Journal of Neuroscience, Psychology, and Economics

Neural Random Utility: Relating Cardinal Neural Observables to Stochastic Choice Behavior

Ryan Webb, Ifat Levy, Stephanie C. Lazzaro, Robb B. Rutledge, Paul W. Glimcher

Funded by:

- National Institute on Drug Abuse

Neural Random Utility:
Relating Cardinal Neural Observables
to Stochastic Choice Behaviour

immediate

November 16, 2018

Abstract

We assess whether a cardinal model can be used to relate neural observables to stochastic choice behaviour. We develop a general empirical framework for relating any neural observable to choice prediction, and propose a means of bench-marking their predictive power. In a previous study, measurements of neural activity were made while subjects considered consumer goods. Here, we find that neural activity predicts choice behaviour, with the degree of stochasticity in choice related to the cardinality of the measurement. However, we also find that current methods have a significant degree of measurement error, severely limiting their inferential and predictive performance.

1 Introduction

Traditional economic methods for establishing a utility representation, such as revealed preference, are now routinely used to identify the anatomical and functional characteristics of "value" signals in the human brain (Fehr and Rangel 2011, Glimcher and Fehr 2013). This suggests a general strategy for eliciting preferences in situations where standard revealed preference methods are problematic or choice data is unavailable: measurements of neural activity can be used to assess valuations of choice alternatives in order to directly predict a subject's choice behaviour (Knutson et al. 2007, Lebreton et al. 2009, Krajbich et al. 2009, Tusche et al. 2010, Levy et al. 2011, Smith et al. 2014, Telpaz et al. 2015). Indeed, such prediction methods have not only been applied within-individual, but also across individuals and across populations (Falk et al. 2012, Smith et al. 2014, Telpaz et al. 2015, Genevsky and Knutson 2015, Genevsky et al. 2017).

The early prediction literature has proceeded along two avenues. The first established the *ordinal* properties of the neural measurement within a deterministic choice model (Tusche et al. 2010, Levy et al. 2011). In effect, it was assumed that the choice alternative associated with the higher measurement of neural activity is always chosen. The second relaxed this assumption of ordinality in an effort to better fit the choice data. It is well-accepted that choice behaviour exhibits stochastic properties (Luce 1959), and *cardinal* methods allow the probability of choosing an item to depend on the difference in measured neural activity between two choice alternatives (e.g. Knutson et al. 2007, Smith et al. 2014, see Section 2 for a full review).

However, in the context of choice prediction, little attention has been paid to the sources of stochasticity — in neural activity and its measurement— which lead to a cardinal choice model. Of course, it is widely held in neuroscience that neural activity is inherently stochastic and the cardinal properties of neural measurements have been routinely described for over half a century (e.g. Rieke et al. 1997, Glimcher 2005). However the sources of this stochasticity, and at what stage of the choice process it might

arise, can have critical implications for how researchers relate stochastic neural data to stochastic choice behaviour. In particular, our understanding of how the statistical properties of neural measurements interact with the experimental paradigm is limited, thus impacting both choice prediction and inference.

In economics, the class of Random Utility Models are routinely applied to capture stochastic choice behaviour within a utility maximization framework (Becker et al. 1963, McFadden 1973, 1981, 2001). Many of these models have the inherently cardinal feature that the probability of choosing an alternative is related to the differences in utilities. Motivated by classic experiments which demonstrate that choice probabilities vary with utility (e.g. Mosteller and Nogee 1951, Hey and Orme 1994), random utility models have been widely applied to experimental data. However experimental studies have also questioned the relevance of a cardinal model for individual stochastic choice behaviour, at least in some cases, instead proposing that utilities might be random but not cardinal (e.g. "random preference" models, Loomes 2005). For instance, Agranov and Ortoleva (2017) present subjects with repeated choice sets over lotteries, and consistent with previous literature, they find a large majority of subjects exhibit stochastic choice. However, they also find no statistically significant relationship between the difference in the (estimated) Expected Utility of any two lotteries and the likelihood that a subject switched their choice on repeated trials. How best to model stochastic choice behaviour is still a question of much debate (Hey 2005).

In this article, we aim to assess whether a cardinal framework can, and should, be used to relate neural observables to stochastic choice behaviour. Below, we outline a broadly applicable econometric framework for relating neural observables to stochastic choice behaviour, which we call the Neural Random Utility Model (NRUM). The NRUM extends familiar aspects of the random utility framework to neural observables, including both the maximization of stochastic decision variables and the possibility that differences in these variables contain information for choice prediction.

Additionally, the NRUM allows the development of hypotheses about the various sources of error present in the measurement of neural activity, an issue that has not

been addressed by previous literature. We demonstrate how these errors interact with stochastic decision variables in a choice prediction exercise, and we examine how features of the experimental design allow separate estimation of measurement error from the stochastic decision variable. This allows us to quantify the relative magnitudes of these errors in a way that is not possible with an ordinal approach, providing an estimate of the variance of measurement error in choice prediction experiments.

Because the model is general purpose, it can be employed with regard to any neural observable to assess whether different experiments – or future measurement techniques – provide true advances in choice prediction. To demonstrate this feature, we apply it to a well-known dataset previously used to establish the ordinal properties of a neural measurement (Levy et al. 2011). In the first stage of the experiment, subjects were shown each of 20 consumer items while they were inside an fMRI scanner. For each item, targeted measurements of neural activity in the medial Pre-Frontal Cortex (mPFC) and Striatum were recorded. In the second stage, subjects were asked to make choices between all pairs of the items, with all choice sets repeated twice. This dataset thus has a crucial feature which can be used to test a cardinal choice model. Since the measurements were made independently of the choice (over the course of an hour), the scale on which the measurements were made must be (at least partially) maintained over measurements for the dataset to have any predictive power. This would provide evidence for one property of a cardinal measurement, namely that each measurement is from a common scale.

The Levy et al. (2011) data also hints at a second property of a cardinal measurement, namely that differences between measurements contain predictive information. While the choice behaviour of subjects was relatively consistent, maintaining transitivity in 96% of eligible triplets, subjects did switch their choices in 9.3% of the repeated choice sets they faced. This is a degree of choice stochasticity typically found in such experiments (e.g. Telpaz et al. 2015). Even though the analysis in Levy et al. (2011) consisted of an ordinal ranking of the BOLD activity for each item (Figure 1) — the item with the higher ranking was predicted to be chosen — the choice prediction rate

was highest for the pairs of items with the largest ordinal "distance" in ranking (83% vs. an overall prediction rate of 56%, across all choice sets and all subjects). This suggest that the distance between neural measurements matters. However the analysis in this widely cited study highlights two issues typical of the neuroscientific literature on choice:

- The sources of stochasticity in neural prediction variables are not modelled at all. Note that the prediction exercise appears to perform *worse* than chance for items with adjacent rankings – which is obviously impossible. This arises because neural measures are constructed via multiple levels of analyses. The interaction between these random variables, the errors in their measurement, and the choice prediction exercise can, and has, led to errors in inference.[1]
- An ordinal analysis does not account for some information, namely the difference in neural activity, that may improve predictive performance for repeated choice sets. For instance, while an ordinal model does not predict that a subject will switch their preference, a cardinal analysis can vary the probability of choice according to the difference in neural activity.

These issues can be addressed by applying the NRUM to this dataset. We find that difference in the neural observable is significantly correlated with choice behaviour — and has choice prediction power beyond chance — however we also find evidence for a startling degree of measurement error in the neural data. This measurement error biases model estimates towards zero, adversely affecting both prediction rates and inference about which brain areas have predictive power. It also leads to (what appear to be at first glance) puzzling features of the choice data. The NRUM allows a partial correction for measurement error, and we examine features of the experimental design which yields identification of this source of error.

[1]Each neural measurement is constructed from numerous independent scans, each comprised of a large amount of data with a spatial and temporal structure, and each measurement is related to repeated choices from one of $\binom{20}{2}$ binary choice sets. The sources of stochasticity present in neural measurement, the experimental paradigm, and how they interact in choice prediction, is clearly an important aspect of any econometric evaluation of such datasets (Harrison 2008).

Finally, a means of benchmarking the predictive power of the measurements, with regard to stochastic choice, is also proposed. Existing prediction methods treat each choice trial independently, even if there are repeated trials from the same set. We propose two methods for assessing the predictive performance of such repeated choice paradigms, and demonstrate that the ordinal prediction methods used to date cannot capture these basic features of the data. We also find that the neural measurements found in Levy et al. (2011) yield choice in-sample prediction results barely on par with standard observables (price and quality ratings), even after accounting for their cardinal features. Combining the neural measurements with these standard economic observables improves predictive performance. While suggestive that these two types of observables contain orthogonal predictive information, it remains to be seen if improvements in measurement technology can achieve improved prediction rates.

2 Related Literature

2.1 Choice Prediction with Neural Observables

Much of the initial excitement in the field of neuroeconomics involved extracting value signals from the vast amount of data produced by fMRI studies. Typically, these studies measured the response in neural activity to some behavourial manipulation or stimulus, such as willingness-to-pay or reward amount, and analyzed models of the form:

$$\text{NeuralActivity} = \beta \text{Behaviour} + \epsilon. \tag{1}$$

Here, neural activity is the dependent variable and β is the parameter of interest, in particular, *which* brain regions significantly code value signals. Recent meta-analyses of this literature (now amounting to over 200 independent datasets) identify that activity in two brain areas, the ventral striatum and the mPFC, is tightly correlated with every known economic method for estimating the values subjects place on choice objects — ranging from consumable goods, to money lotteries, to charitable donations, to durable

goods, to social preferences, to political preferences (Levy and Glimcher 2012, Bartra et al. 2013, Clithero and Rangel 2013).

Building on these advances, neuroeconomists began exploring whether this relationship could be reversed for the purpose of explaining choice behaviour.

$$\Pr(\text{Behaviour}) = \beta \text{NeuralActivity} + \epsilon. \tag{2}$$

In this modelling approach, neural activity (causally) determines the choice probability, and the strength of this relationship is governed by the parameter β. For example, in a landmark study, Knutson et al. (2007) applied a Logit model to an fMRI dataset and found that incentivized purchasing behaviour can be predicted by measures of neural activity in the mPFC and the ventral striatum. Follow up studies have demonstrated similar results in binary choice experiments over disparate objects (FitzGerald et al. 2009), and have even extended this analysis to market level outcomes (Falk et al. 2012, Venkatraman et al. 2015, Genevsky and Knutson 2015, Genevsky et al. 2017).[2]

An obvious concern with (2) is that there are a large number of potential neural variables in an fMRI dataset to use as predictors, much larger than the number of choice observations. This suggests that many neural signals will be correlated with choice by chance. The initial literature took a conservative approach to this issue by defining regions of interest either *a priori* or via independent localization.[3] In an effort to improve prediction rates, more sophisticated methods for model selection have been developed. For example, Smith et al. (2014) use a *shrinkage estimator* to determine which voxels to include as predictors.[4]

Regardless of the estimation method, analysis based on the model (2) is inherently cardinal. The parameter(s) β determine how choice probabilities change in response to neural activity, and these probabilities can then be used to predict choices out-

[2]See Berkman and Falk (2013) for a discussion of more applications of this approach.

[3]For instance, Levy et al. (2011) used an independent sample to identify the regions of the mPFC and ventral striatum to include in the model.

[4]A Logit model comprised of all voxels is evaluated, but the model's likelihood is penalized via a LASSO regression to guard agains over-fitting. This penalization acts as a model selection criterion, with the resulting estimates of $\beta \neq 0$ only for some voxels.

of-sample. For instance, in Smith et al. (2014) the fitted probabilities $\hat{P}$ from the regression (2) are used to code a predicted choice then compared to the choices from a holdout sample, yielding a neuro-choice prediction rate of 61% within-subject.[5] The estimates of β therefore determine the relative weights of different brain areas, voxels, or stimuli, in determining choice. We will use this cardinal framework to consider the stochastic structure which underlies both behaviour and neural activity.

2.1.1 The Impact of Measurement Error

One useful example of why it is important to address the stochastic structure of a choice model is the prevalence of measurement error in fMRI datasets. Since neural activity is not typically observed directly, only an indirect measure of it is available for prediction (say via the BOLD measure from fMRI). To model measurement error, we follow the standard approach of appending a measurement error, μ, to our variable of interest.[6]

$$\text{NeuralMeasure} = \text{NeuralActivity} + \mu. \tag{3}$$

In standard analyses based on equation (1), measurement error is somewhat innocuous: any error in the measurement of neural data will simply end up in the error term of the regression,

$$\begin{aligned} \text{NeuralMeasure} &= (\beta \text{Behaviour} + \epsilon) + \mu \\ &= \beta \text{Behaviour} + (\epsilon + \mu) \end{aligned} \tag{4}$$

Though μ clearly adds noise to the model, thereby increasing standard errors, the estimate of β is not directly affected. A number of methods have been proposed

[5] Machine-learning algorithms can also be used to jointly analyze (or weight) regions of activity within the mPFC to classify whether a particular item was chosen, or not, from a binary choice set (Kahnt 2017, for a review). These weights can then be applied to a test dataset to predict choice behaviour. Krajbich et al. (2009) use such methods to classify valuations in a public goods game with 60% accuracy, while Tusche et al. (2010) observe classification rates upwards of 75% in a choice task over activities. However this binary classification does not provide relative choice probabilities, only predicted outcomes. The former plays a crucial role in modelling stochastic choice behaviour.

[6] See Greene (2003) for a textbook treatment of measurement error.

to address this issue in standard fMRI analysis software, primarily relying on the autocorrelation structure of the measurement error (e.g. Lund et al. 2006).

In the choice prediction model (2), however, the impact of measurement error is more nefarious. Now the measurement error is embedded in the explanatory variables of the model. We can observe this by directly substituting in equation (3).

$$\begin{aligned} \Pr(\text{Behaviour}) &= \beta \text{NeuralActivity} + \epsilon \\ &= \beta \text{NeuralMeasure} + (-\beta\mu + \epsilon) \qquad (5) \\ &= \beta \text{NeuralMeasure} + \tilde{\epsilon}. \qquad (6) \end{aligned}$$

Because the error term $\tilde{\epsilon}$ (which includes the measurement error) is now correlated with the explanatory variable, a critical exogoneity assumption of the regression model is violated. This "error-in-variables" problem biases the estimate of β towards zero (Yatchew and Griliches 1985).[7] Not only does this bias alter the predicted choice probability given a change in the neural measure, but it also means that inference on β for a given brain region or voxel will be too conservative. On average, this will lead to fewer rejections of a false null hypothesis (i.e. increased "Type II Errors"). In Section 4, we will demonstrate how the NRUM can be used to address the measurement error problem.

2.2 Models of Stochastic Choice

The literature on modelling stochastic choice consists primarily of two model classes which fall under the technical definition of a Random Utility Model (RUM; Becker et al. 1963). Consider a set of n items, indexed $i = 1 \dots n$. Denote P_i the probability that alternative i is chosen from this set, or equivalently, the frequency with which i is chosen on repeated trials.

A RUM posits the existence of a vector of random variables $\mathbf{u}$, with element u_i,

[7] Also see Ramsey et al. (2010) for a discussion of this issue in dynamic causal modelling.

such that

$$P_i = \Pr\{u_i > u_j, \quad \forall j \neq i\}. \tag{7}$$

Conditions placed on P_i determine whether observed behaviour is consistent with the principle of utility maximization (Block and Marschak 1959, Falmagne 1978, McFadden 2005).

The two approaches to modelling stochastic choice, while both technically RUMs, are distinct in interpretation. One class, known as *random preference* models, posit that a choice is represented by a preference relation (or utilities $\mathbf{u}$) stochastically drawn from a set $\mathcal{U}$ which obeys some underlying axioms (Loomes and Sugden 1995, Gul and Pesendorfer 2006). Each alternative in a choice set is processed simultaneously according to this realized preference relation. This approach allows for preferences to vary from trial to trial for different realizations of $\mathbf{u}$, but in a manner which is internally consistent with the axioms which determine membership in $\mathcal{U}$. Such models have important implications for both model-testing and normative analysis, since they posit no violations of the underlying axioms due to stochasticity (see, for example, Loomes 2005).

A second approach to modelling choice stochasticity derives from the long literature on stochasticity in sensory perception (Fechner 1860, McFadden 2001).[8] A *Fechnerian* RUM holds that choices can be described by a single "core" valuation v_i that is perceived or represented with error ϵ_i for each item, such that $u_i = v_i + \epsilon_i$. The perturbed value is then compared, and the number of choice errors (in violation of the ordering given by v_i) is governed by the magnitude of the difference $v_i - v_j, \ \forall j \neq i$. Therefore the additive model is described as cardinal (Batley 2008).[9] Empirical studies which utilize the Fechnerian model include Hey and Orme (1994), Hey (1995), Buschena and

[8] See also Weber (1834), Stevens (1961), Falmagne (1985). For applications in the economics literature, see (e.g. Hey and Orme 1994, Camerer and Ho 1994, Loomes 2005, Harrison and Rutstrom 2008, Johnson and Ratcliff 2013, for reviews).

[9] A taxonomical issue currently exists between the theoretical and applied discrete choice literatures in economics. The applied literature classifies the additive model as a RUM since it satisfies the definition (7). However the theoretical literature does not since the stochasticity in the model leads to violations of the axioms underlying membership in $\mathcal{U}$. Here, we return to the standard definition from Becker et al. (1963) used by the applied literature.

Zilberman (2000), Hey (2005), with the negative result found in Agranov and Ortoleva (2017) previously noted. Review articles which contrast the behavioural evidence for the two approaches can be found in Loomes (2005) and Wilcox (2008).[10]

Intriguingly, support for both Random Preference and Fechnerian approaches can be found in the neuroscience evidence (in so far as neural evidence can be used to support an economic model). It is widely held that the activity of a neuron is governed by a fundamentally stochastic (thermodynamic) process, and this stochasticity extends to the populations of neurons which act as basic computational units (Glimcher 2005).[11] It has also been demonstrated empirically that the instantaneous perception of the attributes of a stimulus is stochastic even when all properties of the stimulus and state of the chooser are held constant (Stevens 1961). This stochasticity in subjective perception has been shown to be an obligate feature at all levels of sensory processing (see Glimcher 2011, for an overview; Beck et al. 2012, Woodford 2014, for relation to optimality), and this would necessarily lead to stochasticity in preferences.[12]

However, the stochasticity of neural activity extends beyond sensory processing, particularly to the neural circuitry necessary for comparison and implementation of motor actions. A class of models of this process, referred to as Bounded Accumulation Models (BAM), posit the dynamic accumulation of a decision signal to a threshold given a value input.[13] In the well-known *drift diffusion* model (Ratcliff 1978, Fehr and Rangel 2011), the relative values of the alternatives determines the slope of the accumulation, which determines the choice probabilities. A tight relationship exists between these Bounded Accumulation models and the stochastic choice literature; the choice probabilities of a BAM can be represented by a Fechnarian RUM, therefore

[10] See also Apesteguia and Ballester (2018) for critical issues with estimation.

[11] Neural activity shows significant variation even under conditions in which measurement error can be shown to be near zero (Tolhurst et al. 1983, Churchland et al. 2010, 2011). It is widely held that this is not simply a high dimensional signal of zero stochasticity projected imperfectly into a low dimensional space via limitations in measurement. For more on this issue, see Rieke et al. (1997) and Shadlen and Newsome (1998).

[12] To take one example, variability in the valuation of a sweet tasting liquid can arise from variability in the sensory experience of sweetness, even when the objective sugar concentration is held constant.

[13] Neural evidence for such dynamics in neural activity has been uncovered both in psychophysical and economic choice tasks (Gold and Shadlen 2007, Basten et al. 2010, Hare et al. 2011), as well as behavioural evidence for the role of decision dynamics and attention (Milosavljevic et al. 2010, Krajbich et al. 2010).

imply a cardinal random utility representation Webb (2018).[14]

The mixture of behavioural and neural evidence for both a Random Preference and an Fechnerian RUM approach suggests that the least restrictive econometric specification should be composed of a stochastic valuation (which may be restricted by a particular theory) and a subsequent error term which is, in essence, cardinal and strictly welfare decreasing. We stress that it may, in some cases, be necessary to restrict the econometric problem to a model with only one (or a linear combination) of these sources of stochastic choice for the purpose of identification, depending on the nature of the data being analyzed. Indeed this will be the case for our current dataset. However we present here the more general case as a starting point for theory, and note explicitly our identification assumptions.

3 Neural Random Utility Model

We now adapt the standard framework for stochastic choice in economics, random utility maximization, to a form that explicitly treats subjective value as a stochastic neural observable. We present the model for a binary choice set $\{i, j\}$, and we observe repetitions of all binary sets from the same subject. The extension of the model beyond binary choice is straightforward, though we note special considerations.

The subjective value of item i on trial t is defined to be an observable random variable $v_{i,t} \in \mathbb{R}_+$, with the vector of subjective values denoted $\mathbf{v}_t \in \mathbb{R}^n_+$. In principle, $\mathbf{v}_t$ is observable in the firing rate activity of value-related neurons.[15] We assume $\mathbf{v}_t$ is independent over trials, but not necessarily over items. Although we do not yet formally specify a distribution for $\mathbf{v}_t$, let us define $\nu_{i,t}$ as the difference between $v_{i,t}$

[14]This also clarifies the relationship between BAMs and the NRUM. Accumulation models place restrictions on the form of the NRUM and will prove invaluable for exploring a more structural approach to modelling decisions. However we do note that the NRUM brings a large econometric toolbox to bear for relating neural observables to choice prediction and for testing the predictions of more structural approaches with weaker assumptions on functional forms.

[15]Electrophysiological evidence for such observables can be found in Padoa-Schioppa (2013) and Rich and Wallis (2016).

and its mean $E[v_{i,t}]$, for each item,

$$\nu_{i,t} \equiv v_{i,t} - E[v_{i,t}]. \tag{8}$$

We emphasize that $v_{i,t}$ is the only observable in (8) and we provide a distributional assumption shortly.[16] Note that the distribution of $v_{i,t}$ puts no restrictions on the covariances over items, allowing a random preference formulation.[17] We discuss this issue further in Section 7.

Once subjective values are instantiated in neural activity, they must be compared and a choice executed. This additional neural process, which we refer to as the "choice mechanism", effectively compares subjective values in the requisite circuitry for producing behaviour. The neural evidence suggests this comparison takes place via an accumulation of $v_{i,t}$ to threshold in dorso-medial and parietal regions of cortex (Basten et al. 2010, Hare et al. 2011, Domenech et al. 2017). Webb (2018) demonstrates this process is equivalent to a random utility formulation with an additive noise term $\eta_{i,t} \in \boldsymbol{\eta}_t \in \mathbb{R}^n_+$ which captures stochasticity in this maximization operation. This yields the decision vector

$$\mathbf{u}_t = \mathbf{v}_t + \boldsymbol{\eta}_t, \tag{9}$$

For a binary choice trial t, the subject chooses i from the pair of items $\{i, j\}_t$ if

$$u_{i,t} > u_{j,t}$$

$$v_{i,t} + \eta_{i,t} > v_{j,t} + \eta_{j,t}.$$

[16] One possible interpretation of $E[v_{i,t}]$ is a 'core' value, instantiated noiselessly by some biological mechanism, but represented with error in the neural substrate under observation. This is not a view compatible with the biophysical properties of neural processes. Instead, we interpret $E[v_{i,t}]$ as simply the limiting quantity of the sample mean of $v_{i,t}$ and our definition of $\nu_{i,t}$ in an additive specification is for the purpose of exposition.

[17] However we do restrict the variance to be constant over items. In addition, there is the question of whether the central tendency of subjective value is stable or if it can be manipulated through contextual effects; for the purposes of this experiment, we assume a stable mean over trials.

yielding a probability of choosing i on trial t

$$\begin{aligned} P_{ij,t}\left(v_{i,t}, v_{j,t}\right) &= \Pr\left\{v_{i,t} - v_{j,t} > \eta_{j,t} - \eta_{i,t}\right\} \\ &= \Pr\left\{\tilde{v}_{ij,t} > \tilde{\eta}_{ji,t}\right\}, \end{aligned} \tag{10}$$

where $\tilde{v}_{ij,t} \equiv v_{i,t} - v_{j,t}$. The notation $\tilde{\cdot}_{ij}$ denotes the ijth item-pair difference throughout. Since the differences in measurements of subjective value determine these probabilities, this model now exhibits properties of cardinality (Batley 2008, p47).

Equation (10) is the conditional probability of choosing i given a measurement of subjective value *during a choice*. Before we arrive at a specification suitable for our empirical application, we must take two additional steps. First, we will need to impose some distributional structure on $\boldsymbol{\eta}_t$, therefore we assume that the difference in additive noise is independent over item-pair and trial, and distributed normally $\tilde{\eta}_{ji,t} \sim \mathcal{N}(0, \sigma^2_{\tilde{\eta}})$.[18] This yields a probability of choosing i

$$P_{ij,t}\left(v_{i,t}, v_{j,t}\right) = \Phi\left(\frac{\tilde{v}_{ij,t}}{\sigma_{\tilde{\eta}}}\right), \tag{11}$$

where $\Phi()$ is the standard normal CDF.

Second, our experimental application attempts to relate subjective value measures *in the absence of choice* to subsequent choice behaviour. By design, we do not observe the realization of subjective value $v_{i,t}$ on the trial t in which the choice was made, therefore specification (11) is inappropriate for analysis. Though an observation of $v_{i,t}$ in synchrony with the choice of our subjects would yield both the best predictive results and sharpest inference, the choice probability can also be derived conditional on the mean of subjective value $E[v_{i,t}]$, and not just its realization on a choice trial.

To demonstrate this, let us assume $\nu_t \sim \mathcal{N}(0, \Omega_\nu)$ with covariance matrix Ω_ν. Since our experiment uses a binary choice environment, the realizations of $\tilde{\nu}_{ij,t}$ for different item-pairs must occur on different trials t. Therefore the $\tilde{\nu}_{ij,t}$ are independent over

[18] There is little known about the appropriate distribution of η_t at this level of aggregation, though Webb (2018) provides a derivation directly from bounded accumulation models. The assumption of independence over item-pair is only made for convenience, see footnote 19.

ij due to independence over trials, even for different item-pairs that share an item.[19] Therefore $\tilde{\nu}_{ij,t}$ is distributed $\mathcal{N}(0, \sigma^2_{\tilde{\nu}})$, and this yields a probability of choosing i,

$$\begin{aligned} P_{ij,t}\left(E[v_{i,t}], E[v_{j,t}]\right) &= \Pr\left\{E[v_{i,t}] - E[v_{j,t}] > \nu_{j,t} - \nu_{i,t} + \eta_{j,t} - \eta_{i,t}\right\} \\ &= \Pr\left\{E[\tilde{v}_{ij,t}] > \tilde{\nu}_{ji,t} + \tilde{\eta}_{ji,t}\right\} \qquad (12) \\ &= \Phi\left(\frac{E[\tilde{v}_{ij,t}]}{\sigma_{\tilde{\nu}+\tilde{\eta}}}\right), \qquad (13) \end{aligned}$$

where $\sigma_{\tilde{\nu}+\tilde{\eta}}$ is the standard deviation of the sum of the two neural noise terms $\tilde{\nu_t}$ and $\tilde{\eta_t}$. This term reflects the degree of stochasticity in choice due to stochasticity in neural activity. Clearly, predictive accuracy is worse under this specification since $\sigma_{\tilde{\nu}+\tilde{\eta}} > \sigma_{\tilde{\eta}}$.

However $E[v_{i,t}]$ is not an observable, therefore equations (12) and (13) should be viewed as the limiting probabilities given a sample mean that approaches $E[v_{i,t}]$. The sample analog, derived from repeated measurements of $v_{i,t}$, is

$$\begin{aligned} P_{ij,t}(\bar{v}_i, \bar{v}_j) &= \Pr\left\{\tilde{\bar{v}}_{ij} > \tilde{\bar{\nu}}_{ij} + \tilde{\nu}_{ji,t} + \tilde{\eta}_{ji,t}\right\} \qquad (14) \\ &= \Phi\left(\frac{\tilde{\bar{v}}_{ij}}{\bar{\sigma}_{\tilde{\nu}+\tilde{\eta}}}\right), \qquad (15) \end{aligned}$$

where $\bar{\sigma}_{\tilde{\nu}+\tilde{\eta}} \to \sigma_{\tilde{\nu}+\tilde{\eta}}$ as $\tilde{\bar{\nu}}_{ij} \to 0$. This is the specification we will work from in our empirical setting.

4 Testing a NRUM with Behavioural and Neural Measurements

We now establish the NRUM as an econometric toolset for relating neural observables to choice prediction in an experimental dataset. In section 4.1, we apply the model to

[19] The extension of the model beyond binary choice would have to account for a full covariance matrix for the vector composed of the $\tilde{\nu}_{ij,t}$ on each trial (similarly for the $\tilde{\eta}_{ij,t}$). In principle, a full covariance matrix should be identifiable for such a dataset (Hausman and Wise 1978, Train 2009) and the results that follow would have to be argued in terms of this full matrix. The assumption of normality for $\mathbf{v}_t$ is again made for convenience. To our knowledge no study has yet examined the distribution of the aggregate firing rates that make up subjective value.

a combined dataset of choices and neural measurements from two brain regions known to encode subjective value (mPFC and Striatum) and one control region (OCC). A detailed description of the Levy et al. (2011) experiment, including the BOLD measure of neural activity, can be found in Appendix 10.1. The role measurement error plays in the relationship between the BOLD measure and choice behaviour is examined in section 4.2.

In the analysis, we treat the item-pair and the two choices made in each pair as the dimensions of our behavioural dataset, and pool item-pairs over subjects. For 12 subjects, this yields $n = 4560$ choices grouped into 2280 pairs.[20] Essentially we are treating different subjects viewing the same item-pairs as equivalent to the same subject viewing different item-pairs. While this allows each subject's preferences – therefore subjective valuations – to be idiosyncratic, it does contain the implicit assumption that the relationship between subjective valuation, the BOLD measure, and the choice likelihood is the same across subjects. We relax this assumption in section 4.3 at the expense of a reduced sample size.

4.1 A Cardinal Neural Observable

The random utility model specifies that the difference in utility influences choice likelihood, and therefore posits that utility is a cardinal quantity. To establish that our neural observable is cardinal, we must establish that neural measurements are made on some scale in which the difference between measurements is related to the likelihood that a subject will switch their choice behaviour in repeated choice sets, and that this difference predicts choices beyond a simple ordering.

In the Levy et al. (2011) experiment, measurements of BOLD activity from mPFC and Striatum were taken on 11 scanning trials independently for each good over the course of an hour. The measurements preceded — and were independent of — the two choice trials of interest. We use the time index m to denote these measurement trials,

[20] Striatal activation was not recorded for one subject, so analysis on this brain area will use 4180 choices grouped into 2090 pairs.

and use the general notation $B_{i,m}$ to denote a measurement from one of these regions (we will report results for each brain area separately).

We assume a linear form for the relationship between the BOLD measurement $B_{i,m}$ from a brain region and subjective value $v_{i,m}$.

$$B_{i,m} = a + \gamma v_{i,m} + \mu_{i,m}.$$

The error term $\mu_{i,m} \sim N(0, \sigma_\mu^2)$ reflects the error present in measuring neural activity in an MRI scanner, therefore a neural measure of subjective value $B_{i,m}$ has two sources of variance: the fluctuation in subjective value on our measurement trials, and measurement error. To arrive at a measure for predicting choice between items i and j on an independent trial t, we average over our 11 measurements and then take the difference.

$$\bar{B}_i = a + \gamma \bar{v}_i + \bar{\mu}_i \tag{16}$$

$$\tilde{\bar{B}}_{ij} = \gamma \tilde{\bar{v}}_{ij} + \tilde{\bar{\mu}}_{ij}. \tag{17}$$

Initially, we proceed under the assumption that there is no sampling and measurement error, $\tilde{\bar{B}}_{ij} = \gamma E[\tilde{v}_{ij,t}]$. While this assumption is clearly not valid, it does lead to some useful intuition for the full model in section 4.2. Specifically, assuming an error-free measure of the mean of subjective value allows us to use specification (13). Substituting in (17) yields a probability of choosing i,

$$P_{ij,t}(E[\tilde{v}_{ij,t}]) = \Phi\left(\frac{E[\tilde{v}_{ij,t}]}{\sigma_{\tilde{\nu}+\tilde{\eta}}}\right) \tag{13}$$

$$= \Phi\left(\frac{\gamma^{-1}}{\sigma_{\tilde{\nu}+\tilde{\eta}}} \tilde{\bar{B}}_{ij}\right). \tag{18}$$

Under this specification, the NRUM makes three predictions about the likelihood our subject will choose item i. First, if behaviour was only determined by the ordinal comparison $v_{i,t} > v_{j,t}$ on a given choice trial, then the average measurement of each

good over repeated independent measurement trials should contain no predictive information. By contrast, the NRUM predicts that as $\tilde{\tilde{B}}_{ij}$ increases, the subject should be more likely to choose item i on any given choice trial (see Figure 2.A).

Second, recall that subjects made choices over each item-pair twice. Therefore the likelihood that a subject switches their choice upon repeated trials should decrease with the absolute value of $\tilde{\tilde{B}}_{ij}$.

Third, if we segregate our item-pairs into those pairs in which the subject chose item i twice, once, or never at all as a function of $\tilde{\tilde{B}}_{ij}$, the NRUM would predict $P(twice) > P(once) > P(never)$ for a positive difference in measured subjective value. This prediction is depicted in the right panel Figure 2.A, in which choices were simulated according to the NRUM, then the number of *twice*, *once*, and *never* observations were fit using an Ordered Probit model.

Table 1 presents the estimates from bringing (18) to our dataset with the normalization $\sigma_{\tilde{\nu}+\tilde{\eta}} = 1$. This standard identification assumption means we are estimating only the *relative* relationship between neural activity and the choice probabilities. We also included a specification with a constant term c predicted to be zero by the model: $\Phi\left(c + \frac{\gamma^{-1}}{\sigma_{\tilde{\nu}+\tilde{\eta}}}\tilde{\tilde{B}}_{ij}\right)$. For both the mPFC and the Striatum, the estimate for γ^{-1} is positive, therefore the relationship between the difference in neural measurement ($\tilde{\tilde{B}}_{ij}$) and the probability of choosing an item is indeed monotonic (see Figure 2.B for the mPFC). As might be expected, no such relationship is found in the OCC control region (Table 1).

To test the second prediction, we repeat the analysis conducted by Agranov and Ortoleva (2017) on their lottery choice dataset. An indicator variable codes item-pairs in which subjects switched their choice on repeated trials. Table 2 presents the results of a random-effects GLS regression of this indicator variable on $|\tilde{\tilde{B}}_{ij}|$. A clear negative relationship between the magnitude of the difference in BOLD activity and the likelihood of the subject switching their choice is observed in this sample. This lies in contrast to the results from Agranov and Ortoleva (2017), which found no relationship between a behaviourally-established measurement of subjective value and

choice stochasticity.

However the third prediction of the NRUM does not fare as well, at least at first glance. The fit of the Ordered Probit model to the number of observed choices has a clear mis-ordering; subjects are more likely to choose an item twice, than never, than once for positive $\tilde{\tilde{B}}_{ij}$. We observe too few *once* choices when $\tilde{\tilde{B}}_{ij}$ is small, too many when it is large, and far too many *never* choices when $\tilde{\tilde{B}}_{ij}$ is large and positive (similarly for *twice* when it is large and negative). This apparent contradiction of the NRUM arises because we (like much of the neuroscience literature working with choice data) have so far assumed no error in both our BOLD measurement and the construction of our neural measure $\tilde{\tilde{B}}_{ij}$. The following section addresses this issue.

4.2 Accounting for Measurement Error

We can identify at least three source of measurement error in our dataset. First, since we are not measuring subjective value during a choice trial, the realizations of $v_{i,m}$ we do measure are not the ones related to choice on trial t. This component of our measurement error is the sampling error present in $\bar{v}_i$ and is denoted by $\bar{\nu}_i$ in (8). Second, we should also allow for error in the conservative procedure for identifying and constructing a single neural time-series from the 250,000 we measured. The degree to which the mean activity level of our measure captures the neural encoding of subjective value for consumer items depends on our ex-ante restriction to the mPFC and Striatum and the accuracy with which our first procedure identifies the relevant voxels. This source of variability is captured in $\mu_{i,m}$. A third source of noise doubtlessly results from the technical limitations imposed by measuring neural activation with an fMRI scanner (Logothetis 2002), which is also captured in $\mu_{i,m}$

The effect of measurement error in non-linear models (such as the Probit) is larger than in the linear model, but generally follows the same intuition: the data is over-dispersed along the dimension of the independent variable and the slope parameter is biased towards zero (Yatchew and Griliches 1985). Formally, we can no longer work directly from specification (13) since $P_{ij,t}(\tilde{\tilde{B}}_{ij})$ is no longer equivalent to $P_{ij,t}(E[\tilde{v}_{ij,t}])$.

This means our estimate of γ^{-1} in section 4.1 is biased towards zero and the severity of this bias increases in the degree of measurement error. Since our hypothesis predicts a positive value for γ^{-1}, inference performed on this biased estimate is still valid, though pursuing a less biased estimate will yield improved inference and choice prediction.

Recalling equation (16), measurement error enters our specification as an item-specific i.i.d. error term.[21] If we proceed with a specification derived from substituting in our measured neural activation into the sample analog (14), the conditional probability of choosing i is

$$\begin{aligned} P(y_{ij,t} = i \mid \tilde{\tilde{B}}_{ij}) &= P\left(\gamma^{-1}(\tilde{\tilde{B}}_{ij} - \tilde{\tilde{\mu}}_{ij}) > \tilde{\tilde{\nu}}_{ij} + \tilde{\nu}_{ji,t} + \tilde{\eta}_{ji,t}\right) \\ &= P\left(\gamma^{-1}\tilde{\tilde{B}}_{ij} - e_{ij} > \tilde{\nu}_{ji,t} + \tilde{\eta}_{ji,t}\right), \end{aligned}$$

with the sources of measurement error grouped together in the variable $e_{ij} \equiv \gamma^{-1}\tilde{\tilde{\mu}}_{ij} + \tilde{\tilde{\nu}}_{ij}$.

The fact that subjects chose between each item-pair twice means that e_{ij} is constant over both choice trials. This means we have two independent choices for each realization of the measurement error. Or said another way, the e_{ij} are (perfectly) correlated over repeated choice trials. We can use this correlation pattern to achieve more efficient (and less biased) estimates of γ^{-1} — as well as an estimate of the standard deviation of the measurement error — provided we specify and integrate out a distribution for e_{ij}. We assume $e_{ij} \overset{iid}{\sim} \mathcal{N}(0, \sigma_e^2)$, therefore our specification takes the form of a random-effects Probit model, however with two important caveats that differ from standard applications.[22]

1. $\tilde{\tilde{B}}_{ij}$ and e_{ij} are not independent. This means that the random-effects Probit estimate of γ^{-1} will also be biased towards zero, though not as severely as a

[21] This form of measurement error is referred to as "classical measurement error" since the error is additive and independent of the unobserved quantity (Carroll et al. 2006). It specifies that our neural measurement $\tilde{\tilde{B}}_{ij}$ has a larger variance than the unobserved quantity of interest, a natural assumption in the context of measuring neural activity with a noisy fMRI signal.

[22] A random-effect model is robust to the distributional assumption for the random-effect (here, measurement error) provided it is not highly asymmetric (Neuhaus et al. 2011).

Probit with no random-effect. Therefore, we can only partially correct for the bias introduced by measurement error.

2. The e_{ij} are not independent over choice pairs. Since the neural measurement takes place at the level of the individual item, when differencing the measurement for an item-pair there is correlation in the random effect e_{ij} between item-pairs that share an item. For instance, e_{12} and e_{13} are correlated because they share the measurement of item 1. This means a random-effects estimate is inefficient, and standard errors will be biased towards zero if not controlled for. In addition, the estimate of σ_e will be biased positively (Wang et al. 1998).

To account for these issues, we pursue a hybrid approach in which we estimate the random-effects model clustered at the level of the item-pair (to capitalize on the common measurement error over choice trials within an item-pair, partially reducing the bias and achieving more efficient estimates), then correct our standard errors for inference using a multi-way clustering approach (to account for the non-independence of the differenced measurement errors). The item-pair level likelihood is then given by

$$P\left(y_{ij,1}, y_{ij,2}|\tilde{\tilde{B}}_{ij}\right) = \int_{-\infty}^{\infty} \frac{\mathrm{e}^{-{e_{ij}}^2/2\sigma_e^2}}{\sqrt{2\pi}\sigma_e} \left[\prod_t F\left(y_{ij,t}, \tilde{\tilde{B}}_{ij}\right)\right] de_{ij}, \tag{19}$$

where

$$F(y,x) = \Phi\left(\frac{\gamma^{-1}\left(x - e_{ij}\right)}{\sigma_{\tilde{\nu}+\tilde{\eta}}}\right)^y \left[1 - \Phi\left(\frac{\gamma^{-1}\left(x - e_{ij}\right)}{\sigma_{\tilde{\nu}+\tilde{\eta}}}\right)\right]^{1-y}.$$

We also include a specification with a constant term (predicted to be zero).

Including a correction for measurement error substantially increases the fit of the NRUM (Table 1), with the log-likelihood(s) improving by nearly a factor of $\frac{1}{3}$. The estimated coefficients for γ^{-1} are also substantially higher than our baseline specification, increasing by roughly a factor of 5 in both the mPFC and Striatum. This indicates that the relationship between neural activity and choice probability is severely biased when measurement error is unaccounted for. Figure 3 depicts the fitted probability of choosing item i as a function of the difference in neural activity (generated under the

assumption that the random-effect $e_{ij} = 0$). Accounting for measurement error yields a significant increase in the magnitude of the relationship between neural activity and choice probability compared to our earlier analysis in Section 4.1. Moreover, the difference in neural activity yields improved model fit compared to a simple ordinal ranking of the BOLD activity, again establishing the cardinality of our neural measure.

In both mPFC and Striatum, the standard deviation of the measurement error σ_e is estimated to be ~ 4.7 times $\sigma_{\tilde{\nu}+\tilde{\eta}}$. Therefore in both specifications, over 95% of the variance in the model is attributed to measurement error. To verify that measurement error is generating the results observed in Section 4.1, we introduced measurement error into the simulated data reported in Figure 2.A and repeated the original analysis. These simulated results now match our empirical findings (Figure 2.C). Because measurement error has the effect of "smearing" the observed *once* choices over the range of observed $\tilde{\tilde{B}}_{ij}$, a choice pair in which the distributions of subjective value are close together (small $E[\tilde{v}_{ij,t}]$) – likely resulting in a *once* outcome – could yield a large $\tilde{\tilde{B}}_{ij}$ because of measurement error. This occurs because the degree of measurement error has no effect on the number of *once* choices observed, only on where they appear on the $\tilde{\tilde{B}}_{ij}$ axis. While this degree of measurement error is striking, and verified by simulation (Figure 2.C), we should note again that this estimate is based on a misspecification of the random-effect.

4.3 Subject-Specific Analysis

In principle, a subject-specific analysis is useful to consider. Commensurate with existing data and previous fMRI studies (Logothetis 2003), it is likely that different subjects have a steeper mapping between the BOLD measurement and neural activity than do others. The bulk of this difference is typically held to reflect a technical feature of the interaction between the scanner and the subject: the subject-specific coefficient describing the coupling of neural activity to the blood flow rate measured by fMRI.

We can capture such heterogeneity by allowing equation 16 to vary by subject s,

$$B_{s,i,m} = a + \gamma_s v_{i,m} + \mu_{s,i,m} \tag{20}$$

$$\tilde{\tilde{B}}_{s,ij} = \gamma_s \tilde{\tilde{v}}_{s,ij} + \tilde{\tilde{\mu}}_{s,ij}, \tag{21}$$

The parameter γ_s is therefore a subject specific relationship between the neural measurement and subjective value. We can estimate γ_s^{-1} through a subject-$\tilde{\tilde{B}}_{ij}$ interaction term using specification (19) on the full sample, however with only 380 observations per subject.

Breaking up the sample into so few observations per subject reveals the limits of discrete choice estimation methods in small samples. For the mPFC, six of the subjects yield positive and significant estimates of γ_s^{-1}, while six are not significantly different from zero (Table 5). While under the null hypothesis we should only expect one significant subject, rather than six, this is still a substantial degree of variance in the model.[23] Results from the Striatum display a similar pattern, though with a somewhat larger amount of variation. Nine of eleven estimates are positive, though only two significantly so at the .10 level. In both the mPFC and the Striatum, the AIC is higher than for the pooled estimates, even after correcting for measurement error. This suggests that a subject-specific estimate of the relationship between neural activity and choice is limited in small samples, and that pooling data to estimate this relationship yields improved fit (provided preferences are allowed to vary across subjects).

5 Application: In-Sample Choice Prediction

The NRUM yields an estimated relationship between neural activity (or other observables) and choice behaviour. In the analysis that follows, we compare the performance

[23] Monte carlo simulations verify the loss in efficiency due to reducing observations. Simulated choice and neural data with $\gamma_s^{-1} = 10$ and measurement error from section 4.2 leads to ~5% of the γ_s^{-1} estimates less than, but not significantly different from, zero (from a total of 1000 simulations).

of three models:

- NRUM: subject specific estimates from (Table 1),
- NRUM w/ m.e. correction: subject specific estimates corrected for measurement error (Table 5),
- NRUM + observables: corrected subject specific estimates with additional economic observables included as regressors: the price of the item (a market-based method) and its 'Amazon star' rating (a stated-preference method).[24]

The estimates from each of these models yields a fitted choice probability for each choice pair $\hat{P}_{ij}$, and these probabilities can be combined with a prediction rule in order to assess the model's true predictive performance for any set of neurobiological observables. Numerous methods have been proposed to evaluate the performance of discrete choice models and the literature. For instance, to determine whether a cardinal prediction rule captures stochastic choice behaviour, some method for pooling over discrete choices is required (after all, we wish to compare a probability to a binary outcome). The simplest way to achieve this is to average over repeated choices from the same choice set.[25] For this reason, we will examine the results from different choice prediction rules when repeated choice trials are treated both independently and jointly. For exposition, we will focus on in-sample prediction rates for measurements from mPFC.

[24]The 'Amazon star' rating is the aggregation of user ratings that can be found on the item's description on amazon.com. Both of these measurements have the drawback of being population level variables which represent (to some degree) the aggregation of preference across all consumers, limiting their ability to predict individual choices. However, both of them were significant predictors. The Amazon rating varied positively with the choices of our subjects, suggesting some homogeneity in the preferences of New York University undergrads, while prices varied negatively with choice. One might expect subjects to be choosing high priced goods (which they receive at no monetary cost in the experiment), but likely reflects the popularity of the CDs in our choice set, a relatively inexpensive item.

[25]For another method which pools "locally" over nearby choicesets with similar predicted probabilities (see Smith et al. (2014)).

5.1 Treating each choice trial independently

We first examine in-sample predictive performance in which each trial is treated independently, regardless of whether it comes from the same choice set, and consider the following prediction rules:

- *Bayes Classifier*: This is the prediction rule typically reported in statistical software. The fitted probabilities $\hat{P}_{ij}$ codes a predicted choice $\hat{y}_{ij,t} = 1$ if $\hat{P}_{ij} > 0.5 \; \forall t$, $\hat{y}_{ij,t} = 0$ otherwise. The prediction is then compared to the observed choice $y_{ij,t}$ and the rate of successful predictions reported. The Bayes Classifier essentially nullifies the cardinality of the analysis, and magnifies the sign difference of the neural observable (i.e $\hat{\gamma}^{-1}\tilde{\tilde{B}}_{ij} > 0 \Rightarrow \hat{P}_{ij} > 0.5$) yielding a deterministic prediction (i.e. all trials from the same choice pair will have the same prediction).

- *Bernoulli Prediction Rate*: The predicted probability of the observed outcome for choice pair ij on trial t is $y_{ij,t}\hat{P}_{ij} + (1 - y_{ij,t})(1 - \hat{P}_{ij})$. Averaging this predicted probability over all ij and t gives the proportion of successful predictions if each trial is treated as an independent draw from a Bernoulli distribution with probability $\hat{P}_{ij}$.

- *Cramer's* λ: Let $\bar{P}^{+}$ and $\bar{P}^{-}$ denote the average predicted probability on trials in which $y_{ij,t} = 1$ and $y_{ij,t} = 0$ respectively. Then $\lambda \equiv \bar{P}^{+} - \bar{P}^{-} \in [0, 1]$ reflects the ability of the model to discriminate between outcomes, and measures the proportion of total variation in y that is 'explained' (Cramer 1999). A $\lambda = 0$ represents the null model predicting at chance, while $\lambda = 1$ represents perfect discrimination.

Results from the prediction rules are presented in Figure 4. The improvement granted by the correction for measurement error can be seen in the results for Cramer's λ. The correction improves the discriminability of the NRUM to 0.16, an improvement of 0.10. Compared with the prediction rate of 55.7% in Levy et al. (2011), in-sample prediction rates increase to 57.9% using the NRUM estimates corrected for measurement error. The naive prediction rate for the NRUM estimates is 60.3%, with the

improvement over the Levy et al. (2011) analysis coming from the three subjects with negative estimates of γ^{-1}.[26] This marginal improvement in prediction rates highlights the limitations of assessing a cardinal model when each choice is treated independently.[27]

Moreover, note that these prediction rates are still lower than those derived from a model which only includes the price and quality observables (64.5% and 65.4%, depending on the prediction rule). Combining the neural measurements with these additional variables increases prediction rates further to 69.6% and 71.0% (depending on the prediction rule), and significantly improves the discriminability of the model, suggesting the individually-measured neural activity contains information orthogonal to the aggregate observables.

5.2 Treating repeated choice trials jointly

To assess the ability of a cardinal model to capture stochastic choice behaviour, we propose two possible methods for comparing predictive performance when repeated choice sets are treated jointly.

- *Conditional on choice outcome*: Let the vector $z_{ij} \in \{0, 1, 2\}$ represent whether item i was chosen *never*, *once*, or *twice* from a choice pair ij. Let N denote the total number of choice pairs, and $\alpha_2 = \frac{\sum 1_{z_{ij}=2}}{N}$ the proportion of *twice* observed in the dataset, with α_1, α_0 defined accordingly. The predicted probability of the observed outcome for choice pair ij is given by $P^*_{ij} \equiv (2 - |z_{ij} - 1|)\hat{P}^{z_{ij}}_{ij}(1 - \hat{P}_{ij})^{2-z_{ij}}$. Averaging this predicted probability conditional on the outcome then gives a measure of how well the model predicts the sample of observed outcomes. For example, $\frac{\sum_{\{ij:z_{ij}=2\}} P^*_{ij}}{\#\{ij:z_{ij}=2\}}$ is the average probability of a correct prediction of a *twice* outcome. However this approach ignores the fact that outcomes in the

[26] In comparison, Smith et al. (2014) report a 61.3% out-of-sample prediction rate, while the rate reported here is in-sample. This improvement in their choice prediction likely arises from a more sophisticated aggregation of the BOLD data than used in this study.

[27] All of these results are robust to reserving half of the sample for estimation, then implementing the prediction exercise on the holdout sample.

sample occur in different proportions.[28]

- *Conditional on choice prediction*: Consider a predicted choice $\hat{z}_{ij} \in \{0, 1, 2\}$ drawn from the binomial distribution of size 2, with probability of success $\hat{P}_{ij}$. A correct prediction of, for example, $\hat{z}_{ij} = 2$, would be observed with probability $\hat{P}_{ij}^2$ for choice pairs on which a *twice* outcome occurred, and 0 otherwise. Summing this prediction rate over all choice pairs, and dividing by $\sum_{ij} \hat{P}_{ij}^2$, therefore "weights" predicted probabilities by the proportion in which the outcomes are observed in the data. For intuition, consider taking R draws of $\hat{z}_{ij}$. The measure is equivalent to calculating the number of correct predictions in this simulated sample, conditional on the prediction being *twice*, *once*, or *never*.

The distinction between the two prediction rules is important, because for our entire sample, the frequency of *never* is 46.0%, *once* is 9.3%, and *twice* is 44.8%. If each individual choice were predicted at chance, we would predict *never* on $\frac{1}{4}$ of trials, *once* on $\frac{1}{2}$, and *twice* on $\frac{1}{4}$, and we would be correct on $\frac{1}{4} \times 46.0 + \frac{1}{2} \times 9.3 + \frac{1}{4} \times 44.8 = 27.4\%$ of trials. Therefore the prediction rates arrived at by chance depend on the distribution of *never*, *once*, or *twice* in the dataset. In such a null model, $\hat{P}_{ij} = \frac{1}{2}$, and the predicted probability of a *twice* outcome is $\hat{P}_{ij}^2 = \frac{1}{4}$. Therefore the prediction rate conditional on the prediction yields

$$\begin{aligned} \frac{\sum_{ij} 1_{z_{ij}=2} \hat{P}_{ij}^2}{\sum_{ij} \hat{P}_{ij}^2} &= \frac{\frac{1}{4} \sum_{ij} 1_{z_{ij}=2}}{\sum_{ij} \frac{1}{4}} \\ &= \frac{\frac{1}{4} \alpha N}{\sum_{ij} \frac{1}{4}} \\ &= \alpha, \end{aligned}$$

the proportion of *twice* outcomes observed in the sample. As the predictive power of the model improves, this measure approaches 1.

However an ordinal prediction based solely on the ordered BOLD activity, such as

[28] For a similar argument in the case of an independent binary choice trial in which the observed outcomes are not in equal proportion, see Cramer (1999).

in Levy et al. (2011) predicts an item will be chosen either *twice* or *never*, and can not account for trials in which an item was chosen only *once*. Since the NRUM uses the cardinal difference in valuations to modulate the choice probabilities, it can be combined with the above prediction rules to predict such behaviour.

The results from this exercise are reported in Table 4.[29] We find that the NRUM predicts 28.6% of such trials observed in the dataset, compared to 0% for the ordinal model. Here, again, we see the (often overlooked) effect of measurement error in the observable. Because our measurement error correction increases the discriminability of the predicted probabilities, it reduces the number of a *once* outcomes predicted compared to the uncorrected estimates and the *null* model. But because the correction only measures the variance of the measurement error, as opposed to its realization on any trial, there are still many *once* outcomes observed in which the predicted probability $\hat{P}_{ij}$ is near 0 or 1, decreasing the prediction rate of *once* outcomes from its upper bound of 50%. This improvement in discriminability, however, does improve the number of *twice* and *once* predictions (45.0% vs. 30.8%), and was also more accurate conditional on whether an item was predicted to be chosen *never* or *twice* from a pair (56.8% vs. 51.8%, and 56.1% vs. 50.8%).

6 Application: Estimating Demand

One proposed advantage of neuroeconomic methods is a richer datasource on which to assess the demand for new products (Ariely and Berns 2010). Consider a standard demand forecasting exercise for a new product i. A researcher sets out to assess the change in demand for this product from manipulating a characteristic (e.g. quality or price). Assume that this manipulation increases the underlying valuation of this product from v_i to v_i'. Section 4.2 details how the presence of significant error in neural measurements will bias the estimates of these marginal effects.

[29]For comparison pruposes, we also consider the *null* model (randomly selecting one item from each choice pair) and *known* benchmark model which sets the probability of choosing an item at 1 when it was chosen twice, 0.5 when it was chosen once, and 0 when it was never chosen from a pair.

To clarify this point, suppose the researcher has access to neural measurements B_i and B_i' to assess this manipulation. In addition, they make a neural measure for a reference product j, which for the sake of argument we normalize to B_j=0. The true change in demand, as a function of the neural measurements, is thus $P_i(B_i') - P_i(B_i)$ (Figure 6, solid black line). Note that the change in demand depends on the magnitude of the measurement. This relation between the magnitude of the marginal effect and the location of the measurement is a feature of any demand prediction exercise based on a discrete choice model.

For instance, the NRUM provides a predicted choice probability, $\hat{P}_i(B_i)$, as a function of the neural measure and the estimated marginal effect. Therefore the predicted change in demand from the manipulation is given by $\hat{P}_i(B_i') - P_i(B_i)$ (Figure 6, dashed line). Even a small increase in neural response to the manipulation, B_i', will lead to higher predicted demand $\hat{P}_i(B_i') - P_i(B_i) > 0$.

However the relation between the marginal effect and location of the measurement is also why measurement error can impact a demand prediction exercise. In a "naive" model which does not account for measurement error, the predicted probability is constructed via an estimate $\hat{\gamma}^{-1}$ which is biased towards zero. In the absence of a correction for this error, the predicted demand $\hat{P}_i(B_i')$ is (weakly) smaller than the true demand $P_i(B_i')$. The magnitude of this gap also depends on the difference between the neural measures. For some differences in the neural measurement the naive analysis will *underpredict* the change in demand from the product manipulation (B_i to B_i'). But suppose the researcher further manipulates the product, yielding a larger neural response B_i''. Now the naive analysis *overpredicts* the change in demand $\hat{P}_i(B_i'') - \hat{P}_i(B_i')$ relative to the true change $P_i(B_i'') - P_i(B_i')$. Since the correction for measurement error offered by the NRUM reduces the gap between the true and predicted probabilities, it yields a predicted change in demand that is closer to the true demand.

Our experimental dataset provides an opportunity to quantify the degree of this bias. The sample of choice objects contained five 50/50 lottery tickets over different dollar amounts ($10, $15, $20, $25, and $30 if win, and $0 if lose), so we can analyze

the change in demand as the amount of the winning outcome is increased, relative to the reference $10 lottery. Since the lottery amounts are monotonically increasing, subjects with completely transitive preferences should always choose the higher lottery (relative to the $10 reference lottery). Indeed, this is what we find in our data. Figure 6 also reports the predicted probabilities from the NRUM with and without the measurement error correction, taken at the average BOLD measurement for each lottery (across measurement trials and subjects). As expected, the degree of bias due to measurement error is large. For the baseline Probit model, the predicted change in demand for the larger lotteries is minimal (2% in mPFC and 5% in Striatum), considerably understating the change in demand for the larger lotteries. By contrast, the NRUM with measurement error correction is larger (roughly 10% in mPFC and 20% in Striatum).

This example illustrates a fundamental issue with predicting discrete choice outcomes in the presence of measurement error. Given the degree of measurement error we find in our neural measurement, it is paramount that prediction exercise account for this bias in the estimated relationship between neural activity and choice behaviour. At the very least, the bias correction proposed in Section 4.2 should be considered in future prediction exercises.

7 Normative Implications: Distribution of Subjective Value

The general formulation of a RUM places no *a priori* restriction on the distribution of utilities (Becker et al. 1963). In this version of the NRUM, we have attempted to formulate subjective value as generally as possible so that it might encompass the two predominant views about stochastic choice in the economics literature.

The NRUM is general enough to allow for a random preference interpretation since no restriction is placed on the distribution of $\mathbf{v}_t$, particularly its within-item covari-

ances. Therefore the stochastic valuations of each alternative can be correlated in accordance with the requirements of a particular random preference formulation. Of course, it is also possible to impose independence directly on $\mathbf{v}_t$, yielding a model in which the stochastic valuation of each alternative is processed independently. Since the NRUM renders the covariance matrix of $\mathbf{v}_t$ empirically observable, it is possible to differentiate between these views with an appropriate dataset. In this study, since the subjective values of items were measured independently, in isolation, and on different trials; we can safely assume that $v_{i,m}$ and $v_{j,n}$ are independent over different measurement trials m and n.[30]

Even after allowing for a random preference specification for subjective value, however, the NRUM still incorporates a Fechnerian stochastic element, modelled via the additive random vector $\boldsymbol{\eta}_t$. This error term arises from stochasticity in the choice process downstream from valuation regions. The distinction between these two neural sources of stochasticity has critical normative implications. If $\sigma_{\tilde{\eta}} = 0$, then all choice stochasticity is due to variation in subjective value and choice can be defined as optimal (in the traditional economic sense) because choosers then act to maximize their realized, albeit stochastic, subjective values. However, if $\sigma_{\tilde{\eta}} > 0$, then some choices can be classified as errors arising in the neural implementation of the maximization operation and the execution of the choice behaviour. Thus the relative sizes of $\sigma_{\tilde{\nu}}$ and $\sigma_{\tilde{\eta}}$ reflect the degree to which stochasticity in choice can be strictly viewed as welfare decreasing in a given neural dataset. Evidence from perceptual neuroscience (in which there is an objectively "correct" answer) identifies that most of the variance in choice stochasticity can be attributed to brain areas encoding stimulus value, suggesting less than 10% of choice stochasticity can be attributed to downstream neural circuitry which implements the choice (Michelson et al. 2013, Drugowitsch et al. 2016).

We should note that in all likelihood, $\boldsymbol{\nu}$ and $\boldsymbol{\eta}$ are the product of realizations at multiple points in the human nervous system. While we are unable to fully differentiate

[30]In an alternative dataset in which the subjective values of both items were measured simultaneously (i.e. $m = n$), this assumption would not be feasible, thus random preferences should be accounted for in the modelling. Examples of such studies include Chib et al. (2009) and Levy and Glimcher (2011).

between these two sources of variance in this specific study because we do not make independent measurements at multiple stages along the pathways that represent subjective value, we observe stochastic choice behaviour that has features of an additive random utility specification: a larger difference in subjective value makes an item more likely to be chosen. Our own conviction, which stems from an amalgamation of the economic and neurobiological literature, is that a model which incorporates both classes of stochasticity will most closely approximate the structure of human choice behaviour. We note that anchoring our model to this conviction effectively posits a distinction between the fraction of choice stochasticity that can be attributed to stochasticity in preference and the fraction that can be attributed to errors induced by the choice mechanism. This distinction has clear welfare implications that would necessarily be of interest as more is learned about these sources of stochasticity in choice behaviour (Bernheim 2009).

8 Conclusion

In this article, we have proposed a cardinal econometric framework, the Neural Random Utility Model, for relating neural observables to stochastic choice behaviour. The NRUM specifies the sources of stochasticity present in a measurement of neural activity, incorporating both the *Random Preference* and *Fechnerian* approach to modelling stochastic choice behaviour, and examines how these sources interact within an experimental paradigm for the purposes of choice prediction.

A concrete example of subjects choosing over consumer items was developed in detail. We find that neural activity, measured in isolation, predicts subsequent choice behaviour as has been previously argued, and that the magnitude of the difference in neural activity is positively correlated with the degree of stochasticity in choice (measured via the number of preference switches in repeated trials). These results establish that neural measurements carry cardinal information relevant for choice prediction.

However, we also find that measurement error limits the effectiveness of the neu-

ral observables far more than has been acknowledged in the literature. Econometric techniques available to the NRUM framework mitigate some of the impact of measurement error – yielding less-biased model estimates – provided that the experiment consists of repeated choice trials from the same choice sets. To assess the predictive performance of these measurements, we examined previously proposed prediction rules for choice trials treated independently, and propose new prediction rules appropriate for repeated choice trials from the same choice set. When choice trials are treated independently, the NRUM yields marginal improvements in choice prediction, primarily due to the correction for measurement error. However when repeated choice trials are treated jointly, the cardinality of the NRUM allows the model to better capture the distribution of choice outcomes compared to an ordinal model.

The measurement error correction we explore in this article utilizes a convenient property of the Levy et al. (2011) dataset, namely that each choice was repeated twice. This allows measurement error to be modelled as a random-effect which holds constant over repeated choices. Apart from the improvement in model estimates, this approach has the added benefit of providing identification of the standard deviation of the measurement error. However there are limitations to this method. Since the measurement error is correlated over trials, the random-effect is misspecified and the measurement error estimate will be biased positively. For this reason, the estimate $\hat{\sigma_e}$ provided here should be considered an upper-bound, though we do confirm via simulation that a considerable degree of measurement error is needed to match features of the observed data. In principle, unbiased estimates should be feasible provided that the correct structure of the random effect is specified. This would require devising an estimator which relaxes the independence assumption used here.[31]

Even after the measurement error correction, choice prediction performance barely matches two standard aggregate observables, the price and quality ratings of the items. Combining neural measurements and standard observables further improves choice

[31] In addition, simulation-based techniques for an unbiased estimate exist in the bio-statistics literature (Carroll et al. 2006, Chapter 5). Our simulation results (Figure 2) suggest σ_e is too large by roughly a factor of 2 for them to be applicable, but may soon become practical as technology improves.

prediction, suggesting that the neural observables provide subject-specific information. Of course, this improvement comes at a high implementation cost for brain-scanning technology (roughly $50,000 for Levy and colleagues to produce this dataset), limiting the prevalence and usefulness of current neural measurements.

Our approach to modelling choice prediction from neural observables thus offers four contributions to the literature. It establishes that neural measurements do carry cardinal information about the relative values of alternatives. It establishes the positive performance of neural measurements using fMRI technology, and defines clearly the benchmarking process that will be required for future measurement techniques. It offers a general framework for combined economic-neurobiological modelling from which both richer, more restrictive specifications can be developed. And finally, it lays out the basic welfare structure inherent in a neurobiological decision model.

9 Figures

mPFC (n=4560)				
Coefficient	Probit		Probit w/ m.e. correction	
	No Constant	Constant	No Constant	Constant
γ^{-1}	0.24	0.24	1.16	1.16
	(0.10)	(0.10)	(0.52)	(0.51)
c		-0.01		-0.06
		(0.08)		(0.37)
σ_e			4.73	4.73
			(0.37)	(0.37)
LL	-3140.46	-3140.22	-2272.22	-2272.09
BIC	6290	6297	4561	4570
Striatum (n=4180)				
Coefficient	Probit		Probit w/ m.e. correction	
	No Constant	Constant	No Constant	Constant
γ^{-1}	0.69	0.69	3.32	3.32
	(0.17)	(0.17)	(0.83)	(0.85)
c		-0.01		-0.02
		(0.08)		(0.38)
σ_e			4.67	4.67
			(0.40)	(0.40)
LL	-2841.03	-2840.98	-2063.05	-2063.04
BIC	5690	5699	4143	4151
OCC (n=4560)				
Coefficient	Probit		Probit w/ m.e. correction	
	No Constant	Constant	No Constant	Constant
γ^{-1}	0.05	0.05	0.25	0.25
	(0.08)	(0.08)	(0.36)	(0.36)
c		-0.01		-0.06
		(0.08)		(0.37)
σ_e			4.76	4.76
			(0.37)	(0.37)
LL	-3159.22	-3158.96	-2282.65	-2282.50
BIC	6327	6335	4582	4590

Table 1: NRUM estimates with and without a correction for measurement error. Clustered standard errors are in brackets.

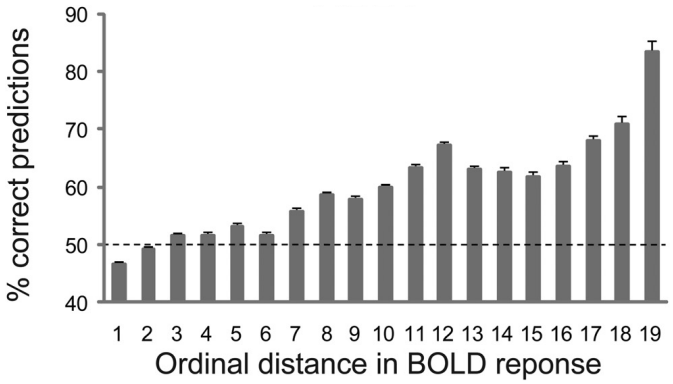

Figure 1: Choice prediction results (across subjects) from ordinal analysis of mPFC activity (Levy et al. 2011). BOLD activity for each item was ranked (within subject). Choice sets with an ordinal distance of 19 consist of the two items with the highest and lowest BOLD measurement, while choice sets with an ordinal distance of 1 consist of items that are adjacent in the ranking.

mPFC (n=2280)		
	Estimate	p-value
constant	.103	0.00
$\lvert\tilde{\tilde{B}}_{ij}\rvert$	-.031	0.01
Striatum (n=2090)		
	Estimate	p-value
constant	.102	0.00
$\lvert\tilde{\tilde{B}}_{ij}\rvert$	-.040	0.09
OCC (n=2280)		
	Estimate	p-value
constant	.087	0.00
$\lvert\tilde{\tilde{B}}_{ij}\rvert$	.013	0.50

Table 2: Estimates of random-effects GLS of stochastic choice indicator on difference in BOLD activity, as in Agranov and Ortoleva (2017). The random-effect and clustered standard errors are implemented at the subject level.

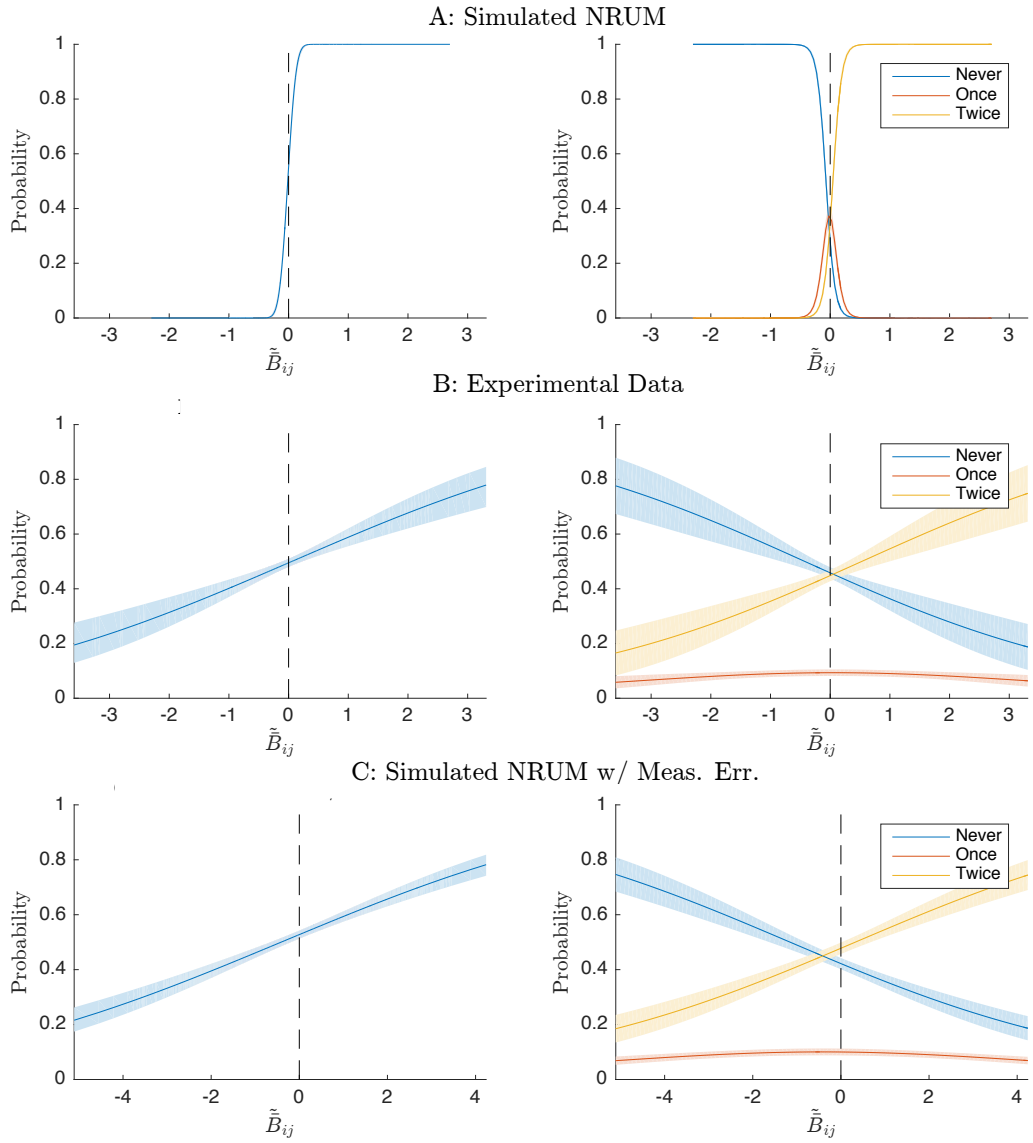

Figure 2: Analysis of (A) a simulated NRUM, (B) the mPFC activity from the experimental dataset, and (C) a simulated NRUM w/ measurement error. Left Panes: The fit of the Probit model from (18), assuming no measurement error (i.e. $\tilde{\tilde{B}}_i = \gamma E[\tilde{v}_{ij,t}]$). Right Panes: Fit of an Ordered Probit model for the probability of observing the ith item in an ij pair chosen *twice*, *once*, and *never*. The NRUM was simulated with $\gamma^{-1} = 10$, $\sigma_{\tilde{\nu}+\tilde{\eta}} = 1$, and $\sigma_e = 0$ or $\sigma_e = 5$ (A or C, respectively).

mPFC ($n = 4560$)							
Coeff	Est.	Std. Err.	P-Val	Coeff	Est.	Std. Err.	P-Val
c_1	0.03	1.14	0.98	γ_1^{-1}	-1.17	1.07	0.27
c_2	-0.15	1.25	0.91	γ_2^{-1}	0.66	2.89	0.82
c_3	-0.07	1.27	0.95	γ_3^{-1}	-3.25	2.36	0.17
c_4	-0.34	1.17	0.77	γ_4^{-1}	10.14	2.90	0.00
c_5	0.08	1.22	0.95	γ_5^{-1}	1.39	0.57	0.02
c_6	-0.07	1.22	0.95	γ_6^{-1}	-3.23	2.50	0.20
c_7	-0.14	1.30	0.91	γ_7^{-1}	2.78	3.30	0.40
c_8	0.41	1.22	0.73	γ_8^{-1}	10.39	3.53	0.00
c_9	-0.18	1.18	0.88	γ_9^{-1}	4.98	2.38	0.04
c_{10}	0.69	1.24	0.58	γ_{10}^{-1}	5.01	1.39	0.00
c_{11}	0.07	1.23	0.95	γ_{11}^{-1}	2.61	3.18	0.41
c_{12}	-0.44	1.14	0.70	γ_{12}^{-1}	13.04	3.80	0.00
σ_e	4.53	0.38					
LL = -2197.56, AIC = 4605							
Striatum ($n = 4180$)							
Coeff	Est.	Std. Err.	P-Val	Coeff	Est.	Std. Err.	P-Val
c_1	-0.05	1.21	0.97	γ_1^{-1}	1.23	2.23	0.58
c_2	-0.20	1.32	0.88	γ_2^{-1}	5.66	4.67	0.23
c_3	0.07	1.31	0.96	γ_3^{-1}	2.88	5.26	0.58
c_4	0.06	1.27	0.96	γ_4^{-1}	9.44	4.59	0.04
c_5	0.45	1.33	0.74	γ_5^{-1}	4.53	1.47	0.00
No Striatal Data for Subject 6							
c_7	-0.08	1.34	0.95	γ_7^{-1}	3.55	1.82	0.05
c_8	0.01	1.30	0.99	γ_8^{-1}	-4.60	5.44	0.40
c_9	-0.01	1.15	0.99	γ_9^{-1}	5.43	3.65	0.14
c_{10}	-0.00	1.28	1.00	γ_{10}^{-1}	3.24	1.90	0.09
c_{11}	0.03	1.27	0.98	γ_{11}^{-1}	-0.19	3.10	0.95
c_{12}	-0.12	1.16	0.92	γ_{12}^{-1}	3.52	4.02	0.38
σ_e	4.60	0.48					
LL = -2046.42, AIC = 4301							

Table 3: Subject-specific estimates from the NRUM (after correcting for measurement error).

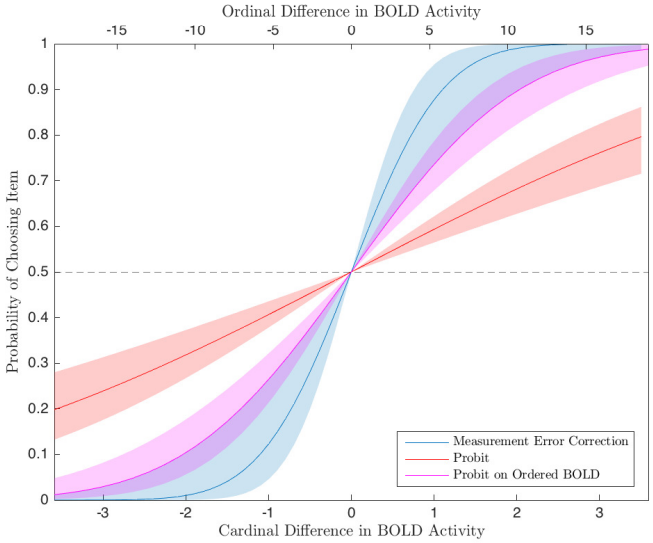

Figure 3: The probability of choosing an item depends on the difference in mPFC activity between items. The fitted probabilities are generated using a standard Probit estimate for γ^{-1}, an estimate for γ^{-1} corrected for measurement error (assuming the random-effect is zero), and a standard Probit estimate of choice on the ordinal difference in the BOLD ranking. The shaded areas depict the fitted probabilities derived from the 95% confidence intervals of the estimates.

	Prediction Rate (%)						
		Conditional on prediction			Conditional on choice outcome		
	Avg	Never	Once	Twice	Never	Once	Twice
null	27.4	46.0	9.3	44.8	25	50	25
Levy et al. (2011)	51.1	51.2	-	50.5	56.3	0	56.2
NRUM	31.6	51.8	9.3	50.8	30.8	47.0	29.7
NRUM w/ m.e. correction	42.8	56.8	9.3	56.2	45.0	28.6	43.2
NRUM + observables	59.5	72.4	11.5	64.8	64.7	20.0	62.2
known		95.2	100	95.1	100	50	100

Table 4: Comparison of choice prediction results for repeated choice trials.

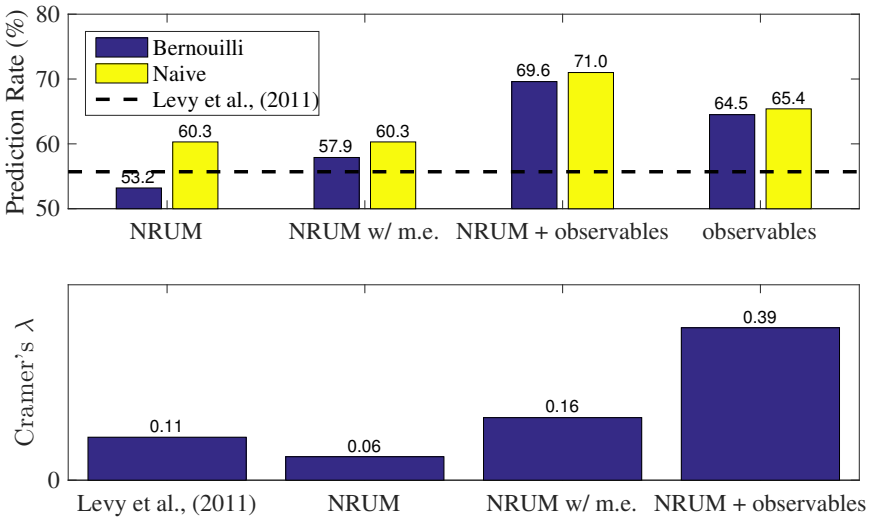

Figure 4: Prediction performance if each choice trial is treated independently.

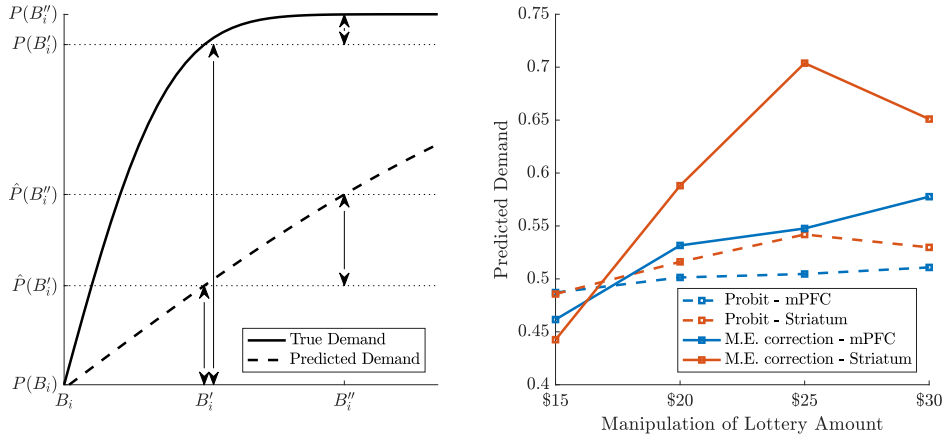

Figure 5: Effect of measurement error on product demand prediction.

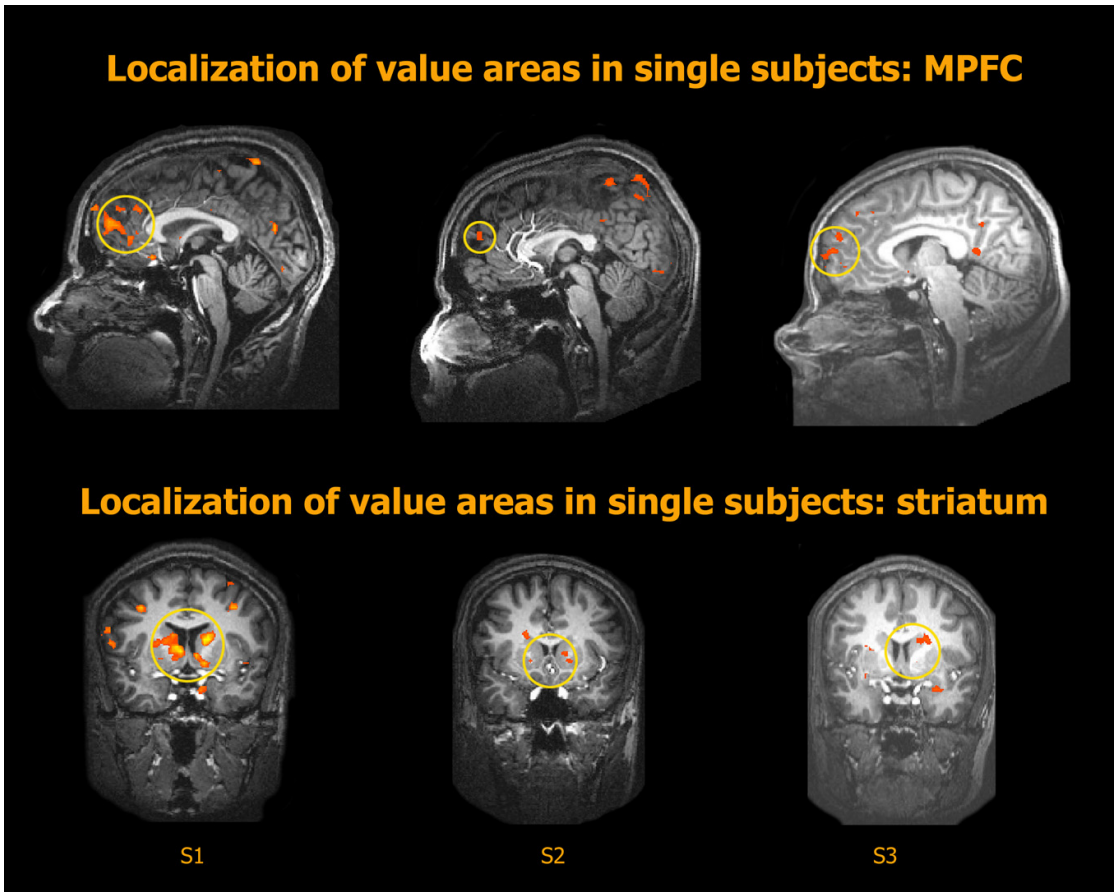

Figure 6: Region-of-interest localizations for subjects 1, 2, and 3. Activity from these regions were used to define $B_{s,i,m}$.

Subject	mPFC	Striatum
1	1985	1258
2	2019	1111
3	370	138
4	130	346
5	1953	415
6	2640	-
7	3040	168
8	1340	410
9	3272	971
10	3262	432
11	3611	604
12	600	284

Table 5: Number of voxels in each ROI.

References

Agranov, Marina and Pietro Ortoleva (2017) "Stochastic Choice and Preferences for Randomization," *Journal of Political Economy*, Vol. 125, pp. 40–68.

Apesteguia, J and M Ballester (2018) "Monotone stochastic choice models: The case of risk and time preferences," *Journal of Political Economy*.

Ariely, Dan and Gregory S Berns (2010) "Neuromarketing: the hope and hype of neuroimaging in business," *Nature Reviews Neuroscience*, Vol. 11, pp. 284–292.

Bartra, Oscar, Joseph T McGuire, and Joseph W Kable (2013) "The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value," *NeuroImage*, Vol. 76, pp. 412–427.

Basten, U, G Biele, H R Heekeren, and C J Fiebach (2010) "How the brain integrates costs and benefits during decision making," *Proceedings of the National Academy of Sciences*, Vol. 107, pp. 21767–21772.

Batley, R (2008) "On Ordinal Utility, Cardinal Utility and Random Utility," *Theory and Decision*, Vol. 64, pp. 37–63.

Beck, Jeffrey M, Wei Ji Ma, Xaq Pitkow, Peter E Latham, and Alexandre Pouget (2012) "Not Noisy, Just Wrong: The Role of Suboptimal Inference in Behavioral Variability," *Neuron*, Vol. 74, pp. 30–39.

Becker, Gordon M, Morris H DeGroot, and Jacob Marschak (1963) "Stochastic models of choice behavior," *Behavioral Science*, Vol. 8, pp. 41–55.

Berkman, E T and Emily Falk (2013) "Beyond brain mapping: Using neural measures to predict real-world outcomes," *journals.sagepub.com*, Vol. 22, pp. 45–50.

Bernheim, B Douglas (2009) "On the potential of neuroeconomics: a critical (but hopeful) appraisal," *American Economic Journal: Microeconomics*, Vol. 1, pp. 1–41.

Block, H D and Jacob Marschak (1959) "Random Orderings and Stochastic Theories of Responses," *Cowles Foundation Discussion Papers*.

Boksem, Maarten A S and Ale Smidts (2015) "Brain Responses to Movie Trailers Predict Individual Preferences for Movies and Their Population-Wide Commercial Success," *Journal of Marketing Research*, Vol. 52, pp. 482–492.

Buschena, David and David Zilberman (2000) "Generalized Expected Utility, Heteroscedastic Error, and Path Dependence in Risky Choice," *Journal of Risk and Uncertainty*, Vol. 20, pp. 67–88.

Camerer, Colin F and Teck-Hua Ho (1994) "Violations of the betweenness axiom and nonlinearity in probability," *Journal of Risk and Uncertainty*, Vol. 8, pp. 167–196.

Caplin, Andrew, Mark Dean, Paul W Glimcher, and Robb B Rutledge (2010) "Measuring beliefs and rewards: a neuroeconomic approach," *Quarterly Journal of Economics*, Vol. 125, pp. 923–960.

Carroll, Raymond J, David Ruppert, Leonard A. Stefanski, and Ciprian M. Crainiceanu (2006) *Measurement Error in Nonlinear Models: A Modern Perspective*: CRC Press, 2nd edition.

Chib, Vikram S, Antonio Rangel, Shinsuke Shimojo, and John P O'Doherty (2009) "Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex," *The Journal of Neuroscience*, Vol. 29, pp. 12315–12320.

Churchland, Anne K, R Kiani, R Chaudhuri, Xiao-Jing Wang, Alexandre Pouget, and Michael N Shadlen (2011) "Variance as a Signature of Neural Computations during Decision Making," *Neuron*, Vol. 69, pp. 818–831.

Churchland, Mark M, Byron M Yu, John P Cunningham, Leo P Sugrue, Marlene R Cohen, Greg S Corrado, William T Newsome, Andrew M Clark, Paymon Hosseini, Benjamin B Scott, David C Bradley, Matthew A Smith, Adam Kohn, J Anthony Movshon, Katherine M Armstrong, Tirin Moore, Steve W Chang, Lawrence H Snyder, Stephen G Lisberger, Nicholas J Priebe, Ian M Finn, David Ferster, Stephen I Ryu, Gopal Santhanam, Maneesh Sahani, and Krishna V Shenoy (2010) "Stimulus onset quenches neural variability: a widespread cortical phenomenon,"

Nature Neuroscience, Vol. 13, pp. 369–378.

Clithero, John A and Antonio Rangel (2013) "Informatic parcellation of the network involved in the computation of subjective value," *Social cognitive and affective neuroscience*, Vol. 9, pp. 1–14.

Cramer, J S (1999) "Predictive Performance of the Binary Logit Model in Unbalanced Samples," *Journal of the Royal Statistical Society. Series D The Statistician*, Vol. 48, pp. 85–94.

Domenech, Philippe, Jérôme Redouté, Etienne Koechlin, and Jean-Claude Dreher (2017) "The Neuro-Computational Architecture of Value-Based Selection in the Human Brain.," *Cerebral Cortex*, pp. 1–17.

Drugowitsch, J, V Wyart, A D Devauchelle, and E Koechlin (2016) "Computational precision of mental inference as critical source of human choice suboptimality," *Neuron*.

Falk, Emily B, Elliot T Berkman, and Matthew D Lieberman (2012) "From Neural Responses to Population Behavior: Neural Focus Group Predicts Population-Level Media Effects," *Psychological Science*, Vol. 23, pp. 439–445.

Falmagne, J C (1978) "A representation theorem for finite random scale systems," *Journal of Mathematical Psychology*, Vol. 18, pp. 52–72.

Falmagne, Jean-Claude (1985) *Elements of Psychophysical Theory*: Oxford University Press, USA.

Fechner, G (1860) *Elements of Psychophysics*, Vol. 1: Holt, Rhinehart and Winston, New York.

Fehr, Ernst and Antonio Rangel (2011) "Neuroeconomic Foundations of Economic Choice—Recent Advances," *Journal of Economic Perspectives*, Vol. 25, pp. 3–30.

FitzGerald, Thomas H B, Ben Seymour, and Raymond J Dolan (2009) "The role of human orbitofrontal cortex in value comparison for incommensurable objects," *The Journal of Neuroscience*, Vol. 29, pp. 8388–8395.

Genevsky, Alexander and Brian Knutson (2015) "Neural Affective Mechanisms

Predict Market-Level Microlending," *Psychological Science*, Vol. 26, pp. 0956797615588467–1422.

Genevsky, Alexander, Carolyn Yoon, and Brian Knutson (2017) "When Brain Beats Behavior: Neuroforecasting Crowdfunding Outcomes," *The Journal of Neuroscience*, Vol. 37, pp. 8625–8634.

Glimcher, Paul W (2005) "Indeterminacy in Brain and Behavior," *Annual Review of Psychology*, Vol. 56, pp. 25–56.

——— (2011) *Foundations of Neuroeconomic Analysis*: Oxford University Press.

Glimcher, Paul W and Ernst Fehr (2013) *Neuroeconomics*, Decision Making and the Brain: Academic Press, 2nd edition.

Gold, Joshua I and Michael N Shadlen (2007) "The neural basis of decision making," *Annual Review of Neuroscience*, Vol. 30, pp. 535–574.

Greene, W H (2003) *Econometric Analysis* : Prentice Hall, 5th edition.

Gul, Faruk and Wolfgang Pesendorfer (2006) "Random Expected Utility," *Econometrica: Journal of the Econometric Society*, Vol. 74, pp. 121–146.

Hare, Todd A, Wolfram Schultz, Colin F Camerer, John P O'Doherty, and Antonio Rangel (2011) "Transformation of stimulus value signals into motor commands during simple choice," *Proceedings of the National Academy of Sciences*, Vol. 108, pp. 18120–18125.

Harrison, Glenn W (2008) "Neuroeconomics: A critical reconsideration," *Economics and Philosophy*, Vol. 24, pp. 303–344.

Harrison, Glenn W and E Elisabet Rutstrom (2008) "Risk Aversion in the Laboratory," in James C Cox and Glenn W Harrison eds. *Risk Aversion in Experiments*, Bingley, UK: Emerald Group Publishing Ltd. pp. 41–155.

Hausman, Jerry A and David A Wise (1978) "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences," *Econometrica: Journal of the Econometric Society*, Vol. 46, pp. 403–426.

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

——— (2005) "Why we should not be silent about noise," *Experimental Economics*, Vol. 8, pp. 325–345.

Hey, John D. and Chris Orme (1994) "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica: Journal of the Econometric Society*, Vol. 62, pp. 1291–1326.

Johnson, Eric J and Roger Ratcliff (2013) "Computational and Process Models of Decision Making in Psychology and Behavioral Economics," in Paul W Glimcher and Ernst Fehr eds. *Neuroeconomics*: Academic Press.

Kahn, Itamar, Mitul Desai, Ulf Knoblich, Jacob Bernstein, Michael Henninger, Ann M Graybiel, Edward S Boyden, Randy L Buckner, and Christopher I Moore (2011) "Characterization of the functional MRI response temporal linearity via optical control of neocortical pyramidal neurons," *The Journal of Neuroscience*, Vol. 31, pp. 15086–15091.

Kahnt, Thorsten (2017) "A decade of decoding reward-related fMRI signals and where we go from here," *NeuroImage*.

Knutson, Brian, Scott Rick, G Elliott Wimmer, Drazen Prelec, and George Loewenstein (2007) "Neural Predictors of Purchases," *Neuron*, Vol. 53, pp. 147–156.

Krajbich, Ian, Carrie Armel, and Antonio Rangel (2010) "Visual fixations and the computation and comparison of value in simple choice," *Nature Neuroscience*, Vol. 13, pp. 1292–1298.

Krajbich, Ian, Colin F Camerer, John Ledyard, and Antonio Rangel (2009) "Using neural measures of economic value to solve the public goods free-rider problem.," *Science*, Vol. 326, pp. 596–599.

Lebreton, Maël, Soledad Jorge, Vincent Michel, Bertrand Thirion, and Mathias Pessiglione (2009) "An automatic valuation system in the human brain: evidence from functional neuroimaging.," *Neuron*, Vol. 64, pp. 431–439.

Levy, Dino and Paul W Glimcher (2011) "Comparing Apples and Oranges: Using Reward-Specific and Reward-General Subjective Value Representation in the Brain," *The Journal of Neuroscience*, Vol. 31, pp. 14693–14707.

Levy, Dino J and Paul W Glimcher (2012) "The root of all value: a neural common currency for choice," *Current Opinion in Neurobiology*, Vol. 22, pp. 1027–1038.

Levy, Ifat, Stephanie Lazzaro, Robb B Rutledge, and Paul W Glimcher (2011) "Choice from non-choice: predicting consumer preferences from blood oxygenation level-dependent signals obtained during passive viewing," *The Journal of Neuroscience*, Vol. 31, pp. 118–125.

Logothetis, N K, H Guggenberger, S Peled, and J Pauls (1999) "Functional imaging of the monkey brain," *Nature Neuroscience*, Vol. 2, pp. 555–562.

Logothetis, N K, J Pauls, M Augath, T Trinath, and A Oeltermann (2001) "Neurophysiological investigation of the basis of the fMRI signal," *Nature*, Vol. 412, pp. 150–157.

Logothetis, Nikos K (2002) "The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal.," *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, Vol. 357, pp. 1003–1037.

——— (2003) "The underpinnings of the BOLD functional magnetic Resonance Imaging Signal," *The Journal of Neuroscience*, Vol. 23, pp. 3963–3971.

Loomes, Graham (2005) "Modelling the stochastic component of behaviour in experiments: Some issues for the interpretation of data," *Experimental Economics*, Vol. 8, pp. 301–323.

Loomes, Graham and Robert Sugden (1995) "Incorporating a stochastic element into decision theories," *European Economic Review*, pp. 641–648.

Louie, Kenway, Lauren E Grattan, and Paul W Glimcher (2011) "Reward value-based gain control: divisive normalization in parietal cortex," *The Journal of Neuroscience*, Vol. 31, pp. 10627–10639.

Luce, R. Duncan (1959) *Indvidual Choice Behaviour*: John Wiley and Sons, Inc.

Lund, Torben E, Kristoffer H Madsen, Karam Sidaros, Wen-Lin Luo, and Thomas E Nichols (2006) "Non-white noise in fMRI: Does modelling have an impact?," *NeuroImage*, Vol. 29, pp. 54–66.

Manski, Charles F (1977) "The structure of random utility models," *Theory and Decision*, Vol. 8, pp. 229–254.

McFadden, Daniel L (1973) "Conditional Logit Analysis of Qualitative Choice Behavior," in Paul Zarembka ed. *Frontiers in Econometrics*: Academic Press, Inc. pp. 105–142.

——— (1981) "Structural Discrete Probability Models Derived from Theories of Choice," in Charles F Manski and Daniel L McFadden eds. *Structural Analysis of Discrete Data and Econometric Applications*: MIT Press, pp. 198–272.

——— (2001) "Economic choices," *The American Economic Review*, Vol. 91, pp. 351–378.

——— (2005) "Revealed stochastic preference: a synthesis," *Economic Theory*, Vol. 26, pp. 245–264.

Michelson, Charles, Jonathan Pillow, and Eyal Seidemann (2013) "Perceptual decisions are limited primarily by variability in early sensory cortex," in *Computational and Systems Neuroscience (Cosyne) Meetings 2013*, Salt Lake City, UT.

Milosavljevic, M, J Malmaud, A Huth, Christof Koch, and Antonio Rangel (2010) "The drift diffusion model can account for value-based choice response times under high and low time pressure," *Judgement and Decision Making*, Vol. 5, pp. 437–449.

Mosteller, Frederick and Philip Nogee (1951) "An Experimental Measurement of Utility," *Journal of Political Economy*, Vol. 59, pp. 371–404.

Neuhaus, John M, Charles E McCulloch, and Ross Boylan (2011) "A note on type II error under random effects misspecification in generalized linear mixed models," *Biometrics*, Vol. 67, pp. 654–6; disucssion 656–60.

Padoa-Schioppa, Camillo (2013) "Neuronal Origins of Choice Variability in Economic Decisions," *Neuron*, Vol. 80, pp. 1322–1336.

Ramsey, J D, S J Hanson, C Hanson, YO Halchenko, Russell A Poldrack, and C Glymour (2010) "Six problems for causal inference from fMRI," *NeuroImage*, Vol. 49.

Ratcliff, Roger (1978) "A theory of memory retrieval," *Psychological Review*, Vol. 85, pp. 59–108.

Rich, Erin L and Jonathan D Wallis (2016) "Decoding subjective decisions from orbitofrontal cortex," *Nature Neuroscience*, Vol. 19, pp. 973–980.

Rieke, Fred, David Warlan, Rob de Ruyter van Steveninck, and William Bialek (1997) *Spikes*: MIT Press.

Shadlen, Michael N and William T Newsome (1998) "The variable discharge of cortical neurons: implications for connectivity, computation, and information coding," *The Journal of Neuroscience*, Vol. 18, pp. 3870–3896.

Smith, Alec, B Douglas Bernheim, Colin F Camerer, and Antonio Rangel (2014) "Neural Activity Reveals Preferences without Choices," *American Economic Journal: Microeconomics*, Vol. 6, pp. 1–36.

Stevens, S S (1961) "To Honor Fechner and Repeal His Law: A power function, not a log function, describes the operating characteristic of a sensory system," *Science*, Vol. 133, pp. 80–86.

Telpaz, Ariel, Ryan Webb, and Dino J Levy (2015) "Using EEG to Predict Consumers' Future Choices," *Journal of Marketing Research*, Vol. 52, pp. 511–529.

Tolhurst, D J, J Anthony Movshon, and A F Dean (1983) "The statistical reliability of signals in single neurons in cat and monkey visual cortex," *Vision research*, Vol. 23, pp. 775–785.

Train, Kenneth E (2009) *Discrete Choice Methods with Simulation*: Cambridge University Press, 2nd edition.

Tusche, Anita, Stefan Bode, and John-Dylan Haynes (2010) "Neural Responses to Unattended Products Predict Later Consumer Choices," *The Journal of Neuroscience*, Vol. 30, pp. 8024–8031.

Venkatraman, V, A Dimoka, P A Pavlou, Khoi Vo, William Hampton, Bryan Bollinger, Hal E Hershfield, Masakazu Ishihara, and Russel S Winer (2015) "Predicting advertising success beyond traditional measures: New insights from neurophysiological methods and market response modeling," *Journal of Marketing Research*, Vol. 52, pp. 436–452.

Vul, Edward, C Harris, Piotr Winkielman, and H Pashler (2009) "Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognitionn," *Perspectives on Psychological Science*, Vol. 4, pp. 274–290.

Wang, N, X Lin, RG Gutierrez, and Raymond J Carroll (1998) "Bias Analysis and SIMEX Approach in Generalized Linear Mixed Measurement Error Models," *Journal of the American Statistical Association*, Vol. 93, pp. 249–261.

Webb, Ryan (2018) "The (Neural) Dynamics of Stochastic Choice," *Management Science*.

Webb, Ryan, Paul W Glimcher, and Kenway Louie (2016) "Rationalizing Context-Dependent Preferences: Divisive Normalization and Neurobiological Constraints on Choice," *SSRN Working Paper*.

Weber, Ernst (1834) *On the Tactile Senses (with translation of De Tactu)*: Experimental Psychology Society, New York.

Wilcox, Nathaniel T (2008) "Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison," in Glenn W Harrison and James C Cox eds. *Risk Aversion in Experiments*: Emerald Group Publishing Ltd. pp. 197–292.

Woodford, M (2014) "Stochastic choice: An optimizing neuroeconomic model," *The American Economic Review*, Vol. 104, pp. 495–500.

Yatchew, A. and Z. Griliches (1985) "Specification Error in Probit Models," *Review of Economics and Statistics*, Vol. 67, pp. 134–139.

10 Appendix

10.1 The Levy et al. (2011) Experiment

The laboratory experiment was divided into three stages. The first two stages were performed inside an MRI scanner. In the first stage, subjects passively viewed the outcome of a series of small lotteries over changes to their wealth. The purpose of this stage was to identify the areas of the brain which encoded the subject's subjective values, $v_{i,t}$. In the second stage, subjects passively viewed 20 consumer items while intermittently performing an incentivized task so as to maintain subject engagement. The purpose of this stage was to repeatedly measure the subjective values of these items. Immediately after the second stage, subjects performed a third stage outside of the scanner in which they made all possible binary choices over this set of items in an incentive compatible fashion. Before leaving the subject also received a $25 show-up fee in cash.

10.1.1 Localization of Subjective Value in Medial Prefrontal Cortex

The first stage of the experiment was designed to identify an area in the brain of each subject which encodes subjective value. For brain measurements, we employed functional MRI (fMRI) using standard techniques (as in Caplin et al. 2010, Levy et al. 2011). These techniques indirectly measure brain activity over a 2 second interval in each of about 250,000 $3mm \times 3mm \times 3mm$ cubes (voxels) tiling the human brain. The product of this process is thus a time-series, in 2 second increments, of activation levels in each voxel.

The measure of activation is derived from the paramagnetic properties of the hemoglobin molecule and is known as the Blood-Oxygenation Level Dependent (BOLD) signal. This measurement has been demonstrated to be strictly monotonic in the average of the neural activity within the voxel, and most studies indicate that BOLD approximates a linear transformation of neural activity (Logothetis et al. 1999, 2001, Kahn et al. 2011).

A statistical challenge arises from the sheer number of time-series fMRI generates imposed by determining which voxels/timeseries to study (Vul et al. 2009). This study restricted analysis to regions of the brain known to encode subjective value-like signals, the medial prefrontal cortex (MPFC) and Striatum.[32] An initial experiment aimed at independently 'localizing' subjective value encoding voxels within the mPFC and Striatum, with the intention of conducting the analysis of the main experiment upon a time-series derived by averaging over these localized voxels.

In this initial stage of the experiment each subject was endowed with $40. On ensuing trials a lottery with equal probability of gaining or losing $2 was presented visually to the subject in the scanner. The outcome of the lottery was then revealed to the subject and the result was added to or deducted from the subject's wealth. In total, 128 trials of this kind were presented.[33] For each mPFC voxel, the difference in average activity between winning and losing was calculated. For each subject, voxels which showed a statistically significant difference were identified as our region of interest for encoding subjective valuation.

10.1.2 Recording the Subjective Value of Items

Immediately following the first stage, subjects completed a second stage in the scanner intended to measure the subjective values of 20 consumer items. Subjects completed six 7-minute brain scans over the course of 45 minutes, each consisting of 40 trials, for a total of 240 trials. In each of these trials, subjects passively viewed an image of one of 20 different items, including four DVD movies, two books, four art posters, three music CDs, two pieces of stationery, and five monetary lotteries represented by pie charts. Each lottery offered a 50% chance of receiving a designated amount of money ($10, $15, $20, $25, $30) and a 50% chance of receiving $0. All items were presented 12 times in a random order to each subject. Subjects were instructed that when they saw an item they should think about how much it was worth to them in a dollar amount.

[32]For reviews relating mPFC and Striatal activity to subjective value see (Levy and Glimcher 2012, Bartra et al. 2013, Clithero and Rangel 2013)

[33]This task is a non-choice version of the task previously developed in Caplin et al. (2010).

To keep subjects alert, on 20 randomly selected trials (one for each of the 20 items), subjects were asked whether they preferred the item they had just seen or a randomly selected amount of money (ranging from \$1 to \$10). Subjects were told that one of these question trials would be randomly realized at the end and they would receive their selection on that trial - the item or the money. These 20 question trials were excluded from all behavioural and neural analysis. During the scanning stage, subjects did not know they would subsequently be offered an opportunity to choose between these same items after the scanning process was complete.

10.1.3 Choice Task

Following the second scanning stage, subjects were asked to perform a choice task outside of the scanner. Subjects were presented with a complete series of binary choices between the 20 items previously presented in the scanner. Each possible binary comparison (190 choices) was presented twice (switching the left-right location on each repetition), in random order, for a total of 380 choices. The result of one of these choices was randomly selected for realization.

The choices of subjects were largely consistent, with $96 \pm 2\%$ of triplets transitive and subjects switching their selection in only $9 \pm 1\%$ of choice repetitions. Choices were also highly idiosyncratic across subjects such that the individual preferences of a given subject could not be predicted from preferences exhibited by other subjects (mean correlation of ranking between pairs of subjects, excluding lotteries: $r = 0.1 \pm 0.3$).[34]

10.2 Comparison with Standard Latent Variable Modelling

The NRUM decomposes the uncertainty present in the standard RUM into biophysically distinct sources, yielding the observable variable v on which to base choice prediction. This allows us to investigate, as a benchmark for our measurement, the potential

[34]We also verified that the random amounts of money used in the question trials in the scanner did not bias subjects' choices outside of the scanner.

benefit of using neural data to predict choices compared to a dataset of only standard economic observables. In particular, we focus on specification error in the standard approach due to the modeller's inability to observe all the attributes (of alternatives and decision makers) that make up utility (Manski 1977).

To cement ideas, suppose on a given trial the econometrician only observes a partition, $X_{i,t} \in \mathbb{R}^k$, of the full vector of attributes, $Z_{i,t} \in \mathbb{R}^l$, which make up subjective value (or utility) for item i (i.e. $k < l$). In the standard formulation of the RUM, this partitioning matters since the econometrician does not observe the utility of item i, instead the latent variable $u_{i,t}$ must be indirectly specified. The components of subjective value that are observed, $X_{i,t}$, are related to this latent variable as a linear combination, $X_{i,t}\beta$, while the components of $u_{i,t}$ that are unobserved are bundled in to an error term $\varepsilon_{i,t}$.

Given our NRUM, we can decompose $\varepsilon_{i,t}$ into three sources. For the sake of this argument, we follow the standard approach and assume that subjective value is related to the arguments Z or X through the linear function $V(X_{i,t};\beta) = X_{i,t}\beta + \nu_{i,t}$.[35] The difference between the full specification $V(Z_{i,t};\beta)$ and the partitioned specification $V(X_{i,t};\beta)$, which we will refer to as specification error, is denoted $\omega_{i,t}$. Together with the stochasticity in subjective value and the choice mechanism, this yields a decision variable in which $\varepsilon_{i,t} \equiv \nu_{i,t} + \omega_{i,t} + \eta_{i,t}$ bundles together the three sources of uncertainty in our NRUM as follows:

$$
\begin{aligned}
v_{i,t} &= V(Z_{i,t}, \beta) \\
v_{i,t} &= V(X_{i,t}, \beta) + \omega_{i,t} \\
v_{i,t} + \eta_{i,t} &= X_{i,t}\beta + \nu_{i,t} + \omega_{i,t} + \eta_{i,t} \\
u_{i,t} &= X_{i,t}\beta + \nu_{i,t} + \omega_{i,t} + \eta_{i,t}.
\end{aligned}
$$

[35] In practice, this function must be non-linear because the neural activity which encodes v is bounded above and below. Additionally, there is evidence that $V()$ takes the entire vector X as its argument, yielding subjective values which depend on the composition of the choice set (Louie et al. 2011, Webb et al. 2016). Both of these issues result in misspecification error if unaccounted for. While the first issue can be easily dealt with in a standard RUM, the second requires careful attention (Webb et al. 2016). Regardless, both of these issues disappear if v is observed directly.

As before, we can derive choice probabilities after imposing normality assumptions to arrive at the familiar textbook specification of the Probit model,

$$P(y_{ij,t} = 1 \mid X_{ij,t}) = P\left(X_{i,t}\beta + \tilde{\omega}_{ij,t} > \tilde{\nu}_{ji,t} + \tilde{\eta}_{ji,t}\right) \tag{22}$$

$$= P\left(X_{ij,t}\beta > \tilde{\varepsilon}_{ij,t}\right)$$

$$= \Phi\left(\frac{X_{ij,t}\beta}{\sigma_{\tilde{\varepsilon}}}\right), \tag{23}$$

where the variable $\tilde{\varepsilon}_{ij,t}$ aggregates all of the differenced error terms and $\sigma^2_{\tilde{\varepsilon}} = \sigma^2_{\tilde{\omega}} + \sigma^2_{\tilde{\nu}+\tilde{\eta}}$.

An obvious implication is that the latent variable model with non-zero specification error (23) will have the worst predictive power relative to the two neural specifications (11) and (13) since $\sigma^2_{\tilde{\eta}} \leq \sigma^2_{\tilde{\nu}+\tilde{\eta}} < \sigma^2_{\tilde{\varepsilon}}$. The latent variable formulation introduces error into the specification due to an inability of the modeller to fully explain subjective value with observables in the dataset (Manski 1977). Observing a neural measure of subjective value removes this source of error, provided we can obtain a suitable neural measurement.