QUANTUM PROBABILITY: A NEW METHOD FOR MODELLING TRAVEL CHOICES

Thomas O. Hancock (Corresponding Author)
Institute for Transport Studies
University of Leeds
tstoh@leeds.ac.uk

Stephane Hess
Institute for Transport Studies
University of Leeds
S.Hess@its.leeds.ac.uk

Charisma F. Choudhury
Institute for Transport Studies
University of Leeds
C.F.Choudhury@leeds.ac.uk

ABSTRACT

There has been an increasing effort to increase the behavioural realism of the mathematical models of choice, resulting in efforts to move away from random utility maximisation (RUM) models. Some new insights have been generated with, for example, models based random regret minimisation (RRM). However, many of the alternatives to RUM tested on real-world data, have looked at only modest departures from RUM, and differences in results have consequently been small. In this paper, we address this research gap by investigating the applicability of models based on quantum theory - which are substantially different from the state-of-the-art choice modelling techniques. These models emphasise the importance of contextual effects, state dependence and the impact of choice or question order. As a result, quantum probability models have had some success in better explaining several phenomena in cognitive psychology. In this paper, we consider how to best operationalise quantum probability into a choice model. Two of our specifications find good model fit across three route choice datasets. Additionally, we test the quantum model frameworks on a best/worst route choice dataset and demonstrate that they find useful transformations to capture differences between the attributes important in a favourite alternative compared to that of the least favourite alternative. Similar transformations can also be used to efficiently capture contextual effects in a dataset where the order of the attributes and alternatives are manipulated. Overall, it appears that models incorporating quantum concepts hold significant promise in improving the state-of-the-art travel choice modelling paradigm through their adaptability and efficient modelling of contextual changes.

1. INTRODUCTION

Random utility maximisation (RUM) framework has dominated the travel choice modelling field for many decades. More recently, RUM has been criticised as being inadequate in explaining the full range of behavioural complexity [Chorus et al., 2008, Guevara and Fukushi, 2016]. This has resulted in many attempts to better incorporate behavioural concepts into travel behaviour models, including regret [Chorus et al., 2008, Chorus, 2010], contextual relative advantages [Leong and Hensher, 2014] and prospect theory [Avineri and Bovy, 2008]. However, none of these developments have yet rivalled RUM as the preferred model in real-world applications. This is due to difficulties that quickly arise once a modeller departs from the firm economic foundations of RUM [Hess et al., 2018]. Consequently, caution is required if we are to step away from random utility models. Departures to models with similar underlying structure, such as random regret minimisation [Chorus et al., 2008, Chorus, 2010], which have the same error structure, result in only small differences whilst facing the same key fault of all departures from RUM, the loss of the ability to calculate welfare measures. Departures to more different models, such as decision field theory [Busemeyer and Townsend, 1992], whilst sometimes finding improvements in model fit, additionally result in models that become computationally infeasible for large-scale datasets [Hancock et al., 2018]. Thus, if we are to move away from RUM, we need to investigate alternative approaches that are computationally simpler - yet, better reflect behavioural realism. This leads us to explore ideas from other disciplines which are further away from the tried and tested. Given the success of using ideas from quantum physics in cognitive psychology, one possible alternative is to see if quantum physics can make a similar step into travel behaviour modelling.

Quantum physics, first considered in the early 20th century, was originally created to explain phenomena and results that could not be explained by classical theories of probability and physics. In particular, physicists noticed that the measurement of one variable could impact the measurement of another. The most famous example of this relates to measuring the position and momentum (mass multiplied by velocity) of a particle. Physicists found that they could not measure both accurately at the same time. This led to 'Heisenberg's uncertainty principle' [Heisenberg, 1927]. Formally, this could be written [Kennard, 1927] as:

$$\sigma_x \cdot \sigma_\rho \geq \frac{\hbar}{2}, \tag{1}$$

where σ_x and σ_ρ are the standard deviations of position and momentum respectively, which multiplied together give the uncertainty, and $\hbar$ is the reduced Planck constant, $\frac{h}{2\pi}$ (with h the Planck constant, a physical constant first used by Planck [1901] that relates the energy carried by a photon to its frequency). To illustrate how this led to the breakdown of classical probability, we first imagine that we have three possible propositional variables:

A : 'the particle has momentum in the interval $[\rho_1,\rho_2]$'

B : 'the particle is in the interval $[x_1,x_2]$'

C : 'the particle is in the interval $[x_2,x_3]$.'

If we have a system where $\hbar = 1$, the minimum allowed uncertainty by Heisenberg's uncertainty principle (Equation 1) is $\frac{1}{2}$. Under this system, we might observe a particle such that the propositional variables (A) and (B or C) are true (where total uncertainty is greater than $\frac{1}{2}$). Simultaneously,

it is possible that both (A and B) and (A and C) each have total uncertainties of less than $\frac{1}{2}$ and are thus false. Consequently, the distributivity law of classical probability ($A(B+C) = AB+AC$) fails to hold.

This resulted in the creation of a new theory of probability, known as quantum logic [Birkhoff and Von Neumann, 1936]. Under quantum logic (which is also known as quantum probability), a new set of probability rules were defined, which crucially did not include the axiom of distributivity. This new theory of probability has subsequently made the transition into cognitive psychology [Bruza et al., 2015] and has also been introduced into transport behaviour modelling. For example, Vitetta [2016] introduced a quantum model based on random utility models with the addition of an interference term for route choice problems. Additionally, Yu and Jayakrishnan [2018] demonstrated that quantum cognition models can efficiently be used to capture the difference in state of mind between choices made under stated preference and revealed preference settings. However, thus far, as far as the authors are aware, there has not been a choice model developed with quantum concepts that incorporates attribute values for individual alternatives and can work for general choices as well as 'changes in perspective'. Thus the focus of this paper is to explore ways to develop a choice modelling framework based on quantum logic that can be used for choices in general, as well as efficiently capturing effects caused by ordering, some form of interference or some change in 'state of mind'.

The rest of this paper is organised as follows. First, we introduce quantum logic and discuss the relative benefits of using such a system. We then mathematically describe quantum logic, giving detailed graphical examples. Next, we discuss how it can be incorporated into a choice model, detailing two different formulations for new models. We then test the performance of our proposed models against typical choice models such as multinomial logit and also random regret minimisation in the context of travel decisions. Finally, we test the models on best-worst and contextual choice data, before drawing some conclusions.

2. QUANTUM LOGIC

In this section, we first give a general overview of quantum logic. We then give the mathematical definitions for how quantum logic works for basic choices. We conclude by describing how it works for a series of related choices. It is in the transformation from one choice task to another that a modelling framework based on quantum logic looks very different to traditional choice models.

2.1. Overview of quantum logic

A simple example of how quantum logic works is given in Figure 1. Initially, a decision-maker might be making a single choice between two alternatives, travelling by car or by train. Each of these alternatives are represented by vectors, $|T\rangle$ and $|C\rangle$ respectively (the axes in the Figure 1). Under quantum logic, the decision-maker has some initial state, denoted $|z\rangle$, regarding whether they will choose car or train.

The action of making a choice (or equivalently making a judgement or coming to some result) results in a 'change of state'. This can be represented graphically by moving from the initial state vector and 'projecting' onto the vector corresponding to the chosen alternative. In this example, ρT, represents the scalar projection of $|z\rangle$ onto a straight line parallel to $|T\rangle$. The length of this projection is then denoted $|\rho T|$. In Figure 1, these projections are directly over the corresponding vectors. The arrows from the initial state to these vectors are not the projections themselves, but help to determine the length of the projections, which are the distances from the origin to the points

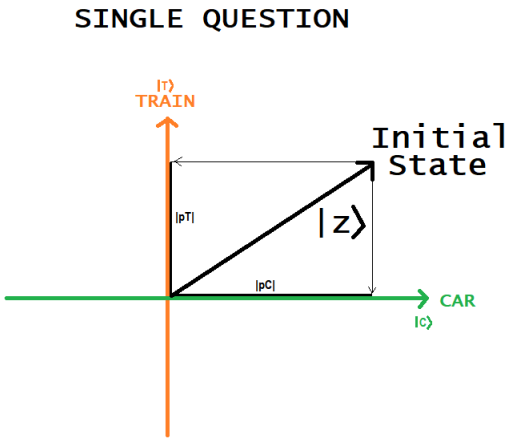

FIGURE 1 : A single question under quantum probability

where the arrows meet the vector. When the state vector is at 45 degrees, the projections are of equal length and the probabilities are thus 50% each. In the example in Figure 1, the car alternative has a higher probability. The full mathematical description for this is given in the following section on a basic choice under quantum logic, which also gives a 3-dimensional example. The longer the projection onto the vector for an alternative, the more likely that alternative is chosen. The crucial difference in using such a system is how an additional question or nudge can then impact the decision-maker's choice for the first question (car or train). If, for example, the decision-maker was asked 'are you environmentally friendly?' before they had made up their mind between the choice of car or train, they would then be initially answering a different question and making a different choice (see Figure 2).

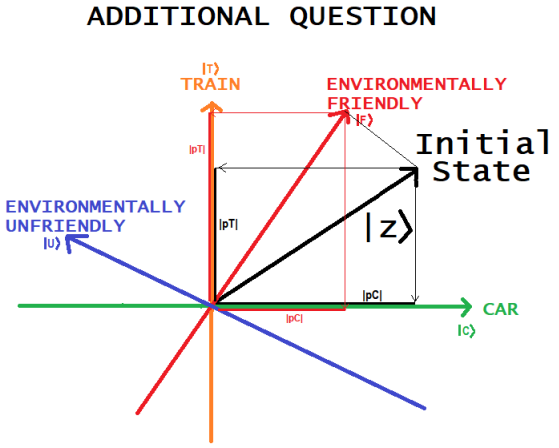

FIGURE 2 : Making two choices under quantum logic

As a result of the decision-maker deciding 'I am environmentally friendly', the decision-maker's state moves from the initial starting state and is projected onto the vector representing

'environmentally friendly' and vice versa if they decide 'I am not environmentally friendly' (see Figure 2). This results in making the choice between car and train from a different state. Consequently the length of the projections ($|\rho C|$ and $|\rho T|$) onto the vectors for car and train have changed. This is graphically represented in Figure 2, with the projection length for $|\rho T|$ longer if the initial state is first projected onto the environmentally friendly vector before being projected onto the train vector, relative to the projection length if train is chosen directly from the initial state. Consequently, the probabilities for choosing car and train are altered.

Cognitive psychologists have many key reasons for using quantum logic [Busemeyer et al., 2011] that are also relevant for transport behaviour modelling. Firstly, a behavioural state is initially 'indefinite' and is often created rather than just recorded by an attempt to measure it. For example, a decision-maker might only start considering how environmentally friendly they are after they have been asked (or reminded) about how environmentally friendly they are. For this reason, it is essential that surveys including both choice tasks and attitudinal questions require the respondent to complete the choice tasks first, if the researcher wishes to avoid bias in the choice task [Ben-Akiva et al., 2019]. However, conversely, a decision-maker may try to 'justify' their choices with their responses to the attitudinal questions [Cunha-e Sá et al., 2012]. Consequently, it is often difficult to measure a decision-maker's *true* attitudes, opinions and preference without some form of bias. It is easy to see how this relates to issues for choice modellers with, for example, analysts often having concerns about the biases or truthfulness within stated preference data [Mahieu et al., 2016]. Secondly, psychologists have put forward the argument that cognition behaves like a wave rather than a particle [Trueblood and Busemeyer, 2012]. A decision-maker might consider the advantages of getting the train but then also consider the advantages of driving. Indeed, many models developed in mathematical psychology assume preferences for alternatives that update stochastically [Busemeyer and Townsend, 1993, Krajbich et al., 2012]. Under quantum logic, their preference over time 'behaves like a wave' and consequently fluctuates over time. It is only when a decision-maker makes up their mind that their preference exists as some measurable definite state. Before an action or choice is made, an observer does not know what the decision-maker will do. There are many preference states within travel behaviour that could similarly be described as 'wave-like', such as anticipating merging onto a new lane when driving, changing travel mode when weather worsens, or choosing which route to take depending on traffic conditions. One of the most crucial quantum concepts, however, is the idea of interferences or nudges (such as the previous example of being asked about the environment whilst in the process of making a mode choice). After the development of quantum physics to explain ordering effects of observed variables [Birkhoff and Von Neumann, 1936], a wide range of quantum models, often based on the idea of quantum interference, have been put forward in cognitive psychology [Bruza et al., 2015]. These include a quantum model to explain ordering effects [Trueblood and Busemeyer, 2011], a quantum similarity model [Pothos et al., 2013] and a quantum judgement model [Busemeyer et al., 2011]. These models perform a similar function to choice models that include state dependence, where a number of different models [Seetharaman, 2003] have been applied to capture the temporal correlation of choices over time. Given the success of quantum models at explaining ordering effects within cognitive psychology, there is ample scope for quantum logic and quantum ideas within travel behaviour modelling and choice modelling in general.

2.2. A basic choice under quantum logic

Under quantum logic, a measurement (or choice scenario), X, can be represented geometrically as a subspace L_x in a multidimensional Hilbert space [Trueblood et al., 2014b]. For each measurement, a number of discrete 'events' are possible. These events, if mutually exclusive, are represented by orthonormal vectors, which are denoted $|x_1\rangle$, $|x_2\rangle$, ... $|x_J\rangle$ (with J the number of alternatives). For these vectors, we use bra-ket notation in keeping with the standard notation used in quantum mechanics and quantum cognition (c.f. Trueblood and Busemeyer 2011). Under bra-ket notation, a column vector in a Hilbert space is represented by a 'ket' vector, $|\cdot\rangle$, with the corresponding row vector (with each element being complex conjugated) a 'bra' vector, $\langle\cdot|$ [Yu and Jayakrishnan, 2018]. These orthonormal vectors then form a basis for the subspace L_x. Consequently, the Hilbert space for a choice task with J alternatives can be represented by a J-dimensional space. This means that for a choice set where there are three alternatives, the Hilbert space is a 3-dimensional space and can be represented as shown in Figure 3.

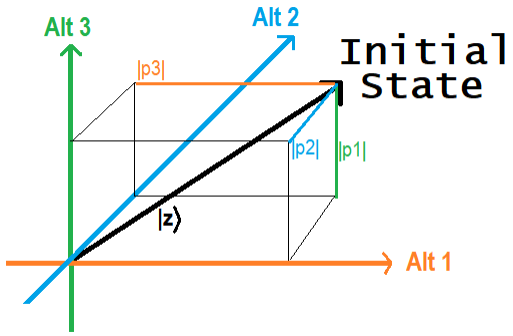

FIGURE 3 : A basic example of 3-dimensional Hilbert probability space

Under quantum logic, a decision-maker has some opinion, initial state or 'indefinite state', denoted $|z\rangle$, which can be represented by a vector of unit length (see Figure 3). When a decision-maker makes a choice, their state goes from 'indefinite' to 'definite', by projecting onto the vector representing the chosen alternative. This means that for each event alternative L_{x_i}, there is a corresponding projection operator ρ_{x_i} that projects $|z\rangle$ onto the vector $|x_i\rangle$.

As the indefinite state vector is of unit length and the subspace is represented by a set of orthonormal vectors, the sum of the squared length of the projections must sum to 1:

$$\sum_{i=1}^{J} |\rho_{x_i}|^2 = 1. \tag{2}$$

A visual proof of this fact is given in Figure 3. The lengths of the three projections can be visualised as the three sides of the cuboid in 3-dimensional space. By Pythagoras' theorem, the fact that the vector of unit length cuts diagonally from one corner to the opposite corner of the cuboid means that the squared lengths of the projections must sum to one.

2.3. A sequence of choices

If a decision-maker makes a second choice across a different set of alternatives, this choice may be influenced by the first. Quantum logic captures this by representing the two events by two separate

subspaces within the Hilbert space, L_x and L_y. Each subspace is separately defined by the set of orthonormal vectors representing the alternatives for each event. This means that L_x is spanned by $|x_1\rangle$, $|x_2\rangle$, ... $|x_J\rangle$ and L_y is spanned by $|y_1\rangle$, $|y_2\rangle$, ... $|y_K\rangle$, where there are J alternatives for choice scenario X and K alternatives for scenario Y.

Revisiting the example presented in Figure 2, a decision-maker might be initially making a choice, X, between commuting by car or train. Under quantum logic, the decision-maker has some initial state (which could be based on past experiences) regarding whether they will choose car or train. All possible states are spanned by the basis vectors $|x_{car}\rangle$, $|x_{train}\rangle$. The closer the vector representing the decision-maker's state is to the vector representing an alternative, the more likely that alternative will be chosen. However, the decision-maker could first be asked some a different question (Y) about whether they consider themselves to be environmentally friendly or not. Under quantum logic, the 'indefinite state' does not change. This means that if the probabilities for alternatives being chosen in question Y are to be different from the probabilities for alternatives being chosen in question X, we need choice Y to be represented[1] by a different set of basis vectors, $|y_{env-friendly}\rangle$, $|y_{env-unfriendly}\rangle$. Consequently, if the decision-maker makes the choice 'I am environmentally friendly', they move through the Hilbert space and their state is projected onto the environmentally friendly vector, $|y_{env-friendly}\rangle$ (see Figure 2). This means that their new state is the vector $|y_{env-friendly}\rangle$ itself. By making choice (Y) first, the original choice X between car and train is made from a different state (or perspective).

Crucially, by moving state (and making what we define a 'quantum rotation'), the squared lengths of the projections onto the vectors for train and for car have changed[2]. As a result, in this example, the decision-maker is more likely to choose to commute by train if they first decide that they are environmentally friendly. This is graphically represented in Figure 2, where the length of the projection onto the vector representing train being chosen has increased (which results in the probability of choosing train increasing relative to the probability of choosing car).

3. BUILDING A CHOICE MODEL FROM QUANTUM LOGIC

Whilst Lipovetsky [2018] has applied quantum models to consumer recall tasks with multi-alternative, multi-attribute alternatives, quantum probability has not ever been applied to multi-alternative, multi-attribute choice scenarios (as far as the authors are aware). In this section, we look at how we can use ideas from quantum probability within a choice model. We do this by first considering what the requirements are for a quantum choice model. We then explore the use of sine and cosine functions, a method which was previously found to be effective for multialternative scenarios [Lipovetsky, 2018]. Next, we consider ideas from two different existing choice models to use within a quantum model framework: regret functions from random regret minimisation [Chorus, 2010] and drift rate functions from the multi-attribute linear ballistic accumulator model [Trueblood et al., 2014a]. Finally, we consider how similar or related choices could be mathematically explained by a 'quantum rotation'.

[1] It is important here to not visualise the state as a point (a,b). It is a vector, $\left|\begin{bmatrix} a \\ b \end{bmatrix}\right\rangle$, meaning that the projection lengths from this state to the set of basis vectors will change if we use a different set of basis vectors.

[2] Two choices that require a different set of basis vectors are known as 'incompatible'. If the choices are in fact compatible and can be represented by the same set of basis vectors, then the order in which the choices are made has no impact on the probabilities of each alternative being chosen. Consequently, quantum probability collapses back into classical probability [Hughes, 1992].

3.1. Requirements

For our choice model to use quantum logic, we need to define a method for calculating an indefinite state vector. If this state vector is of unit length and we take projections from it to a set of orthonormal basis vectors (with one vector for each discrete alternative), then the sum of the squared length of these projections will equal one. Consequently, for each alternative, we need to find the length of the projection, as the square of this length equals the probability with which the alternative is chosen (see Figure 3). This means that if we are to use quantum logic to understand multi-attribute, multi-alternative choices, we must first consider how best to represent the state vector, $|z\rangle$. If, for example, we imagine that we are making a route choice between a cheap and slow alternative and a fast and expensive alternative, the development of a state could be represented by Figure 4.

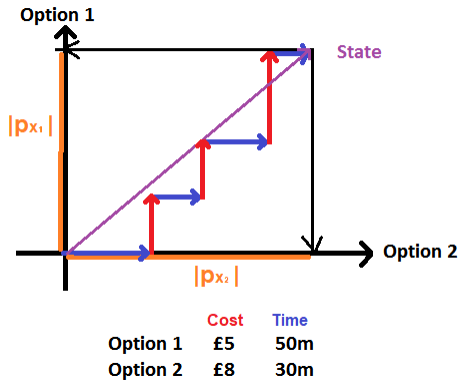

FIGURE 4 : The development of a 'state'

When the decision-maker considers travel time, the second option is better, and consequently the state vector extends in the direction of the vector representing option 2 (and hence increasing the length of the projection onto option 2 whilst not increasing the length of the projection onto option 1). If the decision-maker considers cost, the state vector instead extends in the direction of the vector for alternative 1. At some point the decision-maker makes the choice when they reach some state. To generate this state, we need to know the relative importance of the attributes. This means that one option is to calculate 'utilities' or 'preference values' for each alternative. However, if we used a utility specification from multinomial logit, $U_j = \beta' x_j$, where β is a vector of coefficients and x_j is a vector of observed variables relating to alternative j, then some alternatives could have positive utilities and others might have negative utilities. As the probability of an alternative is the squared length of the projection from the state vector onto the vector for the alternative, positive and negative values would lead to the same result. This means that we require a method for comparing attributes across alternatives that always results in positive (or always negative) values. Additionally, a quantum choice model could have parameters equivalent to alternative specific constants, which would estimate the starting point for the state (which may not necessarily be zero). This in effect acts as an initial preference state in the multidimensional space. Finally, some alternatives may be compared more than others, so weights could be defined for each pair of alternatives (with one fixed for normalisation purposes).

3.2. Cosine and sine functions

It is easy to see from Figure 5 that cosine and sine functions provide one possibility for quantum models. As $sin^2(\theta) + cos^2(\theta) = 1$, one possible method for defining quantum probabilities is to

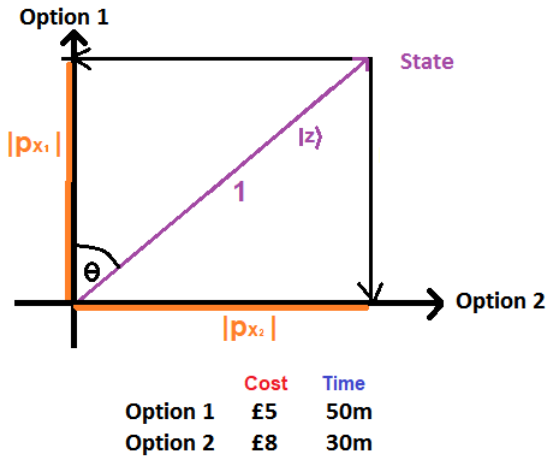

FIGURE 5 : Using geometry within a quantum model

find the angle θ between the state vector $|z\rangle$ and the projection onto the chosen alternative, ρ_{x_j}. For a choice scenario with two alternatives as defined in Figure 5, $cos(\theta) = |\rho_{x_1}|$ and $sin(\theta) = |\rho_{x_2}|$. Consequently, finding the angle θ will give us the probabilities of both alternatives. We thus need to define the angle θ as a weighted sum of the differences between the attributes of the alternatives. This angles can then be used to calculate projections (ρ):

$$\begin{aligned} \theta &= \sum_{k=1}^{K} \beta_k (x_{k_1} - x_{k_2}), \\ \rho(Alt_1) &= sin(\theta), \\ \rho(Alt_2) &= cos(\theta), \\ \rightarrow \rho(Alt_1)^2 + \rho(Alt_2)^2 &= 1, \end{aligned} \tag{3}$$

where $k = 1,...K$ is an index across the attributes, β_k is a coefficient for attribute k and x_{k_i} is the value of attribute k for alternative i. This approach is used by Lipovetsky [2018] to predict which pizza brands participants remember. He demonstrates that for more than two alternatives, we simply need to find additional angles:

$$\begin{aligned} [sin^2(\theta_A)] + [cos^2(\theta_A)] &= 1 \\ \rightarrow sin^2(\theta_A) + cos^2(\theta_A)(sin^2(\theta_B) + cos^2(\theta_B)) &= 1 \\ \rightarrow [sin^2(\theta_A)] + [cos^2(\theta_A) \cdot sin^2(\theta_B)] + [cos^2(\theta_A) \cdot cos^2(\theta_B)] &= 1 \\ \rightarrow [sin^2(\theta_A)] + [cos^2(\theta_A) \cdot sin^2(\theta_B)] + & \\ [cos^2(\theta_A) \cdot cos^2(\theta_B) \cdot sin^2(\theta_C)] + [cos^2(\theta_A) \cdot cos^2(\theta_B) \cdot cos^2(\theta_C)] &= 1. \end{aligned} \tag{4}$$

However, when there are more than two choice alternatives, complications arise. Firstly, it is not clear which differences between alternatives should be used, as we only require $J-1$ differences

where J is the number of alternatives. Crucially, under this system, a small increase in the attribute of one alternative may not have the equivalent impact as a small increase in the attribute of a different alternative. An example of this is given in Table 1. In this set of examples, we have three

TABLE 1 : Examples of probabilities under various coefficients for sine and cosine models

	Parameter/Attribute	Ex. 1 coefs.	Ex. 2 coefs.	Ex. 3 coeffs.
Constant A:	x	0.615	0.615	0.615
Constant B:	y	0.785	0.785	0.785
Attribute for Alt A:	Z_A	3	3.1	3
Attribute for Alt B:	Z_B	3	3	3.1
Attribute for Alt C:	Z_C	3	3	3
Alternative	Projection length	Ex. 1 prob.	Ex. 2 prob.	Ex. 3 prob.
A	$\sin(\theta_A)$	0.333	0.430	0.243
B	$\cos(\theta_A) \cdot \sin(\theta_B)$	0.333	0.285	0.454
C	$\cos(\theta_A) \cdot \cos(\theta_B)$	0.333	0.285	0.304

alternatives {A,B,C} and just a single attribute, Z. We set the relative importance weights $\beta_k = 1$, $\theta_A = x + Z_A - Z_B$ and $\theta_B = y + Z_B - Z_C$. For the first example, we demonstrate that precise constants of $x = 0.615$ and $y = 0.785$ are required to allow for the probability of the three alternatives to be identical if the alternatives have matching value of 3 for attribute Z. If, in example 2, we imagine that the value for attribute Z_A increases to 3.1, then the probabilities adjust appropriately (the probability of alternative B remains the same as the probability of alternative C). This is not the case in example 3, where an equivalent increase in attribute Z_B does not result in alternatives A and C having the same probability. Consequently, we must look for alternative model frameworks for multi-alternative settings.

3.3. Quantum pairwise comparison framework A (QPCA)

We will now show that an alternative possibility for a quantum model framework is to use regret functions from random regret minimisation (RRM) as the key component for the definition of projection lengths for each alternative. The deterministic regret [Chorus, 2010] for respondent n, for alternative i, in choice task t, is given by:

$$R_{int} = \sum_{k=1}^{K} \sum_{j \neq i} ln(1 + e^{\beta_k (x_{jntk} - x_{intk})}) \tag{5}$$

with $k = 1, ..., K$ an index across attributes and β_k a coefficient for the relative importance of attribute k. This has the potential to work for a quantum model as the logarithm guarantees that only positive values are generated from the pairwise comparisons between the alternatives. We can therefore make the following definition for 'quantum pairwise comparison version A (QPCA)' to calculate the length of projection for alternative i (with n the respondent and t the choice task):

$$|\rho_{int}| = \delta_{QPCA,i} + I_0 + \sum_{k=1}^{K} \sum_{j \neq i} wt_{ij} \cdot ln(1 + e^{\beta_k (x_{intk} - x_{jntk})}), \tag{6}$$

where $\delta_{QPCA,i}$ are alternative-specific constants and I_0 is a constant that has the same value across all alternatives[3]. Together, these are used to define the starting point for the state vector in the

[3]Note that x_jntk and x_intk in Equation 6 compared to Equation 5 are reversed, as using negative regret as in RRM models will be ineffective given the projection lengths are later squared.

Hilbert space. wt_{ij} is a weight for the relative importance of the comparison between alternatives i and j, meaning that it is constant across attributes as it is at the alternative level. Once these projection lengths have been calculated, the probability for each alternative can be defined simply as:

$$P_{QPCA,jnt} = \frac{|\rho_{jnt}|^2}{\sum_{i=1}^{J}(|\rho_{int}|^2)}, \tag{7}$$

where $i = 1,...J$ is an index across the possible alternatives. This means that to estimate this model, we require K attribute coefficients, $\frac{(J)(J-1)}{2} - 1$ weights, wt, for the relative importance of comparisons between the different alternatives (where the use of logistic transformations ensures that these weights will sum to one) and J constants to estimate the initial starting state. Whereas adding a constant to the utility of every alternative does not have an impact in random utility models, it is multiplying all of the length of projections by a constant that does not impact the choice probabilities of alternatives under a quantum system (see Equation 2). Consequently, we can have J parameters for J alternatives to define the starting state (either by normalising one attribute specific constant or by not using I_0, the constant that is added to all alternatives). The greater the magnitude of these constants, the less deterministic the choices become.

3.4. Quantum pairwise comparison framework B (QPCB)

The final possibility we consider for a quantum model framework in this paper is the use of drift rate functions from the multi-attribute linear ballistic accumulator model [Trueblood et al., 2014a]. The linear ballistic accumulator (LBA), was originally designed within mathematical psychology, and is a model designed to capture both choices and response times [Brown and Heathcote, 2008]. Under LBA, a decision-maker starts with a random amount of evidence for each alternative. The evidence for each alternative then grows linearly according to a set of drift rates (with one rate for each alternative). The first to reach some threshold is then the chosen alternative. This model was then adjusted for alternatives with multiple attributes (MLBA) and has been used successfully to explain choices between ratings for eyewitness testimony [Trueblood et al., 2014a], consumer and perceptual choices [Turner et al., 2018] and gambling and accommodation choices [Cohen et al., 2017]. Under MLBA, the drift rates are generated from a normal distribution where the mean drift rates are a function of the attributes of the alternatives. Crucially, the mean drift rates are often defined such that they are guaranteed to be positive, meaning that they could be used within a quantum probability framework. The mean drift rate is defined as:

$$d_i = I_0 + \sum_{j \neq i} \sum_{k=1}^{K} \left(w_{x_{k,i,j}} \cdot \beta_k \cdot (x_{k,i} - x_{k,j}) \right), \tag{8}$$

where I_0 is a constant, which can be defined such that $d_j \geq 0$ for all j. Then additionally K is the number of attributes, $w_{x_{k,i,j}}$ is a similarity weighting, β_k is an attribute-specific scaling coefficient for attribute k and $x_{k,i}$ and $x_{k,j}$ are the values for alternatives i and j for attribute k. Whilst similar in appearance to regret functions, rather than using a logarithm, we instead use similarity weightings, which are defined such that they are an exponentially decaying function of distance:

$$w_{x_{k,i,j}} = exp\Big(- \left(\lambda_1 \cdot \left[x_{k,i} \geq x_{k,j} \right] + \lambda_2 \left[x_{k,i} < x_{k,j} \right] \right) \cdot \beta_k \cdot |x_{k,i} - x_{k,j}| \Big). \tag{9}$$

Under MLBA, two different values, λ_1 and λ_2, are used to capture Tversky [1977]'s findings that the subjective similarity between A and B and the subjective similarity between B and A may

not be equal. Given that differences between losses and gains have regularly been shown to be important in a transport context [Hess et al., 2008, Masiero and Hensher, 2010, Stathopoulos and Hess, 2012], this is a useful feature for this quantum model as well. Both λ values should be greater than zero to ensure that attributes that are more similar have a higher similarity value $w_{x_{k,i,j}}$. This results in weights that are between 0 and 1. Finally, we adjust the drift rate specification as before to include weights for the relative importance of comparisons between pairs of alternatives, and to include constants to adjust the initial state vector. Consequently, the length of the projection for an alternative j for decision-maker n in choice task t in our quantum pairwise comparison version B (QPCB) model is:

$$|\rho_{int}| = \delta_{QPCB,i} + I_0 + \sum_{j \neq i} \sum_{k=1}^{K} wt_{i,j} \cdot \left(w_{x_{k,i,j}} \cdot \beta_k \cdot (x_{k,i} - x_{k,j}) \right), \tag{10}$$

where again $\delta_{QPCB,i}$ is used to define the starting point for the state vector and wt_{ij} is a weight for the relative importance of the comparison between alternatives i and j. Once these projection lengths have been calculated, we can use Equation 7 again to calculate the probability of alternatives under this model.

3.5. Quantum rotation

We also look at how quantum models can explain a pair of related choices through a 'quantum rotation'. Under quantum logic, a separate set of basis vectors is required for a different choice. For example, if a decision-maker is selecting their favourite of three alternatives, then the set of basis vectors could be $|Alt1_{best}\rangle$, $|Alt2_{best}\rangle$ and $|Alt3_{best}\rangle$ (see Figure 6 Graph A).

The decision-maker could choose any of the three alternatives and therefore there are three possible projections from the state vector $|z\rangle$ onto the three alternatives (with the corresponding projection lengths labelled $|\rho 1|_{best}$, $|\rho 2|_{best}$ and $|\rho 3|_{best}$ in Figure 6 Graph A). Now suppose the decision-maker chooses alternative 2. Under quantum logic, the state changes from $|z\rangle$ to $|Alt2_{best}\rangle$. However, if the decision-maker was then asked to make a 2nd choice to pick their least favourite, then the projection lengths from $|Alt2_{best}\rangle$ to $|Alt1_{best}\rangle$ or $|Alt3_{best}\rangle$ are both zero, as these are orthogonal vectors. Consequently, to calculate probabilities for the 2nd choice, we must instead assume that the the initial state can still accurately capture the difference between the 2nd and 3rd favoured alternatives. This means the probabilities for the 2nd choice can be found by reducing the dimensionality (moving from the initial state to the new state, $|z_2\rangle$, in Figure 6 Graph B).

Under a typical MNL, the utility (or equivalently regret in RRM) of picking alternative i can easily be adapted to a corresponding utility for not picking alternative i:

$$U(alt_{i-worst}) = -U(alt_{i-best}). \tag{11}$$

For quantum models, however, this translation is not as simple. This is a consequence of using the squared length of the projections to calculate the probability of alternatives (see Equation 7), as using 'negative lengths' for each projection will result in the same probabilities for each alternative. However, when there are only three alternatives (a regular setting for many surveys), there exists a simple transformation. Given that there are two alternatives left, the probability of picking one option as the second best (or second most preferred) equals the probability of picking the other as

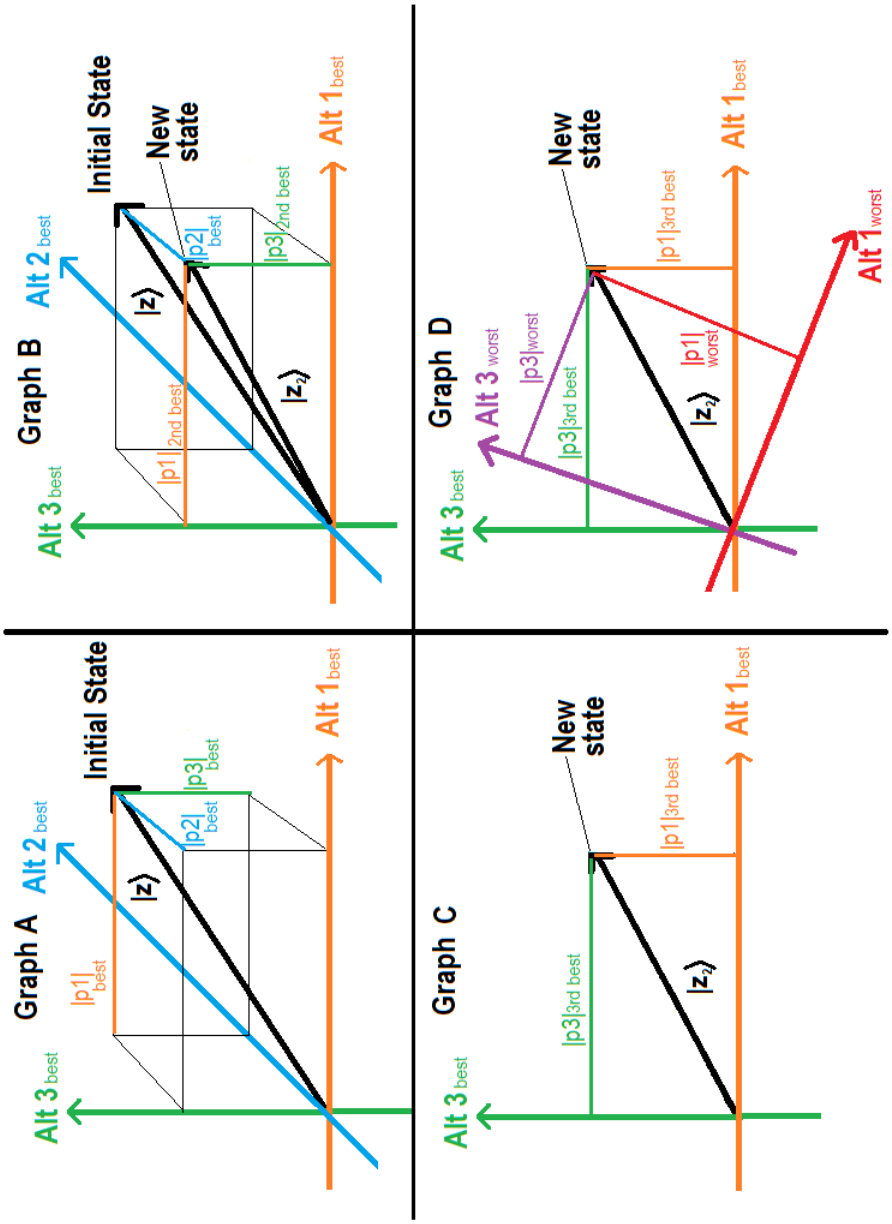

FIGURE 6 : Making best-worst choices under quantum logic

the worst. Consequently given alternatives i and j, we can define the length of the projection for alternative i being the worst as:

$$|\rho_{i_{worst}}| = |\rho_{j_{best}}|. \tag{12}$$

This can be represented graphically by moving from Figure 6 Graph B to Graph C. A different translation would be required if there were more than two alternatives left. For example, one could simply use differences in favour of the alternative rather than against the alternative in the definitions of the projection length for each alternative. For example, Equation 6 would become:

$$|\rho_{int}| = \delta_{QPCA,i} + \sum_{k=1}^{K} \sum_{j \neq i} wt_{ij} \cdot ln(1 + e^{\beta_k (x_{jntk} - x_{intk})}). \tag{13}$$

Equivalently, using $x_{k,j} - x_{k,i}$ instead of $x_{k,i} - x_{k,j}$ in Equation 10 would result in an appropriate translation for QPCB:

$$|\rho_{int}| = \delta_{QPCB,i} + I_0 + \sum_{j \neq i} \sum_{k=1}^{K} wt_{i,j} \cdot \left(w_{x_{k,i,j}} \cdot \beta_k \cdot (x_{k,j} - x_{k,i}) \right), \tag{14}$$

However, Equations 11-14 only work under the assumption that 'worst' is considered directly opposite to 'best' and the sensitivities to attributes is identical when picking favoured and least favoured alternatives.

It is possible that the attributes which are important for a best choice are not necessarily the attributes which are important for a worst choice [Giergiczny et al., 2017]. The simplest way for taking this into account would be to estimate a completely separate set of parameters for worst choices compared to best choices. A different possibility under quantum logic is to instead have a new set of basis vectors, $|Alt1_{worst}\rangle$, $|Alt2_{worst}\rangle$ and $|Alt3_{worst}\rangle$ representing the choice of the worst alternative. This means that if a decision-maker chose alternative 2 as the best alternative, then the probability for choosing, for example, alternative 3 as the worst, is now the squared length of the projection from $|z_2\rangle$ onto $|Alt3_{worst}\rangle$ (rather than $|Alt1_{best}\rangle$, as would be the case under Equation 12). This means that the length of the projection has changed (see Figure 6 Graph D). To calculate the new length of the projection, we would need to know how to make the change of basis from the Hilbert space represented by $|Alt1_{best}\rangle$, $|Alt2_{best}\rangle$ and $|Alt3_{best}\rangle$ to the space represented by $|Alt1_{worst}\rangle$, $|Alt2_{worst}\rangle$ and $|Alt3_{worst}\rangle$. Thus we must estimate a change of basis matrix, M^* that appropriately adjusts the lengths of the projections (with the matrix being of size n, which is the number of alternatives in the initial choice scenario). For all models, this matrix maps original utilities/lengths, p, to new ones, q, simply through matrix multiplication:

$$q_j = \sum_{i=1}^{n} M^*_{j,i} \cdot v_i \tag{15}$$

Mathematically, this is effectively simultaneously a change in both underlying preferences towards alternatives and a change in scale. In our empirical application, we test this idea of a 'quantum rotation' for MNL, RRM, QPCA and QPCB to see whether the rotation can capture as much of a difference in the different sensitivities as a different set of parameters would.

4. EMPIRICAL APPLICATION

In this section, we test our different specifications of quantum models on a number of route choice datasets. First, we describe the five different datasets that we use. Three of these are basic route choice datasets, with the final two providing examples under which we can test 'quantum rotations'. We apply the models to the three basic datasets first, whilst also considering parameter estimates and out-of-sample validation for the most complex of these datasets. Next, we consider a best-worst dataset where we test the ability of our quantum models to capture both best and worst choices. We then also test quantum rotation on a dataset where the order of the alternatives and the attributes is manipulated, thus testing whether this rotation can capture ordering effects. Finally, we consider out of sample validation for quantum rotations. For all models, we use R packages maxLik [Henningsen and Toomet, 2011] and apollo [Hess and Palma, 2019] for estimation of the log-likelihood functions.

4.1. Datasets

4.1.1. Danish dataset

The first dataset we use comes from the Danish value of travel time study [Fosgerau, 2006]. 545 participants completed a total of 4,214 choice tasks. Each task involved a simple choice between two routes, where each task has a cheap but slow alternative and a fast but expensive alternative.

4.1.2. Swiss dataset

The second dataset we use comes from the Swiss value of time study [Axhausen et al., 2008]. 389 participants each make 9 binary route choice tasks. The two alternatives are described by travel cost, travel time, headway and the number of interchanges required to complete the trip.

4.1.3. UK dataset 1

The third dataset that we use is a survey asking public transport commuters living in the UK to make a set of ten choices between three route alternatives in a stated preference survey. A total of 368 participants completed the survey resulting in 3,680 choices. Each choice task involves an invariant reference trip and two hypothetical alternatives. Each alternative is described by seven attributes: travel time (in minutes), fare (£), rate of crowded trips, rate of delays (both out of 10 trips), the average length of delays (across delayed trips) and the availability and cost of a provision of an information service (£). Full details of the dataset are given by Hess and Stathopoulos [2013].

4.1.4. Best-worst dataset

The best-worst dataset we use is similar to our UK dataset involving public transport commuters. A total of 391 participants complete 10 choice tasks described by three alternatives with the same set of attributes as before in the previous UK dataset. As participants choose a best and a worst alternative in each choice task, we have a total of 7,820 choices. For full details of the dataset, readers should refer to Stathopoulos and Hess [2012].

4.1.5. UK dataset 2

The final dataset that we use in this paper comes from the most recent value of travel time study conducted in the UK [Batley et al., 2017]. This dataset comprises of 15,045 choices between two alternatives, one of which is cheaper and the other faster (SP1 in Batley et al. 2017). The advantage of this dataset is that the cheaper alternative is sometimes first and sometimes second, and

additionally the order in which the attributes are displayed is also manipulated across respondents. These ordering effects have previously been found to be significant [Hess et al., 2017], making this an appropriate dataset to test quantum rotations on.

4.2. Basic models

For basic tests of our quantum models, we use the Danish, Swiss and first UK datasets. We test the models on all three datasets before testing out-of-sample performance of the models on the UK dataset.

4.2.1. Estimation results

For all datasets, we compare the quantum models against multinomial logit (MNL). We additionally test the models against random regret minimisation (RRM) for our dataset with more than two alternatives. For all of our models we use an initial parameter search algorithm, based on Bierlaire et al. [2010]'s heuristic for non-linear global optimisation. This reduces the risk of models converging to poor local optima. We test all three quantum specifications. Results for these models are given in Table 2.

For the quantum pairwise comparison version A (QPCA) model, we find that for the Danish dataset, there is no significant difference in model fit by including a number of different starting state vector parameters corresponding to the number of alternatives. Consequently, I_0 and δ_1 are fixed to zero. Additionally, fixing one of the attribute scaling coefficients, β_1, has no impact on model fit either. This is also the case for the Swiss dataset, but not for the UK dataset, which requires a full set of attribute scaling coefficients. Whilst the Swiss model improves with a full sets of constants, there is no significant impact on model fit for the UK dataset when all constants are removed, as the alternative comparison weights, wt, more effectively capture the underlying baseline preferences towards the different alternatives (with worse model fit obtained if alternatives specific weights are used instea of alternative comparison weights).

We observe similar results for quantum pairwise comparison version B (QPCB), which also only requires one parameter to calculate the initial starting state (δ_1 for the Danish and Swiss datasets, I_0 for the UK dataset). For identification purposes, we fix one λ parameter to a value of 1, as dividing λ by some value x and multiplying the β parameters by x results in projection lengths that are also multiplied by x (hence not changing the probability with which each alternative is chosen, see Equation 10). For the Danish dataset, the second λ coefficient can also be fixed to zero without impacting model fit.

Finally, we also test our trigonometric quantum model (TQ) based on the use of sines and cosines. As discussed previously, the TQ model would not work well for more than two alternatives, therefore we only test it on the Danish and Swiss datasets.

Most significantly, the QPCA and QPCB models have better model fit than both MNL and RRM across all three datasets, with a large improvement in particular for the Danish and Swiss datasets. The trigonometric quantum model does not perform as well, having worse fit than MNL for both the Danish and Swiss datasets. This is perhaps unsurprising given that sine and cosine function oscillate. Consequently, there is a restricted range of values that the estimated parameters can take such that the largest difference in attribute value still results in the largest value after a sine or cosine function is applied.

In Table 3, we also give some parameter estimates for models run on the UK data. Whilst the outputs from a quantum model cannot be translated into measures such as the value of travel

TABLE 2 : Results from applying quantum models to the three initial stated preference datasets

Danish				
Model	pars	LL	AIC	BIC
MNL	3	-2,301.25	4,608.50	4,627.54
QPCA	2	-2,012.25	4,028.50	4,041.19
QPCB	3	-2,010.85	4,027.70	4,046.74
TQ	3	-2,490.46	4,986.92	5,005.96
Swiss				
Model	pars	LL	AIC	BIC
MNL	5	-1,667.97	3,345.94	3,376.74
QPCA	5	-1,587.67	3,185.33	3,216.13
QPCB	6	-1,569.46	3,150.92	3,187.88
TQ	5	-1,699.16	3,408.33	3,439.13
UK				
Model	pars	LL	AIC	BIC
MNL	10	-3,360.43	6,740.86	6,802.97
RRM	10	-3,363.91	6,747.82	6,809.93
QPCA	10	-3,339.00	6,698.00	6,760.11
QPCB	12	-3,327.63	6,679.25	6,753.78

time, we can get an indication of the relative importance of the attributes by dividing the parameter estimates by the sum of the absolute value of all eight attribute coefficients[4]. We find that MNL and RRM have near identical relative importances, again suggesting that they are very similar models. QPCA finds relative importances which are largely similar to their respective values under RRM, with the exception of the values for average delay and rate of delays, both of which are more important under QPCA. Under QPCB, rather different values are found. In particular, the relative importance of the fare is higher under QPCB. It is notable that under both quantum models, the comparison of alternatives 1 and 2 is most important, followed by the comparison between 1 and 3. As alternative 1 is the reference alternative, this suggests that decision-makers give more importance to their current alternative. This implies that if they are to choose a new alternative (2 or 3) then it is more important that this alternative is better than the reference alternative than the other new alternative.

4.2.2. Validation results

We also try forecasting for the quantum models to test whether they perform well in out-of-sample validation. In this case, we split the UK dataset into five subsamples. For each subsample, we first estimate the parameters for the MNL, RRM, QPCA and QPCB models on the first 80% of the data before finding the log-likelihood of the remaining 20% validation set under the estimated parameters found for the initial set. For four and three out of five subsets respectively we observe

[4]Note that we use a logarithmic transformation of fare as we find non-linear sensitivities in this data. In line with Hess et al. [2012], a term for reliability is also added by calculating the expected length of delay (rate of delays multiplied by average delay time).

TABLE 3 : Parameter estimates for the models ran on the UK data

UK dataset		MNL	RRM	QPCA	QPCB
Log-likelihood		-3,360.43	-3,363.91	-3,339.03	-3,327.63
AIC		6,740.86	6,747.82	6,698.05	6,679.25
BIC		6,802.97	6,809.93	6,760.16	6,753.78
β_{TT}	estimate	-0.05	-0.03	-0.39	-0.03
	robust t-ratio	-9.50	-9.58	-5.58	-2.64
	rel. importance	0.72%	0.63%	0.68%	0.48%
β_{LFare}	estimate	-6.00	-4.11	-50.29	-6.54
	robust t-ratio	-18.87	-17.66	-7.27	-2.59
	rel. importance	86.08%	86.71%	87.17%	92.70%
β_{Crowd}	estimate	-0.22	-0.15	-1.32	-0.18
	robust t-ratio	-8.58	-8.59	-4.36	-2.60
	rel. importance	3.16%	3.16%	2.29%	2.60%
β_{Delay}	estimate	-0.03	-0.02	-0.82	-0.08
	robust t-ratio	-3.25	-3.06	-2.92	-2.25
	rel. importance	0.43%	0.42%	1.42%	1.12%
β_{Rate}	estimate	-0.19	-0.12	-2.08	-0.07
	robust t-ratio	-5.96	-5.82	-5.38	-1.97
	rel. importance	2.73%	2.53%	3.60%	0.97%
β_{Rel}	estimate	-0.06	-0.04	-0.22	-0.01
	robust t-ratio	-2.64	-2.71	-2.81	-1.82
	rel. importance	0.86%	0.84%	0.38%	0.18%
β_{Inf}	estimate	-0.09	-0.05	-0.36	-0.01
	robust t-ratio	-1.13	-0.95	-0.59	-0.32
	rel. importance	1.29%	1.05%	0.62%	0.13%
β_{InfF}	estimate	0.33	0.22	2.22	0.13
	robust t-ratio	4.95	4.85	4.57	2.67
	rel. importance	4.73%	4.64%	3.84%	1.81%
asc_{alt_1}	estimate	0.39	0.27	0.00	0.00
	robust t-ratio	5.85	4.17	**fixed**	**fixed**
asc_{alt_2}	estimate	0.16	0.17	0.00	0.00
	robust t-ratio	3.3	3.38	**fixed**	**fixed**
wt_{PC12}	estimate			40.58%	43.66%
	robust t-ratio			28.90	14.14
wt_{PC13}	estimate			36.59%	36.09%
	robust t-ratio			25.29	13.31
I_0	estimate			0.00	0.24
	robust t-ratio			**fixed**	3.01
λ_2	estimate				16.17
	robust t-ratio				2.09

that QPCA and QPCB have better out-of-sample log-likelihoods than MNL or RRM (see Table 4) assuring that the improved goodness-of-fit observed in the estimation results are not due to overfitting the data.

TABLE 4 : Results from holdout samples for the different models for the UK dataset

UK dataset		MNL	RRM	QPCA	QPCB
Full Dataset	Estimate	-3,360	-3,364	-3,339	-3,328
Subset 1	Estimate	-2,651	-2,653	-2,633	-2,627
	Forecast	-713	-714	-709	-706
Subset 2	Estimate	-2,721	-2,725	-2,701	-2,688
	Forecast	-642	-641	-639	-642
Subset 3	Estimate	-2,694	-2,697	-2,670	-2,658
	Forecast	-667	-668	-670	-671
Subset 4	Estimate	-2,682	-2,685	-2,672	-2,663
	Forecast	-681	-682	-670	-668
Subset 5	Estimate	-2,684	-2,685	-2,672	-2,662
	Forecast	-679	-680	-670	-668

4.3. Models with quantum rotation

Given that quantum logic can explain ordering effects, it is also important that any quantum choice model retains this property. We therefore also look at the ability of our models to incorporate 'quantum rotations', as defined in Section 3.5.

4.3.1. Best-worst data findings

For our best-worst data, we try five different variations of each model, including three different versions of 'quantum rotations':

1. A basic structure for each model where a 'single' set of parameters are estimated for predicting both best and worst choices, where we simply take negative values (as described in Section 3.5) to translate from describing a best alternative to describing a worst alternative for MNL and RRM. For the quantum models, we try two adjustments for picking the worst alternative as oppose to the best alternative. The first method follows the rotation defined by Equation 12, such that the probability of choosing an alternative as second best is simply the probability that the other unchosen alternative is chosen as the worst. The second method we test is to use different projection lengths as defined by Equations 13 and 14, which uses differences in attributes in favour of the alternative in question (rather than differences against).

2. Allowing for a completely 'separate' set of parameters for the best choices compared to the worst. This is equivalent to running two separate models where the dataset is split into two subsets: one with the only the best alternative choice tasks and one with only the worst alternative choice tasks.

3. A quantum rotation model with a diagonal rotation matrix M_1^* (with zeros off the diagonal). This results in 3 additional parameters for MNL and RRM, and 2 for the quantum models

(as one needs to be fixed for identification purposes, as Equation 7 means that changing all projection lengths by the same constant results in the same probability with which each alternative is chosen).

4. A quantum rotation model with a symmetric rotation matrix M_2^*. This results in 5 additional parameters for all models, as one must now also be fixed for identification purposes for MNL and RRM (increasing all entries of a column of M^* has no impact on the probabilities of choosing an alternative).

5. A quantum rotation model with a fully flexible rotation matrix M_3^*. This results in 6 extra parameters for MNL and RRM and 8 extra parameters for the quantum models.

The results of these models are given in Table 5. For each model, the equation that is used to adjust utilities/lengths for best alternatives to utilities/lengths for worst alternatives is given by the best-worst adjustment equations[5] at the top of Table 5.

Unsurprisingly, every model finds a significant improvement in model fit by having a separate set of parameters for the best alternatives compared to the worst alternatives (in line with the results of Giergiczny et al. 2017). This suggests that the relative sensitivities to the different attributes for a best alternative are not necessarily the same as the relative sensitivities to the different attributes for a worst alternative. Additionally, we also observe that the quantum models that use Equations 13 and 14 for best-worst rotations result in better model fit than models that use Equation 12. This is convenient in that rotations based on Equation 12 are inflexible and cannot be simply expanded for best-worst choice scenarios involving more than three alternatives.

Table 5 also gives log-likelihood gains (LL Gains) for the models using a quantum rotation as a percentage of the improvement gained by using a 'separate' set of parameters compared to a 'single' set. Crucially, we find that both quantum models (and RRM) make vast improvements by incorporating a quantum rotation, compared to models with a single set of parameters. Additionally, models with symmetric or fully flexible rotation matrices (Quantum rotation 2 and 3 in Table 5) provide better model fit for the second version of each quantum model than their respective 'separate' counterparts. Notably, even simple rotations work well for these models, with only two additional parameters required for QPCA and QPCB (2nd versions) to achieve 99% and 95% log-likelihood gains with respect to the improvement of using separate sets of parameters compared to a single set. This suggests that the concept of a 'quantum rotation' is particularly valid for a quantum model. It also demonstrates that the projection onto a basis vector for a best alternative is not equivalent to the projection onto a basis vector for a worst alternative. Consequently, we find that best-worst choices in this dataset are incompatible: a quantum rotation is required to move from a set of basis vectors for best choices to a different set of basis vectors for worst choices.

The impact of the 'quantum rotation' is that the probabilities of each alternative being chosen is altered. For the QPCA and QPCB (2nd version) quantum rotation models, the corresponding (fully flexible matrix) rotation matrices that are estimated are:

$$M_{QPCA} = \begin{bmatrix} 1.000 & -0.018 & -0.004 \\ -0.117 & 2.058 & 0.132 \\ -0.159 & 0.166 & 1.826 \end{bmatrix}, M_{QPCB} = \begin{bmatrix} 1.000 & 0.270 & 0.192 \\ -0.009 & 2.270 & 0.484 \\ -0.396 & 0.416 & 2.393 \end{bmatrix}. \quad (16)$$

[5]Note that these adjustments are not applied to the 'separate' models, for which a completely separate set of parameters is estimated for the best choice compared to the worst choice.

TABLE 5 : Results from models for the best-worst dataset

	models best-worst adjustment	MNL Eq. 11	RRM Eq.11	QPCA Eq.12	 Eq.13	QPCB Eq.12	 Eq. 14
Single	parameters	10	10	11	11	12	12
	Log-likelihood	-5,802.67	-5,803.97	-5,805.38	-5,802.48	-5,943.13	-5,772.78
	BIC	11,695	11,698	11,709	11,704	11,994	11,653
Separate	parameters	20	20	22	22	24	24
	Log-likelihood	-5,666.81	-5,667.24	-5,617.67	-5,617.67	-5,606.92	-5,606.92
	BIC	11,513	11,514	11,433	11,433	11,429	11,429
Quantum Rotation 1	parameters	13	13	13	13	14	14
	Log-likelihood	-5,752.51	-5,681.13	-5,681.86	-5,619.15	-5,692.32	-5,615.13
	BIC	11,622	11,479	11,480	11,355	11,510	11,356
	LL Gain	36.9%	89.8%	65.8%	99.2%	74.6%	95.1%
Quantum Rotation 2	parameters	15	15	16	16	17	17
	Log-likelihood	-5,731.11	-5,671.29	-5,672.43	-5,614.50	-5,662.90	-5,593.01
	BIC	11,597	11,477	11,488	11,372	11,478	11,338
	LL Gain	52.7%	97.0%	70.8%	101.7%	83.3%	108.4%
Quantum Rotation 3	parameters	16	16	19	19	20	20
	Log-likelihood	-5,730.17	-5,671.00	-5,672.43	-5,610.95	-5,662.29	-5,590.82
	BIC	11,604	11,485	11,515	11,392	11,504	11,361
	LL Gain	53.4%	97.3%	70.8%	103.6%	83.5%	109.7%

The impact of these matrices on the probability with which each alternative is chosen is demonstrated in Table 6. We start with a base scenario of two alternatives having a length of projection of 3 and one alternative having a length of 5, and consider how the probability of each alternative changes after a quantum rotation.

TABLE 6 : The impact on the probability of alternatives being chosen after quantum rotations

	Scenario 1	Scenario 2	Scenario 3
	Length of projection		
Alternative 1	5	3	3
Alternative 2	3	5	3
Alternative 3	3	3	5
	Initial Probability		
Alternative 1	58.1%	20.9%	20.9%
Alternative 2	20.9%	58.1%	20.9%
Alternative 3	20.9%	20.9%	58.1%
	New probability under QPCA		
Alternative 1	28.0%	5.6%	6.4%
Alternative 2	41.2%	71.6%	31.3%
Alternative 3	30.9%	22.8%	62.3%
	New probability under QPCB		
Alternative 1	27.2%	9.6%	9.0%
Alternative 2	45.1%	64.6%	33.6%
Alternative 3	27.7%	25.8%	57.4%

For both models, a key change is that the probability of choosing alternative 1 decreases. As alternative 1 is the status quo alternative in this dataset, this suggests that an individual is more likely to choose the status quo as a best choice than a worst choice. Additionally, alternative 2 is more likely to be chosen than alternative 3 (as the worst). The higher values for elements $M_{2,3}$ and $M_{3,2}$ for M_{QPCB} compared to M_{QPCA} results in scenarios 2 and 3 being less deterministic after a rotation for QPCB compared to QPCA.

4.3.2. Contextual and ordering effects

Our next test of quantum rotation is to investigate its ability to capture contextual effects. For example, our second UK dataset has some choice sets with the cheaper alternative shown first, and some with the faster alternative shown first. Additionally, cost is sometimes on the left and sometimes on the right. Whilst we could again use a full set of different parameters for the four different scenarios, quantum rotations could also be used. Two different quantum rotations are required, one for the order of the alternatives and the other for the order of the attributes. If a quantum rotation improves the model, it implies that the choice scenarios cannot be treated as equivalent if the context in which the choice is presented changes. We again test single and separate sets of parameters as well as a two versions of quantum rotations. The basic MNL and

RRM models use parameters for cost and time as well as having one alternative specific constant[6]. QPCA and QPCB additionally have a second alternative specific constant[7] and QPCB additionally has two sensitivity parameters. The separate parameter models simply have four times as many parameters (a set of parameters for each combination of attribute and alternative order). The first set of rotations use diagonal matrices, resulting in one extra parameter per rotation for the quantum models and two extra per rotation for MNL and RRM. The second set uses matrix rotations with full sets of free parameters. The results of these models are given in Table 7.

TABLE 7 : Results from applying quantum rotations to models for UK dataset 2

	models	MNL	RRM	QPCA	QPCB
Single	parameters	3		4	5
	Log-likelihood	-9,603.17		-9,369.61	-9,210.60
	BIC	19,235		18,778	18,469
Separate	parameters	12		16	20
	Log-likelihood	-9,584.18		-9,353.15	-9,189.43
	BIC	19,284		18,860	18,571
Quantum Rotation 1	parameters	7	7	6	7
	Log-likelihood	-9,593.83	-9,591.44	-9,359.81	-9,201.74
	BIC	19,255	19,250	18,777	18,471
	LL Gain	49.2%	61.8%	59.6%	41.9%
Quantum Rotation 2	parameters	11	11	10	11
	Log-likelihood	-9,592.17	-9,588.88	-9,355.14	-9,198.82
	BIC	19,290	19,284	18,806	18,503
	LL Gain	57.9%	75.3%	87.9%	55.7%

For all models, it appears that using separate sets of parameters results in an improvement in model fit. Whilst the quantum rotation models are not as successful as capturing the difference between the contextual situations, full matrix rotations for QPCA result in a model that is only 2 log-likelihood units worse than a model with a full set of separate parameters, which has 6 additional parameters. Consequently, the quantum rotation models achieve the best BIC values. Notably, both quantum models significantly outperform the MNL and RRM models. The parameter estimates for the quantum rotation matrices for changing from having the cheaper alternative on the left (first) to on the right (second) are:

$$M_{QPCA} = \begin{bmatrix} 1.000 & 0.029 \\ -0.014 & 1.130 \end{bmatrix}, M_{QPCB} = \begin{bmatrix} 1.000 & -0.053 \\ 0.102 & 1.225 \end{bmatrix}. \quad (17)$$

The estimates for the quantum rotation matrices for changing from having the travel time first to having the travel cost first are:

$$M_{QPCA} = \begin{bmatrix} 1.000 & 0.101 \\ -0.001 & 1.097 \end{bmatrix}, M_{QPCB} = \begin{bmatrix} 1.000 & 0.043 \\ -0.008 & 0.898 \end{bmatrix}. \quad (18)$$

[6]Note that whilst these models perform identically when there are only two alternatives, this is not the case when a quantum rotation is applied. This is a result of differing levels of impact on the difference in utility of the two alternatives.

[7]This does not cause identification issues as increasing both values here simply results in a less deterministic choice. As a contrast to the previous models in this paper, the 2nd alternative specific constant results in a significant improvement for both quantum models for this dataset.

We use similar scenarios before to test the impact of these rotations on the probabilities of picking the first or second alternatives. Using lengths of 3 and 5 for the base case (with the cheaper alternative first and the travel time first), the respective probabilities for choosing the alternatives are 27% for the alternative with length 3 and 73% for the alternative with length 5 (See Table 8). For all scenarios and all models, alternative 2 is always more likely to be chosen if it is the cheaper alternative, demonstrating that there is a bias towards picking the cheaper alternative. Whilst there is little change in the probabilities under QPCA for having the travel cost or travel time first, QPCB rotation parmeters results in alternative 1 being picked more often if travel cost is listed first. This is a result of having a value lower than 1 in matrix element $M_{QPCB_{2,2}}$.

TABLE 8 : Probabilities of picking alternatives after quantum rotations under both QPCA and QPCB

	QPCA		QPCB	
	Scenario 1	Scenario 2	Scenario 1	Scenario 2
	Length of projection			
Slow & cheap	3	5	3	5
Fast & expensive	5	3	5	3
	Cheaper alternative first, travel time first			
Slow & cheap	26.5%	73.5%	26.5%	73.5%
Fast & expensive	73.5%	26.5%	73.5%	26.5%
	Cheaper alternative second, travel time first			
Slow & cheap	23.9%	70.1%	15.3%	57.2%
Fast & expensive	76.1%	29.9%	84.7%	42.8%
	Cheaper alternative first, travel cost first			
Slow & cheap	29.0%	72.3%	34.1%	78.9%
Fast & expensive	71.0%	27.7%	65.9%	21.1%
	Cheaper alternative second, travel cost first			
Slow & cheap	26.7%	69.0%	21.5%	64.6%
Fast & expensive	73.3%	31.0%	78.5%	35.4%

4.3.3. Validation results for quantum rotations

Finally, we again test for overfitting, this time for models for which full quantum rotations have been applied. We test both the quantum rotations from best to worst choice for our best-worst dataset and both rotations for the 2nd UK dataset. In all cases, the first 80% of the dataset is used for model estimation, with the remaining 20% used for validation. The results of these models are given in Table 9. Crucially, it appears that for both datasets, neither QPCA or QPCB overfits the data. It is in these models that we see more of a difference between the two quantum models, with QPCB outperforming QPCA in 7 out of 10 validation subsets.

TABLE 9 : Results from holdout samples for the different models under a full quantum rotation for the best-worst and 2nd UK datasets

Best-worst dataset					
Model		MNL	RRM	QPCA	QPCB
Full Dataset	Estimate	-5,730	-5,671	-5,611	-5,591
Subset 1	Estimate	-4,563	-4,515	-4,457	-4,439
	Forecast	-1,172	-1,157	-1,157	-1,156
Subset 2	Estimate	-4,606	-4,573	-4,520	-4,505
	Forecast	-1,129	-1,103	-1,096	-1,091
Subset 3	Estimate	-4,534	-4,480	-4,430	-4,431
	Forecast	-1,203	-1,198	-1,186	-1,188
Subset 4	Estimate	-4,628	-4,578	-4,536	-4,519
	Forecast	-1,105	-1,095	-1,078	-1,074
Subset 5	Estimate	-4,570	-4,522	-4,485	-4,473
	Forecast	-1,165	-1,154	-1,131	-1,122
UK dataset 2					
Model		MNL	RRM	QPCA	QPCB
Full Dataset	Estimate	-9,592	-9,589	-9,355	-9,199
Subset 1	Estimate	-7,722	-7,721	-7,491	-7,386
	Forecast	-1,873	-1,871	-1,866	-1,816
Subset 2	Estimate	-7,667	-7,665	-7,475	-7,337
	Forecast	-1,928	-1,927	-1,882	-1,864
Subset 3	Estimate	-7,546	-7,544	-7,445	-7,374
	Forecast	-2,079	-2,077	-1,914	-1,827
Subset 4	Estimate	-7,671	-7,668	-7,484	-7,397
	Forecast	-1,924	-1,924	-1,873	-1,903
Subset 5	Estimate	-7,728	-7,724	-7,516	-7,451
	Forecast	-1,869	-1,870	-1,841	-1,845

5. CONCLUSIONS

In this paper, we move away from the tried and tested alternatives to random utility maximisation by considering ideas first developed in quantum physics. With the probability framework developed in quantum physics having made a successful transition to cognitive psychology, we look at whether it can be operationalised into a choice model framework. Under quantum logic, a decision-maker has some 'indefinite state' regarding their preferences over alternatives, from which the probabilities of each alternative can be inferred. Thus a key component of this paper is the development of specifications for the indefinite state.

We find two possible suitable specifications for the 'indefinite state' which allow us to incorporate quantum logic within a choice model. The first uses an adapted specification based on random regret minimisation. The second uses a variation of the specification of the mean drift rates within a multi-attribute linear ballistic accumulator model. We find that our quantum pairwise comparison version A (QPCA) model provides good model fit and outperforms multinomial logit (MNL) and random regret minimisation (RRM) across three simple route choice datasets as well as providing good out-of-sample fit for the most complex of these datasets. Additionally, the quantum pairwise comparison version B (QPCB) model also provides good fit for all three datasets. Whilst QPCB has better model fit, QPCA provides more robust parameter estimates. The third quantum framework possibility, based on sines and cosines, appears unsuitable. This is due to a restricted range of attribute values allowed for two alternatives, and basic probability issues for three or more alternatives. Nevertheless, the positive results from our initial tests on the QPCA and QPCB models suggest that there is ample scope for models with a quantum framework to be used within travel behaviour modelling. They are simple to run and estimate, meaning that they could be applied to a wide range of choice scenarios. However, for these models to make a transition into large-scale modelling, an alternative specification would need to be defined to avoid the same pitfall of random regret minimisation for large numbers of alternatives: a comparison between every pair of alternatives quickly becomes computationally infeasible.

Another issue with the current specifications of the models is that it could be argued that there is no real 'quantum insight' in the model structure. QPCA is in essence a random regret minimisation model with a different error structure. However, results from our quantum rotation models suggest that there are benefits of bringing quantum logic into choice models. Crucially, our best performing models for the best-worst dataset and the contextual choice dataset, after allowing for model complexity to be taken into account, are the rotated QPCA and rotated QPCB models. This suggests that there is some merit in the concept of quantum rotation, which suggests a different set of basis vectors are required for different choice situations. For example, the rotation works well for capturing the difference between best and worst choice. Despite the fact that the best-worst choices are related, quantum rotation suggests that these choices are in fact incompatible: the choices cannot be made at the same time and consequently they may not follow the classical probability law of distributivity. This means that different choices may be observed depending on whether the decision-maker chooses the best or worst alternative first.

While the quantum rotation findings here are just illustrative examples, these results demonstrate that there is a large amount of scope for future work within travel behaviour modelling. For example, large-scale models frequently aim to understand a series of related, sequential choices. Given the ability of quantum rotation to capture the translation between best and worst choices, it theoretically should also work for a larger sequence of related choices where continuously adding on separate sets of parameters may not be possible. Thus ordering effects and state dependence

may be well captured by models within a quantum framework. Furthermore, it may be possible to mitigate the impacts of contextual effects by applying the appropriate quantum rotation from other models that account for the same effect. Future efforts could also compare quantum frameworks against other models that are specifically designed to deal with contextual effects, such as prospect theory. These future possibilities combined with the positive results in our route choice datasets mean that this paper serves as a proof-of-concept that quantum ideas can be incorporated into choice models aiming to understand travel behaviour.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial support by the European Research Council through the consolidator grant 615596-DECISIONS.

REFERENCES

Avineri, E. and Bovy, P. H. (2008). Identification of parameters for a prospect theory model for travel choice analysis. *Transportation Research Record*, 2082(1):141–147.

Axhausen, K. W., Hess, S., König, A., Abay, G., Bates, J. J., and Bierlaire, M. (2008). Income and distance elasticities of values of travel time savings: New swiss results. *Transport Policy*, 15(3):173–185.

Batley, R., Bates, J., Bliemer, M., Börjesson, M., Bourdon, J., Cabral, M. O., Chintakayala, P. K., Choudhury, C., Daly, A., Dekker, T., et al. (2017). New appraisal values of travel time saving and reliability in great britain. *Transportation*, pages 1–39.

Ben-Akiva, M., McFadden, D., Train, K., et al. (2019). Foundations of stated preference elicitation: Consumer behavior and choice-based conjoint analysis. *Foundations and Trends® in Econometrics*, 10(1-2):1–144.

Bierlaire, M., Thémans, M., and Zufferey, N. (2010). A heuristic for nonlinear global optimization. *INFORMS Journal on Computing*, 22(1):59–70.

Birkhoff, G. and Von Neumann, J. (1936). The logic of quantum mechanics. *Annals of mathematics*, pages 823–843.

Brown, S. D. and Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, 57(3):153–178.

Bruza, P. D., Wang, Z., and Busemeyer, J. R. (2015). Quantum cognition: a new theoretical approach to psychology. *Trends in cognitive sciences*, 19(7):383–393.

Busemeyer, J. R., Pothos, E. M., Franco, R., and Trueblood, J. S. (2011). A quantum theoretical explanation for probability judgment errors. *Psychological review*, 118(2):193.

Busemeyer, J. R. and Townsend, J. T. (1992). Fundamental derivations from decision field theory. *Mathematical Social Sciences*, 23(3):255–282.

Busemeyer, J. R. and Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, 100(3):432.

Chorus, C. G. (2010). A new model of random regret minimization. *EJTIR, 10 (2), 2010*.

Chorus, C. G., Arentze, T. A., and Timmermans, H. J. (2008). A random regret-minimization model of travel choice. *Transportation Research Part B: Methodological*, 42(1):1–18.

Cohen, A. L., Kang, N., and Leise, T. L. (2017). Multi-attribute, multi-alternative models of choice: Choice, reaction time, and process tracing. *Cognitive Psychology*, 98:45–72.

Cunha-e Sá, M. A., Madureira, L., Nunes, L. C., and Otrachshenko, V. (2012). Protesting and justifying: a latent class model for contingent valuation with attitudinal data. *Environmental and Resource Economics*, 52(4):531–548.

Fosgerau, M. (2006). Investigating the distribution of the value of travel time savings. *Transportation Research Part B: Methodological*, 40(8):688–707.

Giergiczny, M., Dekker, T., Hess, S., and Chintakayala, P. K. (2017). Testing the stability of utility parameters in repeated best, repeated best-worst and one-off best-worst studies. *European Journal of Transport and Infrastructure Research*, 17(4).

Guevara, C. A. and Fukushi, M. (2016). Modeling the decoy effect with context-rum models: Diagrammatic analysis and empirical evidence from route choice sp and mode choice rp case studies. *Transportation Research Part B: Methodological*, 93:318–337.

Hancock, T. O., Hess, S., and Choudhury, C. F. (2018). Decision field theory: Improvements to current methodology and comparisons with standard choice modelling techniques. *Transportation Research Part B: Methodological*, 107:18–40.

Heisenberg, W. (1927). Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. *Zeitschrift für Physik*, 43(3):172–198.

Henningsen, A. and Toomet, O. (2011). maxlik: A package for maximum likelihood estimation in R. *Computational Statistics*, 26(3):443–458.

Hess, S., Daly, A., and Batley, R. (2018). Revisiting consistency with random utility maximisation: theory and implications for practical work. *Theory and Decision*, 84(2):181–204.

Hess, S., Daly, A., Dekker, T., Cabral, M. O., and Batley, R. (2017). A framework for capturing heterogeneity, heteroskedasticity, non-linearity, reference dependence and design artefacts in value of time research. *Transportation Research Part B: Methodological*, 96:126–149.

Hess, S. and Palma, D. (2019). Apollo: a flexible, powerful and customisable freeware package for choice model estimation and application, www.apollochoicemodelling.com.

Hess, S., Rose, J. M., and Hensher, D. A. (2008). Asymmetric preference formation in willingness to pay estimates in discrete choice models. *Transportation Research Part E: Logistics and Transportation Review*, 44(5):847–863.

Hess, S. and Stathopoulos, A. (2013). A mixed random utility - random regret model linking the choice of decision rule to latent character traits. *Journal of Choice Modelling*, 9:27–38.

Hess, S., Stathopoulos, A., and Daly, A. (2012). Allowing for heterogeneous decision rules in discrete choice models: an approach and four case studies. *Transportation*, 39(3):565–591.

Hughes, R. I. (1992). *The structure and interpretation of quantum mechanics*. Harvard university press.

Kennard, E. H. (1927). Zur quantenmechanik einfacher bewegungstypen. *Zeitschrift für Physik*, 44(4-5):326–352.

Krajbich, I., Lu, D., Camerer, C., and Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in psychology*, 3:193.

Leong, W. and Hensher, D. A. (2014). Relative advantage maximisation as a model of context dependence for binary choice data. *Journal of choice modelling*, 11:30–42.

Lipovetsky, S. (2018). Quantum paradigm of probability amplitude and complex utility in entangled discrete choice modeling. *Journal of choice modelling*, 27:62–73.

Mahieu, P.-A., Crastes, R., Louviere, J., Zawojska, E., et al. (2016). Rewarding truthful-telling in stated preference studies. Technical report.

Masiero, L. and Hensher, D. A. (2010). Analyzing loss aversion and diminishing sensitivity in a freight transport stated choice experiment. *Transportation Research Part A: Policy and Practice*, 44(5):349–358.

Planck, M. (1901). Ueber das gesetz der energieverteilung im normalspectrum. *Annalen der Physik*, 309(3):553–563.

Pothos, E. M., Busemeyer, J. R., and Trueblood, J. S. (2013). A quantum geometric model of similarity. *Psychological Review*, 120(3):679.

Seetharaman, P. (2003). Probabilistic versus random-utility models of state dependence: an empirical comparison. *International Journal of Research in Marketing*, 20(1):87–96.

Stathopoulos, A. and Hess, S. (2012). Revisiting reference point formation, gains–losses asymmetry and non-linear sensitivities with an emphasis on attribute specific treatment. *Transportation Research Part A: Policy and Practice*, 46(10):1673–1689.

Trueblood, J. S., Brown, S. D., and Heathcote, A. (2014a). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological review*, 121(2):179–205.

Trueblood, J. S. and Busemeyer, J. R. (2011). A quantum probability account of order effects in inference. *Cognitive science*, 35(8):1518–1552.

Trueblood, J. S. and Busemeyer, J. R. (2012). Quantum information processing theory. In *Encyclopedia of the Sciences of Learning*, pages 2748–2751. Springer.

Trueblood, J. S., Pothos, E. M., and Busemeyer, J. R. (2014b). Quantum probability theory as a common framework for reasoning and similarity. *Frontiers in psychology*, 5:322.

Turner, B. M., Schley, D. R., Muller, C., and Tsetsos, K. (2018). Competing theories of multialternative, multiattribute preferential choice. *Psychological Review*, 125(3):329.

Tversky, A. (1977). Features of similarity. *Psychological Review*, 84(4):327.

Vitetta, A. (2016). A quantum utility model for route choice in transport systems. *Travel Behaviour and Society*, 3:29–37.

Yu, J. G. and Jayakrishnan, R. (2018). A quantum cognition model for bridging stated and revealed preference. *Transportation Research Part B: Methodological*, 118:263–280.