbol{QU})$. By virtue of orthogonality of $\boldsymbol{Q}$, we induce the symmetrically symmetric vector $\boldsymbol{Y}$ comprised of the pair of random variables

$$\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \boldsymbol{QU} \tag{2.B.4}$$

Let $C = \left(\begin{smallmatrix} a & b \\ 0 & c \end{smallmatrix}\right)$ for some non-zero constants a, b, c. Thus

$$\boldsymbol{X} = \begin{bmatrix} X_i \\ X_k \end{bmatrix} = \boldsymbol{CY} = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} aY_1 + bY_2 \\ cY_2 \end{bmatrix} \tag{2.B.5}$$

$$\Longrightarrow X_i/X_k = \left(\frac{a}{c}\right)\left(\frac{Y_1}{Y_2}\right) + \left(\frac{b}{c}\right) \tag{2.B.6}$$

According to Theorem 2.2.4, $\frac{Y_1}{Y_2}$ follows a standard Cauchy law, i.e., $\frac{Y_1}{Y_2} \sim \mathscr{C}(0,1)$ by virtue of the spherical symmetric relationship in (2.B.4). Hence in (2.B.6), $X_i/X_k = \frac{a}{c}\mathscr{C}(0,1) + \frac{b}{c}$ which follows a generalized Cauchy law by definition. □

2.B.3 Proof of Theorem 2.2.6 on existence of Generalized Cauchy for MLA index

Proof. The proof follows that for Theorem 2.2.5 in Appendix 2.B.2. In this case we let $X_i = \bar{X}^L$ and $X_k = \bar{X}^G$. Then substitute a for $\frac{b}{c}$ and b for $\frac{a}{c}$. Thus we have $\bar{X}^L/\bar{X}^G = b\mathscr{C}(0,1) + a \sim \mathscr{C}(a,b)$. □

2.B.4 Proof of Theorem 2.3.1 Bifurcated RIH with Consumption Ratchet

Proof. For gain in income we have

$$\frac{C_y}{Y_t} = 1 - \alpha_0 - \alpha_1 \frac{Y_t}{M_t} = 1 - \alpha_0 - \alpha_1(1 + g_t^G) \tag{2.B.7}$$

$$\Longrightarrow C_t = Y_t(1 - \alpha_0 - \alpha_1) - \alpha_1 g_t^G Y_t \tag{2.B.8}$$

$$\text{Similarly, for losses } C_t = Y_t(1 - \alpha_0 - \alpha_1) + \alpha_1 g_t^L \tag{2.B.9}$$

$$\text{For no change } C_t = Y_t(1 - \alpha_0 - \alpha_1) \tag{2.B.10}$$

Let $a(d) = 1 - \alpha_0 - \alpha_1$, $\Delta^b C_t^D = -\alpha_1 g_t^G Y_t$ and $\lambda_t = \frac{|g_t^L|}{g_t^G}$ and the proof is done. □

2.B.5 Proof of Proposition 2.5.3 MLA index iid

Proof. By Lemma 2.5.1 and construction in (2.5.7), the MLA index estimator $\widehat{\lambda}_{k-1}$ is put in correspondence with $\mathscr{B}_{2k-1,2k}$. So $\widehat{\lambda}_{k-2}$ is put in correspondence with $\mathscr{B}_{2k-3,2k-2}$. Suppose that the premise of the proposition is false. Then $\widehat{\lambda}_{k-1}$ and $\widehat{\lambda}_{k-2}$ are correlated, for they are in correspondence with a common element, call it x, in $\mathscr{B}_{2k-1,2k}$ and $\mathscr{B}_{2k-3,2k-2}$. The probability associated with that event is given by

$$\Pr\{x \in \mathscr{B}_{2k-3,2k-2} \cap \mathscr{B}_{2k-1,2k}\} = \Pr\{x \in \mathscr{B}^0\} \tag{2.B.11}$$

$$\Longrightarrow \Pr\{g = 0\} = 0 \tag{2.B.12}$$

Thus, the probability that two MLA index estimates are in correspondence with a common element is zero. This contradicts our assumption. Hence the premise of the proposition stands. □

2.B.6 Proof of Proposition 2.5.6 MLA index is Cauchy rv

Proof. By definition a Cauchy process $\{C_t^u(\omega); \mathscr{F}_t\}$ is comprised of independent increments (Jacobsen, 2006, p. 140). Hence

$$\Pr\left\{\widehat{c}_t = C_{s+t}^u - C_s^u \in (dy) | \widehat{c}_t > 0\right\} = \frac{\sigma t}{\pi(\sigma^2 t^2 + y^2)} \mathbb{I}_{\{\widehat{c}_t > 0\}} dy \tag{2.B.13}$$

This is the equation of a generalized Cauchy distribution with scale parameter σt. Under Theorem 2.2.6 the MLA index estimator $\widehat{\lambda}_t$ also has a generalized Cauchy distribution. Thus, $\widehat{c}_t$ and $\widehat{\lambda}_t$ follow the same law so that $\Pr|\widehat{\lambda}_t - \widehat{c}_t| \xrightarrow{p} 0$. By Slutsky's Theorem (Chow and Teicher, 1988, p. 254) the two random variables converge to the same (Cauchy) distribution on the same probability space. □

2.B.7 Proof of Lemma 2.7.4 consumption growth condition

Proof. Suppose $-g_{t+1}^c > -\varphi_{t+1}(\boldsymbol{g}, \boldsymbol{\lambda})$ so the direction of the inequality $\gg$ is reversed. Hence

$$g_{t+1}^c < \varphi_{t+1}(\boldsymbol{g}, \boldsymbol{\lambda}) \tag{2.B.14}$$

$$= -\lambda_{t+1}\left(\frac{g_{t+1}}{g_t} e^{g_{t+1}^Y}\right) - 1 \tag{2.B.15}$$

Under the myopia Axiom 1 $g_{t+1}^c = g_{t+1}^Y$. A first order approximation of $e^{g_{t+1}^Y}$ produces

$$1 + g_{t+1}^Y < -\lambda_{t+1}\left(\frac{g_{t+1}}{g_t}\right)(1 + g_{t+1}^Y) \tag{2.B.16}$$

$$\Longrightarrow 1 < -\lambda_{t+1}\left(\frac{g_{t+1}}{g_t}\right) \tag{2.B.17}$$

Since all the variables on the right hand side are positive, the inequality is impossible. So

$$-g_{t+1}^c \not> -\varphi_{t+1}(\boldsymbol{g}, \boldsymbol{\lambda})$$

is not reversible. □

2.B.8 Proof of Proposition 2.7.2 MLA index and AP risk nexus

Proof. Let C_r be a reference consumption and U be a twice differentiable utility function at C_r. In an ε-disk $B_\varepsilon(C_r)$ centered at C_r such that for ε sufficiently large $\mu_\eta \leq \varepsilon$ we have a MLA index $\lambda_r(C_r, \mu_\eta)$ and Arrow-Pratt relative risk measure $\gamma_r(\mu_\eta)$ is derived from the following analysis.

$$\gamma_r^D(C_r) = -C_r \frac{U''(C_r)}{U'(C_r)}, \qquad \lambda_r(C_r, \mu_\eta) = \frac{U'_-(C_r - \mu_\eta)}{U'_+(C_r + \mu_\eta)} \tag{2.B.18}$$

$$U'_-(C_r - \mu_\eta) = U'(C_r - \mu_\eta) \qquad U'_+(C_r + \mu_\eta) = U'(C_r + \mu_\eta) \tag{2.B.19}$$

The convention in (2.B.19) is adapted from Wakker (2010, p. 239) who, using different notation, referred to the sub-utility functions U_+ and U_- as intrinsic utility and U as overall utility. Substitution of $C_r + \mu_\eta$ for C_r and $U'_+(C_r + \mu_\eta)$ in the equation for $\gamma_r^D(C_r)$ leads, by abuse of notation, to

$$\begin{aligned}
\gamma_r^D(C_r, \mu_\eta) &= -(C_r + \mu_\eta)\frac{U''_+(C_r + \mu_\eta)}{U'_+(C_r + \mu_\eta)} \\
&= -C_r \frac{U''_+(C_r + \mu_\eta)}{U'_+(C_r + \mu_\eta)} - \mu_\eta \frac{U''_+(C_r + \mu_\eta)}{U'_+(C_r + \mu_\eta)} \\
&= -C_r \frac{U''_+(C_r + \mu_\eta)}{U'_+(C_r + \mu_\eta)} + \mu_\eta \lambda_r(C_r, \mu_\eta) \frac{U''_+(C_r + \mu_\eta)}{U'_-(C_r - \mu_\eta)}
\end{aligned} \tag{2.B.20}$$

$$\Longrightarrow \lim_{C_r \to 0} \gamma_r^D(C_r, \mu_\eta) = \gamma_r^D(0, \mu_\eta) = \mu_\eta \lambda_r(0, \mu_\eta) \operatorname{sgn}(\widetilde{\gamma}(\mu_\eta))\widetilde{\gamma}(\mu_\eta) \tag{2.B.21}$$

$$\widetilde{\gamma}(\mu_\eta) = -\frac{U''_+(\mu_\eta)}{U'_-(-\mu_\eta)}, \qquad \operatorname{sgn}(\widetilde{\gamma}(\mu_\eta)) = \begin{cases} +1 & \widetilde{\gamma}(\mu_\eta) \geq 0 \\ -1 & \widetilde{\gamma}(\mu_\eta) < 0 \end{cases} \tag{2.B.22}$$

Implicit in the analysis is $\partial/\partial \mu_\eta[-U_-(-\mu_\eta)] = U'_-(-\mu_\eta)$ to account for negative utility. Tversky and Kahneman (1992, p. 309) also use a negative argument in the parameterizations of their value function. The sign of $\widetilde{\gamma}(\mu_\eta)$ depends on the sign of $U''_+(\mu_\eta)$. The latter is negative for risk averse DMs and positive for risk seeking DMs. Hence $\gamma_r^D(0, \mu_\eta) > 0$ for risk aversion and $\gamma_r^D(0, \mu_\eta) < 0$ for risk seeking. □

2.B.9 Proof of Proposition 2.7.10 excess volatility from behavioural pricing kernel

Proof. The proof contains two parts. One for bad states, i.e., loss states L; and another for good states, i.e., gain states G.

Bad states L

Let $\bar{r}_L^W$ and $s_L^{W^2}$ be the sample mean and variance in L states.

$$\begin{aligned}\bar{r}_L^W &= (w^+(1-\phi)+\Delta_1)r_{LG}+(w^-(\phi)-\Delta_1)r_{LL}\\ &= w^+(1-\phi)r_{LG}+w^-(\phi)r_{LL}+\Delta_1(r_{LG}-r_{LL}) \\ &= r_{LG}+(w^-(\phi)+\Delta_1)(r_{LL}-r_{LG})\end{aligned} \tag{2.B.23}$$

$$\begin{aligned}s_L^{W^2} &= (w^+(1-\phi)+\Delta_1)(r_{LG}-\bar{r}_L)^2+(w^-(\phi)-\Delta_1)(r_{LL}-\bar{r}_L)^2\\ &= w^+(1-\phi)(r_{LG}-\bar{r}_L)^2+w^-(\phi)(r_{LL}-\bar{r}_L)^2+\Delta_1(r_{LG}^2-r_{LL}^2-2\bar{r}_L(r_{LG}-r_{LL}))\end{aligned} \tag{2.B.24}$$

In the neoclassical model the mean and variance from Q are given by

$$\bar{r}_L = (1-\phi)r_{LG}+(\phi)r_{LL} = r_{LG}-\phi(r_{LG}-r_{LL}) \tag{2.B.25}$$

$$\begin{aligned}s_L^2 &= (1-\phi)(r_{LG}-\bar{r}_L)^2+(\phi)(r_{LL}-\bar{r}_L)^2\\ &= (r_{LG}-\bar{r}_L)^2+(r_{LL}-\bar{r}_L)^2+\phi(r_{LG}^2-r_{LL}^2-2\bar{r}_L(r_{LG}-r_{LL}))\end{aligned} \tag{2.B.26}$$

The expected mean return under probability weighting is larger because

$$\bar{r}_L^W = r_{LG}+(w^-(\phi)+\Delta_1)(r_{LL}-r_{LG}) \ > \ \bar{r}_L = r_{LG}-\phi(r_{LG}-r_{LL}) \tag{2.B.27}$$

by virtue of $w^-(\phi)+\Delta_1>0$. The variance is larger under probability weighting if $s_L^{W^2}-s_L^2>0$. This implies that

$$(w^+(1-\phi)+\Delta_1) > \frac{(\phi-\Delta_1)(r_{LG}^2-r_{LL}^2-2\bar{r}_L(r_{LG}-r_{LL}))+(\phi-w^-(\phi))(r_{LL}-\bar{r}_L)^2}{(r_{LG}-\bar{r}_L)^2} \tag{2.B.28}$$

$$= (\phi-\Delta_1)-\frac{(r_{LL}-\bar{r}_L)^2}{(r_{LG}-\bar{r}_L)^2}(w^-(\phi)-\Delta_1) \tag{2.B.29}$$

Let $0 < \delta < 1$ be the ratio expression in (2.B.29) by virtue of preference for good states under Assumption 2.7.7. Substituting $1 - w^{-}(\phi)$ for $w^{+}(1-\phi)$ we get

$$1 - w^{-}(\phi) + \Delta_1 > (\phi - \Delta_1) - \delta(w^{-}(\phi) - \Delta_1) \tag{2.B.30}$$

$$\Longrightarrow (w^{-}(\phi) - \Delta_1)(1-\delta) < 1 - (\phi - \Delta_1) < 1 \Rightarrow (w^{-}(\phi) - \Delta_1)) < \frac{1}{1-\delta} \tag{2.B.31}$$

The latter expression is true for $0 < \Delta_1 < 1$ uniformly. Hence the probability weighted variance of returns is larger than under the neoclassical model.

Good states G

The mean $\bar{r}_G$ and variance s_G^2 of returns in good states under the neoclassical model, and mean $\bar{r}_G^W$ and variance $s_G^{W^2}$ are given by

$$\bar{r}_G = q_{GG} r_{GG} + q_{GL} r_{GL}, \qquad s_G^2 = q_{GG}(r_{GG} - \bar{r}_G)^2 + q_{GL}(r_{GL} - \bar{r})^2 \tag{2.B.32}$$

$$\bar{r}_G^W = w(q_{GG}) r_{GG} + w(q_{GL}) r_{GL}, \qquad s_G^{W^2} = w(q_{GG})(r_{GG} - \bar{r}_G)^2 + w(q_{GL})(r_{GL} - \bar{r})^2 \tag{2.B.33}$$

Since $w^{-}(q_{GL}) < q_{GL}$ for probabilistic risk seeking, and $w^{+}(q_{GG} > q_{GG}$ for probabilistic risk aversion, assuming that $q_{GG} = q_{LL} = \phi$ and $q_{GL} = q_{LG} = 1 - \phi$ for symmetric Q under the neoclassical model, if $\bar{r}_G^W > \bar{r}_G$, then

$$w^{+}(\phi) r_{GG} + w^{-}(1-\phi) r_{GL} > \phi r_{GG} + (1-\phi) r_{GL} \tag{2.B.34}$$

$$\Rightarrow (w^{+}(\phi) - \phi) r_{GG} + (1 - w^{+}(\phi) + \phi) r_{GL} > r_{GL} \tag{2.B.35}$$

$$\Rightarrow (w^{+}(\phi) - \phi) r_{GG} - (w^{+}(\phi) - \phi) r_{GL} + r_{GL} > r_{GL} \tag{2.B.36}$$

$$\Rightarrow r_{GG} > r_{GL} \tag{2.B.37}$$

This result is contrary to the ranking predicted by the behavioral pricing kernel. Therefore, the hypothesis $\bar{r}_G^W > \bar{r}_G$ is false. Thus, we must have $\bar{r}_G^W \leq \bar{r}_G$ in G states.

If $s_G^{W^2} > s_G^2$, then

$$w^+(\phi)(r_{GG} - \bar{r}_G)^2 + w^-(1-\phi)(r_{GL} - \bar{r}_G)^2 > \phi(r_{GG} - \bar{r}_G)^2 + (1-\phi)(r_{GL} - \bar{r}_G)^2 \tag{2.B.38}$$

$$\Rightarrow w^+(\phi)[(r_{GG} - \bar{r}_G)^2 - (r_{GL} - \bar{r}_G)^2] + (r_{GL} - \bar{r}_G)^2 > \phi[(r_{GG} - \bar{r}_G)^2 - (r_{GL} - \bar{r}_G)^2] + (r_{GL} - \bar{r}_G)^2 \tag{2.B.39}$$

$$\Rightarrow w^+(\phi) > \phi \tag{2.B.40}$$

Since the latter inequality holds for probabilistic risk aversion characteristic of good states, the hypothesis $s_G^{W^2} > s_G^2$ is upheld. So in good states we have low average risk premia and excess volatility relative to the neoclassical model. □

2.C DATA APPENDIX

2.C.1 Loss aversion index estimate around the world

Table 2.7: Loss aversion indexes around the world

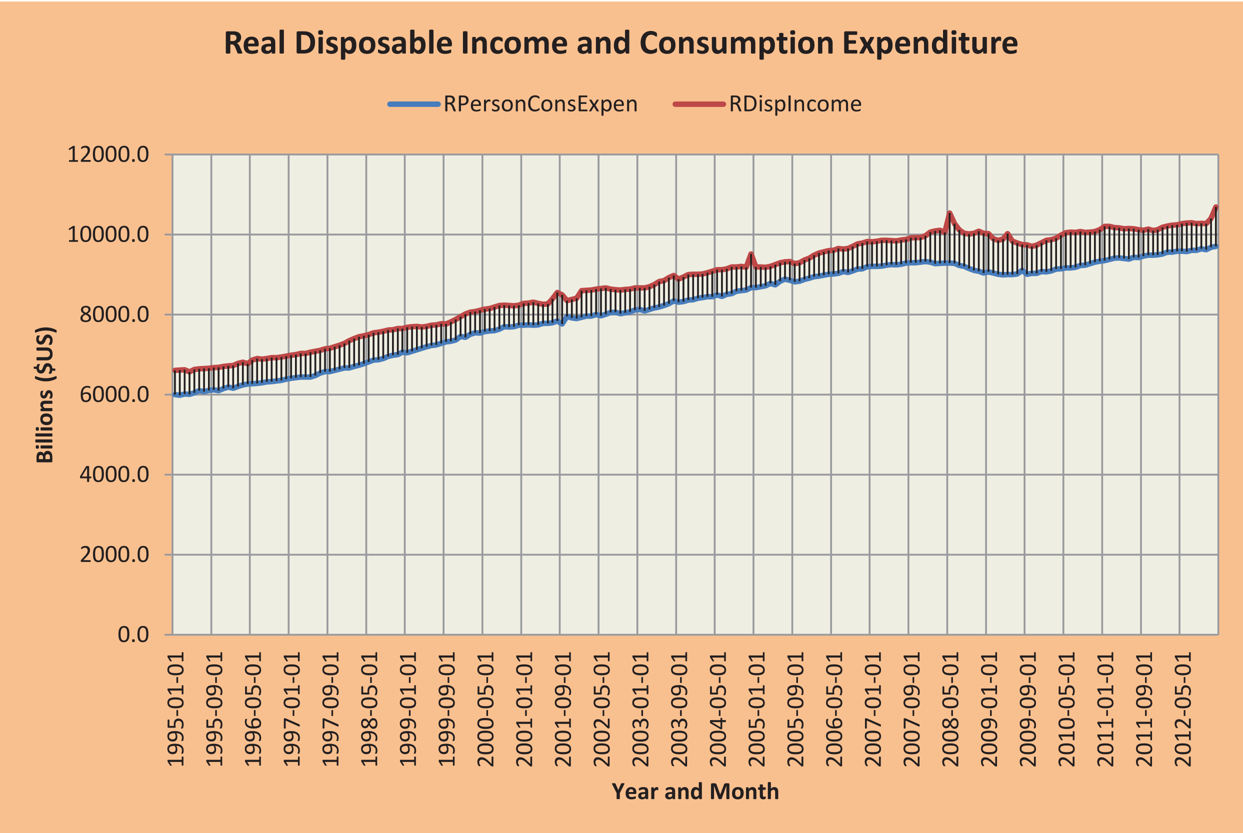
Country	α	β	γ	θ
Angola	0.60	1.00	0.60	1.45
Argentina	0.60	1.00	0.70	1.09
Australia	0.60	0.95	0.60	1.24
Austria	0.40	0.95	0.65	1.62
Azerbaijan	0.60	1.00	0.65	1.23
Bosnia–Herzegovina	0.65	0.90	0.45	1.00
Canada	0.50	1.00	0.50	2.00
Chile	0.55	1.00	0.65	2.00
China	0.60	1.00	0.60	1.83
Colombia	0.40	1.00	0.35	2.00
Croatia	0.60	1.00	0.45	2.33
Czech Republic	0.60	1.00	0.55	2.00
Denmark	0.50	1.00	0.65	2.00
Estonia	0.50	1.00	0.35	4.00
Georgia	0.55	1.00	0.60	5.50
Germany	0.45	1.00	0.50	2.00
Greece	0.65	0.80	0.50	2.00
Hong Kong	0.40	1.00	0.30	2.43
Hungary	0.50	1.00	0.45	2.00
Ireland	0.50	1.00	0.45	2.00
Israel	0.58	0.95	0.35	1.99
Italy	0.45	1.00	0.50	2.46
Japan	0.45	1.00	0.60	2.00
Lebanon	0.53	0.95	0.25	1.74
Lithuania	0.55	1.00	0.35	2.00
Malaysia	0.58	1.00	0.60	1.50
Mexico	0.40	1.00	0.35	1.50

Country	α	β	γ	θ
Moldova	0.65	0.95	0.65	3.44
New Zealand	0.65	0.95	0.50	1.50
Nigeria	0.75	1.00	0.50	2.00
Norway	0.55	1.00	0.55	1.83
Portugal	0.50	1.00	0.65	1.83
Romania	0.50	1.00	0.60	3.33
Russia	0.53	1.00	0.33	3.00
Slovenia	0.55	1.00	0.40	2.12
South Korea	0.60	0.95	0.70	1.37
Spain	0.45	1.00	0.60	2.38
Sweden	0.50	1.00	0.65	2.00
Switzerland	0.45	1.00	0.50	2.00
Taiwan	0.55	0.95	0.53	2.00
Thailand	0.65	0.90	0.55	3.00
Turkey	0.60	1.00	0.65	1.80
UK	0.50	1.00	0.50	1.38
USA	0.58	1.00	0.43	1.65
Vietnam	0.60	1.00	0.55	1.75

Source: Rieger et al. (2011, Table 2, p. 7)
α, β are curvature parameter for power value function;
γ is the curvature parameter for probability weighting function;
θ is Tversky and Kahneman (1992) robust ratio scale loss aversion index.

2.C.2 Plot of US Myopic Consumption Tracking Income

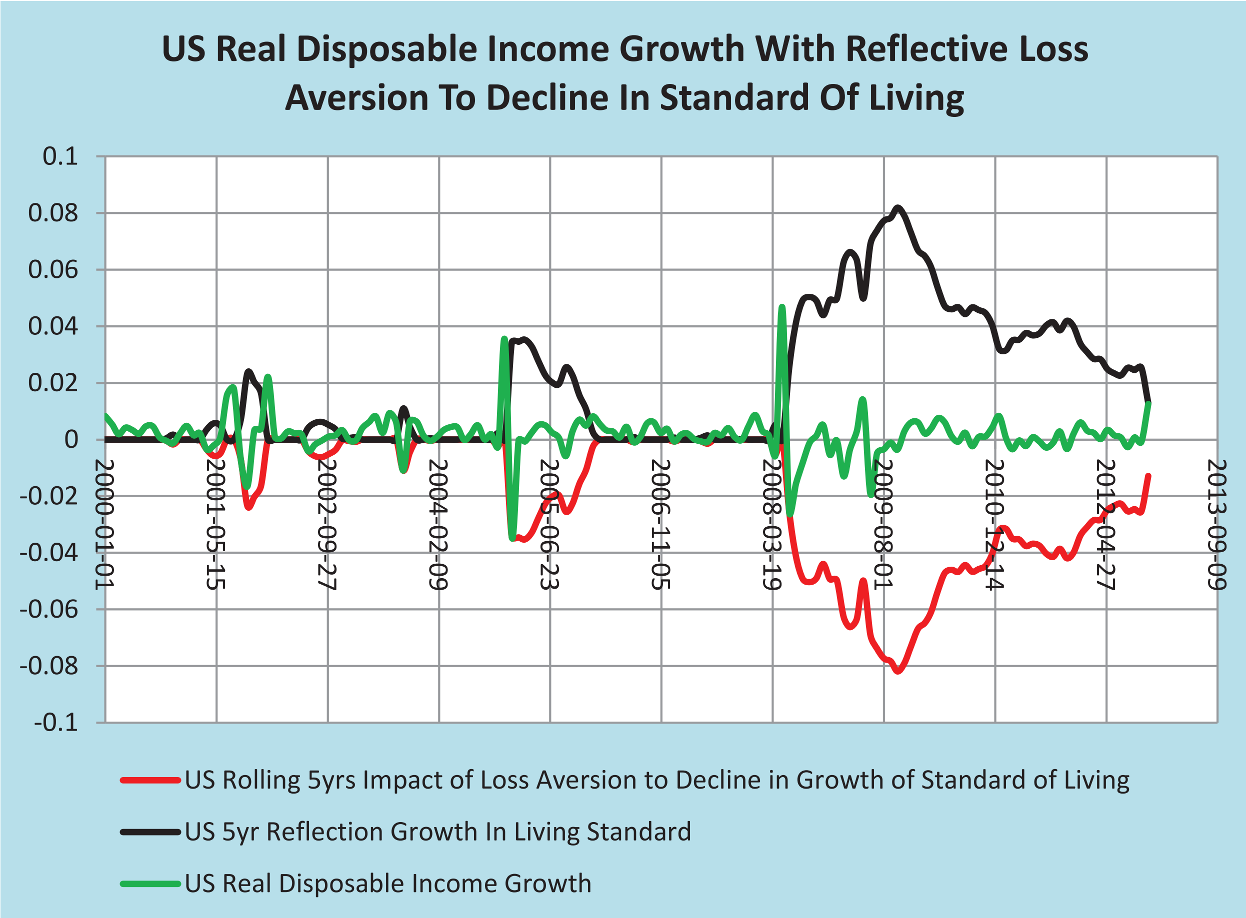
Figure 2.37: US Myopic Consumption Tracking Income



According to Shea (1995, pp. 798-799) “Under myopia, consumption tracks current income. Thus, the failure of the LCH/PIH should be symmetric: consumption should respond equally to predictable income increases and decreases.”

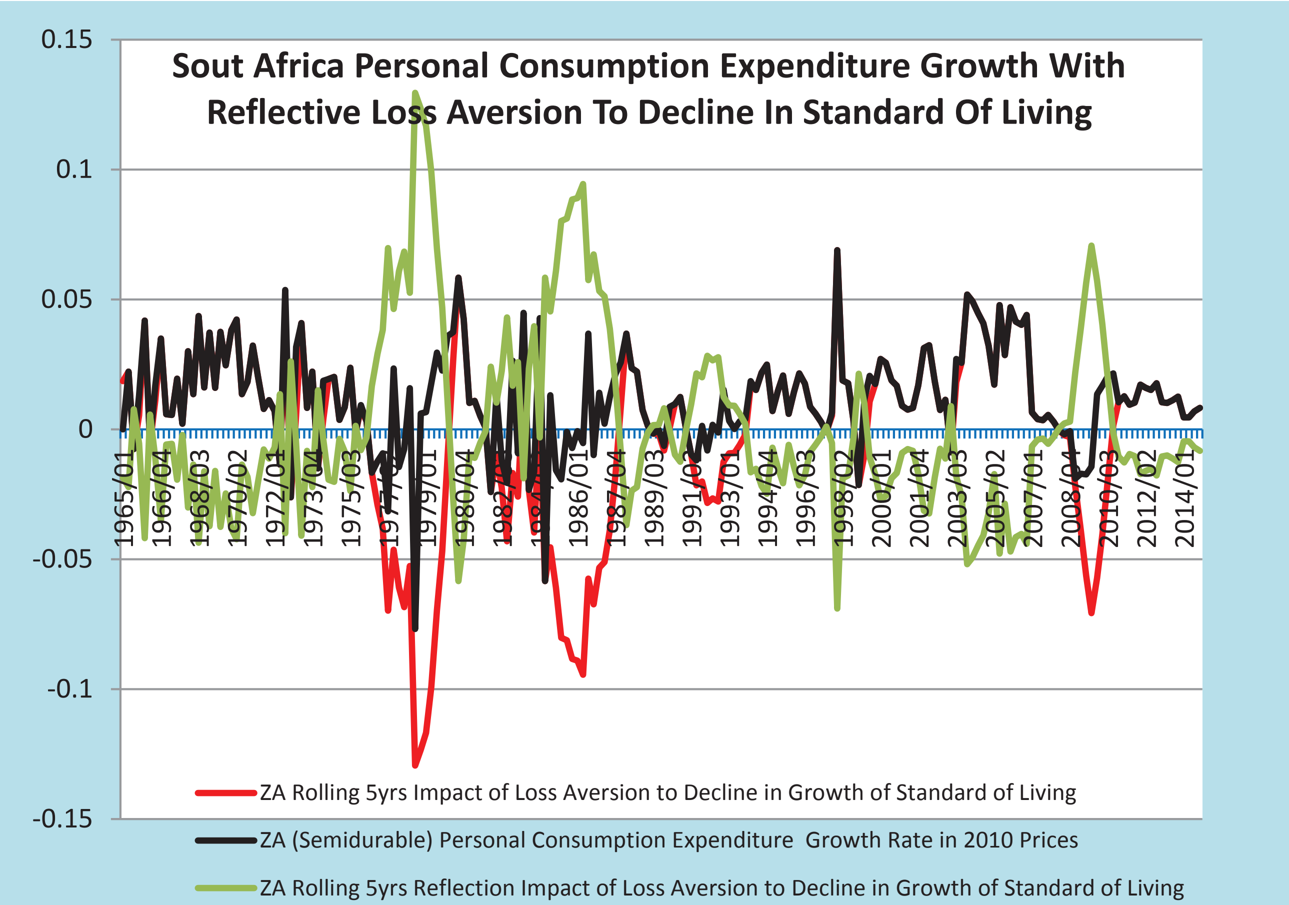
2.C.3 Plot of US Real Disposable Income Growth With MLA Reflection

Figure 2.38: US Real Disposable Income Growth With MLA Effects and Reflection



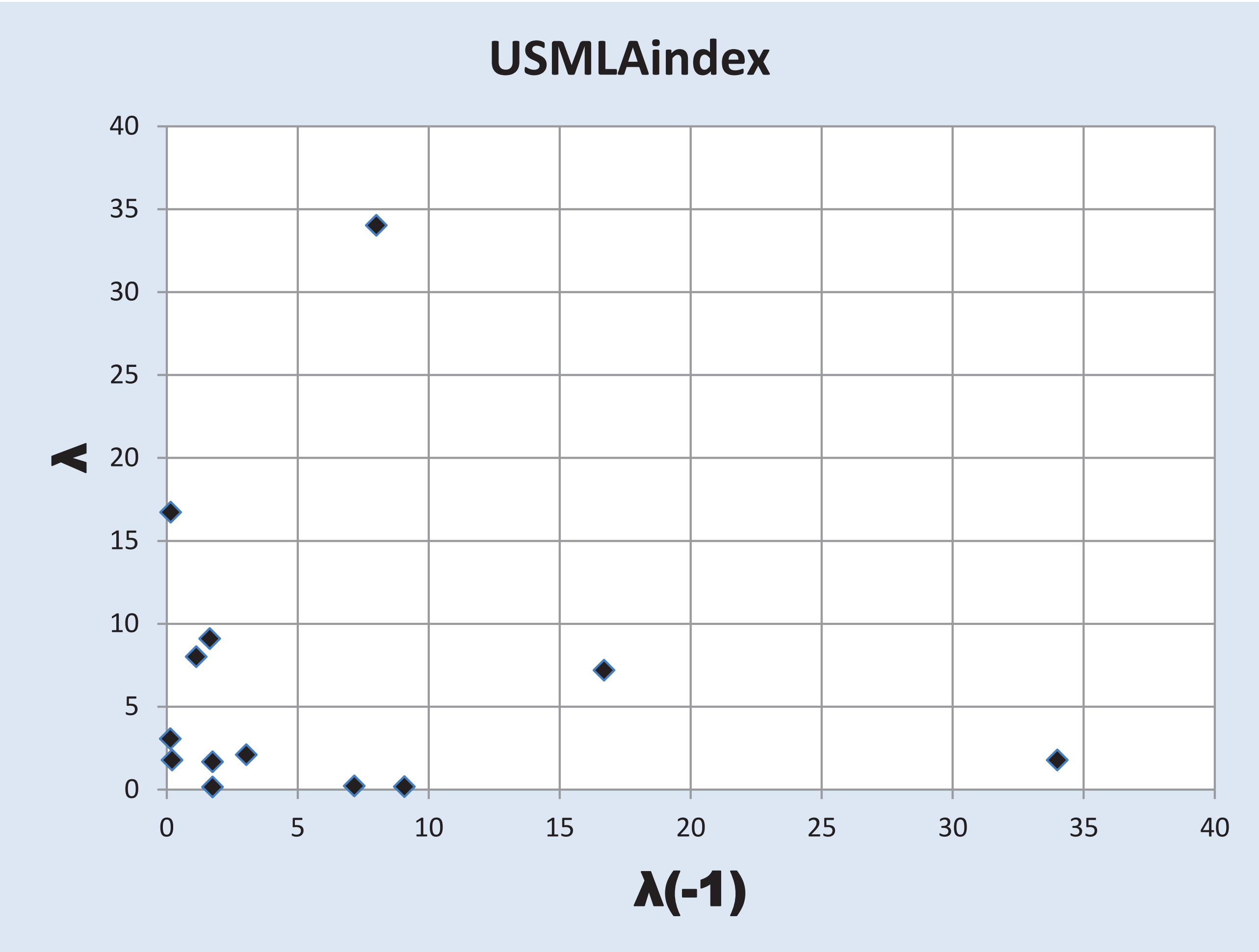
2.C.4 Plot of South Africa PCE Growth With MLA Reflection

Figure 2.39: South Africa PCE Growth With MLA Effects and Reflection



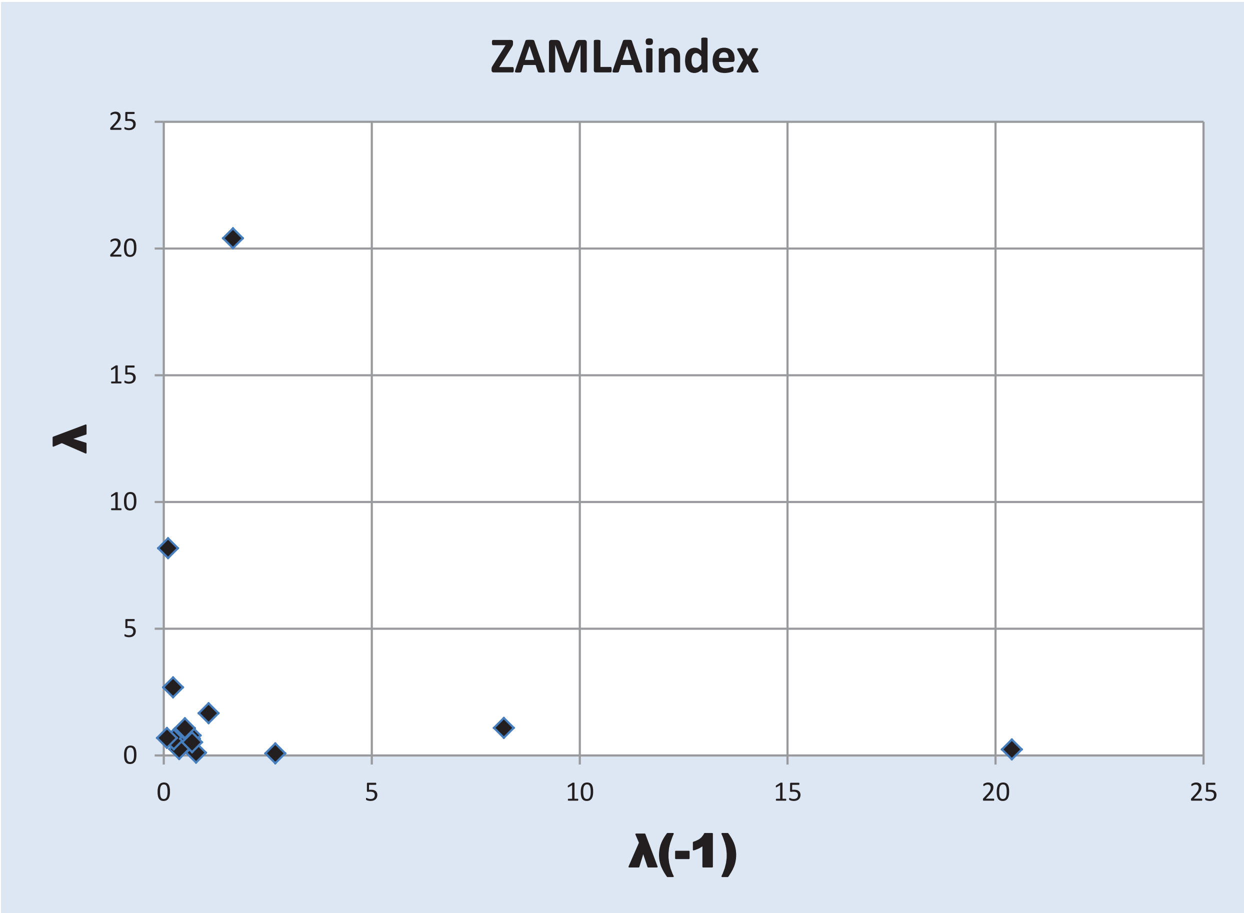
2.C.5 Plots of US MLA index independence

Figure 2.40: US MLA index independence:
$\widehat{\lambda}_{t^{2k-1,2k}}$ *vs.* $\widehat{\lambda}_{t^{2k-3,2k-2}}$



2.C.6 Plots of South Africa MLA index independence

Figure 2.41: South Africa MLA index independence:
$\widehat{\lambda}_{t_{2k-1,2k}}$ vs. $\widehat{\lambda}_{t_{2k-3,2k-2}}$



Autoregression plots support the prediction of Proposition 2.5.3 that the MLA index is independent and identically distributed.

2.D APPENDIX OF FITTED STATISTICAL DISTRIBUTIONS

The distributions were fitted using the maximum likelihood option in EasyFit®. The Anderson and Darling (1954) is a nonparametric test which did not provide *p*-values. The χ^2-test and Kolmogorov-Smirnov goodness of fit tests provide *p*-values as indicated. Additionally, the software ranked the distributions in a tournament for goodness of it as indicated.

2.D.1 Diagnostics for fitted US income growth data

2.D.2 Diagnostics for fitted US standard of living growth data

2.D.3 Diagnostics for fitted South Africa income growth data

2.D.4 Diagnostics for fitted South Africa standard of living growth data

2.D.5 Diagnostics for fitted Fishburn and Kochenberger (1979) Metastudy data

2.D.6 Diagnostics for fitted Rieger et al. (2011) MLA index around the world

2.D.7 Diagnostics for fitted US macroeconomic MLA index

2.D.8 Diagnostics for fitted South Africa macroeconomic MLA index

Chapter 3

A Probability Model of Irrational Exuberance and Financial Market Instability

"A source is a specific set of events". Wakker (2010, pg. 318)

3.1 Introduction

This chapter contributes to the literature on behavioural finance by introducing a behavioral empirical stochastic process based on probabilistic risk attitude factors that affect investor confidence. We show how the process provides early warning for financial market instability and we calibrate the empirical process with published parameter estimates.

Financial markets reflect activity in the real economy. For example, bond markets facilitate the allocation of credit to firms, and stock markets reflect the valuation of shareholder equity in publicly traded firms. Events like the Great Recession of 2008 triggered by uncertainty about credit risk from the US real estate and collateralized debt obligation (CDO) crisis (Adelson, 2013), and the Eurozone financial crisis (Castellacci and Choi, 2014), have eroded confidence in financial markets.[1] This has led to credit freeze, subsequent massive unemployment, social unrest and calls for market reform. Thus, it cannot be gainsaid that financial market stability, and confidence in same, are of utmost importance to market participants, regulators, and policy makers.[2] However, asset pricing bubbles and market crashes are not new.[3] So they are likely to repeat themselves. Refer to Kindleberger and Aliber (2011); Reinhart and Rogoff (2014) for a detailed history of

[1] For example, a March NBC News/Wall Street Journal poll showed 57 percent of surveyed American adults believed the United States was still in a recession (although that is the lowest share of respondents under that impression since early 2008), according to "5 Years After the Great Recession, Our Economy Still Far from Recovered, Huffington Post (updated August 26, 2014) http://www.huffingtonpost.com/andrew-fieldhouse/five-years-after-the-grea˙b˙5530597.html.

[2] For example, the IMF Global Financial Stability Report (https://www.imf.org/external/pubs/ft/gfsr/); Financial Policy Committee, Bank of England; and Financial Stability Committee, Federal Reserve Board in the US are newly formed entities tasked with monitoring financial stability (McGrane and Da Costa, 2014).

[3] See e.g., Shefrin (2015b) who draws parallels between the recent Chinese stock market crash and Minsky's financial instability hypothesis. But compare London (2015) who questions whether the recent Chinese stock market crash "was that big".

such events. This paper provides some new tools, based on a probability model of "irrational exuberance" and financial instability motivated by behavioural finance. It provides an early warning systems for bubbles and crashes predicated on market confidence factors.

Active research on financial market instability is conducted under rubric of market microstructure (Easley et al., 2011; Aldridge, 2014), econophysics (Johnson et al., 2000; Zhou and Sornette, 2006; Quax et al., 2013), macrofinance (Grasselli and Costa Lima, 2012; Keen, 2013; Wigniolle, 2014), agent based models (Thurner et al., 2012; Poledna et al., 2013; Hommes, 2013), experimental finance (Smith et al., 1988; Ashparouva et al., 2014) financial networks (Allen and Gale, 2000; Acemoglu et al., 2015), bank runs (Diamond and Dybvig, 1983), debt-deflation cycles (Fisher, 1933), revolving doors (Charles-Cadogan and Cole, 2014; Shive and Forster, 2014; Lucca et al., 2014; Lambert, 2015; Charles-Cadogan and Cole, 2015), and creative destruction in financial markets (Minsky, 1986). However, none of the foregoing papers use probability weighting functions[4] (pwfs)–which reflect pessimism and optimism in the presence of risk and uncertainty–to characterize financial market instability. Yet investor optimism, pessimism and uncertainty are common cores of all financial crises (Fisher, 1933; Kindleberger and Aliber, 2011). A notable exception is Bhattacharya et al. (2015) who model Minsky's financial instability hypothesis in the context of financial institutions' optimism, leverage and portfolio risk as part of a debt-deflation cycle. However, those authors did not use probability weighting functions to characterize financial market expectations about future states of the economy.

This paper contributes to the literature with a novel behavioural empirical local Lyapunov exponent[5] (BELLE) process that characterizes financial market instability with pwfs implied by index option prices. The latter allow natural experiments on probabilistic risk attitudes because they involve bets on future price movements.[6] The shapes of pwfs for index option prices reflect

[4]A probability weighting function reflects the weight that decision makers give to an otherwise objective cumulative probability measure over corresponding ranked outcomes.

[5]A Lyapunov exponent (λ) is a measure of the rate of convergence or divergence of a trajectory over time relative to two nearby starting points. Refer to Bask (2010); Bask and Widerberg (2012) for application of Lyapunov exponents to financial market instability.

[6]The Black and Scholes (1973, p. 644) European style call option price at time t is a pseudo-lottery $L = \{S(t), P_1;\ -e^{r(T-t)}K, P_2;\ 0, 1-P_1-P_2\}$ for $P_1+P_2<1$ with actuarial value $C(S,\sigma,t) = S(t)P_1 - e^{r(T-t)}KP_2$ where S is the underlying stock price, σ is its volatility, K is strike price, r is a discount rate, and T is expiry date, P_1 and P_2 are probabilities. The pwf w implied by ranked outcomes in L implies $w(P_1) \neq P_1$, $w(P_2) \neq P_2$. Refer to Boyer and Vorkink (2014) for details on stock options as lotteries phenomenon.

investors optimism and pessimism about different sources of credit risk in the market. So pwfs are source functions. In the sequel "source function" refers to a pwf generated by a specific credit risk source whereas "pwf" pertains to an abstract pwf. [7]

Time dependent *behavioural noise* in the orbit of source functions induce a local empirical process[8] for Lyapunov exponents in fixed point probability neighbourhoods. This facilitates estimation of critical values for investor probabilistic risk attitude factors from closed form expressions for the probability that a stable source function becomes unstable and *vice versa*. We calibrate the model with data from (Polkovnichenko and Zhao, 2013), and illustrate its robustness across credit risk sources for the 1997 Asian currency crisis, 2005 US real estate and CDO bubble, and Great Recession of 2008. We prove that credit risk source functions implied by index option prices provide early warning of financial market instability, and that they are sufficient statistics for a behavioural version of Minsky's financial instability hypothesis: An economy has stable and unstable regimes, and it transits from financial relations that make it stable to those that make it unstable (Minsky (1986, pp. 173-174) and Minsky (1994)).

Figure 3.1: Stable pwf

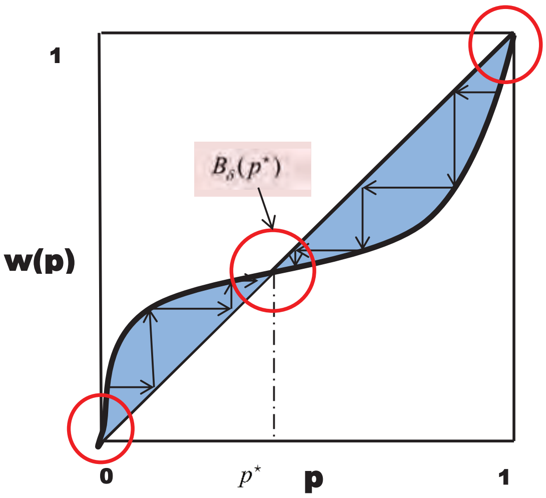
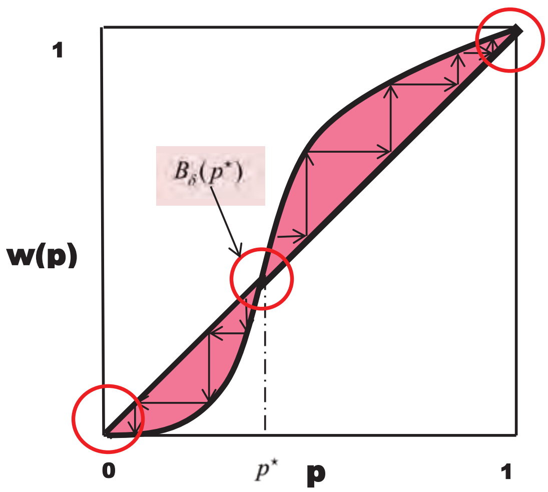


Figure 3.2: Unstable pwf



Phase diagrams for stable and unstable pwfs $w(p)$. The behavioral stochastic Lyapunov exponent process $\{\lambda(t,\omega); t \geq 0\}$ in fixed point $(p^\star)$ probability neighbourhood $B_\delta(p^\star)$ predict the shape of pwfs.

Figure 3.1 and Figure 3.2 depict the topology of our model. The phase diagram in Fig-

[7] Refer to Wakker (2010, pp. 318-321) for a detailed exposition of the source function concept.

[8] A classic empirical process is one comprised of sums of independent and identically distributed (*iid*) random variables (in our case noise) that converge to a limit process (Shorack and Wellner, 1986, pp. 1, 24).

ure 3.1 is stable because perturbed points contract to the restoring fixed point. In contrast, in Figure 3.2 perturbed points diverge from the fixed point. See e.g., Devaney (1989, p. 21). Behavioural dynamics in fixed point probability neighbourhoods $B_\delta(p^\star)$ centered at $p^\star$ with radius δ, control the shape of pwfs, and determine phase transition of stable and unstable pwfs shapes. Specifically, we introduce a local behavioural stochastic Lyapunov exponent (BELLE) process $\{\lambda(t,\omega); t \geq 0\}$ in $B_\delta(p^\star)$ that characterizes stochastic stability of pwfs, and we calibrate it to different credit risk source functions implied by index option prices to identify early warning signals for financial market instability. According to Minsky (1986) new financial products increase leverage and risk in the economy. Since the long term consequences of these financial innovations are unknown they can cause a seemingly stable system to become unstable. We believe that the pwf implied by index option prices is a sufficient statistic for the destabilizing effects of financial innovation since the implied pwf summarizes investor confidence.

3.1.1 Positioning of the paper in context of related literature

A growing literature on behavioural finance features pwfs popularized by Quiggin (1982), (Lopes, 1987, 1990) and Tversky and Kahneman (1992) in the evaluation of investor risk attitudes towards financial decision making (Polkovnichenko and Zhao, 2013; He and Zhou, 2013; Kliger and Levy, 2010; Wigniolle, 2014; Dierkes, 2009, 2013; Chabi-Yo and Song, 2013; Weigert and Ruenzi, 2013). A common theme in those papers is the impact of psychological factors or "sentiment" such as hope, fear, aspirations, underconfidence and overconfidence on investor risk attitudes towards tail events and formation of asset pricing bubbles.[9] Several of those papers use rank dependent utility (RDU), in conjunction with stochastic discount factors or pricing kernels[10] to evaluate risky prospects. Under expected utility theory (EUT), preference for lotteries are linear functionals over probabilities for given utility of outcomes (Von Neumann and Morgenstern, 1953, p. 24, eq(3:1:b)). RDU generalizes EUT by replacing probabilities in the latter with decision weights obtained by transforming a pwf, and it accommodates probabilistic risk attitudes via pwfs

[9] That strand of literature is distinguished from that spawned by Breeden and Litzenberger (1978, pp. 627,630) formulae for recovering state price density from option prices. Refer to Merton (1992, pp. 351-354) for a review.

[10] Roughly, the pricing kernel is the ratio of the price of an Arrow security and an investor's subjective probability about future states of the economy (Breeden et al., 2015). Formally, the pricing kernel incorporate risk, intertemporal substitution, and and time discount factors. Refer to Cochrane (2005, p. 17) for details.

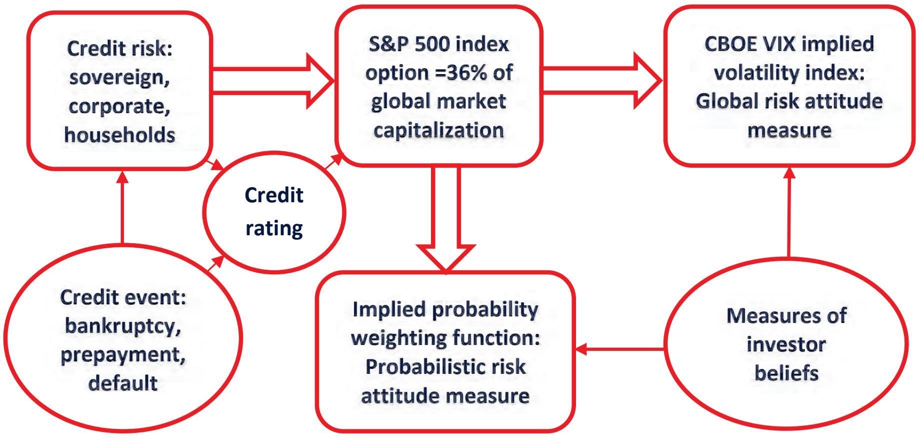
(Quiggin, 1993, p. 76). Appendix 3.B provides examples on how those attitudes are accommodated.

In contrast, a subset of the neuroscience and by extension neuroeconomics literature study probability distortions, and noise induced patterns, in fast and slow dynamical systems of neuronal activity (Tobler et al., 2008; Zhang and Maloney, 2012; Berglund and Gentz, 2010). In particular, the class of random accumulator models (RAMs) contemplate a decision maker's accumulation of noise terms over time whereupon a decision is reached after it attains a given threshold (Resulaj et al., 2009). This implies that *behavioural noise* is inherent in the decision making process and it should be accounted for in pwf dynamics. In the experimental economics literature, among others, Hey (1995, 2005); Blavatskky (2007); Loomes (2005); Loomes and Pogrebna (2014) study the impact of noise on measuring risk preferences.

Even though decision times are an important element of decision making, they seem to be neglected in the behavioural economics and finance literature (Webb, 2015). Our model fills that gap in the literature. It synthesizes pwfs dynamics and behavioural noise to construct a behavioural stochastic Lyapunov exponent process. Even though stochastic Lyapunov exponent processes are known to the statistical theory (Nychka et al., 1992; McCaffrey et al., 1992) and econometrics theory (Whang and Linton, 1999; Shintani and Linton, 2004; Park and Whang, 2012; BenSaïda, 2014) literatures, they have not been extended to pwf dynamics. Thus, our model is new to behavioural finance.[11]

[11]The model is in the spirit of Shefrin (2014, p. 588) who reminds us that the highly interdisciplinary nature of behavioural finance combines aspects of psychology, economics, mathematics, and statistics.

Figure 3.3: Credit risk sources and investor psychology



Our study of pwf dynamics implied by S&P 500 index option prices has implications for understanding global financial market instability.[12] For instance, synergies between risk preferences around the world (Szpiro, 1986; Rieger et al., 2015), integrated global financial markets (King and Wadhwani, 1990; Bongaerts et al., 2014), and contagion from the Great Recession of 2008 triggered by U.S. real estate and CDO markets,[13] imply that market instability is transmitted globally.[14] To be sure, credit spreads are also market indicators of investor fear level.[15] However, the model introduced in this paper is based on a probability model of investor confidence in financial markets.

Figure 3.3 provides a *schema* of how investor beliefs about credit risk sources are transmitted to the CBOE VIX and the pwf implied by S&P 500 index option prices. According to Figure 3.3, the sources of credit risk are prepayment, bankruptcy or default of debt obligations of sovereigns, corporations and households that interrupt the cash flow streams anticipated by in-

[12]The S&P 500 stock market index contains the stocks of 500 Large-Cap corporations, that comprise over 70% of the total market cap of all stocks traded in the U.S. See e.g., http://seekingalpha.com/article/1139431-u-s-as-a-percentage-of-world-market-cap and http://www.world-stock-exchanges.net/indices.html. Furthermore, the US market cap makes up over a third of the world's stock market cap, and **(alias?)** (a measure of the implied volatility of the S&P 500 index) is a sufficient statistic for investor psychology (Whaley, 2000).

[13]Dymski (2010) extends Minsky's FIH to include contagion in a globalized economy and the consideration of subprime lending in banks investment opportunity set.

[14]For example, shocks to the VIX are transmitted to the global economy via flight to quality wherein investors sell relatively risky EME bonds and buy US treasuries (Choi, 2014).

[15]"What is a credit spread, after all, but a measurement of the fear level associated with holding the underlying instrument? As fear spikes, spreads widen. The worry is over a "tail" event. VIX itself moves via much the same dynamic, and often in reaction to the exact same events. And the lion's share of demand lies in hedging against the "tails" . . . in the VIX's case, that translates into out-of-the-money VIX calls." Shaeffer Investment Research 4/2014.

vestors over a given time horizon. This is a manifestation of the uncertainty described in Minsky (1996, p. 362) “Money Manager Capitalism”. It makes investors probabilistic risk attitudes towards credit events time and source (state) dependent. Refer to Baucells and Heukamp (2012); Savadori and Mittone (2015) for time varying probabilistic risk attitudes. The source dependence of pwfs (Tversky and Wakker, 1995; Kilka and Weber, 2001; Abdellaoui et al., 2011) provides insight about probabilistic risk attitudes that is lacking in raw VIX scores.

The CBOE VIX provides a volatility score for credit risk; whereas a credit risk source function, i.e., the pwf corresponding to the source of credit risk, reflects investors probabilistic risk attitudes about the ranked index option prices that produce the VIX score. To see this, suppose we are given a raw VIX score of 30 that correspond to an “objective probability” of 0.1 derived from some forecast model. If the source of credit risk is sovereign default investors could assign a pessimistic probability weight of 0.25 to that event. By the same token if the source of credit risk is corporation default investors may assign to it a pessimistic probability weight of 0.15. Thus, investors give greater weight to the sovereign default source than they do to the corporation default source. Moreover, their assigned weight is controlled by curvature and elevation parameters of a source function.[16]

The rest of the paper proceeds as follows. In section 3.2 we introduce the local empirical process for the Lyapunov exponent which characterizes stochastic stability of pwfs. We provide closed form expressions for tipping points in pwf shape reversal. In section 3.3 we calibrate the model to pwfs corresponding to the source of credit risk implied by S&P 500 index option prices, and show how they predict market crash and provide early warning systems for market instability. In section 5.4 we conclude.

[16] Abdellaoui et al. (2011, p. 704) interpret curvature (α) and elevation (β) parameters for (Prelec, 1998) 2-parameter pwfs, and Abdellaoui et al. (2011, Fig. 9, p. 713) provide an example of source dependent pwfs.

3.2 The behavioural empirical local Lyapunov exponent (BELLE) process for probability weighting functions

In this section, we consider the *finite-time* behaviour of the Lyapunov exponent for the noisy orbit of deterministic subjective probability distributions for a large sample of heterogenous decision makers (DMs) with a noisy addend. Specifically, we characterize stable and unstable pwfs and the large sample probability estimate(s) for *tail event instability* in a seemingly stable system of DMs pwfs. And we apply it in the next section to detect market crash phenomenon in option price data. $B_\delta(p^\star)$ characterizes the stable and unstable pwfs for DMs in the dynamical system (see Appendix Definition 3.C.1) for pwfs based on an invariant manifold theorem that follows.[17]

Stable and unstable probability weighting functions

In nonlinear dynamics the stable manifold theorem (stated in Appendix subsection 3.C.2) plays a key role in identifying stable and unstable fixed points. It essentially decomposes an invariant manifold into stable and unstable components (Chicone, 1999, Ch. 4). This subsection applies the stable manifold theorem in the context of pwfs.

[17] "Informally, a manifold is a subset of $\mathbb{R}^n$ such that, for some fixed integer $k \geq 0$, each point in the subset has a neighborhood that is essentially the same as the Euclidean space $\mathbb{R}^k$ ... Points, lines, planes, arcs, spheres, and tori are examples of manifolds." (Chicone, 1999, p. 28). Alternatively, a manifold implies that every point x in an abstract space X can be mapped into a small ball in $\mathbb{R}^m$. The small ball is a m-manifold (McLennan, 2014, p. 10).

Figure 3.4: Stable pwf fixed points

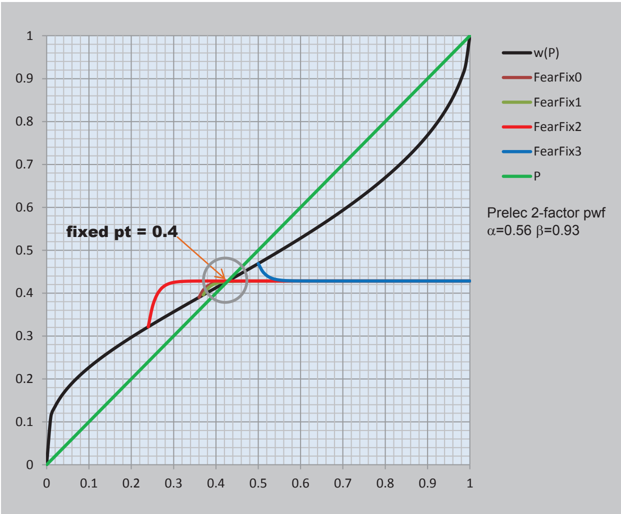


Figure 3.5: Unstable pwf fixed points

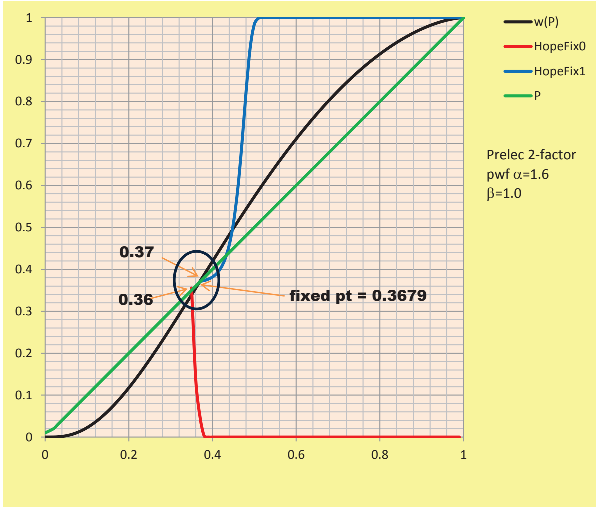


Figure 3.4 depict stable fixed points for Prelec's 2-factor pwf calibrated to parameter estimates for $\alpha = 0.56, \beta = 0.93$ for monthly index option prices data for 1996-2008 in Polkovnichenko and Zhao (2013). The orbit generated by the iterated function system (IFS) $p, w(p), \ldots, w^n(p)$ for that plot converges to the fixed point $p = 0.428292$. This limiting value of p is close to values reported in the experimental economics literature (Prelec, 1998, p. 506). This is a stable attractor because all starting points for the trajectories $FearFix, FearFix1$, $FearFix2, FearFix3$ converge to $p^\star$. This is an empirical realization of the phase portrait in Figure 3.1. By contrast, Figure 3.5 depicts an unstable fixed point for $\alpha = 1.6, \beta = 1$. A small perturbation of the fixed point $0.36 < p^\star = e^{-1} \approx 0.3679 < 0.37$ causes the orbit to jump to $p = 0$ ($HopeFix0$ trajectory) which contains no information or $p = 1$ ($HopeFix1$ trajectory) which represents full certainty. This is an empirical realization of the phase portrait in Figure 3.2. Thus, iterative pwf dynamics is dispositive of pwf [in]stability.

Proposition 3.2.1 (Invariant manifold for probability weighting functions)**.**

Let F be a cumulative probability distribution, and $w(F)$ be a probability weighting functional. Define the set

$$C(F) = \{F \mid -w(F)\ln(w(F)) = F,\ 0 \le F \le 1\} \tag{3.2.1}$$

Then $C(F)$ is an invariant set of fixed point functions for probability weighting. Moreover, in the restricted case when $w(F) = F$ we get $C(F) = \{0, e^{-1}, 1\}$ where $F = p^\star = e^{-1}$.

Proof. See Appendix 3.D □

Remark 3.2.1. McLennan (2014, Fig. 1.2, pg. 8) identified sets like $C(F)$ as an *essential set* of fixed points. Alternatively, $C(F)$ is an invariant subspace of $[0,1]$. That is, $w : C(F) \to C(F) \subset [0,1]$ and

$C(F)$ is an invariant manifold.

Corollary 3.2.2 (Invariant manifold decomposition)**.** *$C(F)$ is decomposable into stable (S) and unstable (U) submanifolds such that $U \oplus S = C(F)$.*

Proof. Apply the stable manifold theorem in section 3.C to $C(F)$ in Propsition 3.2.1. □

Remark 3.2.2. Since w is defined on $C(F)$ the invariance decomposition property implies $w(S(F)) \subset S(F)$ and $w(U(F)) \subset U(F)$. Refer to section 3.C for further details. □

Corollary 3.2.3 (Linear Probability Weighting Operator)**.** *There exist a linear probability weighting operator $\widehat{T}$ on $C(F)$ separating expected and nonexpected utility theories.*

Proof. See Appendix 3.E □

The phase diagrams in Figure 3.1 and Figure 3.2 depict the stable manifold theorem's decomposition of the invariant manifold in Proposition 3.2.1. Figure 3.4 and Figure 3.5 depict the stable manifold theorem at work in the stable and unstable probability weighting functions implied by index options.

Local Lyapunov exponent for probability weighting functions

In this subsection we formally define the Lyapunov exponent for pwfs. Intuitively, a Lyapunov exponent λ is a measure of the rate of system divergence or convergence when two nearby initial conditions are compared.

Definition 3.2.1 (Lyapunov exponent)**.** Jost (2005, pg. 31). Let $w(p)$ be a probability weighting function such that the first derivative w' exist. The Lyapunov exponent of the orbit $p_n = w(p_{n-1})$, $n \in \mathbb{N}$ for $p_0 = p$ is

$$\lambda(p) := \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \ln|w'(p_j)| \tag{3.2.2}$$

provided the limit exist.

This definition implies that the Lyapunov exponent is an invariant of the Jacobian $w'(p_j) = \frac{\partial w}{\partial p_j}$ that determines local stability of the points that satisfy (3.2.2). It is the average rate of divergence for the *iterative function system* $p, w(p), w \circ w(p), \ldots, w \underbrace{\circ \ldots \circ}_{(n-1)\,\text{times}} w(p)$ where $w \circ w(p) = w^2(p)$, $(w \circ w \circ w)(p) = w \circ w^2(p) = w^3(p)$ and so on.

For purposes of exposition, we consider the 2-parameter probability weighting function introduced by Prelec (1998, Prop. 1, pg. 503):[18]

$$w(p) = \exp(-\beta(-\ln(p))^{\alpha}),\ 0 < \alpha < 1,\ \ \beta > 0 \tag{3.2.3}$$

After log differentiation we get

$$\ln[w'(p)] = a(p;\alpha,\beta) = \ln(\alpha\beta) + (\alpha - 1)\ln(-\ln(p)) - \ln(p) - \beta(-\ln(p))^{\alpha} \tag{3.2.4}$$

Monotonicity of $w(p)$ guarantees that $w'(p) > 0$ so the absolute value requirement in (3.2.2) is satisfied. However, the true probability weighting function $w(p)$ is unknown, so the parameters α and β are unobservable in phase space.

3.2.1 Stochastic Lyapunov exponent process in econometrics theory

The stochastic Lyapunov exponent concept was presented in nonlinear time series analysis in the early 1990s via its estimation with nonparametric regressions. Refer to Nychka et al. (1992); McCaffrey et al. (1992) and references therein. Important papers by Bougerol and Picard (1992); Whang and Linton (1999); Shintani and Linton (2004) extended the concept to the econometrics theory literature. Recently, Park and Whang (2012, p. 64) introduced a nonparametric test for random walk against a chaos alternative. In their model the sample estimate for the Lyapunov exponent process of interest is a Brownian functional[19]

$$\lambda_n(t) = \int_0^t \ln|m_n^o(\sqrt{n}B_n(s))|ds,\ \ t \in [0,1] \tag{3.2.5}$$

[18] In Appendix 3.F we show how the robustness of Prelec's model can be extended to other popular 2-parameter models like that by Goldstein and Einhorn (1987) (which is analyzed extensively in Lattimore et al. (1992)). We note in passing that Stott (2006) found that Prelec's 2-parameter pwf is "the best model" for pwfs in his metastudy of function forms in CPT.

[19] Refer to Karatzas and Shreve (1991, p. 185) for technical details on this concept.

where m_n^o is the first derivative of a Nadaraya-Watson kernel estimator for the nonparametric nonlinear function $m_n(\cdot)$, defined on the space of continuous function $C[0,1]$ on the closed unit interval, endowed with the sup norm metric, and $B_n(t) \in C[0,1]$ is approximate Brownian motion. BenSaïda (2012); BenSaïda (2014) applied the tests above to financial time series that include the S&P 500 index and failed to find chaos in the data. In the sequel, our sample function for the Lyapunov exponent process is also a Brownian functional of $B_n(t)$ but its "kernel estimator" m_n^o is parametric.

3.2.2 Representation theorem for behavioural Lyapunov exponent process

Consider a large sample of N heterogenous decision makers (DMs). Let ε_i, $i = 1,\ldots,N$ be the measurement error associated with the choice made by the i-th DM. Furthermore, assume that $\varepsilon_i \sim iid(0,\sigma^2)$. We assume a common core belief in (3.2.4) so heterogeneity in the model is represented by appending ε_i to (3.2.4)[20] so that

$$a^j(p;\alpha,\beta) = a(p;\alpha,\beta) + \varepsilon_j \tag{3.2.6}$$

$$\ln[w'^j(p,\alpha,\beta)] = a^j(p;\alpha,\beta) \tag{3.2.7}$$

Let $[0,T]$ be the finite time interval for which Lyapunov exponents are observed for each DM. Without loss of generality we normalize the time interval to coincide with $[0,1]$ and let $\Pi^{(n)} = \{0,t_1^{(n)},t_2^{(n)},\ldots,t_k^{(n)},\ldots,1\}$ partition $[0,1]$ into dyadic intervals such that $t_k^{(n)} = k.2^{-n}$.

Consider the cumulative effect of DMs measurement errors at time $t \in [t_k^{(n)},t_{k+1}^{(n)})$ defined

[20] Technically, this is more like measurement error. However, we could model heterogeneity as a random effect such as $\alpha_i = \mu_i + \varepsilon_{ij}$ where $\mu_i \sim (0,\sigma^2)$ and ε_{ij} is a "treatment effect" (e.g., Kutner et al. (2005, pg. 1031); Shi et al. (2012, p. 323)). Nonetheless, in the sequel we provide interval estimates for α and β which satisfy the heterogeneity criterion *a fortiori* (e.g., Hey and Orme (1994, p. 1301); Loomes (2005, p. 305)).

by the partial sums[21]

$$S_{nt}^{j}=\sum_{k=1}^{nt}\varepsilon_j(t_k^{(n)}),\qquad S_{[nt]}^{j}=\sum_{k=1}^{[nt]}\varepsilon_j(t_k^{(n)}) \tag{3.2.8}$$

where $[nt]$ is the integer part of nt. In the psychology and neuroscience literature (3.2.8) is the basis for a random accumulator model (RAM) of decision making over time. Most important, RAMs admit "changes of mind" or reversal when the partial sums hit a given threshold (Resulaj et al., 2009). The random broken line connecting the points $([nt], S_{[nt]}^{j})$ and (nt, S_{nt}^{j}) is given by

$$W_n^j(t)=S_{[nt]}^{j}+(nt-[nt])\varepsilon_j([nt]+1) \tag{3.2.9}$$

By Donsker's Theorem, i.e., functional central limit theorem, we assume $W_n^j(t)$ is an approximate Brownian motion in the space of continuous functions $C[0,1]$.[22] Let $w(t;p,\alpha,\beta)$ be the state of the core pwf at time t. According to Gikhman and Skorokhod (1969, pp. 370-371), by virtue of (3.2.4), the incremental change in time dependent pwf for the j-th DM at time t can be written as

$$\Delta\ln[w'^{j}(t;p,\alpha,\beta)]=a^j(p;\alpha,\beta)\Delta t+\sigma\Delta W_n^j(t) \tag{3.2.10}$$

This paves the way for the following behavioural empirical local Lyapunov exponent (BELLE) process

Theorem 3.2.4 (BELLE process)**.** *Assume that DMs probabilistic risk attitudes are characterized by Prelec (1998) 2-parameter pwf* $w(p)=\exp(-\beta(-\ln(p))^{\alpha})$*. For a given sample size N of DMs whose preferences are measured with behavioural noise* $\varepsilon\sim iid\,N(0,1)$*, the behavioural stochastic*

[21] The partial sums allow us to construct an approximate random function as follows. Suppose $\varepsilon\sim(0,\sigma^2)$. Divide the interval $[0,1]$ into n equal parts i/n, $i=1,\ldots,n$. Define $S_i=\varepsilon_1+\cdots+\varepsilon_n$. Let $W_n(i/n)=\frac{1}{\sigma\sqrt{n}}S_i$. For $t\in[(i-1)/n,\, i/n]$ we interpolate to get

$$W_n(t)=\frac{i/n-t}{1/n}W_n((i-1)/n)+\frac{(t-(i-1)/n)}{1/n}W_n(i/n)=\frac{1}{\sigma\sqrt{n}}S_{i-1}+n(t-(i-1)/n)\,\frac{1}{\sigma\sqrt{n}}\varepsilon_i$$

For t in the half open interval $[(i-1)/n,\, t)$ we have $i-1=[nt]$. So we get

$$W_n(t)=\frac{1}{\sigma\sqrt{n}}S_{[nt]}+(nt-([nt])\,\frac{1}{\sigma\sqrt{n}}\varepsilon_{[nt]+1}$$

[22] This is a common assumption in econometrics theory (White, 2001, Ch. 7) and probability theory (e.g., Serfling (1980, p. 41); Knight (1962); Gikhman and Skorokhod (1969, pp. 452-453); and Karatzas and Shreve (1991, pg. 66)).

Lyapunov exponent process $\bar{\lambda}_N(t;p,\alpha,\beta)$ has the following representation

$$d\bar{\lambda}_N(t;p,\alpha,\beta) = \bar{a}_{m,N}(p;\alpha,\beta)dt + \sigma d\overline{W}_{n,N}(t), \tag{3.2.11}$$

where $\bar{a}_{m,N}(p;\alpha,\beta)$ is a drift term, σ is a volatility coefficient, and $\overline{W}_{n,N}(t)$ is an approximate Brownian motion.

Proof. See Appendix 3.G. □

Remark 3.2.3. This is a behavioural manifestation of the Park and Whang (2012) Brownian functional result in (3.2.5). The *existence and uniqueness* (Gikhman and Skorokhod, 1969, Ch. VIII, §3) of $\bar{\lambda}_N(t;p,\alpha,\beta)$ is implied by Definition 3.2.1.

Stochastic stability condition

Definition 3.2.2 (Stochastic stability)**.** (Gihman and Skorohod, 1972, p. 145). A stationary point $p^\star$ will be called stable if for any $\varepsilon > 0$, there exist $\delta > 0$ such that for $B_\delta(p^\star) = \{p : |p - p^\star| < \delta\}$

$$\Pr\{\lim_{t\to\infty} \xi_p(t) = p^\star | \xi_p(t) \in B_\delta(p^\star)\} \geq 1 - \varepsilon \tag{3.2.12}$$

where $\xi_p(t)$ is a process (possibly stochastic) starting at p. □

Remark 3.2.4. In our model, the fixed point probability $p^\star$ is a stationary point. (3.2.12) imply that initial values for $\xi_p(t)$ converge uniformly to $p^\star$ over time (Arnold, 1984, p. 210). □

After integrating the stochastic differential equation in Theorem 3.2.4 we get the behavioural Brownian functional

$$\bar{\lambda}_N(t;p,\alpha,\beta) = \int_0^t \bar{a}_{m,N}(p;\alpha,\beta)du + \sigma \int_0^t d\overline{W}_{n,N}(u) \tag{3.2.13}$$

$$= \bar{a}_{m,N}(p;\alpha,\beta)t + \sigma(\overline{W}_{n,N}(t) - \overline{W}_{n,N}(0)) \tag{3.2.14}$$

The stochastic Lyapunov exponent stability condition implies negative eigenvalues (e.g., (Leonov and Kuznetsov, 2007; Hommes and Manzan, 2006), and Wiggins (2003, p. 7)) and it is given by the folowing

Lemma 3.2.5 (Stochastic Lyapunov exponent stability condition)**.**

$$\sup_t \bar{\lambda}_N(t; p, \alpha, \beta) < 0 \implies \sup_t \overline{W}_{n,N}(t) < \overline{W}_{n,N}(0) - \frac{1}{\sigma}\bar{a}_{m,N}(p; \alpha, \beta)t, \tag{3.2.15}$$

where $\overline{W}_{n,N}(0) = 0$. □

Thus, in the context of Definition 3.2.2 the stochastic stability condition for pwfs is characterized by (3.2.15).

3.2.3 Estimating the probability of tail event instability

To estimate the probability of stability we rewrite $\overline{W}_{n,N}(t)$ in Lemma 3.2.5 as an approximate Brownian motion

$$\overline{W}_{n,N}(t) \equiv W_n\left(\frac{t}{N}\right) \tag{3.2.16}$$

If $\phi(\cdot)$ is the probability density function for $\overline{W}_{n,N}(t)$, then the probability density function for $\sup_t \overline{W}_{n,N}(t)$ is proportional to $\phi(\cdot)$ (e.g., Gikhman and Skorokhod (1969, pg. 286) and Karatzas and Shreve (1991, pg. 96, Prob. 8.2)). So that the stochastic stability condition in Lemma 3.2.5 is characterized by

$$\Pr\left\{\sup_t \overline{W}_{n,N}(t) < -\frac{1}{\sigma}\bar{a}_{m,N}(p; \alpha, \beta)t\right\} = \Pr\left\{\sup_t W_n\left(\frac{t}{N}\right) < -\frac{1}{\sigma}\bar{a}_{m,N}(p; \alpha, \beta)t\right\} \tag{3.2.17}$$

$$= c_0\Phi\left(-\frac{\bar{a}_{m,N}(p; \alpha, \beta)}{\sigma}\sqrt{Nt}\right) \tag{3.2.18}$$

$$\Rightarrow \Pr\left\{\sup_t \overline{W}_{n,N}(t) \geq -\frac{1}{\sigma}\bar{a}_{m,N}(p; \alpha, \beta)t\right\} = 1 - c_0\Phi\left(-\frac{\bar{a}_{m,N}(p; \alpha, \beta)}{\sigma}\sqrt{Nt}\right) \tag{3.2.19}$$

where c_0 is a constant of proportionality, and $\Phi(\cdot)$ is the cumulative normal distribution that controls numerical probability. Here, $\overline{W}_{n,N}(t)$ induces a Lyapunov-Perron effect[23] with tail event large deviation probability (Dembo and Zeitouni, 1998) of instability given by (3.2.19) in a seemingly

[23] This effect stems from the notion of hyperbolic fixed points and unstable manifolds (Wiggins, 2003, pp. 12, 50). Cursory inspection of Figure 3.1 and Figure 3.2 show that $w(F)$ is hyperbolic in a sufficiently large neighbourhood $B_\delta(p^\star)$ of the fixed point $p^\star$. So it contains an invariant manifold in $\mathbb{R}^2$.

stable system (Leonov and Kuznetsov, 2007, p. 1079).

Theorem 3.2.6 (Probability of tail event instability for a seemingly stable pwf)**.** *Assume that a large sample size N of DMs have local stochastic Lyapunov exponent process represented by Theorem 3.2.4. Then the large deviation probability of tail event instability for a seemingly stable pwf is given by*

$$Pr\Big\{\sup_t \overline{W}_{n,N}(t) \geq -\frac{1}{\sigma}\bar{a}_{m,N}(p;\alpha,\beta)t\Big\} = 1 - c_0\Phi\Big(-\frac{\bar{a}_{m,N}(p;\alpha,\beta)}{\sigma}\sqrt{Nt}\Big)$$

$$\Longrightarrow \lim_{N\to\infty}\frac{1}{N}\log Pr\Big\{\sup_t \overline{W}_{n,N}(t) \geq -\frac{1}{\sigma}\bar{a}_{m,N}(p;\alpha,\beta)t\Big\} = -c_0\frac{\bar{a}_{m,N}(p;\alpha,\beta)}{\sigma}t$$

where Φ is the cumulative normal and c_0 is a constant of proportionality. □

Cursory inspection of (3.2.19) shows that given α,β,σ at time t the probability of instability increases since $\Phi(\cdot)$ gets smaller as N gets larger. The same also holds for fixed N and increasing t. For technical details on large deviation analysis see Dembo and Zeitouni (1998).

Chen and Xuefeng (2003, pp. 420, 423) introduced a market microstructure model of liquidity in a stock market with fundamentalist and chartists who have different confidence levels in prices reflected in their subjective probabilities. They show that the Lyapunov exponent, which characterize chaotic dynamics in their paper, depends critically on the number of traders in the system. Day and Huang (1990, Prop. 3, p. 319) refer to a critical mass of noise traders in such a system as "market sheep". Thus, there is *inherent heterogeneity* associated with N in financial markets. In this paper the Lyapunov exponent process also depends on the number N of DMs and probabilistic risk attitude factors. Thus, we establish a nexus between behavioural Lyapunov exponent and probabilistic risk attitudes.

3.2.4 Impact of the drift term on sign reversal of BELLE process

To evaluate the impact of the other control variables on Lyapunov-Perron type probability of instability we turn to comparative statics. Rewrite the drift term in Theorem 3.2.4 for given p so that

$$f(\alpha,\beta;p) = \ln(\alpha\beta) + (\alpha-1)\ln(-\ln(p)) - \ln(p) - \beta(-\ln(p))^{\alpha} \tag{3.2.20}$$

$$\frac{\partial f}{\partial \alpha} = \alpha^{-1} + \ln(-\ln(p)) - \beta(-\ln(p))^{\alpha+1} \tag{3.2.21}$$

$$\frac{\partial f}{\partial \alpha} > 0 \Rightarrow 0 < \beta < \frac{\frac{1}{\alpha} + \ln(-\ln(p))}{(-\ln(p))^{\alpha+1}} \tag{3.2.22}$$

Similarly,

$$\frac{\partial f}{\partial \beta} = \frac{1}{\beta} - (-\ln(p))^{\alpha}, \quad \frac{\partial f}{\partial \beta} > 0 \Rightarrow 0 < \beta < (-\ln(p))^{-\alpha} \tag{3.2.23}$$

The first order effects for increasing drift (and hence increased probability of instability) in (3.2.22) and (3.2.23) is given by

$$0 < \beta < \max\left\{\frac{\alpha^{-1} + \ln(-\ln(p))}{(-\ln(p))^{\alpha+1}},\ (-\ln(p))^{-\alpha}\right\} \tag{3.2.24}$$

Since α controls the curvature of $w(p)$, it determines the degree of DM's confidence. Whereas β is an elevation parameter that controls (1) the location of fixed point probability in the underlying probability distribution, and (2) degree of "cautiously hopeful" behaviour (Lopes, 1995, p. 187). So (3.2.24) depicts the range of elevation that control the hyperbolic fixed points for invariant manifolds that support stability and instability. Refer to Chicone (1999, p. 28) and Wiggins (2003, Ch. 3) for details on hyperbolic fixed points. Chen and Xuefeng (2003, p. 423) used a similar analytic apparatus to identify the conditions under which chaos appear in their model. In the case of Prelec (1998) single factor model, i.e., $\beta = 1, 0 < \alpha < 1$, we find that the set of feasible values in (3.2.24) for curvature α are solutions to the nonlinear equation

$$\frac{\alpha^{-1} + \ln(-\ln(p))}{(-\ln(p))^{\alpha+1}} > 1 \tag{3.2.25}$$

$$\Rightarrow (-\ln(p))^{\alpha+1} - \alpha^{-1} - \ln(-\ln(p)) < 0 \tag{3.2.26}$$

We summarize the result above in

Proposition 3.2.7 (Critical values of probabilistic risk factors)**.** *Given a large sample N of DMs* *with core Prelec (1998) 2-parameter pwfs (α and β) in a dynamical system with behavioural stochastic Lyapunov exponent process in psychological space characterized by Theorem 3.2.4, the tail event probability of instability in Theorem 3.2.6 depends on either of the following*

1. *growth in sample size N or time t or both;*
2. *curvature (α) and elevation (β) of pwfs that induce the range of critical probabilistic risk attitude factors*
$$0 < \beta(p) < \max\{\varphi_1(\alpha, p), \varphi_2(\alpha, p)\};$$
3. *increased precision in the diffusion coefficient σ for classifying measurement error or behavioural noise by DMs;*

where $\varphi_1(\alpha, p) = \dfrac{\alpha^{-1} + \ln(-\ln(p))}{(-\ln(p))^{\alpha+1}}$ *and* $\varphi_2(\alpha, p) = (-\ln(p))^{-\alpha}$. □

3.3 Market instability identified by behavioural Lyapunov exponent process

In this section we plot and describe the source functions, i.e., pwfs, implied by option prices from parameter estimates in Polkovnichenko and Zhao (2013). We calibrate analytic expressions from our criterion function for market instability in Proposition 3.2.7, and compare the predictions of the theory to historic events in option price behavior.[24] The following definition is adapted from Wakker (2010, p. 320) and it plays a key role below.

Definition 3.3.1 (Source function)**.** Assume that all uncertainties can be quantified in terms of probabilities, and that a source is a specific set of events. For each event E define $W(E)$ as $w_{\mathbb{S}}(P_{\mathbb{S}}(E))$ where $\mathbb{S}$ is the source from which E obtains, $P_{\mathbb{S}}$ is a probability measure on $\mathbb{S}$, and $w_{\mathbb{S}}$ is the pwf corresponding to E. We call $w_{\mathbb{S}}$ a *source function*. □

[24] Constantinides (1990) and Abel (1990) used a similar calibration methodology for analytic expressions in their models to illustrate their resolution of the equity premium puzzle.

Figure 3.6: Duke/CFO Magazine Global Optimism Index

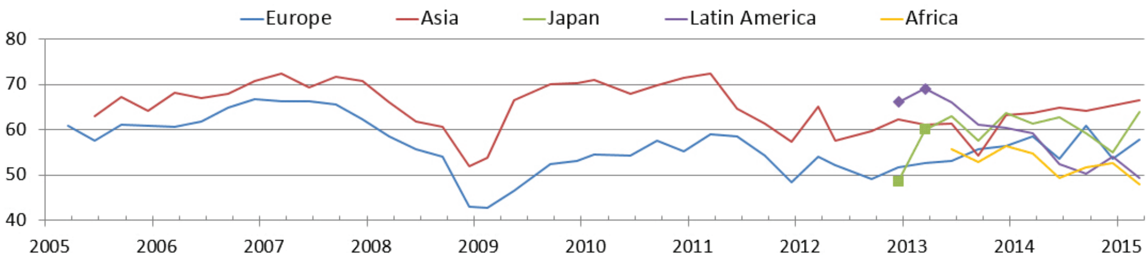
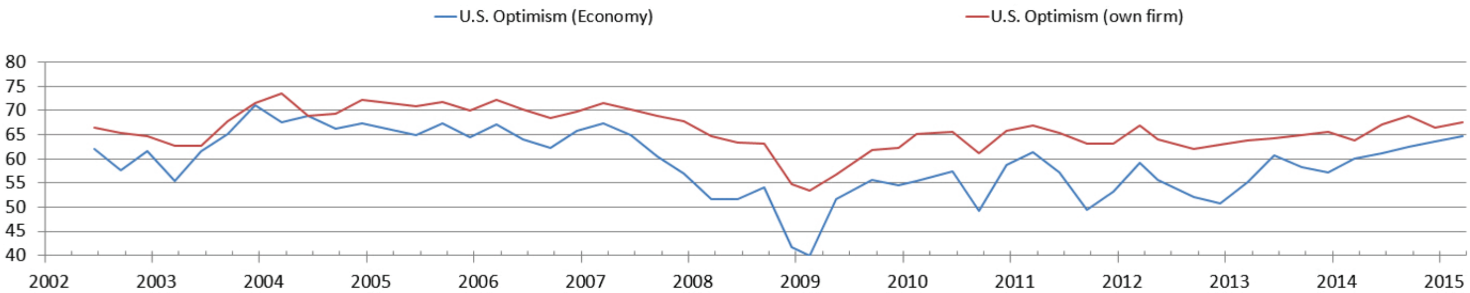


Figure 3.7: Duke/CFO Magazine US Optimism Index



Source: Duke/CFO Business Outlook Survey at http://www.cfosurvey.org/.

Cursory inspection of Figure 3.7 plotted from the Duke/ CFO Magazine Survey on optimism shows that US CFOs optimism scores are uniformly higher for their firms' prospects compared to that for the economy. Evidently CFOs are more uncertain about the economy than they are for their own firms. Therefore, we would expect CFO source function for the economy to be above that for their firms. Cf. Hogarth and Einhorn (1990). CFO Global optimism in Europe, Japan, Latin America and Africa also show distinct patterns in Figure 3.6. Therefore, the source functions for those CFOs will be different.

3.3.1 Calibrating credit risk sources of pwfs implied by S&P 500 index option

Polkovnichenko and Zhao (2013, pg. 595, Fig. 5) derived estimates of Prelec (1998) 2-parameter pwf for shape parameter α and elevation parameter β assuming CRRA utility $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$ and $\gamma = 2.0$, for $S\&P500$ option price data over the sample period January 1996 to December 2008. The source functions with those α and β values are plotted in Figure 3.8 for $p = 0$ to $p = 1$ in steps of 0.01. We describe them as follows.

On June 19, 1997 there was pessimism in the option market depicted by the concave-convex inverted S-shape curve $(\alpha = 0.56, \beta = 0.93)$ in Figure 3.8. This was around the time of the Asian currency crisis (Mishkin, 1999; Corsetti et al., 1999) so we label that *credit risk source function* $W^{\text{Asia}}_{\text{currency}} = w^{\text{Asia}}_{1997}(p; \alpha = 0.56, \beta = 0.93)$.

By April 21, 2005 the state changed to optimism depicted by the convex-concave skewed S-shape curve $(\alpha = 1.6, \beta = 1)$ in Figure 3.8. This time there was a real estate bubble in the US (Zhou and Sornette, 2006) driven by sub-prime loans and asset securitization.[25] So we label that *credit risk source function* $W^{\text{US}}_{\text{RealEstateCDO}} = w^{\text{US}}_{2005}(p; \alpha = 1.6, \beta = 1.0)$.

Early warning critical values of probabilistic risk factors

Undeniably, option market sentiment had a *phase transition from pessimism to optimism* between 1997 and 2005. The $\beta(p)$-instability distribution predicted by Proposition 3.2.7 is plotted in Figure 3.9 for $\beta(p) = \min\{\max\{\varphi_1(\alpha, p), \varphi_2(\alpha, p)\}, \beta_c\}$ where $\varphi_1(\alpha, p) = \dfrac{\alpha^{-1} + \ln(-\ln(p))}{(-\ln(p))^{\alpha+1}}$ and $\varphi_2(\alpha, p) = (-\ln(p))^{-\alpha}$ and β_c is the observed pwf elevation parameter, i.e., the β parameter estimated by Polkovnichenko and Zhao (2013) for $W^{\text{Asia}}_{\text{currency}}$ and $W^{\text{US}}_{\text{RealEstateCDO}}$. We construe β_c as the "true value". The plot excludes values of $\beta(p) > \beta_c$ since they are inadmissible, so it is truncated accordingly as depicted. A quadratic curve in p was fitted for $\beta(p)$ as indicated. According to Figure 3.9, the region of instability for both curves in Figure 3.8 is supported by probability values less than the fixed point probability $p^* = 0.4$. For example, $0 < p \leq 0.4$ for $W^{\text{Asia}}_{\text{currency}}$ and $W^{\text{US}}_{\text{RealEstateCDO}}$. This is the probability support for low ranked option prices. It reflects low quality assets, long shot bias, and high risk assets. Figure 3.9 shows that DMs with $\alpha = 1.6$ and $0.26 \leq \beta(p) < 0.74$ are prone to induce instability in states of optimism or overconfidence when $\beta_c = 1$. In contrast, DMs

[25] Adelson (2013) argues that sub prime loans and mortgages "may have served as the spark that ignited the powder keg" of underlying causes of the Great Recession of 2008 but it was not the cause. The Financial Crisis Inquiry Comission (2011, Ch. 10) referred to this period as "The Madness".

Figure 3.8: Source functions implied by S&P index option prices

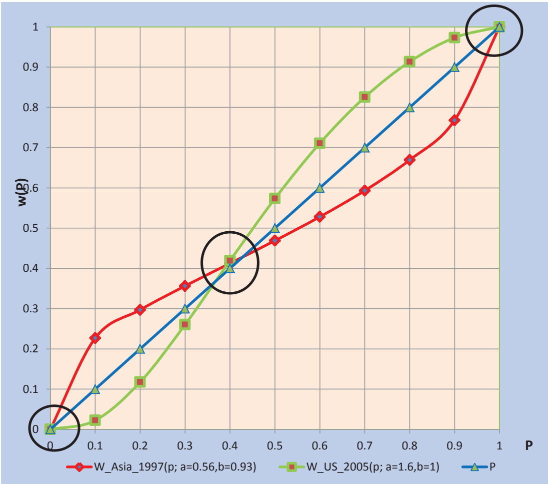


Figure 3.9: $\beta(p)$-instability distribution with tipping points at p=0.4

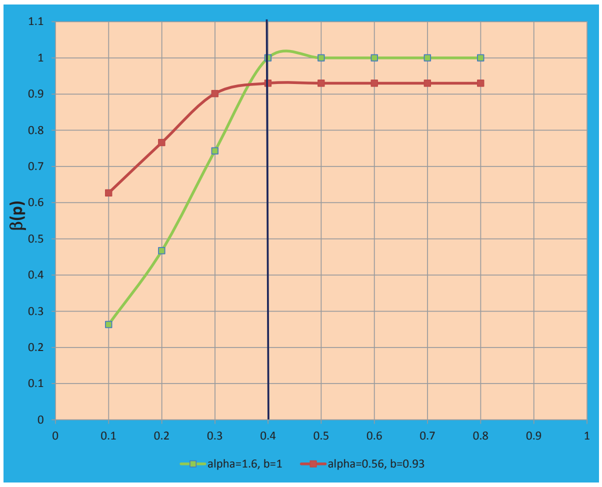


Figure 3.10: $\beta(p)$ instability (0.6, 0.2) for 1997 Asia and 2005 US risk sources

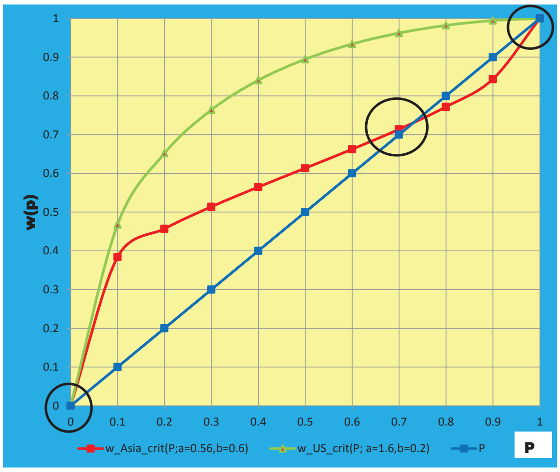


Figure 3.11: $\beta(p)$-instability (0.7, 0.5) for 1997 Asia and 2005 US risk sources

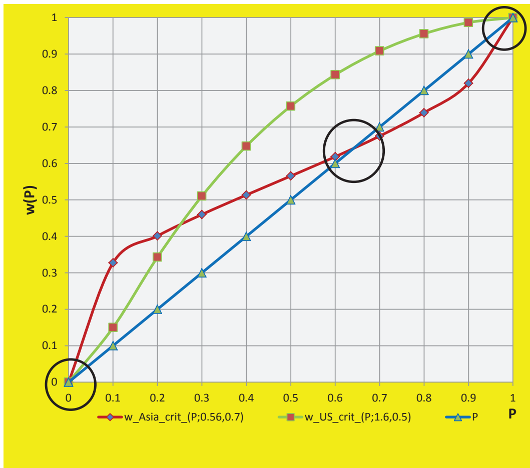


Figure 3.8 is calibrated with Polkovnichenko and Zhao (2013) estimates for Prelec (1998) 2-parameter source functions for 1997 Asian currency crisis $(\alpha = 0.56, \beta = 0.93)$ and 2005 US real estate and CDO bubble $(\alpha = 1.6, \beta = 1)$ for CRRA parameter $\gamma = 2$. Figure 3.9 depicts the distribution of critical values for early warning and tipping point for market crash when p=0.4 for each source function. Figure 3.10 and Figure 3.11 depict source functions for market crash for 2005 US real estate and CDO $\beta(p) = 0.2$ and 0.5 resp., and market instability for 1997 Asian currency $\beta(p) = 0.6$ and 0.7 resp.

with $\alpha = 0.56$ and $0.63 \leq \beta(p) < 0.93$ are prone to induce instability in states of pessimism or underconfidence when the true $\beta_c = 0.93$. According to Polkovnichenko and Zhao (2013, p. 585) β shifts the distribution and α mainly affects the tails.

3.3.2 Calibrated source functions predicted by tipping point value

In the analysis that follows we reiterate that pwfs are identified with the source of credit risk so they are source functions in accord with Definition 3.3.1. Table 3.1 presents parameter values used in the analysis below.

Table 3.1: Parameter values for credit risk source functions

Parameter	Asia 1997 Currency Crisis	US 2005 Real Estate & CDO Bubble
α-curvature	0.56	1.60
β-elevation	0.93	1.00
$p^{\star}$-fixed point	0.40	0.40
$\beta(p)$-instability	0.60	0.20
	0.70	0.50
$\beta(p)$-tipping point	0.93	0.67

The plots corresponding to the parameter values above are depicted in Figure 3.8, Figure 3.9, Figure 3.10 and Figure 3.11. Tipping point values for $\beta(p^{\star})$ are depicted in Figure 3.9.

Source function for US real estate and CDO bubble circa 2005

To illustrate our theory for US real estate and CDO source functions, we select $\beta(p)$-instability values $\beta = 0.25$ and $\beta = 0.5$ from the distribution of critical values predicted by Proposition 3.2.7 to characterize dynamics of the underlying source functions. The orientation of the skew S-shaped source function $W^{\text{US}}_{\text{RealEstateCDO}}$ for $(\alpha = 1.6, \beta = 1.0)$ in Figure 3.8 switched to an all concave shape in Figure 3.10 depicting *ex ante* $W^{\text{US, crash}}_{\text{RealEstateCDO}}(p; crit(\alpha = 1.6, \beta = 0.2))$ for fixed $\alpha = 1.6$, and critical value $\beta = 0.2$. So the relative strength of DMs confidence is such that they are now uniformly fearful and pessimistic over the entire range of rank ordered option prices. Similarly, in Figure 3.11 the *ex ante* source function $W^{\text{US, crash}}_{\text{RealEstateCDO}}(p; crit(\alpha = 1.6, \beta = 0.5))$ is concave for $\alpha = 1.6$ and critical value $\beta = 0.5$. Cursory inspection of Figure 3.10 and Figure 3.11 show $W^{\text{US, crash}}_{\text{RealEstateCDO}}(p; crit(\alpha = 1.6, \beta = 0.2)) > W^{\text{US, crash}}_{\text{RealEstateCDO}}(p; crit(\alpha = 1.6, \beta = 0.5))$. According to Hogarth and Einhorn (1990, Fig. 2, Fig. 4, pp. 786-787) the higher curve implies greater ambiguity in the market. Nonetheless, each of those functions symbolize market crash since DMs are uniformly pessimistic.

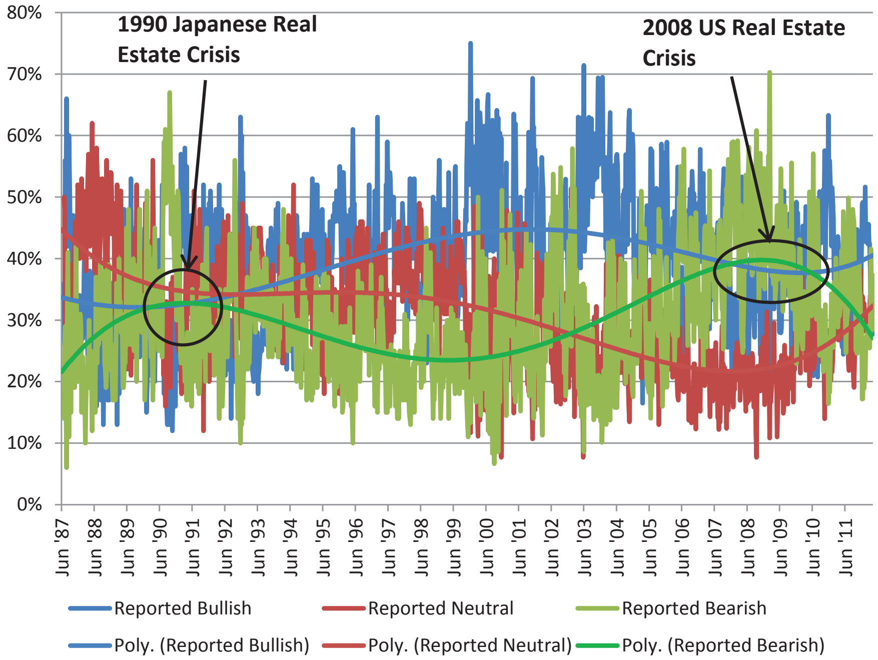
Source function for Asian currency crisis circa 1997

The elevation of the inverse S-shaped source function $W^{\text{Asia}}_{1997}(p;\,\alpha = 0.56, \beta = 0.93)$ in Figure 3.8 for $\alpha = 0.56, \beta = 0.93$ has a fixed point probability $p^* = 0.4$. However, when $\beta = 0.6$ the fixed point probability jumped from 0.4 to about $p^* = 0.75$ in Figure 3.10. According to Hogarth and Einhorn (1990, Fig. 2, pp. 785-786), this northeast movement of the fixed point $p^* = 0.75$, associated with *ex ante* source function $W^{\text{Asia}}_{1997}(p;\, crit(\alpha = 0.56, \beta = 0.6))$, implies larger anticipated losses. In this case the market was more cautious than hopeful. In contrast, when $\alpha = 0.56$ and $\beta = 0.7$ the fixed point probability for *ex ante* source function $W^{\text{Asia}}_{1997}(p;\, crit(\alpha = 0.56, \beta = 0.7))$ falls to about $p^* = 0.7$ in Figure 3.11. In this case, the market was less cautious and more hopeful compared to when $p^* = 0.75$. Each one of the source function plots in Figure 3.10 and Figure 3.11 depict probabilistic preference reversal relative to the corresponding plots in Figure 3.8. Evidently, $\beta(p)$ elevation shifts the underlying distribution for given sentiment reflected by α curvature.

Out-of-sample prediction of 2008 Great Recession market crash

Figure 3.10 and Figure 3.11 show that DMs in the market are pessimistic over most or all option prices. Financial market crash is predicted by the *ex ante* concave source functions $W^{\text{US, crash}}_{\text{RealEstateCDO}}(p;\, crit(\alpha = 1.6, \beta = 0.2))$ and $W^{\text{US, crash}}_{\text{RealEstateCDO}}(p;\, crit(\alpha = 1.6, \beta = 0.5))$ for the critical $\beta(p)$ values $0.2, 0.5$. In Figure 3.5 those fixed points are unstable attractors. They coincide with the case of fixed point probabilities 0 or 1 in Proposition 3.2.1. So we would

Figure 3.12: AAII weekly sentiment survey: June 1987–June 2011



"The AAII Investor Sentiment Survey measures the percentage of individual investors who are bullish, bearish, and neutral on the stock market for the next six months; individuals are polled from the ranks of the AAII membership on a weekly basis. Only one vote per member is accepted in each weekly voting period." Source: http://www.aaii.com/SentimentSurvey?adv=yes. A 5th-degree polynomial smoother was used to generate sentiment waves.

expect a small volume of trade or no trade at all, because the market breaks down since all traders are uniformly pessimistic about asset quality (e.g., Akerlof (1970); Harris and Raviv (1993) and Stiglitz and Grossman (1976, p. 250)). The uniform fear predicted by our $\beta(p)$-instability criterion function is characteristic of market crashes of the type that led to the Great Recession of 2008. Alternatively, it signifies heightened uncertainty in the market.[26] For example, on or about September 13, 2008 just before Lehman Brothers filed for bankruptcy there was tremendous uncertainty about whether the Federal Reserve Bank of New York would step in and bail them out and credit markets froze to a halt (De Haas and Van Horen, 2012).

Figure 3.12 supports the market crash scenarios represented by our source function analysis. It depicts 5th-degree polynomial smoothers for weekly sentiment data for Bulls, Bears, and

[26] Abdellaoui et al. (2011, p. 706) depicted this phenomenon with a convex dashed line curve over gains while Hogarth and Einhorn (1990, Fig. 4, p. 787) depicted this phenomenon as a concave function over losses. The latter depiction is consistent with $W_{2005}^{US}\text{crit}(\alpha, \beta(p))$ in Figure 3.10 and Figure 3.11. Charles-Cadogan (2016b, Appendix A.1) shows how the curvature of pwfs reflect probabilistic risk attitudes like loss aversion for pessimists (concave) and risk seeking for optimists (convex).

Risk neutral survey participants for American Association of Individual Investors data for the period June 1987 to June 2011. During the Great Recession of 2008, induced by the US Real Estate market failure, there was very little difference of opinion between Bulls and Bears. In fact, the Bears were slightly more bullish than the Bulls! Similarly, during the 1990 Japanese Real estate crisis (Peek and Rosengren, 2000) there was no difference of opinion between Bulls and Bears and markets crashed. Our $\beta(p)$-instability analysis predicts that those scenarios are represented by uniformly concave source functions like those in Figure 3.10 and Figure 3.11.

Polkovnichenko and Zhao (2013, Fig. 6, p. 593) provide a time series plot of Prelec's (α, β) reproduced in Figure 3.13. By eyeballing the plot around when the market crashed in 2008, one finds that $\alpha \approx 1.2$ and $\beta \approx 0.93$ for $\gamma = 0$ which corresponds to risk neutrality over outcomes. So the bulk of the risk preference load is carried by the shape of the underlying source function(s). According to Proposition 3.2.7, Figure 3.15 predicts that the tipping point value for $\beta(p)$-instability is $\beta(p) = 0.904095$ when $\alpha = 1.2$ and $p = 0.4$. Furthermore, the fixed point probability $p^* = 0.4$ in Figure 3.8 moved South-West along the diagonal towards $p = 0$. This is consistent with the prediction for unstable pwf in Figure 3.5 where the fixed point $p^* = 0.4$ was perturbed and the iterative function system converged to the South-West corner $p = 0$. Our $\beta(p)$-instability distribution makes the striking *out-of-sample* prediction that for $\alpha = 1.2$ the market will crash at tipping point $\beta(p) = 0.9$.[27]

[27]That tipping point value is different from the tipping point value $\beta(p) = 0.65$ when $\alpha = 1.6$ in Figure 3.9 since β depends on α as input in Proposition 3.2.7.

Figure 3.13: Time series plot for Prelec (1998) (α,β)

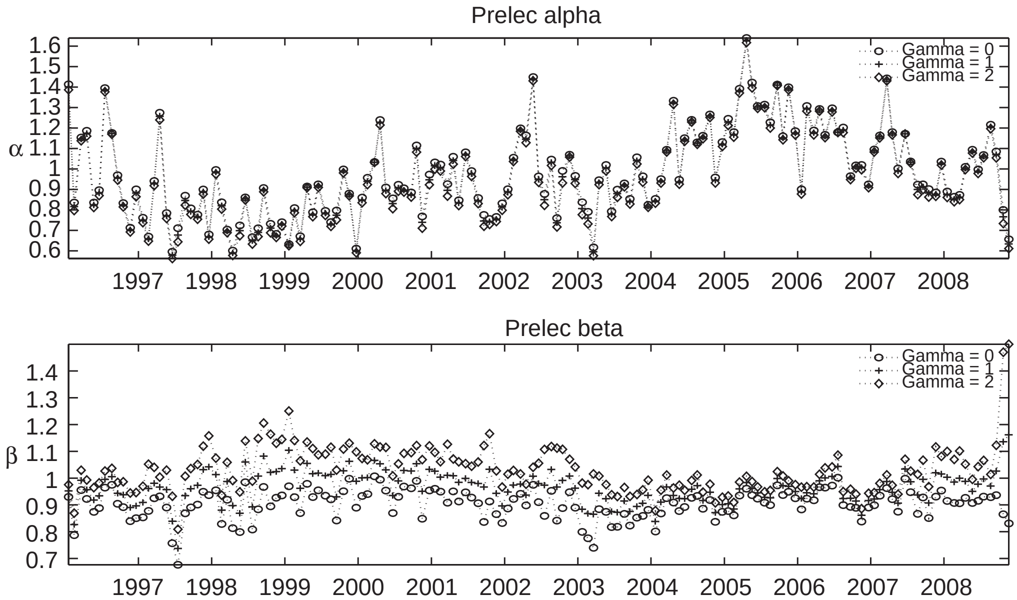


Figure 3.13 is a reproduction of Polkovnichenko and Zhao (2013, Fig. 6, p. 593) time series plot for Prelec (1998) (α, β). The value $\alpha = 1.2$ in 2008 is for CRRA parameters $\gamma = 0, 1$.

Figure 3.14: Source function for 2008 Great Recession Market Crash

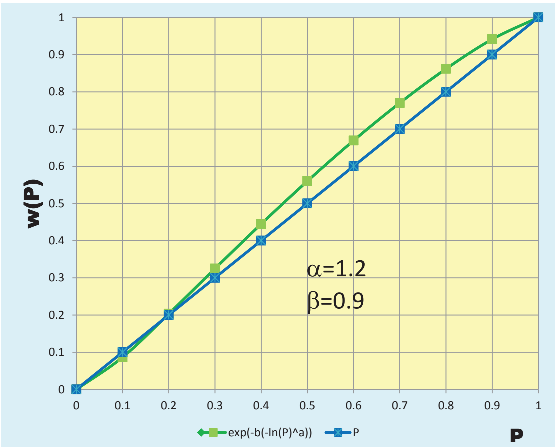


Figure 3.15: $\beta(p)$-instability for 2008 Great Recession Market Crash

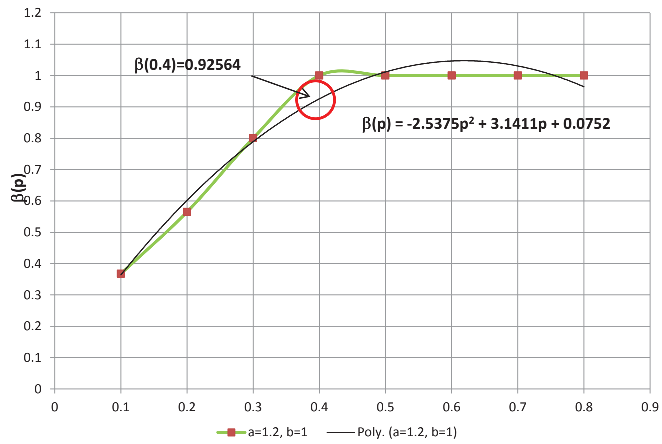


Figure 3.15 depicts the predicted $\beta(p)$-instability *out-of-sample* early warning and tipping point market crash value of $\beta(p) = 0.9$ at p=0.4 for W^{US}_{2005}crit$(\alpha = 1.2, \beta = 1)$. Figure 3.14 depicts the concave source function for the crash value of $\beta(p) = 0.9$ for $\alpha = 1.2$. The fixed point shifted to the left (Cf. Figure 3.5). Eyeballing the plot in Figure 3.13 shows that when the market crashed in 2008 the *ex-post* value was $\beta(p) \approx 0.93$ for $\alpha = 1.2$ and $\gamma = 0$. So our $\beta(p)$-instability *out-of-sample* tipping point value 0.9 in 2005 predicted the crash value 0.93 realized in 2008.

The *ex-post* value in Figure 3.13 is $\beta \approx 0.93$ when the market crashed in 2008. Figure 3.14 depicts the source function for $W^{\text{US}}_{2008}\text{crit}(\alpha = 1.2, \beta = 0.9)$. Qualitatively, the curve is almost uniformly concave–a prerequisite for market crash.[28] Moreover, the market appears to be well calibrated for low ranked option prices since the source function coincides with linear probability weighting for $p \leq 0.2$. Cf. Corollary 3.2.3. We reiterate that Figure 3.13 provides time series plots whereas our $\beta(p)$-instability number was generated by the closed form expression in Proposition 3.2.7.

3.4 Conclusion

We contribute to the literature on financial market instability with a behavioural Lyapunov exponent process for stochastic stability of probability weighting functions (pwfs) implied by index option prices. We show how the shape of pwfs depend on the prevailing source of credit risk in the market so they are "source functions". While our model cannot predict the precise date of a market crash, it predicts critical out-of-sample tipping point values for a single output parameter that predict market crash for given market sentiment input parameter(s). We illustrated the model's robustness across different credit risk sources, e.g., the Asian currency crisis in 1997, US real estate and CDO bubble in 2005, and its striking out-of-sample performance which would have predicted the Great Recession of 2008. Thus, we provide new tools for identifying early warning signals for market instability. Further research includes calibrating the model to provide estimates for the probability of market crash.

3.A APPENDIX

3.B Constructing probabilistic risk attitudes with pwfs: Example

Let $U(x) = \sqrt{x}$ be a concave utility function depicted in Figure 3.16, $\boldsymbol{x} = (4, 9, 16)$ be a distribution of outcomes ranked from worst to best, and $\boldsymbol{p} = (1/2,\ 3/8,\ 1/8)$ be a probability distribution over $\boldsymbol{x}$. Let $L = (4, \frac{1}{2}; 9, \frac{3}{8}; 16, \frac{1}{8})$ be a lottery constructed from the pair $(\boldsymbol{x}, \boldsymbol{p})$.

[28] Our model is based on the assumption that $\gamma = 2$. However, Polkovnichenko and Zhao (2013) reports that there is very little qualitative difference in their empirical pwfs for $0 \leq \gamma \leq 2$.

Figure 3.16: Concave utility function

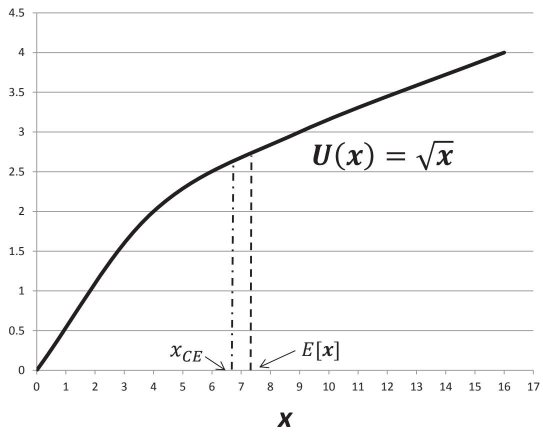
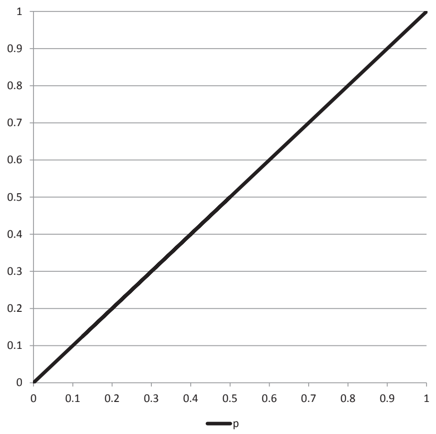


Figure 3.17: Linear probability weighting



Choice under expected utility theory

The von Neumann-Morgenstern utility (Von Neumann and Morgenstern, 1953) is the expected utility over the linear probability distribution in Figure 3.17 and it is given by

$$V(\boldsymbol{p}) = \sum_{i=1}^{3} p_i U(x_i) = (1/2)(2) + (3/8)(3) + (1/8)(4) = 21/8 = 2.625 \tag{3.B.1}$$

In this contrived example, the actuarial value of L is $E[\mathbf{x}] = 59/8$ and $U(E[\mathbf{x}]) \approx 2.7$. So $U(E[\mathbf{x}]) > E[U(\boldsymbol{x})] = V(\boldsymbol{p})$ and $U(x_{CE}) = 2.625 \Rightarrow x_{CE} \approx 6.89$ where x_{CE} is certainty equivalent. Risk averse decision makers (DMs) prefer to receive 6.89 for certain rather than play the lottery L. Their required risk premium to play L is $\varpi = E[\mathbf{x}] - x_{CE} = 0.485$.

Choice under rank dependent utility

Rank dependent utility (RDU) by Quiggin (1982, 1993) posits the existence of a probability weighting function (pwf) $w(p)$ that does not sum to 1 and from which *decision weights* π are constructed as follows. $\pi_1 = w(1/2)$; $\pi_2 = w(1/2 + 3/8) - w(1/2) = w(7/8) - w(1/2)$; and $\pi_3 = 1 - w(7/8)$. Wakker (2010, Ch. 5) provides details on the decision weight procedure. The rank dependent expected utility is given by

$$RDU(\boldsymbol{p}) = \sum_{i=1}^{3} \pi_i U(x_i) = \pi_1 \times 2 + \pi_2 \times 3 + \pi_3 \times 4 \tag{3.B.2}$$

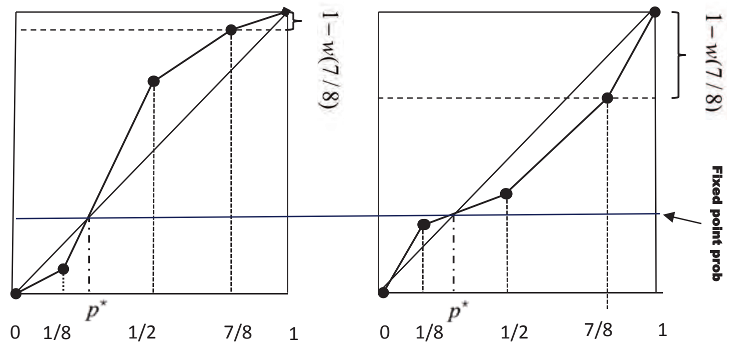
If we sum the decision weights we get $\pi_1 + \pi_2 + \pi_3 = w(1/2) + w(7/8) - w(1/2) + 1 - w(7/8) = 1$. However nonlinearity of w implies $w(7/8) + w(1/8) \neq 1$.

Assumption 3.B.1. *Assume $w(7/8) + w(1/8) > 1$ and either (A) $w(1/8) < 1/8$ or (B) $w(1/8) > 1/8$.*

Under assumption 3.B.1(A) we have $1 - w(7/8) < w(1/8) < 1/8$ and $w(7/8) > 7/8$. Under assumption 3.B.1(B) we have $w(1/8) > \max\{1/8, 1 - w(7/8)\} \Longrightarrow w(1/8) > 1/8$ and $w(7/8) < 7/8$.

The inequality in assumption 3.B.1(B) implies that decision makers (DMs) are *pessimistic* about the lower ranked outcomes because they weight the cumulative probability of those outcomes occurring as if they are greater than they really are. However, even though $w(7/8)$ weights the sum of $p_1 = 1/2$ and $p_2 = 3/8$ it tells us nothing about the individual weights $w(1/2)$ and $w(3/8)$. For example, it is possible that even though $w(7/8) > 7/8$ in assumption 3.B.1(A) we can still have $w(3/8) < 3/8$ and $w(1/2) > 1/2$ or $w(3/8) > 3/8$ and $w(1/2) < 1/2$. (Wakker, 2010, Ch. 7) refers to related scenarios as pessimism and likelihood insensitivity. Substitution of $w(1/8)$ for $\pi_3 = (1 - w(7/8))$ in (3.B.2) highlights the implication of nonlinearity. So now $RDU(\boldsymbol{p}) = 2w(1/2) + 3(w(7/8) - w(1/2)) + 4w(1/8)$ implies $RDU(\boldsymbol{p}) > 3(w(7/8) + w(1/8)) + w(1/8) - w(1/2) > 3 + 1/8 - w(1/2)$ under assumption 3.B.1(B). If $RDU(\boldsymbol{p}) < E[U(\boldsymbol{x})] = V(\boldsymbol{p}) < U(E[\boldsymbol{x}])$, then we must have the valuation $3 + 1/8 - w(1/2) < 21/8 \Rightarrow w(1/2) > 1/2$. Thus, DMs are *pessimistic* and *risk averse* about $x = 4$. If $RDU(\boldsymbol{p}) > U(E[\boldsymbol{x}]) > E[U(\boldsymbol{x})] = V(\boldsymbol{p})$, then $w(1/2) < 1/2$. Thus, DMs are

Figure 3.18: Empirical probability weighting functions



Convex-concave S-shape
Comparatively risk seeking A.1(A)

Concave-convex inverse S-shape
Comparatively risk averse A.1(B)

optimistic and *risk seeking* about $x = 4$ even though the underlying concave utility function $U(x)$ only supports risk averse preferences over x. Thus, DM *sentiment*, i.e., pessimism and or optimism, is driven by the pwf which picks up convexity in preferences whereas $U(x)$ in Figure 3.16 does not. Thus, for $\pi_2 = w(7/8) - w(1/2)$ we have either (a) $\pi_2 < 3/8$ or (b) $\pi_2 > 3/8$. For scenario (a) we have $w(7/8) - w(1/2) < 3/8 \Rightarrow w(7/8) < 3/8 + w(1/2)$. Recall that $w(7/8) > 7/8$ under assumption 3.B.1(A). Thus, we get $7/8 < w(7/8) < 3/8 + w(1/2) \Rightarrow w(1/2) > 1/2$. So now DMs are pessimistic about $x = 4$. However, by hypothesis $w(1/8) > 1 - w(7/8)$ satisfies assumption 3.B.1(A) when $w(1/8) < 1/8$. So we have $\{w(1/8) < 1/8,\ w(1/2) > 1/2,\ w(7/8) > 7/8\}$.

For scenario (b) we have $w(7/8) - w(1/2) > 3/8 \Rightarrow w(7/8) > 3/8 + w(1/2)$. Under assumption 3.B.1(B) $w(7/8) < 7/8$ and $w(1/8) > 1/8$. So combining scenario (b) with assumption 3.B.1(B) we get $3/8 + w(1/2) < w(7/8) < 7/8$. This implies that $w(1/2) < 1/2$. So we have $\{w(1/8) > 1/8,\ w(1/2) < 1/2, w(7/8) < 7/8\}$. The fixed point probability $p^\star$ is the same for each scenario above. Scenarios (a) and (b) imply that the underlying shape of the pwf is as follows:

(a) Concave-convex inverse S-shape: $\{w(1/2) > 1/2;\ \ w(7/8) > 7/8;\ \ w(1/8) < 1/8\}$

(b) Convex-concave S-shape: $\{w(1/2) < 1/2;\ \ w(7/8) < 7/8;\ \ w(1/8) > 1/8\}$

These pwfs shapes in Figure 3.18 are not exhaustive. They are simply two of many that reflect the latent heterogeneity in probabilistic risk attitudes towards the lottery L for given utility function U.

3.C The stable manifold theorem and preliminaries

This appendix provides some preliminaries and an elementary statement of the stable manifold theorem.

3.C.1 Preliminaries

Definition 3.C.1 (Qualities of a dynamical system)**.**

> A dynamical system is a system that evolves in time through the iterated application of an underlying dynamical rule. That transition rule describes the change of the actual state in terms of itself and possibly also previous states. The dependence of the state transitions on the states of the system itself means that the dynamics is recursive. In particular, a dynamical system is not a simple input-output transformation, but the actual states depend on the systems own history. In fact, an input need not even be given to the system continuously, but rather it may be entirely sufficient if the input is only given as an initial state and the system is then allowed to evolve according only to its internal dynamical rule. This will represent the typical paradigm of a dynamical system. (Jost, 2005, p. 1). □

Definition 3.C.2 (Dynamical system)**.** (Rebaza, 2012, p. 327).
Let E be an open set in $\mathbb{R}^n$, i.e., $E \subset \mathbb{R}^n$. The function $\phi : \mathbb{R} \times E \to E$ defined by $\phi(t, \boldsymbol{x}) = \exp(At)\boldsymbol{x}$ defines a dynamical system on E. Specifically, if $\dot{\boldsymbol{x}} = A\boldsymbol{x}$ with initial value $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$, then its solution $\boldsymbol{x}(t) = \exp(At)\boldsymbol{x}_0$ defines how a state $\boldsymbol{x} \in E$ evolves over time. If $\boldsymbol{x} = f(\boldsymbol{x})$ is a nonlinear system, then $A = \boldsymbol{D}f(\boldsymbol{x})$ is the Jacobian for a local linearization. □

Definition 3.C.3 (Invariant subspace)**.** (Rebaza, 2012, p. 338).
A subset $S \subset \mathbb{R}^n$ is called invariant with respect to the system $\dot{\boldsymbol{x}} = A\boldsymbol{x}$ if $\exp(At)S \subset S$. For instance, for any initial value $\boldsymbol{x}(t_0) = \boldsymbol{x} \in S$, the solution $\boldsymbol{x}(t) = \exp(At)\boldsymbol{x}_0$ stays in S for all $t \geq 0$. □

Definition 3.C.4 (Eigenvalue criterion for stability)**.** (Rebaza, 2012, p. 338).
Let $\lambda_j = a_j + ib_j, \quad j = 1, \ldots, n$ be a complex valued eigenvalue of A in Definition 3.C.2 with eigenvectors $\boldsymbol{u}_j = \boldsymbol{v}_j + i\boldsymbol{w}_j$. Then we define

$$E^s = \text{span}\{\boldsymbol{v}_j, \boldsymbol{w}_j : a_j < 0\} \qquad \text{(stable subspace)}$$
$$E^u = \text{span}\{\boldsymbol{v}_j, \boldsymbol{w}_j : a_j > 0\} \qquad \text{(unstable subspace)}$$
$$E^c = \text{span}\{\boldsymbol{v}_j, \boldsymbol{w}_j : a_j = 0\} \qquad \text{(center subspace)}$$

All three subspaces are invariant with respect to the system. Furthermore, they induce the identity

$$E^s \oplus E^u \oplus E^c = \mathbb{R}^n \tag{3.C.1}$$

The $\oplus$ symbol means that for any $\boldsymbol{x} \in \mathbb{R}^n$ we have the decomposition $\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{v} + \boldsymbol{w}$ with $\boldsymbol{u} \in E^s$, $\boldsymbol{v} \in E^u$, $\boldsymbol{w} \in E^c$. □

Definition 3.C.5 (Homeomorphism)**.** A homeomorphism is a function $h : A \to B$ that is a bijection, is continuous and whose inverse is also continuous. □

Definition 3.C.6 (Differentiable manifold)**.** A differentiable manifold of dimension n as a set that is *locally* homeomorphic to the usual Euclidean space $\mathbb{R}^n$. A differentiable manifold is in fact a topological space that generalizes the intuitive and geometric notion of a curve or a surface. Consider for example the one-dimensional space $\mathbb{R}$. (say, the usual x axis). Then a differentiable manifold homeomorphic to it, is the cubic parabola $y = x^3$: It is a continuous deformation of the x axis. □

3.C.2 Statement of stable manifold theorem

In the case of pwfs in this paper the fixed point probability $p^\star$ is an equilibrium point and the neighbourhood $B_\delta(p^\star)$ is differentiable manifold. Without loss of generality, we consider the equilibrium point to be the origin in the following.

Theorem 3.C.1 (Stable manifold theorem)**.** *(Rebaza, 2012, p. 343).*
Let $E \subset \mathbb{R}^n$ be open containing the origin, let $f \in C^1(E)$, and let ϕ_t be the flow $\dot{\mathbf{x}} = f(\mathbf{x})$. Suppose the origin is a hyperbolic equilibrium point and that $A = Df(\mathbf{0})$ has k eigenvalues with negative real part and the remaining $n-k$ eigenvalues have positive real part. Then,

(a) *There exists a k-dimensional differentiable manifold S tangent to E^s at the origin, such that $\phi_t(S) \subset S, \forall t \geq 0$ and $\lim_{t\to\infty} = 0, \forall x \in S$.*

(b) *There exists a $(n-k)$-dimensional differentiable manifold S tangent to E^u at the origin, such that $\phi_t(U) \subset U, \forall t \leq 0$ and $\lim_{t\to\infty} = 0, \forall x \in U$.*

□

3.D Proof of invariant manifold Proposition 3.2.1

Proof. By construction we can rewrite $w(F) = \exp(-\frac{F}{w(F)})$. The distribution function(s) F which solves that nonlinear equation represents fixed probability distributions that satisfy the equation. By definition, F represents a continuum of probabilities. Specifically,

1. let ξ_p be the p-quantile of F. Define the set of probabilities $X(p) = \{p| -w(p)\ln(w(p)) = p,\ F(\xi_p) = P(X \leq \xi_p) = p,\ F \in C(F)\}$. By construction $X(p)$ is a cluster set of probabilities since it contains the accumulation or fixed points p that satisfy the entropy equation, and by construction $X(p) \subseteq C(F)$. Hence $C(F)$ is a hereditary cluster set.

2. Suppose $F \in C(F)$, and $p \notin X(p)$. The latter relation implies that $-w(F(\xi_p))\ln(w(F(\xi_p))) \neq F(\xi_p)$ and $F(\xi_p) \notin C(F)$. This contradicts our incipient hypothesis $F \in C(F)$. In which case $F(\xi_p) = p \in X(p)$ and $C(F) \subseteq X(p)$.

The results of 1. and 2. imply that $C(F) = X(p)$. In which case $C(F)$ is a cluster set of probabilities. The restriction $W(F(\xi_p)) = F(\xi_p) = p$ produces the fixed point solution $W(F) = F = \exp(-1)$. □

3.E Proof of linear probability weighting operator Corollary 3.2.3

Proof. Let $C(F)$ be the domain of definition of an operator $\widehat{T}$ that maps $C(F)$ into itself. Vizly, $\mathscr{D}(\widehat{T}) = C(F)$ and $\widehat{T} : C(F) \to C(F)$. Let $h_1,\ h_2 \in C(F)$ and choose constants θ_1 and θ_2 such that $\theta_1 h_1 + \theta_2 h_2 \in C(F)$. Thus $\theta_1 h_1 + \theta_2 h_2$ has a fixed point representation so by hypothesis

$$\widehat{T}(\theta_1 h_1 + \theta_2 h_2) = \theta_1 h_1 + \theta_2 h_2 = \theta_1 \widehat{T}(h_1) + \theta_2 \widehat{T}(h_2) \tag{3.E.1}$$

Since $\theta_1 h_1 + \theta_2 h_2 \in \mathscr{D}(\widehat{T})$, the relationship in (3.E.1) implies that $\widehat{T}$ is a linear operator by definition. □

3.F Robustness of Prelec (1998) 2-parameter pwf and other 2-parameter pwfs

In this appendix we provide a simple first order correspondence exercise between Prelec (1998) 2-parameter pwf and other popular 2-parameter pwfs by Goldstein and Einhorn (1987) (GE) and its functional equivalent by Lattimore et al. (1992) (LBW).

$$w(p_m) = \frac{\tau p_m^\gamma}{\tau p_m^\gamma + \sum_{k \neq m} p_k^\gamma} = \left(1 + \frac{1}{\tau} \sum_{k \neq m} \left(\frac{p_k}{p_m}\right)^\gamma\right)^{-1} \tag{Goldstein Einhorn}$$

$$w(p) = \exp(-\beta(-\ln p)^\alpha) \tag{Prelec}$$

For a given probability $p = p_m$, a first order expansion of (Goldstein Einhorn) and (Prelec) specifications produces:

$$w(p_m) \approx 1 - \frac{1}{\tau} \sum_{k \neq m} \left(\frac{p_k}{p_m}\right)^\gamma \tag{3.F.1}$$

$$w(p_m) \approx 1 - \beta(-\ln p_m)^\alpha \tag{3.F.2}$$

$$\Longrightarrow 1 - \beta(-\ln p_m)^\alpha \equiv 1 - \frac{1}{\tau} \sum_{k \neq m} \left(\frac{p_k}{p_m}\right)^\gamma \tag{3.F.3}$$

The relationship in (3.F.1) and (3.F.2) imply that

$$\beta = \frac{1}{\tau}, \qquad \alpha = \frac{\ln\left(\sum_{k \neq m} \left(\frac{p_k}{p_m}\right)^\gamma\right)}{\ln((-\ln p_m))} \tag{3.F.4}$$

where β and α are elevation and curvature parameters. Lattimore et al. (1992, Figure 1a, pp.˙380-381) provide examples of how the shapes generated by (Goldstein Einhorn) model correspond to those in Prelec (1998, Figure 2). al Nowaihi and Dhami (2006) show how (3.F.4) can be modified to find α values for what they call "higher order Prelec" pwfs that cut the diagonal (45^o line) more than once. Cavagnaro et al. (2013) provide a mechanism for selecting among competing pwfs. Additionally, Blavatskyy (2013, p. 15) conducted robustness checks between a 2-parameter cubic pwf he introduced, a 1-parameter Prelec pwf, power function for pwf, Goldstein-Einhorn pwf, and Tversky-Kahneman pwf. He found that "[f]or a great majority of subjects the goodness of fit” between between his cubic function “is not significantly different from that of [the other] functions”. Thus, we conclude that Prelec (1998) 2-parameter specification is robust.

3.G Proof of representation theorem for behavioural stochastic Lyapunov exponent process Theorem 3.2.4

Proof. The aggregate change in pwf for heterogeneous DMs in the sample size N is given by

$$\sum_{j=1}^{N} \Delta \ln[w'^{j}(t;p)] = \sum_{j=1}^{N} a^{j}(p;\alpha,\beta)\Delta t + \sigma \sum_{j=1}^{N} \Delta W_{n}^{j}(t) \tag{3.G.1}$$

Substituting $\Delta \ln[w'^{j}(t;p,\alpha,\beta)]$ for $\ln|w'(p_j)|$ in (3.2.2), and by virtue of the continuous mapping theorem (White, 2001, Thm. 7.20, p. 178) replacing Δt and ΔW_n with dt and dW_n, respectively, we get in the limit[29]

$$\frac{1}{m}\sum_{j=1}^{N}\sum_{r=1}^{m} d\lambda^{j}(t;p,\alpha,\beta) = \frac{1}{m}\sum_{j=1}^{N}\sum_{r=1}^{m} a^{j}(p_r;\alpha,\beta)dt + \frac{1}{m}\sum_{j=1}^{N}\sum_{r=1}^{m} dW_{n}^{j}(t) \tag{3.G.2}$$

Dividing left hand side (LHS) and right hand side (RHS) by N and using “bar” to represent sample average, we get the stochastic Lyapunov exponent process

$$d\bar{\lambda}_N(t;p,\alpha,\beta) = \bar{a}_{m,N}(p;\alpha,\beta)dt + \sigma d\overline{W}_{n,N}(t), \tag{3.G.3}$$

$$\bar{\lambda}_N(\cdot) = \frac{1}{N}\sum_{j=1}^{N}\lambda^{j}(\cdot), \quad \bar{a}_{m,N}(\cdot) = \frac{1}{N}\frac{1}{m}\sum_{j=1}^{N}\sum_{r=1}^{m} a^{j}(p_r;\alpha,\beta), \quad \overline{W}_{n,N}(t) = \frac{1}{N}\sum_{j=1}^{N} W_{n}^{j}(t) \tag{3.G.4}$$

□

[29]There is no p_r term on the LHS by definition.

Chapter 4

Harmonic Probability Weighting Functions And The Preference Reversal Puzzle

> "The gamble has been to decision research what the fruit fly has been to biology - a vehicle for examining fundamental processes with presumably important implications outside the laboratory". Paul Slovic.

4.1 Introduction

The contributions of this chapter are threefold. First, we extend Hogarth and Einhorn (1990) seminal descriptive probability weighting function to a coherent harmonic probability weighting function (HPWF) via maximum entropy methods. Second, we show how that probability model theoretically resolves the preference reversal puzzle by identifying an observer effect in the preference reversal phenomenon. That is, observers report preference reversal even when the true state is procedure invariance *and* transitive preferences because their (observers) very act of measurement disturbs subjects probability cycles before they are complete. This is a manifestation of the "uncertainty principle" or "observer effect" explained at length in Von Neumann (1955, Ch. VI). Third, we show how the HPWF decomposes regret theory and rank dependent utility theory into core expected utility theory (EUT) plus functionally equivalent stochastic error addends.

Preference reversal (PR) is a phenomenon which arises when subjects in an experiment are asked to choose between pairs of bets with similar expected values, one bet has a high probability of winning a relatively small sum of money (the *P*-bet), and the other has a low probability of winning a large amount of money (the $-bet). Subjects typically choose the *P*-bet but when they are shown each bet in insolation and asked to state a reservation price if they were to sell each bet, they typically choose to sell the $-bet for more.[1]

[1]The PR phenomenon is not exclusive to binary choice over bets. Irwin et al. (1993) find PR in environmental choices, and Amiel, Cowell, Davidovitz, and Polovin (Amiel et al.) find it in policy decisions involving income inequality. Hinvest et al. (2014) collects references to non economics contexts where the PR phenomenon is observed.

The following example is taken from Goldstein and Einhorn (1987, pp. 236-237). Suppose subjects choose between the following bets, one of which has a high probability of winning a small amount of money (*P*-bet) and the other, which has a low probability of winning a large amount of money ($-Bet):

P-bet: Win $4 with $p = 0.97$, Lose $l with $p = 0.03$

$-bet: Win $16 with $p = 0.31$, Lose $1.5 with $p = 0.69$

When subjects are asked to choose which gamble they prefer to play, approximately half choose the *P*-bet over the $-bet. When each gamble is presented singly and subjects are asked to state the lowest selling price for the gamble if they owned it, the $-bet receives a higher price than the *P*-bet. If selling prices reflect preferences, then preference reverses depending on whether subjects chooses or states selling prices.

The PR phenomenon is of concern because of its implication for policy and resource allocation. For example, imputed values for non-tradeable items may be elicited from surveys and policymakers may use the responses to inform resource allocation. If the preference ranking extrapolated from survey responses do not reflect the choices that respondents would actually make, then the survey elicited valuation is unreliable and policy is misinformed. According to Seidl (2002) the literature identifies four causes of preference reversal:

- elicitation mode of certainty equivalent (CE);
- intransitivity of preferences;
- overpricing of $-bet and/or underpricing of *P*-bet;
- nonlinear probabilities.

In this paper our focus in on (1) intransitivity of preferences, and (2) nonlinear probabilities. We do not consider elicitation mode such as probability equivalent, and pricing issues, and we know of no single model that addresses all four causes of PR simultaneously. Specifically, we prove that even if preferences are transitive, and procedure invariance is not violated, an observer may still observe preference reversal if she measures choice behaviour before subjects' probability cycles are completed. This theoretical result is in contrast to Tversky et al. (1990) who find that violation of procedure invariance is the leading cause of observed preference reversal.

PR was first published by Lichtenstein and Slovic (1971); Lindman (1971) and confirmed in an "artefactual field study" with casino players in real betting situations in Lichtenstein and Slovic (1973). It has proven to be quite robust (Slovic, 1995).[2] Interest in the phenomenon lies in its relation to intransitivity of preferences. The latter implicates the transitivity axiom–a cornerstone of decision theory and preference based economic theory[3] in particular (Tversky (1969, p. 31), Tversky et al. (1990), Loomes et al. (1991, p. 430) and Wilcox (2008, pp. 198-200)).

In the [in]transitivity context, PR implies that given a binary preference relation $\succeq$ (where $\succ$ means strictly preferred and $\sim$ means indifferent to), a set of outcomes x, y, z such that a decision maker (DM) expresses $x \succ y$ and $y \succ z$, if she is prepared to pay a fee ε to acquire x, and if her elicited preferences are $x \succ y \succ z \succ x$, then that DM is vulnerable to a money pump or Dutch book that extracts an amount ε at the end of each cycle (Fishburn, 1988, p. 44). Evidently, violation of the transitivity axiom involves some kind of cyclic behaviour.[4] For example, Tversky (1969, p. 31) surmised "*that the observed inconsistencies reflect inherent variability or momentary fluctuation in the evaluative process. This consideration suggests that preference should be defined in a probabilistic fashion*." (emphasis added). Loomes and Sugden (1982); Bell (1982); Loomes and Sugden (1983) introduced regret theory which accommodates *some* intransitive preferences (Loomes et al., 1991). Lichtenstein and Slovic (1971) proffered an *information processing* hypothesis as the cause of PR. Even though several other hypotheses have been offered to explain the PR phenomenon (see Lichtenstein and Slovic (2006) for a review), to the best of our knowledge, the "information processing" hypothesis has not been explored in "a probabilistic fashion" in the context of information theory.[5]

This paper fills a gap in the literature by presenting a new approach to Seidl (2002, p. 637)

[2]The opening paragraph of the oft cited paper by Grether and Plott (1979) states that that "paper reports the results of a series of experiments designed to discredit the psychologists' works as applied to economics.". Instead, their experiments reaffirmed preference reversal.

[3]For instance, if extant preference based economic theory is unable to resolve personal equilibrium pricing (Kőszegi and Rabin, 2006) within a DM, how much more likely is it to resolve the indifference pricing puzzle among DMs? Cf. Mataramvura (2015).

[4]Rubinstein and Segal (2012, p. 2485) claim the maximal probability that a random sampling procedure yields such cycle is $\frac{8}{27}$.

[5]In cognitive science, Jerome Busemeyer and his co-workers resolve the preference reversal puzzle in the context of quantum information processing which employs quantum probability tools borrowed from quantum mechanics (Busemeyer and Diederich, 2002; Busemeyer and Wang, 2007; Pothos and Busemeyer, 2009; Busemeyer et al., 2011).

observation that "it seems that the explanation of preference reversals by intransitive preferences is still an open issue." It introduces an information theory of preference reversal by an through a coherent harmonic probability weighting function (HPWF) derived from applying Jaynes (1968) principle of maximum entropy to a distribution of outcomes.[6] Conceptually, it synthesizes the information processing, probabilistic choice, and momentary fluctuation hypotheses of preference reversal and intransitive preferences. The HPWF extends Hogarth and Einhorn (1990) descriptive outcome dependent pwf as indicated in the sequel. We use simulation to confirm that a simple affine transformation of our maximum entropy induced HPWF produces a sinusoidal inverted S-shaped probability weighting functional consistent with likelihood insensitivity[7] reported in Tversky and Wakker (1995), and in recent source function theory of uncertainty by Abdellaoui et al. (2011).[8] However, to the best of our knowledge the inherent HPWF result predicted by maximum entropy is new.[9]

We assume two procedurally invariant experiments with the same DMs are temporally spaced. Preference functions are the same in each experiment. Thus the preference reversal burden is carried by probability distributions over outcomes. In that way, "preferences are defined in a probabilistic fashion". We prove that DMs appear to violate the transitivity axiom when the probability cycles[10] of their HPWF are incomplete. To the best of our knowledge, the probability cycles resolution of the PR puzzle is new. We prove that experimenters interfere with the *inherent* HPWF induced by a statistical ensemble of outcomes in a lottery, when they assign probabilities to outcomes *ex ante*. The interference causes them to observe PR in the lab when the true state is no PR. This is a manifestation of the "uncertainty principle" or "observer effect" explained at length

[6]In the neuroeconomics and psychophysics literature Takahashi (2006) introduced a stylized pwf based on the difference between perceived probability and Claude Shannon's entropy measure of uncertainty for probabilities. The pwf is maximized when entropy is minimum.

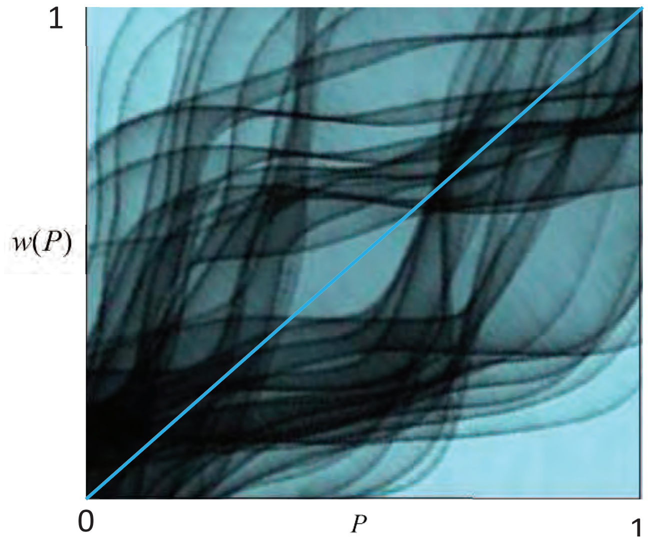
[7]Geometrically, this concept refers to flatness in the central portion of a pwf relative to its extremes.

[8]Harrison (2011) derived maximum likelihood estimates of parameters from a structural model with normally distributed measurement error which rejected the findings in this paper.

[9] al Nowaihi and Dhani (2010) introduced *composite cumulative prospect theory* (CCP), and *composite Prelec probability weighting functions* (CPF) to "address events close to the boundary of the probability interval $[0,1]$". Their Figure 5.1 depicts a piecewise harmonic probability weighting function based on heuristics of observed behaviour. Maccheroni et al. (2006, pg. 1448) used the concept of *relative entropy* to address issues related to variational utility representation of probabilistic preferences. A paper with tangentially related results like that reported here is Busemeyer et al. (2006) who used a quantum "superposition state" with respect to basis functions to analyze their quantum wave functions.

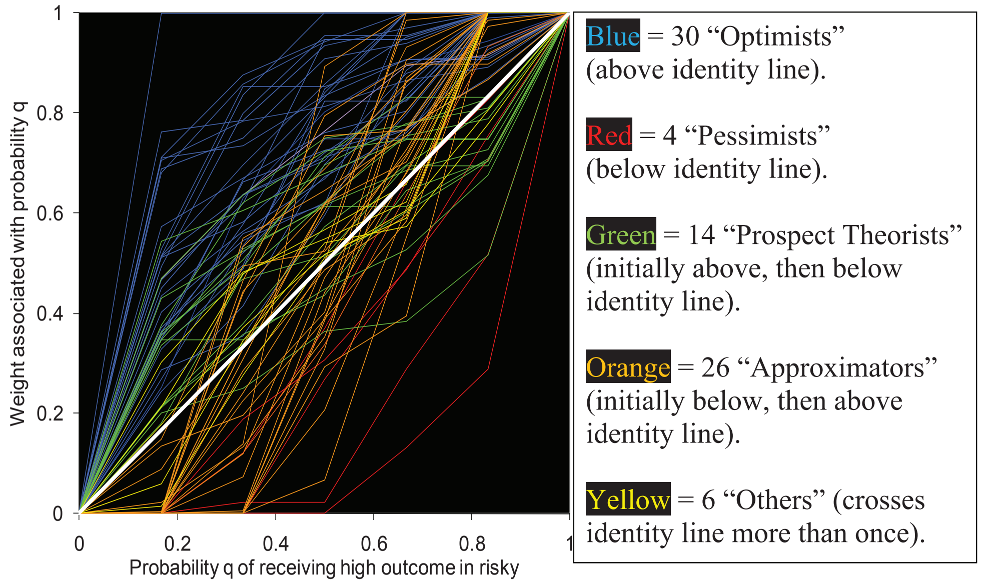
[10]A probability cycle is the greatest common divisor of the set of periods for a periodic phase function.

Figure 4.1: Harmonic strings in a unit square



String harmonics adapted to unit square to mimic a HPWF

Figure 4.2: Plots of PWFs observed in Wilcox (2011) experiments



The assorted shapes of pwfs observed in Wilcox (2011) resemble piecewise string harmonics.

in Von Neumann (1955, Ch. VI).[11]

Loomes and Pogrebna (2015) conducted experiments which seem to support aspects of our model. Their model involved Weak Stochastic Transitivity[12] (WST) by eliciting certainty equivalents (CEs) from a set of candidate CEs for given bets. Thus, subjects underlying preferences were modeled as a probability distribution over the set of CEs. They find that "when certainty equivalent values are inferred from repeated binary choices, the classic PR phenomenon largely disappears". In other words, subjects eventually complete their inherent probability cycles in a repeated choice context over a distribution of certainty equivalents.

The intuition behind our model is depicted in Figure 4.1 and Figure 4.2. The latter figure depicts plots of pwfs, in a unit square $[0,1]\times[0,1]$, observed in an experiment conducted by Wilcox (2011) to compare the predictive performance of three probabilistic choice models: two context dependent,[13] and one context free. One of his main results is that context dependent probabilistic choice models explains "switching behavior" among subjects. Figure 4.1 depicts string harmonics in a unit square $[0,1]\times[0,1]$. One can imagine a dense function space of pwfs in Figure 4.2 mimicking the string harmonics in Figure 4.1. So probability weighting functions admit harmonics. This paper provides a coherent model that supports that observation and its implication for preference reversal.

4.1.1 Positioning the paper in the literature on probabilistic preferences

The literature on preference reversal is huge (see e.g., Slovic and Lichtenstein (1983); Slovic (1995); Seidl (2002); Lichtenstein and Slovic (2006) for reviews, and Hinvest et al. (2014) for a succinct collection of the literature). Thus, as a practical matter we review select papers to put the instant paper in the perspective of probabilistic preferences.

Karni and Safra (1987, p. 679) introduced an EUT based model that predicts preference

[11] Stewart et al. (2015) conducted a set of experiments which confirm the experimenter effect on the instability of subjects' choices. They did not proffer an information theory.

[12] Let $\succ$ be a preference order relation and $\{a,b,c\}$. WST implies that if $\Pr(a\succ b)\geq 0.5$ and $\Pr(b\succ c)\geq 0.5$, then $\Pr(a\succ c)\geq 0.5$.

[13] "Context dependence" implies that "the relative attractiveness of x compared to y often depends on the presence or absence of a third option z" For example, an individual who prefers x over y in a binary choice cannot select y from the set $\{x,y,z\}$. But even if each individual satisfies [value maximization] we could obtain $P(x;y)>P(y;x)$ and $P(x;y,z)<P(y;x,z)$ if those who prefer x over y also prefer z over x." where $P(x,S)$ is the proportion of subjects choosing x from a set S (Tversky and Simonson, 1993). This contextual influence on preferences implies that DMs construct preferences during an experiment (Irwin et al., 1993, p. 6).

reversal. They used a piecewise linear probability weighting function (PWF), and piecewise linear utility function to estimate the maximum rank dependent expected utility (RDEU) due to Quiggin (1982) for several lotteries. They showed that the maximum RDEU for say lottery B exceeded that of lottery A even though the certainty equivalent (CE) or selling price of A exceeded that of B in the Becker et al. (1964) (BDM) auction mechanism for selling P-bets (high p low x) and \$-bets (low p and high x).[14] Tversky et al. (1990, pp. 207, 210) criticized the Karni and Safra (1987) model on grounds that experiments that employ elicitation schemes which do not depend on the independence axiom also find preference reversal. They argue that violation of procedure invariance[15] is the source of PR. Loomes et al. (1991, p. 430) show that regret theory can explain *some* PR patterns but not all. They made the grim conclusion that the results of their experiments "raise serious doubts about the descriptive validity of the transitivity axiom," *ibid* at 438.

Regenwetter et al. (2011) dispute reported findings of preference reversal in laboratory experiments. They conducted a meta study of the PR literature and re-examined several data sets. They found that PR reported in the papers they examined disappeared up to a Type I error rate of 5% when a mixture model was used to analyze the data. In fact, they state (p. 43):

> The mixture model assumes that choices vary because the decision maker is in different mental states * * * at different points in time. That is, the decision maker has some probability distribution over mental states and chooses x over y if and only if her or his current mental state is one in which she or he prefers x to y.

The authors' statement is motivated by the following model. Let $\mathscr{C}$ be a set of choice alternatives, $(A, \succ)$ be a strict partially ordered set[16] on $\mathscr{C}$ which binary relation is a strict partial order $\succ$, and $\mathscr{T}$ be the collection of all binary preference relations on $\mathscr{C}$. Define P_{xy} to be the probability that x is strictly preferred to y, and $P_{\succ}$ is the probability that a person is in transition state of preference $\succ$ in $\mathscr{T}$. Additionally, $P_{\succ}$ need not be a homogenous probability measure for it admits mixture

[14]The overlapping endpoints in Karni-Safra piecewise PWF and utility function induce bias in estimates of their criterion functions. Furthermore, Harrison (1986, p. 8) shows that DMs could strategically game the sequential pricing feature of the BDM in early rounds and bias experimental results.

[15]Roughly, "equivalent procedures for assessing preference should yield the same choices" Hinvest et al. (2014).

[16]section 4.B provides a brief review of the concepts associated with partially ordered sets.

measures. Thus

$$P_{xy} = \sum_{\substack{\succ \in \mathscr{T} \\ x \succ y}} P_{\succ} \tag{4.1.1}$$

Regenwetter et al. (2011, p. 47) points out "that $\mathscr{T}$ includes, for example, transitive, but incomplete, partial orders". Furthermore, "[a] mixture model of transitivity states that an axiom-consistent person's response at any time point originates from a transitive preference but that responses at different times need not be generated by the same transitive preference state," Regenwetter et al. (2011, p. 47).[17] In other words, transitive preference orders include incomplete partial orders. They did not provide a parametric form for the probability distribution over mental states.[18]

Cubitt et al. (2004) modified the "standard" PR experiment based on a monetary valuation (MV) task by adding a probabilistic valuation (PV) task. They used an ordinal payoff scheme in their experimental design, and subjects were randomly assigned to MV and PV. In the latter, subjects choose the probability to be associated with a fixed monetary value such that it is the certainty equivalent of a gamble. They found that subjects in the PV task experiment exhibited "marked difference" in PR patterns compared to those in the MV task. Furthermore, *more reversals were found within the PV group compared to within the MV group*. None of the psychology or economic hypotheses they reviewed could explain the phenomenon.[19] Butler and Loomes (2007) proposed a model of "probability equivalent" (PE) that extended the Cubitt et al. (2004) study. Theirs is a "model of imprecision" that "provides a framework that can accomodate the "strong" [certainty equivalent] reversals observed in many studies" (p. 293).[20]

Blavatskky (2009) extended Butler and Loomes (2007) with a model that accounts for PE and CE. In his model certainty equivalents are random variables. For example, the certainty

[17] Loomes and Sugden (1995, p. 643) introduce a tangentially related model, which they call "random preference model" that contains "core theory" (i.e., an axiom-consistent response) plus a stochastic addend.

[18] Theorem 4.H.1 in section 4.H shows how our model applies to mental states by its extension to quantum information processing and quantum probability models championed by Pothos and Busemeyer (2009) and Busemeyer et al. (2011).

[19] A recent study by Viechnicki (2015) use eye-tracking experiments which confirm that preference reversals are accompanied by differential attention to gamble attributes (probability of winning versus amount to win).

[20] Strong preference reversal occurs "when an individual chooses the P-bet over the \$-bet in a direct binary choice even though the certainty equivalent of the \$-bet is strictly greater than the highest possible outcome of the P-bet" (Blavatskky, 2009, p. 238). See also, Fishburn (1988, §2.8 p. 46).

equivalent (CE) of a bet is distributed over the maximal range of its prize. The certainty equivalent of the \$-bet is positively skewed but the certainty equivalent of the P-bet is negatively skewed. Hence, a DM always values the \$-bet more frequently than the P-bet. The diametric skewness allows the model to explain overpricing of the \$-bet.

Hogarth and Einhorn (1990) developed a descriptive model of probability weighting that depends on the sign and size of payoffs. They assumed that DMs first anchor on a stated probability p_A, and then adjust by mentally simulating other values. In their model (p. 783) the pwf is:

$$w(p_A) = p_A + k, \quad \text{where} \quad k = k_g - k_s \tag{4.1.2}$$

k_g represents the weights given to possible decision weights (in this case probability weights) above p_A, and k_s represents weighted values below p_A, and k is the net adjustment factor. Hogarth and Einhorn surmised:

> For decisions involving good or positive payoffs, values greater than the anchor are given less weight than those below; moreover, the degree of differential weighting increases with the size of the payoffs. Conversely, bad or negative payoffs imply that greater weight is accorded to values above rather than below the anchor, and the extent of differential weighting increases with the absolute size of the negative payoffs.

They describe the weight functions as follows:

$$k_g = f(\sigma, \theta, p_A, v(x)), \qquad k_s = g(\sigma, \theta, p_A, v(x)) \tag{4.1.3}$$

$$\frac{\partial k_g}{\partial \sigma} > 0, \quad \frac{\partial k_s}{\partial \sigma} > 0, \qquad \frac{\partial k_g}{\partial \theta} > 0, \quad \frac{\partial k_s}{\partial \theta} > 0 \tag{4.1.4}$$

$$\frac{\partial k_g}{\partial p_A} < 0, \quad \frac{\partial k_s}{\partial p_A} > 0, \qquad \frac{\partial k_g}{\partial |v(x)|} > 0, \quad \frac{\partial k_s}{\partial |v(x)|} > 0 \tag{4.1.5}$$

where σ is a measure of uncertainty, i.e., standard deviation of outcomes; θ is a measure of ambiguity; p_A is anchor probability; and $v(x)$ is a value function preferred by Kahneman and Tversky (1979). Hogarth and Einhorn state that "[t]he absolute size of $v(x)$ increases both k_g and k_s and together with its sign, determines the extent to which more weight is given in imagination to values above or below the anchor". Thus, in (4.1.2) $w(p_A)$ is a pwf that depends on the distribution of outcomes x, its first and second moments through σ, and for which k *fluctuates* around p_A.

Our HPWF is a coherent version of the Hogarth and Einhorn (1990) model because it was derived from maximum entropy analysis of the distribution of outcomes x. For instance, we derive the following outcome dependent HPWF[21] by employing the maximum entropy principle,[22] popularized by Jaynes (1968), to a distribution of outcomes construed as a realization of a random field or statistical ensemble.

$$w(\boldsymbol{x},\boldsymbol{p}) = \eta_0 p + \eta_1 \tan(\psi(\boldsymbol{z})) \tag{4.1.6}$$

where η_0 and η_1 are elevation and curvature parameters (Abdellaoui et al., 2010), $\boldsymbol{x} = [x_1,\ldots,x_n]$ is a vector valued rank ordered statistical ensemble or random field of outcomes, $\boldsymbol{p}$ is the corresponding set or probability distribution, $p \in \boldsymbol{p}$, and $\boldsymbol{z} = \{z_1,\ldots,z_n\}$ is a corresponding set of Z-scores for $\boldsymbol{x}$, $x_j = \mu_x + z_j\sigma_x$, and $\psi(\boldsymbol{z})$ is a phase function. Imposing the identifying restriction $0 \leq w(\boldsymbol{x},\boldsymbol{p}) \leq 1$ on (4.1.6) guarantees that $\tan(\psi(\boldsymbol{z}))$ does not blow up even though the function is not well behaved at the endpoints, i.e., it fluctuates–an empirical regularity of pwfs.[23] The fixed point probability in our model is $p^\star = e^{-1}$–the same fixed point derived in Prelec (1998) axiomatized model.[24] So our model is admissible.

In (4.1.6) p corresponds to the anchor probability p_A in Hogarth and Einhorn (1990) model in (4.1.2). The fluctuations in k in the latter model is captured by the phase function relation $\psi(\boldsymbol{z}) \sim (2k-1)\pi + \psi(\boldsymbol{z})$ in our model.[25] Moreover, the distribution of outcomes and their variance, represented by $v(x)$ and σ in (4.1.2), are incorporated in $\boldsymbol{z}$ and $\psi(\boldsymbol{z})$ in (4.1.6). Our model is distinguished from Hogarth and Einhorn (1990) by its identification of probability cycles in so caled "mental simulation". A probability cycle is the greatest common divisor $d(k)$ of the set $\mathscr{K}$

[21] Dillenberger et al. (2013) also introduced outcome dependent probability weighting functions but theirs are not harmonic.

[22] Good (1963, pg. 911) asserts:

> PRINCIPLE OF MAXIMUM ENTROPY. Let X be a random variable whose distribution is subject to some set of restraints. Then entertain the null hypothesis that the distribution is the one of maximum entropy, subject to these restraints.

[23] Kahneman and Tversky (1979, pp. 282-283) characterized this as follows: "Because people are limited in their ability to comprehend and evaluate extreme probabilities, highly unlikely events are either ignored or overweighted, and the difference between high probability and certainty is either neglected or exaggerated. Consequently, [the pwf] is not well-behaved near the end-points." (emphasis added).

[24] Lattimore et al. (1992) also introduced a flexible 2-parameter pwf. However, the fixed point probability in their model seems to be $p = 0.5$ (pp. 381, 382) and it does not fluctuate near the end point (p. 382).

[25] Here k is a number different from the adjustment factor in (4.1.2).

of values k on which the phase function relation holds. The intuition behind our model is that if preferences are observed before a cycle is complete, i.e., at times other than $d(k)$, then probabilistic preferences will be intransitive. For example, MacCrimmon (1968) found that a sample of business executives exhibited preference reversal in experiments he designed to test several postulates of Savage (1972) subjective expected utility (SEU) theory. When confronted with their behaviour in a *post-experiment* interview, the executives attributed their behaviour to carelessness or not reading the instructions carefully, and most indicated that they would *reverse* their behaviour. In other words, their probability cycles were completed during the *post-experiment* interview.[26]

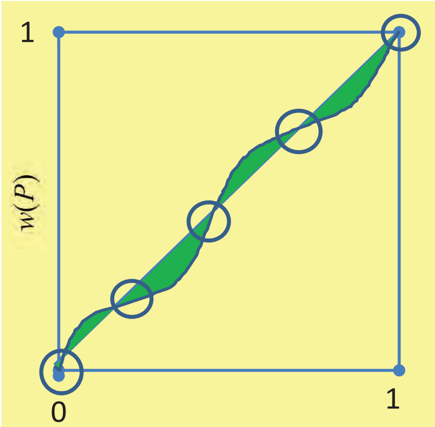
We use the RDEU transformation scheme from Quiggin (1982) to compute decision weights

$$\pi_j^{(k)}(\boldsymbol{x},\boldsymbol{p}) = \eta_0 p_j + \eta_1 \varphi^{(k)}(\mathfrak{z}_j), \qquad \sum_j \pi_j^{(k)}(\boldsymbol{x},\boldsymbol{p}) = 1 \tag{4.1.7}$$

where p_j is the probability associated with the j-th ranked outcome, and $\mathfrak{z}_j$ is a vector of Z-scores for the first j ranked outcomes, and $\varphi^{(k)}(\mathfrak{z}_j)$ is a harmonic *outcome weighting function* controlled by jumps $\sin(\Delta\psi(\boldsymbol{z}))$ in the phase function. The identifying restriction in (4.1.7) implies that $\sum_j \varphi^{(k)}(\mathfrak{z}_j) = (1-\eta_0)/\eta_1$. In that setup, $RDEU = \sum_j \pi_j^{(k)}(\boldsymbol{x},\boldsymbol{p})U(x_j) = EU \oplus WU$. When DMs focus only on linear probabilities, i.e., probability cycles are complete, they are EU maximizers. When they focus only on outcome weighting they are weighted utility (WU) maximizers (Chew and Waller (1986, p. 59)). Jointly, they are RDU maximizers. Figure 4.3 provides a sketch of a decision weight in (4.1.7). It depicts fluctuations of $\varphi^{(k)}(\mathfrak{z}_j)$ around an "observed" probability p_j. The former is based on probability cycles that vanish when a cycle is complete at which point $\pi_j^{(k)}(\boldsymbol{x},\boldsymbol{p}) = p_j$ (when $\eta_0 = 1$) at the fixed points where DMs have Von Neumann and Morgenstern (1953) type preferences. See also, al Nowaihi and Dhami (2006, pp. 5-7) who propose a higher order Prelec pwf that also fluctuates around the diagonal.

[26] Resulaj et al. (2009) study on bounded accumulation time mimics this process. Over time as a subject accumulates information in a stochastic choice experiment. She may change her mind from the initial decision when the accumulation of noisy evidence reaches a certain threshold. See Webb (2015) for a review of the literature.

Figure 4.3: Harmonic decision weights π



Sketch of Quiggin-type decision weights $\pi_j^{(k)}(\boldsymbol{x},\boldsymbol{p}) = \eta_0 p_j + \eta_1 \varphi^{(k)}(\mathfrak{z}_j)$ assuming $\eta_0 = 1$ for a HPWF. The fixed point probabilities (enclosed by circles) coincide with EUT. Preference reversal is supported by incomplete probability cycles (dark regions) of decision weights.

According to the Becker et al. (1964) (BDM) auction mechanism, under EUT the expected value of a bet should be the same in the choice phase (C), and bid phase (B), i.e., $EUT_C = EUT_B$. We prove that experimenters induce the decision weight relationship

$$\pi_j^{o1^{(k)}} = \pi_j^{o2^{(k)}} + \eta_1 \left(\varphi_1^{(k)}(\mathfrak{z}_j^o) - \varphi_2^{(k)}(\mathfrak{z}_j^o) \right) \tag{4.1.8}$$

where the superscripts 1 and 2 correspond to C and B respectively, and o is for "observation". The decision weight relationship induce the measurement

$$EUT_C = EUT_B + \Delta_\varphi^{(k)} WU \tag{4.1.9}$$

where $\Delta_\varphi^{(k)} WU$ is a quantity that vanishes when probability cycles are complete. $\Delta_\varphi^{(k)} WU$ constitutes the *inherent variability or momentary fluctuation in the evaluative process*" anticipated by Tversky (1969, p. 31).[27] It also mimics Regenwetter et al. (2011, p. 47) "incomplete partial order" hypothesis because the DM's response originates from a transitive preference state before the experimenter broke her probability cycle and observed a different state.

[27] Tversky (1969, p. 43) introduced a theorem based on an "additive difference model" which shows that the transitivity axiom is satisfied for a multidimensional utility function if the difference in utility for each component part is linear. (4.1.9) implies some of those features.

Our analysis of Karni and Safra (1990, p. 491) (KS1990) show that their preference reversal model contains a harmonic function $\tan(\alpha'') - \tan(\alpha')$, where $\tan(\alpha'') = (1 - w(p'')/(1 - p'')$ and $\tan(\alpha') = (1 - w(p')/(1 - p')$ and $w(\cdot)$ is a "standard" PWF. Since $\tan(\alpha'') - \tan(\alpha') = \sin(\alpha'' - \alpha')\sec(\alpha'')\sec(\alpha')$, preference reversal is controlled by the *angular jump* $(\alpha'' - \alpha')$ in $\sin(\alpha'' - \alpha')$, which is similar to $\sin(\Delta\psi(\mathbf{z}))$ in our model. The KS1990 model is distinguished from ours in several ways.

First, the underlying PWF in their model is not harmonic. The harmonic feature of their result is based on our interpretation of their analysis. Second, they restrict the fixed point θ in their model to the interval $\frac{1}{2} < \theta < 1$ and assume that "a typical probability transformation function" is inverted S-shape. Refer to (Quiggin, 1993, pp. 60-61) for pwf shapes. The interval for θ implies that DMs in their model may be inordinately pessimistic. If so, then the selling price in the BDM auction mechanism used in their model maybe inflated due to probabilistic loss aversion (Schmidt and Zank, 2008). Third, the KS1990 model resolves preference reversal in the BDM auction framework. In contrast, our model resolves it in the context of choice over at least three outcomes when the same experiment is repeated with the same DMs. So our result is robust to experimental procedures. Moreover, we make no assumptions about the shape of the PWF or the location of the fixed point. The fixed point in our model depends on curvature and elevation parameters of the underlying PWF. Finally, our HPWF is a bivariate function of outcomes and their *inherent* probabilities.

There is a burgeoning literature on quantum cognition in which methods of quantum physics are applied to address, *inter alia*, preference reversal.[28] For example, Trueblood and Busemeyer (2011) use a quantum probability paradigm to show how order effects introduce a dimension of uncertainty because they influence the computed probability of an event. See also, Busemeyer

[28]The problem addressed in this paper was described thus:

> Interference is the effect that is typical of all those phenomena which are described by wave equations. Following the Bohrs idea [] of describing mental processes in terms of quantum mechanics, one is immediately confronted with the interference effect, since the physical states in quantum mechanics are characterized by wave functions. The possible occurrence of interference in the problems of decision making has been discussed before on different grounds []. However, *no general theory has been suggested, which would explain why and when such a kind of effect would appear, how to predict it, and how to give a quantitative analysis of it that can be compared with empirical observations*. In our approach, interference in decision making arises only when one takes a decision involving composite intentions. The corresponding mathematical treatment of these interferences within QDT is presented . . .

[Citations omitted, emphasis added]. Yukalov and Sornette (2009, pp. 1088-1089). I thank Jerome Busemeyer for directing me to this reference. See also, Khrennikov and Haven (2009).

et al. (2011, pg. 198). This is similar to our "experimenter interference" result described above. In fact, Theorem 4.H.1 in section 4.H characterizes DMs preference states in a HPWF setting. However, the results produced in this paper are distinguished because they were derived from microfoundations of prospect theory using a bottom-up approach.[29] As opposed to extant literature which use a top down approach by fitting quantum mechanics models to behavioural phenomena.

The rest of the paper proceeds as follows. In section 4.2 we set up the maximum entropy problem of probabilistic preferences over probability distributions. There, we derive an abstract harmonic probability weighting function (HPWFs) under rubric of representation Theorem 4.2.2. That is followed by section 4.3 where we identify the HPWF by identifying restrictions and use it in section 4.4 to resolve preference reversal phenomena. In section 4.5 we provide a simple calibration and simulation exercise to provide a rough estimate of our HPWF model. In section 4.6 we conclude with perspectives for further research.

4.2 Maximum entropy and the inherent distribution of outcomes

This section provides preliminary motivation in subsection 4.2.1 for our use of the entropy device to characterize the degree of uncertainty associated with a gamble or bet. In subsection 4.2.2 we employ the canonical principle of maximum entropy to derive an abstract HPWF. In subsequent sections identifying restrictions allow us to provide applications.

4.2.1 Preliminaries

In this paper, entropy is comprised of the expected value (average) of the information contained in a sample drawn from a distribution or data stream. Entropy thus characterizes our uncertainty[30] about our source of information, and increases for more sources of greater randomness. The source is also characterized by the probability distribution of the samples drawn from it. The idea here is that the less likely an event is, the more information it provides when it occurs.

[29] Stutzer (1996) also used a bottoms up maximum entropy approach to asset pricing theory. However, he did not posit asset prices as random fields nor did he identify a behavioural quantum wave as we do here.

[30] We use the term "uncertainty" in its most general sense. The economics literature distinguishes between risk and uncertainty. The former pertains to the case when the underlying probability distribution is known. The latter, when it is unknown (Knight, 1921).

Consider the following example adapted from Khincin (1957, p. 3). Let L_1 and L_2 be two bets comprised of common events A_1 and A_2 but with separate probability distribution $0.5, 0.5$ and $0.999, 0.001$, such that $L_1 = (A_1, 0.5; A_2, 0.5)$ and $L_2 = (A_1, 0.999, 0.001)$. In L_2 the outcome of the bet is almost surely event A_1 while the outcome in L_1 is less certain. A third bet $L_3 = (A_1, 0.3; A_2, 0.7)$ contains an intermediate level of uncertainty about outcomes. In information theory, "entropy" is a measure of the uncertainty in each bet. It is an "inverse probability" measure. That is, the lower the probability, the more uncertain the event.

The principle of maximum entropy popularized by Jaynes (1968) provides a mechanism for deriving a coherent prior probability distribution for a distribution of outcomes when the corresponding probability distribution is unknown. The mechanism is characterized by identifying restrictions based on what information, if any, is known about the unknown probability distribution. This typically involves assumptions about moment conditions. The position taken in this paper is that a decision maker has a subjective probability distribution that she associates with a given distribution of outcomes. And that she uses a computation mechanism, in this case maximum entropy, to arrive at that distribution. This hypothesis finds some support in the neuroeconomics literature where Hinvest et al. (2014) surmised that "[t]he combination of behavioral and neuroimaging data may suggest that participants used a mathematical approach to formation of valuations but choices were subject to emotional influence. These findings may provide support for a biological procedural invariance view of gambling preference". Moreover, since outcomes a measured with respect to some scale, the inherent probability distribution can be derived from maximum entropy considerations (Frank and Smith, 2010).

Motivated by the above, we provide a formal definition for entropy.

Definition 4.2.1 (Entropy)**.** Cover and Thomas (1991, pg. 13)
The entropy $H(X)$ of a discrete random variable X is defined by

$$H(X) = -\sum_{x \in X} p(x) \ln(p(x))$$

Remark 4.2.1. $H(X) = 0$ implies either certainty, i.e. $p(x) = 1$ for all x, or null event, i.e. $p(x) = 0$ for all x. So large H implies uncertainty, and small H implies more certainty. □

Good (1963, pg. 911) also gives an implicit definition where he states "Let X be a random variable whose (physical) probability distribution is not completely given. Of all the available distributions there will usually be one of maximum entropy, i.e. maximum uncertainty".

4.2.2 Maximum entropy with partial information about distribution of outcomes

Let X be a space of outcomes and f be an unknown continuous probability distribution over X. Since f is unknown, it admits complex values. The canonical constrained maximum entropy problem over probability densities is stated as follows:[31]

$$\begin{aligned} &H(f) = -f\ln(f) \\ &\max_{f} H(f) \\ &\text{s.t } f(x) \geq 0, \qquad \int_K f(x)dx = 1 \\ &\int_K r_i(x) f(x)dx = \alpha_i, \;\; 1 \leq i \leq m \end{aligned} \tag{4.2.1}$$

where K is a compact subset of X such that $f(x)$ vanishes on X/K. (4.2.1) represent m *moment conditions* imposed on the optimality problem. In Jaynes (1968, pg. 234) moment conditions are used to establish a *correspondence principle* for maximum entropy distributions and "frequency distributions". To account for the probability weighting phenomenon in Tversky and Kahneman (1992) cumulative prospect theory, following Bochner (1955, pg. 12), and by abuse of notation, for random variable X, we define a set function $F(A)$ as follows

$$F(A) = P(X \leq x|\, x \in A) = \int_{\substack{\{x\in A\} \\ \{X \leq x\}}} f(x)dx \tag{4.2.2}$$

$$(w \circ P)(X \leq x|\, x \in A) = (w \circ F)(A) \tag{4.2.3}$$

[31] See Cover and Thomas (1991, Ch. 11, pp. 266-267) and Avellaneda (1998) for further details on constrained maximum entropy. Stutzer (1996) credited Josiah Willard Gibbs and Edwin T. Jaynes with incipient formulation and subsequent popularization, respectively, of the optimality approach. See Abbas (2006) for applications to utility theory.

Furthermore, on the whole space of all gambles $\mathscr{X}$ we write symbolically

$$F = F(X) = P(X \leq x|\, x \in \mathscr{X}) = \int_{\substack{\{X \leq x\} \\ \{x \in \mathscr{X}\}}} f(x)dx \tag{4.2.4}$$

Assume that the composite relation holds under the integral so that we assign the functional restricted to A the value

$$(w \circ F)(A) = (w \circ P)(X \leq x|\, x \in A) = \int_A (w \circ f)(x)dx \tag{4.2.5}$$

Thus we write the entropy function over some compact set K in $\mathscr{X}$ as

$$\tilde{H}(F) = -w(F)\ln(w(F)) \tag{4.2.6}$$

In the sequel we write $w(F)$ for $w(F(x))$ and so on, unless indicated otherwise. The optimality problem is one of maximum entropy with respect to probability weighting. Let J be the Lagrangian or criterion function and λ_i be the Lagrange parameters coresponding to m moments restrictions on f. So that for K compact in X

$$J(w(F)) = \tilde{H}(F) + \lambda_0 \int_K f(x)dx + \sum_{i=1}^{m} \lambda_i \int_K r_i(x) f(x)dx \tag{4.2.7}$$

Let μ be a measure of information on some measureable set $A \in \mathscr{B}(X)$. So that for integration over the set A we have

Lemma 4.2.1 (Functional Gâteaux derivative)**.**

$$F(A) = \int_A f\mu(dx) \Rightarrow \frac{\partial F}{\partial f} = \mu(A) \Rightarrow \frac{\partial w(F)}{\partial f} = w'(F)\mu(A) \tag{4.2.8}$$

Proof. See Appendix 4.C. □

Remark 4.2.2. Kullback (1968, pg. 5) uses measures like $F(A) = \frac{1}{\mu(A)} \int_A f(x)\mu(dx)$ to represent "averages" over the given set A in his information theory models.

By abuse of notation, to find the inherent probability weighting function on X we solve the problem

$w(F) = \arg\max_f J(w(F))$. The results are formalized in the following:

Theorem 4.2.2 (Maximum Entropy Harmonic Probability Weighting Functional Representation)**.**

Let X be an outcome space, $\mathscr{B}(X)$ be the σ-field of Borel measureable subsets of X, and μ be σ-finite measure. Let F be the distribution function for the underlying probability density in the maximum entropy problem, $w(F)$ be the corresponding maximum entropy probability weighting functional, E be an expectation operator, and $\{E[r_i(x)]\}_{i=1}^m$ be a moment generating sequence defined on X. Then for arbitrary $A \in \mathscr{B}(X)$, the maximum entropy probability weighting functional $w(F) \in L^2(X, \mathscr{B}(X), \mu)$ has representation

$$w(F) = e^{-1} \exp\Big(-\frac{\tilde{w}_m(F)}{w'(F)\mu(A)}\Big)$$

with first order approximation

$$w(F) = e^{-1} - e^{-1}\{w'(F)\}^{-1} \sum_{j=1}^{m} \tilde{a}_j e^{i\omega_j x}$$

where e^{-1} is the fixed point probability, $\tilde{w}_m(F) = \sum_{j=1}^m a_j e^{i\omega_j x}$, and $\tilde{a}_j$ is a scaled coordinate with respect to the measure $\mu(A)$ and the orthogonal basis $\{e^{i\omega_j x}\}_{j=1}^\infty$ in $L^2(X, \mathscr{B}(X), \mu)$.

Proof. See Appendix 4.D. □

Remark 4.2.3. The functional form for $w(F)$ is well defined. Notice that it includes the fixed point probability e^{-1} axiomatized by Prelec (1998); the harmonic or sinusoidal component $\tilde{w}_m(F)$ which approximates the moment conditions imposed on the underlying nonparametric probability density function; the *given* slope or rate at which $w(F)$ changes with respect to underlying probabilities; and the information scale $\mu(A)$ for "spreading information mass" over the set A. This suggests that the analysis extends to Lebesgue-Stieltjes measures.

Corollary 4.2.3 (Phase functional representation)**.**

There exist phase functionals $\vartheta(F)$ such that

$$\tilde{w}_m(F) = \|\tilde{w}_m(F)\|_{L^2_\mu} \exp(-i\vartheta(F))$$

and

$$w(F) = e^{-1} \exp\left(-\frac{\|\tilde{w}_m(F)\|_{L^2_\mu} \exp(-i\vartheta(F))}{w'(F)\mu(A)} \right)$$

Proof. Substitute the De Moivre representation of $\tilde{w}_m(F)$ in the exponent for $w(F)$ in Theorem 4.2.2. □

4.3 Identifying restrictions for harmonic probability weighting function

In this section we show how the abstract theory of section 4.2 applies to the Z-transformation of the joint distribution of outcomes $x_1, \ldots, x_n$ in a lottery. Whereupon we derived the parametric representation of the HPWF employed later in the paper.

4.3.1 HPWF and Z-transformation of outcomes

Consider the following application. Without making any *a priori* assumptions about functional form, let X be a random variable with mean μ and variance σ^2 and unknown distribution function F. Among several other examples, Cover and Thomas (1991, pg. 268) shows that the normal distribution has maximum entropy for probability density functions with finite first and second moment. However, under auspice of the Central Limit Theorem, we can always construct a standard normal random variable $\mathfrak{z} = (x - \mu_x)/\sigma_x$ for any distribution with finite first and second moments, and extend the problem to the distribution function F. Thus, our constrained optimization problem in (4.2.1) turns to characterizing maximum entropy probability weighting functions $w(F)$ when the underlying outcome distribution has finite mean and variance. [32] In that case, *conditioned* on $w'(F)$ and the inherent probability distribution of x, we can write (4.D.3) as

$$w(F) = e^{-1} \exp\left(-\frac{(x-\mu)^2}{2w'(F)\sigma^2} \right) \tag{4.3.1}$$

[32] Stutzer (1996) used maximum entropy methods to show that $\lambda_0, \lambda_1, \ldots$ are the market price of risk in a mean-variance efficient portfolio.

Application of the inequality $0 \leq w(F) \leq 1$ to (4.3.1) gives us

$$w'(F) \geq -\frac{(x-\mu)^2}{2\sigma^2} = -\frac{\mathfrak{z}^2}{2} \tag{4.3.2}$$

where $\mathfrak{z} = (x-\mu)/\sigma$ is the standard normal random variable and $F = \Phi(\mathfrak{z})$ is the cumulative normal distribution.[33] So $\Phi(\mathfrak{z})$ mimics rank ordered probability quantiles. For some constant k added to equalize the right hand side of (4.3.2), after substitution and integration of both sides of the equality we get

$$d[w(\Phi(\mathfrak{z}))] = w'(\Phi(\mathfrak{z}))\Phi'(\mathfrak{z})d\mathfrak{z} \Rightarrow \int d[w(\Phi(\mathfrak{z}))] = \int w'(\Phi(\mathfrak{z}))\Phi'(\mathfrak{z})d\mathfrak{z} \tag{4.3.3}$$

$$\Rightarrow \int_{-\infty}^{\mathfrak{z}} d[w(\Phi(y))] = \int_{-\infty}^{\mathfrak{z}} (k - \frac{y^2}{2})\Phi'(y)dy \tag{4.3.4}$$

$$\Rightarrow w(\Phi(\mathfrak{z})) - w(0) = k\Phi(\mathfrak{z}) - \frac{1}{2}\int_{-\infty}^{\mathfrak{z}} y^2\Phi'(y)dy \tag{4.3.5}$$

Even though the domain of definition of the probability weighting function $w(\Phi(\mathfrak{z}))$ is $[0,1]$ in (4.3.5), the righthand side equation is a *functional* in $\mathfrak{z}$. So we have to transform the $\mathfrak{z}$-axis isomorphically into the unit interval $[0,1]$ in order for (4.3.5) to be a harmonic pwf.

To make that transformation, we impose the following restrictions. By virtue of Lindeberg's condition[34] applied to the tails of a normal distribution, we use an arbitrary cutoff $\mathfrak{z}$-score of -4 to approximate $-\infty$; and +4 to approximate $+\infty$.

Assumption 4.3.1 (Lindberg normal approximation)**.** *We assume the normal distribution is a mol-*

[33] As a practical matter the Z-statistic can be replaced by the T-statistic $T_j = \frac{x_j - \bar{x}_n}{\frac{s}{\sqrt{n-2}}} \sim t_\alpha^{(n-2)}$, $n \geq 3$ in small samples of size n where $\bar{x}_n$ is sample mean, s is sample standard deviation, and $(n-2)$ is the degrees of freedom (df), and $t_\alpha^{(n-2)}$ is the t-distribution with $(n-2)$ df for a one-sided α level test. Refer to Frank and Smith (2010) for a rationalization of this result.

[34] This is a uniform integrability condition that characterizes the tail of a distribution. According to Gikhman and Skorokhod (1969, pg. 452), let $F_{ni}(x)$ denote the distribution function of a random variable ξ_{ni}. Then, if for every $\varepsilon > 0$

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \int_{|u|>\varepsilon} u^2 dF_{ni}(u) = 0$$

the random variable ξ_{ni} is said to satisfy Lindeberg's condition.

lifier or bump function on the compact set $[-4,4] \subset [-\infty,+\infty]$ *such that*

$$\Phi(4) = 1, \quad \Phi(-4) = 0, \quad w(1) = 1, \quad \textit{and} \quad w(0) = 0 \tag{4.3.6}$$

$$I(\mathfrak{z}) = \int_{-4}^{\mathfrak{z}} y^2 \Phi'(y) dy, \quad |\mathfrak{z}| \leq 4 \tag{4.3.7}$$

$$\Phi(\mathfrak{z}) = 0, \quad |\mathfrak{z}| > 4 \tag{4.3.8}$$

□

Since the variance of a standard normal distribution is 1, by virtue of Assumption 4.3.1

$$\int_{-4}^{+4} y^2 \Phi'(y) dy \approx 1 \tag{4.3.9}$$

After substituting $\mathfrak{z} = 4$ in (4.3.5) we get

$$k = \frac{3}{2}, \qquad w(\Phi(\mathfrak{z})) = \frac{1}{2}(3\Phi(\mathfrak{z}) - I(\mathfrak{z})) \tag{4.3.10}$$

Assumption 4.3.1 "calibrates" the HPWF to a fixed range of values for $\mathfrak{z}$, and that led to (4.3.10). However, (4.3.10) should extend to values of $\mathfrak{z}$ that are not restricted by Assumption 4.3.1. We rectify that in the following

Assumption 4.3.2 (Generalized Lindberg normal approximation)**.** *We assume that the normal distribution is a mollifier or bump function on the interval* $|\mathfrak{z}| \leq \mathfrak{z}_h < \infty$ *so that*

$$\Phi(\mathfrak{z}_h) = 1, \quad \Phi(-\mathfrak{z}_h) = 0, \quad w(1) = 1, \quad \textit{and} \quad w(0) = 0 \tag{4.3.11}$$

$$I(\mathfrak{z}) = \int_{-\mathfrak{z}_h}^{\mathfrak{z}} y^2 \Phi'(y) dy, \quad \int_{-\mathfrak{z}_h}^{+\mathfrak{z}_h} y^2 \Phi'(y) dy \approx 1 \tag{4.3.12}$$

$$\Phi(\mathfrak{z}) = 0, \quad |\mathfrak{z}| > \mathfrak{z}_h \tag{4.3.13}$$

□

On the basis of Assumption 4.3.2 we generalize (4.3.10) which process is summarized in the following

Theorem 4.3.3 (HPWF–Harmonic Probability Weighting Function). *Let* $\boldsymbol{x} = [x_1, \ldots, n_n]$ *be a vector valued statistical ensemle of ranked outcomes, and* $\boldsymbol{z}$ *be the corresponding vector of Z-scores for the ranked outcomes such that* $x_j = \mu_x + z_j \sigma_x$ *where* μ_x *and sigma*$_x$ *are the mean and standard deviation of* $\boldsymbol{x}$*, respectively. The inherent HPWF of a DM is given by*

$$w(p,x) = \eta_0 p + \eta_1 \tan(\psi(\mathfrak{z})) \tag{4.3.14}$$

where $\psi(\mathfrak{z})$ *is a phase function and* $-\frac{\eta_0}{\eta_1} \leq \tan(\psi(\mathfrak{z})) \leq \frac{1-\eta_0 p}{\eta_1}$.

Proof. See Appendix 4.E. □

4.3.2 Reconciling HPWF fixed point and endpoints

The phase function $\psi(\mathfrak{z})$ is a realization of the phase function $\vartheta(F)$ in Corollary 4.2.3 on page 170. According to Theorem 4.3.3, for fixed point probability $w(p^\star) = p^\star$ we have

$$p^\star = e^{-1} \Rightarrow \tan(\psi(\mathfrak{z})) = \left(\frac{1-\eta_0}{\eta_1}\right) e^{-1} \tag{4.3.15}$$

Thus, the fixed point depends on elevation η_0 and curvature η_1. Restrictions imposed on the HPWF in Theorem 4.3.3 imply that it fluctuates near its end points. That is,

$$p = 0 \Rightarrow \eta_1 \tan(\psi(\mathfrak{z})) = 0 \Rightarrow \psi(\mathfrak{z}) = n\pi, \;\; n = 0, 1, \ldots \tag{4.3.16}$$

$$p = 1 \Rightarrow \tan(\psi(\mathfrak{z})) = \frac{1-\eta_0}{\eta_1} \Rightarrow \psi(\mathfrak{z}) = (2n-1)\pi + \psi(\mathfrak{z}) = \tan^{-1}\left(\frac{1-\eta_0}{\eta_1}\right) \tag{4.3.17}$$

So we provide a theoretical explanation for the observation that probability weighting is not well behaved near the end points. [35]

Additionally, $\tan(\psi(\mathfrak{z}))$ is an outcome or *prize weighting function* that overweighs outcomes over suitably defined "small probabilities"; and underweighs outcomes over "large proba-

[35] Because people are limited in their ability to comprehend and evaluate extreme probabilities, highly unlikely events are either ignored or overweighted, and the difference between high probability and certainty is either neglected or exaggerated. Consequently, $[w]$ *is not well behaved near the end points.*

Kahneman and Tversky (1979, pp. 282-283).

bilities". In other words, $\tan(\psi(\mathfrak{z}))$ is a "quantal effect'[36] that captures diminishing sensitivity (in the middle part of the curve) reported in Tversky and Wakker (1995, pg. 1257). Figure 4.5 depicts the impact of $\tan(\psi(\mathfrak{z}))$ in a calibration of our HPWF. Substitution of the relationships

$$\tan(\psi(\mathfrak{z})) = -i\tanh(i\psi(\mathfrak{z})) \tag{4.3.18}$$

$$w(p,x) = \eta_0 p - i\eta_1 \tanh(i\psi(\mathfrak{z})) \tag{4.3.19}$$

show that $w(p,x)$ has the form of a complex valued trigonometric polynomial consistent with the prediction of Theorem 4.2.2.

4.4 Rank dependent utility and preference reversal over probability cycles

In this section we employ Quiggin (1993, §5.2, p. 57) RDEU model and our HPWF to resolve the PR puzzle. We provide some preliminaries that include definition of the probability cycle concept. In subsection 4.4.1 we decompose the decision weights obtained from the HPWF by using Quiggin (1982) transformation procedure. In subsection 4.4.2 we introduce a model in which procedure invariance is imposed on an experiment and show how the HPWF resolves PR in that context. In subsection 4.4.3 we provide analytics which show that experimenters interference with a DM's probability cycle lead them to misperceive PR.

Preliminaries

Consider the probabilistic rank dependent expected utility (RDEU) specification[37] in Wakker (1994, pg. 10) for a simple rank ordered lottery $p_1, x_1; \ldots, x_n, p_n$. For utility $U(x_j)$ of outcome

[36] Kahneman and Tversky (1979, pg. 282)

[37] In order not to overload the paper with issues pertaining to framing effects in decision weights, we did not employ cumulative prospect theory (CPT). Compare Tversky and Wakker (1995, pg. 1259).

x_j we have

$$RDEU(p_1,x_1;\ldots,x_n,p_n) = \sum_{j=1}^{n} \pi_j U(x_j) \tag{4.4.1}$$

$$\pi_j = w(p_1+\ldots+p_j) - w(p_1+\ldots+p_{j-1}) \tag{4.4.2}$$

$$\pi_1 = w(p_1), \qquad \sum_j \pi_j = 1 \tag{4.4.3}$$

where $w : [0,1] \to [0,1]$ is a probability transform function, and π_j is a *decision weight*. In this case, we let w be our inherent HPWF. Since $\tan(\psi(\mathfrak{z}))$ is cyclic in Theorem 4.3.3, we have the periodic relationship

$$\begin{aligned} &\tan(\psi^{(k)}(\mathfrak{z}))) = \tan(\psi(\mathfrak{z})), \quad k = 1,2,\ldots \\ &\text{where } \ \psi^{(k)}(\mathfrak{z}) = (2k-1)\pi + \psi(\mathfrak{z}) \end{aligned} \tag{4.4.4}$$

Unless otherwise stated in the sequel π is 180^o in radians. π with subscripts (or superscripts) is a decision weight. This leads to the following

Definition 4.4.1 (Probability cycle)**.** Kemeny et al. (1976, p. 144). Let $w(p,x)$ be as in Theorem 4.3.3 and $\mathscr{K} = \{k|\ \psi^{(k)}(\mathfrak{z}) = (2k-1)\pi + \psi(\mathfrak{z})\}$ be a set of periodic phase functions. The probability cycle for $w(p,x)$ is the greatest common denominator $d(k)$ of $\mathscr{K}$. □

The identity $p_j = \Phi(z_j)$ implies that there exist $\tilde{\mathfrak{Z}}_j$ such that $\sum_{k=1}^{j} p_j = \Phi(\tilde{\mathfrak{Z}}_j)$. It follows from Theorem 4.3.3 that for some pair of values $(\tilde{\mathfrak{Z}}_{j-1}, \tilde{\mathfrak{Z}}_j)$ to be determined, we have the decision weight

$$\begin{aligned} \pi_j^{(k)} &= \pi(\tilde{\mathfrak{Z}}_{j-1}, \tilde{\mathfrak{Z}}_j;\, k) = \\ & w\Big(p_1+p_2+\ldots p_j\Big) - w\Big(p_1+p_2+\ldots+p_{j-1}\Big) = \\ & \eta_0\Big(p_1+p_2+\ldots p_j - p_1 - p_2 - \ldots - p_{j-1})\Big) + \\ & \eta_1\Big(\tan(\psi^{(k)}(\tilde{\mathfrak{Z}}_j)) - \tan(\psi^{(k)}(\tilde{\mathfrak{Z}}_{j-1}))\Big) \end{aligned} \tag{4.4.5}$$

$$= \eta_0 p_j + \underbrace{\eta_1\Big(\tan(\psi^{(k)}(\tilde{\mathfrak{Z}}_j)) - \tan(\psi^{(k)}(\tilde{\mathfrak{Z}}_{j-1}))\Big)}_{\text{harmonic component of decision weight}} \tag{4.4.6}$$

Thus, the decision weight $\pi_j^{(k)}$ is cyclic by virtue of (4.4.4) and the harmonic term in (4.4.6).[38] Even if it was not cyclic, it depends nonlinearly on $\tilde{\mathfrak{Z}}_{j-1}$ and $\tilde{\mathfrak{Z}}_j$ as indicated. Additionally, if k is not an integer in (4.4.4), i.e. the cycle is incomplete, equality does not hold so probabilistic choice is different for the same set of stimuli. The results above extend naturally to Tversky and Kahneman (1992) cumulative prospect theory (CPT) because the latter employs the same decision weight scheme as RDU. We note that in (4.4.6) the harmonic component of decision weights vanishes when $\tilde{\mathfrak{Z}}_j = \tilde{\mathfrak{Z}}_{j-1}$ so $\pi_j^{(k)} = \pi_j$ and RDU collapses to EUT due to $\eta_0 p_j$. We say more on that next.

4.4.1 Decomposition of decision weights obtained from HPWF

By virtue of (4.4.6), we decompose $\pi_j^{(k)}$ as follows. Let

$$\psi^{(k)}(\tilde{\mathfrak{Z}}_j) = \psi^{(k)}(\tilde{\mathfrak{Z}}_{j-1}) + \Delta\psi^{(k)}(\tilde{\mathfrak{Z}}_j), \;\; \tilde{\mathfrak{Z}}_j = \Phi^{-1}\left(\sum_{r=1}^{j} p_r\right) = \Phi^{-1}\left(\sum_{r=1}^{j} \Phi(\mathfrak{z}_r)\right) \tag{4.4.7}$$

$$\begin{aligned} &\tan(\psi^{(k)}(\tilde{\mathfrak{Z}}_j)) - \tan(\psi^{(k)}(\tilde{\mathfrak{Z}}_{j-1})) \\ &= \tan(\Delta\psi^{(k)}(\tilde{\mathfrak{Z}}_j))\left[1 + \tan(\psi^{(k)}(\tilde{\mathfrak{Z}}_j))\tan(\psi^{(k)}(\tilde{\mathfrak{Z}}_{j-1}))\right] \end{aligned} \tag{4.4.8}$$

$$= \sin(\Delta\psi^{(k)}(\tilde{\mathfrak{Z}}_j))\sec(\psi^{(k)}(\tilde{\mathfrak{Z}}_j))\sec(\psi^{(k)}(\tilde{\mathfrak{Z}}_{j-1})) \tag{4.4.9}$$

$$\pi_j^{(k)} = \eta_0 p_j + \eta_1 \sin(\Delta\psi^{(k)}(\tilde{\mathfrak{Z}}_j))\sec(\psi^{(k)}(\tilde{\mathfrak{Z}}_j))\sec(\psi^{(k)}(\tilde{\mathfrak{Z}}_{j-1})) \tag{4.4.10}$$

In (4.4.10) the decision weight is decomposed into a part $(\eta_0 p_j)$ that reflects DM confidence (due to elevation η_0) about the inherent probability p_j associated with outcome x_j, and a harmonic part $\eta_1 \varphi^{(k)}(\boldsymbol{\mathfrak{z}}_j)$ (defined next) controlled by the curvature parameter η_1 and jump in phase function $\Delta\psi^{(k)}(\tilde{\mathfrak{Z}}_j)$ at x_j based on the Z-scores $\mathfrak{z}_1, \ldots, \mathfrak{z}_j$ that comprise $\tilde{\mathfrak{Z}}_j$. Because $\sin(\Delta\psi^{(k)}(\tilde{\mathfrak{Z}}_j))$ is cyclic, the decision weight is cyclic. Let

$$\varphi^{(k)}(\boldsymbol{x}_j|\mu_x, \sigma_x) = \varphi^{(k)}(\boldsymbol{\mathfrak{z}}_j) = \sin(\Delta\psi^{(k)}(\tilde{\mathfrak{Z}}_j))\sec(\psi^{(k)}(\tilde{\mathfrak{Z}}_j))\sec(\psi^{(k)}(\tilde{\mathfrak{Z}}_{j-1})) \tag{4.4.11}$$

[38] Karni and Safra (1990, p. 491, eq(5)) also depends on a harmonic component like the one in (4.4.6).

where $\boldsymbol{x}_j = (x_1, \ldots, x_j)$ and $\boldsymbol{\mathfrak{z}}_j = (\mathfrak{z}_1, \ldots, \mathfrak{z}_j)$. Rewrite (4.4.1) and (4.4.10), respectively, as

$$\pi_j^{(k)} = \eta_0 p_j + \eta_1 \varphi^{(k)}(\boldsymbol{x}_j|\mu_x, \sigma_x) \tag{4.4.12}$$

$$RDEU(\boldsymbol{x}, \boldsymbol{p}) = \sum_{j=1}^{n} \pi_j^{(k)} U(x_j) = \eta_0 EU(\boldsymbol{x}, \boldsymbol{p}) + \eta_1 WU^{(k)}(\boldsymbol{x}, \boldsymbol{p}) \tag{4.4.13}$$

$$\text{where } \; EU(\boldsymbol{x}, \boldsymbol{p}) = \sum_{j=1}^{n} p_j U(x_j) \; \text{ and } \; WU^{(k)}(\boldsymbol{x}, \boldsymbol{p}) = \sum_{j=1}^{n} \varphi^{(k)}(\boldsymbol{x}_j|\mu_x, \sigma_x) U(x_j) \tag{4.4.14}$$

where EU is Von Neumann and Morgenstern (1953) utility functional and $WU^{(k)}$ is a weighted utility expression (Chew and Waller (1986, pg. 59, eq (2.6))) that depends on the k-cycle for $\varphi^{(k)}(\boldsymbol{x}_j|\mu_x, \sigma_x)$ in (4.4.11). When $\Delta\psi^{(k)}(\tilde{\mathfrak{Z}}_j) = 0$, $WU^{(k)} = 0$ so there is no weight given to the inherent distribution of outcomes and $RDEU$ is reduced to EU. Thus, we have just proven the following

Theorem 4.4.1 (Inconsistent probabilistic preferences)**.** *Let* $x_j, \; j = 1, 2, \ldots, n$ *for* $n \geq 3$ *be a statistical ensemble of outcomes, and* $\Phi(\cdot)$ *be the inherent cumulative normal distribution function predicted by the Good-Jaynes Max-Ent principle (Cover and Thomas (1991, p. 268)). Let* $\mathfrak{z}_j = \frac{x_j - \mu_x}{\sigma_x}$ *be a standardized score and* μ_x *and* σ_x *be the mean and standard deviation of the distribution* x_j*'s. Let*

$$w(p) = \eta_0 p + \eta_1 \tan(\psi(\mathfrak{z}))$$

be a representation of the inherent *HPWF predicted by Theorem 4.2.2 and Corollary 4.2.3 for standardized score* $\mathfrak{z}$ *and* monotone phase function $\psi(\mathfrak{z})$, *where* $\tan(\psi(\mathfrak{z}))$ *is a weighting function for outcomes satisfying Theorem 4.3.3. So that* $w(\cdot)$ *operates on a "k-cycle" for* $\tan((2k-1)\pi + \psi(\mathfrak{z})) = \tan(\psi(\mathfrak{z})), \quad k = 1, 2, \ldots.$ *As in (4.4.12), define* decision weights

$$\pi_j^{(k)} = \eta_0 p_j + \eta_1 \varphi^{(k)}(\boldsymbol{\mathfrak{z}}_j), \quad \sum_j \pi_j^{(k)} = 1$$

Suppose that probabilistic preferences *are represented by rank dependent utility (RDU) so that for von-Neuman utility* $U(x_j)$ *and inherent* prior probability $p_j = \Phi(\mathfrak{z}_j)$, *we have, for the correspond-*

ing gamble or lottery,

$$RDEU(\boldsymbol{x},\boldsymbol{p}) = \sum_{j=1}^{n} \pi_j^{(k)} U(x_j) = \eta_0 \sum_{j=1}^{n} p_j U(x_j) + \eta_1 \sum_{j=1}^{n} \varphi^{(k)}(\mathfrak{z}_j) U(x_j)$$

Then subjects will make different probabilistic choices when faced with the same stimuli because decision weights are cyclic unless their choices coincide with a probability k-cycle. Furthermore, we have the decomposition $RDEU(\boldsymbol{x},\boldsymbol{p}) = EU(\boldsymbol{p}) \oplus WU(\boldsymbol{x})$. □

Remark 4.4.1. The RDU decomposition result above was anticipated by Wakker (1994) who axiomatized decomposition of RDU preferences into probabilistic risk attitude and utility based components. In Appendix 4.I we show how the decomposition analysis extends to regret theory as well (Loomes and Sugden, 1983). Thus, the HPWF decomposes RDEU and regret theory into core EUT plus functionally equivalent addends.

Theorem 4.4.2 (Almost sure inconsistent probabilistic preferences)**.** *Probabilistic preferences induced by non-expected utility decision weights are different for the same stimuli almost surely.* □

Proof. According to Theorem 4.4.1 probabilistic preferences are consistent iff $\pi_j^{(k)} = \pi_j$ for every j in a "k-cycle". However, for fixed j the Lebesgue measure of the set $\{k : \pi_j^{(k)} = \pi_j, \forall k \in \mathbb{R}_+\}$ is zero. Thus, for each j probabilistic preferences are inconsistent except on a set with Lebesgue measure zero. □

If probabilistic preferences are cyclic, then subjects will make different choices with different probabilities when repeatedly presented with the same or similar stimuli over time by breaking the cycle. In which case, choice depends on subjects' location in the probability cycle and not so much on stimuli. This point was made by Regenwetter et al. (2011, pg. 43) who introduce a mixture model in which choice vary because DMs are in different mental states. Our result has implications for the "probabilistic loss aversion" concept, introduced in Schmidt and Zank (2008, pg. 213), based on a ratio of decision weights.

4.4.2 Resolution of preference reversal phenomenon in experiments

In this section we show how PR is characterized and resolved in an abstract experiment. Let $\boldsymbol{x} = \{x, y, z\}$ be a set of non negative outcomes, $U(\cdot)$ be a utility function that satisfies Von Neumann and Morgenstern (1953) axioms, and $\boldsymbol{p} = \{p_x, p_y, p_z\}$ be the corresponding probability distribution of $\boldsymbol{x}$. We make the following assumptions.

Assumption 4.4.3 (Experimental tasks)**.** *$E1$ and $E2$ are two temporally spaced experimental procedures that elicit preferences. Subjects choose in $E1$ and bid in $E2$ over the same lotteries.* □

Assumption 4.4.4 (Procedure invariance)**.** *The expected value of a lottery under a choice or bid procedure is the same.* □

Assumption 4.4.5 (Observed preference reversal)**.** *In $E1$ our DM expresses the preference $x \succ y \succ z$, and in $E2$ she reverses preference such that $y \succ z \succ x$.* □

Let π_j^i, $i = 1, 2$; $j = x, y, z$ be the decision weights computed via Quiggin's transformation procedure in each experiment. For example, π_x^1 and π_x^2 are a DM's decision weights for x in $E1$ and $E2$, respectively. In principle, DMs choose under $E1$ and bid under $E2$ from each of three lotteries which can be displayed as follows:

$$L_1 \equiv (x, p_x; \;\; y, p_y; \;\; 0, 1 - p_x - p_y) \tag{4.4.15}$$

$$L_2 \equiv (y, p_y; \;\; z, p_z; \;\; 0, 1 - p_y - p_z) \tag{4.4.16}$$

$$L_3 \equiv (x, p_x; \;\; z, p_z; \;\; 0, 1 - p_x - p_z) \tag{4.4.17}$$

There are 3! ways in which the lotteries can be ordered. A simple experiment design[39] for our purposes may be to choose a compound binary choice experiment $E1$ as a baseline such that

$$E1 \equiv \{A_s, \theta; \quad L, 1 - \theta\} \tag{4.4.18}$$

where A_s is a conveniently selected reward with probability θ, and $L \equiv \{L_1, L_2, L_3\}$ is a lottery to be played out if selected with probability $1 - \theta$. We assume that if L is chosen, then DMs

[39] See Harrison and Rutström (2008) for a comprehensive survey and taxonomy of experimental designs and econometric approaches.

choose $x \succ y$ in L_1; choose $y \succ z$ in L_2; choose $x \succ z$ in L_3. This choice pattern is consistent with transitivity. Let $A_v = xp_x + yp_y + zp_z$ be the actuarial value of the lottery so that the probability $\theta = \frac{A_v}{A_v + A_s}$ equalizes the expected value of DMs choices between A_s and the lottery L.

In the second experiment procedure $E2$, the rewards can be scaled by a common factor c and subjects can be randomly assigned to any of the following experiments:

$$E2 \equiv \{cA_s, \theta; \quad c\widehat{L}, 1-\theta\}, \text{ where} \tag{4.4.19}$$

$$\widehat{L} \in \Big\{ \{L_1, L_2, L_3\}, \{L_1, L_3, L_2\}, \{L_2, L_1, L_3\}, \\ \{L_2, L_1, L_3\}, \{L_3, L_1, L_2\}, \{L_3, L_2, L_1\} \Big\} \tag{4.4.20}$$

$c\widehat{L}$ means that the outcomes are scaled by c. Since a constant scale does not affect transitivity $E2$ will be effectively "identical" to $E1$ up to randomization.[40] For the purpose of exposition we assume that for a given $\widehat{L}$ in $E2$ DMs reverse the choices they made in L_1 in L_3, i.e., they bid such that $z \succ x$. The analysis that follow applies only to those DMs that choose L in $E1$ and $\widehat{L}$ in $E2$. In principle, preference reversal is also manifest if a DM choose A_s in E_1 and $c\widehat{L}$ in E_2 or vice versa. We address that in (4.4.22) and (4.4.23) below.

Preference reversal in E2

(i) The simple lottery representation. In $E2$ DMs weigh probilities by $w(\theta)$ and $(1 - w(1-\theta))$ (Quiggin, 1993, p. 57). Since we assumed that $\{\theta, cA_s\} \sim \{1-\theta, c\widehat{L}\}$, under EUT we expect $\theta U(cA_s) = (1-\theta)U(c\widehat{L})$. In particular, according to RDU (a generalization of EUT (Quiggin, 1993)) under the experimental design in $E2$

$$w(\theta)U(cA_s) = (1 - w(1-\theta))U(c\widehat{L}) \tag{4.4.21}$$

In the context of our HPWF $w(\theta) = \eta_0\theta + \eta_1 + \tan(\psi(\mathfrak{z}))$. So (4.4.21) reduces to

$$\eta_1[\tan(\psi(Z_{cA_s}))U(cA_s) - \tan(\psi(Z_{cA_v}))U(c\widehat{L})] = \eta_0[\theta U(cA_s) - (1-\theta)U(c\widehat{L})] \tag{4.4.22}$$

[40] Holt and Laury (2002) report an increase in risk aversion when stake size is scaled.

After applying EUT to $E2$ in (4.4.19), the RHS in (4.4.22) vanishes. So we are left with

$$\tan(\psi(Z_{cA_s}))U(cA_s) = \tan(\psi(Z_{cA_v}))U(c\widehat{L}) \Rightarrow \tan(\psi(Z_{cA_v})) = c_u \tan(\psi(Z_{cA_s})) \tag{4.4.23}$$

where $c_u = U(c\widehat{L})/U(cA_s)$. The harmonic relationship in (4.4.23) holds only for complete probability cycles. That is, $k \in \mathscr{K}$ must be divisible by $d(k)$ according to Definition 4.4.1. If that cycle is broken, then we would expect PR in $E2$ for those DMs who chose L in $E1$ and bid A_s in E2. However, that type of PR does not address Assumption 4.4.5. Thus, our interest lies in those DMs who bid $c\widehat{L}$ in $E2$ and satisfy Assumption 4.4.5.

(ii) Compound lottery representation. Under the transitivity axiom hypothesis, DMs ordinal selection in L_1 should be preserved in $\widehat{L}$. So we can set $c = 1$ without loss of generality. We assume that the decision weights for the rank ordered outcomes in $E1 : x \succ y \succ z$ in (4.4.18) and $E2 : y \succ z \succ x$ in (4.4.19) are computed as follows:

$$\pi_z^1 = w(p_z); \quad \pi_y^1 = w(p_z + p_y) - w(p_z); \quad \pi_x^1 = 1 - w(p_z + p_y) \tag{4.4.24}$$

$$\pi_x^2 = w(p_x); \quad \pi_z^2 = w(p_z + p_x) - w(p_x); \quad \pi_y^2 = 1 - w(p_z + p_x) \tag{4.4.25}$$

According to received theory, under the principle of procedure invariance (Tversky et al., 1990, p. 204) RDEU should be the same for $E1$ and $E2$ by virtue of Assumption 4.4.3. Let $RDEU_1$ and $RDEU_2$ be the valuation for rank ordered choices in $E1$ and $E2$ respectively. Thus, we have

$$RDEU_1(\boldsymbol{x}, \boldsymbol{p}) = \pi_x^1 U(x) + \pi_y^1 (U(y) + \pi_z^1 U(z) \tag{4.4.26}$$

$$RDEU_2(\boldsymbol{x}, \boldsymbol{p}) = \pi_x^2 U(x) + \pi_y^2 (U(y) + \pi_z^2 U(z) \tag{4.4.27}$$

Under Assumption 4.4.5, preference reversal implies $RDEU_1 \neq RDEU_2$, In which case

$$(\pi_x^1 - \pi_x^2)U(x) + (\pi_y^1 - \pi_y^2)U(y) + (\pi_z^1 - \pi_z^2)U(z) \neq 0 \tag{4.4.28}$$

Since U preserves ordinality, under $E1$ we can normalize U by setting $U(y) = 0$ and $U(x) = 1$ to simplify the analysis (Anscombe and Aumann (1963, p. 201), Karni and Safra (1990, p. 493), Quiggin (1993, p. 63)). Thus, (4.4.28) reduces to

$$(\pi_x^1 - \pi_x^2) + (\pi_z^1 - \pi_z^2)U(z) \neq 0 \Rightarrow U(z) \neq -\frac{(\pi_x^1 - \pi_x^2)}{(\pi_z^1 - \pi_z^2)} > 0 \tag{4.4.29}$$

Since the underlying probabilities p_x and p_z do not change in $E1$ and $E2$, we have for

$$(\pi_x^1 < \pi_x^2 \quad \text{and} \quad \pi_z^1 > \pi_z^2) \text{ or } (\pi_x^1 > \pi_x^2 \quad \text{and} \quad \pi_z^1 < \pi_z^2) \text{ and given } k \text{ from } \quad (4.4.12) \tag{4.4.30}$$

$$\pi_x^1 = \eta_0 p_x + \varphi_{1,x}^{(k)}(\mathfrak{z}_x) < \pi_x^2 = \eta_0 p_x + \varphi_{2,x}^{(k)}(\mathfrak{z}_x) \Rightarrow \varphi_{1,x}^{(k)}(\mathfrak{z}_x) < \varphi_{2,x}^{(k)}(\mathfrak{z}_x) \tag{4.4.31}$$

$$\pi_z^1 = \eta_0 p_z + \varphi_{1,z}^{(k)}(\mathfrak{z}_z) > \pi_z^2 = \eta_0 p_z + \varphi_{2,z}^{(k)}(\mathfrak{z}_z) \Rightarrow \varphi_{1,z}^{(k)}(\mathfrak{z}_z) > \varphi_{2,z}^{(k)}(\mathfrak{z}_z) \tag{4.4.32}$$

where $\varphi_{1,x}^{(k)}(\mathfrak{z}_x)$, $\varphi_{1,z}^{(k)}(\mathfrak{z}_x)$, $\varphi_{1,z}^{(k)}(\mathfrak{z}_z)$, $\varphi_{2,z}^{(k)}(\mathfrak{z}_z)$ are the cyclic components of π in Theorem 4.4.1, and PR is driven by those components. In (4.4.10), the decision weight is controlled by $\varphi^{(k)}$ through the jump $\sin(\Delta\psi^{(k)}(\tilde{\mathfrak{z}}_j))$. So under Theorem 4.4.2 and the inequalities in (4.4.31), (4.4.32) we have

$$\sin(\Delta\psi_{i,j}^{(k)}(\tilde{\mathfrak{z}}_j)) \neq \sin(\Delta\psi_{i,j}^{(k)}((2k-1)\pi + \tilde{\mathfrak{z}}_j)), \;\; i = 12;\; j = x, y, z \tag{4.4.33}$$

According to (4.4.33) $d(k)$ is not a gcd of $k \in \mathscr{K}$ in Definition 4.4.1, i.e., probability cycles are broken. Because we assumed procedure invariance between $E1$ and $E2$, the preference reversal in (4.4.30) is due to incomplete probability cycles. Moreover, for small jumps $\sin(\Delta\psi_{i,j}^{(k)}(\tilde{\mathfrak{z}}_j)) \approx \Delta\psi_{i,j}^{(k)}(\tilde{\mathfrak{z}}_j)$. So the PR phenomenon is a momentary fluctuation in the evaluative process. Thus, PR is eventually resolved as $\Delta\psi_{i,j}^{(k)}(\tilde{\mathfrak{z}}_j) \to 0$. We summarize this result in

Theorem 4.4.6 (Temporal PR)**.** *Preference reversal is due to momentary fluctuations of the evaluative process and it is resolved when probability cycles are complete.* □

4.4.3 Experimenter interference of HPWF and misperception of preference reversal

In practice, the experimenter *assigns* an *observed* probability distribution to $\boldsymbol{x}$. Call it $\boldsymbol{p}^o$. In which case we have $p_j^o = p_j + e_j^o$. So the inherent probability p_j is unobserved and disturbed by e_j^o. Cf. Busemeyer et al. (2011, p. 193)("drawing a conclusion from one judgment changes the context, which disturbs the state of the cognitive system"). Substitution of the observed values in the equations above do not alter the analysis since we simply substitute $p_j = p_j^o - e_j^o$. However, the inherent probability distribution is *disturbed* and that may cause experimenters to report PR when there is none (Regenwetter et al., 2011, p. 44). To evaluate that hypothesis we make the following

Assumption 4.4.7 (Transitivity)**.** *The transitivity axiom holds.*

The choice of p_j^o induces a Z-score different from $\mathfrak{z}_j$. That is, $p_j^o = \Phi(\mathfrak{z}_j^o)$ for some Z-score $\mathfrak{z}_j^o \neq \mathfrak{z}_j$. Thus, the HPWF in Theorem 4.3.3 is altered by imposition of *ex ante* probabilities p_j^o as follows (with $\boldsymbol{x}$ suppressed)

$$w(p_j) = \eta_0 p_j + \eta_1 \tan(\psi(\boldsymbol{z}_j)) \tag{4.4.34}$$

$$w(p_j^o) = \eta_0 (p_j^o - e_j^o) + \eta_1 \tan(\psi(\mathfrak{z}_j^o)) \tag{4.4.35}$$

Let π_j^o be the observed decision weight, and π_j be the true decision weight. In the sequel superscript (k) implies that k-cycles are in play for a given variable. According to Theorem 4.4.1 and (4.4.35) we have

$$\pi_j^{(k)} = \eta_0 p_j + \eta_1 \varphi^{(k)}(\boldsymbol{\mathfrak{z}}_j) \tag{4.4.36}$$

$$\pi_j^{o(k)} = \eta_0 (p_j^o - e_j^o) + \eta_1 \varphi^{(k)}(\boldsymbol{\mathfrak{z}}_j^o) \tag{4.4.37}$$

$$\Rightarrow \pi_j^{o(k)} = \pi_j^{(k)} + \eta_1 (\varphi^{(k)}(\boldsymbol{\mathfrak{z}}_j^o) - \varphi^{(k)}(\boldsymbol{\mathfrak{z}}_j)) \tag{4.4.38}$$

Let π_j^{o1}, π_j^{o2}, $j = x, y, z$ be the observed decision weights in $E1$ and $E2$ respectively. Under Assumption 4.4.7 for a given k we have $\pi_j^{1^{(k)}} = \pi_j^{2^{(k)}}$ so there is no preference reversal. By (4.4.38), the corresponding relationship for observed decision weights is

$$\pi_j^{o1^{(k)}} - \eta_1 (\varphi_1^{(k)}(\boldsymbol{\mathfrak{z}}_j^o) - \varphi_1^{(k)}(\boldsymbol{\mathfrak{z}}_j)) = \pi_j^{o2^{(k)}} - \eta_1 (\varphi_2^{(k)}(\boldsymbol{\mathfrak{z}}_j^o) - \varphi_2^{(k)}(\boldsymbol{\mathfrak{z}}_j)) \tag{4.4.39}$$

where φ_1 and φ_2 are the corresponding phase functions in $E1$ and $E2$. Under Assumption 4.4.7

$$\pi_j^{1^{(k)}} = \pi_j^{2^{(k)}} \Rightarrow \varphi_1^{(k)}(\boldsymbol{\mathfrak{z}}_j) = \varphi_2^{(k)}(\boldsymbol{\mathfrak{z}}_j) \tag{4.4.40}$$

After $\eta_1 \varphi_1^{(k)}(\mathfrak{z}_j)$ terms cancel, (4.4.39) reduces to

$$\pi_j^{o1^{(k)}} = \pi_j^{o2^{(k)}} + \eta_1 \left(\varphi_1^{(k)}(\mathfrak{z}_j^o) - \varphi_2^{(k)}(\mathfrak{z}_j^o) \right) \tag{4.4.41}$$

Because of experimenter interference, there is no guarantee that the expression in brackets in the right hand side of (4.4.41) is 0. In fact, Theorem 4.4.2 implies that more often than not $\varphi_1^{(k)}(\mathfrak{z}_j^o) \neq \varphi_2^{(k)}(\mathfrak{z}_j^o)$. Thus, our experimenter will report preference reversal because she observes

$$\pi_j^{o1^{(k)}} \neq \pi_j^{o2^{(k)}} \tag{4.4.42}$$

even though the true but unobserved relationship in (4.4.40) is based on the transitivity axiom Assumption 4.4.7. Thus we conclude with

Theorem 4.4.8 (Observer effect of experimenter misperception of PR)**.** *Experimenter assignment of ex ante probabilities to the elements of a statistical ensemble or random field of outcomes interferes with the inherent PWF for those outcomes, and induces observed PR when the true state is no PR.* □

Theorem 4.4.8 manifests the "uncertainty principle" or "observer effect" articulated in Von Neumann (1955, pp. 418-420). Specifically, "we must always divide the world into two parts, the one being the observed system, the other the observer. In the former, we can follow up all physical processes (in principle at least) arbitrarily precisely. In the latter, this is meaningless. The boundary between the two is arbitrary to a very large extent," *ibid* p. 420.

4.5 Model simulations and estimates

In this section we provide a simple calibration exercise and simulation of the HPWF, and report on observed characteristics including but not limited to likelihood insensitivity, and fix point probability dynamics. In subsection 4.5.1 we show how our model accommodates likelihood insensitivity. In subsection 4.5.2 we analyze the entropy implications of fixed point probabilities.

4.5.1 Likelihood insensitivity of HPWF

For the purpose of exposition we calibrate $w(\Phi(\mathfrak{z}))$ in (4.3.10). As shown in (4.E.9) in the appendix, for $|\mathfrak{z}| \leq 4$, we derive the form

$$w(p,x) = \frac{3p}{2} + \frac{3}{8\sqrt{2\pi}}\tan(\theta(\mathfrak{z})) - \frac{3}{8\sqrt{2\pi}}\mathfrak{z} - \frac{7}{24\sqrt{2\pi}}\mathfrak{z}^3 + \frac{608}{15\sqrt{2\pi}} \tag{4.5.1}$$

where $\mathfrak{z} = (x-\mu_x)/\sigma_x$. The equivalent HPWF representation consistent with Theorem 4.2.2 is given by

$$w(\Phi(\mathfrak{z})) \approx \frac{3\Phi(\mathfrak{z})}{2} + \frac{3}{8\sqrt{2\pi}}\tan(\mathfrak{z}) - \frac{3}{8\sqrt{2\pi}}\mathfrak{z} - \frac{7}{24\sqrt{2\pi}}\mathfrak{z}^3 + \frac{608}{15\sqrt{2\pi}} \tag{4.5.2}$$

A plot of the probability weighting function in (4.5.2) is depicted in Figure 4.4 on page 199. To convert the Z-axis to a *psuedo-probability axis* Z^* we make the following scale and affine transformations:

$$w(Z^*) = \frac{w(\phi(Z))}{32} \tag{4.5.3}$$

$$Z^* = \frac{1}{2} + \frac{Z}{8} \tag{4.5.4}$$

$$\widetilde{Y} = \max(3.7 \times Z^*,\ 1) \tag{4.5.5}$$

We use $\mathfrak{z}$ to denote given values of Z. The last equation (4.5.5) is a ray from the origin or probability adjustment that cuts through the fixed point e^{-1}. The inverted S-shape is retained so that $w(\mathfrak{z}^*)$ in (4.5.3) is indeed a probability weighting function over the *psuedo-probability axis* Z^* in (4.5.4)–as depicted in Figure 4.5 on page 200. The estimated HPWF has the following characteristics.

- The shape of the curve is similar to Tversky and Fox (1995, Fig. 5), and Abdellaoui et al. (2011, Fig. 1c) likelihood insensitivity source function. Specifically, the relatively flat portion of the HPWF (less than 45° incline from horizontal) compared to the 45° line for linear probabilities, suggests that prospect theory is insensitive, in this case very insensitive, to changes in uncertainty compared to expected utility theory (Tversky and Fox (1995, pg. 276)). Thereby suggesting that entropy methods can be used to derive probability weighting functions in ambiguous situations that include unobserved inherent probability distributions.

- The probability weighting plots indicate that the process is slow varying in roughly the middle third, and heavily weighted in each of the two other thirds that correspond to over weighting low ranked outcomes and underweighting high ranked outcomes. This depicts "the principle of bounded subadditivity" introduced in Tversky and Fox (1995, pg. 270) and Tversky and Wakker (1995). In this case, the lower third "possibility gap", and upper third "possibility gap", are consistent with empirical research (Tversky and Fox (1995, pg. 276)).

We note that the class of flexible two-parameter pwfs introduced by Lattimore et al. (1992, p. 382) is also able to generate likelihood insensitivity.

4.5.2 Fixed point probability of pwfs has maximum entropy

The maximum entropy associated with the underlying unobserved probability distribution is depicted in Figure 4.6 on page 201. The value for $P = e^{-1} \approx 0.38$ depicts maximum entropy. This is the fixed point probability that separates over weighting and under weighting regimes in cumulative prospect theory. By comparison, the maximum entropy for the affine transformation of probability weighted functionals is as indicated in Figure 4.7 on page 202. Here the maximum entropy point $P = 0.38$ is the same as before. The fact that maximum entropy occurs at the fixed point e^{-1} is puzzling. Because it implies that that invariant point axiomatized by (Prelec, 1998) is the point of highest uncertainty or least information in a given distribution.

Nonetheless, we superimpose entropy and probability weighting functionals on the *psuedo-probability axis* as depicted in Figure 4.8 on page 203. There, the *psuedo-probability* Z^* has to be adjusted as indicated in (4.5.5) so that it cuts the probability weighting functional at the fixed point e^{-1}. Otherwise, the fixed point for the psuedo-probability is derived from $w(\mathfrak{z}^*) = \mathfrak{z}^*$. This suggests that the psuedo-probability fixed point occurs around $\mathfrak{z}^* = 0.5125$ with entropy around 0.34, which corresponds to $Z = 0.1$ and $-p\ln(p) = 0.33$ in Table 4.1. This result is still consistent with prospect theory because there exist "a more general form, $w(p) = \exp(-\beta(-\ln p)^{\alpha})$ due to (Prelec, 1998, pg. 499), which is not constrained to the $1/e$

fixed point value".[41] It should be noted that "[a]n inverse S-shaped weighting function can be completely below the 45-degree line depicting non-transformed probabilities, or it can be completely above it," Abdellaoui et al. (2010, pg. 41). Also, "a linear weighting function is seen as the benchmark for measuring optimism and pessimism," *ibid* at p. 49. So the linear probability weighting function in (4.5.5) is that of a pessimist. To see this consider a subject with expected utility $E[U]$ and probability weighting function

$$\tilde{w}(p) = \max\{\alpha p, 1\}, \quad 0 < p < 1, \;\; \alpha > 1 \tag{4.5.6}$$

The transformed expected utility $\tilde{E}[U]$ along the line above the diagonal is given by

$$\tilde{E}[U] = \alpha E[U] > E[U] \tag{4.5.7}$$

This shows that under EUT, for a given probability distribution P a subject with utility function U and weighting function $\tilde{w}(P) > P$ will have more probabilistic risk aversion (Wakker, 1994, p. 10), i.e., the pwf in (4.5.5) and (4.5.6) is piecewise concave. So (4.5.5) is the weighting function for a pessimist. Even so, the *psuedo-probability* weighting function is slightly greater than the pessimist weighting function over low ranked outcomes, and they coincide at the point of maximum entropy as indicated in Figure 4.8. Evidently, maximum entropy provides a robust estimate of the fixed point probability in cumulative prospect theory. This result is not surprising because our model predicted that maximum entropy of unconstrained unobserved inherent probability distributions yields Von Neuman-Morgenstern preferences which are linear in probabilities–in this case even by a constant scale factor. Thus, we close with the following

Proposition 4.5.1 (Maximum entropy and cluster set of fixed points)**.**
Let F be a cumulative probability distribution, and $w(F)$ be a probability weighting functional. Define the set

$$C(F) = \{F \mid -w(F)\ln(w(F)) = F,\ 0 \leq F \leq 1\} \tag{4.5.8}$$

[41] Refer to Abdellaoui et al. (2010, pg. 49) ("Empirical studies, using more general weighting functions, suggests that this intersection is around 1/3").

Then $C(F)$ is a cluster set of fixed point functions for probability weighting. Moreover, in the restricted case when $w(F) = F$ the set $C(F)$ is a singleton comprised of the Tversky and Kahneman (1992) fixed point e^{-1}.

Proof. See Appendix section 4.F. □

Remark 4.5.1. McLennan (2012) identified sets like $C(F)$ as an *essential set* of fixed points. Alternatively, $C(F)$ is an invariant subspace of $[0,1]$. That is, $w : C(F) \to C(F) \subset [0,1]$.

Corollary 4.5.2 (Linear probability weighting on essential set of fixed points)**.** *The probability weighting function w is a* linear operator *on the essential cluster set of fixed points $C(F)$ separating expected and nonexpected utility theories.*

Proof. See Appendix section 4.G. □

This result suggests that there is an invariant linear segment of every probability weighting scheme which separates "optimists" from "pessimists" in the sense of Tversky and Wakker (1995, pg. 1264) and Abdellaoui et al. (2010, pg. 41). Perhaps more important, it provides a theoretical foundation for mixture preference models (Harrison and Rutström (2009)). For the cluster set of fixed point probabilities is a linear subspace that support von-Neuman-Morgenstern (VNM) utility, see Von Neumann and Morgenstern (1953). Whereas the complementary subspace is nonlinear so it supports nonexpected utility theories.

4.6 Conclusions

This paper treats the outcome dimension of a gamble or lottery as an underlying statistical ensemble. Whereupon it employs the principle of maximum entropy to identify the probability weighting function (pwf) based on a DM's information about the moments of the outcome. By so doing, it characterizes a DM's inherent pwf for a distribution of outcomes. We find that probabilistic preference is explained by an inherent harmonic probability weighting function (HPWF). Furthermore, we introduce an abstract representation theorem for HPWF that has implications for the quantum probability strand of decision theory literature outside the scope of this paper. Our theory also shows how decision weights, obtained from Quiggin's rank dependent utility transformation procedure, can be decomposed into a part due to linear probabilities and a part due to harmonic weighting of outcomes. That sets the stage for the identity $RDU = EU \oplus WU$. RDU collapses to EU when probability cycles are completed.

The HPWF theory is applied to resolve the preference reversal puzzle. We find that experimenters interfere with the inherent HPWF when they assign probabilities to outcomes *ex ante*, and that that interference breaks the probability cycles of subjects in the experiment. This causes experimenters to misreport intransitivity of preferences, i.e., preference reversal, when the true state is no preference reversal. Because subjects behaviours are transitive once their probability cycles are completed. We show that the HPWF fixed point probability is e^{-1}–the same value axiomatized in the literature. However, that fixed point has maximum entropy. Our preliminary analysis suggests that that result is explained by the existence of a cluster set of fixed points or fixed probability distribution. And, preferences over the cluster set of fixed points are linear in probabilities. In other words, every point on the diagonal of a unit square is an admissible fixed point probability. Thus casting doubt on the axiomatized fixed point of e^{-1}. The HPWF in this paper is limited to concave-convex shape. However, for further research, in principle one should be able to derive convex-concave shape by imposition of suitable identifying restrictions on the HPWF. This paper adds to the literature by virtue of extending probability weighting functionals to the complex domain by and through harmonics in probability weighting. It provides "new" analytic tools that portend further research on the issue of probability cycles and intransitivity of preferences.

4.A APPENDIX

4.B Order relations

The material in this section is drawn from Willard (1970, p. 5). A binary relation $\mathscr{R}$ on a set A is any subset of $A \times A$. The relation $(a,b) \in \mathscr{R}$ is denoted $a\mathscr{R}b$. A relation $\mathscr{R}$ is *reflexive* iff $a\mathscr{R}a$ for $a \in A$, *symmetric* iff $a\mathscr{R}b$ implies $b\mathscr{R}a$ for all $a,b \in A$, *antisymmetric* iff $a\mathscr{R}b$ and $b\mathscr{R}a$ implies $a = b$ for all $a,b \in A$, and transitive iff $a\mathscr{R}b$ and $b\mathscr{R}c$ implies $a\mathscr{R}c$ for all $a,b,c \in A$.

Partial order. A relation $\mathscr{R}$ on A is a partial order provided $\mathscr{R}$ is reflexive, antisymmetric and transitive. Thus, $\succeq$ is a partial order on $\mathbb{R}$.

4.C Proof of Lemma 4.2.1 functional Gâteaux derivative

Proof. Technically, we are using a Gâteaux type derivative (Siddiqui, 2004, §5.2). Let $T_\mu(f) = \int_A f dmu(dx)$ be a linear operator, $\| \ \|$ be a norm defined on T_μ, and $DT_\mu(f)h$ be the value of the Gâteaux derivative at f in direction h.

$$\lim_{g\to 0}\left\|\frac{T_\mu(f+gh)-T_\mu(f)h}{g}-DT_\mu(f)h\right\|=0$$

implies $DT_\mu(f) = T_\mu$ for every f (Siddiqui, 2004, Rem. 5.2.1). Choose h to be the identity function direction so that $T_m u(h) = \int_A h\mu(dx) = \mu(A)$. Thus, $DT_\mu(f) = \mu(A)$ where $\dfrac{\partial F}{\partial f} = DT_\mu(f)$. □

4.D Proof of Theorem 4.2.2 HWPF existence

The proof of the theorem rests on first and second order optimality criteria and transformation of the underlying moment problem.

Proof. The first order necessary conditions for a maximum for criterion function $J(w(F))$ requires

$$\frac{\partial J}{\partial f} = 0 \text{ and } \quad \frac{\partial \tilde{H}(F)}{\partial f} = w'(F)\mu(A)[1+\ln(w(F))] \tag{4.D.1}$$

$$\text{which implies } w'(F)\mu(A)[1+\ln(w(F))] + \lambda_0\mu(A) + \sum_{i=1}^{m}\lambda_i r_i(x) = 0 \tag{4.D.2}$$

$$w(F) = e^{-1}\exp\left(-\{w'(F)\mu(A)\}^{-1}\left\{\lambda_0\mu(A)+\sum_{i=1}^{m}\lambda_i r_i(x)\right\}\right) \tag{4.D.3}$$

A first order approximation expansion of the latter equation reveals

$$w(F) = e^{-1} - e^{-1}\frac{\lambda_0\mu(A)+\sum_{i=1}^{m} r_i(x)}{w'(F)\mu(A)} \tag{4.D.4}$$

Undeniably, the quantity e^{-1} constitutes the fixed point for compound invariance in Prelec (1998, pp. 503-504). Furthermore, $\lambda_0\mu(A) + \sum_{i=1}^{m}\lambda_i r_i(x)$ represents an affine transformation of the moment conditions imposed on the nonparametric probability density function f in (4.2.1). And $w'(F)$ and $\mu(A)$ are scale parameters. The shape of w is characterized by $\text{sign}(w'(F))$.

It is known that for $\mu(A) < \infty$, the moments $1, x, x^2, \ldots$ can be orthogonalized to form a basis for the canonical Hilbert space $L^2_\mu(A)$ with norm $\| \ \|_{L^2_\mu}$. And that the harmonic sequence $\{e^{i\omega_j x}\}_{j=1}^{\infty}$ is a canonical basis for that space(Akheizer and Glazman (1961, pp. 27-28)). Therefore we can replace $\lambda_0\mu(A) + \sum_{i=1}^{m}\lambda_i r_i(x)$ with some complex valued function $\tilde{w}(F)$ induced by orthogonalized moments defined over that basis. More relevant here its *m-order approximation*

$$\tilde{w}_m(F) = \sum_{j=1}^{m} a_j e^{i\omega_j x} \tag{4.D.5}$$

where a_j is the coordinate of $\tilde{w}(F)$ with respect to the orthogonal basis. It is a so called abstract

Fourier coefficient derived from the canonical inner product for L^2 (Dunford and Schwartz (1957, p. 858)) given by

$$< f,g> = \int_X f\bar{g}\mu(dF), \qquad a_j = \frac{<\tilde{w}(F), e^{i\omega_j x}>}{<e^{i\omega_j x}, e^{i\omega_j x}>} \tag{4.D.6}$$

So the first order approximation is given by

$$w(F) \approx e^{-1} - e^{-1}\{w'(F)\mu(A)\}^{-1}\tilde{w}_m(F) \tag{4.D.7}$$

$$= e^{-1} - e^{-1}\{w'(F)\}^{-1}\sum_{j=1}^{m}\tilde{a}_j e^{i\omega_j x} \tag{4.D.8}$$

where $\tilde{a}_j = \frac{a_j}{\mu(A)}$ is the scaled coordinate with respect to the measure $\mu(A)$. □

4.E Proof of Theorem 4.3.3 HPWF identification

The proof of the theorem is organized as follows. We calibrate a HPWF with numerical values and then generalize the result.

Proof. The restriction $0 \leq w(\Phi(\mathfrak{z})) \leq 1$ implies that (4.3.10) reduces to

$$(3\Phi(\mathfrak{z}) - 2) \leq I(\mathfrak{z}) \leq 3\Phi(\mathfrak{z}) \tag{4.E.1}$$

Cursory inspection indicates that for $P = \Phi(\mathfrak{z}) = \frac{1}{3}$, in (4.E.1) we have $-1 \leq I(\mathfrak{z}) \leq 1$. Moreover, $P = \frac{1}{3}$ is near to the fixed point e^{-1} so equation (4.3.10) is well defined. Negative values of $\mathfrak{z}$ suggests that the cubic term in the evaluated integral $I(\mathfrak{z})$ is dominant negative, and vice versa for positive values.[42] So in (4.3.10) w tends to be higher for small negative $\mathfrak{z}$. For positive values of $\mathfrak{z}$ the cubic term is dominant positive and w tends to be lower. So the weighting function in (4.3.10) [weakly] retains inverted S-shape characteristics of the probability weighting function (Prelec (1998) and Tversky and Kahneman (1992)).

We turn now to evaluate (4.3.10). We begin with a very rough linear approximation for the normal density $\Phi'(y)$ in (4.3.4) to evaluate $I(\mathfrak{z})$ in (4.3.7) as follows.

$$\Phi'(y) \approx \frac{1}{\sqrt{2\pi}}\left(1 - \frac{y^2}{2}\right) \tag{4.E.2}$$

$$I(\mathfrak{z}) \approx \frac{1}{\sqrt{2\pi}}\int_{-4}^{\mathfrak{z}}\left(y^2 - \frac{y^4}{2}\right)dy = \frac{1}{\sqrt{2\pi}}\left[\frac{y^3}{3} - \frac{y^5}{10}\right]_{-4}^{\mathfrak{z}} \tag{4.E.3}$$

Substitution of these expressions in and algebraic manipulation of (4.3.10) leads to

$$w(\Phi(\mathfrak{z})) = \frac{3\Phi(\mathfrak{z})}{2} - \frac{1}{60\sqrt{2\pi}}(10\mathfrak{z}^3 - 3\mathfrak{z}^5 - 2432) \tag{4.E.4}$$

(4.E.5)

[42] Blavatskyy (2013, p. 15) introduced a 2-parameter cubic pwf which performs just as well as several other popular pwfs.

The cubic and quintic terms in (4.E.4) suggest expansion of $\tan(\mathfrak{z})$ to 3 terms to get

$$\tan(\mathfrak{z}) \approx \mathfrak{z} + \frac{\mathfrak{z}^3}{3} + \frac{2\mathfrak{z}^5}{15} \Rightarrow \mathfrak{z}^5 \approx \frac{15}{2}\left(\tan(\mathfrak{z}) - \mathfrak{z} - \frac{\mathfrak{z}^3}{3}\right) \tag{4.E.6}$$

$$\Rightarrow w(\Phi(\mathfrak{z})) \approx \frac{3\Phi(\mathfrak{z})}{2} + \frac{3}{8\sqrt{2\pi}}\tan(\mathfrak{z}) - \frac{3}{8\sqrt{2\pi}}\mathfrak{z} - \frac{7}{24\sqrt{2\pi}}\mathfrak{z}^3 + \frac{608}{15\sqrt{2\pi}} \tag{4.E.7}$$

$$\tag{4.E.8}$$

Writing $p = \Phi(\mathfrak{z})$, $\mathfrak{z} = \Phi^{-1}(p)$, and replacing $\mathfrak{z}$ with a "residual" *monotone phase function* $\theta(\mathfrak{z})$ predicted by Corollary 4.2.3, we replace $\approx$ with $=$ to get

$$w(p,x) = \frac{3p}{2} + \frac{3}{8\sqrt{2\pi}}\tan(\theta(\mathfrak{z})) - \frac{3}{8\sqrt{2\pi}}\mathfrak{z} - \frac{7}{24\sqrt{2\pi}}\mathfrak{z}^3 + \frac{608}{15\sqrt{2\pi}} \tag{4.E.9}$$

Recall that $x = \mu_x + \sigma_x\mathfrak{z}$ so (4.E.9) is a *bivariate* function of p and x. A more parsimonious representation based on some [hereditary monotone] phase function $\psi(\mathfrak{z})$ would be

$$w(p,x) = \frac{3p}{2} + \tan(\psi(\mathfrak{z})), \text{ where} \tag{4.E.10}$$

$$\tan(\psi(\mathfrak{z})) = \frac{3}{8\sqrt{2\pi}}\tan(\theta(\mathfrak{z})) - \frac{3}{8\sqrt{2\pi}}\mathfrak{z} - \frac{7}{24\sqrt{2\pi}}\mathfrak{z}^3 + \frac{608}{15\sqrt{2\pi}} \tag{4.E.11}$$

Since the choice of $z = 4$ in Assumption 4.3.1 is arbitrary, in its most general form (4.E.10) can be rewritten to accomodate all values of $\mathfrak{z}$ as follows

$$w(p,x) = \eta_0 p + \eta_1 \tan(\psi(\mathfrak{z})) \tag{4.E.12}$$

where η_0 and η_1 are elevation and curvature parameters respectively. In the sequel we use this more general form. The term $\tan(\psi(\mathfrak{z}))$ dominates the harmonic PWF in (4.E.10). It is known that that term is cyclic, and that it has an inverse S-shape consistent with the probability weighting function popularized by Tversky and Kahneman (1992) cumulative prospect theory (CPT). For (4.E.12) to be a pwf we impose the following identifying restrictions:

$$0 \le \eta_0 p + \eta_1 \tan(\psi(\mathfrak{z})) \le 1 \Rightarrow -\frac{\eta_0}{\eta_1} \le \tan(\psi(\mathfrak{z})) \le \frac{1-\eta_0 p}{\eta_1} \tag{4.E.13}$$

$$p = 0 \Rightarrow \eta_1 \tan(\psi(\mathfrak{z})) = 0 \Rightarrow \psi(\mathfrak{z}) = n\pi, \; n = 0, 1, \ldots \tag{4.E.14}$$

$$p = 1 \Rightarrow \tan(\psi(\mathfrak{z})) = \frac{1-\eta_0}{\eta_1} \Rightarrow \psi(\mathfrak{z}) = \tan^{-1}\left(\frac{1-\eta_0}{\eta_1}\right) \tag{4.E.15}$$

□

4.F Proof of Proposition 4.5.1 maximum entropy of fixed pt cluster set

Proof. By construction we can rewrite $w(F) = \exp(-\frac{F}{w(F)})$. The distribution function(s) F which solves that nonlinear equation represents fixed probability distributions that satisfy the equation. By definition, F represents a continuum of probabilities. Specifically,

1. let ξ_p be the p-quantile of F. Define the set of probabilities $X(p) = \{p| -w(p)\ln(w(p)) =$

$p,\ F(\xi_p) = P(X \le \xi_p) = p,\ F \in C(F)\}$. By construction $X(p)$ is a cluster set of probabilities since it contains the accumulation or fixed points p that satisfy the entropy equation, and by construction $X(p) \subseteq C(F)$. Hence $C(F)$ is a hereditary cluster set.

2. Suppose $F \in C(F)$, and $p \notin X(p)$. The latter relation implies that $-w(F(\xi_p))\ln(w(F(\xi_p))) \neq F(\xi_p)$ and $F(\xi_p) \notin C(F)$. This contradicts our incipient hypothesis $F \in C(F)$. In which case $F(\xi_p) = p \in X(p)$ and $C(F) \subseteq X(p)$.

The results of 1. and 2. imply that $C(F) = X(p)$. In which case $C(F)$ is a cluster set of probabilities. The restriction $W(F(\xi_p)) = F(\xi_p) = p$ produces the fixed point solution $W(F) = F = \exp(-1)$. □

4.G Proof or Corollary 4.5.2 linear weighting on essential cluster set of fixed points

Proof. Let $h_1,\ h_2 \in C(F)$ and choose constants α and β such that $\alpha h_1 + \beta h_2 \in C(F)$. Thus by hypothesis $\alpha h_1 + \beta h_2$ is a fixed point if

$$w(\alpha h_1 + \beta h_2) = \alpha h_1 + \beta h_2 = \alpha w(h_1) + \beta w(h_2)$$

In which case w is a linear operator by definition. □

4.H Harmonic pwf and quantum probability models of decision

In this appendix we briefly digress from the main theme in the paper to show how our HPWF theory extends to the quantum probability models popularized by Jerome Busemeyer and his co-workers. See e.g., Busemeyer and Wang (2007); Pothos and Busemeyer (2009); Busemeyer et al. (2011). Cursory inspection of the functional form of $w(F)$ in Theorem 4.2.2 indicates that we can write

$$\mu(A) = \frac{\tilde{w}_m(F)}{\left(\log\left(\frac{1}{w(F)}\right) - 1\right) w'(F)} \tag{4.H.1}$$

Since $\tilde{w}_m(F)$ is complex valued, for internal consistency of real $(\mathscr{R}e)$ and imaginary $(\mathscr{I}m)$ parts, we must have for the numerator above

$$\mu(A) = \mathscr{R}e(\tilde{w}_m(F)) > 0, \qquad \mathscr{I}m(\tilde{w}_m(F)) = 0 \tag{4.H.2}$$

$$\log\left(\frac{1}{w(F)}\right) - 1 = -(1 + \log(w(F))) > 0 \tag{4.H.3}$$

However, there is a quantum probability interpretation as well. Let $\psi(x)$ be a wave equation such that in accord with Corollary 4.2.3

$$\psi(x) = -\left\{(1+\ln(w(F)))w'(F)\right\}^{-1} \|\tilde{w}_m(F)\|_{L^2_\mu} \exp(-i\vartheta(F)) \tag{4.H.4}$$

$$\|\psi(x)\|^2_{L^2_\mu} = \int_X |\psi(x)|^2 \mu(dx) = \|A(F)\|^2_{L^2_\mu} = \int_X |A(F)|^2 \mu(dF), \text{ where} \tag{4.H.5}$$

$$A(F) = -\left\{(1+\ln(w(F)))w'(F)\right\}^{-1} \|\tilde{w}_m(F)\|_{L^2_\mu} \tag{4.H.6}$$

Specifically, there exist some unit state function $\widehat{\psi}$ which satisfies the Schrödinger equation (Pothos and Busemeyer, 2009, §2.). Namely

$$i\frac{\partial \widehat{\psi}(x)}{\partial x} = \widehat{\vartheta}(F)\widehat{\psi}(x) \;\Rightarrow \widehat{\psi}(x) = \exp(-i\widehat{\vartheta}(F)) \tag{4.H.7}$$

$$\widehat{\psi}(x) = \frac{\psi(x)}{\sqrt{\int_X |A(F)|^2 \mu(dF)}}, \quad \int_X \|\widehat{\psi}(x)\|^2_{L^2_\mu} \mu(dx) = 1 \tag{4.H.8}$$

The underlying “Born rule” probability density is given by

$$\rho(x) = \|\widehat{\psi}(x)\|^2_{L^2_\mu} \tag{4.H.9}$$

In which case the probability of finding the wave function in some set B is given by

$$P(B) = P\{\widehat{\psi} \in B\} = \int_B \rho(x)\mu(dx) = \int_B \|\widehat{\psi}(x)\|^2_{L^2_\mu} \mu(dx) \tag{4.H.10}$$

The foregoing analysis gives rise to the following

Theorem 4.H.1 (Quantum preference states)**.**
Let $(X, \mathscr{B}(X), \mu)$ be outcome measure space, F be an unknown probabilistic distribution, and $\tilde{w}_m(F)$ be a complex valued functional of F. For some $B \in \mathscr{B}(X)$, let $\psi(x)$ represent the state of uncertainty about outcome $x \in B$, and $w(F)$ represent the maximum entropy probability weighting functional (MaxEnt-HPWF) for F over X. Then the state dependent MaxEnt-HPWF is given by

$$\psi(x) = -\left\{(1+\ln(w(F)))w'(F)\right\}^{-1} \|\tilde{w}_m(F)\|_{L^2_\mu} \exp(-i\vartheta(F))$$

Furthermore, the probability that uncertain outcome x is in B is given by

$$P(B) = \frac{1}{\int_X |A(F)|^2 \mu(dF)} \int_B |A(F)|^2 \mu(dF)$$

where $A(F) = -\left\{(1+\ln(w(F)))w'(F)\right\}^{-1} \|\tilde{w}_m(F)\|_{L^2_\mu}$.

□

Remark 4.H.1. The simplest time separable Schrödinger wave equation is typically written as $\psi(x,t) = \psi(x)f(t)$ where $H\Psi = i\hbar\frac{\partial \psi}{\partial t}$ for some [Hamiltonian] operator H (Szekeres, 2004, p. 383).

And the general solution is written $\psi(x,t) = \psi(x)\exp(-\frac{i}{\hbar}Et)$ where $\frac{H\psi(x)}{\psi(x)} = \frac{i\hbar\frac{\partial f(t)}{\partial t}}{f} = E$. So even though our equation does not include time it can be extended to do so. □

Remark 4.H.2. The "information" restrictions imposed in (4.H.2) implies that we can replace $\tilde{w}_m(F)$ with $\mu(B)$ in the theorem. In that case, $\mu(B)$ is a measure of the information content of the set $B \in \mathscr{B}(X)$. Thus, $P(B)$ is a subjective probability estimate based on the perceived information content of $\mu(B)$. □

4.I Harmonic pwf, regret theory, and core theory plus noise

This appendix sketches the relationship between regret theory and the HPWF. To fix ideas we sketch the Regret Theory (RT) model described in Loomes and Sugden (1983). Let A_i and A_k be two *actions* available in a *binary choice* situation. The *consequence* of choosing A_i if state j occurs is x_{ij}, and the *state dependent probability* is p_j. Let $C(\cdot)$ be a *choiceless* utility function such that $C(x_{ij}) = c_{ij}$. That is, c_{ij} is the utility a DM would experience if she did not have to make any choices. Assume that our DM is faced with the choice of either accepting A_i and *simultaneously* rejecting A_k or else accept A_k and simultaneously reject A_j. According to this model, should the j-th statee occur, our DM experiences x_{ij} but she misses out on x_{kj}. Recall that $C(x_{ij}) = c_{ij}$ and $C(x_{kj}) = c_{kj}$, so relative to c_{ij} she derives an incremental increase in utility if $c_{ij} > c_{kj}$ and rejoices or she experiences a decrement in utility if $c_{ij} < c_{kj}$ and regrets. Let the rejoice-regret function be $R(\cdot)$. We can think of $R(c_{ij} - c_{kj})$ as a function whose argument *switches* between positive (if rejoice) and negative (if regret). In other words, $R(\cdot)$ is a *weak rejoice-regret switching-function* where $R(0) = 0$. We write the modified utility function that accounts for rejoice or regret as:

$$m_{ij}^k = c_{ij} + R(c_{ij} - c_{kj}) \tag{4.I.1}$$

The expected utility associated with (4.I.1) over the set of all possible states $j = 1, \dots, n$ is given by:

$$EURT_i^k = \sum_{j=1}^n p_j m_{ij}^k = \underbrace{\sum_{j=1}^n p_j c_{ij}}_{\text{EUT Component}} + \underbrace{\sum_{j=1}^n p_j R(c_{ij} - c_{kj})}_{\text{Switch Component}} \tag{4.I.2}$$

The following assumptions are required for internal consistency, and for reconciling empirical reality with our theoretical specifications.

Assumption 4.I.1 (Exchangeability)**.** *Let Q be a probability measure which characterizes rank dependent utility and P be a probability measure that characterize other generalized expected utility models. We assume that there exist a probability weighting function $w(Q)$ which generates Quiggin (1982) type decision weights $\Pi = \{\pi_j\}_{1 \leq j \leq n}$ such that Π coincides with P.* □

This exchangeability assumption is technical. It allows us to take expectations with respect to the same probability measure under rank dependent utility and other generalized expected utility theory specifications. Refer to Berger (1985, p. 105) for further details on exchangeability concept.

In Theorem 4.4.1 we proved that

$$RDEU^{(k)}(\boldsymbol{x},\boldsymbol{p}) = \sum_{j=1}^{n} \pi_j^{(k)} U(x_j) = \eta_0 \underbrace{\sum_{j=1}^{n} p_j U(x_j)}_{\text{EUT Component}} + \eta_1 \underbrace{\sum_{j=1}^{n} \varphi^{(k)}(\mathfrak{z}_j) U(x_j)}_{\text{Harmonic Component}}, \qquad \sum_j \pi_j^{(k)} = 1 \quad (4.I.3)$$

where k controls the period of the harmonic component of the decision weight(s). The following lemma states a known relationship between harmonic analysis and *binary* (e.g. rejoice-regret) switching functions. It is the essential ingredient that ties the addends in the decomposed generalized EUT functions above together.

Lemma 4.I.2 (Switching Functions and Harmonic Analysis). *Let $Q_k(x) = \exp(2\pi ikx/2) = (-1)^{kx} = \pm 1$ for $k = 0, 1$ and integer values of x, be a Fourier kernel or harmonic basis function. For binary choice, the abstract Fourier expansion of the switching function $R(\cdot)$ is given by*

$$R(x) = 2^{-n} \sum_{j=1}^{n} R_j^{\star} Q_j(x) \qquad (4.I.4)$$

where $R_j^{\star} = \sum_x R(x) Q_j(x)$ is an abstract Fourier coefficient.

Proof. See Lechner (1971, pp. 130-131) □

According to Lemma 4.I.2, and exchangeability Assumption 4.I.1, the harmonic component in (4.I.3) is related to the switching component in (4.I.2) up to a constant. Thus, we prove the following:

Proposition 4.I.3 (HPWF and Rejoice-Regret). *Harmonic addend for HPWF decomposition of RDEU is functionally equivalent to weak switching function addend for RT decomposition.* □

The question arises as to whether or not any of the representations above are consistent with a "core theory" plus stochastic error. Cf. Loomes and Sugden (1995); Hey (2005). For example, (4.I.2) and (4.I.3) are examples of "generalizations of the expected utility preference functional" tested in Hey and Orme (1994). They find that there is not much statistical difference between the performance of those models and classic EUT in economic experiments. Thus, we embark on a theoretical explanation of their results. That is, we claim that if EUT is a core theory in all of the above, i.e., RDEU and RT, then the switching addend component for RT decomposition, and the harmonic component addend for RDEU decomposition, are functionally equivalent to stochastic error terms. To support this claim we introduce the following

Definition 4.I.1 (Radamacher function). (Katznelson, 2004, p. 276). Radamacher functions $\boldsymbol{r}_n$ are a sequence of independent random variables taking the values 1 and -1 with probability 1/2 for each. □

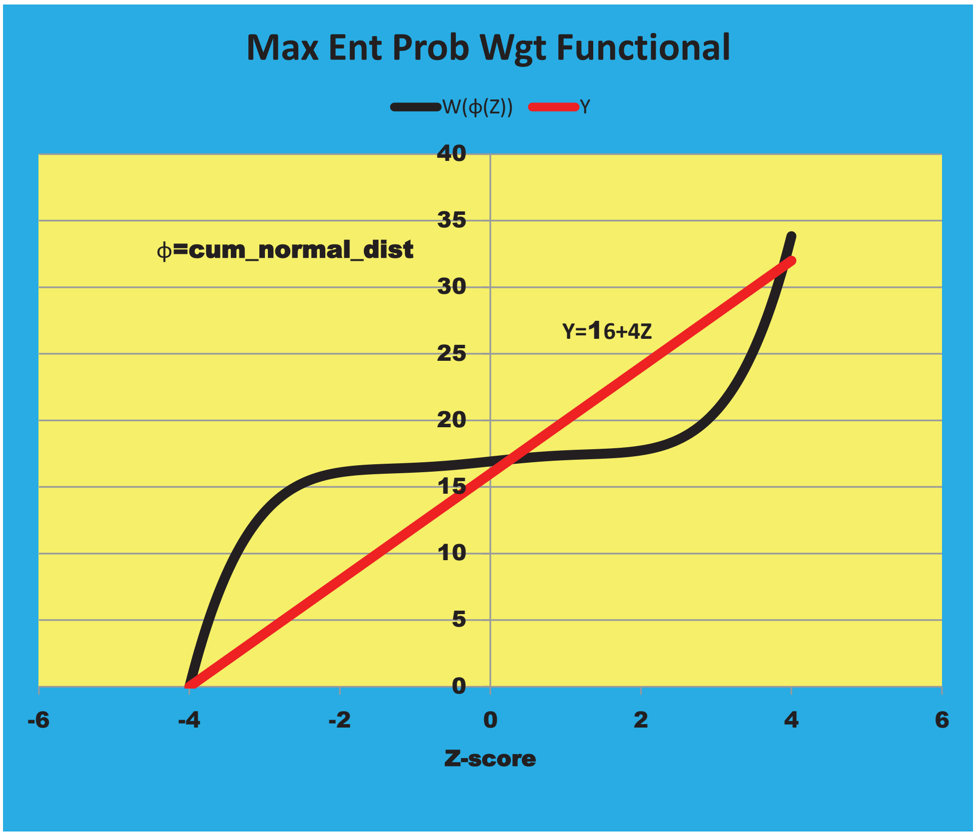
By definition, a Radamacher function (1) is a binary switching function so it satisfies Lemma 4.I.2; (2) has a harmonic representation; and (3) is a random variable, i.e., it can be treated as a stochastic error. Refere to De Guzmán (1981, pp. 23-24) and Folland (1992, p. 201) for technical details. For the sake of exposition one can assume that the stochastic error term in say, Hey (2005), has a Radamacher function representation. In which case the next proposition follows from Proposition 4.I.3:

Proposition 4.I.4 (EUT core plus stochastic error)**.** *The harmonic addend of HPWF decomposition of RDEU, and the switching function addend for RT decomposition, satisfy a core expected utility theory (EUT) plus stochastic error decomposition.* □

This proposition shows that the HPWF produces a theoretical explanation for why Hey and Orme (1994, p. 1301) failed to find statistically significant difference between core EUT and generalized EUT models in their study. According to our HPWF analysis, the difference between core EUT and generalized EUT models is functionally equivalent to stochastic error. A result verified empirically by the Hey-Orme experiments. The HPWF is axiomatized in Charles-Cadogan (2015a) where the generalized EUT decomposition result is further illustrated with prospect reference theory (Viscusi, 1989), disappointment aversion (Gul, 1991), reference dependent preferences (Kőszegi and Rabin, 2006) and weighted expected utility (Chew, 1983; Chew and Waller, 1986). Cox et al. (2013) also decomposed generalized EUT in the context of stake and probability calibration exercises to examine Rabin (2000) calibration critique of EUT. Fehr-Duda and Epper (2012) decomposed disappointment aversion and prospect reference theory. However, none of those authors considered probability harmonics.

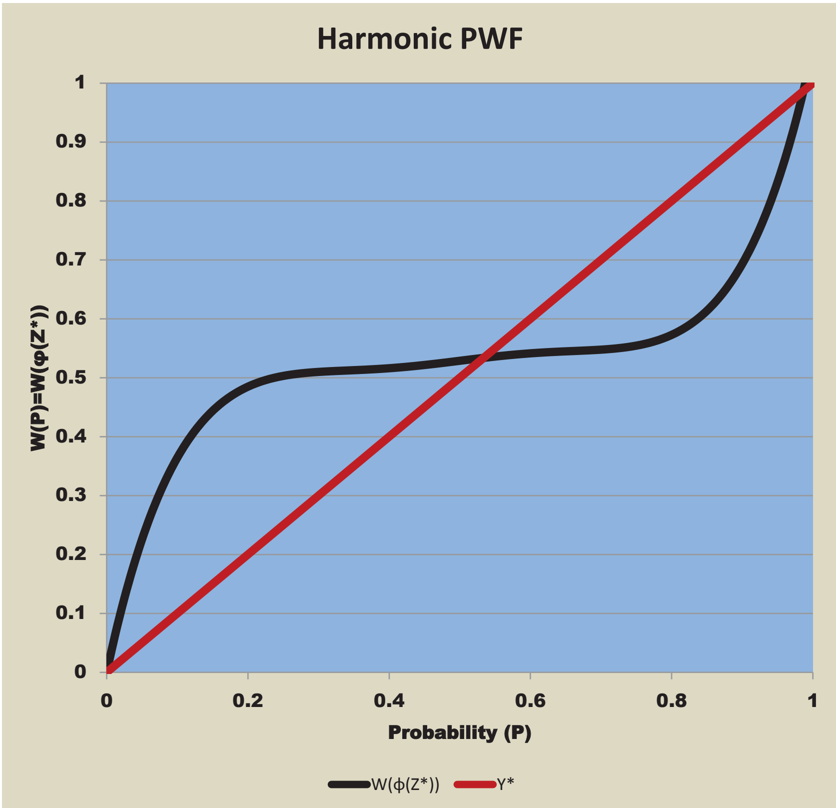
4.J APPENDIX OF PLOTS AND SIMULATED HPWF DATA

Figure 4.4: Maximum Entropy PW-Functional on Z-transform space



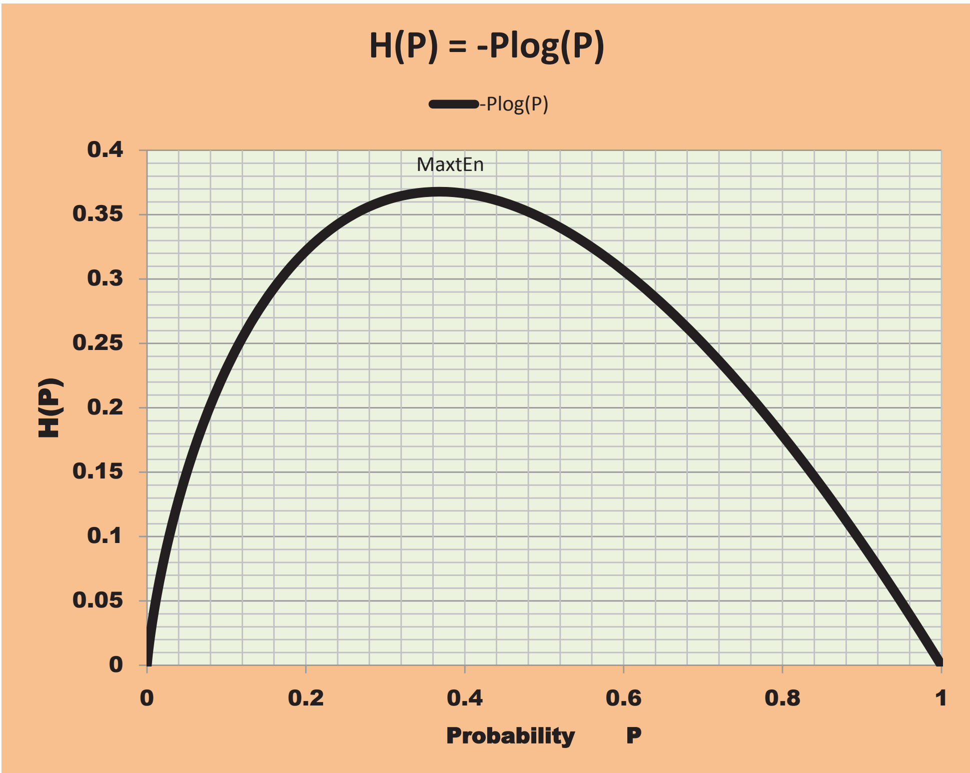
Plot of $w(\phi(\mathfrak{z}))$ *vs.* $\mathfrak{z}$ with respect to Z-transform of outcome space X. Negative values of Z corresponds to points below the mean 0. Those points are over weighted. Values above the mean are under weighted. Likelihood insensitivity is depicted by the relatively flat portion of the plot.

Figure 4.5: Affine Transform of Max Ent PW-Functional



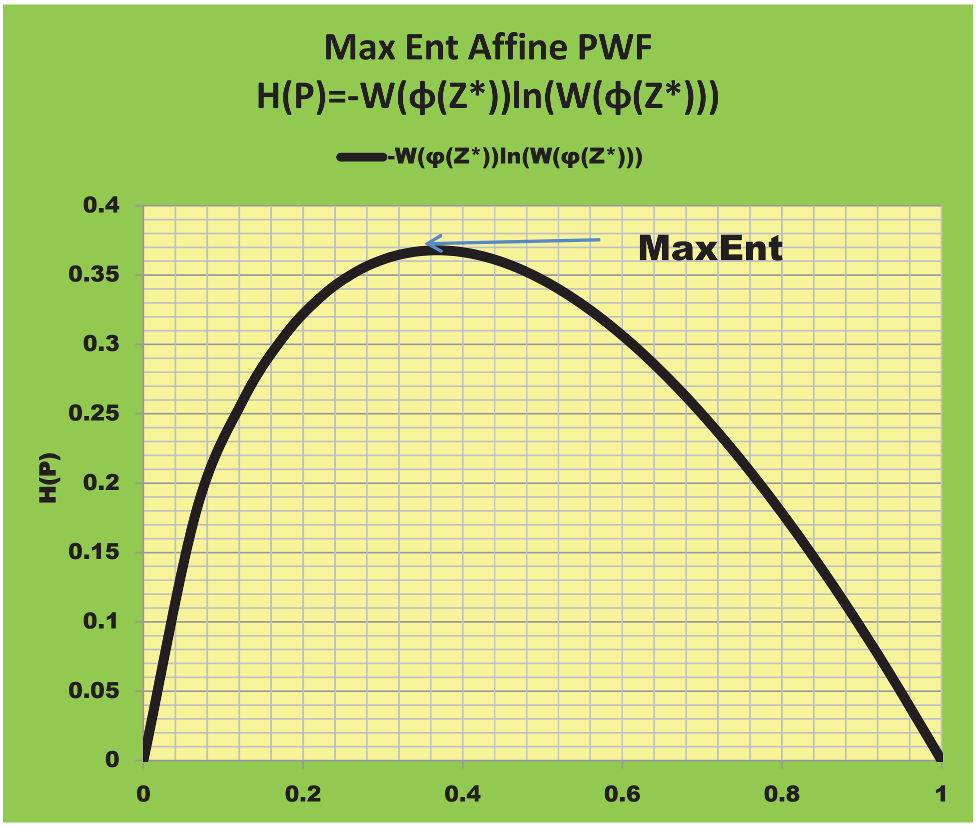
This is a scale and affine transformation of the plot in Figure 4.4. The horizontal axis is a pseudo-probability measure from that transformation.

Figure 4.6: Entropy Distribution for Z-transform of Prize Distribution



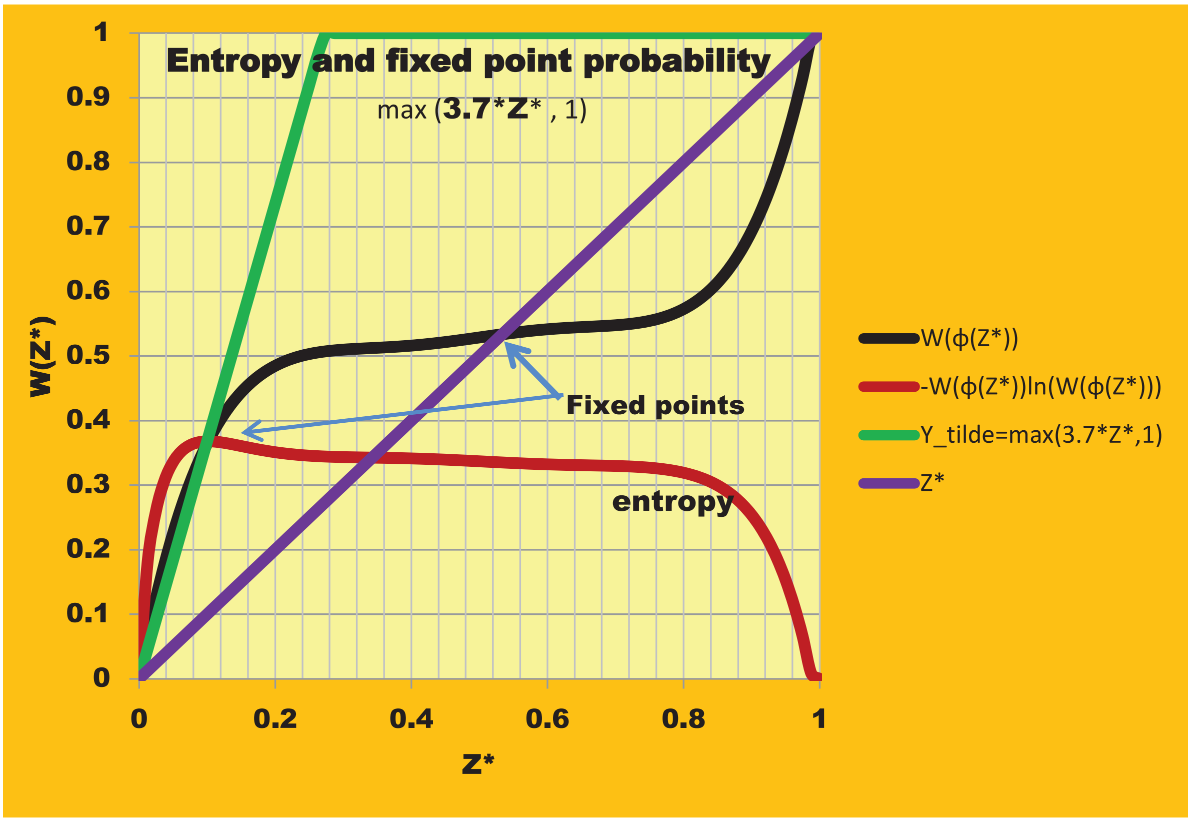
This is a plot of the maximum entropy associated with the probability weighting functional in Figure 4.4. Maximum entropy occurs at the point e^{-1}. This corresponds to the fixed point probability $p^{\star} = e^{-1}$.

Figure 4.7: Entropy Distribution for Affine Transformation of Z-transform of Prize Distribution



This is a plot of the maximum entropy associated with the probability weighting functional in Figure 4.5. Maximum entropy occurs at the point e^{-1}. This corresponds to the fixed point probability $p^{\star} = e^{-1}$.

Figure 4.8: Entropy and fixed point probability relationship



This is a cross plot of the maximum entropy associated with the probability transformation in (4.5.4) and (4.5.5). Maximum entropy occurs at the point e^{-1}. This corresponds to the fixed point probability $p^{\star} = e^{-1}$.

Table 4.1: Table of Simulated Vales and Estimates

Z	W(ϕ(Z))	Y	Z*	W(ϕ(Z*))	Y*	P	-Plog(P)	W(ϕ(Z*))	-W(ϕ(Z*))ln(W(ϕ(Z*)))
-4	4.75069E-05	0	0	1.48459E-06	0	3.16712E-05	0.000328117	1.48459E-06	1.99237E-05
-3.9	2.117551584	0.4	0.0125	0.066173487	0.0125	4.80963E-05	0.000478188	0.066173487	0.179692476
-3.8	4.013899784	0.8	0.025	0.125434368	0.025	7.2348E-05	0.000689768	0.125434368	0.260398314
-3.7	5.706441185	1.2	0.0375	0.178326287	0.0375	0.0001078	0.000984776	0.178326287	0.307459544
-3.6	7.211617197	1.6	0.05	0.225363037	0.05	0.000159109	0.001391552	0.225363037	0.335800544
-3.5	8.54498613	2	0.0625	0.267030817	0.0625	0.000232629	0.00194619	0.267030817	0.352585143
-3.4	9.721247894	2.4	0.075	0.303788997	0.075	0.000336929	0.002693964	0.303788997	0.361940866
-3.3	10.75426889	2.8	0.0875	0.336070903	0.0875	0.000483424	0.003690758	0.336070903	0.366462843
-3.2	11.6571071	3.2	0.1	0.364284597	0.1	0.000687138	0.005004409	0.364284597	0.36786182
-3.1	12.44203736	3.6	0.1125	0.388813667	0.1125	0.000967603	0.006715832	0.388813667	0.367294796
-3	13.12057683	4	0.125	0.410018026	0.125	0.001349898	0.008919757	0.410018026	0.365553274
-2.9	13.70351075	4.4	0.1375	0.428234711	0.1375	0.001865813	0.011724879	0.428234711	0.36317894
-2.8	14.20091822	4.8	0.15	0.443778694	0.15	0.00255513	0.015253239	0.443778694	0.360538804
-2.7	14.6221982	5.2	0.1625	0.456943694	0.1625	0.003466974	0.01963858	0.456943694	0.357876064
-2.6	14.97609552	5.6	0.175	0.468002985	0.175	0.004661188	0.025023518	0.468002985	0.35534559
-2.5	15.27072677	6	0.1875	0.477210211	0.1875	0.006209665	0.031555335	0.477210211	0.353039251
-2.4	15.51360606	6.4	0.2	0.484800189	0.2	0.008197536	0.03938032	0.484800189	0.351004283
-2.3	15.71167037	6.8	0.2125	0.490989699	0.2125	0.01072411	0.048636636	0.490989699	0.349256749
-2.2	15.87130432	7.2	0.225	0.49597826	0.225	0.013903448	0.059445837	0.49597826	0.347791455
-2.1	15.99836419	7.6	0.2375	0.499948881	0.2375	0.017864421	0.071903296	0.499948881	0.346589274
-2	16.09820102	8	0.25	0.503068782	0.25	0.022750132	0.086067943	0.503068782	0.345622528
-1.9	16.1756825	8.4	0.2625	0.505490078	0.2625	0.02871656	0.101951866	0.505490078	0.344858913
-1.8	16.23521352	8.8	0.275	0.507350423	0.275	0.035930319	0.119510486	0.507350423	0.344264326
-1.7	16.28075537	9.2	0.2875	0.508773605	0.2875	0.044565463	0.138634068	0.508773605	0.343804855
-1.6	16.31584332	9.6	0.3	0.509870104	0.3	0.054799292	0.159141418	0.509870104	0.343448137
-1.5	16.34360287	10	0.3125	0.51073759	0.3125	0.066807201	0.180776572	0.51073759	0.343164251
-1.4	16.36676466	10.4	0.325	0.511461396	0.325	0.080756659	0.203209181	0.511461396	0.342926258
-1.3	16.38767829	10.8	0.3375	0.512114946	0.3375	0.096800485	0.226039129	0.512114946	0.342710484
-1.2	16.40832551	11.2	0.35	0.512760172	0.35	0.11506967	0.248805655	0.512760172	0.342496641
-1.1	16.43033319	11.6	0.3625	0.513447912	0.3625	0.135666061	0.27100094	0.513447912	0.342267814
-1	16.45498658	12	0.375	0.514218331	0.375	0.158655254	0.292087757	0.514218331	0.342010384
-0.9	16.48324354	12.4	0.3875	0.515101361	0.3875	0.184060125	0.311520438	0.515101361	0.341713906
-0.8	16.51575033	12.8	0.4	0.516117198	0.4	0.211855399	0.32876808	0.516117198	0.341370966
-0.7	16.5528596	13.2	0.4125	0.517276862	0.4125	0.241963652	0.343338622	0.517276862	0.340977026
-0.6	16.59465094	13.6	0.425	0.518582842	0.425	0.274253118	0.354802304	0.518582842	0.340530271
-0.5	16.64095469	14	0.4375	0.520029834	0.4375	0.308537539	0.362812921	0.520029834	0.340031437
-0.4	16.69137895	14.4	0.45	0.521605592	0.45	0.344578258	0.367125409	0.521605592	0.339483634
-0.3	16.74534007	14.8	0.4625	0.523291877	0.4625	0.382088578	0.367608498	0.523291877	0.338892134
-0.2	16.80209641	15.2	0.475	0.525065513	0.475	0.420740291	0.364251498	0.525065513	0.33826413
-0.1	16.86078497	15.6	0.4875	0.52689953	0.4875	0.460172163	0.357164738	0.52689953	0.337608447
0	16.92046043	16	0.5	0.528764389	0.5	0.5	0.34657359	0.528764389	0.336935192
0.1	16.9801359	16.4	0.5125	0.530629247	0.5125	0.539827837	0.332806566	0.530629247	0.336255359
0.2	17.03882446	16.8	0.525	0.532463264	0.525	0.579259709	0.316278323	0.532463264	0.335580378
0.3	17.0955808	17.2	0.5375	0.5342369	0.5375	0.617911422	0.297468838	0.5342369	0.33492161
0.4	17.14954192	17.6	0.55	0.535923185	0.55	0.655421742	0.276900198	0.535923185	0.334289825
0.5	17.19996617	18	0.5625	0.537498943	0.5625	0.691462461	0.255112596	0.537498943	0.333694655

Table 4.2: Table of Simulated Vales and Estimates (cont'd)

(continued)

Z	W(ϕ(Z))	Y	Z*	W(ϕ(Z*))	Y*	P	-Plog(P)	W(ϕ(Z*))	W(ϕ(Z*))ln(W(ϕ(Z*))
0.6	17.24626992	18.4	0.575	0.538945935	0.575	0.725746882	0.232641046	0.538945935	0.333144051
0.7	17.28806127	18.8	0.5875	0.540251915	0.5875	0.758036348	0.209994217	0.540251915	0.332643768
0.8	17.32517053	19.2	0.6	0.541411579	0.6	0.788144601	0.187636503	0.541411579	0.332196888
0.9	17.35767733	19.6	0.6125	0.542427417	0.6125	0.815939875	0.165974091	0.542427417	0.331803391
1	17.38593429	20	0.625	0.543310446	0.625	0.841344746	0.145345484	0.543310446	0.331459793
1.1	17.41058767	20.4	0.6375	0.544080865	0.6375	0.864333939	0.126016501	0.544080865	0.331158842
1.2	17.43259535	20.8	0.65	0.544768605	0.65	0.88493033	0.108179512	0.544768605	0.330889265
1.3	17.45324258	21.2	0.6625	0.545413831	0.6625	0.903199515	0.091956371	0.545413831	0.330635564
1.4	17.47415621	21.6	0.675	0.546067381	0.675	0.919243341	0.077404337	0.546067381	0.330377811
1.5	17.49731799	22	0.6875	0.546791187	0.6875	0.933192799	0.064524175	0.546791187	0.330091437
1.6	17.52507755	22.4	0.7	0.547658673	0.7	0.945200708	0.053269607	0.547658673	0.329746955
1.7	17.5601655	22.8	0.7125	0.548755172	0.7125	0.955434537	0.043557333	0.548755172	0.329309564
1.8	17.60570734	23.2	0.725	0.550178354	0.725	0.964069681	0.035276952	0.550178354	0.328738594
1.9	17.66523837	23.6	0.7375	0.552038699	0.7375	0.97128344	0.028300235	0.552038699	0.327986687
2	17.74271984	24	0.75	0.554459995	0.75	0.977249868	0.022489363	0.554459995	0.326998671
2.1	17.84255667	24.4	0.7625	0.557579896	0.7625	0.982135579	0.017703893	0.557579896	0.325710003
2.2	17.96961655	24.8	0.775	0.561550517	0.775	0.986096552	0.013806344	0.561550517	0.324044714
2.3	18.12925049	25.2	0.7875	0.566539078	0.7875	0.98927589	0.0106664	0.566539078	0.321912727
2.4	18.3273148	25.6	0.8	0.572728588	0.8	0.991802464	0.008163844	0.572728588	0.319206466
2.5	18.5701941	26	0.8125	0.580318566	0.8125	0.993790335	0.006190345	0.580318566	0.31579664
2.6	18.86482535	26.4	0.825	0.589525792	0.825	0.995338812	0.004650308	0.589525792	0.311527127
2.7	19.21872266	26.8	0.8375	0.600585083	0.8375	0.996533026	0.003460957	0.600585083	0.306208881
2.8	19.64000264	27.2	0.85	0.613750083	0.85	0.99744487	0.002551863	0.613750083	0.299612822
2.9	20.13741011	27.6	0.8625	0.629294066	0.8625	0.998134187	0.001864072	0.629294066	0.291461711
3	20.72034403	28	0.875	0.647510751	0.875	0.998650102	0.001348987	0.647510751	0.281421046
3.1	21.39888351	28.4	0.8875	0.66871511	0.8875	0.999032397	0.000967135	0.66871511	0.269089057
3.2	22.18381376	28.8	0.9	0.69324418	0.9	0.999312862	0.000686902	0.69324418	0.253985943
3.3	23.08665197	29.2	0.9125	0.721457874	0.9125	0.999516576	0.000483307	0.721457874	0.235542497
3.4	24.11967297	29.6	0.925	0.75373978	0.925	0.999663071	0.000336872	0.75373978	0.213088333
3.5	25.29593474	30	0.9375	0.79049796	0.9375	0.999767371	0.000232602	0.79049796	0.185839907
3.6	26.62930367	30.4	0.95	0.83216574	0.95	0.999840891	0.000159096	0.83216574	0.152888529
3.7	28.13447968	30.8	0.9625	0.87920249	0.9625	0.9998922	0.000107794	0.87920249	0.113188567
3.8	29.82702108	31.2	0.975	0.932094409	0.975	0.999927652	7.23454E-05	0.932094409	0.065545972
3.9	31.72336928	31.6	0.9875	0.99135529	0.9875	0.999951904	4.80952E-05	0.99135529	0.008607236
4	33.84087336	32	1	1.057527292	1	0.999968329	3.16707E-05	1.057527292	-0.059151139

Chapter 5

Prospect Theory's Cognitive Error About Bernoulli's Utility Function

5.1 Introduction

A recent survey by Barberis (2013, p. 173) describes Kahneman and Tversky (1979) original version of prospect theory (OPT), and its amendment, cumulative prospect theory (CPT) (Tversky and Kahneman, 1992) thusly. "Prospect theory is still widely viewed as the best available description of how people evaluate risk in experimental settings", while duly noting that "there are relatively few well-known and broadly accepted applications of prospect theory in economics". Prospect theory was proposed in response to purported anomalies from experiments in psychology and behavioral economics which led to revisions of Von Neumann and Morgenstern (1953) expected utility theory (EUT) model, and utility theory more generally.

In the context of applications, Blavatskyy (2005); Rieger and Wang (2006); Pfiffelmann (2011) reexamined the St. Petersburg Paradox (first featured in Bernoulli (1738) EUT setting) in a CPT framework. This paper goes a step further. It argues that Bernoulli (1738) *original utility function*, which "resolved" the St Petersburg Paradox and which falls under rubric of EUT, explicitly satisfies several characteristics of CPT's value function. In a recent paper, Charles-Cadogan (2016b) provides a more sophisticated analysis which shows how one can derive a loss aversion index in an EUT framework. We claim "cognitive error"[1] as the cause of some analysts' misperception of Bernoulli's utility function specification. We also provide evidence that under mild assumptions Bernoulli's utility function accommodates global loss aversion, and a Fisher z-transformation test for loss aversion index. Moreover, the index follows an α-stable law. These findings are important in their own right.

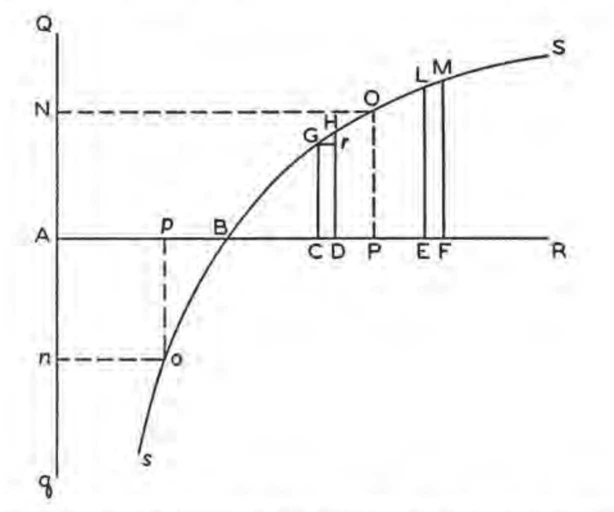
[1] Also known as a cognitive bias, a cognitive error is a pattern of deviation in judgment that occurs in particular situations, which may sometimes lead to perceptual distortion, inaccurate judgment, illogical interpretation, or what is broadly called irrationality. See e.g., http://en.wikipedia.org/wiki/Cognitive_bias.

This paper is motivated by Daniel Kahneman Nobel Prize lecture, a significant part of which is devoted to what he deemed "Bernoulli's error". He states in relevant part:

> Perception is *reference-dependent*: the perceived attributes of a focal stimulus reflect the contrast between that stimulus and a context of prior and concurrent stimuli.
>
> * * * * * * * * *
>
> [Amos Tversky and I] noted, however, that reference-dependence is incompatible with the standard interpretation of Expected Utility Theory, the prevailing theoretical model in this area. This deficiency can be traced to the brilliant essay that introduced the first version of expected utility theory (Bernoulli, 1738).
>
> One of Bernoulli's aims was to formalize the intuition that it makes sense for the poor to buy insurance and for the rich to sell it. He argued that the increment of utility associated with an increment of wealth is inversely proportional to initial wealth, and from this plausible psychological assumption he derived that *the utility function for wealth is logarithmic*. He then proposed that a sensible decision rule for choices that involve risk is to maximize the expected utility of wealth (the moral expectation). This proposition accomplished what Bernoulli had set out to do: it explained risk aversion, as well as the different risk attitudes of the rich and of the poor. The theory of expected utility that he introduced is still the dominant model of risky choice. The language of Bernoulli's essay is prescriptive it speaks of what is sensible or reasonable to do but the theory is also intended to describe the choices of reasonable men (Gigerenzer et al., 1989). As in most modern treatments of decision making, there is no acknowledgment of any tension between prescription and description in Bernoulli's essay. The idea that decision makers evaluate outcomes by the utility of final asset positions has been retained in economic analyses for almost 300 years. This is rather remarkable, because the idea is easily shown to be wrong; I call it *Bernoulli's error*.
>
> *Bernoulli's model is flawed because it is reference-independent: it assumes that the value that is assigned to a given state of wealth does not vary with the decision makers initial state of wealth*[Footnote in original][What varies with wealth in Bernoulli's theory is the response to a given change of wealth. This variation is represented by the curvature of the utility function for wealth. Such a function cannot be drawn if the utility of wealth is reference-dependent, because utility then depends not only on current wealth but also on the reference level of wealth.]. This assumption flies against a basic principle of perception, where the effective stimulus is not the new level of stimulation, but the difference between it and the existing adaptation level. *The analogy to perception suggests that the carriers of utility are likely to be gains and losses rather than states of wealth, and this suggestion is amply supported by the evidence of both experimental and observational studies of choice* (see Kahneman & Tversky, 2000). [Emphasis added]. (Kahneman, 2002, pp. 460-461).

This paper provides a critical review of the Bernoulli (1738) model, and compares it to the claims made against it in the Kahneman lecture above. In section 5.2 we compare the geometry of

Figure 5.1: Bernoulli Utility of Wealth Function



Reproduction of utility function sketched in Bernoulli (1738, p. 26).

Bernoulli's utility function to that of Kahneman-Tversky skew S-shape value function. For example, Bernoulli's sketch of his log concave utility function cuts the horizontal axis at an incipient wealth level B in Figure 5.1. He analyzed log ratios of wealth relative to the incipient wealth level AB, i.e., $\log\frac{AC}{AB}, \log\frac{AD}{AB}$ and so on. Thus, generating log growth in wealth relative to AB whereupon the log concave utility function cuts the horizontal at relative wealth level equal to $\frac{AB}{AB} = 1$. It is known that the logarithm of 1 is zero. So the log of points to the left of 1 is negative, and log of points to the right of 1 is positive. Thus, 1 (the relative wealth level point B) is a *de facto* reference point. So loss aversion is latent in Bernoulli's specification. Approximation of Bernoulli's specification also accommodates higher order risk attitudes that include a preference for skewness. In section 5.3 we show how Bernoulli's specification supports a closed form global loss aversion index, that the index is α-stable, and we characterize its relation to Fisher's z-transformation test. We conclude in section 5.4.

5.2 Prospect theory value function vs Bernoulli utility function

In this section we emphasize the geometric properties of Bernoulli (1738) utility function, identify its *de facto* reference point, and contrasts it to the qualitative and geometric properties of Kahneman and Tversky (1979) value function.

5.2.1 The *de facto* reference point in Bernoulli's log concave utility function

We begin this subsection with Bernoulli (1738, pg. 26) description of the *geometry* of his utility of wealth function reproduced in Figure 5.1:

> "[L]et *AB* represent the quantity of goods initially possessed. Then after extending AB, a curve BGLS must be constructed, whose ordinates CG, DH, EL, FM, etc., designate utilities corresponding to the abscissas BC, BD, BE, BF, etc., designating *gains in wealth*. Further, let m, n, p, q, etc., be the numbers which indicate the number of ways in which *gains in wealth* BC, BD, BE, BF, etc., can occur". [Emphasis added]

The point *B* in Figure 5.1 is *Daniel Bernoulli's de facto reference point* against which other wealth levels are compared. Furthermore, Bernoulli (1738, pg. 29) states:

> First, it appears that in many games, even those that are absolutely fair, both of the players may expect to suffer a loss; indeed this is Nature's admonition to avoid the dice altogether ... This follows from the *concavity of curve sBS* to *BR*. For in making the stake, *Bp*, equal to the *expected gain*, *BP*, it is clear that the *disutility po* which results from a *loss will always exceed the expected gain in utility, PO*. [Emphasis added]

It is indisputable that the italicized text in Bernoulli's analysis above involves gains and losses relative to the *de facto* reference point *B*. Furthermore, he compared "utility" of expected gain *BP* to the "disutility" of a loss of an equal amount *Bp*, and plainly concludes that "loss will always exceed the expected gain in utility". In other words, "losses loom larger than gains" (Kahneman and Tversky, 1979, p. 279) in Bernoulli's utility function specification. Nonetheless, Kahneman and Tversky (1979, pg. 276) states:

> "[Markowitz (1952)] was the first to propose that utility be defined on gains and losses rather than on final asset positions, an assumption which has been implicitly accepted in most experimental measurements of utility".

5.2.2 Bernoulli's forgotten numeraire wealth level and risk-return tradeoff

Given Bernoulli's log-concave specification, the reference wealth level can only cut the horizontal axis at $x = 1$. In other words, Bernoulli normalized wealth levels so that a given wealth level W_x, say, is *numeraire*–the reference wealth. In which case, any other wealth level, say W_z, is represented by $\frac{W_z}{W_x}$. Thus, the points in his graph are *changes in wealth* relative to the *numeraire*. This fact may have been obscured by his use of "analytic geometry" as opposed to "algebraic geometry" to represent the geometric mean. Stigler (1950, pg. 374) also analyzed Bernoulli's utility function by introducing the notion of a "subsistence level at c" where

$$U(c) = k\ln(c) + a = 0 \Rightarrow a = -k\ln(c) \Rightarrow U(x) = k\ln(\frac{x}{c}) \tag{5.2.1}$$

Using Stigler's interpretation, $U(c) = 0$ at precisely where "subsistence wealth level" $x = c$ and *relative wealth* $\frac{x}{c} = 1$. Ironically, there is evidence that those at "subsistence levels" of income are more prone to purchasing lottery tickets (Friedman and Savage, 1948; Light, 1977; Beckert and Lutter, 2012; Scott and Barr, 2012). More on point, Bernoulli (1738, pg. 28) writes:

$$b\log\frac{AP}{AB} = \frac{mb\log\frac{AC}{AB} + nb\log\frac{AD}{AB} + pb\log\frac{AE}{AB} + qb\log\frac{AF}{AB} + \dots}{m+n+p+q+\dots} \tag{5.2.2}$$

That equation can be rewritten as

$$\frac{AP}{AB} = \left[\left(\frac{AC}{AB}\right)^m \left(\frac{AD}{AB}\right)^n \left(\frac{AE}{AB}\right)^p \left(\frac{AF}{AB}\right)^q \dots\right]^{\frac{1}{m+n+p+q+\dots}} \tag{5.2.3}$$

which is a *weighted geometric mean* relative to the *reference wealth level AB* (see e.g., Stearns, 2000, p. 221)("Bernoulli taught us how to measure risk with the geometric mean"). Since $AP < AB$ in Figure 5.1, $\frac{AP}{AB} < 1$ if and only if at least one or all of the fractions on the right hand side in (5.2.3) is much smaller than 1 ($\ll 1$). Let $\{W_P, W_B, W_C, W_D, W_E, W_F, \dots\}$ be a *ranking of nominal wealth* where the subscripts coincide with the corresponding letters in Bernoulli's model. So that we have the strict partial preference order $W_P \prec W_B \prec W_C \prec \dots$. Choose W_B as *numeraire* so that *relative wealth*

$$\begin{aligned}\frac{W_P}{W_B} \sim AP, \frac{W_B}{W_B} = 1 \sim AB, \frac{W_C}{W_B} \sim AC, \frac{W_D}{W_B} \sim AD,\\ \frac{W_E}{W_B} \sim AE, \frac{W_F}{W_B} \sim AF, \dots\end{aligned} \tag{5.2.4}$$

This is an implicit assumption in Bernoulii's model. Let $N = m+n+p+q+\ldots$ and $\frac{W_C}{W_B} = (1+r_C)$ and so on. We rewrite (5.2.3) as

$$1+r_P = \left[(1+r_C)^m(1+r_D)^n(1+r_D)^p(1+r_E)^p(1+r_F)^q\ldots\right]^{\frac{1}{N}} \tag{5.2.5}$$

$$\Rightarrow r_P = \exp\left(\frac{1}{N}\sum_{\substack{j\in\{C,D,E,F,\ldots\}\\ k\in\{m,n,p,q,\ldots\}}} \ln(1+r_j)^k\right) - 1 \tag{5.2.6}$$

$$\approx r_N = \frac{1}{N}\sum_{\substack{j\in\{C,D,E,F,\ldots\}\\ k\in\{m,n,p,q,\ldots\}}} k(r_j - \frac{1}{2}r_j^2) \Rightarrow r_P \approx \lim_{N\to\infty} r_N = r^\star == \mu_r - \frac{\sigma_r^2}{2} < \infty \tag{5.2.7}$$

up to a second order approximation. Stearns (2000, p. 224) provides applications of Bernoulli's utility model to genetics and posits that $r^\star = \mu_r - \frac{\sigma_r^2}{2\mu_r}$ is the most widely used approximation of the geometric mean in (5.2.7) where μ_r and σ_r^2 are the mean and variance of the distribution of $r_j,\ j\in\{C,D,E,F,\ldots\}$.

Lemma 5.2.1 (Bernoulli's reference dependent change in wealth)**.**

Bernoulli normalized his wealth levels with a numeraire so that reference wealth $x = 1$ where his log-concave utility cuts the horizontal at the de facto numeraire. All other points on the axis represent percent changes in wealth relative to $x = 1$. So that points to the left of $x = 1$ correspond to percent loss in wealth and points to the right correspond to percent gain in wealth. □

Lemma 5.2.2 (Bernoulli's risk-return tradeoff)**.**

The numeraire wealth level in Bernoulli utility function induces a geometric mean $r^\star$ that is approximated by risk-return tradeoffs in (5.2.7). □

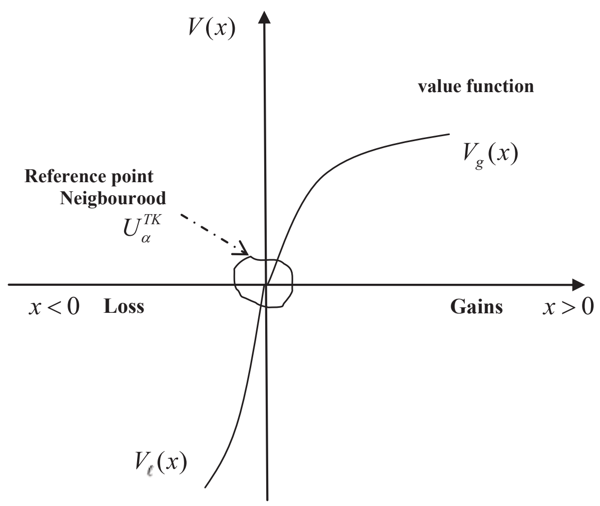
Remark 5.2.1. This seemingly over looked result has implications for asset pricing models as well (Campbell et al., 1997, §1.4).

5.2.3 Bernoulli utility function vs. Kahneman-Tversky latent skew S-shaped utility function with loss aversion

Next we examine the shape of Kahneman and Tversky (1979, pg. 279) value function sketched in Figure 5.2, and the specification in Tversky and Kahneman (1992) and its implication for the ubiquitous loss aversion index. According to Kahneman and Tversky (1979, pg. 279) (KT79)

In summary, we have proposed that the value function is *(i) defined on deviations from the reference point; (ii) generally concave for gains and commonly convex for losses; (iii) steeper for losses than for gains*. A value function which satisfies these properties is displayed in Figure 3. Note that the proposed S-shaped value function is steepest at the reference point, in marked contrast to the utility function postulated by Markowitz which is relatively shallow in that region. [Emphasis added]

Figure 5.2: Prospect theory's value function



Sketch of Prospect Theory's bifurcated utility function with kink at reference point at the origin, and value functions V_g concave over gain domain and V_ℓ convex and steeper over loss domain.

We now show that Bernoulli utility function satisfies $(i)-(iii)$ in the KT79 quote above.

Case (i): Deviations from the reference point

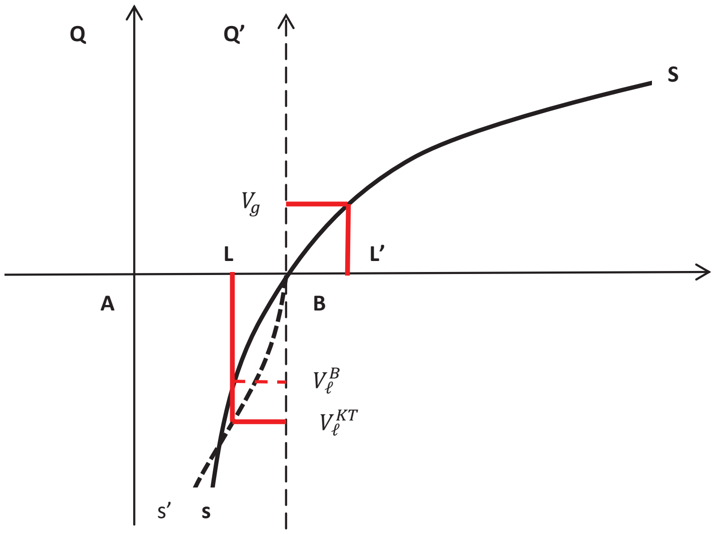
From the outset, we note that Figure 5.3 depicts Bernoulli's utility function with reference point B. The Bernoulli (1738, pg. 29) quote above satisfies KT79 condition (i). Bernoulli conducted his analysis over positive values *AB* and beyond in Figure 5.3. However, if we shift (i.e. translate) Bernoulli's axis *Qq* in Figure 5.1 so that it passes through *B*, then his loss *Bq* is negative while his gain *BQ* is positive. Ironically, this change of axes involves a translation[2] of relative wealth by -1 so that the reference point *B* is now 0–the same as Kahneman and Tversky's reference point–as shown in Figure 5.3. To see this, note that according to Stigler (1950, pg. 374) Bernoulli's

[2]Eeckhoudt et al. (1995, Fig. 1, p. 335) shows that Bernoulli type utility is a concave function of translation in wealth.

specification for *relative wealth* is of type

$$u(x) = \begin{cases} \ln(1+x) & \text{for gain} \\ \ln(1-x) & \text{for loss} \end{cases} \tag{5.2.8}$$

Figure 5.3: Bernoulli Utility Function With Reference Point



Sketch of Bernoulli's original utility function Ss for $\ln(1+x)$ for change x in reference wealth at reference point B and sub-utility V_g over gain domain. BQ' is a translation of axis so that B coincides with Prospect Theory's reference point at the origin. Now BL is a loss relative to B and BL' is a gain. V_ℓ^B is Bernoulli's negative sub-utility function (steeper than V_g) over loss domain represented by the arc Bs or $\ln(1-x)$ without kink at the origin. V_ℓ^{KT} is Prospect Theory's negative sub-utility (value function) over loss domain represented by the dashed arc Bs' which is Bernoulli's arc Bs induced by multiplying Bernoulli's B with a loss aversion index to induce a kink at B and convexity over loss domain.

In the sequel x *is change in relative wealth*. Replacing x by $X = x-1$ generates $u(X) = \ln(X)$ for gains. However, that change of axis induces undefined concepts like the logarithm of negative terms when $X < 0$ for losses. So in Bernoulli's model, a "pure utility" of loss is undefined under the translation scheme and the *reflection effect* is unobservable *under that transformation* (Kahneman and Tversky, 1979). Tversky and Kahneman (1992) "resolved" this problem by writing $-X$ when $X < 0$ and they "hardcoded" a loss aversion index λ to account for the skew. However, the *utility of loss can be recovered in Bernoulli's specification* without resort to "hardcoding" as evidenced by the following approximations when *return on [reference] wealth* $x > 0$:

$$u_g(x) = \ln(1+x) \approx x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4}, \qquad \text{for gains} \tag{5.2.9}$$

$$u_\ell(x) = \ln(1-x) \approx -x + \frac{x^2}{2} - \frac{x^3}{3} + \frac{x^4}{4}, \qquad \text{for loss} \tag{5.2.10}$$

In each case, $u_g(0) = u_\ell(0) = 0$ when the reference point is zero. This is semantics because under the original specification $\ln(1+0) = 0$. However, we introduce (5.2.9) and (5.2.10) for comparison with Tversky and Kahneman (1992) model of loss aversion.[3] Note that the curve is *smooth* at the reference point 0–*there is no kink*–contrary to Kahneman and Tversky's specification, infra. The translation does not affect the [absolute] magnitude (x) of Bernoulli's relative gains or losses in wealth. But it does change the orientation of his curve from the solid concave portion *Bs* to the dotted *convex portion* Bs' as shown in Figure 5.3 to accommodate the fact that now points to the left of *B* are negative, and so the curve is in the negative quadrant with a longer tail because of the asymmetric response to gains and losses. This point is explained in detail later in the paper. For the purpose of illustration, we provide a numerical example in Table 5.1 which brings us to Case (ii).

Case(ii): Generally concave for gains and commonly convex for losses

The asymptotic expansion of Bernoulli's utility function in (5.2.9) and (5.2.10) to higher moments introduces risk attitude concepts like prudence, temperance, etc (e.g., Eeckhoudt et al., 1995). For example, $u_g''''(x) < 0$ implies that a decision maker (DM) with Bernoulli utility over gains is temperate (Eeckhoudt et al., 1995, Cor. 2). Applying an expectation operator E to (5.2.9) and (5.2.10) implies the following higher risk attitudes (Noussair et al., 2014, p. 326, fn.1). For example, the higher order terms are analogized to measures of dispersion and $\oplus$ is treated as an abstract conjoint operation.

[3] Kelly (1956) used a product specification $V_N = u_g^W u_\ell^L V_0$ involving (5.2.9) and (5.2.10) which relates V_N to Shannon's entropy, where V_N is a gambler's capital after N bets, V_0 is initial capital and W and L signify numbers of wins and losses.

Table 5.1: Hierarchical risk attitudes:
Prospect value vs. Bernoulli utility

x	$Y_1(x)$	$Y_2(x)$	$Y_3(x)$	$Y_4(x)$	$Y_{PT}(x)^a$
-1.0	-1.0	-1.500	-1.833	-2.083	-2.250
-0.9	-0.9	-1.305	-1.548	-1.712	-2.051
-0.8	-0.8	-1.120	-1.291	-1.393	-1.849
-0.7	-0.7	-0.945	-1.059	-1.119	-1.644
-0.6	-0.6	-0.780	-0.852	-0.884	-1.435
-0.5	-0.5	-0.625	-0.667	-0.682	-1.223
-0.4	-0.4	-0.480	-0.501	-0.508	-1.005
-0.3	-0.3	-0.345	-0.354	-0.356	-0.780
-0.2	-0.2	-0.220	-0.223	-0.223	-0.546
-0.1	-0.1	-0.105	-0.105	-0.105	-0.297
0	0	0	0	0	0
0.1	0.1	0.095	0.095	0.095	0.132
0.2	0.2	0.180	0.183	0.183	0.243
0.3	0.3	0.255	0.264	0.266	0.347
0.4	0.4	0.320	0.341	0.348	0.446
0.5	0.5	0.375	0.417	0.432	0.543
0.6	0.6	0.420	0.492	0.524	0.638
0.7	0.7	0.455	0.569	0.629	0.731
0.8	0.8	0.480	0.651	0.753	0.822
0.9	0.9	0.495	0.738	0.902	0.911
1.0	1.0	0.500	0.833	1.083	1.000

[a] $Y_{PY} = x^{0.88}$, $x > 0$; $Y_{PY} = -2.25(-x)^{0.88}$, $x < 0$.

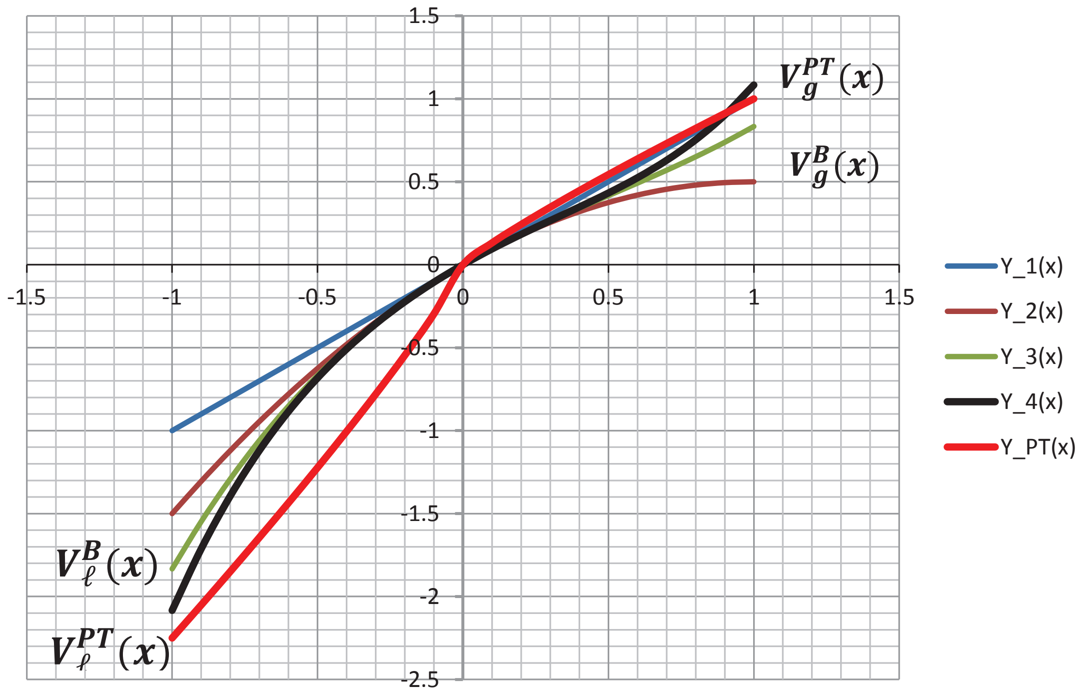
Higher order risk attitudes implied by approximation of Bernoulii's sub-utility functions $u_g(x) = \ln(1+x) \approx x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4}$ over gain domain, and $u_\ell(x) = \ln(1-x) \approx -x + \frac{x^2}{2} - \frac{x^3}{3} + \frac{x^4}{4}$ over loss domain, for *return on wealth* x relative to a reference wealth level B in Figure 5.3.

Hierarchical higher order risk attitudes

- $E[Y_1(x)] \equiv mean \Longrightarrow$ "risk neutrality";
- $E[Y_2(x)] \equiv mean \oplus variance; \Longrightarrow$ "risk aversion";
- $E[Y_3(x)] \equiv mean \oplus variance \oplus skewness \Longrightarrow$ "prudence";
- $E[Y_4(x)] \equiv mean \oplus variance \oplus skewness \oplus kurtosis \Longrightarrow$ "temperance"

Table 5.1 contains the distribution of those values for equally spaced values of x. The underlying premise is that the range of wealth is normalized by its maximum value. A plot of Bernoulli's value

Figure 5.4: Bernoulli vs. Kahnemen-Tversky Value Function



Plot of higher order risk attitudes implied by approximations for Bernoulli's sub-utility functions V_g^B over gain domain and V_ℓ^B over loss domain relative to reference point 0 for return on wealth x. Compared to Prospect Theory's value function V_g^{PT} over gain domain and V_ℓ^{PT} over loss domain. Y_1 is risk neutral. V_ℓ^B is steeper than Y_1 in loss domain even though V_ℓ^{PT} is uniformly steeper than V_ℓ^B. In contrast, V_g^B is less steeper than Y_1 and V_g^{PT} over gain domain.

function with hierarchical higher order risk attitudes, and the Kahneman-Tversky value function (in red) is superimposed in Figure 5.4.

Prospect theory's artificial kink at the origin is due to experimenter bias

Kahneman-Tversky value function is a power function x^α which is concave over gains and convex over losses. So Figure 5.4 is a *qualitative* comparison. Each of the curves in Table 5.1 represent hierarchical risk attitudes, and they are subsets of (5.2.9) and (5.2.10). The "loss tail" for Bernoulli's construct is depicted by $V_\ell^B(x)$ in Figure 5.4. Whereas the value function over losses, $V_\ell^{PT}(x)$, is obtained by pre-multiplying the loss component by $\lambda = 2.25$–the median value of the loss aversion index reported in Tversky and Kahneman (1992). That introduces an *artificial kink* at the origin where there is no kink in the Bernoulli construct. Undeniably, for $x < 0$ we have $E[x] > V_\ell^B(E[x])$.

Thus, a DM would prefer to take risk over losses rather than receive the sure payment (Machina, 1987b, p. 538). This implies that V_ℓ^B should be convex over losses despite its appearance in the sketch. This is depicted by Bs' in Figure 5.3. Because Kahneman-Tversky hard coded -2.25 or $-\lambda$ if you will in their (loss) tail specification their model is confounded.

Preference for skewness

In gain domains Bernoulli preferences are strictly risk averse. The schema for that is depicted in Figure 5.5. However, the exception is Y_4 (with kurtosis) which exceeds the expected value near the tail, i.e., for $0.9 < x \leq 1$ as indicated in the last row of Table 5.1 and the black curve in Figure 5.4. This suggests the existence of a *convex* segment of the curve consistent with gambling over gain domain depicted in Figure 5.6.[4] So an otherwise risk averse DM faced with the prospect of doubling her wealth[5] is willing to take risks if she has kurtosis preferences. Interestingly, $E[Y_4(x)]$ implies that a DM presented with a high kurtosis gamble would be gain seeking in gain domains for a sufficiently large increase in relative wealth. [S]he is less temperate when given an opportunity to double her wealth.

According to Menezes and Wang (2005) DMs with skewness preference (like in Figure 5.6) take probability mass from the mean and transfers it to the tail of the distribution. Recent experiments by Ebert and Wiesen (2012) confirms the existence of the behaviours characterized above for higher moments. In particular, Bernoulli's DMs are risk seeking over gain and loss domains in this veritable "preference for skewness" setting that has implications for asset pricing outside the scope of this paper (Kraus and Litsenberger, 1976; Harvey and Siddique, 2000; Dittmar, 2002; Barberis and Huang, 2008; Polkovnichenko and Zhao, 2013).

[4] Markowitz (1952, p. 155) extended Friedman and Savage (1948) utility function to include a convex-concave segment in gain domain to account for insurance (concave) and gambling (convex) choices of DMs. Our analysis is more primitive since the Bernoulli approximations in Table 5.1 do not generate convex-concave curves.

[5] Recall that the numeraire wealth level is 1.0 so $0.9 < x \leq 1$ in (5.2.9) almost doubles or doubles the numeraire wealth level.

Figure 5.5: Risk averter utility function

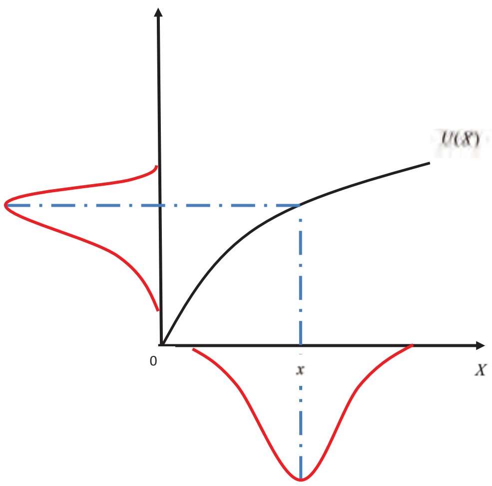
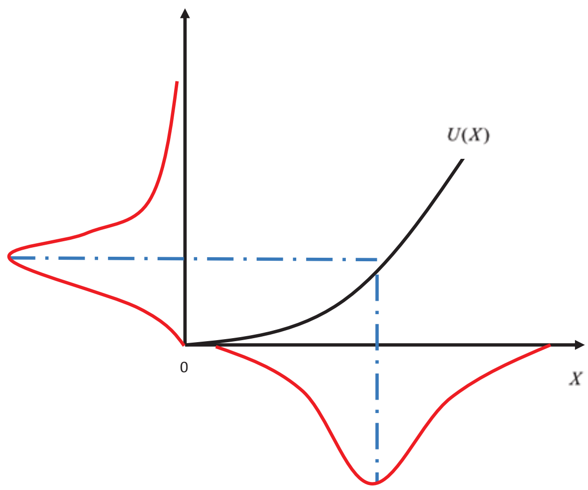


Figure 5.6: Risk seeker utility function



Left tail skew distribution projected by risk averter in Figure 5.5. It implies risk averter avoids losses from tail risk. Right tailed skew distribution projected by risk seeker in Figure 5.6. It implies risk seeker attracted to gains from long shot bias.

Case(iii): Steeper for losses than for gains

According to H́ospital’s rule (Apostol, 1967, pp. 292-293) Bernoulli’s loss aversion index at the origin is

$$\lambda^B = \lim_{x\to 0} \frac{u'_\ell(-x)}{u'_g(x)} = 1 \tag{5.2.11}$$

When we impose Tversky and Kahneman (1992) reported median loss aversion index estimate of 2.25 we introduce a kink at the “reference point” where there was none before. The tails are now longer and the loss aversion index is a lot larger than it is under Bernoulli’s specification. Perhaps most important, Tversky and Kahneman (1992) “interference” with the value function induced a kink at 0 so local loss aversion is now undefined there. Instead of λ^B above, we now get the Köbberling and Wakker (2005) estimate

$$\lambda^{KW} = \frac{v'_\ell(0^-)}{v'_g(0^+)} \tag{5.2.12}$$

which spawned a whole literature around estimation of the loss aversion index (Wakker, 2010). In the sequel we show how a simple global loss aversion index can be derived and provide some estimates.

More on point, in Figure 5.3 *Bernoulli's analysis is invariant to the orientation of Bs or Bs' since wealth levels are not affected by the orientation*. In a nutshell, the curve Bs' is latent in Bernoulli's analysis because he did not place his *reference point for changes in wealth* at the origin. Ironically, Tversky and Kahneman (1992, p. 309) proffered the following specification for their value function:

$$v(x) = \begin{cases} x^{\alpha}, & x > 0 \\ -\lambda(-x)^{\beta}, & x < 0 \end{cases} \tag{5.2.13}$$

There, α and β are shape parameters and λ is the celebrated loss aversion index that captures asymmetric responses to symmetric gain and loss. More on point, if $x < 0$ then $-x > 0$. So the value function segment $(-x)^{\beta}$ is positive and concave for $0 < \beta < 1$ ($\beta = 0.5$ is Gabriel Cramer's specification) and functionally equivalent to Bernoulli's *Bs*. Thus, Bernoulli satisfied KT79 condition (ii). By pre-multiplying the concave segment (Bs) by $-\lambda$ (for $\lambda > 1$), Tversky and Kahneman in effect reoriented *Bs* so that it is convex (say Bs'). In the context of *behavioural operator theory*, that is a "spin"–no pun intended. Thus, but for the power law specification, the value function in (5.2.13) is qualitatively the same as Bernoulli's utility function. In that case, if a sufficiently small gain and loss are equidistant from the reference point B, i.e., BL' and BL in Figure 5.3, we would expect to find $\alpha \approx \beta$.

Lemma 5.2.3 (Equality of shape parameters)**.** *For a sufficiently small symmetric gain and loss equidistant from the reference point B in Bernoulli's utility function shape parameters are approximately equal.* □

Thus, it is no surprise that Tversky and Kahneman (1992, pg. 311) upholds Lemma 5.2.3 when they report in the italicized text below:

> [I]t is common to assume a parametric form (e.g., a power utility function), but this approach confounds the general test of the theory with that of the specific parametric form. For this reason, we focused here on the qualitative properties of the data rather than onparameter estimates and measures of fit. However, in order to obtain a parsimonious description of the present data, we used a nonlinear regression procedure to

estimate the parameters of equations (5) and (6), separately for each subject. *The median exponent of the value function was 0.88 for both gains and losses, in accord with diminishing sensitivity*. [Emphasis added]

Perhaps more important, one need not premultiply the curve by $-\lambda$ to capture the effect of the loss aversion parameter λ since the reoriented curve represented by the dotted line Bs' already captures asymmetric skew evidenced by $V_\ell^B(-x) > V_g^B(x)$, where the former is Bernoulli's value for loss (ℓ) and the later is value for gain (g). That is, λ is a latent parameter in Bernoulli's model. Thereby, causing Bernoulii's latent analysis of Bs' under auspice of Bs to satisfy KT79 (iii). In fact, a cursory inspection of Table 5.1 shows that at points equidistant from the origin 0, Bernoulli utility is larger in loss domain compared to gain domain. This is evident in Figure 5.2 where $V_\ell(-x) > V_g(x)$. By virtue of satisfying $(i) - (iii)$ in Kahneman and Tversky (1979, pg. 279) we proved

Proposition 5.2.4 (Bernoulli's value function). *Bernoulli (1738) value function is* qualitatively equivalent *to Kahneman and Tversky (1979) value function.* □

The geometry of Proposition 5.2.4 is depicted in Figure 5.7 and Figure 5.8. It is based on juxtaposition of Fishburn and Kochenberger (1979, Fig. 1, pg. 504), and an annotated sketch of Bernoulli (1738, pg. 26) representation in the context of our prior analysis.

Figure 5.7: Fishburn-Kochenberger utility with reference point

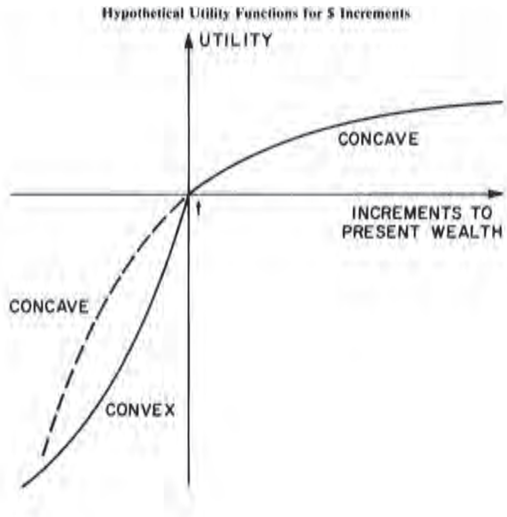


Figure 5.8: Bernoulli Utility with reference wealth

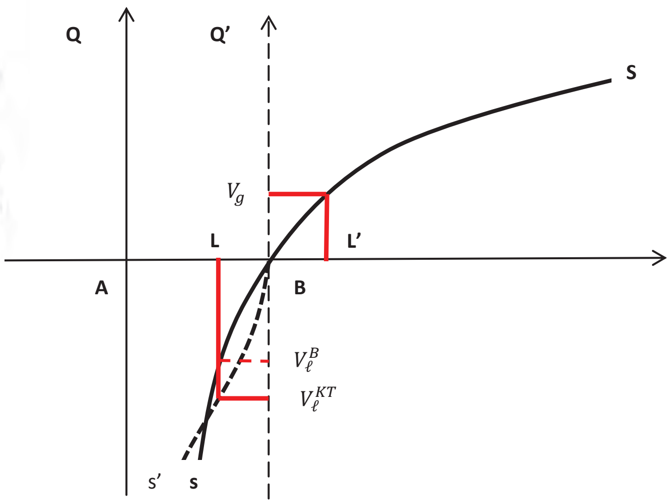


Figure 5.7 is a reproduction of Fishburn and Kochenberger (1979, Fig. 1, p. 504) juxtaposed to Figure 5.8 which reproduces Figure 5.3 for convenience and comparison. Fishburn and Kochenberger (1979) conducted a metastudy based on "changes in wealth or return on investment" (similar to the x in (5.2.8)) by fitting separate utility functions in loss and gain domains using linear, power and exponential specifications. In the context of Figure 5.8, they multiplied Bs by -1 to get Bs' and concluded that Bs' was convex and steeper than BS. However, they noticed that regardless of specification "below-target utility ... is almost always steeper than above-target utility". This is consistent with Bs being steeper than BS in Bernoulli's specification.

5.3 A global loss aversion index formula from Bernoulli utility

In this section, we use the geometry of Bernoulli's utility function to derive a global utility loss aversion index and establish its relation to Fisher's z-transform statistic.[6] Cursory inspection of Figure 5.8 shows that the value V_ℓ^{KT} for Kahneman and Tversky's skew is such that

$$V_\ell^{KT} > V_\ell^B > V_g \Rightarrow \frac{V_\ell^{KT}}{V_g} > \frac{V_\ell^B}{V_g} > 1 \tag{5.3.1}$$

Thus, the "disutility" of loss in either case is such that it is greater than the "utility" of an equal nominal gain. Let $A_{V_\ell^{KT}}, A_{V_g^{KT}}, A_{V_\ell^B}, A_{V_g^B}$ be the *impact of an incremental change in wealth* η under the Kahneman-Tversky (KT) and Bernoulli (B) value functions. Thus, from (5.2.8) and (5.2.13),

[6] Cohen (2014, Fig. 1, p. 9) introduced a "state dependent loss aversion Bernoulli utility" function that is kinked at the reference point. However, he used a Kőszegi and Rabin (2006) type specification to characterize his loss aversion index.

for a *symmetric deviation* η from the "reference point" we get the impacts:

$$A_{V_\ell^{KT}} = \int_{-\eta}^{0} (-\lambda(-x)^\beta)dx = -\lambda\left\{\frac{(-\eta)^{\beta+1}}{\beta+1}\right\}, \quad \eta > 0 \tag{5.3.2}$$

$$A_{V_g^{KT}} = \int_{0}^{\eta} x^\alpha dx = \frac{\eta^{\alpha+1}}{\alpha+1} \tag{5.3.3}$$

$$A_{V_\ell^{B}} = \int_{x-\eta}^{x} \frac{dy}{y}dy = \ln x - \ln(x-\eta) \tag{5.3.4}$$

$$A_{V_g^{B}} = \int_{x}^{x+\eta} \frac{dy}{y}dy = \ln(x+\eta) - \ln x \tag{5.3.5}$$

Inasmuch as Lemma 5.2.3 suggests that shape parameters are equal for sufficiently small symmetric gain and loss, the loss aversion index λ^{KT} derived from the impact of a change in wealth in (5.3.2) and (5.3.3); and λ^B derived from the same change in (5.3.4) and (5.3.5) are given by

$$\lambda^{KT} = \frac{|-\lambda\frac{(-\eta)^{\beta+1}}{\beta+1}|}{\frac{\eta^{\alpha+1}}{\alpha+1}} = \lambda\frac{\alpha+1}{\beta+1}|\eta|^{\beta-\alpha} \approx \lambda, \quad \text{when } \alpha \approx \beta \tag{5.3.6}$$

$$\lambda^B(x,\eta) = \frac{\ln x - \ln(x-\eta)}{\ln(x+\eta) - \ln x} \tag{5.3.7}$$

Our ratio-of-areas approach to deriving the loss aversion index differs from that in the literature which favors first derivatives and ratio of utilities (Wakker, 2010, p. 239). Nonetheless, it is a valid measure as shown by (5.3.6).[7] Undeniably, the loss aversion index in (5.3.7) depends on the reference wealth level x (depicted by B in Figure 5.3) and the amount of loss η (depicted by BL in Figure 5.3). Thus we proved

Theorem 5.3.1 (A reference dependent loss aversion index for Bernoulli)**.** *The loss aversion index $\lambda^B(x,\eta)$ in (5.3.7) computed from Bernoulli's value function is reference dependent. In particular, $\lambda^B(x,\eta)$ is a global loss aversion index over the distribution of change in relative wealth x and loss η.* □

5.3.1 Global loss aversion index, its conjugate, and Fisher's z-transform

Without loss of generality, in the sequel we replace *change in relative wealth* x with the corresponding gain and loss amount η. Given the relationships in (5.3.1), (5.3.6) and (5.3.7) we have

[7] Wakker (2010, p. 268) addresses the effect on local utility loss aversion index measures when $\alpha \neq \beta$.

numeraire or reference point driven loss aversion index relationships:

$$\lambda^{KT} > \lambda^{B} > 1 \Rightarrow \lambda^{B} = \frac{\ln 1 - \ln(1-\eta)}{\ln(1+\eta) - \ln 1} = \frac{-\ln(1-\eta)}{\ln(1+\eta)} > 1 \tag{5.3.8}$$

$$\Rightarrow \ln(1-\eta^{2}) < 0 \Rightarrow |\eta| < 1 \tag{5.3.9}$$

Since $\eta > 0$, the operational inequality is $0 < \eta < 1$. That inequality implies that the absolute nominal change in wealth must be less than the reference wealth level for loss aversion to be upheld. Beyond that point, the formula breaks down. It is an open question as to what happens when wealth increases by a factor of 2 or more. Perhaps most important, Kahneman and Tversky's loss aversion index λ^{KT} in (5.3.6) is uniformly distributed over their value function for all changes in wealth. So in theory, there is no bound on the magnitude of changes relative to the reference point–it is irrelevant. By contrast, the loss aversion index λ^{B} in (5.3.7) and (5.3.8) is reference dependent and responsive to all changes in wealth less than the reference wealth level. To wit, it is "global". Thus, we proved

Theorem 5.3.2 (A Global Loss Aversion Index Formula)**.**
A global loss aversion index formula for a loss η (expressed as a percent change in wealth relative to a numeraire), when utility is log concave, is given by

$$\lambda^{B}(\eta) = -\frac{\ln(1-\eta)}{\ln(1+\eta)}$$

where $0 < \eta < 1, \quad 0 \leq \lambda^{B} \leq \infty$. □

The conjugate loss aversion index formula is derived when utility is not logconcave but when it is the "antilog". That is Bernoulli's logconcave function is now transformed to $U(y) = \exp(y)$. See Figure 5.11, *infra*. Recall that if y_r is numeraire wealth then $y = (1 \pm \eta)y_r$. In which case for a nominal symmetric gain\loss η in a neighbourhood of reference wealth level y_r, the conjugate loss aversion index formula is

$$\lambda^{*^{B}}(\eta) = \frac{\exp(1-\eta)y_r}{\exp(1+\eta)y_r} = \exp(-2\eta\, y_r) \tag{5.3.10}$$

Here, $0 < \lambda^{*^B} \leq 1$ for $0 \leq y_r < \infty$. In this case, for Bernoulli's canonical reference point 1 we have

Corollary 5.3.3 (Conjugate global loss aversion index formula).
The conjugate global loss aversion index formula for a loss η (expressed as a percent change in wealth relative to a numeraire) is given by

$$\lambda^{*^B}(\eta) = \frac{\exp(1-\eta)}{\exp(1+\eta)} = \exp(-2\eta)$$

where $0 \leq \eta \leq 1$. □

The literature shows that Kőszegi and Rabin (2006, 2007) formulated stylized models of reference dependent preferences, based on concave utility functions, but failed to proffer a closed form loss aversion index formula. Thus, the loss aversion index is robust to criticism against the application of logconcave utility to all wealth levels. Because a log concave utility function is just a special case of the abstract concave utility in Kőszegi and Rabin (2006). Perhaps more important, assuming bivariate normality between reference wealth Theorem 5.3.2 suggests that λ^B is related to Fisher's z-transformation[8] if we treat symmetric gain and loss (η) as a pseudo correlation coefficient, i.e., $|\eta| \leq 1$.[9] That is, it transforms the (truncated) interval, normalized by "reference wealth", from $[0,1]$ for η to $[0,\infty]$, so that $0 \leq \lambda^B \leq \infty$. We summarize this artefact in the following

Theorem 5.3.4 (Fisher z-transform test for loss aversion). *Assume that $W_1,\ldots,W_n$ are rank ordered independent identically distributed wealth levels for $n > 3$. Let W_r be a reference wealth level* $1 < r < n$. *Assume that W_j, W_r are* iid bivariate lognormal *and that* $\frac{W_j}{W_r} = 1 + \widehat{\eta}_j,\ j \neq r$. *Let* $\widetilde{\eta} = \frac{1}{n}\sum_{j=1}^{n} \widehat{\eta}_j^2 - 1$ *be a random variable, z be Fisher's z-transform, and* $|\widetilde{\eta}| < 1$ *be a* given truncated symmetric gain and loss *relative to W_r. It is known that if $E[\widetilde{\eta}] = 0$, then the population parameter η with sample estimate $\widetilde{\eta}$ is such that z is normally distributed with mean* $\frac{1}{2}\ln\left(\frac{1+\eta}{1-\eta}\right)$ *and variance* $(n-3)^{-1}$, *i.e.,* $z \sim N\left(\frac{1}{2}\ln\left(\frac{1+\eta}{1-\eta}\right),\ (n-3)^{-1}\right)$ *(Anderson, 2003, p. 134). Then from Theorem 5.3.2*

$$\widehat{z} = \frac{1}{2}(1+\widehat{\lambda}^B)\ln(1+\eta)$$

[8] Refer to Cramér (1962, p. 241).

[9] (Tversky and Kahneman, 1992, Table 5) provides means of the correlations between high and low probability gains and losses after transformation to Fisher's z statistic.

where $\widehat{z}$ is the sample Fisher z-transform for sample estimates $\widehat{\lambda}^B$ and η with test statistic

$$\widehat{Z} = \left((1+\widehat{\lambda}^B)\ln(1+\eta) - \ln\left(\frac{1+\eta}{1-\eta}\right) \right) \frac{\sqrt{n-3}}{2} \sim N(0,1)$$

Proof. See Appendix 5.B. □

Remark 5.3.1. Berry and Mielke Jr (2000) show that the Fisher z-transform test would be valid here only when $E[\eta] = 0$. That condition is satisfied here *a fortiori*.

Example–Fisher z-transform test for utility loss aversion index

According to Theorem 5.3.4 we would reject an observed value $\widehat{\lambda}$ at η percentage points from the reference point if it was statistically significant different from the λ^B expected in Table 5.2. For example, at sample size $n = 30$ if we observed a loss aversion estimate $\widehat{\lambda} = 0.9$ located at $\eta = 0.8$, i.e. 80% distance away from the reference point, then the test statistic is $\widehat{Z} = -2.807$. So we would reject $\widehat{\lambda} = 0.9$ as being too small since we would expect $\lambda^B = 2.738$ in Table 5.1 when $\eta = 0.8$. For small samples we use a t-test with $n-2$ degrees of freedom, and sample variance $s_n = [n(n-3)]^{-1}$. We obtain the sample T-statistic by substituting s_n for $\sqrt{n-3}$ in the formula for $\widehat{Z}$. For the $\widehat{\lambda}, \eta$ pair above, suppose $n = 10$. We get the sample statistic T=-4.520 which is statistically significant at $p = 0.005$ at $\nu = 10-2 = 8$ degrees of freedom $(t^{10-2}_{.005} = -3.355)$. Again, we would reject $\widehat{\lambda} = 0.9$ as being too small, compared to $\lambda^B = 2.738$ in the table, for location distance $\eta = 0.8$ from the reference point. □

Table 5.2: Sample Distribution of Loss Aversion Index for Bernoulli Utility

Loss η	Loss Aver. Index $\lambda^B(\eta)$	Conj. Loss Aver. Index $\lambda^{B^*}(\eta)$
0	0	1
0.00001	1.000010000	0.999980
0.05	1.051303993	0.904837
0.1	1.105448714	0.818731
0.2	1.223901086	0.67032
0.3	1.359464654	0.548812
0.4	1.518180605	0.449329
0.5	1.709511291	0.367879
0.6	1.949539695	0.301194
0.7	2.268957226	0.246597
0.8	2.738132742	0.201897
0.9	3.587397603	0.165299
0.99	6.692251671	0.138069
0.99999	16.60976029	0.135338
1.0	∞	0.135335
Mean	2.18367494*	0.433408308
STDev	1.617554778	0.317314522

* This value is for $0 < \lambda^B(\eta) \leq 0.99$.

Loss column η represents relative loss as a fraction of Bernoulli's reference wealth level. $\lambda^B(\eta)$ is the global loss aversion index we would expect to find for a subject with Bernoulli preferences. $\eta^{B^*}(\eta)$ is the corresponding global gain seeking index for conjugate utility for that subject.

Figure 5.9 is an unscaled plot of Fisher z-transform for the loss data η in Table 5.2. The z-transform is approximately linear for $-0.5 < z < 0.5$ and it steepens fairly rapidly after that. It has asymptotes at $\eta = \pm 1$. Table 5.2 provides a sample distribution for $\lambda^B(\eta)$ and its conjugate $\lambda^{B^*}(\eta)$ based on equally spaced intervals between 0.1 and 0.9. The points 0.00001, 0.05, 0.99, 0.99999 were inserted to highlight the behavior of the distribution near

Figure 5.9: Distribution of Fisher z-transform

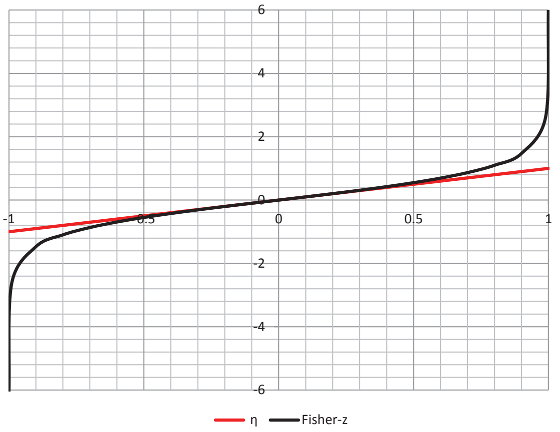
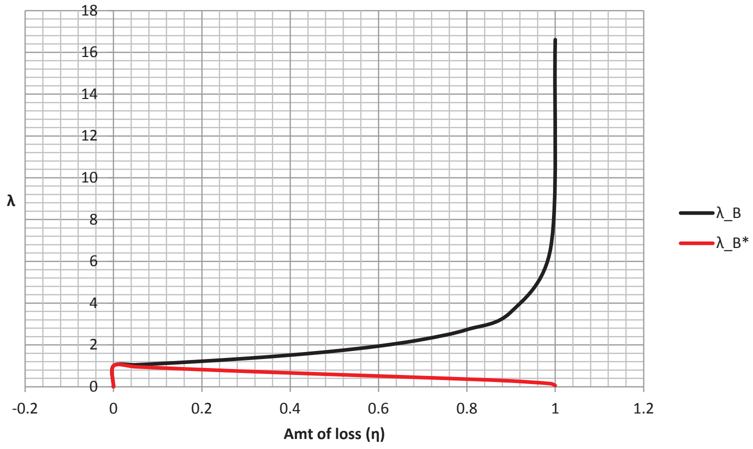
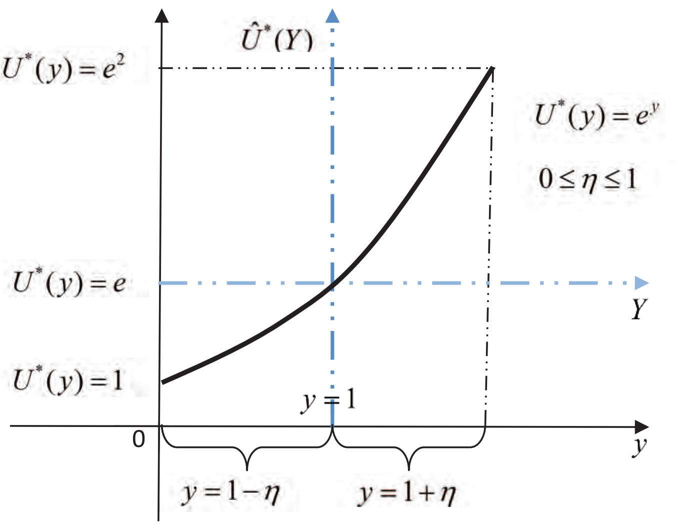


Figure 5.10: Distribution of Global Loss Aversion Index for Bernoulli Utility



In Figure 5.9 $|\eta| \uparrow 1 \Rightarrow$ Fisher $z \to \infty$. In Figure 5.10 $\eta \to 1 \Rightarrow \lambda^B \to \infty$. Intuitively, this implies that as an agent approaches losing it all her loss aversion becomes quite high and explosive upon losing it all. In contrast λ^{B^*} exhibits gain seeking. The smaller the number, the higher the gain seeking. So gain seeking is highest when an agent is faced with losing it all, and it approaches the limiting value $\lambda^{B^*} = 0.135335$ for relative wealth. In (5.3.10) as reference wealth $y_r \to \infty$, we have $\lambda^{B^*} \to 0$.

Figure 5.11: Conjugate Bernoulli Utility Function



Conjugate utility function $U^*(y) = \exp(y)$ for risk seeking is the inverse of Bernoulli utility function $U(y) = \ln(y)$ for risk aversion. For numeraire wealth y_r, *relative wealth* is $y \backslash y_r = 1 + \eta, y > y_r$; $y \backslash y_r = 1 - \eta, y < y_r$, $0 \le \eta \le 1$. Without loss of generality, in Figure 5.11 we assume y is relative wealth. A change of axes to "reference point" $(1, e)$ induces $Y = y - 1$, and $\widehat{U}^*(Y) = U^*(y) - e$ for gains $Y > 0$, $\widehat{U}^*(Y) = e - U^*(y)$ for losses $Y < 0$. So $\widehat{U}^*(\eta) = U^*(1+\eta) - e$ for gains and $\widehat{U}^*(-\eta) = e - U^*(1-\eta)$ for losses with no kink at the reference point $(1, e) \sim (0, 0)$.

the edge. A plot of $\lambda^B(\eta)$ is depicted in Figure 5.10. In Figure 5.11 a conjugate Bernoulli function is plotted to depict convex utility and risk seeking over gains. Köbberling and Wakker (2005, p. 128) also constructed a normalized *concave* utility function with a *convex conjugate* to show how Arrow-Pratt risk aversion index relates to the loss aversion index. Here, we are only interested in the loss aversion index. The plot shows that $\eta \to 1 \Rightarrow \lambda^B(\eta) \to \infty$ and that $\lambda^B(\eta)$ is slowly varying as η approaches 1.

Theory and evidence of α-stable loss aversion

Definition 5.3.1 (Regularly varying). $\lambda^B(\eta)$ is regularly varying (RV) with index α if

$$\lim_{t\to\infty} \frac{\lambda^B(t\eta)}{\lambda^B(\eta)} = \eta^\alpha \tag{5.3.11}$$

We use the notation $\lambda^B \in RV_\alpha$ to represent this phenomenon. The inverse function $\lambda^{B^{-1}} \in RV_{1\backslash\alpha}$ for $\eta^{\frac{1}{\alpha}}$. □

The following representation theorem implies that $\lambda^B(\eta)$ is α-stable. It is a restatement of De Haan and Ferreira (2007, Thm. B.1.6).

Theorem 5.3.5 (α-stable loss aversion). *If $\lambda^B \in RV_\alpha$, then there exist measurable functions $a : \mathbb{R}_+ \to \mathbb{R}$ and $c : \mathbb{R}_+ \to \mathbb{R}$ with*

$$\lim_{t\to\infty} c(t) = c^*, \quad 0 < c^* < \infty, \textit{ and } \lim_{t\to\infty} a(t) = \alpha \tag{5.3.12}$$

and $t_0 \in \mathbb{R}_+$ such that for $t > t_0$

$$\lambda^B(t) = c(t) \exp\left(\int_{t_0}^{t} \frac{a(s)}{s} ds\right) \tag{5.3.13}$$

Conversely, if (5.3.13) holds with a and c satisfying (5.3.12), then $\lambda \in RV_\alpha$.

Proof. See De Haan and Ferreira (2007, Thm. B.1.6, p. 365) or Feller (1970, p. 282). □

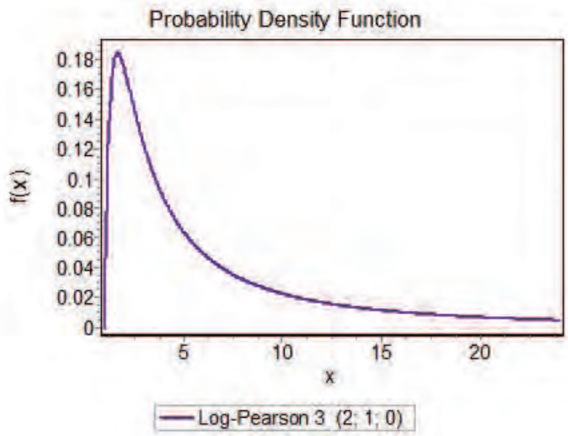
We use sample data for the loss aversion indexes in Fishburn and Kochenberger (1979) metastudy to test the efficacy of Theorem 5.3.5. Table 5.3 presents the descriptive statistics. It shows that the underlying distribution has excess kurtosis (right skew) with large variance, relatively narrow interquartile range $Q1 - Q3$ and a long tail between $Q3 - Q4$. Table 5.12 depicts a Log Pearson Type III (Log Gamma) distribution which was fitted to the data. Refer to Kleiber and Kotz (2003, §5.3 Log–Gamma Distribution) for details on Log Pearson Type III distributions. The specific parameterization fitted via maximum likelihood estimation (MLE) is

$$p(\lambda) = \begin{array}{l} \frac{1}{\lambda\,|\beta|\,\Gamma(\alpha)}\left(\frac{\ln\lambda - \gamma}{\beta}\right)\exp\left(-\left(\frac{\ln\lambda-\gamma}{\beta}\right)\right), \quad 0 < \lambda < \exp(\gamma)\ \ \beta < 0; \\ \qquad\qquad\qquad \exp(\gamma) < \lambda < \infty\ \ \beta > 0 \end{array} \tag{5.3.14}$$

Table 5.3: Diagnostics for Fishurn-Kochenberger loss aversion index data

Statistic	Value	Percentile	Value
Sample Size	30	Min	0.8
Range	164.4	5%	1.295
Mean	12.34	10%	1.83
Variance	876.65	25% (Q1)	2.825
Std. Deviation	29.608	50% (Median)	4.85
Coef. of Variation	2.3994	75% (Q3)	7.725
Std. Error	5.4057	90%	22.96
Skewness	5.0684	95%	87.925
Excess Kurtosis	26.809	Max	165.2

Figure 5.12: Fitted Log Pearson 3



The descriptive statistics in Table 5.3 are for the loss aversion index estimates for two-piece linear $(L^- L^+)$ local utility function in Fishburn and Kochenberger (1979, Tales 1A, 1B, pp. 508-509). They show that the index is leptokurtic and right skewed. The data were fitted to a Log Pearson Type III distribution ($p = 0.32514$ for Anderson-Darling goodness of fit statistic) plotted in Table 5.12. Fishburn and Kochenberger (1979) also reported extreme values for the loss aversion index, i.e., $\lambda = 3300$, $\lambda = \infty$ for a two-piece exponential $(E^- E^+)$ local utility function.

The fitted values are $\overset{\triangle}{\alpha} = 2$, $\overset{\triangle}{\beta} = 1$, and $\overset{\triangle}{\gamma} = 0$ where α and β are shape parameters. The distribution is characterized by its mode which has the form $\exp((\alpha - 1)\backslash(\beta + 1))$. Perhaps most important, if we put (5.3.14) in correspondence with Theorem 5.3.5 we find that

$$c(t) \equiv \frac{1}{t|\beta|\Gamma(\alpha)} \left(\frac{\ln t - \gamma}{\beta}\right) \tag{5.3.15}$$

$$\exp\left(\int_{t_0}^{t} \frac{a(s)}{s} ds\right) \equiv \exp\left(-\left(\frac{\ln t - \gamma}{\beta}\right)\right) \tag{5.3.16}$$

Thus, empirical evidence shows that the α-stable feature of the loss aversion index extends to its probability distribution. Vizly, the composite function $(p \circ \lambda)(\eta) \in [0,1]$ implies $(p \circ \lambda)(\eta) = \eta \in [0,1]$, and $\lambda(\eta) = p^{-1}(\eta)$. So that p is an inverse function in the class $RV_{1\backslash\alpha}$. See e.g., De Haan and Ferreira (2007, p. 368).

Curiously, $\lambda^B(\eta) + \lambda^{B^*}(\eta) \approx 2.0,\ 0 < \eta \leq 0.5$. Perhaps more important, the *mean value* for $\lambda^B(\eta) \approx 2.18,\ 0 < \eta < 0.99$ in Table 5.2 is close to the median value in Tversky and Kahneman (1992, pg. 311)("[t]he median λ was 2.25, indicating pronounced loss aversion"). Unlike Tversky and Kahneman (1992) who posit a constant local loss aversion index λ, the Bernoulli loss aversion index $\lambda^B(\eta)$ is monotone increasing in the amount of loss η. One major restriction of the Fisher z-transform test is that it does not accommodate $|\eta| > 1$. So it only addresses gains or loss which size is less that one hundred percent (100%) of the numeraire reference level.

The conjugate loss aversion $\lambda^{*^B}(\eta)$ in Corollary 5.3.3 depicts the case in when the slope of the curve in Figure 5.11 at points $y > 1$ is greater than the slope at points $y < 1$. That is, when the utility curve is convex–the conjugate of concave–over gains and losses. This is an important result because Kahneman and Tversky (1979) only accounts for those cases when the utility curve is concave over $y > 1$ and convex when $y < 1$. However, laboratory results often report loss aversion index estimates $0 < \lambda < 1$ which connotes gain seeking (Wakker, 2010, p. 239) and convexity over gains and losses.[10] It implies that low income are more prone to gamble since their utility function is convex near reference wealth (Light, 1977; Scott and Barr, 2012; Beckert and Lutter, 2012).

5.4 Conclusion

Close inspection of the geometry of Bernoulli (1738) original utility function specification show that it accommodates gains and losses relative a reference point. Furthermore, it has a reference

[10] Refer to Harrison and Rutström (2008, Fig. 14, pg. 97) for a a distribution of λ with values greater than and less than 1.

dependent feature that supports a global loss aversion index that is monotone increasing in the magnitude of loss. Moreover, we introduce theory and evidence that this loss aversion index is slowly varying so it is α-stable. So we extend the solution space for loss aversion to α-stable laws. We show how Fisher's z-transform test can be used in econometric tests of the loss aversion index. According to our analysis, many of the results explained by prospect theory's skew S-shaped value function are explicable with Bernoulli's incipient utility function specification or at least a suitable high order approximation of it. Further research is needed to see whether results in this paper extend to utility functions in general.

5.A APPENDIX

5.B Proof of Fisher z-transform test for loss aversion Theorem 5.3.4

Proof. By hypothesis if W_j, W_r are bivariate iid lognormal, then $\widetilde{\eta}$ is a truncated random variable (Greene, 2003, p. 757). To see this, let $W_j = \exp(X_j)$ and $W_r = \exp(X_r)$ where X_j, X_r are normally distributed. Write $W_j = W_r(1+\widehat{\eta}_j)$ so that $\ln\left(\frac{W_j}{W_r}\right) = \ln(1+\widehat{\eta}_j) \approx \widehat{\eta}_j = X_j - X_r$, and the latter difference is iid normally distributed such that $E[\widehat{\eta}_j] = E[X_j] - E[X_r] = 0$. By hypothesis, $\frac{1}{n}\sum_{j=1}^{n}\widehat{\eta}_j^2 = 1+\widetilde{\eta}$. Thus $E\left[\frac{1}{n}\sum_{j=1}^{n}\widehat{\eta}_j^2\right] = 1 \Longrightarrow E[\widetilde{\eta}] = 0$. Recall that $|\widetilde{\eta}| < 1$ is a truncated random variable by hypothesis. In which case, according to Cramér (1962, eq(18.3.3), p. 242) we can write the z-transform $e^{2z} = \dfrac{1+\eta}{1-\eta}$ for the population parameter η with sample estimate $\widetilde{\eta}$. However from Theorem 5.3.2 we get $-\ln(1-\eta) = \widehat{\lambda}^B \ln(1+\eta)$. After taking log of e^{2z} and substituting the foregoing expression for $\ln(1-\eta)$ in the formula we get the desired result. □

Bibliography

The CBOE Volatility Index - VIX®. Undated Report. Chicago Board Option Exchange.

Abbas, A. (2006, March-April). Maximum Entropy Utility. *Operations Research 54*(2), 277–290.

Abdellaoui, M., A. Baillon, L. Placido, and P. Wakker (2011, Apr). The Rich Domain of Uncertainty: Surce Functions and their Experimental Implementation. *American Economic Review 101*(2), 695–723.

Abdellaoui, M., H. Bleichrodt, and H. Kammoun (2013). Do financial professionals behave according to prospect theory? An experimental study. *Theory and Decision 74*(3), 411–429.

Abdellaoui, M., H. Bleichrodt, and C. Paraschiv (2007). Loss Aversion Under Prospect Theory: A Parameter Free Measurement. *Management Science 53*(10), 1659–1674.

Abdellaoui, M., O. LHaridon, and H. Zank (2010). Separating curvature and elevation: A parametric probability weighting function. *Journal of Risk and Uncertainty 41*, 39–65.

Abel, A. B. (1990, May). Asset Prices Under Habit Formation and Keeping Up With The Jonses. *American Economic Review 80*(2), 38–42. Papers and Proceedings of the 102nd Annual Meeting of AEA.

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2015). Systemic risk and stability in financial networks. *American Economic Review 105*(2), 564–608.

Aczel, J. (1966). *Lectures in Functional Equations and Their Applications*. New York. Academic Press.

Adelson, M. (2013, Spring). The Deeper Causes of The Financial Crisis: Mortgages Alone Cannot Explain It. *Journal of Portfolio Management*, 1–16.

Akerlof, G. A. (1970, Aug). The market for lemons: quality uncertainty and the market mechanism. *Quarterly Journal of Economics 84*(3), 488–500.

Akheizer, N. and I. Glazman (1961). *Theory of Linear Operators in Hilbert Space*. Frederick Ungar Publ. Co.:New York. Dover reprint 1993.

al Nowaihi, A. and S. Dhami (2006, Sep). Insurance and Probability Weighting Functions. University of Leicester, Dep't of Economics, Working Paper 05/19.

al Nowaihi, A. and S. Dhani (2010, April). Composite Prospect Theory: A Proposal to Combine 'Prospect Theory' and 'Cumulative Prospect Theory'. Working Paper No. 10/11, Univ. Leicester, Dep't Econ.

Aldridge, I. (2014). High-Frequency Runs and Flash-Crash Predictability. *Journal of Portfolio Management 40*(3), 113–123.

Ali, M. M. (1977). Probability and utility estimates for racetrack bettors. *Journal of Political Economy 85*(4), 803.

Allen, F. and D. Gale (2000, Feb). Financial Contagion. *Journal of Political Economy 108*(1), 1–33.

Amiel, Y., F. Cowell, L. Davidovitz, and A. Polovin. Preference reversals and the analysis of income distributions. *Social Choice and Welfare 30*(2), 305–330.

Anderson, L. R. and J. M. Mellor (2009). Are risk preferences stable? Comparing an experimental measure with a validated survey-based measure. *Journal of Risk and Uncertainty 39*(2), 137–160.

Anderson, T. W. (2003). *Introduction to Multivariate Analysis* (3rd ed.). Wiley Interscience. New York, NY: John Wiley & Sons, Inc.

Anderson, T. W. and D. A. Darling (1954, Dec). A Goodness of Fit Test. *Journal of the American Statistical Association 49*(268), 765–769.

Andries, M. (2014, December). Consumption-based Asset Pricing with Loss Aversion. Working Paper. Dep't of Economics, Tolouse University.

Anscombe, F. J. and R. J. Aumann (1963, Mar.). A Definition of Subjective Probability. *Annals of Mathematical Statistics 34*(1), 199–205.

Apostol, T. M. (1967). *Advanced Calculus*, Volume I. New York: John Wiley & Sons, Inc.

Arnold, B. C. and P. L. Brockett (1992). On Distributions Whose Component Ratios are Cauchy. *American Statistician 46*(1), 25–26.

Arnold, V. I. (1984). *Ordinary Differential Equations* (3rd ed.). Universitext. New York, NY: Springer-Verlag.

Ashparouva, E., P. Bossaerts, and A. Tran (2014, May). Market Bubbles and Crashes as an Expression of Tension Between Social and Individual Rationality. Working Paper, Dep't of Finance, Eccles School of Business, Univ. of Utah.

Åstebro, T., J. Mata, and L. Santos-Pinto (2015). Skewness seeking: risk loving, optimism or overweighting of small probabilities? *Theory and Decision 78*(2), 189–208.

Avellaneda, M. (1998). The Minimum-Entropy Algorithm And Related Methods For Calibrating Asset-Pricing Models. In *Proceedings of the International Congress of Mathematicians*, Volume 3, pp. 545–563. IMU, Berlin: Doc. Math.

Backus, D. K., A. W. Gregory, and S. E. Zin (1989). Risk premiums in the term structure: Evidence from artificial economies. *Journal of Monetary Economics 24*(3), 371 – 399.

Baker, M., X. Pan, and J. Wurgler (2012). The effect of reference point prices on mergers and acquisitions. *Journal of Financial Economics 106*(1), 49–71.

Bandi, F. M. and A. Tamoni (2015, May). Business-cycle Consumption Risk and Asset Prices. Working Paper, John Hopkins Univ., Department of Finance.

Barberis, N. and M. Huang (2001, Aug). Mental Accounting, Loss Aversion, and Individual Stock Returns. *Journal of Finance 56*(4), 1247–1292.

Barberis, N. and M. Huang (2008). Stocks as Lotteries: The Implications of Probability Weighting for Security Prices. *American Economic Review 98*(5), 2066–2100.

Barberis, N., M. Huang, and T. Santos (2001). Prospect Theory and Asset Prices. *Quarterly Journal of Economics 116*(1), 1–53.

Barberis, N. C. (2013, September). Thirty Years of Prospect Theory in Economics: A Review and Assessment. *Journal of Economic Perspectives 27*(1), 173–96.

Barro, R. J. (2006, August). Rare Disasters and Asset Markets in the Twentieth Century. *Quarterly Journal of Economics 121*(3), 823–866.

Barro, R. J. (2009, March). Rare Disasters, Asset Prices, and Welfare Costs. *American Economic Review 99*(1), 243–262.

Bask, M. (2010). Measuring potential market risk. *Journal of Financial Stability 6*(3), 180 – 186.

Bask, M. and A. Widerberg (2012). Actual and potential market risks during the stock market turmoil 2007–2008. *Applied Financial Economics 22*(5), 339–349.

Battistin, E., R. Blundell, and A. Lewbel (2009, Dec). Why Is Consumption More Log Normal than Income? Gibrats Law Revisited. *Journal of Political Economy 117*(6), 1140–1154.

Baucells, M. and F. H. Heukamp (2012). Probability and time trade–off. *Management Science 58*(4), 831–842.

Becker, G. M., M. H. Degroot, and J. Marschak (1964). Measuring utility by a single-response sequential method. *Behavioral Science 9*(3), 226–232.

Becker, R. A. (2012, Mar). The Variance Drain and Jensen's Inequality. Indiana Univ.-Bloomington, CAEPR Working Paper No. 2012-004. Available at SSRN: http://ssrn.com/abstract=2027471.

Beckert, J. and M. Lutter (2012). Why the Poor Play the Lottery: Sociological Approaches to Explaining Class-based Lottery Play. *Sociology*. Forthcoming.

Bell, D. E. (1982). Regret in Decision Making Under Uncertainty. *Operations Research 30*(5), 961–981.

Benartzi, S. and R. E. Thaler (1995, Feb.). Myopic Loss Aversion And The Equity Premium Puzzle. *Quarterly Journal of Economics 110*(1), 73–92.

Benartzi, S. and R. H. Thaler (1999). Risk Aversion or Myopia? Choices in Repeated Gambles and Retirement Investment. *Management Science 45*(3), 364–381.

BenSaïda, A. (2012). Are financial markets stochastic: A test for noisy chaos. *Am Int J Contemp Res 2*(8), 57–68.

BenSaïda, A. (2014). Noisy chaos in intraday financial data: Evidence from the American index. *Applied Mathematics and Computation 226*(0), 258 – 265.

Berger, J. (1985). *Statistical Decision Theory and Bayesian Analysis* (2nd ed.). Springer Series in Statistics. New York, N.Y.: Springer-Verlag.

Berglund, N. and B. Gentz (2010). Stochastic dynamic bifurcations and excitability. In C. Laing and G. J. Lord (Eds.), *Stochastic Methods in Neuroscience*, Chapter 3, pp. 64–93. New York, NY: Oxford University Press.

Bernoulli, D. (1738). Specimen Theoriae Novae de Mensura Sortis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae 5*, 175–192. English translation (1954): Exposition of a new theory on the measurement of risk. *Econometrica*, 22(1):23-36.

Berry, K. J. and P. W. Mielke Jr (2000). A monte carlo investigation of the fisher z transformation for normal and nonnormal distributions. *Psychological Reports 87*(3f), 1101–1114.

Bhattacharya, S., C. A. E. Goodhart, D. P. Tsomocos, and A. P. Vardoulakis (2015, Aug). A Reconsideration of Minsky's Financial Instability Hypothesis. *Journal of Money Credit and Banking 47*(5), 931–973.

Birnbaum, M. H. (2004). Tests of rank-dependent utility and cumulative prospect theory in gambles represented by natural frequencies: Effects of format, event framing, and branch splitting. *Organizational Behavior and Human Decision Processes 95*(1), 40–65.

Birnbaum, M. H. (2005). Three new tests of independence that differentiate models of risky decision making. *Management Science 51*(9), 1346–1358.

Birnbaum, M. H. (2008). New paradoxes of risky decision making. *Psychological Review 115*(2), 463.

Birnbaum, M. H. and J. B. Navarrete (1998). Testing descriptive utility theories: Violations of stochastic dominance and cumulative independence. *Journal of Risk and Uncertainty 17*(1), 49–79.

Birnbaum, M. H., J. N. Patton, and M. K. Lott (1999). Evidence against rank-dependent utility theories: Tests of cumulative independence, interval independence, stochastic dominance, and transitivity. *Organizational Behavior and human decision Processes 77*(1), 44–83.

Black, F. and M. Scholes (1973, May-Jun). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy 81*(3), 637–654.

Blavatskky, P. (2009, Dec). Preference Reversal and Probabilistic Decisions. *Journal of Risk and Uncertainty 39*(3), 237–250.

Blavatskky, P. R. (2007). Stochastic Expected Utility Theory. *Journal of Risk and Uncertainty 34*, 259–286.

Blavatskyy, P. (2013). A Probability Weighting Function for Cumulative Prospect Theory and Mean-Gini Approach to Optimal Portfolio Investment. Working paper. Available at http://ssrn.com/abstract=2380484.

Blavatskyy, P. R. (2005). Back to the St. Petersburg Paradox? *Management Science 51*(4), 677–678.

Blavatskyy, P. R. (2011). Loss aversion. *Economic Theory 46*(1), 127–148.

Bochner, S. (1955). *Harmonic Analysis and the Theory of Probability*. Berkeley, CA: Univ. California Press.

Bongaerts, D., R. Roll, D. Rösch, M. A. Van Dijk, and D. Yuferova (2014, Aug). The Propagation of Shocks Across International Equity Markets: A Microstructure Perspective. Available at SSRN: http://ssrn.com/abstract=2475518.

Booij, A. S., B. M. S. van Praag, and G. van de Kuilen (2010). A parametric analysis of prospect theorys functionals for the general population. *Theory and Decision 68*(1-2), 115–148.

Borowiak, D. S. and A. F. Shapiro (2014). *Financial and Actuarial Statistics: An Introduction* (2nd ed.). Boca Raton, FL: CRC Press: A Chapman & Hall Book.

Bougerol, P. and N. Picard (1992). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics 52*(12), 115 – 127.

Bowman, D., D. Minehart, and M. Rabin (1999). Loss Aversion in a Consumption Saving Model. *Journal of Economic Behavior and Organizations 38*, 155–178.

Boyer, B. H. and K. Vorkink (2014). Stock Options as Lotteries. *Journal of Finance 69*(4), 1485–1527.

Breeden, D. T. (1979). An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities. *Journal of Financial Economics 7*, 265–296.

Breeden, D. T. and R. H. Litzenberger (1978, Oct). Prices of State-Contingent Claims Implicit in Option Prices. *Journal of Business 51*(4), 621–651.

Breeden, D. T., R. H. Litzenberger, and T. Jia (2015). Consumption-Based Asset Pricing: Research and Applications. *Annual Review of Financial Economics 7*(1), Forthcoming.

Busemeyer, J. and Z. Wang (2007, July). Introduction to Quantim Information Processing for Cognitive Scientists: Quantum Dynamics. Working Paper, Dep't Psychology, Indiana University-Bloomington.

Busemeyer, J. R. and A. Diederich (2002). Survey of Decision Field Theory. *Mathematical Social Sciences 43*, 345–370.

Busemeyer, J. R., E. M. Pothos, R. Franco, and J. S. Trueblood (2011, April). A Quantum Theoretical Explorations for Probability Judgment Errors. *Psychological Review 118*(2), 193–218.

Busemeyer, J. R., Z. Wang, and J. T. Townsend (2006, June). Quantum Dynamics of Human Decision Making. *Journal of Mathematical Psychoology 50*(3), 220–241.

Butler, D. J. and G. C. Loomes (2007). Imprecision as an account of the preference reversal phenomenon. *American Economic Review*, 277–297.

Campbell, J. Y. (2003). Consumption-based asset pricing. In *Handbook of the Economics of Finance*, Volume 1, pp. 803–887. Amsterdam, Noth-Holland),: Elsevier.

Campbell, J. Y. (2015). Emerging Trends: Asset Pricing. In Robert A. Scott and Stephen M. Kosslyn (Ed.), *Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource*. New York, NY: John Wiley & Sons, Inc.

Campbell, J. Y. and J. H. Cochrane (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy 107*(2), 205–251.

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

Cantril, H. (1965). *The Pattern Of Human Concerns*. New Brunswick, NJ: Rutgers University Press.

Carr, P. and L. Wu (2004). Time-changed Lévy processes and option pricing. *Journal of Financial Economics 71*(1), 113–141.

Cartea, A. and S. Jaimungal (2013). Modelling Asset Prices for Algorithmic and High Frequency Trading. *Applied Mathematical Finance 20*(6), 512–547. Available at http://dx.doi.org/10.1080/1350486X.2013.771515.

Castellacci, G. and Y. Choi (2014). Modeling Contagion in the Eurozone Crisis via Dynamical Systems. *Journal of Banking & Finance*. Forthcoming.

Cavagnaro, D. R., M. A. Pitt, R. Gonzalez, and J. I. Myung (2013). Discriminating among probability weighting functions using adaptive design optimization. *Journal of Risk and Uncertainty 47*, 255–289.

Çinlar, R. (2011). *Probability and Stochastics*, Volume 261 of *Graduate Tests in Mathematics*. New York, NY: Springer.

Chabi-Yo, F. and Z. Song (2013, June). Recovering The Probability Weights of Tail Events With Volatility Risk From Option Prices. Working Paper, Department of Finance, Fisher School of Business, Ohio State U.

Charles-Cadogan, G. (2015a). Harmonic Probability Weighting Functions and Mental States of Decision Makers. Paper prepared for Foundations of Utility and Risk Conference, June 2016 (FUR XVII), University of Warwick, Coventry.

Charles-Cadogan, G. (2015b, Aug). Loss Aversion Index Recovery in The Cross-Section of Subjective Well Being and Happiness On Economic Growth. Paper prepared for NIBS Conferece April 2016 on "Assessing well-being when preferences are incoherent", Univ. East Anglia, Norwich.

Charles-Cadogan, G. (2016a). Diffusing Explosive Portfolio Performance Evaluation Of High Frequency Traders. *Journal of Investment Strategies 5*(1), 1–25.

Charles-Cadogan, G. (2016b, Mar). Expected Utility Theory and Inner and Outer Measures of Loss Aversion. *Journal of Mathematical Economics 63*(1), 10–20. 10.1016/j.jmateco.2015.12.001.

Charles-Cadogan, G. and J. A. Cole (2014). Bankruptcy Risk Induced By Career Concerns of Regulators. *Financial Research Letters 11*, 259–271.

Charles-Cadogan, G. and J. A. Cole (2015). A Regulator's Exercise of Career Option To Quit and Join A Regulated Firm's Management with Applications to Financial Institutions. Working paper. Available at *SSRN eJournal*.

Chen, Y. and S. Xuefeng (2003). Study on the chaos model of liquidity in the stock market. *Systems Research and Behavioral Science 20*(5), 419–425.

Chew, S. H. (1983, Jul). A Generalization of the Quasilinear Mean with Applications to Measurable of Income Inequality and Decision Theory Resolving The Allais Paradox. *Econometrica 51*(4), 1065–1092.

Chew, S. H. and W. S. Waller (1986). Empirical tests of weighted utility theory. *Journal of Mathematical Psychology 30*(1), 55–72.

Chicone, C. C. (1999). *Ordinary Differential Equations With Applications*, Volume 34 of *Texts in Applied Mathematics*. New York, NY: Springer.

Choi, S. (2014, Dec.). The Impact of VIX Shocks on Emerging Market Economies: Flight to Quality Mechanism. Working Paper, Dep't Economics, UCLA.

Choi, S., R. Fisman, D. Gale, and S. Kariv (2007). Consistency and Heterogeneity of Individual Behavior under Uncertainty. *American Economic Review 97*(5), 1921–1938.

Choquet, G. (1954). Theory of Capacities. *Annales de L'institut Fourier 5*, 131–295.

Chow, Y. S. and H. Teicher (1988). *Probability Theory: Independence, Interchangeability, Martingales* (2nd ed.). Springer Texts in Statistics. New York, NY: Springer-Verlag.

Clark, P. K. (1973, Jan). A Subordinated Stochastic Process Model With Finite Variance for Speculative Prices. *Econometrica 41*(1), 135–155.

Cochrane, J. (2005). *Asset Pricing* (Revised edition ed.). Princeton, NJ: Princeton University Press.

Cohen, O. (2014, Nov). COnsumption Smoothing with State-Dependent Loss-Aversion Preferences. Working paper. Department of Economics, Northwestern Univ.

Constantinides, G. M. (1990, Jun.). Habit formation: A resolution of the equity premium puzzle. *Journal of Political Economy 98*(3), 519–543.

Coppola, M. (2014). Eliciting risk-preferences in socio-economic surveys: How do different measures perform? . *Journal of Socio-Economics 48*, 1 – 10.

Corsetti, G., P. Pesenti, and N. Roubini (1999). What caused the Asian currency and financial crisis? *Japan and the World Economy 11*(3), 305–373.

Cover, T. M. and J. A. Thomas (1991). *Elements of Information Theory*. Wiley Series in telecommunications. New York: John Wiley & Sons, Inc.

Cox, D. R. and V. Isham (1980). *Point Processes*. London, UK: Chapman and Hall.

Cox, J. C., V. Sadiraj, B. Vogt, and U. Dasgupta (2013). Is there a plausible theory for decision under risk? A dual calibration critique. *Economic Theory 54*(2), 305–333.

Cramér, H. (1962). *Mathematical Methods of Statistics*. Bombay, India: Asia Publishing House. First Indian Edition of 1946 volume.

Cubitt, R. P., A. Munro, and C. Starmer (2004, Jul). Testing Explanations of Preference Reversal. *Economic Journal 114*(497), 709–726.

Curatola, G. (2015). Loss Aversion, Habit Formation and the Term Structures of Equity and Interest Rates. *Journal of Economic Dynamics and Control 53*, 103–122.

Cyert, R. M. and J. G. March (1992). *A Behavioral Theory of The Firm* (2nd ed.). Hoboken, NJ: WIley-Blackwell.

Day, R. H. and W. Huang (1990). Bulls, bears and market sheep. *Journal of Economic Behavior & Organization 14*(3), 299 – 329.

De Guzmán, M. (1981). *Real Variable Methods in Fourier Analysis*, Volume 46 of *North-Holland Mathematics Studies*. North-Holland, Amsterdam: North-Holland Publishing Co.

De Haan, L. and A. Ferreira (2007). *Extreme Value Theory: An Introduction*. Springer Series in Operations Research and Financial Engineering. New York, NY: Springer Science & Business Media.

De Haas, R. and N. Van Horen (2012). International Shock Transmission after the Lehman Brothers Collapse: Evidence from Syndicated Lending. *American Economic Review 102*(3), 231–37.

De Neve, J.-E., G. W. Ward, F. De Keulenaer, B. Van Landeghem, G. Kavetsos, and M. I. Norton (2015, Mar). The Asymmetric Experience of Positive and Negative Economic Growth: Global Evidence Using Subjective Well-Being Data. Technical report, Institute for the Study of Labor. IZA Discussion Paper No. 8914.

Dembo, A. and O. Zeitouni (1998). *Large Deviations Techniques and Applications* (2nd ed.). Number 38. New York, NY: Springer-Verlag.

Devaney, R. L. (1989). *An Introduction To Chaotic Dynamical Systems* (2nd ed.). Studies in Nonlinearity. Reading, MA: Addison-Wesley Publishing Co., Inc.

Diamond, D. W. and P. H. Dybvig (1983, Jan). Bank Runs, Deposit Insurance, and Liquidity. *Journal of Political Economy 91*(3), 401–419.

Dierkes, M. (2009, Nov). Option Implied Risk Attitude Under Rank Dependent Utility. Working Paper, Finance Center, Univ. Münster, Germany.

Dierkes, M. (2013, Apr). Probability Weighting And Asset Prices. Working Paper, Finance Center, Univ. Münster, Germany.

Dillenberger, D., A. Postlewaite, and K. Rozen (2013). Optimism and Pessimism with Expected Utility. *Cowles Foundation Discussion Papers No. 1829, Yale University*.

Dittmar, R. F. (2002). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *Journal of Finance 57*(1), 369–403.

Doerrenberg, P., D. Duncan, and C. Zeppenfeld (2015). Circumstantial risk: Impact of future tax evasion and labor supply opportunities on risk exposure. *Journal of Economic Behavior & Organization 109*, 85 – 100.

Duesenberry, J. S. (1949). *Income, Saving, and The Theory of Consumer Behavior*. Cambridge, MA: Harvard University Press. 4th Printing, 1962.

Duffee, G. R. (2005, Aug). Time Variations in The Covariance Between Stock Returns and Consumption Growth. *Journal of Finance 60*(4), 1673–1712.

Dugundji, J. (1966). *Topology*. Boston, MA: Allyn and Bacon, Inc.

Dumontier, A. (2012, Sep). Best Practices for Long-Term Investors in a Microsecond Market. *The Motley Fool*. Available at http://www.fool.com/investing/general/2012/09/14/bestpracticesforlongterminvestorsinamicros.aspx.

Dunford, N. and J. T. Schwartz (1957). *Linear Operators Part I: General Theory*. Pure and Applied Mathematics. New York: Interscience Publishers, Inc.

Dunn, L. F. and I. A. Mirzaie (2015). Consumer Debt Stress, Changes In Household Debt, And The Great Recession. *Economic Inquiry*. In press.

Dybvig, P. (1995). Dusenberry‘s Ratcheting of Consumption: Optimal Dynamic Consumption and Investment Given Intolerance for any Decline In The Standard of Living. *Review of Economic Studies 62*, 287–313.

Dymski, G. A. (2010). Why the subprime crisis is different: A Minskyian approach. *Cambridge Journal of Economics 34*(2), 239–255.

Easley, D., M. M. López De Prado, and M. O'Hara (2011). The Microstructure of the "Flash Crash": Flow Toxicity, Liquidity Crashes, and the Probability of Informed Tradding. *Journal of Portfolio Management 37*(2), 118–128.

Easley, D. and L. Yang (2014, Jan). Loss Aversion, Survival and Asset Prices. Working Paper. Available at http://ssrn.com/abstract=1784367.

Ebert, S. and D. Wiesen (2012). Joint measurement of risk aversion, prudence, and temperance. Working paper. Available at SSRN eJournal http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1975245.

Eeckhoudt, L., C. Gollier, and T. Schneider (1995). Risk-aversion, prudence and temperance: A unified approach. *Economics Letters 48*(3), 331–336.

Epstein, L. G. and S. E. Zin (1989, July). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica 57*(4), 937–969.

Fehr-Duda, H. and T. Epper (2012). Probability and Risk: Foundations and Economic Implications of Probability Dependent Risk Preferences. *Anual Review of Economics 4*, 563–593.

Feller, W. (1970). *An Introduction to Pobability Theory and Its Applications*, Volume II. New York, NY: John Wiley & Sons, Inc.

Fellner, W. (1961, Nov.). Distortion of Subjective Probabilities as a Reaction to Uncertainty. *Quarterly Journal of Economics 75*(4), pp. 670–689.

Financial Crisis Inquiry Comission (2011, Jan). The Financial Crisis Inquiry Report: Final Report Of The National Commission On The Causes Of The Financial And Economic Crisis In The United States. Technical report, National Commission on the Causes of the Financial and Economic Crisis in the United States, Washington, DC. Superintendent of Documents, U.S. Government Printing Office.

Fishburn, P. (1988). *Nonlinear Preferences and Utility Theory*. John Hopkins Studies in Mathematical Sciences. Baltimore, MD: John Hopkins University Press.

Fishburn, P. C. and G. A. Kochenberger (1979). Two-Piece Von Neumann-Morgenstern Utility Functions. *Decision Sciences 10*(4), 503–518.

Fisher, I. (1933, Oct). The Debt-Deflation Theory of Great Depression. *Econometrica 1*(4), 337–357.

Folland, G. (1992). *Fourier Analysis and Its Aplications*. Wadsworth & Brook/Cole Mathematics Series. Pacific Grove, CA: Brooks/Cole Publishing Co.

Fox, C. R., B. A. Rogers, and A. Tversky (1996). Options traders exhibit subadditive decision weights. *Journal of Risk and Uncertainty 13*(1), 5–17.

Frank, S. A. and E. D. Smith (2010). Measurement Invariance, Entropy and Probability. *Entropy 12*, 289–303. http://dx.doi.org/doi:10.3390/e12030289.

Friedman, M. and L. J. Savage (1948, Aug). The Utility Analysis of Choice Involving Risk. *Journal of Political Economy 56*(4), 279–304.

Gentle, J. E. (2007). *Matrix Algebra: Theory, Computations and Applications*. Springer Texts in Statistics. New York, NY: Springer.

George, T. J. and C.-Y. Hwang (2004). The 52-week high and momentum investing. *Journal of Finance 59*(5), 2145–2176.

Gihman, I. I. and A. V. Skorohod (1972). *Stochastic Differential Equations*. Number 72 in Mathematics and Its Applications. New York, NY: Springer-Verlag.

Gikhman, I. I. and A. V. Skorokhod (1969). *Introduction to The Theory of Random Processes*. Phildelphia, PA: W. B. Saunders, Co. Dover reprint 1996.

Gneezy, U. and J. Potters (1997, May). An Experiment on Risk Taking and Evaluation Periods. *Quarterly Journal of Economics 112*(2), 631–645. In Memory of Amos Tversky (1937-1996).

Goldstein, W. M. and H. J. Einhorn (1987). Expression theory and the preference reversal phenomena. *Psychological Review 94*(2), 236 – 254.

Golec, J. and M. Tamarkin (1998, Feb). Bettors Love Skewness, Not Risk, At The Horse Race Track. *Journal of Political Economy 106*(1), 205–225.

Good, I. J. (1963). Maximum Entropy for Hypothesis Formulation, Especially for Multidimensional Contingency Tables. *Annals of Mathematical Statistics 34*(3), pp. 911–934.

Grasselli, M. and B. Costa Lima (2012). An analysis of the Keen model for credit expansion, asset price bubbles and financial fragility. *Mathematics and Financial Economics 6*(3), 191–210.

Greene, W. H. (2003). *Econometric Analysis* (5th ed.). Upper Saddle Rd., N. J.: Prentice-Hall, Inc.

Grether, D. M. and C. R. Plott (1979, Sep). Economic Theory of Choice and the Preference Reversal Phenomenon. *American Economic Review 69*(4), 623–638.

Grinblatt, M. and J. T. Linnainmaa (2011). Jensen's inequality, parameter uncertainty, and multiperiod investment. *Review of Asset Pricing Studies 1*(1), 1–34.

Grüne, L. and W. Semmler (2008). Asset pricing with loss aversion. *Journal of Economic Dynamics and Control 32*(10), 3253 – 3274.

Guggenheimer, H. W. (1977). *Differential Geometry*. Mineola, New York: Dover Publications, Inc.

Gul, F. (1991, May). A Theory of Disappointment Aversion. *Econometrica 59*(3), 667–686.

Guo, J. and X. D. He (2015, June). Equilibrium Asset Pricing with Epstein-Zin and Loss Averse Investors. Working Paper, Dep't IEOR, Columbia Univ. Available at http://ssrn.com/abstract=2615235.

Guvenen, F., S. Ozkan, and J. Song (2014). The Nature of Countercyclical Income Risk. *Journal of Political Economy 122*(3), 621–660.

Haigh, M. S. and J. List (2005, Feb.). Do Traders Exhibit Myopic Loss Aversion? An experimental analysis. *Journal of Finance 60*(1), 523–534.

Hansen, L. P. and R. Jagannathan (1991, April). Implications of Security Market Data for Models of Dynamic Economies. *Journal of Political Ecoonomy 99*(2), 225–262.

Hansson, B. (1988). *Decisions, Probability, and Utility*, Chapter 8: Risk Aversion as a Problem of Conjoint Measurement, pp. 136–158. NY: Cambridge University Press.

Hardie, B. G. S., E. J. Johnson, and P. S. Fader (1993). Modeling loss aversion and reference dependence effects on brand choice. *Marketing Science 12*(4), 378–394.

Harris, M. and A. Raviv (1993). Differences of Opinion Make a Horse Race. *Review of Financial Studies 6*(3), 473–506.

Harrison, G. W. (1986). An experimental test for risk aversion. *Economics Letters 21*(1), 7–11.

Harrison, G. W. (2011, May). The Rich Domain of Uncertainty: Comment. Working Paper WP2011-13, Georgia State Univ., Center for Economic Analysis of Risk (CEAR).

Harrison, G. W. and E. E. Rutström (2008). *Risk Avrsion in Experiments*, Volume 12 of *Research in Experimental Economics*, Chapter 3: Risk Aversion in the Laboratory, pp. 41–196. Bingley, UK: Emerald Group Publishing Limited.

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

Hartzmark, S. M. (2015). The Worst, the Best, Ignoring All the Rest: The Rank Effect and Trading Behavior. *Review of Financial Studies 28*(4), 1024–1059.

Harvey, C. R. and A. Siddique (2000). Conditional skewness in asset pricing tests. *Journal of Finance 55*(3), 1263–1295.

He, X. D. and X. Y. Zhou (2013). Hope, fear, and aspirations. *Mathematical Finance*. Forthcoming.

Hey, J. D. (1995). Experimental investigations of errors in decision making under risk. *European Economic Review 39*(34), 633 – 640. Papers and Proceedings of the Ninth Annual Congress European Economic Association.

Hey, J. D. (2005). Why we should not be silent about noise. *Experimental Economics 8*(4), 325–345.

Hey, J. D. and C. Orme (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica*, 1291–1326.

Hinvest, N. S., M. J. Brosnan, R. D. Rogers, and T. L. Hodgson (2014). fMRI evidence for procedural invariance underlying gambling preference reversals. *Journal of Neuroscience, Psychology, and Economics 7*(1), 48.

Hirschleifer, D. (2001, Aug). Investor Psychology and Asset Pricing. *Journal of Finance 56*(4), 1533–1597.

Hirshleifer, D. (2015). Behavioral Finance. *Annual Review of Financial Economics 7*(1), Forthcoming.

Hofstede, G. (1980). *Cultures Consequences*. Beverly Hills, CA: Sage.

Hofstede, G. (1983, Spring - Summer). National Cultures in Four Dimensions: A Research-Based Theory of Cultural Differences among Nations. *International Studies of Management & Organization 13*(1/2), 46–74. Cross-Cultural Management: II. Empirical Studies.

Hogarth, R. M. and H. J. Einhorn (1990, Jul). Venture Theory: A Model of Decision Weights. *Management Science 36*(7), 780–803.

Holt, C. A. and S. K. Laury (2002). Risk aversion and incentive effects. *American Economic Review 92*(5), 1644–1655.

Hommes, C. (2013). *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems*. New York, NY: Cambridge University Press.

Hommes, C. H. and S. Manzan (2006). Comments on "Testing for nonlinear structure and chaos in economic time series". *Journal of Macroeconomics 28*(1), 169 – 174. Special Issue: Nonlinear Macroeconomic Dynamics.

Hoyle, A. (2010). *Information-Based Models for Finance and Insurance*. Ph. D. thesis, Department of Mathematics, Imperial College London, London, UK.

Hoyle, E., L. P. Hughston, and A. Macrina (2011). Lévy random bridges and the modelling of financial information. *Stochastic Processes and their Applications 121*(4), 856 – 884.

Huang, C. and R. H. Litzenberger (1988). *Foundations for Financial Economics*. Englewood Cliffs, NJ: Prentice-Hall.

Huddart, S., M. Lang, and M. H. Yetman (2009). Volume and price patterns around a stock's 52-week highs and lows: Theory and evidence. *Management Science 55*(1), 16–31.

Hung, M.-W. and J.-Y. Wang (2011). Loss aversion and the term structure of interest rates. *Applied Economics 43*(29), 4623–4640.

Ikpe, D., S. Mataramvura, and R. Becker (2014). Modelling Financial Information by conditioning. *Communications on Stochastic Analysis 8*(1), 99–110.

Inada, K.-I. (1963, Jun). On a Two-Sector Model of Economic Growth: Comments and a Generalization. *Review of Economic Studies 30*(2), 119–127.

Irwin, J. R., P. Slovic, S. Lichtenstein, and G. H. McClelland (1993). Preference reversals and the measurement of environmental values. *Journal of Risk and Uncertainty 6*(1), pp. 5–18.

Jacobsen, M. (2006). *Point Processs Theory and Applications: Marked Point and Piecewise Determiitic Processes*. Probability and Its Applications. Boston, MA: Birkhäuser.

Jaynes, E. T. (1968). Prior Probabilities. *IEEE Transactions On Systems Science and Cybernetics 4*(3), 227–241.

Jensen, M. H., A. Johansen, and J. Simonsen (2003). Inverse Statistics in Economics: The Gain-Loss Asymmetry. *Physica A: Statistical Mechanics and its Applications 324*, 338–343.

Johnson, A., O. Ledoit, and D. Sornette (2000). Crashes as Critical Points. *International Journal of Theoretical and Applied Finance 3*(2), 219–255.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1994). *Continuous univariate distributions* (2nd ed.), Volume 1 of *Wiley Interscience Publication*. New York, NY: New York: John Wiley & Sons.

Jost, J. (2005). *Dynamical Systems: Examples of Complex Behavior*. Universitext. New York, NY: Springer.

Kahneman, D. (2002, Dec.). Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice. In *Nobel Prize Lecture*, Nobel Laureate Lectures, Stockholm, Sweden, pp. 449–489. http://www.nobelprize.org/. Available at http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2002/kahnemann-lecture.pdf.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decisions under risk. *Econometrica 47*(2), 263–291.

Karatzas, I. and S. E. Shreve (1991). *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. New York, N.Y.: Springer-Verlag.

Karni, E. and Z. Safra (1987, May). "Preference Reversal" and the Observability of Preferences by Experimental Methods. *Econometrica 55*(3), 675–685.

Karni, E. and Z. Safra (1990, Jun). Rank-Dependent Probabilities. *Economic Journal 100*(401), pp. 487–495.

Katznelson, Y. (2004). *An Introduction to Harmonic Analysis* (3rd ed.). New York, NY: Cambridge University Press. Corrected edition.

Keen, S. (2013, Feb.). A monetary Minsky model of the Great Moderation and the Great Recession. *Journal of Economic Behavior & Organization 86*, 221–235.

Kelly, J. L. (1956). A new interpretation of information rate. *Information Theory, IRE Transactions on 2*(3), 185–189.

Kemeny, J. G., A. W. Knapp, and J. L. Snell (1976). *Denumerable Markov Chains*. New York, NY: Springer.

Khincin, Y. (1957). *Mathematical Foundations of Information Theory*. Mineola, NY: Dover Publications, Inc. Translated from Russian by R. A. Silverman and M. D. Friedman.

Kőszegi, B. and M. Rabin (2006). A Model of Reference-Dependent Preferences. *Quarterly Journal of Economics 121*(4), 1133–1165.

Kőszegi, B. and M. Rabin (2007). Reference-dependent risk attitudes. *American Economic Review 97*(4), 1047–1073.

Khrennikov, A. Y. and E. Haven (2009). Quantum mechanics and violations of the sure-thing principle: The use of probability interference and other concepts. *Journal of Mathematical Psychology 53*(5), 378 – 388. Special Issue: Quantum Cognition.

Kilka, M. and M. Weber (2001, Dec). What determines The Shape of The Probability eighting Function Under Uncertainty? *Management Science 47*(12), 1712–1726.

Kindleberger, C. P. and R. Z. Aliber (2011). *Manias, panics and crashes: A history of financial crises*. Palgrave Macmillan.

King, M. A. and S. Wadhwani (1990). Transmission of volatility between stock markets. *Review of Financial studies 3*(1), 5–33.

Kleiber, C. and S. Kotz (2003). *Statistical Size Distributions in Economics and Actuarial Sciences*. Wiley Series in Probability and Statistics. Hoboken, NJ: John Wiley & Sons, Inc. Wiley-Interscience.

Kliger, D. and O. Levy (2010). Overconfident investors and probability misjudgments. *Journal of Socio-Economics 39*(1), 24–29.

Knight, F. B. (1962, May). On the Random Walk and Brownian Motion. *Transactions of the American Mathematical Society 103*(2), 218–228.

Knight, F. H. (1921). *Risk, Uncertainty and Profit*. New York, NY: Hart, Shaffner & Marx.

Köbberling, V. and P. Wakker (2005). An index of loss aversion. *Journal of Economic Theory 112*, 119–131.

Komlos, J. (2014, June). Behavioral Indifference Curves. Working Paper 20240, National Bureau of Economic Research.

Kraus, A. and R. H. Litsenberger (1976, September). Skewness Preference and Valuation of Risk Assets. *Journal of Finance 31*(4), 1085–1100.

Kullback, S. (1968). *Information Theory and Statistics*. Mineola, NY: Dover Publications, Inc. Dover reprint 1978.

Kumar, A. (2009). Who gambles in the stock market? *Journal of Finance 64*(4), 1889–1933.

Kutner, M. H., C. J. Nachtsheim, J. Neter, and W. Li (2005). *Applied Statistical Models* (5th ed.). New York: McGraw-Hill International Edition.

Lambert, T. (2015, May). Lobbying on Regulatory Enforcement Actions: Evidence from Banking. UC-Louvain Working Paper.

Landsman, Z. M. and E. A. Valdez (2003). Tail conditional expectations for elliptical distributions. *North American Actuarial Journal 7*(4), 55–71.

Lattimore, P. K., J. R. Baker, and A. D. Witte (1992). The influence of probability on risky choice: A parametric examination. *Journal of Economic Behavior & Organization 17*(3), 377 – 400.

Lechner, R. J. (1971). Harmonic analysis of switching functions. In A. Mukhopadhyay (Ed.), *Recent Developments in Switching Theory*, Chapter V, pp. 121 – 228. Academic Press.

Leonov, G. A. and N. V. Kuznetsov (2007). Time-Varying Linearization And The Perron Effects. *International Journal of Bifurcation and Chaos 17*(04), 1079–1107.

Lichtenstein, S. and P. Slovic (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology 89*(1), 46–55.

Lichtenstein, S. and P. Slovic (1973, Nov.). Response-induced reversals of preference in gambling: An extended replication in Las Vegas. *Journal of Experimental Psychology 101*(1), 16–20.

Lichtenstein, S. and P. Slovic (2006). The construction of preference: An overview. In S. Lichtenstein and P. Slovic (Eds.), *The Construction of Preference*, Chapter 1, pp. 1–41. New York, NY: Cambridge University Press.

Light, I. (1977, Dec.). Numbers Gambling Among Blacks: A Financial Institution. *American Sociological Review 42*(6), 892–904.

Lindman, H. (1971). Inconsistent Preferences Among Gambles. *Journal of Experimental Psychology 89*, 390–397.

London, D. D. (2015, Aug 28th). China's Stock Market: Was The Crash That Big? *Economist*. Available at http://www.economist.com/blogs/freeexchange/2015/08/chinasstockmarket0.

Loomes, G. (2005). Modelling the stochastic component of behaviour in experiments: Some issues for the interpretation of data. *Experimental Economics 8*(4), 301–323.

Loomes, G. and G. Pogrebna (2014, May). Measuring Individual Risk Attitudes When Preferences Are Imprecise. *Economic Journal 124*, 569–593.

Loomes, G. and G. Pogrebna (2015, Jul). Do Preference Reversals Disappear When We Allow for Probabilistic Choice? Working Paper. Dep't Economics, Univ. Warwick. Available at http://wrap.warwick.ac.uk/70154/.

Loomes, G., C. Starmer, and R. Sugden (1991). Observing Violations of Transitivity by Experimental Methods. *Econometrica*, 425–439.

Loomes, G. and R. Sugden (1982, Dec.). Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty. *Economic Journal 92*(368), 805–824.

Loomes, G. and R. Sugden (1983, Jun). A Rationale for Preference Reversal. *American Economic Review 73*(3), 428–432.

Loomes, G. and R. Sugden (1995). Incorporating a stochastic element into decision theories. *European Economic Review 39*(3), 641–648.

Lopes, L. L. (1981). Decision making in the short run. *Journal of Experimental Psychology: Human Learning and Memory 7*(5), 377.

Lopes, L. L. (1987). Between Hope and Fear: The Psychology of Risk. *Advances in Experimental Social Psychology 20*(3), 255–295.

Lopes, L. L. (1990). Re-Modeling Risk Aversion: A Comparison Of Bernoullian and Rank Dependent Value Approaches. In G. M. von Furstenberg (Ed.), *Acting Under Uncertainty: Multidisciplinary Conceptions*, Theory and Decision Library. Series A., Philosophy and Methodology of The Social Science, Chapter 11, pp. 267–299. Dordrecht, Netherlands: Kluwer Academic Publishers.

Lopes, L. L. (1995). Algebra and Process in The Modeling of Risky Choice. In J. Busemeyer, R. Hastie, and D. L. Medin (Eds.), *Decision Making From A Cognitive Perspective*, Volume Advances in Research and Theory of *The Psychology of Learning and Motivation*, pp. 177–220. San Diego, CA: Academic Press, Inc.

Lopes, L. L. and G. C. Oden (1999). The role of aspiration level in risky choice: A comparison of cumulative prospect theory and sp/a theory. *Journal of Mathematical Psychology 43*(2), 286 – 313.

Lucas, R. E. (1978, Nov.). Asset Prices In An Exchange Economy. *Econometrica 46*(6), 1429–1445.

Lucca, D., A. Seru, and F. Trebbi (2014). The Revolving Door and Worker Flows in Bank regulation. *Journal of Monetray Economics 65*, 17–32.

Luce, R. D. (2000). *Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches*. Scientific Psychology Series. Mawah, NJ: Lawrence Erlbaum Associates, Publishers.

Luce, R. D. (2001). Reduction invariance and Prelec's weighting functions. *Journal of Mathematical Psychology 45*, 167–179.

Ludvigson, S. C. (2013). Advances in Consumption-Based Asset Pricing: Empirical Tests. In *Handbook of the Economics of Finance*, Volume 1, Chapter 12, pp. 799–906. Amsterdam, Noth-Holland),: Elsevier B. V.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006, Nov.). Ambiguity Aversion, Robustness, and the Variational representation of Preferences. *Econometrica 74*(6), 1447–1498.

MacCrimmon, K. R. (1968). Descriptive and normative implications of the decision-theory postulates. In K. Borch and J. Mossin (Eds.), *Risk and Uncertainty*, Chapter 1, pp. 3–23. London, UK: Macmillan Publishers, Inc.

Machina, M. J. (1987a). Decision-making in the presence of risk. *Science 236*(4801), 537–543.

Machina, M. J. (1987b). Decision-making in the presence of risk. *Science 236*(4801), 537–543.

Markowitz, H. (1952, April). The Utility of Wealth. *Journal of Political Economy 40*(2), 151–158.

Mataramvura, S. (2015, April). Indifference pricing of contingent claims in incomplete markets. Working Paper, University of Cape Town, Faculty of Commerce.

McCaffrey, D. F., S. Ellner, A. R. Gallant, and D. W. Nychka (1992, Sep). Estimating the Lyapunov Exponent of a Chaotic System With Nonparametric Regression. *Journal of the American Statistical Association 87*(419), 682–695.

McGrane, V. and P. N. Da Costa (2014, Sept. 12). Federal Reserve Creates Financial Stability Committee. *Wall Street Journal*. Available at http://www.wsj.com/articles/federal-reserve-creates-financial-stability-committee-1410543972.

McLennan, A. (2012, November). Advanced Fixed Point Theory for Economics. Mimeo. Available at http://cupid.economics.uq.edu.au/mclennan/Advanced/advanced˙fp.pdf. Last visited 11/23/2012.

McLennan, A. (2014, April). Advanced Fixed Point Theory for Economics. Mimeo. Available at http://cupid.economics.uq.edu.aumclennanAdvancedadvanced_fp.pdf. Last visited 10/1/2014.

Mehra, R. and E. C. Prescott (1985, March). The Equity Premium: A Puzzle. *Journal of Monetary Economics 15*(2), 145–161.

Menezes, C. F. and X. H. Wang (2005). Increasing outer risk. *Journal of Mathematical Economics 41*(7), 875 – 886.

Merkle, C. (2015, May). Financial Loss Aversion Illusion. Working Paper, Dep't of Finance & Banking, University of Mannheim. Available at SSRN: http://ssrn.com/abstract=2445941.

Merton, R. C. (1992). *Continuous Time Finance*. Boston, MA: Blackwell.

Minsky, H. (1986). *Stabilizing an Unstable Economy*. New Haven, CT: Yale Univ. Press.

Minsky, H. (1994). The financial instability hypothesis. In M. Arestis and M. C. Sawyer (Eds.), *The Elgar Companion To Radical Political Economy*, pp. 153–157. Brookfield, VT: Elgar. Available at Jerome Levy Economics Institute, Bard College. Working Paper No. 74. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=161024.

Minsky, H. (1996). Uncertainty And The Institutional Structure of Capitalist Economies: Remarks upon receiving the Veblen-Commons Award. *Journl of Economic Issues 30*(2), 357–368.

Mishkin, F. S. (1999). Global Financial Instability: Framework, Events, Issues. *Journal of Economic Perspectives 13*(4), 3–20.

Montes, I., E. Miranda, and S. Montes (2013). Decision making with imprecise probabilities and utilities by means of statistical preference and stochastic dominance. *European Journal of Operational Research*. Forthcoming.

Mosteller, F. and P. Nogee (1951, Oct). An Experimental Measure of Utility. *Journal of Political Economy 59*(5), 371–404.

Noussair, C. N., S. T. Trautmann, and G. van de Kuilen (2014). Higher order risk attitudes, demographics, and financial decisions. *Review of Economic Studies 81*(1), 325–355.

Nychka, D., S. Ellner, A. R. Gallant, and D. McCaffrey (1992). Finding Chaos in Noisy Systems. *Journal of the Royal Statistical Society. Series B (Methodological) 54*(2), 399–426.

Owen, J. and R. Rabinovitch (1983, Jun). On the Class of Elliptical Distributions and their Applications to the Theory of Portfolio Choice. *Journal of Finance 38*(3), pp. 745–752.

Park, J. Y. and Y. Whang (2012). Random walk or chaos: A formal test on the Lyapunov exponent. *Journal of Econometrics 169*(1), 61 – 74. Recent Advances in Panel Data, Nonlinear and Nonparametric Models: A Festschrift in Honor of Peter C.B. Phillips.

Parker, J. A. and C. Julliard (2005). Consumption risk and the cross section of expected returns. *Journal of Political Economy 113*(1), 185–222.

Peek, J. and E. S. Rosengren (2000). Collateral Damage: Effects of the Japanese Bank Crisis on Real Activity in the United States. *American Economic Review 90*(1), 30–45.

Pfanzagl, J. (1967). Subjective probability derived from the Morgenstern-von Neumann utility concept. In *Essays in Mathematical Economics in Honour of Oskar Morgenstern*, Chapter 18. Princeton, NJ: Princeton University Press.

Pfiffelmann, M. (2011). Solving the St. Petersburg Paradox in cumulative prospect theory: The right amount of probability weighting. *Theory and Decision 71*(3), 325–341.

Poledna, S., S. Thurner, J. D. Farmer, and J. Geanakoplos (2013). Leverage-induced systemic risk under Basle II and other credit risk policies. *arXiv preprint arXiv:1301.6114. Available at http://arxiv.org/pdf/1301.6114v1.pdf*.

Polkovnichenko, V. and F. Zhao (2013). Probability weighting functions implied in options prices. *Journal of Financial Economics 107*(3), 580 – 609.

Por, H.-H. and D. V. Budescu (2013). Revisiting the gain–loss separability assumption in prospect theory. *Journal of Behavioral Decision Making 26*(4), 385–396.

Pothos, E. M. and J. R. Busemeyer (2009, June). A Quantum Probability Explanation for Violation of 'Rational' Decision Theory. *Proc. R. Soc. B 276*, 2171–2178.

Prelec, D. (1998). The probability weighting function. *Econometrica 60*, 497–528.

Preston, M. G. and P. Baretta (1948, April). An Experimental Study of the Auction Value of an Uncertain Outcome. *American Journal of Psychology 61*(2), 183–193.

Quax, R., D. Kandhai, and P. M. Sloot (2013, May). Information dissipation as an early-warning signal for the Lehman Brothers collapse in financial time series. *Scientific Reports 3*. Article no. 1898.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behaviour and Organization 3*(4), 323–343.

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

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica 68*(5), 1281–1292.

Rebaza, J. (2012). *A First Course In Applied Mathematics*. Hobokken, NJ: John Wiley & Sons, Inc.

Regenwetter, M., J. Dana, and C. Davis-Stober (2011). Transitivity of preferences. *Psychological Review 118*(1), 42.

Reinhart, C. M. and K. S. Rogoff (2014). This time is different: A panoramic view of eight centuries of financial crises. *Annals of Economics and Finance 15*(2), 1065–1188.

Resulaj, A., R. Kiani, D. M. Wolpert, and M. N. Shadlen (2009). Changes of mind in decision-making. *Nature 461*(7261), 263–266.

Riedel, F. (2009). Optimal consumption choice with intolerance for declining standard of living. *Journal of Mathematical Economics 45*(7), 449–464.

Rieger, M. O. and M. Wang (2006). Cumulative prospect theory and the st. petersburg paradox. *Economic Theory 28*(3), 665–679.

Rieger, M. O., M. Wang, and T. Hens (2011, Aug). Prospect theory around the world. NHH Dept. of Finance & Management Science Discussion Paper No. 2011/19.

Rieger, M. O., M. Wang, and T. Hens (2015). Risk preferences around the world. *Management Science 61*(3), 637–648.

Rietz, T. A. (1988). The equity risk premium a solution. *Journal of Monetary Economics 22*(1), 117 – 131.

Rosenberg, J. V. and R. F. Engle (2002). Empirical Pricing Kernel. *Journal of Financial Economics 64*, 341–372.

Rubinstein, A. and U. Segal (2012). On the likelihood of cyclic comparisons. *Journal of Economic Theory 147*(6), 2483 – 2491.

Sahm, C. R. (2012). How Much Does Risk Tolerance Change? *Quarterly Journal of Finance 02*(04), 1250020.

Samoradnitsky, G. and M. S. Taqqu (1994). *Stable Non-Gaussian Random Processes: Stochastic Models With Infinite Variance*, Volume 1 of *Stochastic Modelling*. Boca Raton, FL: CRC Press.

Samuelson, P. (1963). Risk and Uncertainty: A Fallacy of Large Numbers. *Scientia XCVIII*, 108–113.

Savadori, L. and L. Mittone (2015). Temporal distance reduces the attractiveness of p-bets compared to $-bets. *Journal of Economic Psychology 46*(0), 26 – 38.

Savage, F. H. (1972). *Foundations Of Statistics* (2nd rev ed.). Mineola, NY: Dover Publications, Inc.

Schmeilder, D. (1989, May). Subjective Probability and Expected Utility without Additivity. *Econometrica 57*(3), 571–587.

Schmidt, U. and H. Zank (2005). What Is Loss Aversion. *Journal of Risk and Uncertainty 30*(2), 157–167.

Schmidt, U. and H. Zank (2008, Jan.). Risk Aversion in Cumulative Prospect Theory. *Management Science 54*(1), pp. 208–216.

Schneeweiss, H. (1974). Probability and utility–dual concepts in decision theory. In Günter Menges (Ed.), *Information, inference and decision*, Volume 1 of *Theory and Decision Library*, pp. 113–144. New York, NY: Springer.

Scott, L. and G. Barr (2012). Unregulated Gambling in South African Townships: A Policy Conundrum? *Journal of Gambling Studies*, 1–14. Forthcoming.

Seidl, C. (2002). Preference reversal. *Journal of Economic Surveys 16*(5), 621–655.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. New York, NY: John Wiley & Sons.

Shea, J. (1995, Aug). Myopia, Liquidity Constraints, and Aggregate Consumption: A Simple Test. *Journal of Money, Credit and Banking 27*(3), 798–805.

Shefrin, H. (2008). Risk and return in behavioral SDF-based asset pricing models. *Journal of Investment Management 6*(3), 1–18.

Shefrin, H. (2009). Behavioralizing Finance. *Foundations and Trends in Finance 4*(1-2), 1–184.

Shefrin, H. (2014). Foreword: Special issue of quantitative finance on 'behavioral finance'. *Quantitative Finance 14*(4), 587–588. Available at http://dx.doi.org/10.1080/14697688.2014.896570.

Shefrin, H. (2015a, June). Assessing the Contribution of Hyman Minsky's Perspective to Our Understanding of Economic Instability. Available at SSRN: http://ssrn.com/abstract=2311045.

Shefrin, H. (2015b, Aug). china today: What minsky would say. *Forbes*. Available at http://www.forbes.com/sites/hershshefrin/2015/08/19/china-today-what-minsky-would-say/.

Shi, X., H. Tsuji, and S. Zhang (2012). Introducing Heterogeneity of Managers' Attitude into Behavioral Risk Scoring for Software Offshoring. *Systems Research and Behavioral Science 29*(3), 299–316.

Shintani, M. and O. Linton (2004). Nonparametric neural network estimation of Lyapunov exponents and a direct test for chaos. *Journal of Econometrics 120*(1), 1 – 33.

Shive, S. and M. Forster (2014, Sep). The Revolving Door For Financial Regulators. Working Paper, Dep't of Finance, Univ. of Notre Dame. Available at http://ssrn.com/abstract=2348968.

Shorack, G. R. and J. A. Wellner (1986). *Empirical Processes with Applications to Statistics*. New York: John Wiley & Sons, Inc.

Siddiqui, A. (2004). *Applied Functional Analysis: Numerical Methods, Wavelets Methods, and Image Processing*, Volume 258 of *Pure and Applied Mathematics*. New York, NW: Marcel Dekker.

Slovic, P. (1995, May). The Construction of Preferences. *American Psychologist 50*(5), 364–371.

Slovic, P. and S. Lichtenstein (1983, Sep). Preference reversals: A broader perspective. *American Economic Review 73*(4), 596–605.

Smith, V. L., G. L. Suchanek, and A. W. Williams (1988, Sep). Bubbles, Crashes, and Endogeneous Expectations in Experimental Spot Asset Markets. *Econometrica 56*(5), 1119–1151.

Springer, M. D. (1979). *The Algebra of Random Variables*. Wiley Series in Probability and Mathematical Statistics. New York, NY: John Wiley & Sons, Inc.

Stearns, S. (2000). Daniel Bernoulli (1738): Evolution and economics under risk. *Journal of Biosciences 25*(3), 221–228.

Stewart, N., S. Reimers, and A. J. L. Harris (2015). On the Origin of Utility, Weighting, and Discounting Functions: How They Get Their Shapes and How to Change Their Shapes. *Management Science 61*(3), 687–705.

Stigler, G. J. (1950, Oct). The Development of Utility Theory. II. *Journal of Political Economy 58*(5), pp. 373–396.

Stiglitz, J. E. and S. J. Grossman (1976). Information and competitive price systems. *American Economic Review 66*(2), 246–253.

Stott, H. (2006). Cumulative Prospect Theory's Functional Menagerie. *Journal of Risk and Uncertainty 32*, 101–130.

Stutzer, M. J. (1996). *Fundamental Theories of Physics*, Volume 62, Chapter Maximum Entropy and Bayesian Methods, pp. 375–390. Kluwer Academic.

Szekeres, P. (2004). *A Course in Modern Mathematical Physics: Groups, Hilbert Space and Differential Geometry*. New York, NY: Cambridge University Press.

Szpiro, G. G. (1986). Relative risk aversion around the world. *Economics Letters 20*(1), 19 – 21.

Takahashi, T. (2006). A mathematical framework for probabilistic choice based on information theory and psychophysics. *Medical Hypotheses 67*(1), 183 – 186.

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

Thaler, R. H. (1990). Anomalies: Saving, fungibility, and mental accounts. *Journal of Economic Perspectives 4*(1), 193–205.

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.

Thurner, S., J. D. Farmer, and J. Geanakoplos (2012). Leverage causes fat tails and clustered volatility. *Quantitative Finance 12*(5), 695–707.

Tobler, P. N., G. I. Christopolous, J. P. O'Doherty, R. J. Dolan, and W. Schultz (2008, Nov). Neuronal Distortions of Reward Probability Without Choice. *Journal of Neuroscience 28*(45), 11703–11711.

Trueblood, J. S. and J. R. Busemeyer (2011). A Quantum Probability Account of Order Effects in Inference. *Cognitive Science 35*(8), 1518–1552.

Tversky, A. (1969). Intransitivity of Preferences. *Psychological Review 76*, 31–48.

Tversky, A. and C. R. Fox (1995). Weighting Risk and Uncertainty. *Psychological Review 102*(2), 269–283.

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.

Tversky, A. and I. Simonson (1993). Context-Dependent Preferences. *Management Science 39*(10), 1179–1189.

Tversky, A., P. Slovic, and D. Kahneman (1990). The Causes of Preference Reversal. *American Economic Review 80*(1), 204–17.

Tversky, A. and P. Wakker (1995, Nov.). Risk Attitudes and Decision Weights. *Econometrica 63*(6), 1255–1280.

Viechnicki, B. (2015, Jan). *Cognitive and Neural Processes of Constucted Preferences*. Ph. D. thesis, University of Pennsylvania. Department of Psychology. Available at : http://repository.upenn.edu/edissertations/115.

Viscusi, W. K. (1989). Prospective Reference Theory: Toward an Explanation of the Paradoxes. *Journal of Risk and Uncertainty 2*(3), 235–263.

Von Neumann, J. (1955). *Mathematical Foundations of Quantum Mechanics*. Princeton, NJ: Princeton University Press.

Von Neumann, J. and O. Morgenstern (1953). *Theory of Games and Economic Behavior* (3rd ed.). Princeton University Press.

Wachter, J. A. (2013). Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility? *Journal of Finance 68*(3), 987–1035.

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

Wakker, P., I. Erev, and E. U. Weber (1994). Comonoonic Indeoendence: The Critical Test Between Classical and Rank-Dependent Utility Theories. *Journal of Risk and Uncertainty 91*, 195–230.

Wakker, P. and A. Tversky (1993). An Axiomatization of Cumulative Prospect Theory. *Journal of Risk and Uncertainty 7*(2), 147–175.

Wakker, P. A. (2010). *Prospect Theory for Risk and Ambiguity*. New York, NY: Cambridge University Press.

Walck, C. (2007). Handbook On Statistical Distributions for Experimentalists. Technical Report SUF-PRY/96-0, University of Stockholm.

Webb, R. (2015, July). The Dynamics of Stochastic Choice. Working Paper, Rotman School of Management, U. Toronto.

Weigert, F. and S. Ruenzi (2013, Mar). Crash Sensitivity And The Cross Section Of Expected Stock Returns. Swiss Inst. Banking and Finance, WP No. 2013/24, Univ. St. Gallen.

Wessels, D. R. (2010, Dec). Dividends: The Major Source of Real Equity Returns. Technical report, DRW Investment Research.

Whaley, R. E. (2000). The Investor Fear Gauge: Explication of the CBOE VIX. *Journal of Portfolio Management 26*(3), 12–17.

Whang, Y. and O. Linton (1999). The asymptotic distribution of nonparametric estimates of the Lyapunov exponent for stochastic time series. *Journal of Econometrics 91*(1), 1 – 42.

White, H. A. (2001). *Asymptotic Theory for Econometricans* (2nd ed.). San Diego, CA: Academic Press, Inc.

Wickens, M. (2011). *Macroeconomic Theory: A Dynamic General Equilibrium Approach* (2nd ed.). Princeton, NJ: Princeton University Press.

Wiggins, S. (2003). *Introduction to Applied Nonlinear Dynamical Systems and Chaos* (2nd ed.), Volume 2 of *Texts in Applied Mathematics*. New York, NY: Springer.

Wigniolle, B. (2014, Apr). Optimism, pessimism and financial bubbles. *Journal of Economic Dynamics and Control 41*, 188–208.

Wilcox, N. (2008). Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison. *Research in Experimental Economics 12*, 197–292. Special Issue: Risk Aversion in Experiments.

Wilcox, N. T. (2011, March). Comparison of Three Probabilistic Models of Binary Discrete Choice Under Risk. Work-in-Progress, Economic Science Institute, Chapman University.

Willard, S. (1970). *Gneral Topology*. Reading, MA: Addison-Wesley Publishing Co.

Wu, G. (1994). An empirical test of ordinal independence. *Journal of Risk and Uncertainty 9*(1), 39–60.

Wu, G. and A. B. Markle (2008). An Empirical Test of Gain-Loss Separability in Prospect Theory. *Management Science 54*(7), 1322–1335.

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

Yogo, M. (2008). Asset prices under habit formation and reference-dependent preferences. *Journal of Business & Economic Statistics 26*(2), 131–143.

Yukalov, V. I. and D. Sornette (2009, Dec.). Processing Information in Quantum Decision Theory. *Entropy 11*(4), 1073–1120. Special Issue: Entropy and Information.

Zeisberger, S., D. Vrecko, and T. Langer (2012). Measuring the time stability of Prospect Theory preferences. *Theory and Decision 72*(3), 359–386.

Zhang, H. and L. T. Maloney (2012, Jan). Ubiquitous Log Odds: A Common Representation Of Probability and Frequency Distortions in Perception, Action, and Cognition. *Frontiers in Neuroscience 6*(1), 1–14.

Zhou, W.-X. and D. Sornette (2006). Is there a real-estate bubble in the US? *Physica A: Statistical Mechanics and its Applications 361*(1), 297–308.