Title
Choosers Adapt Value Coding to the Environment, But Do Not Attain Efficiency

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Kurtz-David, Vered
Alladi, Vinayak
Bucher, Stefan
et al.

Publication Date
2023

Peer reviewed

Choosers Adapt Value Coding to the Environment, But Do Not Attain Efficiency

Vered Kurtz-David (kurtzv02@nyu.edu), Kenway Louie (kl837@nyu.edu), Ambuj Dewan (ambuj.dewan@gmail.com) & Paul W. Glimcher (paulg@nyu.edu)
Grossman School of Medicine, New York University, 435 E 30th Street
New York, NY 10016 USA

Vinayak Alladi (v.alladi@gmail.com) & Agnieszka Tymula (agnieszka.tymula@sydney.edu.au)
School of Economics, University of Sydney, A02 – Social Sciences Building
Sydney, NSW 2006 Australia

Stefan Bucher (stefan.bucher@tuebingen.mpg.de)
Tübingen AI Center and Max Planck Institute for Biological Cybernetics, Maria-von-Linden-Straße 6,
Tübingen 72076 Germany

Adam Brandenburger (adam.brandenburger@stern.nyu.edu)
Stern School of Business, NYU Tandon, and NYU Shanghai, New York University, 44 West 4th Street
New York, NY 10012 USA

Abstract

We investigate how human choosers adapt their value encoding strategy to the statistics of the choice environment. Specifically, we ask whether the human value encoding mechanism exhibits divisive normalization only in the Pareto-distributed environments in which it is information-maximizing. To test this theory, we conduct a risky choice experiment in which subjects are presented with blocks of choice stimuli drawn from either a Pareto-distributed environment or a uniform-distributed environment. Our results show that subjects exhibit some degree of normalization regardless of whether it is efficient or not, but do adapt the curvature of their encoding function to the environment. These findings suggest that human value coding mechanisms are flexible but biologically constrained to be perfectly efficient only in specific environments. This study provides new insights into the neural mechanism of human decision-making and the role of environmental statistics in shaping it.

Keywords: efficient coding; divisive normalization; stochastic choice; neuroeconomics; risky choice

Introduction

The divisive normalization (DN) encoding function is often viewed as a canonical encoding mechanism (Carandini & Heeger, 2012) and has been implicated in a wide array of cognitive functions from visual stimuli (Carandini et al., 1997; Heeger, 1992), through auditory (Schwartz & Simoncelli, 2000), olfactory (Olsen et al., 2010) sensory processing, and even in high cognitive functions, such as value and reward representation (Louie et al., 2011; Webb et al., 2021). The main advantage of this encoding mechanism is that it allows an informationally-limited cognitive system to efficiently encode input stimuli from naturalistic environments (Heeger, 1992), and it is thus often considered a form of efficient code (Schwartz & Simoncelli, 2001; Steverson et al., 2019). Furthermore, the divisive formulation allows for scale-invariance in the representation of natural environments (Chater & Brown, 1999; Kello et al., 2010).

DN maximizes mutual information, and is therefore efficient, only for environments characterized by certain classes of input stimulus distributions (Bucher & Brandenburger, 2022). Therefore, different optimization criteria and stimuli distributions (Heng et al., 2020; Steverson et al., 2019), may imply the implementation of other encoding mechanisms (Bhui & Gershman, 20181115; Heng et al., 2020). Whether human behavior relies on a divisive form of normalization regardless of the environment, or whether it only arises in environments in which it is efficient, remains unknown.

Here, using behavioral experiments and computational modeling, we asked whether humans use a DN-like mechanism to encode the values of rewards when making choices among classical economic lotteries. Similarly to Kutzner et al. (2017), we manipulate the statistics of the choice environment. Subjects made choices when these lotteries were drawn from environmental distributions for which DN is provably efficient, and for environmental distributions for which it is inefficient. This allowed us to ask if choices are constrained by an encoding mechanism that can only operate using some divisive form, or whether the encoding mechanism adapts across different choice environments to whatever form is maximally efficient (Figure 1A).

Our main findings suggest that subjects exhibit a divisive form of normalization both when it is, *and when it is not*, efficient. At least over the timescale we examined, subjects employed normalized value encoding regardless of the distribution of the input stimuli they were facing. Nonetheless, the precise curvature of the DN function did

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

vary as a function of the environment. This indicates that subjects do adapt to the different choice environments, although that adaptation appears constrained by the divisive mechanism even when that is not an efficient strategy.

Methods

Experimental Design

The task consisted of two stages, which took place on separate days.

BDM Task (STAGE I). To eliminate the effect that risk preferences may induce on subjects' choices, we aimed to generate the environmental statistics in the expected utility – or subjective-value (SV) – space (as opposed to, say, the monetary prizes). Thus, we first utilized a bidding experiment in STAGE I for eliciting subjects' risk preferences. STAGE I comprised 33 bid trials, in which subjects reported their willingness to pay to participate in a lottery. On each bid trial, subjects were presented with a visualization of a 50-50 lottery on the computer screen and had to type in their willingness to pay to participate in it in a dollar amount (Figure 1B). For each lottery, the bid could range between its minimal and maximal payoff in $0.1 increments. All subjects completed the same 33 trials in an order randomized at the subject level. At the end of the session, the realization of a randomly selected bid trial was implemented for payment, using a Becker–DeGroot–Marschak (BDM) auction, a mechanism designed to elicit an individual's truthful SV for an item (Becker, DeGroot & Marschak 1964).

Choice Task (STAGE II). In STAGE II of the study, our aim was to test whether the distribution of lottery valuations in the choice environment affected choice mechanisms. Subjects were asked to choose the 50-50 lottery they preferred to play from two available options that varied from trial-to-trial. Lottery payoffs ranged between $0 and $60 in $0.1 increments. Overall, subjects made 640 binary choices that were divided into two blocks of 320 trials each and presented on subsequent days. Our experimental manipulation was that in each block, the lotteries were drawn either from a uniform distribution (all valuations equally likely) or from a Pareto type III distribution of SVs for which DN is efficient-code (Bucher & Brandenburger, 2022). The order of the treatments was counter-balanced across subjects. One trial was randomly selected for payment at the end of each experimental session.[1]

Sessions. Experimental sessions were carried out online via Zoom while subjects completed the task on a website. We ran eight sessions of the experiment between May 2022 and August 2022. After instruction, subjects had to successfully answer a set of comprehension questions about the procedure before starting STAGE I. They could participate in STAGE II of the study only if they completed all trials in STAGE I.

Subjects received all payments after completing both STAGE I and STAGE II. Subjects received a $10 participation fee and on average $24.5 in STAGE I (range $0-60) and $76.02 in STAGE II (range $7.3-120)[1] from the decision task. All parts of the experiment were self-paced. Both the BDM and the choice tasks were programmed in the *oTree* software package (Chen et al., 2016).

Participants. We recruited 130 participants from various departments at the University of Sydney. Subjects gave informed written consent before participating in the study, which was approved by the local ethics committee at the University of Sydney. Fourteen subjects failed comprehension questions and were dropped from the study. Twenty-eight participants did not show up to STAGE II, and thus were also dropped from the study. Five additional subjects started STAGE II but decided to drop out. Hence, we report the results from the remaining 83 participants (49 females, mean age=21.8, std: 3.34, range: 18-30).

Risk Preferences Estimation. Following STAGE I, we used subjects' bids to estimate their risk preferences using a power utility function:

$$(1)\quad SV_t = E\left[V_t^{\rho}\right], \qquad 0 \le \rho,$$

such that the subjective valuation for a lottery in trial $t = 1, \dots, 33$ is given by the expected utility of the lottery. If the risk curvature parameter $\rho < 1$, this indicates that the subject is risk-averse. When $\rho = 1$, the subject is risk-neutral. If $\rho > 1$, the subject is risk-seeking.

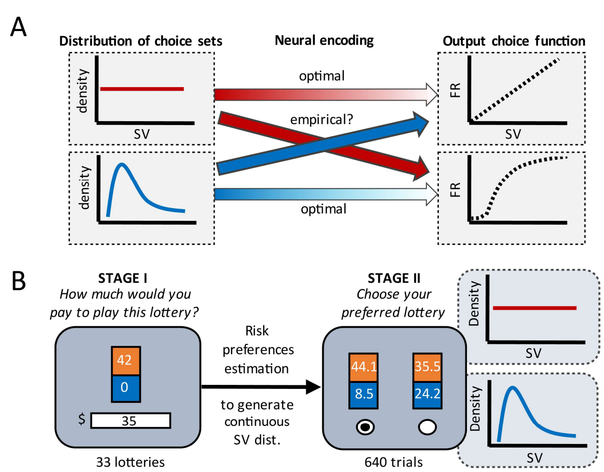

Figure 1: (A) Research question. (B) Experimental design.

[1] Subjects also faced an additional 640-trials with six-option choice sets. Thus, in each environment subjects encountered two 320 choice blocks presenting two-option and six-option conditions. The six-option blocks were designed to examine another research question that goes beyond the scope of the current study and will be reported in a separate paper. Blocks were presented in random order across subjects but on a given day, all blocks were drawn from the same distribution. Payments for STAGE II also included a realization of one choice from the six-options sets.

We used a non-linear least-squares (nls) regression to estimate the ρ parameter separately for each subject, using the following specification –

$$(2)\quad bid_t = \left(0.5\left(x_{1,t}^{\rho} + x_{2,t}^{\rho}\right)\right)^{\frac{1}{\rho}},$$

where bid_t is a subject's bid in trial t (the certainty equivalent of the lottery in trial t), and $x_{1,t}$ and $x_{2,t}$ are the lottery payoffs. Subjects in our sample exhibited heterogeneity in their risk preferences, as evident in the distribution of estimated ρ's in Figure 2.

Using estimated ρ's, we could compute the SV for any combination of x, which allowed us to generate sets of lotteries whose implied SV distributions matched our target distributions (see below), regardless of individual differences in risk attitudes. This step was crucial for STAGE II of the study, where we aimed to control the environmental statistics.

Generating Uniform Distributions of SVs. For each subject, we computed the upper bound of the distribution as the SV of the maximal possible winning amount in the study, which was \$60 (i.e., $\boldsymbol{sv^{max} = 60^{\rho}}$). We then divided the range $[\boldsymbol{0, sv^{max}}]$ into 40 equally-spaced SV increments. For each of the increments, we created eight different lotteries, which would give the subject this exact subjective value (for a total of 320 lotteries). Since the joint distribution of a two-dimensional uniform distribution is independent, hence determined by its marginals, it follows that we could replicate and randomize the SVs-grid twice for generating binary choice sets.

Generating Pareto Type III Distributions of SVs. The DN encoding function is information-maximizing for a bivariate Pareto distribution with a joint pdf (Bucher & Brandenburger, 2022; Eq. 7 with $\boldsymbol{\mu_i = 0}$)[2]:

$$(3)\quad f_{sv}(sv_1, sv_2) = \beta^2 \frac{2\left(\prod_{i=1}^{2}\frac{1}{\sigma_i}\left(\frac{sv_i}{\sigma_i}\right)^{\beta-1}\right)}{\left(1+\sum_{i=1}^{2}\left(\frac{sv_i}{\sigma_i}\right)^{\beta}\right)^{3}},$$

and the marginal pdf being a univariate Pareto type III pdf:

$$(4)\quad f_{sv_i} = \beta \frac{\frac{1}{\sigma_i}\left(\frac{sv_i}{\sigma_i}\right)^{\beta-1}}{\left(1+\left(\frac{sv_i}{\sigma_i}\right)^{\beta}\right)^{2}},$$

where $i \in \{1,2\}$ indicates the dimension of interest, and the number of choice options in the choice sets ($n = 2$) determines the dimension of the distribution. $\beta > 0$ is a shape parameter, while σ_i are scale parameters. We set $\mu_i = 0$, and based on previous empirical estimates (Webb et al., 2021),

[2] We set the location parameter $\mu_i = 0$ to match the lower bound of the uniform distribution, and to avoid negative valuations.

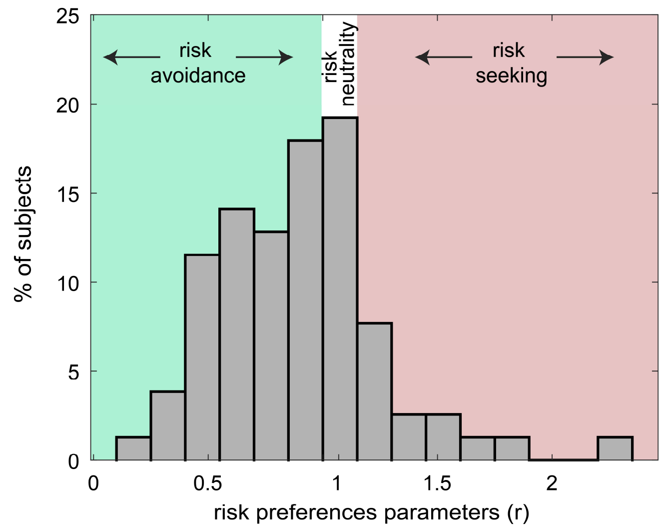

Figure 2. Distribution of the ρ parameter, capturing elicited risk preferences from STAGE I of the study.

we set $\beta = 3$. Note that for this parametrization, the correlation coefficient across dimensions is analytically given (Bucher & Brandenburger, 2022; eq. 11) and equal to 0.7049. Scale parameters σ_i were chosen at the subject-level in a manner such that the conditional expectation of the Pareto distribution, which is given by (Bucher & Brandenburger, 2022):

$$(5)\quad E\left(sv_i \middle| sv_j\right) = \sigma_i\left[1 + \left(\frac{sv_j}{\sigma_j}\right)^{\beta}\right]^{1/\beta} \frac{\Gamma(2-\frac{1}{\beta})\Gamma(\frac{\beta+1}{\beta})}{\Gamma(2)},$$

would match the expectation of the uniform distribution, where Γ indicates the gamma function.

Following Proposition 4 of Bucher & Brandenburger (2022), as well as Arnold (Arnold, 2015), and using the subject-specific parameterization, we generated the Pareto type III distributions as a scale mixture of transformed exponential (or Weibull) random variables, so that:

$$(6)\quad sv_i = \sigma_i\left(\frac{U_i}{Z}\right)^{\frac{1}{\beta}}, \qquad for\ i = 1, \ldots n,$$

where $U_i \sim Exp(\lambda = 1)$ and $Z \sim Exp(\lambda = 1)$ independently of all U_i. Note that small 320-draws experimental sets lead to under-sampling of the distributions. Therefore, to fully capture the shape of the distribution, for each subject, we first generated joint Pareto distributions with 100K draws. We then created small 600-draw experimental distributions that matched the large 100k-draw distributions, allowing a deviation of up to 0.2 utils from the actual first and second moments (mean and standard deviation) of the large 100k-draws sets. Finally, we truncated the long tail of the Pareto type III distributions at sv^{max} (eliminating 6.5 to 23.83

percent of the distribution, depending on the risk parameter ρ), to match the upper bound of the uniform distribution and to avoid extreme reward amounts. We then cast 320 SVs at random from the remaining valuations, which constituted the experimental subject-specific Pareto distributions.

Generating Lotteries from SV Distributions. The final step in our design was to generate lotteries in dollar amounts from the SV distributions (which were created in utility space). For each lottery $\boldsymbol{i = 1, 2}$ in trial $\boldsymbol{t = 1, \ldots, 320}$ with a given $\boldsymbol{sv_{i,t}}$, valuation, $\boldsymbol{x_{1_{i,t}}}$ was randomly drawn uniformly in the range $[\boldsymbol{0, 60}]$ with 0.1 increments. We then solved for $\boldsymbol{x_{2_{i,t}}}$, giving rise to the desired $\boldsymbol{sv_{i,t}}$, and rounded to one decimal place, using the following equation:

$$(7)\quad x_{2_{i,t}} = \left(2sv_{i,t} - x_{1_{i,t}}^{\rho}\right)^{\frac{1}{\rho}}.$$

In practice, we had to restrict the upper range of x_1's to $\left(2sv_{i,t}\right)^{\frac{1}{\rho}}$ to avoid negative values within the parentheses in eq. 7, and then determined the maximal value of $x_{1_{i,t}}$ using the minimum function:

$$(8)\quad x_{1_{i,t}}^{max} = \min\{\left(2sv_{i,t}\right)^{\frac{1}{\rho}}, 60\}.$$

We restricted the share of trials with first-order stochastic dominance (FOSD) (trials on which both amounts of one lottery were higher or equal to the other lottery) to 45 percent. For subjects with $\rho \to 0$, we could not generate experimental sets with only 45 percent of the trials taking this easy form. Therefore, we fixed ρ=1, for all subjects with $\rho < 0.1$ (a total of 6 subjects), limiting the interoperability of data from this small number of subjects.

Model Fitting

DN models. To test whether subjects employed normalization under both experimental treatments, we considered a class of divisive functions, which capture both temporal and spatial contexts of choice (Khaw et al., 2017; Schwartz et al., 2007). All functions share the following general formulation:

$$(9)\quad z_i(sv) = \frac{sv_i^{\alpha}}{\delta + D},$$

such that z_i is the normalized value of the subjective valuation sv_i. The parameter $\alpha \geq 0$ is a curvature parameter, while the parameter D$\in \mathbb{R}$ is the normalization factor, which could be either temporal, spatial, or both (see below). δ is the so-called "semi-saturation parameter" (Heeger, 1992).

We estimated twelve specifications of eq. 9, which varied by their definition of the normalization factor D, and then divided into two groups, each comprising six different models. In Group 1, we set $\alpha = 1$, assuming no additional curvature to the function. In Group 2, we estimated α as a free parameter. In the econometric estimation, a coefficient $\omega \in [0,1]$ determines the degree of the normalization. When $\omega \to 0$, there is no normalization. Within each group, in models I-II, we focus on spatial normalization. In model I, we estimated symmetric ω weights and in model II, we estimated asymmetric weights across the choice set elements within a trial, sv_i and sv_j:

$$\text{I.}\quad z_i(sv) = \frac{sv_i^{\alpha}}{\delta + \omega(sv_i^{\alpha} + sv_j^{\alpha})}, \forall i \neq j,$$

$$\text{II.}\quad z_i(sv) = \frac{sv_i^{\alpha}}{\delta + \omega_i(sv_i^{\alpha}) + \omega_j(sv_j^{\alpha})}, \forall i \neq j,$$

In model III, we relaxed the spatial normalization in the denominator, and solely examine temporal normalization (Khaw et al., 2017):

$$\text{III.}\quad z_i(sv) = \frac{sv_i^{\alpha}}{\delta + \omega_t(M^{\alpha})}.$$

The parameter M captures the temporal normalization across trials (Khaw et al., 2017). We do not estimate M, instead plugging-in the median of each distribution, which was uniquely defined for each subject and distribution.

In models IV-V, we tested both spatial and temporal normalizations, with symmetric and asymmetric weights on the denominator:

$$\text{IV.}\quad z_i(sv) = \frac{sv_i^{\alpha}}{\delta + \omega(M^{\alpha} + sv_i^{\alpha} + sv_j^{\alpha})}, \forall i \neq j,$$

$$\text{V.}\quad z_i(sv) = \frac{sv_i^{\alpha}}{\delta + \omega_t(M^{\alpha}) + \omega_i(sv_i^{\alpha}) + \omega_j(sv_j^{\alpha})}, \forall i \neq j,$$

Finally, in model VI, we also directly estimated a classical random utility model (RUM), which precludes any type of normalization, where:

$$\text{VI.}\quad z_i(sv) = \frac{sv_i^{\alpha}}{\delta}.$$

Maximum Likelihood Estimation of the DN Function. Models I-VI were estimated via a maximum likelihood procedure (MLE). Our main goal was to test whether the $\boldsymbol{\omega}$ parameters were significant, implying normalization, under both experimental treatments, and across several specifications of the divisive model.

Table 1: DN pooled estimates, Group 1, $\alpha = 1$. Up: Pareto distribution, bottom: uniform distribution. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001. $N_{uniform}$ = 24,527. N_{Pareto} = 24,405.

Group 1	(I)	(II)	(III)	(IV)	(V)	(VI)
			Pareto			
δ	-0.0781***	-0.0791***	-0.2555**	-0.1799***	-0.2539**	202.9793
ω	0.1168***			0.0785***		
ω_i		0.1257***			0.0257	
ω_j		0.1080***			0.0246	
ω_t			0.2471***		0.1927***	
BIC	20699.1501	20704.4042	20111.2270	20268.0400	20088.8229	33418.2735
			Uniform			
δ	-0.2211***	-0.2211***	-0.4191***	---	-0.4319***	62.5630
ω	0.1311***			---		
ω_i		0.1313***		---	-0.0342+	
ω_j		0.1308***		---	-0.0314	
ω_t			0.2361***	---	0.3046***	
BIC	15070.0649	15080.1609	13724.2675		13696.8862	31085.6892

Table 2: DN pooled estimates, Group 2, α is a free parameter. Up: Pareto distribution, bottom: uniform distribution. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001. $N_{uniform}$ = 24,527. N_{Pareto} = 24,405.

Group 2	(I)	(II)	(III)	(IV)	(V)	(VI)
			Pareto			
α	5.4073***	5.4151***	0.7377***	2.0270***	5.5053***	----
δ	0.0832	0.0829	-0.1697*	-0.5181**	0.0872	---
ω	0.3682***			0.1337***		
ω_i		0.3784***			0.3813***	
ω_j		0.3585***			0.3613***	
ω_t			0.1919***		-0.0000	
BIC	20222.0127	20252.5092	20040.7668	20126.495	20230.0258	
			Uniform			
α	9.7419***	9.7388***	0.9267***	---	---	-0.0556
δ	0.3901	0.3908	-0.3816***	---	---	-0.0126
ω	0.3400***					
ω_i		0.3411***		---	---	
ω_j		0.3389***			---	
ω_t			0.2254***		---	
BIC	13430.1533	13440.167	13728.6665			15889.877

For this aim, we imposed stochasticity in the choice process (Webb, 2019; Webb et al., 2021) with an additive noise term η_i, drawn from a Gumbel distribution:

$$(10)\, u_i = z_i(sv) + \eta_i.$$

Subjects chose the option with the highest utility u_i:

$$(11)\, u_i > u_j \leftrightarrow z_i(sv) + \eta_i > z_j(sv) + \eta_j,\ \forall i \neq j.$$

The probability of choosing option i is given by the logit choice probability, denoted by $P_i(z(sv))$.

We estimated $P_i(z(sv))$ for subjects' aggregated choice data via a logistic function with MLE (Harrison, 2008), clustering standard errors on the subject level. Thus, in our estimation, subjects are treated as one representative decision-maker. Since specifications with spatial normalizations are defined for all $\forall i \neq j$, we omitted trials in which $sv_i = sv_j$. We evaluated goodness-of-fit using the Bayesian information criterion (BIC).

Finally, to test whether the statistical environment had an effect on subjects' coding mechanisms, we estimated another set of regressions with random effects (RE) by distributions for α, the functional "predisposition" curvature (Glimcher & Tymula, 2023).

Results

Tables 1-2 summarize the pooled estimates for Groups 1 and 2 (respectively). For some specifications, the MLE algorithm did not converge. There, we report null findings. In all, we report a total of seventeen regressions of DN models and additional three regressions of the classical Random Utility model (RUM) (McFadden, 2001).

Our results show that across all divisive models, regardless of the exact specification examined, the normalization coefficients (ω's) are significant and positive ($p<0.001$), indicating normalized value coding. This is true for both the Pareto and uniform distributions. The degree of normalization is of the same magnitude in both environments.

Next, we turn to look at model fits for the classical RUM. We find that model parameters are not significant, and that goodness-of-fit measures (BIC scores) of all the DN models are substantially lower (better) than those of the RUM estimates. This suggests, again, that divisive coding is a better fit for subjects' choices even in environments where it is not an efficient strategy.

Taken together, these findings imply that subjects' value coding is bounded by cognitive constraints, since subjects exhibit divisive value coding even in statistical environments for which division comes at the expense of efficiency.

Finally, Table 3 shows the results from a combined regression of the full sample, and reports estimates for constant and random effects of α, the functional predisposition curvature. Our main goal here was to test whether the value coding mechanism responds at all to the statistical structure of the environment. We thus ran the random effects model only for specifications I, II, and III for which the MLE algorithm converged in the separate distribution-specific regressions (presented in Tables 1 and 2). Our results show that in spatial normalization models, the parameter α is significantly higher by 5.3 in the uniform distribution compared with the Pareto distribution ($p<0.001$).

In comparison, when modeling temporal normalization across trials, the parameter α is lower by 0.18 in uniform than in Pareto ($p<0.001$). This shows that even though subjects are constrained to a divisive coding mechanism, they do still adapt their value representations to the statistical structure of their environment, suggesting that value coding is context dependent (Louie et al., 2013; E. Shafir et al., 1993; S. Shafir et al., 2002), if not always efficient.

Conclusions

In a choice experiment, we tested whether decision-makers' value coding is obliged to employ a divisive coding mechanism even when it is inefficient. Leveraging recent theoretical results (Bucher & Brandenburger, 2022), we examined how two environmental statistics influenced value coding in risky choice. We estimated subjects' risk preferences in a biding experiment, and used the elicited risk preferences to generate continuous Pareto and uniform distributions of valuations for each individual subject. This unique design allowed us to eliminate effects that risk preferences might have had on subjects' choices. We then used these distributions to create binary lottery choice sets, where we compared subjects' choices across two statistical environments. While for Pareto-distributed environments, DN is an efficient code, in uniformly-distributed environments we would expect to encounter a near-linear coding of inputs without any divisive element. Using MLE model-fitting, we found evidence for DN value coding regardless of the statistics of the environment. Nevertheless, subjects did adapt to the statistical environment by calibrating the curvature of the encoding function, showing that they were clearly sensitive to environmental statistics, even if they could not adopt a fully efficient encoding strategy.

Our results suggest that subjects are obliged to employ a divisive form of normalized value coding, even at the cost of an embedded inefficiency in choice. These findings are in line with previous empirical results (Carandini & Heeger, 2012; Louie & Glimcher, 2012), which suggest that DN is a canonical encoding mechanism.

In fact, many real-world naturalistic stimuli have long-tailed asymmetric distributions (Simoncelli & Olshausen, 2001), and hence are, at least approximately, encoded efficiently with a DN-like function. This perhaps implies an evolutionary origin of the value-encoding mechanism.

In our model of choice under risk, contextuality is achieved by treating the lotteries as ordinary economic "goods" and then applying a DN valuation to these goods. That is, the inner structure of a lottery is not used in the contextualization step. This is different from how Frydman and Jin (2022) handle context dependency in that their model makes use of this inner structure.

We note that our findings do not necessarily imply the classical form of the DN function, suggested by Heeger (Heeger, 1992), since some of the specifications tested in this study diverge from the classical formulation. Thus, our results rather suggest a more general form of divisive computations (Steverson et al., 2019).

The main limitation of our study is its relatively short timescale. It is reasonable to wonder whether adaptation on a longer timescale might have allowed the subjects to more closely approximate a non-divisive representation when that representation was efficient. Future longitudinal studies will be required to address this important issue.

Table 3: RE of α by distribution for Models I, II and III. Standard errors are clustered at the subject-level, + $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. N=48,932

	(I)	(II)	(III)
Constant α	4.8215***	4.8350***	0.9785***
α Uniform (RE)	5.3152***	5.2980***	-0.1827*
BIC	33678.97	33676.53	33890.89

Acknowledgments

We thank Jerome Sidney Campos and Lexin Liang for their help in running experimental sessions. We also thank two anonymous referees, the Tymula and Glimcher lab members, and the attendees of the Annual Meeting of The Society for Neuroeconomics for invaluable comments. This work was funded with a grant from the Australian Research Council (DP 190100489 to Tymula and Glimcher). Financial support from NYU Stern School of Business, NYU Shanghai, and J.P. Valles (Brandenburger) is gratefully acknowledged.

References

Arnold, B. C. (2015). Pareto Distribution. In *Wiley StatsRef: Statistics Reference Online* (pp. 1–10). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118445112.stat01100.pub2

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, *9*(3), 226–232. https://doi.org/10.1002/bs.3830090304

Bhui, R., & Gershman, S. J. (20181115). Decision by sampling implements efficient coding of psychoeconomic functions. *Psychological Review*, *125*(6), 985. https://doi.org/10.1037/rev0000123

Bucher, S. F., & Brandenburger, A. M. (2022). Divisive normalization is an efficient code for multivariate Pareto-distributed environments. *Proceedings of the National Academy of Sciences*, *119*(40), e2120581119. https://doi.org/10.1073/pnas.2120581119

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*(1), 51–62. https://doi.org/10.1038/nrn3136

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and Normalization in Simple Cells of the Macaque Primary Visual Cortex. *Journal of Neuroscience*, *17*(21), 8621–8644. https://doi.org/10.1523/JNEUROSCI.17-21-08621.1997

Chater, N., & Brown, G. D. A. (1999). Scale-invariance as a unifying psychological principle. *Cognition*, *69*(3), B17–B24. https://doi.org/10.1016/S0010-0277(98)00066-3

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. https://doi.org/10.1016/j.jbef.2015.12.001

Frydman, C., & Jin, L. J. (2022). Efficient Coding and Risky Choice*. *The Quarterly Journal of Economics*, *137*(1), 161–213. https://doi.org/10.1093/qje/qjab031

Glimcher, P. W., & Tymula, A. A. (2023). Expected subjective value theory (ESVT): A representation of decision under risk and certainty. *Journal of Economic Behavior & Organization*, *207*, 110–128. https://doi.org/10.1016/j.jebo.2022.12.013

Harrison, G. W. (2008). *Maximum Likelihood Estimation of Utility Functions Using Stata*.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, *9*(2), 181–197. https://doi.org/10.1017/S0952523800009640

Heng, J. A., Woodford, M., & Polania, R. (2020). Efficient sampling and noisy decisions. *ELife*, *9*, e54962. https://doi.org/10.7554/eLife.54962

Kello, C. T., Brown, G. D. A., Ferrer-i-Cancho, R., Holden, J. G., Linkenkaer-Hansen, K., Rhodes, T., & Van Orden, G. C. (2010). Scaling laws in cognitive sciences. *Trends in Cognitive Sciences*, *14*(5), 223–232. https://doi.org/10.1016/j.tics.2010.02.005

Khaw, M. W., Glimcher, P. W., & Louie, K. (2017). Normalized value coding explains dynamic adaptation in the human valuation process. *Proceedings of the National Academy of Sciences*, *114*(48), 12696–12701. https://doi.org/10.1073/pnas.1715293114

Kutzner, F. L., Read, D., Stewart, N., & Brown, G. (2017). Choosing the Devil You Don't Know: Evidence for Limited Sensitivity to Sample Size–Based Uncertainty When It Offers an Advantage. *Management Science*, *63*(5), 1519–1528. https://doi.org/10.1287/mnsc.2015.2394

Louie, K., & Glimcher, P. W. (2012). Efficient coding and the neural representation of value. *Annals of the New York Academy of Sciences*, *1251*(1), 13–32. https://doi.org/10.1111/j.1749-6632.2012.06496.x

Louie, K., Grattan, L. E., & Glimcher, P. W. (2011). Reward Value-Based Gain Control: Divisive Normalization in Parietal Cortex. *Journal of Neuroscience*, *31*(29), 10627–10639. https://doi.org/10.1523/JNEUROSCI.1237-11.2011

Louie, K., Khaw, M. W., & Glimcher, P. W. (2013). Normalization is a general neural mechanism for context-dependent decision making. *Proceedings of the National Academy of Sciences*, *110*(15), 6139–6144. https://doi.org/10.1073/pnas.1217854110

Mcfadden, D. (2001). Economic Choices. *The American Economic Review*, *91*(3).

Olsen, S. R., Bhandawat, V., & Wilson, R. I. (2010). Divisive Normalization in Olfactory Population Codes. *Neuron*, *66*(2), 287–299. https://doi.org/10.1016/j.neuron.2010.04.009

Schwartz, O., Hsu, A., & Dayan, P. (2007). Space and time in visual context. *Nature Reviews Neuroscience*, *8*(7), Article 7. https://doi.org/10.1038/nrn2155

Schwartz, O., & Simoncelli, E. (2000). Natural Sound Statistics and Divisive Normalization in the Auditory System. *Advances in Neural Information Processing Systems*, *13*. https://proceedings.neurips.cc/paper/2000/hash/96c5c28becf18e71190460a9955aa4d8-Abstract.html

Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nature Neuroscience*, *4*(8), 819–825. https://doi.org/10.1038/90526

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, *49*(1), 11–36. https://doi.org/10.1016/0010-0277(93)90034-S

Shafir, S., Waite, T. A., & Smith, B. H. (2002). Context-dependent violations of rational choice in honeybees (Apis mellifera) and gray jays (Perisoreus canadensis). *Behavioral Ecology and Sociobiology*, *51*(2), 180–187. https://doi.org/10.1007/s00265-001-0420-8

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural Image Statistics and Neural Representation. *Annual Review of Neuroscience*, *24*(1), 1193–1216. https://doi.org/10.1146/annurev.neuro.24.1.1193

Steverson, K., Brandenburger, A., & Glimcher, P. (2019). Choice-theoretic foundations of the divisive normalization model. *Journal of Economic Behavior & Organization*, *164*, 148–165. https://doi.org/10.1016/j.jebo.2019.05.026

Webb, R. (2019). The (Neural) Dynamics of Stochastic Choice. *Management Science*, *65*(1), 230–255. https://doi.org/10.1287/mnsc.2017.2931

Webb, R., Glimcher, P. W., & Louie, K. (2021). The Normalization of Consumer Valuations: Context-Dependent Preferences from Neurobiological Constraints. *Management Science*, *67*(1), 93–125. https://doi.org/10.1287/mnsc.2019.3536