Quantum Type Indeterminacy in Dynamic Decision-Making: Self-control Through Identity Management

Jérôme Busemeyer, Ariane Lambert-Mogiliansky

PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

WORKING PAPER N° 2012 – 20

Quantum Type Indeterminacy in Dynamic Decision-Making: Self-control Through Identity Management

Ariane Lambert-Mogiliansky
Jérôme Busemeyer

JEL Codes: F31, F35, F41

Keywords: Indeterminacy ; Decision-making ; Self-control ; Identity

PARIS-JOURDAN SCIENCES ECONOMIQUES
48, BD JOURDAN – E.N.S. – 75014 PARIS
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10
www.pse.ens.fr

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE – ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES
ÉCOLE DES PONTS PARISTECH – ECOLE NORMALE SUPÉRIEURE – INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE

Quantum Type Indeterminacy in Dynamic Decision-Making: Self-control Through Identity Management

A. Lambert-Mogiliansky*and Jerome Busemeyer†

April 23, 2012

Abstract

The Type Indeterminacy model is a theoretical framework that uses some elements of quantum formalism to model the constructive preference perspective suggested by Kahneman and Tversky. In a dynamic decision context type indeterminacy induces a game with multiple selves associated with a state transition process. We define a Markov perfect equilibrium among the selves with individual identity (preferences) as the state variable. The approach allows to characterize generic personality types and derive some comparative static results.

"The idea of self-control is paradoxical unless it is assumed that the psyche contains more than one energy system, and that these energy systems have some degree of independence from each others" (McIntosh 1969)

1 Introduction

Everyone has been confronted with an ambiguous picture. A famous one allows one to either see the profile of Sigmund Freud's head or the naked body of a little women. The remarkable thing is that one cannot see both simultaneously. The two pictures are true but they are incompatible. This reminds of "Bohr complementarity" in Quantum Physics.[1] The complementarity principle states that some objects have multiple properties that appear to be contradictory. Sometimes it is possible to switch back and forth between different views of an object to observe these properties, but in principle, it is

*Paris School of Economics, alambert@pse.ens.fr

†Indiana University USA, jbusemey@indiana.edu

[1] Niels Bohr is one of the founders of Quantum Mechanics. After intense dicussions with Heisenberg and Pauli, he introduces the fundamental concept of complementarity at the Côme conference in 1927 which was followed by numerous publications.

impossible to view both at the same time, despite their simultaneous existence.[2] Similarities between human sciences and quantum physics were early recognized by the founders of Quantum Mechanics, including Bohr and Heisenberg.[3] In particular Bohr was heavily influenced by the psychology and philosophy of knowledge of Harald Höffding.[4] A fundamental similarity stems from the fact that in both fields *the object of investigation cannot (always) be separated from the process of investigation.* Quantum Mechanics and in particular its mathematical formalism was developed to respond to that epistemological challenge (see the Introduction in [9]). In our view this makes it fully legitimate to explore the value of the mathematical formalism of quantum mechanics to the study of human behavioral phenomena.

Under the last decade scholars from social sciences, psychology, physics and mathematics have contributed to the development of a "quantum-like" decision theory based on the premises of (non-classical) indeterminacy (see e.g., [19, 13, 14, 15, 17, 18, 22, 34, 36]). This line of research has shown itself very fruitful to explain a wide variety of behavioral phenomena ranging from cognitive dissonance to preferences reversal, the inverse fallacy or the disjunction effect. A central feature of (non-classical) indeterminacy is, according to G. Mackey ([38]), that it places limitations on the system, "The laws of quantum mechanics place certain restrictions on the possible simultaneous probability distribution of various observables" (p. 61). Similarly, indeterminacy in decision theory captures cognitive limitations of the individual in the following sense. The individual is not simultaneously endowed with a preference order over all possible subsets of alternatives. Instead, as one elicits the individual's preferences with respect to one subset of alternatives (in a choice experiment), his preferences with respect to another subset of alternatives (associated with an incompatible choice experiment) are modified so behavior can e.g., exhibit preference reversal.[5] This cognitive limitation implies that individual behavior is boundedly rational in the sense that it is not consistent with the existence of a complete ordering over the universal set of alternatives.[6]

The starting point for our approach is that we depart from the classical dogma that individuals are endowed with preferences and attitudes that motivate their behavior. Instead, we propose that the motivational underpinning of behavior is intrinsically uncertain, *i.e.*, indeterminate. It is only at the moment the individual selects an action that a specific type (preferences) is actualized. It is not merely revealed but rather determined in the sense that prior to the choice, there is an irreducible

[2] Ambiguous pictures do not feature all characteristics of Bohr complementarity. For a rigorous analysis of quantum like phenomena in human perception see e.g., [2]. They study oscillations in bi-stable perception of the Necker cube.

[3] Heisenberg [30] formulates three regions of knowledge depending on the degree of separability between the object and the process of investigation. The second region corresponds to the case where we have non-separabilities. Heisenberg puts quantum physics in that region but also psychology and biology.

[4] Bohr considered introspective psychology not as an illustration but as the paradigmatic description of the epistemological limitations in modern physics. [11]

[5] Preference ordering over different subsets may be complementary properties in the sense of Bohr complementarity.

[6] For a comparison between the behavior of a classical rational man and that of a type indeterminate agent see [36] Section 3.1.

multiplicity of potential types. This idea is very much in line with Tversky and Simonson according to whom "There is a growing body of evidence that supports an alternative conception according to which preferences are often constructed not merely revealed in the elicitation process. These constructions are contingent on the framing of the problem, the method of elicitation, and the context of the choice"(in [33], p. 525–526).

The basic model of static decision-making with Type Indeterminate agents, the TI-model, is formulated in [36]. As we consider dynamic individual optimization, the TI-model induces a game among potential incarnations of the individual. In each period these potential incarnations represent conflicting desires or propensities to act. We formulate the decision problem in terms of a game between a multiplicity of (one-period lived) players, the selves.[7] They are linked to each other through two channels: (i) the selves share a common interests in the utility of the future incarnations of the individual and (ii) they are connected to each other in a process of state transition (which captures indeterminacy). In each period the current selves form intentions to act. One action is played by the individual but the whole profile of (intended) actions matters to tomorrow's identity by force of the state transition process. This creates a strategic concern among contemporaneous selves. In particular when the selves pool, the individual's preferences are unchanged while if they choose different actions preferences are modified. We define a Markov Perfect Equilibrium among the selves where the state variable is the individual's identity. In our model behavior affects future preferences (identity) and in particular a concern for identity (self-image) arises endogenously because identity determines future expected utility. Choice behavior exhibits deviations from standard utility maximization. It is characterized by some extent of self-control: some selves may refrain from short-run gains (and pool with others) to secure a desirable identity. It can also feature dynamic inconsistency because as preferences are modified, the choices made by the individual through time are not consistent with a stable preference order. The model delivers some of the predictions of Benabou and Tirole' ([6, 7]) in particular regarding the impact of the concern for identity on choice behavior. It also generates novel predictions. We characterize generic classes of personality/behavior: a little conflicted, weakly decisive but behaviorally stable character and a highly conflicted, strongly decisive and behaviorally unstable character.

We recently witnessed a renewed interest among prominent economic theorists for the issue of self-control and dynamic inconsistency in decision-making (see e.g., [27, 28, 29, 24, 25]). There exists a significant theoretical literature pioneered by Strotz [45] dealing with various forms of dynamic inconsistency. A larger share of this literature focuses on inconsistency that arises because the individual does not discount the future at a constant rate. Some form of myopia is assumed instead

[7] Although this paper uses elements of the quantum formalism in games, we are not dealing with so-called quantum games which study how the extension of classical moves to quantum ones can affect the analysis of the game. That approach consists in changing the strategy space, see for instance [20]. We model a game where the agent is characterized by (quantum) indeterminate preferences.

(e.g., quasi-hyperbolic discounting). Another approach to the planning problem, first proposed by Peleg and Yaari [43], models individual decision-making in terms of multiple selves. Various ways to model those selves and interaction between them have been investigated. Fudenberg and Levine [24, 25] develop a dual-self model of self-control with a long-term benevolent patient self and a multiplicity of impulsive short-term selves. This particular structure allows them to write the game as a decision problem and they can explain a number of behavioral paradoxes. In the present paper, we argue that the quantum approach to decision-making offers a suitable framework to the McIntosh's paradox of self-control because the indeterminacy of individual preferences is equivalent to the multiplicity of the selves (the potential eigentypes). Our approach contributes to the literature on self-control by investigating a mechanism of self-control based on identity management that appeals to (intrinsic) type indeterminacy. The technology for the evolution of identity reflects the dynamics of state transition under non-classical indeterminacy and individual identity is the equilibrium outcome of the interaction between the selves. A contribution is to formalize internal conflicts and explain features of self-management without appealing to time preferences that have been the quasi exclusive focus of earlier works on dynamic inconsistency. Moreover we can connect to another branch of research related to identity and self-image extensively investigated in psychology (in particular within self-perception theory, see next section) and more recently in economics see e.g., Benabou and Tirole [6, 7].

The paper is organized as follows. In the next section we expose some motivating puzzles and argue that our modelling is closely linked to the so called self-perception theory in psychology. Next we present a general model of dynamic optimization by a type indeterminate agent. We illustrate main features in an example. We define generic personality types in terms of the fundamentals of the model and derive some comparative statics results. Finally, we discuss links between our model and the economic literature on self-control and identity.

Recent interest among prominent economic theorists for the issue of self-control (see e.g., Gul and Pesendorfer (2001, 2004, 2005), Fudenberg and Levine (2006, 2010)), often builds on the intuition that an individual may be better described by a multiplicity of selves who may have diverging interests and intentions than as a single piece of coherent intentions. Various ways to model those selves and interaction between them have recently been investigated. Often they amount to enriching the standard model by adding short-run impatient selves. In this paper, we argue that the quantum approach to decision-making provides a suitable framework to the McIntosh's paradox of self-control because the indeterminacy of individual preferences precisely means multiplicity of the selves (the potential eigentypes).

To many people it may appear unmotivated or artificial to turn to Quantum mechanics when investigating human behavioral phenomena. However, the founders of QM, including Bohr [11] and Heisenberg [30] early recognized the similarities between the two fields. In particular Bohr was influenced by the psychology and philosophy of knowledge of Harald Höffding. The similarity stems from the fact that in both fields *the object of investigation cannot (always) be separated from the process of investigation*.[8] Quantum Mechanics and in particular its mathematical formalism was developped to respond to that epistemological challenge(see the introduction in [?] for a enlighting presentation). In the next section we provide an extensive record of the revlevant psychological literature.

Introducing indeterminacy in social sciences and psychology has shown itself very fruitful to explain a wide variety of behavioral phenomena ranging from cognitive dissonance to preferences reversal, the inverse fallacy or the disjunction effect. Indeterminacy is formalized in quantum physics, some of its tools are used in modelling decision-making and to modelling behavior more generally ((see e.g., Deutsch (1999), Busemeyer et al. (2006, 2007, 2008), Danilov et al. (2008), Franco (2007), Danilov et al. (2008), Khrennikov (2010), Lambert-Mogiliansky et al. (2009)). As soon as we consider dynamic individual optimization the quantum approach to decision-making opens up for the issue of self-control or, as we prefer to call it self-management, In contrast with the recent papers on self-control, we can address these issues without introducing the time dimension but focusing instead on the sequential character of decision-making. In this paper we propose an introduction to dynamic optimization using the Type indeterminacy model (Lambert-Mogiliansky et al. 2009). The basic assumption will be that the agent is aware of his type indeterminacy, that is of the way his decisions have impact on his future type and consequently on future choices and (expected) outcomes. We show that, in a TI-model, dynamic optimization translates into a game of self-management among multiple selves. Its natural solution concept is Bayes-Nash equilibrium i.e., a decentralized equilibrium among the selves.

We are used to situations where current decisions affect future decisions. This is the case whenever the decisions are substitutes or complements. A choice made earlier changes the value of future choices by making them more valuable when the choices are complements or less valuable when they are substitutes. The preferences are fixed over time but the endowment changes. The theories of addiction address the case when a current decision impact on future *preferences*.[9] Generally however, the decision theoretical literature assumes that preferences are fixed unless a special additional structure is provided. Yet, these theories fail to explain many phenomena related to action that seems justified by a concern to affect the player's own identity or preferences or that of other players. Examples are

[8] in the words of Bohr "the impossibility of a sharp separation between the behavior of atomic object and the interaction with the measuring instruments which serves to define the condition under which the phenomena appears". In psychology investigating a person's emotional state affects the state of the person. In social sciences "revealing" one's preferences in a choice can affect those preferences: "*There is a growing body of evidence that supports an alternative conception according to which preferences are often constructed – not merely revealed – in the elicitation process. These constructions are contingent on the framing of the problem, the method of elicitation, and the context of the choice". [?] p.525.*

[9] Consuming drugs today makes you more willing to consume tomorrow and you may end up as a drug addict. Knowing that, a rational agent may refrain from an even small and pleasant consumption today in order not to be trapped in addiction.

self-esteem, anticipatory utility, self control.

Closely related to this paper is Benabou Tirole as well as articles by Fudenberg and Levine (2006). They develop a dual self model of self-control that can explain a large variety of behavioral paradoxes. In their model there is a long-term benevolent patient self and a multiplicity of impulsive short-term selves - one per period. This particular structure allows them to write the game as a decision problem. In contrast, we are dealing with a full-fledged game involving a multiplicity of simultaneous (symmetric) selves in each period. All selves are equally rational and care about the future expected utility of the individual. The dual self model is designed to capture the management of impatience and it has a strong predictive power. Interestingly, both the dual self model and the TI-model can show that (apparent) dynamic inconsistency may arise as a result of rational self-control. We trust that the quantum approach has the potential to capture self-management issues reflecting a wide range of conflicting interests within the individual. We aim at testing its predictive power along a variety of issues in future research.

2 Motivating Puzzles

The idea that an individual's choice of action (behavior) determines her inner characteristics (preference, attitudes and beliefs) rather than (exclusively) the other way around has been present in people's mind throughout history and has been addressed in philosophy, psychology as well as more recently in economics.

Nevertheless the dominating view in particular in economics, is based on a postulate: individuals are endowed with an identity (preferences, attitudes and beliefs) that explain their behavior. This postulate is hard to reconcile with a host of experimental evidence including at the more basic level of perception. It has long been known that the perception of pain is only partly a function of the pain stimulus. Zimbardo *et al.* in [48] demonstrated that individuals who had volunteered to continue participating in an experiment using painful electric shocks, reported the shocks to be less painful and were physiologically less responsive than individuals who were given no choice about continuing. Valins and Ray [46] conducted an experiment where snake-phobic subjects were presented pictures of snakes and were falsely reported that their heart beat was calm. Subsequently they exhibited significantly reduced fear for snakes. In another experiment, subject were cued to identify the same physiological arousal as either anger or euphoria [44]. Cognitive dissonance experiments in e.g., the Carlsmith and Festinger classical experiment [21] also show how behavior affects attitudes. For a systematic review of experimental evidences (see [4]). All these evidences led Weick [47] to propose that: "Attribution and attitudes may follow behavior and not precede it". Similarly Berkowitz [8] remarks " We generally assume as a matter of course that the human being acts because of the wants arising from his understanding of the environment. In some cases, the understanding may develop *after* stimuli have evoked the action so that the understanding justifies but does not cause the behavior" (p. 308).

Psychologists developed several theories to account for these experimental facts. According to the famous James-Lang theory of emotion, when an emotional event occurs, our behavioral reaction determines our subjective experience of the event (see [32]). Closely related is self-perception theory. As expressed by Bem [4] self-perception theory is based on two postulates: 1. *"individuals come to 'know' their own attitude and other internal states partially by inferring them from observations of their own behavior and/or the circumstances in which behavior occurs*; 2. *Thus the individual is functionally in the same position as an outside observer, an observer who must necessarily rely upon those same external cues to infer the individual inner state."* (p. 2 in [4]). Self-perception theory does not clearly give up the classical postulate. Nevertheless its own postulates are fully consistent with the hypothesis of (non-classical) indeterminacy which overturns the classical postulate of pre-existing identity, attitudes and preferences. With indeterminacy of the inner state, behavior (the action chosen in a decision situation, see below) shapes the state of preferences/attitude by force of a state transition process (see next section). Indeterminacy means intrinsic uncertainty about individual identity so the individual may not know his *own* attitudes, preferences and beliefs. And as in self-perception theory, it is by observing his own action that he infers (learns) his state (of beliefs and preferences). While self-perception theory emphasizes the similitude between outside observation and self observation, quantum decision theory puts emphasis on the fact that observation is structured. As recognized in self-perception theory inner states are not accessible without some training and instrument to measure them. To observe one needs an "appropriate descriptor" ([4] p. 3). Such "descriptor" includes "cues" that can be manipulated to obtain widely different perceptions cf anger *versus* euphoria above. This is consistent with the most basic feature of indeterminacy namely that the property of a system does not pre-exist observation. Therefore different measurement instruments may give various incompatible but equally true accounts of the same state. As we shall see this is also at the heart of the state transition process and delivers our theory of self-control.

3 The Model

We shall describe the dynamic decision problem as a simple separable dynamic game between the selves of an individual with the state variable identified as the identity(type) of the individual. The equilibrium concept we shall be using is that of Markov Perfect Equilibrium.

The kind of situations we have in mind is a sequence of (at least) two consecutive decision situations (DS). An example is as follows. Bob has just inherited some money from his aunt and the first decision is between buying state obligations or risky assets. The second decision situation is between a stay at home evening or taking his wife to a party. The two situations appeal to different but related (in a sense to be made clear below) type characteristics: the first *DS* appeals to his preference toward risk: cautious (θ_1) risk loving (θ_2). The second decision situation appeals to his attitude toward others: (τ_1) egoistic *versus* generous/empathetic (τ_2). An alternative story that brings us closer to Tirole and Benabou and to the literature on self-control, is to define DS1 as a choice between exercising or sleeping late. DS2 is a choice between watching a good movie and helping your mother with practical things. The idea is that both *DS* involve a choice between tempting (immediate) gratification and more sophisticated satisfaction.

We next develop the general theory and illustrate it in the above mentioned example.

3.1 The players

In each period the individual faces a Decision Situation (*DS*) A^t corresponding to the finite set of available actions in period t. We restrict the one-period players' strategy set to pure actions. The possible preferences over the profiles of actions (one action for each self) are denoted by $e_{M,i}$ where M defines the complete measurement corresponding to A^t—see below. Consider $A^t = \{a_1, a_2\}$ and assume that there exists (only) three possible preferences: prefer the action chosen by the other selves (in case the two others play different actions, randomize) or a strict preference for a_1 (a_2) irrespective of others' choice. $M \in \mathcal{M}$, where $\mathcal{M}$ is the set of all complete measurements, corresponds to an elicitation procedure that fully reveals the preferences in A_t. A choice in *DS* A^t is a coarse measurement e.g., in our example the choice of a_1 by a self does not allows to distinguish between the pooling type and the a_1 dominant action type if at least one of the other selves also choose a_1. We refer to the $e_{M,i}$ as the selves or the "eigentypes" of M.[10] They are the players of our game.

In each period t the individual is represented by his state or type (we use the terms interchangeably), a vector $|s^t\rangle \in \mathbb{S}$, where $\mathbb{S}$ is a (finite) $n-$dimensional Hilbert space and the bracket $|.\rangle$ denotes a (ket) vector in Dirac's notation which is standard when dealing with indeterminacy.[11] The eigentypes $e_{M,i}$ of M are associated with the eigenvectors $|e_{M,i}\rangle$ of the operator which form a basis of the state space. The state vector can therefore be expressed as a superposition[12]: $|s^t\rangle = \sum_{i=1}^n \lambda_i^t |e_{M,i}\rangle$, $\lambda_i \in \mathbb{R}$, $\sum_i (\lambda_i^t)^2 = 1$, where $e_{M,i}$ are the (potentials) selves relevant to *DS* A_t. This formulation means that the individual cannot generally be identified with a single true self. He does not have a single true preference, instead he is intrinsically "conflicted" which is expressed by the multiplicity of the potential selves.[13] The coefficients λ_i, also called amplitude of probability, provides a measure

[10] An eigentype corresponds to an eigenvalue of the operator.

[11] The mathematical concept of a Hilbert space generalizes the notion of Euclidean space. Its significance was underlined with the realization that it offers one of the best mathematical formulations of quantum mechanics. In short, the states of a quantum mechanical system are vectors in a certain Hilbert space, the observables are hermitian operators on that space, and measurements are orthogonal projections

[12] A superposition is simply a linear combination such that the square of the coefficients sum up to 1.

[13] In the human mind conflicting propensities to act co-exist until a choice is made. For instance a person may be hesistating between two deserts :a chocalate cake or a frozen yogurt. With the actual choice of the frozen yogurt, she

of the relative strength of potential self $e_{M,i}$, more precisely the square of the coefficient gives the probability that self $e_{M,i}$ will determine the behavior of the individual in *DS* A_t. As a special case we have $|s^t\rangle = \sum_i \lambda_i^t |e_{M,i}\rangle$ with $\lambda_i = 0, i = 2, \ldots, n$ implying $|s^t\rangle = |e_{M,1}\rangle$. In that special case the individual is, at time t, identified with self $e_{M,1}$.[14] As we shall see below, if the individual preferences are fully determined in *DS* A^t, they are by necessity indeterminate in any non compatible *DS* A^{t+1}.

We assume throughout the paper that there is common knowledge among the players (selves) about the current state, the utility function of all players and about the state transition process (see below). As argued in the introduction, type indeterminacy implies bounded rationality *at the individual level* in the sense that individual behavior is not consistent with a well-defined complete ordering over the universal set of alternatives. As we shall see in the next section, indeterminacy also means that preferences are unstable. However, we assume that the individual is aware of his own indeterminacy and act consistently within the corresponding cognitive limitations. This hypothesis is captured in an assumption of rationality (in a way to be defined below) and common knowledge of rationality at the level of the selves. We do not pretend that we should always expect such a degree of sophistication from the side of the selves but our approach is a natural first step

3.2 Indeterminacy: Decision-making as a state transition process

In each period, the selves form intentions to play and eventually *one* action is taken by the individual. Decision-making is modelled as the measurement of the preferences (cf. the revelation principle) and it is associated with a transition process from the initial state and (intended) actions to a new state. The rules that govern the state transition process reflect the intrinsic indeterminacy of the individual's type or preferences. It features the minimal perturbation principle that defines a measurement operation which is formalized by the von Neumann projection postulate[15]: if the initial state is $|s\rangle$ and the chosen action is a_1 then the new state is the normalized projection of $|s\rangle$ onto the eigenspace belonging to a_1.[16]

Formally, a transition process is a function from the initial state and (intended) actions to a new state. It can be decomposed into an outcome mapping $\mu_A : \mathbb{S} \to \Delta A$ where ΔA is the unit simplex of actions and a transition mapping $\tau_{M,a} : \mathbb{S} \to \mathbb{S}$. The first mapping defines the probability for the possible choices of action when an individual in state $|s\rangle$ is confronted with *DS* A. The second mapping $\tau_{M,a}$ indicates where the state transits as we confront the individual with *DS* A and obtain outcome a.

Let the initial state be $|s^t\rangle = \sum_i \lambda_i^t |e_{M,i}\rangle$. The standard Hilbert space formulation yields that if

becomes a person capable of resisting the temptation to eat a chocolate cake. This identity can be upset by later choices however.

[14] In that special case the individual's preferences in A_t preexist the measurement *i.e.*, the act of choice.

[15] Or the postulate's more stringent version defined by Luder.

[16] We talked about "eigenspace" associated with an eigenvalue "a" of a measurement operator if the eigenvalue is degenerated *i.e.*, if several linearly independent vectors yield the same outcome of the measurement.

we, for instance, observe action a_1, the state transits onto

$$\left|s^{t+1}\right\rangle = \sum_{j=1} \lambda'_j \left|e_{M,,j}\right\rangle \tag{1}$$

where $\lambda'_j = \frac{\lambda^t_j}{\sqrt{\sum_{k^t}\left(\lambda^t_k\right)^2\left(s^*_k=a_j\right)}}$ and $\sum_{k^t}\left(\lambda^t_k\right)^2\left(s^*_k=a_j\right)$ is the sum over the probabilities for the selves who pool in choosing a_j. This is of course equivalent to Bayesian updating *i.e.*, the state transition seems purely informational. The value of this more general formulation comes when dealing with a sequence of non-commuting DS. To see that the formal equivalence breaks down, we have to express $\left|s^{t+1}\right\rangle$ in terms of $\left|e_{N,i}\right\rangle$ where N is the new (non-commuting) measurement in period $t+1$ corresponding to DS A^{t+1} and $\left|e_{N,i}\right\rangle$ are its eigenvectors. The eigenvectors of N also form a (alternative) basis of the state space. And this is where the earlier mentioned correlations between selves from different periods enter into play. The correlations link the two sets of basis vectors: the eigenvectors of M can be written as linear combinations of the eigenvectors of N with the correlations as the coefficients of superposition—see below.

These correlations captures the extent of overlap between the states.[17] In a classical world *all* distinct *atomic* states are orthogonal. So in a classical world either type characteristics are mutually exclusive: so for instance Bob is either of the risk-loving type or of the cautious type, or the types can be combined: Bob can be of the risk-loving type *and* of the egoistic type. But in the later case, the type characteristics "risk-loving" is not be a complete characterization *i.e.*, not an atomic state. The novelty with indeterminacy is that type characteristics can overlap in the sense that they are non-orthogonal atomic states. For instance, in our example the risk-loving type and the cautious type are orthogonal but the risk-loving type and the egoistic type are not. Nevertheless, the three are complete descriptions of the individual *i.e.*, they are atomic states. The risk-loving type overlaps with the egoistic type. This means that if Bob is of the risk-loving type, there is some probability that in his second choice he will reveal egoistic preferences (with the complementary probability he reveals generous preferences) and his type will be modified, he will no longer be of the risk-loving type. Instead, he will be fully characterized as an egoistic type. And, if tested again with respect to cautious/risk-loving characteristics, the state will transit again and he may end up as a cautious type. The correlations are a measure of this overlap.

Let B_{MN} denote the basis transformation matrix that links the two non-compatible type characteristics M and N: $\left|e_{M,i}\right\rangle = \sum_j \gamma_{ij}\left|e_{N,j}\right\rangle$ where γ_{ij} are the elements of the basis transformation matrix

[17] In their remarkable book on quantum logic [3], Beltrametti and Cassinelli write "In physics, the expression transition probability generally refers to dynamical instability. Our use of the term is not directly related to instability rather we follow von Neuman's terminology. The transition probability between two states is meant to represent intuitively a measure of their overlapping. Actual transition from one state to another is triggered by a measurement."

$\gamma_{ij} = \langle e_{N,j} | e_{M,i} \rangle$.[18] Substituting into Equation (1) and collecting the terms we write

$$\left|s^{t+1}\right\rangle = \sum_j \left(\sum_i \lambda_i' \gamma_{ij} \right) \left|e_{N,i}\right\rangle = \sum_i \eta_i^{t+1} \left|e_{N,i}\right\rangle$$

According to Bohr's rule the probability for eigentype $\left|e_{N,1}\right\rangle$ (if the agent is confronted with *DS* A^{t+1} that (coarsely) measures type characteristics N) is

$$TP : p\left(e_{N,1} \middle| \left|s^{t+1}\right\rangle\right) = \left(\sum_i \lambda_i' \gamma_{1i} \right)^2 \tag{2}$$

This is a crucial formula that captures the key distinction between the classical and the type indeterminacy approach. TP is *not* a conditional probability formula where the γ_{ij}^2 are statistical correlations between the eigentypes at the two stages. The probabilities for the N-eigentypes depend on the M-eigentypes' *play* in *DS* A^t. When no player chooses the same action, the choice of a_i^t separates out a single player (some $e_{M,i}$), the sum in parenthesis involves *one* term only. While when several players pool in choosing the same action, the term in parenthesis involves several terms. As a consequence, the probabilities for the different players are given by the square of a sum, implying cross terms called interference effects—and not the sum of squares (as we would have in a classical setting). Since the amplitudes of probability can be negative numbers, the interference effect may be negative or positive.

We note that the state transition process is deterministic by the, earlier mentioned, von Neumann's postulate which says that under the impact of a measurement a pure state transit into another pure state. In this paper we are only dealing with pure types. If we observe a_1^t (as the result of applying A_t) the state

$$\left|s^t\right\rangle = \sum_i \lambda_i^t \left|e_{M,i}\right\rangle \text{ transits onto } \left|s^{t+1}\right\rangle = \sum_{j=1} \lambda_j' \left|e_{M,,j}\right\rangle = \sum_i \eta_i^{t+1} \left|e_{N,i}\right\rangle$$

that is $\left|s^{t+1}\right\rangle$ is a pure state. Yet, predictions on the outcome of (applying) A^{t+1} are probabilistic because of indeterminacy *i.e.*, $\left|s^{t+1}\right\rangle$ is a superposed state.

3.3 Utility

When dealing with multiple selves, the question as to how to relate the utility of the selves (here the players) to that of the individual has no self-given answer.[19] We adopt the following definition of the utility of self (or player) $e_{M,i}$ of playing of a_i^t when the $-i$ other $t-$period players play $\mathbf{a}_{-i}^t$

$$U_{e_{M,i}}(a_i^t) + \delta_{e_{M,i}} \sum_{i=t}^{T} EU(s^{i+1}\left(a_i^t, \mathbf{a}_{-i}^t; s^t \middle| a^t = a_i^t\right)) \tag{3}$$

[18] $\left\langle e_{N,,j} \middle| e_{M,i} \right\rangle$ is a scalar product.

[19] One reason is that while the selves are incarnations of the same individual, they are short-lived. Another is that they might not recognize the "legitimacy" of some future possible incarnations. For instance a current compassionate self may not value the utility of a future spiteful incarnation.

where a^t denotes the actual play of the individual.

The utility for $e_{M,i}$ of playing a_i^t is made of two terms. The first term is the utility in the current period evaluated by player $e_{M,i}$. This term only depends on the action chosen by $e_{M,i}$. The second term is the expected utility of the individual evaluated by the future selves conditional on $a^t = a_i^t$. The second term depends indirectly on the whole profile of (intended) actions in the current period through the state transition process $s^{t+1}(\mathbf{a}_i^t; s^t)$. The next period expected utility is $EU(s^{t+1}(\mathbf{a}^t; s^t)) = \sum_i \eta_i(a_i^t | \mathbf{a}^t) U(s^{i+1}(a_i^t; s^t))$, where $\eta_i(a_i^t | \mathbf{a}^t)$ are the coefficient of superposition relevant to the next period DS. It is the weighted sum of the utility of all the possible resulting types following $\mathbf{a}^t$, where the weights are given after updating and expressing the new state according to TP. The possible resulting states $s^{t+1}(a_i^t; s^t)$ are expressed in terms of the eigentypes relevant to A^{t+1} and the expected utility of period $t+1$ is calculated given the optimal choice of eigentypes relevant in period $t+1$, e.g., $e_{N,j}$.

The current action profile only influences tomorrow's state, the summation term in Equation (3) can therefore be collapsed into a single term $EU^T(s^{t+1}(\mathbf{a}^t; s^t)) = \sum_{i=t}^T EU^*(s^{i+1}(\mathbf{a}^t; s^t))$. Which is the expected utility when all future selves in all periods play an equilibrium pure strategy.[20]

Utility thus writes

$$U_{e_{M,i}}(a_i^t; s^t) + \delta_{e_{M,i}} EU^T(s^{t+1}(\mathbf{a}^t; s^t | a^t = a_i^t)) \tag{4}$$

in each period the payoff relevant history of play is captured by the state variable representing the current state or identity.

The utility function may remind of a Bernoulli function in the following sense. With some probability the self survives (his preferred action is played by the individual) and with the complementary probability, he is "out of the game". The formulation in Equation (4) means that he maximizes utility conditional on surviving. The probability for survival depends on the initial coefficients of superposition as well as on his own and other selves's choice. But the selves do not take that into account. The approach is justified on the following ground: being "out of the game" cannot be valued. The self seizes to "exist" which is neither good or bad. To put it differently there is no reason to assume that selves have a "survival instinct", they are simply mental constructions. A self is defined as rational when he maximizes his conditional utility which is well-defined for any sequence of DS.

3.4 The Equilibrium

In each period, the current selves move simultaneously. They know the current state resulting from the previous (actual and intended) play. We have common knowledge among the selves about the payoff functions of all selves current and future and common knowledge of rationality. The selves'

[20] For the case when there exist mutiple equilibria, we assume that the current selves share the same beliefs about which equilibrium is played.

payoffs are functions of the current actions and the current state as defined in the previous section.[21] Together this means that we are dealing with a separable dynamic game of complete information and that it seems most appropriate to restrict ourselves to Markov strategies: a strategy for *a self* is a function $\mathbb{S} \to A^t$ from the current state to the set of actions available at period t. We shall accordingly focus on Markov Perfect Equilibria.

Definition 1 *A Markov Perfect Equilibrium of the game is characterized by a_i^{t*} :*

$$a_i^{t*} = \arg\max_{a_i \in A^t} U_{e_{M,i}}(a_i^t; s^t) + \delta_{e_{M,i}} \sum_{\tau=t+1}^{T} EU^*(s^\tau\left(\mathbf{a}^t; s^t\right))$$

in all periods $t = 1, \ldots, T$ and for all $e_{M,i},\ M \in \mathcal{M},\ i = 1, \ldots, n.$

So we see that a self "only" needs to worry about his current utility and the expected utility value of his action via the resulting type. The equilibrium is found by backward induction in a standard way.[22] The novelty lies in the technology for the state transition process which captures indeterminacy. So in particular the state variable are the preferences themselves and they evolve in a non-monotonic way reflecting the dynamics of measurement operations and the correlations between non-commuting *DS*.

Remark 1 *For the case all DS commute with each other, the model is the one of an individual who does not initially know his preferences and learns through Bayesian updating as he observes the actions he takes.*

If *all DS* commute, the state variable evolves through Bayesian updating. The individual eventually learns who he is and behaves as a classical decision-maker who maximizes discounted expected utility. In the TI-model, the concern for identity arises exclusively as a consequence of the non-commutativity of successive *DS*. The general case is a one where some *DS* commute and some do not. We below focus on non-commuting *DS* which allows us to address the issue of identity management in each period. But we should keep in mind that the kind of preference instability that we describe in the next section does not apply within a sequence of commuting *DS*.

Definition 2 *We say that the MPE is characterized by self-control when contemporaneous selves with conflicting short-run preferences pool to select the same action.*

When the action set is sufficiently rich to fully sort out the preferences, all selves have "conflicting" preferences with respect to the short-run choice. If the MPE is characterized by pooling, some selves

[21] Maskin and Tirole [40] develop a general approach to Markov Perfect Equilibrium where strategies may depend on history in a more elaborate way. A distinguishing feature of a quantum state is that a measurement erases information about the previous state. Expressed differently all relevant information for predicting the outcome of any measuremnt is contained in the current state. The history of the state has no relevance. This feature invites the restriction to simple Markov strategies.

[22] Although we know that a MPE exists in mixed strategies (*cf* theorem 13.1 in [26]), we have no proof of existence for the case we restrict ourselves to pure strategies as we do here.

must be exercising self restraint, refraining from immediate reward for the sake of the individual future utility—this is an instance of self-control. In a standard *DS,* the set of actions is limited relative to the possible preferences (because a *DS* is generally a coarse measurement), we talk about self-control only when selves with short-run conflicting interest with respect to the *DS* pool. In the next section we consider an example where the set of action is sufficiently rich relative to preferences.

Remark 2 *For the special case with $\delta_{e_{M,i}} = 0$ for all selves in all periods, we are back in the basic TI-model. There is no self-control. For $\delta_{e_{M,i}} \neq 0$ for some selves in some periods, the equilibrium path of action may exhibit some extent of self-control. The model suggests a classification of individual traits and behavior as we show below.*

For the case the selves are short-sighted (or unaware of the impact of action on the future *i.e.*, the case of unsophisticated selves), we are back in the simple decision-making model formulated in [36]. It has been used to explain behavioral anomalies in decision theory from cognitive dissonance to framing effect and preference reversal.

The case with $\delta_{e_{M,i}} = 1$, for all selves in all periods, is interesting because a classical agent would not face any self-control problem. In contrast, for a type indeterminate decision-maker, the issue of self management arises because today's intended actions affect future identity—see example below. For the case, all but one self in each period are short-sighted ($\delta_{e_{M,i}} = 0$) we have a model with the dual structure reminding of Fudenberg and Levine but the means of controlling the behavior are very different (see Discussion section below). We leave to future research the investigation of a type indeterminate individual with such a dual structure.

4 An Illustrative Example

Bob faces two consecutive non commuting decisions: *DS1* with action set $\{a_1, a_2\}$ and *DS2*: $\{x_1, x_2\}$. The story is as follows. Bob just inherited some money from his aunt. The first decision situation involves a choice between buying state obligations (a_1) and risky assets (a_2). The second *DS* is a choice between a stay-at-home evening (x_1) and taking his wife to a party (x_2). The relevant type characteristics to *DS1* have two values (eigentypes): cautious (θ_1) risk loving (θ_2). In *DS2* the type characteristics has two values as well: (τ_1) egoistic *versus* generous/empathetic (τ_2). We have on purpose selected two *DS* that are independent from each other in the sense that there is no complementarity or substitutability between the choices in *DS1* and *DS2*.

We below define the utility associated to the different choices. An assumption that we make is that τ_2 experiences a high utility from x_2 while τ_1 experiences a low utility whatever he chooses. To put it differently, it is better for Bob to be of the τ_2 type. We next provide the classical representation of the decision problem.

4.1 Classical optimization

Let us first characterize the set of types. Since both type characteristics each have two values, Bob may be any of the following four types $\{\theta_1\tau_1, \theta_1\tau_2, \theta_2\tau_1, \theta_2\tau_2\}$.

The utility is described by table 1 and 2 below

Tab. 1

	a_1	a_2
Tab. 1	$U(a_1;\theta_1\tau_1) = U(a_1;\theta_1\tau_2) = 4$	$U(a_2;\theta_1\tau_1) = U(a_2;\theta_1\tau_2) = 2$
	$U(a_1;\theta_2\tau_1) = U(a_1;\theta_2\tau_2) = 2$	$U(a_2;\theta_2\tau_1) = U(a_2;\theta_2\tau_2) = 3$

,

so only the θ value matters for the $a-$choice.

	x_1	x_2
Tab.2	$U(x_1;\theta_1\tau_1) = U(x_1;\theta_2\tau_1) = 2$	$U(x_2;\theta_1\tau_1) = U(x_2;\theta_2\tau_1) = 0$
	$U(x_1;\theta_1\tau_2) = U(x_1;\theta_2\tau_2) = 1$	$U(x_2;\theta_1\tau_2) = U(x_2;\theta_2\tau_2) = 8$

so here only the τ value matters for the $x-$choice.

The tables above give us immediately the optimal choices:

$\theta_1\tau_1 \to (a_1, x_1)$	$\theta_2\tau_1 \to (a_2, x_1)$
$\theta_1\tau_2 \to (a_1, x_2)$	$\theta_2\tau_2 \to (a_2, x_2)$

Using the values in table 1 and 2, we note that type $\theta_1\tau_2$ achieves the highest total utility of 12. the lowest utility is achieved by $\theta_2\tau_1$.[23] While Bob knows his type, we do not. We know that "the population of Bobs" is characterized by the following distribution of types:

$\theta_1\tau_1 \to 0.15$	$\theta_2\tau_1 \to 0.35$
$\theta_1\tau_2 \to 0.35$	$\theta_2\tau_2 \to 0.15$

.

We note that the distribution of types in the population of Bobs exhibit a statistical correlation between the θ and τ type characteristics.

4.2 A TI-Model of Dynamic Optimization

By definition the type characteristics relevant to the first DS1 is $\theta, \theta \in: \{\theta_1, \theta_2\}$. Subjecting Bob to the $a-$choice is a measurement of his θ characteristics. The outcome of the measurement maybe θ_1 or θ_2 and Bobs collapses on an eigentype or the outcome may be null (when both θ_1 and θ_2 choose

[23] Note that we here assume that we can compare the utility of the different types of Bob. This goes beyond standard assumption in economics that preclude inter personal utility comparisons. But is in line with inter personal comparisons made in the context of social choice theory.

the same action).[24] The type characteristics relevant to DS2 is $\tau, \tau \in \{\tau_1, \tau_2\}$. Since the two DS do not commute we can write

$$|\theta_1\rangle = \alpha_1 |\tau_1\rangle + \alpha_2 |\tau_2\rangle$$
$$|\theta_2\rangle = \beta_1 |\tau_1\rangle + \beta_2 |\tau_2\rangle$$

where $\alpha_1^2 + \alpha_2^2 = 1 = \beta_1^2 + \beta_2^2$. For the sake of comparison between the two models we let $\alpha_1 = \beta_2 = \sqrt{.3}$ and $\alpha_2 = \beta_1 = \sqrt{.7}$. Bob's initial type or state is

$$|t\rangle = \lambda_1 |\theta_1\rangle + \lambda_2 |\theta_2\rangle, \ \lambda_1^2 + \lambda_2^2 = 1$$

with $\lambda_1 = \lambda_2 = \sqrt{.5}$.

When discussing utility in a TI-model one should always be careful. This is because in contrast with the classical model, there is not one single "true type" who evaluates the utility value of all choice options. A key assumption is (as in TI-game see Lambert-Mogiliansky 2010) that *all the reasoning of the agent is made at the level of the eigentype* who knows his preferences (type), has full knowledge of the structure of the decision problem and cares about the expected payoff of Bob's future incarnations (type). The utility value for the current decision is evaluated by the eigentype who is reasoning. So for instance when Bob is of type t, two reasonings take place. One performed by the θ_1 eigentype and one performed by θ_1 eigentype. The $\theta-$types evaluate the second decision, using the utility of the type resulting from the first decision. The utility of a superposed type is the weighted average of the utility of the eigentypes where the weights are taken to be the square of the coefficient of superposition.[25] The utility of the eigentypes are depicted in the table 3 and 4 below

Tab. 3

$U(a_1;\theta_1) = 4$	$U(a_2;\theta_1) = 2$
$U(a_1;\theta_2) = 2$	$U(a_2;\theta_2) = 3$

, and Tab. 4

$U(x_1;\tau_1) = 2$	$U(x_2;\tau_1) = 0$
$U(x_1;\tau_2) = 1$	$U(x_2;\tau_2) = 8$

.

As earlier noted Bob in state t performs two (parallel) reasonings. We proceed by backward induction to note that trivially since the "world ends after DS2", τ_1 chooses x_1 and τ_2 chooses x_2 (as in the classical model). We also note that: $U(x_1;\tau_1) = 1 < U(x_2;\tau_2) = 8$. The τ_2 incarnation of Bob always experiences higher utility than τ_1.

The TI-model has the structure of a two-stage maximal information[26] TI-game as follows. The set of players is $N : \{\theta_1, \theta_2, \tau_1, \tau_2\}$, the θ_i have action set $\{a_1, a_2\}$ they play at stage 1. At stage

[24] More correctly when both our eigentypes choose the same action in DS1, DS1 is a null measuremnt i.e., it does not allows to distinguish between the eigentypes.

[25] We note that in the TI-model we cannot escape inter type utility comparison. We must aggregate the utilies over different selves to compute the optimal decisions. However just as in social choice theory there is no unique way of aggregating individual utility into a social value. We return this issue in the discussion.

[26] Maximal information TI-game are the non-classical counter-part of classical complete information games. But in a context of indeterminacy, it is not equivalent to complete information because there is an irreducible uncertainty. It is impossible to know all the type characteristics with certainty.

2, it is the τ_i players' turn, they have action set $\{x_1, x_2\}$. There is an initial state $|t\rangle = \lambda_1 |\theta_1\rangle + \lambda_2 |\theta_2\rangle$, $\lambda_1^2 + \lambda_2^2 = 1$ and correlation between players at different stages: $|\theta_1\rangle = \alpha_1 |\tau_1\rangle + \alpha_2 |\tau_2\rangle$ and $|\theta_2\rangle = \beta_1 |\tau_1\rangle + \beta_2 |\tau_2\rangle$. The utility of the players is as described in tables 3 and 4 when accounting for the players' concern about future selves. So for a $\theta-$player, the utility is calculated as the utility from the choice in DS1 plus the expected utility from the choice in DS2 where expectations are determined by the choice in DS1 as we shall see below.

The question is how will Bob choose in DS1, or how do his different $\theta-$eigentype or selves choose? We here need to do some simple equilibrium reasoning.[27] Fix the strategy of pure type θ_1, say he chooses "a_1".[28] What is optimal for θ_2 to choose? If he chooses "a_2" the resulting type after DS1 is $|\theta_2\rangle$. The utility, in the first period, associated with the choice of "a_2" is $u(a;\theta_2) = 3$. In the second period Bob's type is $|\theta_2\rangle = \beta_1 |\tau_1\rangle + \beta_2 |\tau_2\rangle$ which, given what we know about the optimal choice of τ_1 and τ_2, yields an expected utility of $\beta_1^2 [U(x_1;\tau_1) = 1] + \beta_2^2 [U(x_2;\tau_2) = 8] = .7 + 8(.3) = 3.1$. The total (for both periods) expected utility from playing "a_2" for θ_2 is

$$EU(a_2;\theta_2) = 3 + 3.1 = 6.1$$

This should be compared with the utility, for θ_2, of playing "a_1" in which case he pools with θ_1 so the resulting type in the first period is the same as the initial type i.e., $|t\rangle = \lambda_1 |\theta_1\rangle + \lambda_2 |\theta_2\rangle$. The expected utility of playing a_1 is $u(a_1;\theta_2) = 2$ in the first period plus the expected utility of the second period. To calculate the latter, we first express the type vector $|t\rangle$ in terms of $|\tau_i\rangle$ eigenvectors:

$$|t\rangle = \lambda_1 (\alpha_1 |\tau_1\rangle + \alpha_2 |\tau_2\rangle) + \lambda_2 (\beta_1 |\tau_1\rangle + \beta_2 |\tau_2\rangle) = (\lambda_1\alpha_1 + \lambda_1\beta_1) |\tau_1\rangle + (\lambda_1\alpha_2 + \lambda_2\beta_2) |\tau_2\rangle.$$

The second period's expected utility is calculated taking the optimal choice of τ_1 and τ_2:

$$\left(\lambda_1^2\alpha_1^2 + \lambda_2^2\beta_1^2 + 2\lambda_1\alpha_1\lambda_2\beta_1\right) 1 + \left(\lambda_1^2\alpha_2^2 + \lambda_2^2\beta_2^2 + 2\lambda_1\alpha_2\lambda_2\beta_2\right) 8 = 0.959 + 7.669 = 8,63.$$

which yields

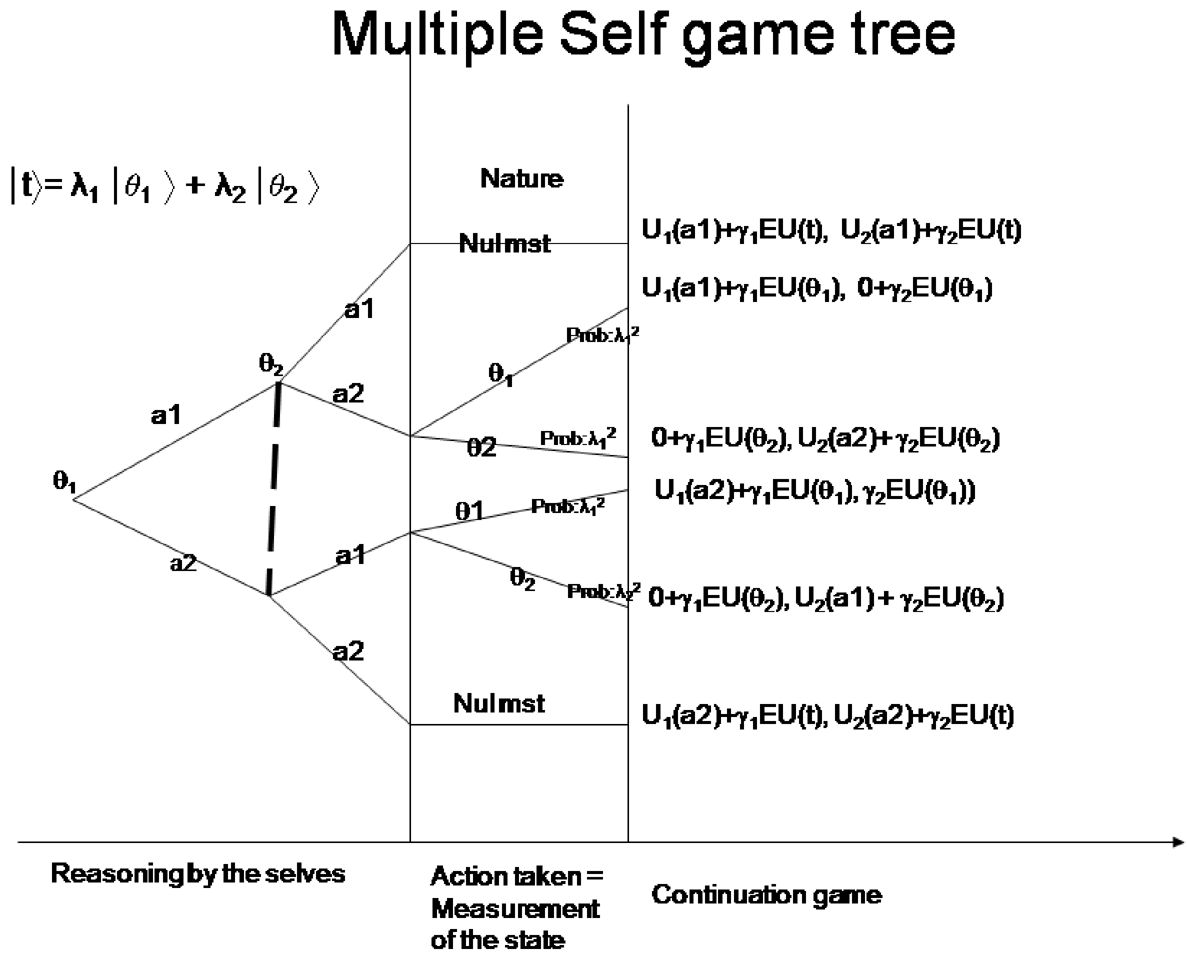
$$EU(a_1;\theta_2) = 2 + 8,63 = 10,63 > EU(a_2;\theta_2) = 3 + 3.1 = 6.1$$

So we see that there is a gain for θ_2 of preserving the superposition i.e., it is optimal for pure type θ_2 to forego a unit of utility in DS1 and play a_1 (instead of a_2 as in the classical model). It can also be verified that given the play of θ_2 it is indeed optimal for θ_1 to choose a_1. The solution to dynamic

[27] Under equilibrium reasoning, an eigentype is viewed as a full valued player. He makes assumption about other eigentypes' play at difference stages and calculate his best reply to the assumed play. Note that no decision is actually made so no collapse actually takes place. When he finds out what is optimal for him, he checks whether the assumed play of others is actually optimal for them given his best response. We have an equilibrium when all the eigentypes are best responding to each others.

[28] We note that the assumption of "a_1" is not fully arbitrary since a_1 gives a higher utility to θ_1 than a_2. However, we could just as well have investigated the best reply of θ_1 after fixing (making assumption) the choice of θ_2 to a_2. See further below and note 12 for a justification of our choice.

optimization is an "inner" Bayes-Nash equilibrium where both θ_1 and θ_2 to play a_1.[29]



The interpretation is that Bob's θ_2 type understands that buying risky assets appeals to his risk-loving self which makes him tense. He knows that when he is tense, his egoistic self tends to take over. So, in particular, in the evening he is very unlikely to feel the desire of pleasing his wife - his thoughts are simply somewhere else. But Bob also knows that when he is in the empathetic mood i.e., when he enjoys pleasing his wife and he does it, he always experiences deep happiness. So his risk-loving self may be willing to forego the thrill of doing a risky business in order to increase the chance for achieving a higher overall utility.

4.3 Generic Classes of Behavior

The 2 types, two actions and two periods case allows illustrating some basic comparative statics results. The TI-model invites us to distinguish between two situations characterized by the sign of the

[29] The equilibrium need not be unique. A similar reasoning could be made for both $\theta-$type pooling on a_2. The inner game is a coordination game. It make sense to assume that coordination is indeed achieved since all the reasoning occurs in one single person.

interference effect applying to the high utility option. Interference effects are the signature of indeterminacy. When the individual is in the superposed state $|s^0\rangle$ both θ eigentypes are simultaneously present in his mind and they interact. The interference effect captures the impact of the interaction between the θ eigentypes in the determination of the probabilities for the outcome in *DS2*. The propensity to be of the type τ_2 who experiences the high second period utility is present in both θ_1 ($\alpha_2 |\tau_2\rangle$) and in θ_2 ($\beta_2 |\tau_2\rangle$). We have positive interference effects when those propensities reinforce each other and increase the chance that the superposed individual will turn out to be of that type. The sign of interferences effect depends on the operators associated with the decisions. These operators and the correlations between them are structural properties of the state space. Our view is that those properties capture neurological and psychological regularities common to all individuals or at least a larger group of individual (e.g., from the same sociocultural background). Whether interference effects are positive or negative is an empirical question.[30] While these are features common to all individuals, each individual is characterized by his state, a vector in a potentially very high dimensional space. Moreover, individuals can differ in the intensity of the utility experienced by their selves and in the value their selves put on the utility of future incarnations. We next investigate how these characteristics together with type indeterminacy gives rise to patterns of behavior or "personality traits".

4.3.1 Self-Control by Inner Agreement

In this section we assume, as in the example, that the interference effect (IE) favors the high utility option x_2 (and automatically reduced the probability for the low utility alternative). As we shall see a positive IE is a factor that favors behavioral (and intertemporal) consistency. Assume that we have

$$EU(\theta_2) < EU(\theta_1) < EU(s^0) \tag{5}$$

where $EU(s)$ is the expected utility in period 2 when the state is s. We recall that the first *DS* is a complete measurement of type characteristics θ[31] so we must have

$$U_{\theta_1}(a_1) > U_{\theta_1}(a_2) \text{ and } U_{\theta_2}(a_1) < U_{\theta_2}(a_2)$$

i.e., the $\theta-$selves have conflicting short-run interests.

When considering a sequence of two non-commuting *DS*, the model distinguishes between two classes of individuals: the balanced individual, an individual whose selves manage to agree on a common choice, and the conflicted individual whose selves make separating choice, more precisely:

Definition 3 *A balanced individual is characterized by a MPE that is a pooling equilibrium. It obtains whenever*

$$U_{\theta_1}(a_1) + \delta_1 EU(\theta_1) \leq U_{\theta_1}(a_2) + \delta_1 EU(s^0) \tag{6}$$

[30] This is also the case in Quantum Mechanics.

[31] This means that when considered in isolation, DS1 separates between the $\theta-$types.

or

$$U_{\theta_2}(a_2) + \delta_2 EU(\theta_2) \leq U_{\theta_2}(a_1) + \delta_2 EU(s^0) \tag{7}$$

or both. Otherwise, the individual is conflicted i.e., her inner equilibrium is characterized by separation.

The Equations (6) and (7) capture the selves' incentives to refrain from choosing their preferred action (exerting self restraint) given that the other self chooses his preferred first period action. When an inequality is falsified it is a dominating strategy for that self to choose his preferred first period action. Since we have a conflict of interest, when neither of them holds, the choices are separating. This means that if invited to choose, the individual promptly incarnates either one or the other self. She shows clear-cut preferences, determination. This also means that the first period action triggers state transition onto one of the eigentypes *i.e.*, identity is modified. As a consequence, behavior will exhibit inconsistency (e.g., preference reversal). So this suggests that individuals who are quite extreme in their judgment and have clear-cut preferences also exhibit behavioral inconsistency. The identity of a conflicted individual jumps from one period to another together with the decision made. Recall that this applies to non-commuting *DS*. So in particular in a sequence *DS1-DS3-DS1* where *DS3* commutes with *DS1*, an individual that we characterize as conflicted, will not exhibit any behavioral instability or inconsistencies.

The balanced individual is characterized by selves who are willing to reach an agreement, they make a pooling choice. This occurs at the expense of one of the selves who chooses to forego his preferred option in period 1. This is an instance of self-control. The balanced individual has no clear-cut preferences. we could say that she retains the freedom to value options from different perspectives. The pooling equilibrium obtains when the inequality in (6) or in (7) or both are true. We have pooling on a_1 for the case, at least, Equation (7) holds. When both inequalities hold, we could have pooling on either action. But since θ_1 is closer (highly correlated) to τ_2 and there is agreement on the advantage of identity τ_2, it is reasonable to expect pooling on a_1. For the case only Equation (6) holds, the MPE yields pooling on a_2. Interestingly, in two of three cases, the pooling MPE yields the "good" action a_1 in period 1. But it is not always the case. We may have that the "good" or more forward-looking self chooses to refrain from his preferred action—when Equation (6) does not hold. This can capture a situation when the individual feels that being too demanding with herself fires back. In the example, if the cautious type insists on being cautious, there is a 50% chance that the individual becomes a risk loving type who will have a high chance to be of the egoistic type which is costly since Bob then will only get a low second period utility.

A pooling MPE, triggers no state transition. If the selves were pooling in all periods, the individual would simply behave as a an individual endowed with stable but stochastic preferences. She does not qualify as behaviorally or dynamically inconsistent.

We would like to emphasize that our model features self-control by *means of identity management.* In this respect we stand closer to Benabou Tirole [7]. In particular, we do not address the question related to taking actions (commitment) to limit future behavior as in Gul and Pesendorfer [27] and Fudenberg and Levine [24]. In the next section we briefly return to this question.

Definition 3 allows us to derive some simple comparative statics. For that purpose we write the inequalities in Definition 3 as follows: $U_{\theta_i}(a_i) - U_{\theta_i}(a_j) \leq \delta_i \left[EU\left(s^0\right) - EU(\theta_i)\right]$.

4.3.2 Negative Interference Effect: Agreeing to Disagree

Consider the case when

$$EU(t) < EU(\theta_2) < EU(\theta_1) \tag{8}$$

which obtains in the example by inverting the signs of β_1 and α_2. The selves' incentives are described unambiguously by the inequalities

$$U_{\theta_1}(a_1) + \delta_1 EU(\theta_1) > U_{\theta_1}(a_2) + \delta_1 EU(t) \tag{9}$$

and

$$U_{\theta_2}(a_2) + \delta_2 EU(\theta_2) > U_{\theta_2}(a_1) + \delta_2 EU(t) \tag{10}$$

This implies that both types prefer separation in *DS1.* We have a case of "agreement to disagree". Identity management concerns do not alter the selves' short-run incentives. In the terms defined in the previous section we say that with respect to *DS* linked by negative interference effects, the individual behaves as a conflicted person whose identity keeps on changing so he exhibits intertemporal and behavioral inconsistencies.

We thus find that identity management concerns cannot promote self-restraint and consistency when IE are negative. In such cases, we may like to consider actions that limit future behavior. But that is outside the scope of the present paper.

5 Discussion

In this section we discuss the relation between the present work and some of the literature in economics. As mentioned in the Introduction, there exists a vast theoretical literature pioneered by Strotz [45] dealing with various type of time inconsistency. A large share of this literature has focused on inconsistency that arises because the individual does not discount the future at a constant rate. A contribution of this paper is to demonstrate that there exists other sources of inner conflict *i.e.*, not related to time preferences. A type indeterminate individual is in each period characterized by a multiplicity of conflicting selves (competing desires). All selves are equally rational and care about the future expected utility of the individual. We formalize the "inner bargaining" formulated by Ainslie [1] as a sequential game and characterize the circumstances when individual behavior exhibits preference instability and intertemporal inconsistency. Our approach also differ from the economic literature on self control in another important respect. We do not consider actions whose primary objective is self-control like putting one's money on an account costly to access in order to limit one's spending opportunities. An important contribution of e.g., Gul and Pesendorfer [27] and Fudenberg and Levine [24] is to provide a rationale for such behavior. The present paper does not address this issue. One reason is that we may expect a commitment decision to commute with decisions related to immediate responses to temptations. Those decisions appeal to separate functions in the brain *i.e.*, higher cognitive functions as opposed to more visceral emotions and desires. If that is the case which can be established in experiments, some of the reasoning of e.g., Fudenberg and Levine [24] would simply carry over. From the perspective of indeterminacy, a central question would instead be related to agenda setting. Indeed, the choice of confronting oneself with a DS about commitment (or avoiding it) is an important one which has consequences for identity.[32]

The type indeterminacy approach brings us close to the work by Benabou and Tirole ([5, 6, 7]). They write "When contemplating choices, they then take into account what kind of a person each alternative would make them and the desirability of those self-views" ([7], p. 806–807). Further they write " Two related forms of behavioral instability are history dependence and non-monotonicity. When a person has been induced to behave prosocially or selfishly, or just provided with signals presumed to be informative about his morality, *his choices in subsequent, unrelated interactions are significantly affected*. Moreover, this reaction *sometimes amplifies* the original manipulation, and is *sometimes in opposition* to it." ([7], p. 810). A common feature to Benabou and Tirole's approach and the TI-model is a that today's' behavior affects tomorrows' identity (or self-image) *i.e.*, effective preferences. With type indeterminacy, individual identity is subject to a state transition process so future identity is a function of past actions. In Benabou and Tirole the basic mechanism is incomplete information about own preferences associated with incomplete recall and incomplete self-control. More precisely, they depart from homo economicus by assuming instead (1) imperfect self-knowledge; (2) imperfect recall; (3) imperfect willpower. With these three imperfections they can derive the value of self-esteem (concern for identity), self-monitoring behavior and reconcile with intertemporal inconsistency in behavior.

As with the postulates of self-perception theory (see Section 2), we argue that the three assumptions in Benabou and Tirole are in many respects equivalent with giving up the classical dogma of a pre-existing (deterministic) individual identity and replacing it by indeterminacy. Indeterminacy implies imperfect knowledge because of intrinsic uncertainty: there does no exist any set of "true preferences" (to be learned). Instead, an individual is represented by a superposition of potential types. Indeterminacy implies imperfect recall because no type is the true type forever. The (preference) state keeps transiting with the action taken so yesterday's correctly inferred information about oneself may simply not be valid tomorrow. Indeterminacy implies "imperfect willpower" because it implies multiple selves both simultaneously (multiplicity of potentials) and dynamically (by force of the non-commutativity of decision situations). Therefore, there are necessarily conflicting desires and issues

[32] It has implications for other decisions that do not commute with the commitment decision.

of self-control and self monitoring. Moreover, in a world of indeterminate agents, actions aimed at shaping one's identity are fully justified from an instrumental point of view (it determines future expected utility). In particular there is no need to add any additional concerns for self-image (as in Benabou and Tirole), or diagnostic utility (as in [10]). The TI-model provides a simple and rigorous setting relying on one single departure from the standard setting.[33] Some of our comparative static results (see Section 4.3) are similar to those in Benabou and Tirole and consistent with a host of empirical data including those mentioned in Section 2. Our contribution is to propose an alternative explanation in terms of a fundamental characteristics of the mind: its intrinsic indeterminacy. Indeed, we find that the postulate of the existence of a true self that we ignore and keep forgetting about, is in not very convincing.[34]

References

[1] Ainslie, G. *The Breakdown of Will*; Cambridge University Press: Cambridge, UK, 2001.

[2] Atmanspacher, H.; Filk, T.; Romer, R. Quantum Zeno features of bistable perception. *Biol. Cybern.* **2004**, *90*, 33–40.

[3] Beltrametti, E.G.; Cassinelli, G. The logic of quantum mechanics. In *Encyclopedia of Mathematics and its Applications*; Rota, G.C., Ed.; Addison-Wesley: Boston, MA, USA, 1981; Volume 15.

[4] Bem D. J. Self Perception theory, *Advances in Experimental Social Psychology*, **1972**, 6, 1-62.

[5] Benabou R., and J. Tirole, Self confidence and personal motivation, *Q. J. of Economics*, **2002**, 117/3, 871-915.

[6] Benabou R., and J. Tirole, Willpower and personal rules" *J. of Pol. Economy* **2004** *112/4*, 848-886.

[7] Benabou R., and J. Tirole, Identity, morals and taboos: beleifs as assets, *The Quarterly Journal of Economics,* **2011**,126, 805-855.

[8] Berkovitz, L., *The motivational status of cognitive consistency theorizing, in Theory of Cognitive consistency: a sourcebook.* Chicago: Rand McNally, **1968**,303-310.

[9] Bitbol M., *Physique Quantique et Sciences Humaines*, Edition CNRS, Paris, 2009.

[10] Bodner, R.; Prelec, D. Self-signaling and diagnostic utility in everyday decision making. *Psychol. Econ. Decis.* **2003**, *1*, 105–126.

[33] The elements of quantum formalism that we use are not complex, they are just novel to most readers.

[34] Moreover, the links between action and self image is pretty weak. Benabou and Tirole recognize that when information concerns oneself, there is an important issue of attribution: is my choice a expression of the weakness of my character or are there situational reasons that justifies it?

[11] Bohr, N. Physique. In *Atomique et Connaissance Humaine*; Edition Gallimard: Paris, France, 1991.

[12] Barkan, R.; Busemeyer, J.R. Modeling dynamic inconsistency with a changing reference point. *J. Behav. Decis. Mak.* **2003**, *16*, 235–255.

[13] Busemeyer, J.R., Wang, Z.; Townsend, J.T. Quantum dynamics of human decision-Making. *J. Math. Psychol.* **2006**, *50*, 220–241.

[14] Busemeyer, J.R. Quantum information processing explanation for interaction between inferences and decisions. In *Proceedings of the Quantum Interaction Symposium*, Arlington, Virginia fall 2010; AAAI Press: Menlo Park, CA, USA 2010.

[15] Busemeyer, J.; Lambert-Mogiliansky, A. An Exploration of type indeterminacy in strategic decision-Making. **2009** In *Quantum Interaction*; Springer: Berlin, Heidelberg, Germany; Volume 5494, pp. 113–128.

[16] Cubitt, R.P.; Starmer, C.; Sugden, R. Dynamic choice and the common ratio effect: An experimental investigation. *Econ. J.* **1998**, *108*, 1362–1380.

[17] Danilov, V.I.; Lambert-Mogiliansky, A. Measurable systems and behavioral sciences. *Math. Soc. Sci.* **2008**, *55*, 315–340.

[18] Danilov, V.I.; Lambert-Mogiliansky, A. Expected utility under non-classical uncertainty. *Theory Decis.* **2009**, *2010*, 25–47.

[19] Deutsch, D. Quantum theory of propability and decisions. *Proc. R. Soc. Lond. A* **1999**, *455*, 3129–3137.

[20] Eisert, J.; Wilkens, M.; Lewenstein, M. Quantum games and quantum strategies. *Phys. Rev. Lett.* **1999**, *83*, 3077–3080.

[21] Festinger L., and Carlsmith J. M. Cognitive consequence of forced compliance, *Journal of Abnormal and Social Psychology*, **1959,** 58 203-201

[22] Franco, R. The conjunction fallacy and interference effects. **2007**, arXiv:0708.3948v1.

[23] Franco, R. The Inverse fallacy and quantum formalism. In *Quantum Interaction* **2008**, College Publications, London 94–98.

[24] Fudenberg, D.; Levine, D.K. A dual self model of impulse control. *Am. Econ. Rev.* **2006**, *96*, 1449–1476.

[25] Fudenberg, D.; Levine, D.K. Timing and Self-Control; *American Economic Review* **2011,** 3, 34-68. .

[26] Fudenberg D. and J. Tirole, *Game Theory*, MIT Press cambridge Massachussets London England, 1991.

[27] Gul, F.; Pesendorfer, W. Temptation and self control. *Econometrica* **2001**, *69*, 1403–1436.

[28] Gul, F.; Pesendorfer, W. Self control and the theory of consumption. *Econometrica* **2004**, *72*, 110–158.

[29] Gul, F.; Pesendorfer, W. The revealed preference theory of changing tastes. *Rev. Econ. Stud.* **2005**, *72*, 429–448.

[30] Heisenberg W., *Philosophie: Le manuscript de 1942*, Edition Seuil, 1998.

[31] Hey, J.D.; Knoll, J.A. How far ahead do people plan? *Econ. Lett.* **2007**, *96*, 8–13.

[32] James, W. The physical basis of emotion. *Psychol. Rev.* **1994**, *1*, 516–529.

[33] Kahneman D., and A. Tversky *Choices, values and Frames*, Cambridge University Press 2000.

[34] Khrennikov, A. *Ubiquitous Quantum Structure—From Psychology to Finance*; Springer: Berlin, Heidelberg, Germany, 2010.

[35] Kirsteiger, G.; Rigotti, L.; Rustichini, A. Your morals might be your moods. *J. Econ. Behav. Organ.* **2006**, *59*, 155–172.

[36] Lambert-Mogiliansky, A.; Zamir, S.; Zwirn, H. Type indeterminacy—A model of the KT(Khaneman Tversky)- man. *J. Math. Psychol.* **2009**, *53*, 349–361

[37] Lambert-Mogiliansky, A. Endogenous Preferences in Games with Type-Indeterminate Players **2010** in AAAI Press: Menlo Park, CA, USA, 2010; Volume FS 10-08, 70–77.

[38] Mackey, G. *Mathematical Foundations of Quantum Mechanics*; Dover Publication Mineola: New York, NY, USA, 1963.

[39] McIntosh, D. *Foundation of Human Society*; University of Chigago Press: Chicago, IL, USA, 1969.

[40] Maskin, E.S. Tirole, J. Markov perfect equilibrium. *J. Econ. Theory* **2001**, *100*, 191–219.

[41] la Mura, P. Correlated equilibria of classical strategies with quantum Signals. *Int. J. Quantum Inf.* **2005**, *3*, 183–188.

[42] Machina, M. Dynamic inconsistency and non-expected utility models of choice under uncertainty. *J. Econ. Lit.* **1989**, *27*, 1622–1668

[43] Peleg, B.; Yaari, M. On the existence of a consistent Course of Action when tastes are changing. *Rev. Econ. Stud.* **1973**, *40*, 391–401.

[44] Schater, A Singer J.E., Congitive, social, and physiological determianants fo emotional states, *Psychological Review*, **1962**, 379-399.

[45] Strotz, R.H. Myopya and time inconsistency in dynamic utility maximization. *Rev. Econ. Stud.* **1956**, *23*, 165–180.

[46] Valins S. and Ray A. A., Effects of cognitive desensistization on avoidance behavior, *Journal of Personality and Social Psychology*, **1967**, 4, 400-408.

[47] Weick K. E. (1967), Dissonance and task enhancement: a problem for compensation theory?, *Organizational Behavior and Human Performence,* **1967,** 7, 345-350."

[48] Zimbardo, P.G.; Cohen, A.; Weisenberg, M.; Dworking, L.; Firestone, I. The control of experimental pain. In *The Congitive Control of Motivation*; Zimbardo, P.G., Ed.; Scotts Foresman: Glenview, IL, USA, 1969; pp. 100–125.