Heterogeneity in Loss Aversion Estimates across Modeling Approaches

Akhil Abburu
Dept. of Humanities and Social Sciences
Indian Institute of Technology Delhi
New Delhi, India
akhil.abburu@hss.iitd.ac.in

Sumeet Agarwal
Yardi School of AI
Indian Institute of Technology Delhi
New Delhi, India
sumeet@scai.iitd.ac.in

Sumitava Mukherjee
Dept. of Humanities and Social Sciences
Indian Institute of Technology Delhi
New Delhi, India
sm1@hss.iitd.ac.in

Abstract

This study investigates how different modeling approaches affect the measurement of loss aversion, a fundamental concept in psychology and economics. Analyzing 10 datasets comprising over 140,000 trials from 686 participants, we compared four prominent methods: Maximum Likelihood Estimation with Prospect Theory, Bayesian Prospect Theory, Generalized Linear Models, and Drift Diffusion Models. While group-level median loss aversion estimates showed consistency across methods, significant differences emerged at the individual level. The analysis revealed substantial individual-level methodological variability in both the magnitude of loss aversion estimates and participant classification. These findings demonstrate the impact of methodological choices on loss aversion measurement and underscore the need for careful consideration when comparing results across studies using different estimation techniques.

Keywords: loss aversion; decision making; modeling; prospect theory

Introduction

Loss aversion, the robust tendency for individuals to perceive losses as more impactful than equivalent gains, is a cornerstone principle in psychology. Prospect Theory (PT) (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) provides a foundational framework for understanding this phenomenon, positing an asymmetric value function that weighs losses more heavily. While the theory has been highly influential in describing loss aversion, computational modeling has also offered a crucial next step—investigating the cognitive mechanisms that give rise to this phenomenon (Zhao, Walasek, & Bhatia, 2020).

The quantification of loss aversion, however, has been far from uniform. A multitude of estimation techniques and model types have emerged. Traditional approaches often employ frequentist methods, such as maximum likelihood estimation (MLE) within Prospect Theory frameworks, or simpler methods like logistic regression (Tom, Fox, Trepel, & Poldrack, 2007). More recently, Bayesian methods have gained traction, offering benefits in handling parameter uncertainty and incorporating prior knowledge (Lee & Wagenmakers, 2014). Additionally, process-based models, such as Drift Diffusion Models (DDMs) (Ratcliff, 1978), provide a dynamic, mechanistic account of decision-making, allowing for the potential decomposition of loss aversion into distinct cognitive components related to valuation and decision processes (Zhao et al., 2020; Sheng et al., 2020; Singhi, Agarwal, & Mukherjee, 2023).

This methodological pluralism raises a critical concern: Does the choice of estimation approach systematically influence our measurement and understanding of loss aversion? Different methods embody distinct assumptions and may, consequently, yield divergent estimates of key parameters like loss aversion (λ). Such methodological variability could undermine the comparability of findings across studies and hinder a cumulative understanding of loss aversion. Indeed, recent meta-analyses (Brown, Imai, Vieider, & Camerer, 2023; Walasek, Mullett, & Stewart, 2024) have revealed substantial heterogeneity in reported λ values across the literature, with estimates ranging from 1.10 to 2.10. While some of this variation stems from differences in experimental settings, populations, and task designs, the role of estimation methodology itself remains insufficiently understood.

To address this critical gap, the present study undertakes a systematic analysis of prominent modeling approaches for estimating loss aversion. Across 10 datasets encompassing over 140,000 trials, we investigate how methodological choices influence the estimation of λ and assess the consistency of these estimates across different model types. By directly comparing estimates derived from frequentist, Bayesian, and process-based models, we aim to shed light on the extent to which methodological factors contribute to the observed variability in loss aversion measurement.

Methods

To investigate the impact of methodological choices on quantifying loss aversion, we compared four prominent approaches using a common experimental paradigm.

Data

We conducted secondary data analysis on 10 datasets (total N = 686 participants, >140,000 trials) sourced from studies employing a sequential, equiprobable mixed-gamble paradigm. In this paradigm, participants choose between a series of gambles offering a 50% chance of gaining G and a 50% chance of losing L versus a sure option of 0. This is a widely used and well-validated method for eliciting risk preferences and loss aversion (e.g., Tom et al., 2007). We specifically selected datasets utilizing this methodology to minimize variability arising from task differences and isolate the impact of analytical approaches. The datasets varied in sample size, number of trials per participant, and the specific ranges of

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

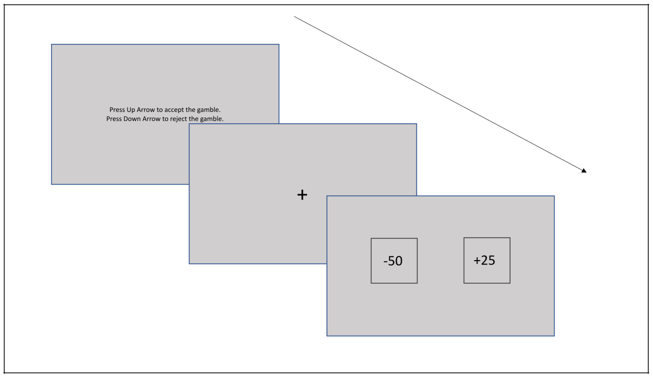

Figure 1: Typical Trial Structure

gain and loss magnitudes used. Table 1 provides a summary of the key characteristics of each dataset. Figure 1 illustrates a typical trial structure faced by participants in these studies.

Modeling approaches

To assess the influence of methodological choices, we compared four computational approaches that are frequently used to estimate loss aversion (λ): PT with Maximum Likelihood Estimation (MLE), Bayesian Prospect Theory, Frequentist Generalized Linear Model (GLM), and Drift Diffusion Models (DDM). As highlighted by Stott (2006), Cumulative Prospect Theory allows for a "functional menagerie" of implementations due to various choices in value functions, weighting functions, and stochastic choice rules. To ensure a focused and interpretable comparison of estimation methodologies, we primarily concentrated on a simplified, 1-parameter PT model in this study.

In this 1-parameter specification, we isolated the loss aversion parameter (λ) by fixing the curvature parameters in the PT value function ($\alpha = \beta = 1$). This simplification, using a linear value function for gains and losses, allowed us to more directly compare estimates of λ across different estimation techniques while maintaining model parsimony. Additionally, this approach also aligns with the fact that GLM and DDM (as implemented here) provide a single loss aversion-like parameter through the ratio of loss and gain coefficients.

While our primary analyses and pairwise comparisons were conducted using this consistent 1-parameter PT framework, the theory can be extended to models with 2 or more parameters. To assess the robustness of our group-level findings, we explored a 2-parameter PT model, by allowing for a curvature parameter (ρ) shared for both gains and losses ($\alpha = \beta = \rho$), at the group level. Our primary analyses focus on the 1-parameter model for comparability.

These models represent a spectrum of methodological sophistication, ranging from traditional value-based models with frequentist estimation to process-based models with Bayesian inference, allowing us to capture a broad range of common analytical strategies. Critically, while each model varies in its underlying assumptions and estimation techniques, our primary focus across all approaches was to derive a comparable measure of loss aversion to directly assess the impact of methodological choice on the quantification of this key parameter. While GLM and DDM intrinsically provide separate coefficients for gains and losses, we specifically focused on the ratio of these coefficients as our measure of λ to ensure comparability with the PT-based models, where λ is directly parameterized within the value function.

Prospect Theory with Maximum Likelihood Estimation (MLE) We based the models on a simplified version of Cumulative Prospect Theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992).

The CPT models decision-making under risk by incorporating a value function and a probability weighting function. For comparability across models and to maintain focus on loss aversion, we concentrate on the value function in this study. We modeled subjective utility using a power function with a parameter for loss aversion and a fixed parameter for risk aversion:

$$v(x) = \begin{cases} x^{\alpha} & \text{if } x \geq 0, \\ -\lambda \cdot x^{\beta} & \text{if } x < 0. \end{cases}$$

where x represents the outcome relative to a reference point (assumed to be zero). λ represents loss aversion, capturing asymmetry in the weighting of losses relative to gains. α and β are the curvature parameters representing the diminishing sensitivity to gains and losses, respectively. As mentioned, we set $\alpha = \beta = 1$ in this study to isolate the effect of λ.

For binary gambles with equal probabilities, the expected utility becomes: $EU = 0.5 \cdot v(G) + v(L)$

Choices were modeled using a softmax function with an inverse temperature parameter (τ):

$P(\text{accept}) = \frac{1}{1+e^{-\tau \cdot \text{EU}}}$

where τ governs the stochasticity of choices (higher τ indicates more deterministic choices). While not a core component of PT's value function, the temperature/bias parameter is a standard addition in applied models to capture the probabilistic nature of human decision-making and helps avoid conflating choice stochasticity with λ.

We employed MLE to fit the CPT models to each participant's choices individually. For each model, we defined a negative log-likelihood function that was minimized using the optim function in R (R Core Team, 2024). The L-BFGS-B optimization algorithm was used with box constraints to ensure parameters remained within plausible ranges (lower bound of .01 and upper bound of 5). Initial parameter guesses for optimization were set to $\lambda = 1.24$ and $\tau = 1$. The optimization was performed with a maximum of 1000 iterations.

Bayesian Prospect Theory Model The Bayesian Prospect Theory model largely mirrors the simplified CPT model described in the MLE section, employing the same core components: a subjective utility function, calculation of expected utility, and a probabilistic choice rule using the softmax function. The key distinction lies in the Bayesian approach to pa-

Table 1: Dataset Characteristics

Dataset name	n_obs	n	gain_low	gain_mean	gain_high	loss_low	loss_mean	loss_high	mean_accept
psh2009	7199	30	2	7	12	-24	-8	-0.5	0.45
psh2015a	5517	37	2	7	12	-24	-8	-0.5	0.34
psh2015b	11217	47	2	7	12	-24	-8	-0.5	0.37
psh2016	28725	120	2	7	12	-24	-8	-0.5	0.24
khan2024_high	21600	108	50	275	500	-500	-275	-50.0	0.43
khan2024_low	20000	100	5	28	50	-50	-28	-5.0	0.47
sheng2020	18800	94	1	6	10	-10	-6	-1.0	0.31
zhao2020_s1	8930	49	10	55	100	-100	-55	-10.0	0.30
zhao2020_s2_hp	10078	52	60	105	150	-100	-55	-10.0	0.50
zhao2020_s2_lp	9192	49	10	55	100	-150	-105	-60.0	0.12

rameter estimation, which allows for incorporating prior beliefs and quantifying uncertainty in parameter estimates.

In the Bayesian framework, we treat the model parameters, loss aversion (λ), and temperature (τ) as random variables with prior distributions. For each participant i, loss aversion λ_i follows a uniform prior distribution, $\lambda_i \sim \text{Uniform}(0,5)$, ensuring it remains positive and within a plausible range. Similarly, the temperature parameter τ_i follows $\tau_i \sim \text{Uniform}(0,10)$, allowing for both noisy and deterministic choice behavior. These uniform priors are uninformative, enabling the data to primarily determine the posterior estimates while maintaining reasonable constraints on parameter values.

The model was implemented in Stan probabilistic programming language (Carpenter et al., 2017). Model fitting utilized the `cmdstanr` R package (Gabry, Češnovar, Johnson, & Bronder, 2024). We assessed MCMC convergence using trace plots and $\hat{R}$ statistics.

Generalized Linear Model (GLM) To provide a comparison with a simpler, widely accessible statistical approach, we employed a Frequentist Generalized Linear Model (GLM). For each participant, we fitted a logistic regression model using the `glm` function, predicting the binary response (gamble accepted or not) from the gain (G) and loss (L) values in each trial. While less directly derived from PT, it serves as a useful benchmark for comparison. The model was specified as:

$$log\left(\frac{P(\text{accept})}{1-P(\text{accept})}\right) = \beta_\text{G} \cdot G + \beta_\text{L} \cdot L$$

From the fitted GLM for each participant, we extracted the coefficients for the intercept, L, and G. Loss aversion was then calculated for each participant as a ratio of the coefficients:

$$\lambda_\text{GLM} = \left|\frac{\beta_\text{L}}{\beta_\text{G}}\right|$$

In the context of a linear logistic regression predicting gamble acceptance, a larger coefficient for loss and a positive coefficient for gain are expected. The ratio of coefficients λ_{GLM} provides an often-used metric for loss aversion, indicating the relative impact of losses compared to the gains.

Drift Diffusion Models To obtain a process-based measure of loss aversion, we utilized the Drift Diffusion Model (DDM). DDM is a prominent computational model that describes decision-making as a process of noisy evidence accumulation towards a decision threshold (Ratcliff, 1978). DDMs are increasingly used to investigate the cognitive dynamics of decision-making and offer a mechanistic perspective beyond static choice models. While the DDM framework allows for the investigation of various cognitive processes (e.g., speed-accuracy trade-offs, response biases), in this study, our primary interest in applying DDM was to derive an estimate of loss aversion (λ_DDM) at the individual level and to assess whether this approach yields different estimates compared to value-based approaches.

In DDM, the decision process is driven by the drift rate (v), which represents the average direction of evidence accumulation. For each gamble, we defined v as a linear function of the gain (G) and loss (L) amounts:

$\nu = \beta_\text{G} \cdot G + \beta_\text{L} \cdot L$

And λ is quantified by:

$$\lambda_\text{DDM} = \left|\frac{\beta_\text{L}}{\beta_\text{G}}\right|$$

λ_DDM provides a process-based measure of loss aversion, reflecting the relative impact of losses compared to gains on the evidence accumulation process underlying choice.

While our primary focus was on loss aversion, DDM also includes other parameters that characterize the decision process. Boundary Separation (a) represents the level of caution in decision-making, with wider boundaries indicating more cautious decisions. Starting Point (z) captures potential biases in the starting point of evidence accumulation. And Ter reflects time spent on processes other than decision-making itself (e.g., perception, motor response).

DDM was implemented using the `hddm` Python package (Wiecki, Sofer, & Frank, 2013). HDDM was chosen for its hierarchical Bayesian approach, which aids robust individual parameter estimation via Markov Chain Monte Carlo

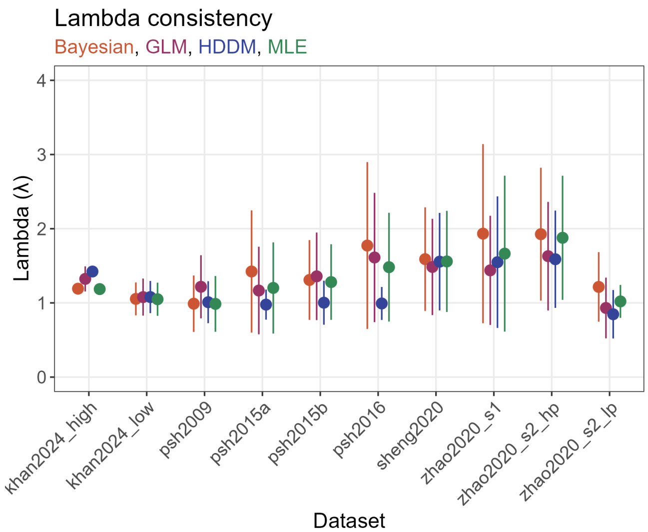

Figure 2: Points represent median loss aversion (λ) estimates for each method within each dataset.

(MCMC) sampling. We used weakly informative priors as implemented by default. MCMC convergence was assessed using trace plots and $\hat{R}$ statistics.

Results and Discussion

We analysed 10 datasets (n = 686), examining both the magnitude and consistency of loss aversion estimates derived from four different computational methods: MLE-Prospect Theory, Bayesian Prospect Theory, GLM, and DDM.

Group-level Analysis

To assess the consistency of loss aversion estimates across methods at the group level, we examined the median lambda values for each method per dataset. Figure 2 displays these median lambda estimates along with their median absolute deviation (MAD) as a measure of variability. The figure suggests a degree of consistency in median lambda values across methods within each dataset, although some variation is apparent, particularly for the HDDM method in some datasets.

To test for differences in group-level median λ estimates across methods, we performed a Kruskal-Wallis test. The test revealed no statistically significant difference in lambda values across the four methods (χ^2 (3) = 3.2254, p = 0.358). Essentially, median λ estimates aggregated at the dataset level find that the choice of estimation method does not significantly impact the central tendency of loss aversion estimates.

To examine the robustness of this finding, we also conducted a separate analysis using the two-parameter PT model for MLE and the Bayesian approach. Results similarly indicated no significant difference (χ^2 (3) = 3.0606, p = 0.382) in λ estimates across methods, even with this alternative model specification. While these group-level tests indicated no significant difference in median lambda values, they do not directly assess the individual lambda estimates across methods.

Individual-Level Variability

We next assessed whether there were significant differences in individual λ estimates across the four methods. Applying a Kruskal-Wallis test to the individual-level lambda values, we found a statistically significant difference in estimates across methods (χ^2 (3) = 36.76, $p < .001$). This indicates that, despite group-level similarities, the choice of method does have a significant impact on individual-level λ estimates.

This pattern of individual-level divergence was robust when using λ estimates from the 2-parameter PT models (allowing curvature ρ estimation) for comparison against GLM and HDDM (Kruskal-Wallis χ^2 (3) = 33.31, $p < .001$). The median absolute difference in λ from 1- to 2-parameter fits was minimal (MLE: 0.12; Bayes: 0.14). A strong correlation between MLE and Bayesian λ estimates also held (ρ = 0.87 with 2-param models).

The median estimated curvature parameter ρ from these 2-parameter fits was 0.50 for MLE and 0.42 for Bayesian models. In both cases, the median ρ was less than 1 (*ps* $< .001$, Wilcoxon signed-rank tests), with approx. 72% (MLE) and 85% (Bayes) of participants exhibiting $\rho < 1$. Within these 2-param models, we found negligible correlation between estimated λ and ρ (e.g., Bayes: $\rho_{\lambda,\rho} = -0.07$), but a weak negative correlation between λ and τ (e.g., Bayes: $\rho_{\lambda,\tau} = -0.19$).

Post-hoc pairwise comparisons using Dunn's test (Holm correction) identified significant differences between methods. Specifically, HDDM showed different medians compared to Bayesian ($p < 0.001$), GLM ($p < .001$), and MLE ($p < .01$). Additionally, the Bayesian model showed significantly different medians compared to MLE ($p < .05$). However, no significant differences were found between Bayesian and GLM (p = 0.2), or between GLM and MLE (p = 0.3).

To investigate whether differences persist even when excluding HDDM—which stands out due to its hierarchical nature—we repeated the Kruskal-Wallis and Dunn's tests, removing HDDM estimates from the analysis. Despite the exclusion, the test still indicated a significant difference in individual λ estimates across the remaining three methods (χ^2 (2) = 7.17, p = .028). These results suggest that while HDDM contributes to the variability, significant differences in λ estimates can also arise even among the value-based methods.

To examine inter-method consistency at the individual level and assess rank-order agreement, we calculated Spearman rank correlations between λ estimates derived from each pair of methods. The analysis revealed a near-perfect correlation between λ estimates from the 1-parameter MLE and Bayesian PT models (ρ = 0.995), indicating that these two methods rank individuals almost identically despite the minor distributional differences noted earlier. Consequently, discrepancies in classification between these are likely minimal, primarily affecting individuals with λ estimates close to one.

In contrast, correlations involving other methods were lower. The PT models (MLE/Bayes) showed moderate correlations with GLM (with MLE, $\rho = 0.688$; Bayes, $\rho = 0.682$). Correlations with HDDM were weaker (with MLE,

ρ = 0.404; with Bayes, ρ = 0.396). GLM and HDDM showed a moderate correlation (ρ = 0.561). This varying rank-order consistency, especially with HDDM, may explain the substantial re-ranking and classification switching. Thus, the observed differences are not just in the central tendency or spread, but in the relative ordering of individuals, suggesting methods may capture distinct facets of loss aversion, particularly value-based versus process-based (DDM) approaches.

To further explore the nature of these differences and to partition the sources of variability, we employed a Bayesian mixed-effects model. The model estimated the overall average λ across all datasets and methods to be approximately 1.54 (95% Credible Interval: 1.12 to 1.96), confirming a general tendency towards loss aversion. Critically, the model also allowed us to partition the variance in λ estimates. The estimated SD of dataset-specific intercepts was 0.40 (95% CI: 0.24 to 0.70), indicating substantial heterogeneity in baseline lambda levels across different datasets. This suggests that factors intrinsic to each dataset, such as specific participant populations or task parameter ranges, contribute significantly.

The SD of method-specific intercepts was estimated at 0.30 (95% CI: 0.10 to 0.91), revealing a degree of variation in baseline λ levels attributable to the choice of estimation method, even after accounting for dataset-level differences. While the credible intervals for dataset and method variance overlap, the point estimate for dataset variability is slightly larger, suggesting that dataset characteristics may exert a somewhat stronger influence on the variability than method alone, although both factors contribute demonstrably.

There are a few potential sources of this variation. The hierarchical approach of DDM, for instance, shrinks individual estimates toward group-level priors, potentially attenuating extreme λ values compared to non-hierarchical methods. GLM's linear approximation of prospect theory's nonlinear value function may lead to differing estimates from the MLE. Bayesian methods, which account for parameter uncertainty, may produce more conservative estimates than frequentist MLE in cases of limited data. These methodological divergences highlight how model assumptions (e.g., hierarchical structure, linearity) shape λ estimation, even when group-level medians appear similar. Given these observed differences in λ magnitudes and the potential influence of modeling, a critical next step is to examine the practical consequences for classifying individuals as loss-averse.

Individual Variation and Classification of Loss Aversion

Building on the observed individual variability in λ estimates, we next examined the consequences of methodological choice for classifying participants as loss averse. Specifically, we tested whether the proportion of participants classified as loss averse ($\lambda > 1$) differed significantly across the four methods. To do this, we conducted a Chi-Squared Test of Homogeneity, using a contingency table summarizing the counts of loss-averse and non-loss-averse classifications for each method, which revealed significant differences (χ^2 (3) = 22.76, $p < .001$). This indicates that the choice of method

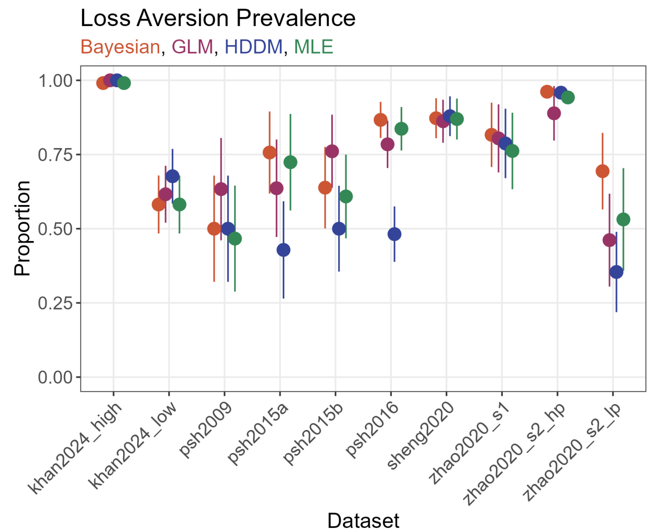

Figure 3: Proportion of Participants with $\lambda > 1$ by Method

systematically influences the estimated prevalence of loss aversion in our sample. Figure 3 shows the proportions of participants classified as loss averse by each method.

Notably, averaging across methods, the proportion of participants not classified as loss averse ($\lambda < 1$) varied substantially across datasets, ranging from 0.5% in *khan2024_high* to 48.8% in *zhao2020_s2_lp*. This variability aligns with prior evidence that loss-aversion is dependent on stake size (Mukherjee, Sahay, Pammi, & Srinivasan, 2017). For instance, figure 4 shows that datasets with smaller stakes (e.g., *sheng2020*: gains/losses of 1-10) exhibited a higher prevalence of non-loss-averse participants (12.9%) compared to high-stakes datasets (e.g., *khan2024_high*: 50-500; 0.5%).

Individual Lambda Variation across Methods

To quantify the magnitude of within-participant variation, we calculated the range of λ estimates for each participant (i.e., the difference between the max and min λ value obtained across the four methods). Across all participants, the mean range was 0.76, with a 95% CI of [0.696, 0.833]. This suggests a non-negligible degree of within-participant variability.

Perhaps more importantly, from a practical perspective, we assessed whether methodological choice leads to changes in loss aversion classification at the individual level. We identified participants for whom the classification as loss-averse ($\lambda > 1$) or not ($\lambda \leq 1$) varied depending on the method used. These are highlighted in red in Figure 5. We found that approximately 30% of participants were classified as loss-averse by some methods but not by others. For a substantial proportion of individuals, the seemingly binary question of whether they exhibit loss aversion or not is contingent on the methodological approach. This individual-level analysis underscores the impact of this choice, emphasizing the need for caution when interpreting and comparing loss aversion estimates across studies employing different methodologies.

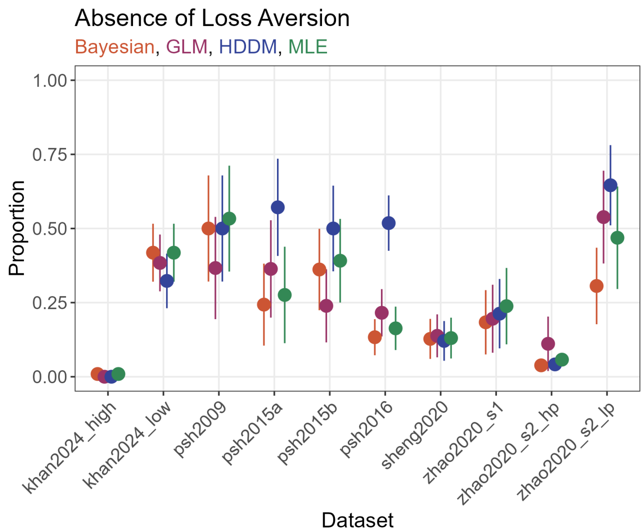

Figure 4: Proportion of Participants Classified as Not Loss Averse ($\lambda \leq 1$) by Method

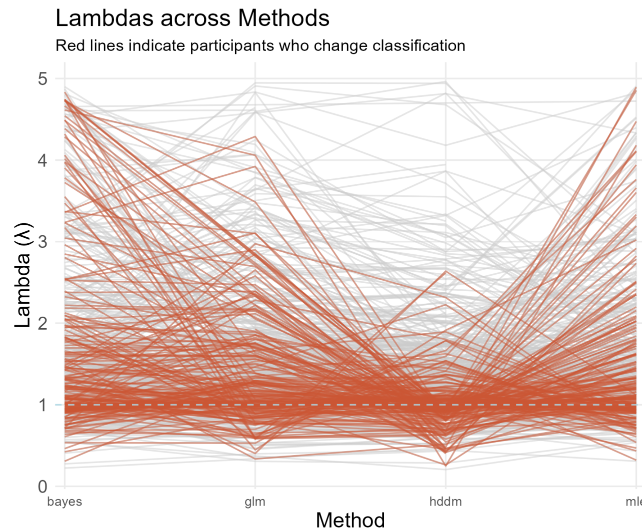

Figure 5: Individual Participant λ Values Across Methods.

Limitations

While this study seeks to provide novel insights into the impact of methodological choices on loss aversion measurement, it is important to acknowledge several limitations. First, our analysis focused primarily on the loss aversion parameter (λ) as the key metric of comparison. Our primary analysis fixed PT curvature parameters to 1. While supplementary 2-parameter PT analyses showed robust core findings and limited $\lambda-\rho$ trade-offs, fully exploring models with more free parameters across all methods remains future work.

Second, our analysis was based on datasets using a specific paradigm. While this is widely used, there are criticisms (Walasek & Stewart, 2021), and the generalisability to other tasks or elicitation methods warrants further investigation.

Third, while we compared λ from PT with GLM and DDM analogues, these measures stem from different theoretical assumptions about the underlying decision processes. Future simulation studies, beyond the scope of this paper, could help establish quantitative relationships between parameters recovered by these distinct model frameworks.

Finally, our DDM analysis employed the hierarchical Bayesian package `hddm`, chosen for its ability to enhance parameter stability, particularly when individual-level data across datasets is limited or noisy. Comparing with non-hierarchical DDM fits could reveal further variability. However, it is important to note that methodological variability and its practical consequences, such as the classification switching of participants as loss averse or not, persisted even when HDDM was excluded from our analyses. This suggests that while the hierarchical approach is a relevant factor, the differences are not solely attributable to it. Despite these limitations, our study provides evidence for the meaningful impact of estimation choices on loss aversion measurement, highlighting the need for methodological rigor and transparency in future research within this domain.

Conclusion

We systematically examined the influence of methodological choices on loss aversion estimation by comparing four approaches across ten datasets. The findings reveal that while median loss aversion estimates are consistent across methods, significant differences emerge at the individual level.

Individual-level analyses demonstrated that the choice of estimation method significantly influences the distribution of λ estimates and contributes to overall variability. We found that the estimated prevalence of loss aversion varied across methods, and a considerable proportion of participants were classified differently depending on the method applied. This underscores a broader challenge in computational modeling: parameters such as λ, while offering quantification, may not consistently represent the same underlying construct across modeling approaches, especially when theories guiding the model construction allow for flexibility (see Eronen & Bringmann, 2021, for a broader discussion). This, in turn, can constrain the direct contribution to cumulative theory building.

These method-driven differences have important practical implications that extend beyond statistical variations. The estimated prevalence of loss aversion, a key metric in psychology and economics, varies depending on the method used, and individual participant classification as loss averse or not is contingent on methodological choice for a substantial proportion of individuals. These results suggest that caution is needed when interpreting and comparing loss aversion estimates across studies using different methods. Such methodological variations potentially represent a significant source of heterogeneity and could be a source of inconsistency hindering a unified understanding. A greater awareness of method-specific influences and limitations will be vital for advancing the field and ensuring validity.

Acknowledgments

We would like to thank the reviewers and the meta-reviewer for their helpful comments. SA is supported by funding from the Yardi School of AI, IIT Delhi. The authors thank the IIT Delhi HPC facility for computational resources.

References

Brown, A. L., Imai, T., Vieider, F., & Camerer, C. F. (2023). Meta-analysis of empirical estimates of loss aversion. *Journal of Economic Literature*.

Eronen, M. I., & Bringmann, L. F. (n.d.). The Theory Crisis in Psychology: How to Move Forward. *Perspectives on Psychological Science*, *16*(4), 779–788. doi: 10.1177/1745691620970586

Gabry, J., Češnovar, R., Johnson, A., & Bronder, S. (2024). *Cmdstanr: R Interface to 'CmdStan'*.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291. doi: 10.2307/1914185

Khan, O., Mukherjee, S., Srinivasan, N., & Abburu, A. (2024). *Role of Magnitude in Loss Aversion.* OSF. doi: 10.17605/OSF.IO/4CHP2

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge, United Kingdom New York: Cambridge University Press. doi: 10.1017/CBO9781139087759

Mukherjee, S., Sahay, A., Pammi, V. S. C., & Srinivasan, N. (2017, January). Is loss-aversion magnitude-dependent? Measuring prospective affective judgments regarding gains and losses. *Judgment and Decision Making*, *12*(1), 81–89. doi: 10.1017/S1930297500005258

Pan, W., Geng, H., Zhang, L., Fengler, A., Frank, M., Zhang, R.-Y., & Chuan-Peng, H. (2022, November). *dockerHDDM: A user-friendly environment for Bayesian Hierarchical Drift-Diffusion Modeling.* PsyArXiv. doi: 10.31234/osf.io/6uzga

R Core Team. (2024). *R: A language and Environment for Statistical Computing.* R Foundation for Statistical Computing. Vienna, Austria.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological review*, *85*(2), 59.

Sheng, F., Ramakrishnan, A., Seok, D., Zhao, W. J., Thelaus, S., Cen, P., & Platt, M. L. (2020, May). Decomposing loss aversion from gaze allocation and pupil dilation. *Proceedings of the National Academy of Sciences*, *117*(21), 11356–11363. doi: 10.1073/pnas.1919670117

Singhi, N., Agarwal, S., & Mukherjee, S. (2023). Using Computational Models to Understand the Role and Nature of Valuation Bias in Mixed Gambles. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *45*(45).

Sokol-Hessner, P., Hartley, C. A., Hamilton, J. R., & Phelps, E. A. (2015, June). Interoceptive ability predicts aversion to losses. *Cognition & Emotion*, *29*(4), 695–701. doi: 10.1080/02699931.2014.925426

Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., & Phelps, E. A. (2009, March). Thinking like a trader selectively reduces individuals' loss aversion. *Proceedings of the National Academy of Sciences*, *106*(13), 5035–5040. doi: 10.1073/pnas.0806761106

Sokol-Hessner, P., Lackovic, S. F., Tobe, R. H., Camerer, C. F., Leventhal, B. L., & Phelps, E. A. (2015, July). Determinants of Propranolol's Selective Effect on Loss Aversion. *Psychological Science*, *26*(7), 1123–1130. doi: 10.1177/0956797615582026

Sokol-Hessner, P., Raio, C. M., Gottesman, S. P., Lackovic, S. F., & Phelps, E. A. (2016, December). Acute stress does not affect risky monetary decision-making. *Neurobiology of Stress*, *5*, 19–25. doi: 10.1016/j.ynstr.2016.10.003

Stott, H. P. (2006, March). Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty*, *32*(2), 101–130. doi: 10.1007/s11166-006-8289-6

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007, January). The Neural Basis of Loss Aversion in Decision-Making Under Risk. *Science*, *315*(5811), 515–518. doi: 10.1126/science.1134239

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. doi: 10.1007/BF00122574

Walasek, L., Mullett, T. L., & Stewart, N. (2024, August). A meta-analysis of loss aversion in risky contexts. *Journal of Economic Psychology*, *103*, 102740. doi: 10.1016/j.joep.2024.102740

Walasek, L., & Stewart, N. (2021, January). You cannot accurately estimate an individual's loss aversion using an accept–reject task. *Decision*, *8*(1), 2–15. doi: 10.1037/dec0000141

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. doi: 10.21105/joss.01686

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in Neuroinformatics*, *7*. doi: 10.3389/fninf.2013.00014

Zhao, W. J., Walasek, L., & Bhatia, S. (2020, December). Psychological mechanisms of loss aversion: A drift-diffusion decomposition. *Cognitive Psychology*, *123*, 101331. doi: 10.1016/j.cogpsych.2020.101331