An Econometric Analysis of Salience Theory

Kelvin Balcombe[1], Iain Fraser*[2], and Abhijit Sharma[3]

[1] *School of Agriculture, Policy and Development, University of Reading. Email: K.G.Balcombe@reading.ac.uk*

[2] *School of Economics, University of Kent. Email: I.M.Fraser@kent.ac.uk*

[3] *Accounting, Finance and Economics Department, Huddersfield University Business School. Email: A.Sharma2@hud.ac.uk*

Sept 2020

Abstract

In this paper, we econometrically examine the performance of Salience Theory (ST) for explaining observed behavior outside of a fully defined state contingent setting. Using a well known data set, we find that only a minority of people act consistently in the way proposed by ST when confronted with lottery choices for which only marginal probabilities are presented. By estimating the implied dependence structure of payoffs consistent with ST, only a minority of people infer independent payoffs when attaching probabilities to states, a finding at odds with ST. Instead, a majority treat lotteries as having positively correlated payoffs which raises questions about the independence assumption in ST. Finally, we also find that ST explains choice behaviour less consistently than Expected Utility. Thus, ST should not be assumed to be superior to the most prominent models within the literature when employed outside of particular contexts.

Keywords: Salience Theory, choice under risk, expected utility

JEL Classifications: C11, C52, D81

Acknowledgements: The authors thank three anonymous referees for helpful and insightful comments and observations on an earlier version of the this manuscript

Accepted Version- Bulletin of Economic Research.

*Corresponding author: School of Economics, University of Kent, Canterbury, Kent, CT2 7FS

1 Introduction

Bordalo et al. (2012) offer a psychologically grounded theory of choice under risk generally referred to as Salience Theory (ST). Under ST, 'states' that have discordant payoffs receive more attention and greater weight in the determination of people's choices relative to states that have similar payoffs. Bordalo et al. (2012) suggests people are 'local thinkers', who re-weight probabilities based on comparison of payoffs under different states. Importantly, local thinkers do not value a lottery in isolation, but relative to alternatives. In this sense ST, unlike Expected Utility (EU) theory, is a state contingent theory of choice under risk.

In this paper, we examine ST using von Neumann and Morgenstern (vNM) lottery pairs. vNM lottery pairs are not framed such that each payoff, or pair of payoffs, is associated with a state. Rather, they are associated only with probabilities and associated payoffs. We refer to these as non-state contingent lottery pairs. Local thinkers (as defined by Bordalo et al., 2012) compare payoffs within states that have associated state probabilities. Consequently, local thinkers require not only that the marginal distributions of payoffs be defined, but also their joint distribution across lotteries. If the joint distributions are not defined objectively then local thinkers must determine them subjectively. Thus, a local thinker faces not only 'risk' but ambiguity when choosing between non-state contingent lottery pairs.[1]

When applying ST, it is commonly assumed that local thinkers treat all possible combinations of payoffs across lotteries as potential states, with payoffs being treated independently. We refer to this model property as the Independent Gambles Assumption (IGA). Specifically, under IGA, each subjective state probability is the product of the two payoff probabilities defining that state. However, quite reasonably, payoffs over two lotteries may be perceived as 'correlated'. For example, a local thinker might believe that a state defined by the highest payoff of one lottery in conjunction with the lowest payoff of another lottery is unlikely. Therefore, there is little justification for assuming the IGA when treating agents as local thinkers.

A simple example can illustrate the importance of the IGA. Assume an employee needs to get to work on time. If they arrive on time they get paid \$10, and zero otherwise. They can go to work by bus or train. The train costs \$3 and there is a 90% they will get there. The bus is cheaper at \$1 but there is only a 80% chance they will get to work on time. Which mode of transport to choose? We can think of the choice of transport as a lottery, so that lottery A is train and lottery B is bus. For these lotteries, we know the respective payoffs and associated probabilities. For a fully state contingent description of this problem there would be four possibilities: train and bus on time; train on time and bus late; train

[1] It is possible that there may be a possible preference for skewness related to the influence of salience of the payoffs on the choice between lotteries.

late and bus on time; and train late and bus late. With the information provided, we know the marginal probabilities but to apply a state contingent theory we need work out the state probabilities. The easiest way of generating these probabilities is to assume that whether the train is on time and the bus is on time are completely independent. This assumption, which avoids the need to generate joint distributions is what we refer to as the IGA. However, the employee might reasonably or unreasonably think that the train arriving on time and bus arriving on time are not necessarily independent because there is a chance that they are correlated because of some exogenous event like the weather. Thus, if this is the case then the IGA is not necessarily valid.

Currently, there is limited econometric evidence in the literature as to the relative predictive power of ST. For example, Booth and Nolen (2012) demonstrate that salience affects men and women differently, with men being affected more. Königsheim et al. (2019) provide estimates for parameters of ST, finding results in line with Bordalo et al. (2012), though their results show substantial heterogeneity. Furthermore, Kontek (2016) and Kontek (2018) observe that ST has a number of serious limitations including violation of the monotonicity assumption and undefined certainty equivalent for certain probability ranges.

Importantly, the antecedent literature does not focus on the applicability of ST outside of the fully specified state contingent domain. To address this, we employ data generated by Stott (2006) (see Appendix A for the full data set) to econometrically examine the relative predictive performance of ST. Specifically, we compare the performance of ST relative to EU theory. The comparison with EU theory has been done as a means to consider the predictive power of ST vis-a-vis EU theory. We use the Stott (2006) data set because it is freely available and as such the results we present can be readily replicated. Also, an attractive feature of the data set is that it was designed in such a way that it does not implicitly favour any specific decision theory. In summary, our results contribute to the existing econometric literature that has examined ST as well as the literature that considers various theoretical limitations of ST.

In terms of our econometric analysis, we present results for a range of model specifications and behavioral assumptions including the IGA following Balcombe and Fraser (2015). Overall, we find less support for ST than Bordalo et al. (2012), consistent with Kontek (2016) and Kontek (2018). Specifically, we do not find evidence supporting the IGA with this data set. In terms of model comparison, we find less support for ST than EU theory. Importantly, EU theory is known to perform worse than Cumulative Prospect Theory (CPT) for this data set (see Stott (2006); Balcombe and Fraser (2015)) and we use identical model specifications to Balcombe and Fraser (2015) such that our results transitively imply that ST is outperformed econometrically by CPT.

2 An outline of Salience Theory

Assume there are two lotteries (i, j)

$$l_i = (p_i, x_i) \text{ and } l_j = (p_j, x_j) \tag{1}$$

where $p_i = (p_{i1,....,}p_{iM})$ and $p_j = (p_{j1,....,} p_{jN})$, are the probabilities, and $x_i = (x_{i1,....,} x_{iM})$ and $x_j = (x_{j1,....,} x_{jN})$ are the payoffs of the lottery, of dimension M and N respectively. Given lotteries, i and j, BGS proceed by hypothesizing a state space D composed of the set of states $\{\xi_{mn}\}_{m,n}$ such that ξ_{mn} is uniquely associated with the payoffs (x_{im}, x_{jn}). Next define x_{im} as the payoff associated with the event ξ_{im} and x_{jn} as the payoff associated with ξ_{jn}, the state being defined by $\xi_{mn} = \xi_{im} \cap \xi_{jn}$ which has an associated probability π_{mn} for which ${}_{nm}\pi_{mn} = 1$. The dimension of D (which may include probability zero states) is, therefore, $S = M \times N$. An economic agent evaluates each lottery (i) relative to lottery (j) according to the value of each payoff $v(x_{im})$ such that

$$U_i =_n \sum_m \pi_{mn} v(x_{im}) \tag{2}$$

The salience function [equation (5) in BGS] is:

$$\sigma_{mn} = \frac{|x_{in} - x_{jm}|}{|x_{in}| + |x_{jm}| + \theta} \tag{3}$$

assuming some $\theta > 0$. The numerator captures the ordering property of the lotteries i and j, and the denominator takes account of the diminishing sensitivity property whereby salience decreases the further a state's payoff is from zero. Given salience values are used to transform probabilities denoting r_{mn} as the rank of σ_{mn} (with one being highest, and ties receiving the same rank) the "weight" is:

$$w_{mn} = \frac{\delta^{r_{mn}}}{{}_{nm}\delta^{r_{mn}}\pi_{mn}} \tag{4}$$

where $\delta^{r_{mn}}$ measures how much salience "distorts" decision weights. If the $\delta^{r_{mn}}$ equal one then we have the typical EU decision maker. Finally, ω_{mn}, are computed from the probabilities that a specific state occurs, which are computed as:

$$\omega_{mn} = w_{mn}\pi_{mn} \tag{5}$$

This is equivalent to equation (8) in Bordalo et al (2012). Unlikely events will be overweighted if they have salient payoffs, and they will be under-weighted otherwise. BGS argue that empirical evidence supports the construction of π_{mn} from independent lotteries using p_i and p_j:

$$\pi_{mn} = p_{im}p_{jn} \tag{6}$$

However, alternative non-independent constructions are plausible. For example,

$$\pi_{mn} = P\left(\xi_{im} \cap \xi_{jn}\right)$$

such that we have $M \times N$ potential states requiring $(M \times N) - 1$ unknowns to be constructed. This is done by ordering states with the highest to lowest where given $\tau \in [-1, 1]$, a split region can be constructed as:

$$\pi_{mn} = p_{jn}p_{im} + \tau \times (p_{jn}p_{im})\, I\,(\tau < 0) + \tau \times (\min(p_{jn}, p_{im}) - (p_{jn}p_{im}))\, I\,(\tau > 0) \tag{7}$$

At $\tau = 0$, we have independence; $\tau = -1$ represents the most extreme negative correlation; and, $\tau = 1$ gives the most extreme positive correlation. τ can be estimated alongside θ and δ.

As outlined in Balcombe and Fraser (2015), the parameters of interest in the models (i.e., $\Omega_s \in \Omega$) are in only one of two forms. That is, we parameterize our models by using a normal variate ϑ_s that is then transformed according to

$$\begin{aligned}
\Omega_s &= t_1\left(\vartheta_s; \varphi_{l_s}, \varphi_{u_s}\right) = \varphi_{l_s} + \left(\varphi_{u_s} - \varphi_{ls}\right)\frac{e^{\vartheta_s}}{1 + e^{\vartheta_s}} \\
&\text{or} \\
\Omega_s &= t_2\left(\vartheta_s\right) = \exp\left(\vartheta_s\right) \\
\text{where } &\vartheta \in \mathbb{R}
\end{aligned} \tag{8}$$

In the case of $t_1(\vartheta; \varphi_l, \varphi_u)$ the transformed parameter lies within the interval (φ_l, φ_u). We set the values for the interval *a priori* in accordance with the inequality constraints outlined above for the three model specifications.

In general, the priors for parameters of the form $t_1(\vartheta; \varphi_l, \varphi_u)$ are $(\vartheta \sim N(0, \zeta))$ where they are assigned a variance ζ equal to $\frac{9}{4}$. This means that there is an approximately uniform prior within the specified interval, although there is marginally less mass at the very extremes. This implies that we are being 'non-informative' about the values except that we have specified the interval over which the parameters lie.

Taking each parameter in turn we first consider the parameters priors that are transformed according to t_1. In this case we set $\delta = t_1(\vartheta_\delta; 0, 1)$ and $\tau = t_1(\vartheta_\tau; -1, 1)$. For the other two parameters we employ the t_2 transformation. For the power value function, which at the individual level must allow scope for some individuals to display convexity, we set our prior such that $\alpha = t_2(\vartheta_\alpha)$ $Pr(\alpha < 0.1) = 0.10$ and $Pr(\alpha < 2) = 0.90$. This equates to having 75% of the prior mass in the concave region. For our final parameter, we have $\phi = t_2(\vartheta_\phi)$ where ϑ_ϕ has a prior standard deviation of 3. By undertaking sensitivity analysis, we established that doubling or halving this standard deviation had little impact on the overall results we present.

How Much Does Correlation Matter in Salience Models?

To motivate this issue we take the case that Bordalo et al. (2012) use as an illustration in their paper. First, we examine the lottery pair $L_1(2,400)$ $L_2(2,400)$ (page 1247). They write this as:

$$L_1(2,400) = \left\{ \begin{array}{cc} Pay & Prob \\ 2,500 & 0.33 \\ 0 & 0.01 \\ 2,400 & 0.66 \end{array} \right\} : L_2(2,400) = \left\{ \begin{array}{cc} Pay & Prob \\ 2,400 & 1 \end{array} \right\}$$

In a state form representation we could write this as

	Payoff 1	Payoff 2	Prob
State 1	2,500	2,400	0.33
State 2	0	2,400	0.01
State 3	2,400	2,400	0.66

Bordalo et al. (2012) use this as an example of how salience can be constructed and understood. However, the fact that there was a sure option (2,400) distributed across all three states means that there was an automatic mapping from payoff probabilities to state probabilities. The fact that this is not always the case can be made clear by examining the second lottery pairing $L_1(0)$ $L_2(0)$.

$$L_1(0) = \left\{ \begin{array}{cc} Pay & Prob \\ 2,500 & 0.33 \\ 0 & 0.67 \end{array} \right\} : L_2(0) = \left\{ \begin{array}{cc} Pay & Prob \\ 2,400 & 0.34 \\ 0 & 0.66 \end{array} \right\}$$

Here, we can construct three particularly interesting cases. There is an independence construction (a), then (b) is under the assumption that the high payments are "mutually exclusive", and (c) when payoffs are in a sense the most positively correlated.

	Payoff 1	Payoff 2	Prob (a)	Prob (b)	Prob (c)
State 1	2,500	2,400	0.113	0	0.33
State 2	2,500	0	0.218	0.33	0
State 3	0	2,400	0.227	0.34	0.01
State 4	0	0	0.442	0.33	0.66

Our main point here is that the potential for economic agents to construct (b) or (c) is not at all implausible, and this leads to very different results.

For example, under linear utility and assuming the parameters used in Bordalo et al. (2012) with respect to the salience function, assuming that $\theta = 0.1$, with $\delta = 0.73$, the weighting parameter (as defined in Bordalo et al. by their equation (7) and the relative utilities of (a), (b) and (c) above are 245.7, 304.1 and 0.1 respectively. Interestingly, these results are in agreement that the third state version gives almost indifference between the two gambles.[2]

Finally, it only needs a minimal shift in the parameters of the salience function to yield negative values for (c) and positive values for (a) and (b). For example, if we employ $\delta = 0.67$ we have (a) 330.8, (b) 394.1 and (c) -3.5. Therefore, how people interpret correlation is material and as such needs to be considered when examining relative model performance.

3 Model specification

The Bordalo et al. (2012) ST model can be estimated by treating the difference in utilities between two lotteries as a deterministic signal with a stochastic 'link'. We specify the utility function as $v(x) = x^{\alpha}$ only, given its support in previous studies (see Stott (2006) and Balcombe and Fraser (2015)) as well as in Bordalo et al. (2012). The ST model is implemented using (3), together with (4). θ is only weakly identified since multiple values for θ give identical salience ranks for states Therefore, we set $\theta = 0.1$ throughout. This choice of parameter value is identical to Bordalo et al. (2012, p. 1264). The dependence relationship in equation (7) is specified by τ. Finally, we embed the utility difference between lotteries i and j, $\Delta_{ij} = U_i - U_j$ within a link function

$$F(\Delta_{ij}, \phi)$$

[2]For example, using the last column probabilities (Prob(c)) ordered according to the State 1 to State 4, the calculations are under equation (3) $\sigma' \simeq (0.0204, 0.99996, .999958, .00000)$, therefore $r' \simeq (3, 1, 2, 4)$, and under equation (4) $\omega' \simeq (0.399, 0.00, 0.017, 0.584)$. Under linear utility $v_1' = (2500, 2500, 0, 0)$ and $v_2' = (2400, 0, 2400, 0)$ then $U_1 = \omega' x_1 \simeq 999.396; U_2 = \omega' x_2 \simeq 999.247; U_1 - U_2 = 0.15$.

that gives the probability that lottery i is chosen over lottery j. We employ four alternative link functions (beta, probit, logit and constant probability). Therefore, we have the parameter set $\Omega = (\alpha, \delta, \tau, \phi)$ and three specifications:

1. EU-Model: Expected Utility - $\delta = 0, \tau = 0, \alpha > 0$
2. SWI-Model: Salience With Independence - $\delta \in (0, 1)$, $\tau = 0$ and $\alpha > 0$
3. SWD-Model: Salience With Dependence - $\delta \in (0, 1)$, $\tau = (-1, 1)$ and $\alpha > 0$

4 Estimation, inference and results

Estimation uses adaptive Markov Chain Monte Carlo for each individual. After convergence, the mean and standard deviation for the parameters are recorded. The Logged Marginal Likelihood (LML) is computed for individual k for model t, denoted as l_{kt}. Larger $l'_{kt}s$ indicate greater support for model t for individual k. At the aggregate level, the marginal likelihood for model t over all individuals is obtained by:

$$m_{kt} = \exp\left(l_{kt} - \sum_{t=1}^{T} \frac{\exp(l_{kt})}{T}\right) \tag{9}$$

Responses were elicited using a gamble format used by Stott (2006) where all respondents were asked to choose between 90 lottery pairs, each with two payoffs. Estimates are generated for all 90 individuals using the EU, SWI and SWD models covered above. We present the aggregate LMLs for each model in Table 1.

Table 1: Logged Marginal Likelihoods by model type

	Beta	Probit	Logit	Constant Probability
EU	-3728.59	-3813.39	-3860.66	-4025.37
SWI	-3745.17	-3836.82	-3878.54	-4172.07
SWD	-3740.78	-3829.99	-3873.52	-4166.33

Higher LMLs in Table 1, indicate more support for a model specification when imposed on all individuals. We find that: EU $\succeq$ SWD $\succeq$SWI (where $\succeq$ denotes preference). Thus, at the aggregate level the EU model has the highest support. The salience model that allows for dependence across the states ($\tau \neq 0$) is preferred. For links, the preference is: Beta $\succeq$Probit$\succeq$Logit$\succeq$Constant Probability. The model rankings of the EU, SWD and SWI are invariant to the link.

That the EU is supported in aggregate, does not mean that it best characterises everyone. Given that SWI has less support, and the Beta link had the most support, we further compared the EU and SWD models with the Beta link. Model probabilities of the SWD model relative to the EU model by individual are presented in Figure 1. Most individuals have a low probability for SWD, but fourteen respondents have model probabilities almost equal to one. Thus, as in Balcombe and Fraser (2015), there is considerable heterogeneity in model support across individuals. We found that choosing $\theta = 0.1$ is not pivotal to model selection since rankings of lotteries are relatively insensitive to θ.

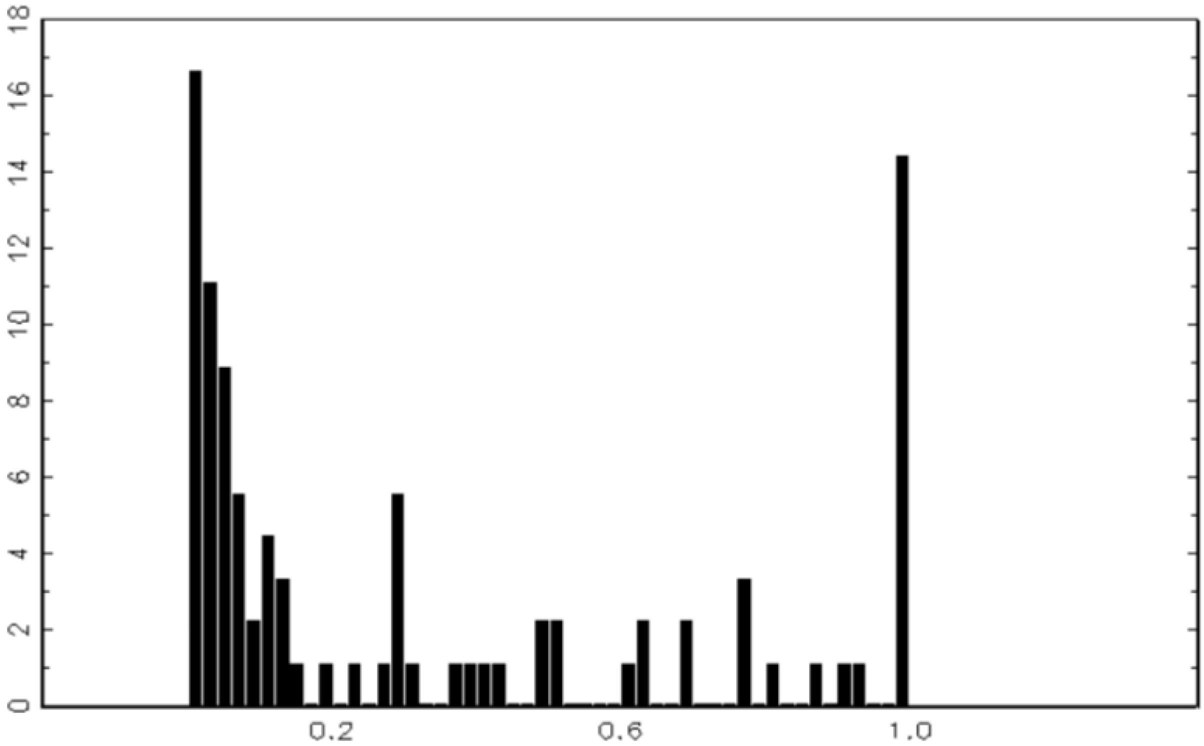

Figure 1: Individual Model Probabilities of the SWD model

We also report estimates for δ (posterior means) for the SWD specification. A histogram for these is given in Figure 2. The greater mass of the histogram is in the upper region (greater than 0.7). While these mean values give the impression of being somewhat below one, most have 95% credible intervals (i.e. higher density regions) with upper bounds close to one. This suggests most respondents behaved in a manner that is broadly consistent with EU, with a minority of participants behaving in accordance with ST.

Finally, we consider the robustness of the IGA assumption. The histogram for τ (individual posterior means) is presented in Figure 3. In this case0, the mass of the distribution is above zero, indicating that, under ST, the majority of individuals perceive payoffs in the

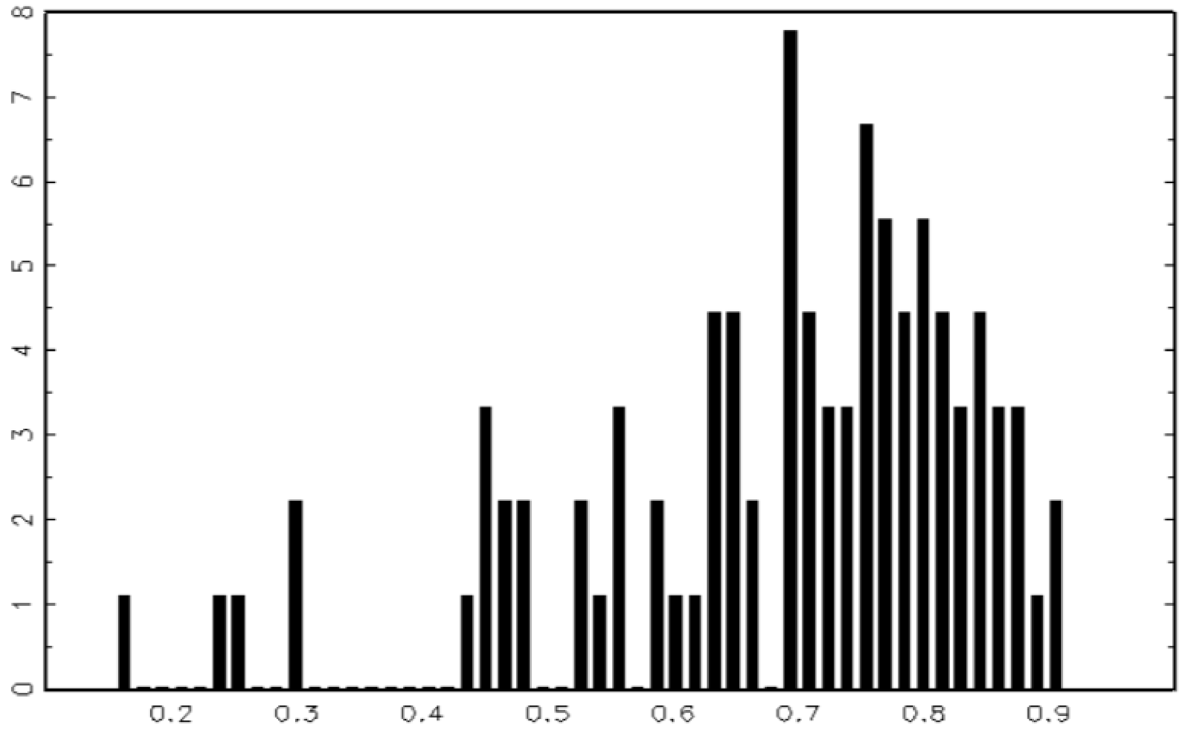

Figure 2: δ by individual

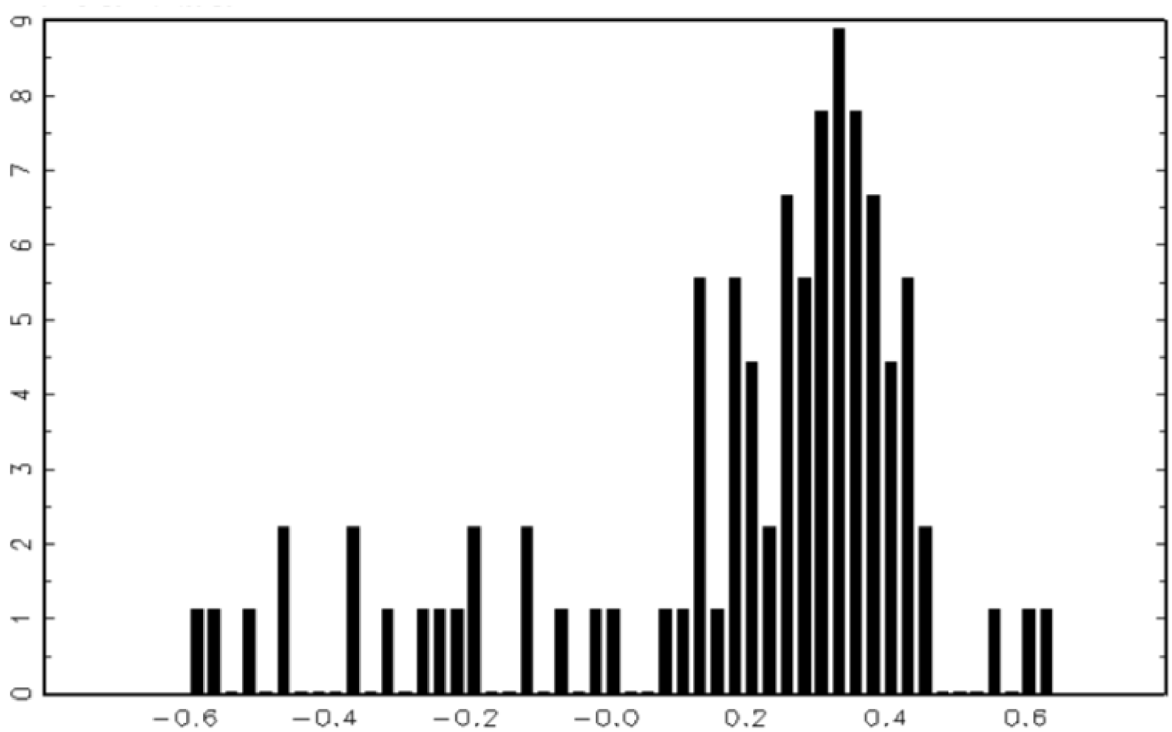

Figure 3: τ by Individual

lottery pairs to have positively correlated payoffs across possible states. Specifically, if a state contains two relatively large payoffs, or two relatively small payoffs, the subjective probability of these states is higher than would be obtained by multiplying the respective payoff probabilities. This finding is at odds with Bordalo et al. (2012) who report results (in their Appendix 2) supporting IGA.

5 Conclusion

We examine the performance of the salience specification under risk introduced by Bordalo et al. (2012) outside of a fully defined state contingent setting. We extend their empirical implementation by analysing lottery preferences from Stott (2006) where only the marginal distributions of the lottery pairs are defined *a priori*. Under salience, payoffs in vNM lottery pairs are commonly treated by local thinkers as being positively correlated, when no such assumption is made explicit. In terms of overall support for salience, a minority of respondents behave consistently with ST compared to EU. In addition, previous research has found that EU could not explain the Stott (2006) data relative to CPT. Thus, ST should not be assumed to be superior to the most prominent models within the literature when employed outside of particular contexts. Our work does not preclude the existence of 'local thinkers' since local thinking requires an additional cognitive burden when the states and their probabilities are not fully defined objectively. We suggest, however, that local thinking is perhaps a context specific strategy employed by some respondents only.

Finally, our results are based on the examination of a single data set that has relatively low average payoffs. **In theory, for both EU and ST, increasing the magnitude of real payoffs should not lead to a change in behaviour. However, in practice this need not be the case when empirical examining risky choices that entail high real payoffs. For example, the degree of state contingent correlation may change when economic agents are making decisions under risk with higher real payoffs than those examined here. Additionally, the parametric assumptions we have employed for both EU and ST may break down under extremely large real payoffs, accentuating errors that are implicit in the econometric model specifications, and reducing the predictive power of both theories.** As such, further examinations of ST using lottery data that include larger stakes would be a useful exercise to undertake in an effort to further our understanding of the strengths and weaknesses of ST.

6 Appendix A

A copy of the Stott (2006) data that has been employed in this paper is provided. For details regarding data collection and associated experimental design please refer to Stott (2006).

References

Balcombe, K. and Fraser, I. (2015). Parametric preference functionals under risk in the gain domain: A bayesian analysis. Journal of Risk and Uncertainty, 50(2):161–187.

Booth, A. and Nolen, P. (2012). Salience, risky choices and gender. Economics Letters, 117(2):517–520.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2012). Salience theory of choice under risk. The Quarterly Journal of Economics, 127(3):1243–1285.

Königsheim, C., Lukas, M., and Nöth, M. (2019). Salience theory: Calibration and heterogeneity in probability distortion. Journal of Economic Behavior & Organization, 157:477–495.

Kontek, K. (2016). A critical note on salience theory of choice under risk. Economics Letters, 149:168–171.

Kontek, K. (2018). Boundary effects in the marschak-machina triangle. Judgement and Decision Making, 13(6):587–606.

Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. Journal of Risk and Uncertainty, 32(2):101–130.