Dual Decision Processes: Retrieving Preferences when Some Choices are Intuitive

Francesco Cerigioni

September 2016

Barcelona GSE Working Paper Series

Working Paper nº 924

Dual Decision Processes:
Retrieving Preferences when some Choices are Intuitive

Francesco Cerigioni*
Universitat Pompeu Fabra and Barcelona GSE

September 14, 2016

Abstract

Evidence from cognitive sciences shows that some choices are conscious and reflect individual preferences while others tend to be intuitive, driven by analogies with past experiences. Under these circumstances, usual economic modeling might not be valid because not all choices are the consequence of individual tastes. We here propose a behavioral model that can be used in standard economic analysis that formalizes how conscious and intuitive choices arise by presenting a decision maker composed by two systems. One system compares past decision problems with the one the decision maker faces, and it replicates past behavior when the problems are similar enough (Intuitive choices). Otherwise, a second system is activated and preferences are maximized (Conscious choices). We then present a novel method capable of finding conscious choices just from observed behavior and finally, we provide a choice theoretical foundation of the model and discuss its importance as a general framework to study behavioral inertia.
(JEL D01, D03, D60)

Keywords: Dual Processes, Fast and Slow Thinking, Similarity, Revealed Preferences, Memory, Intuition

*Department of Economics and Buisness, Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, 08005, Barcelona, Spain. E-mail: francesco.cerigioni@upf.edu. I am grateful to the Ministerio de Ciencia y Innovacion for the FPI scholarship connected with the project ECO2010-17943. I thank Miguel A. Ballester for his insightful comments that were essential for the development of this paper. Furthermore, I want to acknowledge the incredible environment I found in The Eitan Berglas School of Economics (TAU). In particular, I am in debt to Kfir Eliaz, Itzhak Gilboa, Ariel Rubinstein and Rani Spiegler for their help and suggestions. I am also grateful to David Ahn, Jose Apesteguia, B. Douglas Bernheim, Caterina Calsamiglia, Andrew Caplin, Laurens Cherchye, Vincent Crawford, Tugce Cuhadaroglu, Matthew Ellman, Fabrizio Germano, Johannes Gierlinger, Benjamin Golub, Edgardo Lara-Cordova, David Laibson, Marco Mariotti, Daniel Martin, Yusufcan Masatlioglu, Antonio Miralles, Isabel Melguizo, Muriel Nierderle, Efe Ok, Dilan Okcuoglu, Marco Pelliccia, Pedro Rey-Biel, Tomás Rodríguez-Barraquer, Michael Richter, Guillem Roig-Roig, John Tsoukalis and participants at SITE for their comments and suggestions. All mistakes are mine.

1 Introduction

Behavioral economics has posed a serious challenge to standard economic theory since it has documented and studied numerous behaviors inconsistent with preference maximization. Many of these inconsistencies arise because maximizing preferences is effortful for decision makers, as it requires to consciously analyze the problem at hand in order to find the best course of action. Under these conditions, standard revealed preference analysis is misleading. The fact that x is being chosen from menu A does not imply that x is preferred to any other alternative in the menu, unless x was *consciously* chosen from A.

Few of the choices we make every day are the result of a conscious decision making process. Automatic or intuitive decisions are ubiquitous. Think about daily routines, many times we just intuitively stick to decisions we have only consciously thought of once. This mechanism is true not only for marginal decisions. As highlighted by Simon (1987), managers take many intuitive decisions when deciding investment strategies for their firms. Moreover, according to Kahneman (2002), *experts* such as managers, doctors or policy makers often make intuitive decisions or follow their *hunches* due to their extensive experience. Nevertheless, the role of intuition is still understudied in economics. Markets might behave quite differently once intuitive behavior is taken into account. In this paper we provide a first theoretical framework to think about the coexistence of conscious and intuitive behavior and we provide a method to understand what drives individual behavior in different situations once this duality is taken into account.

The dichotomy between fast and intuitive decisions versus slow and conscious ones has been formally studied in cognitive sciences at least since the seminal contributions by Schneider and Shiffrin (1977) and Evans (1977). The successive models and findings in McAndrews, Glisky and Schacter (1987), Evans (1989), Reber (1989), Epstein (1994) and Evans and Over (1996) stimulated the creation of a coherent theory of the individual mind as interaction of two systems called Dual Process Theory that is described in Evans and Frankish (2009) and Kahneman (2011). System 1 is associative and unconscious. System 2 is analytic, conscious and demanding of cognitive capacity.[1] Using analogies, System 1, source of intuitive choices, draws from past behavior to influence decisions. System 2, source of conscious choices, is costly and hence is only activated to solve problems for which past experience cannot be used by System 1.

Past behavior can influence present choices because intuitive decisions are driven by the analogies System 1 makes.[2] Observed behavior is not always indicative of the objectives the decision makers is pursuing, that is choices are not always the outcome of the maximization of preferences. Thus, how can we obtain effective economic modeling if we are unable to understand which are the objectives driving individual behavior?

If we want to understand the role of intuitive decisions in markets, we need first (i) to have

[1]The names of the two systems appeared for the first time in Stanovich (1999). See Evans and Frankish (2009) to get a deeper description of the two different systems and of the historical development of the theory.

[2]See section 2 for a discussion of the different but complementary use of analogies in conscious decisions as formalized in Gilboa and Schmeidler (1995).

a model of intuitive choices that takes into account the role of past choices on present ones and (ii) to analyze whether it is possible from observed behavior to distinguish between conscious and intuitive choices to understand the preferences that drive observed behavior.[3] These are the research questions we address in this paper.

To answer the first question we propose a simple formalization of Dual Process Theory whose main contribution is to make possible for the first time to understand when and how choices should be conscious or intuitive. In section 3, we model a decision maker composed by the two described systems. System 1 compares every decision environment with the ones that the decision maker has already faced according to some given similarity measure. Behavior is replicated from those past problems that are similar *enough* to the present one, i.e. the similarity between the problems passes some threshold. If there are no such problems, System 2 chooses the best available option by maximizing a rational preference relation.[4] Think for example of a consumer that buys a bundle of products from a shelf in the supermarket. The first time he faces the shelf, he tries to find the bundle that best fits his preferences. Afterwards, if the price and arrangement of products do not change *too much*, he will perceive the two decision problems as if they are the same and so he will intuitively stick to the past chosen bundle. If on the contrary, the change in price and arrangement of the products is evident to him, he will choose by maximizing his preferences again.

Even in such a simple framework, there is no trivial way to distinguish which choices are made intuitively and which ones are made consciously. Following the example, suppose our consumer faces again the same problem but this time a new bundle is available and he sticks with the old choice. Is it because the old bundle is preferred to the new one? Or is it because he is choosing intuitively? Notice that the problem can have important economic implications. If the choice is intuitive, the consumer might be missing some better bundles that would increase his utility. If people intuitively stick to suboptimal decisions because of analogies, market outcomes could be far from being efficient. Without a method to distinguish between conscious and intuitive choices it would be very difficult to find a solution to these problems.

We show how to find conscious choices and thus restore standard revealed preference analysis by understanding in which decisions System 2 must have been active. Section 4 assumes that (i) the decision maker behaves according to our model and (ii) the similarity function is known while the threshold is not.[5] We then show that, for every sequence of decision problems, it is possible to identify by means of an algorithm a set of conscious observations and the interval in which the similarity threshold should lie. That is, we provide a novel method to restore revealed preference analysis. First notice that *new* observations, i.e. those in which the choice is an alternative that

[3]See Rubinstein (2007) and Rubinstein (2015) for a similar distinction between conscious and intuitive strategic choices for players participating in a game. See also the distinction in Cunningham and de Quidt (2015) between implicit and explicit attitudes.

[4]Obviously, we are making two strong simplifying assumptions. First, intuitive choices come only from replication of past behavior. Second, System 2 is perfectly rational and does not make mistakes. This is a simplification due to the decision theoretic environment we use. Nevertheless, as the discussion in section 5.1 highlights, it is much more general than what it appears at a first glance.

[5]See sections 3 and 4 for a justification of the latter hypothesis.

had never been chosen before, must be generated by System 2. No past behavior could have been replicated. Starting from these observations, the algorithm iterates the following idea. If an observation is generated by System 2, any other *more novel* observation, that is any problem which is less similar to those decision problems that preceded it, must be also generated by System 2.[6] Returning to our consumer, if we know that after a change in the price of the products on the shelf, the consumer chose consciously, then he must have done so also in all those periods where the change was even more evident.

The algorithm identifies a set of intuitive decisions, i.e. those made by System 1, in a similar fashion, that is, first it highlights some observations that have to be intuitive and then uses this information to reveal other intuitive observations. Notice that understanding if some decisions were made intuitively is very important to understand how analogies are made. Even if intuitive choices do not reveal individual preferences, they tell us what problems are considered similar enough by the decision maker hence allowing for the identification of the interval in which the similarity threshold should lie. We consider *cyclical* datasets, i.e. datasets in which the standard revealed preference relation is cyclic. Any cycle must contain at least one intuitive observation, given that observations generated by System 2 cannot create cycles in the revealed preference. Then, a *least novel observation in a cycle*, i.e. one that is more similar to its past among those forming a cycle, must be intuitive. Once we know that one observation is generated by System 1, so must be all observations that are even *less novel*. Thus, the algorithm finds a set that contains only conscious observations and, in a dual way, another one that contains only intuitive ones. Interestingly enough, if the sequence meets some richness conditions, such sets contain all conscious and intuitive observations respectively.

The algorithm assumes that the decision maker behaves following our model, hence falsifiability of the model becomes a central concern. In section 5 we propose a testable condition that is a weakening of the Strong Axiom of Revealed Preference that characterizes our model and thus renders it falsifiable. Moreover, we show that if the data are rich enough, in particular if we observe choices made by an homogeneous population, two simple consistency requirements not only characterize the model but also allow us to uniquely identify individual preferences and how analogies between decision problems are made.

Section 6 discusses some possible extensions of the base model. First, we analyze the possibility of a decision maker with imperfect memory, that is, that recalls only the m most recent decision problems. We show that such extension does not hinder our algorithmic analysis. In fact, not only the analysis can be reproduced but also, with rich enough data, it is possible to identify the preferences and how similarity comparisons are made by analyzing just one sequence of observations, that is, without recurring to social data. Second, we study the impact of a weaker assumption regarding the similarity of different problems. We allow for the possibility of having only ordinal and partial information on it. Interestingly enough, we find the logic behind the algorithm to be

[6]The idea that conscious behavior is activated in novel or unusual environments is in line with the evidence presented in Alter et al. (2007), Oppenheimer (2008) and Diemand-Yauman, Oppenheimer and Vaughan (2011).

robust to this extension.

Section 7 discusses further implications of the model for the understanding of *sticky behavior*. In particular the model can be seen as a general framework capable of formalizing the coexistence of behavioral inertia and adaptive behavior.[7] The appendix contains all figures, proofs and the estimation of the similarity function from an heterogeneous population of individuals sharing it.

2 Related Literature

In our model the presence of similarity comparisons makes behavior more sticky, that is, if two environments are similar enough then behavior is replicated. This is a different approach with respect to the theory for decisions under uncertainty proposed in Gilboa and Schmeidler (1995) and summarized in Gilboa and Schmeidler (2001). In case-based decision theory, as in our model, a decision maker uses a similarity function in order to assess how much alike are the problem he is facing and the ones he has in his memory. In that model the decision maker tends to choose the action that performed better in past similar cases. There are two main differences with the approach we propose here. First, from a conceptual standpoint, our model relies on the idea of two systems interacting during the decision making process. Second, from a technical point of view, our model uses the similarity in combination with a threshold to determine whether the individual replicates past behavior or maximizes preferences while in Gilboa and Schmeidler (1995) preferences are always maximized. Thus, as section 5.1 highlights, case-based decision theory can be ingrained in the more general structure proposed here. The model in Gilboa and Schmeidler (1995) can be seen as a particular way of making conscious decisions. Nevertheless, both models agree on the importance of analogies for human behavior. In Gilboa and Schmeidler (1995) analogies are used to find the action that maximizes individual utility in a world without priors, here analogies can be potentially dangerous because they determine whether the decision maker thinks about his choices or just chooses automatically without analyzing the problem consciously.

We would like to stress that even if the behavioral model we propose is new and it is a first formalization of Dual Process Theory, nonetheless the idea that observed behavior can be the outcome of the interaction between two different selves is not new and it dates back at least to Strotz (1955). Strotz kind of models, such as Gul and Pesendorfer (2001) or Fudenberg and Levine (2006), are different from the behavioral model we introduce here, since they represent the two selves as two agents with different and conflicting preferences, usually *long-run* vs *short-run* preferences.[8] In our approach however, the two systems are *inherently* different one from the other. One uses analogies to deal with the environment in which the decision maker acts, while the other

[7] See Chetty (2015) for a discussion of the importance of understanding behavioral inertia for public policy.

[8] In some models the difference between the two selves comes from the fact that they have different information. See for example Cunningham (2015) that proposes a model of decision making where the two selves hierarchically aggregate information before choosing an alternative. Some other papers use more than two selves to rationalize individual choices, but still all selves are represented by different preference relations, see for example Kalai, Rubinstein and Spiegler (2002), Manzini and Mariotti (2007) and Ambrus and Rozen (2015).

one uses a preference relation to consciously choose among the alternatives available to the decision maker. Furthermore, which system is activated in a particular decision problem depends on the problems that have been experienced and how similar they are with the present one and thus, it does not depend on whether the decision is affecting the present or the future.[9]

Finally, it is important to notice that the preference revelation strategy we use in the paper agrees with the one used in Bernheim and Rangel (2009). They analyze the same problem of eliciting individual preferences from behavioral datasets, and they do this in two stages. In a first stage they take as given the *welfare relevant domain*, that is the set of observations from which individual preferences can be revealed, and then in a second stage they analyze the welfare relevant observations and propose a criterion for the revelation of preferences that does not assume any particular choice procedure to make welfare judgments.[10] Even if similar, our approach differs in two important aspects. First, by modeling conscious and intuitive choices, we propose a particular method to find the welfare relevant domain, i.e. the algorithm highlighting a set of conscious choices. Second, by proposing a specific choice procedure, we use standard revealed preference analysis on the relevant domain, thus our method, by being behaviorally based, is less conservative for the elicitation of individual preferences. In this sense, our stance is also similar to the one proposed in Rubinstein and Salant (2012), Masatlioglu, Nakajima and Ozbay (2012) and Manzini and Mariotti (2014) that make the case for welfare analysis based on the understanding of the behavioral process that generated the data.

3 Dual Decision Processes

Let X and E be two finite sets. The decision maker (DM) faces at every moment in time t a *decision problem* (A_t, e_t) with $A_t \subseteq X$ and $e_t \in E$. The set of alternatives A_t that is available at time t, and from which the DM has to make a choice, is usually called the *menu*. An alternative is any element of choice like consumption bundles, lotteries or even streams of consumption. The *environment* e_t is a description of the possible characteristics of the problem that the DM faces at time t.[11] We simply denote by $a_t \in A_t$ the chosen alternative at time t. With little abuse of the notation, we refer to the couple formed by the decision problem (A_t, e_t) and the chosen alternative a_t as *observation* t. We denote the collection of observations in the sequence $\{(A_t, e_t, a_t)\}_{t=1}^{T}$ as D, i.e. $D = \{1, ..., T\}$.

The DM is composed by two systems, System 1 (S1) and System 2 (S2) and the chosen alternative is determined by either one of them. S1 is constantly operating and uses *analogies* to relate the decision environment the DM is facing with past ones. If a past environment is similar enough

[9] Nevertheless, we do not exclude the possibility that the fact that a decision affects the present or the future has some kind of influence on how analogies are made.

[10] Notice that Apesteguia and Ballester (2015) propose an approach to measure the welfare of an individual from a given dataset that is also choice-procedure free. They do so by providing a model-free method to measure how close actual behavior is to the preference that best reflects the choices in the dataset.

[11] See below for some examples of environments.

to the one the DM is facing, then S1 replicates the choice made in the past whenever available. If no past environment is similar enough or replication is not possible, S2 is activated. S2 uses a *preference* relation to compare the alternatives in the menu and chooses the best one.[12] We call this behavioral model a *dual decision*(DD) process.

Formally, let $\sigma : E \times E \to [0,1]$ be the *similarity function*. The value $\sigma(e, e')$ measures how similar environment e is with respect to e'. S1 is endowed with a *similarity threshold* $\alpha \in [0,1]$ that delimits which pairs of environments are similar enough. Whenever $\sigma(e, e') > \alpha$ the individual considers e to be similar enough to e'. At time t and facing the decision problem (A_t, e_t), S1 executes a choice if it can replicate the choice of a previous period $s < t$ such that $\sigma(e_t, e_s) > \alpha$. The choice is the alternative a_s chosen in one such period. That is, if the DM faces a decision environment e_t that is similar enough to a decision environment e_s he has already faced and the alternative chosen in s is present in t, the DM chooses it again in t.[13] S2 is endowed with a *preference relation* $\succ$ over the set of alternatives.[14] At time t, if S2 is activated, the most preferred available alternative is chosen, that is, S2 chooses the alternative a_t that maximizes $\succ$ in A_t. Summarizing:

$$a_t = \begin{cases} a_s \text{ for some } s < t \text{ such that } \sigma(e_t, e_s) > \alpha \text{ and } a_s \in A_t, \\ \text{the maximal element in } A_t \text{ with respect to } \succ, \text{ otherwise.} \end{cases}$$

Two remarks are useful here. First, notice that intuitive and conscious decisions are separated by the behavioral parameter α. In some sense α is summarizing the cost of using the cognitive demanding system, i.e. the *laziness* of S2 as it is described in Kahneman (2011). The higher the cost, the lower the threshold. Thus, parameter α captures individual heterogeneity on S1. In fact, following the common interpretation in cognitive sciences, we take the similarity function as given, i.e. as an innate component of all individuals, while the similarity threshold, by representing cognitive costs, is what makes analogy comparisons different at the individual level.[15] Notice that while the similarity function has been widely studied in cognitive sciences, e.g. Tversky (1977), Medin, Goldstone and Gentner (1993) and Hahn (2014), the cognitive costs of activating S2 are still an unknown, thus the method we propose in section 4 can be seen as a first attempt of identifying from observed behavior the interval in which such costs should lie, given the similarity function.

As a second remark, notice that we are describing a class of models because we do not impose

[12]The idea that conscious behavior comes from the maximization of a preference relation is a simplification we use to focus the analysis on the main novelties of the framework presented in this paper. For a more detailed discussion regarding this point, see section 5.1.

[13]Notice that we assume that the DM has perfect memory. There is evidence in favor of perfect memory for choices that are unconscious. See Duhigg (2012) for an informal discussion of how *consciously* forgotten psychological cues can still affect our behavior and decisions. However, as shown in section 6, this assumption does not weaken the analysis.

[14]For ease of exposition, we assume that $\succ$ is a strict order, i.e. an asymmetric, transitive and complete binary relation, defined over X.

[15]In the appendix we show how it is possible to recover the similarity function when different individuals share the same similarity function.

any particular structure on the replicating behavior. We do not specify which alternative would be chosen when more than one past choice can be replicated. Many different behaviors can be part of this class, e.g choosing the alternative that was chosen in the most similar past environment or choosing the alternative that maximizes the preference relation over the ones chosen in similar enough past environments, etc.[16] All the analysis that follows is valid for the class as a whole.

As a final remark, given the centrality of the similarity function for the model, we propose here some examples of environments that are relevant for economic applications and a possible similarity function that can be used in such cases.

Environments as Menus: *In many economic applications it seems sensible to see the whole menu of alternatives, e.g. the budget set, as the main driver of analogies. That is, E could be the set of all possible menus and two decision problems are perceived as similar as their menus are. In this framework, $E = 2^X$.*

Environments as Attributes: *Decision makers many times face alternatives that are bundles of attributes. In those contexts, it is reasonable to assume that the attributes of the available alternatives determine the decision environment. If A is the set containing all possible attributes, then $E = 2^A$.*

Environments as Frames: *We can think of the set E as the set of frames or ancillary conditions as described in Salant and Rubinstein (2008) and Bernheim and Rangel (2009). Frames are observable information that is irrelevant for the rational assessment of alternatives, for example how the products are disposed on a shelf. Every frame can be seen as a set of irrelevant features of the decision environment. Thus, if the set containing all possible irrelevant features is F, we have $E = 2^F$.*

In all the previous examples it is natural to assume that the similarity function relates different environments depending on their commonalities and differences. For example, $\sigma(e, e') = \frac{|e \cap e'|}{|e \cup e'|}$, that is, two environments are more similar the more characteristics they share relative to all the characteristics they have.[17] Although, it is sometimes not possible to have all the information regarding the similarity function, a case we analyze in section 6, from now on we take E and σ as given.[18]

We now provide an example to illustrate the behavioral process we are modeling.

Example 1 *Let $X = \{1, 2, 3, ..., 10\}$. We assume that environments are menus, i.e. E is the set of all subsets of X and we assume that $\sigma(A, A') = \frac{|A \cap A'|}{|A \cup A'|}$. Suppose that S1 is described by $\alpha = .55$ and that the preference $1 \succ 2 \succ 3 \succ \cdots \succ 10$ describes S2. We now explain how our DM makes choices from the following list of ordered menus:*

[16] A formal analysis of these possibilities is available upon request.

[17] Such function is just a symmetric specification of the more general class considered in Tversky (1977).

[18] In those cases where the similarity function cannot be completely known, weaker assumptions can be made. For example, one could think that the similarity function respects the symmetric difference between sets. That is environment e and environment e' are more similar than g and g' if $e \cup e' \setminus e \cap e' \subseteq g \cup g' \setminus g \cap g'$.

$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$	$t=8$
$3,4,5,6$	$1,2,3,4,5,6$	$1,3,4,7$	$2,4,7$	$1,3,6$	$1,2,3,4$	$2,4,8$	$2,4,8,9,10$

In the first period, given the DM has no prior experiences, S2 is going to be active. Thus, the choice comes from the maximization of preferences, that is, $a_1 = 3$*. Then, in the following period, given that we have* $\sigma(A_2, A_1) = \frac{4}{6} > .55$*, S1 is active and so we would observe a replication of past choices, that is,* $a_2 = 3$*. Now, in period 3, notice that the similarity between* A_3 *and* A_2 *or* A_1 *is always below the similarity threshold and this makes S2 to be active. The preference relation is maximized and so* $a_3 = 1$*. A similar reasoning can be applied for the fourth and fifth periods to see that* $a_4 = 2$ *and* $a_5 = 1$*. Then, in period six, S1 is active given that* $\sigma(A_6, A_3) = \frac{3}{5} > .55$*, leading to* $a_6 = 1$*. In period seven, given no past environment is similar enough, S2 is active and so* $a_7 = 2$*. Finally, in the last period S1 is active again given that* $\sigma(A_8, A_7) = \frac{3}{5} > .55$ *and so behavior will be replicated, i.e.* $a_8 = 2$*.*

One may wonder what an external observer would understand from this choice behavior. How would the observer determine which choices were done by S1 or S2 and hence which observations are informative on our DM's preferences? Is it possible to retrieve the similarity threshold? In section 4 we propose an algorithm that allows us to disregard the preferential information coming from the choices in periods two, six and seven, as it should be, while maintaining all the preferential information coming from the remaining choices.

4 The Revealed Preference Analysis of Dual Decisions

In this section, we discuss how to recognize which observations were generated by either S1 or S2 in a DD process. This information is crucial to elicit the unobservables in the model that are the sources of individual heterogeneity, that is the preference relation and the similarity threshold. As we previously discussed in section 3, we take the similarity function, common across individuals, as given, while we want to elicit from observed behavior the similarity threshold α.[19]

It is easy to recognize a set of observations that is undoubtedly generated by S2. Notice that all those observations in which the chosen alternative was never chosen in the past must belong to this category. This is so because, as no past behavior has been replicated, S1 could not be active. We call these observations *new* observations.

In order to identify a first set of observations generated by S1, notice that S2 being rational, it cannot generate cycles of revealed preference.[20] Clearly, for every cycle there must be at least

[19]Notice that, as discussed in section 5, the similarity function and the similarity threshold define a binary similarity function that is individual specific. Thus, by separating the similarity function from the similarity threshold we are able to associate individual heterogeneity to a parameter related with individual cognitive costs without having too many degrees of freedom to properly run the technical analysis.

[20]As it is standard, a set of observations $t_1, t_2, \ldots, t_k$ forms a *cycle* if $a_{t_{i+1}} \in A_{t_i}$, $i = 1, \ldots, k-1$ and $a_{t_1} \in A_{t_k}$, where all chosen alternatives are different.

one observation that is generated by S1. Intuitively, the one corresponding to the most familiar environment should be a decision mimicking a past behavior. The *unconditional familiarity* of observation t is

$$f(t) = \max_{s<t, a_s \in A_t} \sigma(e_t, e_s).^{21}$$

That is, unconditional familiarity measures how similar observation t is to past observations from which behavior *could be* replicated, i.e. those past decision problems for which the chosen alternative is present at t. Then, we say that observation t is a *least novel in a cycle* if it is part of a cycle of observations, and within it, it maximizes the value of the unconditional familiarity.

The major challenge is to relate pairs of observations in a way that allows to transfer the knowledge of which system generated one of them to the other. In order to do so, we introduce a second measure of familiarity of an observation t, that we call *conditional familiarity*. Formally,

$$f(t|a_t) = \max_{s<t, a_s = a_t} \sigma(e_t, e_s).^{22}$$

That is, conditional familiarity measures how similar observation t is with past observations from which behavior *could have been* replicated, i.e. those past decision problems for which the choice is the same as the one at t. The main difference between $f(t)$ and $f(t|a_t)$ is that the first one is an ex-ante concept, i.e. before considering the choice, while the second one is an ex-post concept, i.e. after considering the choice. Our key definition uses these two measures of familiarity to relate pairs of observations.

Definition 1 (Linked Observations) *We say that observation t is* linked *to observation s, and we write $t \in L(s)$, whenever $f(t|a_t) \leq f(s)$. We say that observation t is* indirectly linked *to observation s if there exists a sequence of observations $t_1, \ldots, t_k$ such that $t = t_1$, $t_k = s$ and $t_i \in L(t_{i+1})$ for every $i = 1, 2, \ldots, k-1$.*

Denote by D^N the set of all observations that are indirectly linked to new observations and by D^C the set of all observations to which least novel observations in a cycle are indirectly linked.[23] We are ready to present the main result of this section. It establishes that observations in D^N are generated by S2, while observations in D^C are generated by S1. As a consequence, it guarantees that the revealed preference of observations in D^N, i.e. $R(D^N)$, is useful information regarding the preferences of the individual.[24] Moreover, an interval in which the similarity threshold has to lie is identified. Such interval provides bounds for the individual specific cognitive costs of activating S2.

[21] W.l.o.g., whenever there is no $s < t$ such that $a_s \in A_t$, we say $f(t) = 0$.

[22] W.l.o.g., whenever there is no $s < t$ such that $a_s = a_t$, we say $f(t|a_t) = 0$.

[23] The binary relation determined by the concept of linked observations is clearly reflexive, thus, new observations and least novel observations in a cycle are contained in D^N and D^C respectively.

[24] We say that x is revealed preferred to y in a set of observations O, and write $xR(O)y$, if there is a sequence of different alternatives $x_1, x_2, \ldots, x_k$ such that $x_1 = x, x_k = y$ and for every $i = 1, 2, \ldots, k-1 \in O$, it is $x_i = a_t$ and $x_{i+1} \in A_t$ for some t.

Proposition 1 *For every collection of observations D generated by a DD process:*

1. *all observations in D^N are generated by S2 while all observations in D^C are generated by S1,*
2. *if x is revealed preferred to y for the set of observations D^N, then $x \succ y$,*
3. $\max_{t \in D^N} f(t) \leq \alpha < \min_{t \in D^C} f(t|a_t)$.

To understand the reasoning behind Proposition 1, consider first an observation t that we know is new, and hence generated by S2. We have learnt that its corresponding environment is not similar enough to any other previous environment. In other words, $f(t) \leq \alpha$. Then, any observation s for which the conditional familiarity is less than $f(t)$ must be generated by S2 too. In fact, $f(s|a_s) \leq f(t) \leq \alpha$ implies that no past behavior that could have been replicated in s comes from an environment that is similar enough to the one in s. Thus, any observation linked with a new observation must be generated by S2. It is easy to see that this reasoning can be iterated, in fact, any observation linked with an observation generated by S2 must be generated by S2 too.

Similarly, consider a least novel observation in a cycle t, that we know is generated by S1. Any observation s for which the unconditional familiarity is greater than the conditional familiarity of t must be generated by S1 too. In fact, we know that $\alpha < f(t|a_t)$ because t is generated by S1. Then, any observation s to which t is linked has an unconditional familiarity above α, which implies that some past behavior *could be* replicated by S1, and so such observation must be generated by S1 too. Again, the reasoning can be iterated. Thus, we can start from a small subset of observations undoubtedly generated by either S1 or S2, inferring from there which other observations are of the same type.

We now use Example 1 to illustrate the algorithm. In doing so, we show that it is possible to see our algorithmic analysis in terms of graph theory. In fact, when two observations are linked we can think of them as connected by an undirected edge. Then, we can clearly see that D^N and D^C have to be the sets containing all those nodes that belong to the connected components of new and least novel observations in a cycle respectively.

Graphical Intuition: Example 1

Suppose that we observe the decisions made by the DM in Example 1, without any knowledge on his preferences $\succ$ or similarity threshold α. The following table summarizes the different observations.

$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$	$t=8$
$3,4,5,6$	$1,2,3,4,5,6$	$1,3,4,7$	$2,4,7$	$1,3,6$	$1,2,3,4$	$2,4,8$	$2,4,8,9,10$
$a_1=3$	$a_2=3$	$a_3=1$	$a_4=2$	$a_5=1$	$a_6=1$	$a_7=2$	$a_8=2$

We can easily see that the only new observations are observations 1, 3 *and* 4*, and hence we can directly infer that S2 was active in making the corresponding choices. It is immediate to see that*

new observations are always linked to each other and hence, observations 1, 3 *and* 4 *are connected by undirected edges as in the graph of Figure 1.*[25]

We can go one step further and consider observation 5*. From observed behavior we cannot understand whether the choice comes from maximization of preferences or the replication of past behavior in period* 3*. Nevertheless, S2 was active in period* 3 *and one can easily see that* $f(5|a_5) = \frac{2}{5} \leq \frac{3}{7} = f(3)$*, making observation* 5 *linked with observation* 3 *and according to Proposition 1, making it generated by S2 too. This is represented in Figure 2.*

Consider now observation 7*. We cannot directly link observation* 7 *to either observations* 1, 3 *or* 4*, because* $f(7|a_7) = \frac{1}{2} > \max\{f(1), f(3), f(4)\}$*. However, we can indirectly link observation* 7 *to observation* 3 *through observation* 5*, because* $f(7|a_7) = \frac{1}{2} \leq \frac{1}{2} = f(5)$*, thus making* 7 *an element of* D^N*. See Figure 3 for a graphical description. No other observation is indirectly linked to observations* 1, 3 *or* 4 *and hence,* $D^N = \{1, 3, 4, 5, 7\}$*. The method rightfully excludes all S1 observations from* D^N*.*

The example presents inconsistencies in the revealed preference. Observation 3 *and* 6 *are both in conflict with observation* 2*. That is, observations* 2 *and* 3 *and* 2 *and* 6 *form cycles. Then, noticing that* $\max\{f(2), f(3)\} = f(2)$ *and that* $\max\{f(2), f(6)\} = f(2) = f(6)$ *we can say that observations* 2 *and* 6 *are generated by S1 thanks to Proposition 1, given they are least novel in a cycle. It is immediate again to see that* 2 *and* 6 *are connected by an undirected edge. See Figure 4 for a graphical description.*

But then, notice that observation 6 *is linked to observation* 8 *given that* $f(6|a_6) = \frac{3}{5} \leq f(8) = \frac{3}{5}$ *revealing that the latter must have been generated by S1 too. Figure 5 shows this idea graphically. Thus, we get* $D^C = \{2, 6, 8\}$ *that were the observations rightfully excluded from* D^N*. No decision made by S2 has been cataloged as intuitive. Thanks to the algorithm, we found the two connected components as shown in Figure 6.*

The modified revealed preference exercise helps us determine that alternative 1 *is better than any alternative from* 2 *to* 7*, alternative* 3 *is better than any alternative from* 4 *to* 6*, and alternative* 2 *is better than alternatives* 4, 7 *and* 8 *as it is indeed the case. The value of the similarity threshold* α *by Proposition 1 can be correctly determined to be in the interval* $[0.5, 0.6)$ *thanks to the information retrieved from observations* 7 *and* 8 *respectively.*

Notice that in general D^N and D^C will be proper subsets of the whole S2 and S1 sets of observations, respectively. The set D^N is built upon the set of new observations and those indirectly linked to them.[26] It may be the case that some S2 observations are not linked to other S2 observations.[27] Nonetheless, if the observations are *rich enough*, it is possible to guarantee that D^N and D^C coincide with the sets of conscious and intuitive decisions. In the following section we analyze a rich dataset, where richness is achieved through social data, that allows for the identification of

[25] Notice that for any new observation t, $f(t|a_t) = 0$.

[26] D^N is never empty because it always contains the first observation.

[27] For this reason, nothing guarantees that $D \setminus D^N$ are intuitive observations and hence, Proposition 1 needs to show how to dually construct a set of intuitive decisions D^C.

the two sets of conscious and intuitive observations.[28] More importantly, notice that Proposition 1 relies on one important assumption, that is, the collection of observations is generated by a DD process. The following section addresses this issue.

5 A Characterization of Dual Decision Processes

In section 4 we showed how to elicit the preferences and the similarity threshold of an individual that follows a DD process. Here, building upon the results of that section, we provide a necessary and sufficient condition for a set of observations to be characterized as a DD process with a known similarity function. In other words, we provide a condition that can be used to falsify our model. Finally, we provide an alternative characterization of the model whenever the similarity function is also unknown and observations generated by an homogeneous population are available. This alternative characterization allows us to uniquely identify the preferences of the DM and also how similarity comparisons are made.

From the construction of the set D^N, we understand that a necessary condition for a dataset to be generated by a DD process is that the indirect revealed preference we obtain from observations in D^N, i.e. $R(D^N)$, must be asymmetric. It turns out that this condition is not only necessary but also sufficient to represent a sequence of decision problems as if generated by a DD process. One simple postulate of choice characterizes the whole class of DD processes. Interestingly enough though, it is possible to characterize such class with a condition that is computationally easier to test.

Axiom 1 (Link-Consistency) *A sequence of observations* $\{(A_t, e_t, a_t)\}_{t=1}^T$ *satisfies* Link-Consistency *if, for every* $t \in D^N$, $xR(t)y$ *implies not* $yR(L(t))x$.

This is a weakening of the Strong Axiom of Revealed Preference. In fact it imposes that no cycle can be observed when looking at an observation t in D^N and those directly linked to it. This condition is easy to test computationally because it only asks to check for acyclicity between linked observations. The next theorem shows that such condition is indeed necessary and sufficient to characterize DD processes with known similarity.

Theorem 1 *A sequence of observations* $\{(A_t, e_t, a_t)\}_{t=1}^T$ *satisfies Link-Consistency if and only if there exist a preference relation* $\succ$ *and a similarity threshold* α *that characterize a DD process.*

The theorem is saying that Link-Consistency makes possible to determine whether the DM is following a DD process or not. In particular, when the property is satisfied, we can characterize the preferences of the DM with a completion of $R(D^N)$ which is asymmetric thanks to Link-Consistency and use the lower bound of α as described in Proposition 1 to characterize the similarity

[28] In section 6 we study the case of a forgetful decision maker and we show how in such case it is possible to identify the two sets when one sequence of observations is rich enough.

threshold. In fact, by construction, for any observation t outside D^N it is possible to find a preceding observation that can be replicated, i.e. the one defining $f(t|a_t)$. Clearly the familiarity of an observation can be defined only if the similarity function is known. In the appendix we show how to identify such function from a population of heterogeneous individuals following a DD process. Here we propose an alternative approach, that also uses social data, that allows us to jointly determine preferences and similarity comparisons.

Notice that we do not assume any particular structure for the sequence of observations we use as data and hence, the characterization of preferences does not have to be unique, even when the similarity is known. One way to uniquely characterize the preferences of the DM that allow us also to determine how similarity comparisons are made, is to get different sequences of observations generated by an homogeneous population. That is, suppose we observe many and different sequences of decision problems and choices generated by a population of individuals sharing the same preferences, similarity threshold and similarity function.[29] Let $\mathcal{D}$ be the set containing such sequences and $D \in \mathcal{D}$ be the collection of observations composing one of them. Then, if $\mathcal{D}$ is rich enough we can perfectly identify not only the preference relation but also, for every decision environment, which other environments are considered similar enough and which are not. Notice that this completely identifies the model. In fact, the similarity threshold and the similarity function define a binary similarity function, i.e. a Boolean similarity function. That is, given any environment, the combination of similarity function and similarity threshold partitions the set of environments in two sets of similar and dissimilar environments. Thus, knowing for every decision environment which other environments are considered similar enough and which are not, completely identifies such function.

We say that $\mathcal{D}$ is rich if:

- for any $x, y \in X$ there exists a $D \in \mathcal{D}$ such that there is some $t \in D$ where $x, y \neq a_s$ for $s < t$ and $A_t = \{x, y\}$.
- for any $x, y, z \in X$ there exists a $D \in \mathcal{D}$ such that there is some $t \in D$ where $x, y, z \neq a_s$ for $s < t$ and $A_t = \{x, y, z\}$.
- for any $e, e' \in E$ and any $x, y \in X$, there exists a $D \in \mathcal{D}$ such that there is some $t \in D$ where $x, y \neq a_s$ for $s < t - 1$ and $A_{t-1} = \{y\}$, $e_{t-1} = e'$, $A_t = \{x, y\}$ and $e_t = e$.

The first two requirements impose that for every pair and triple of alternatives, there is some collection D in which until some moment in time t, they have never been chosen.[30] That is, for any pair and triples of alternatives there is a sequence of decision problems in which for some moment in time t they were part of a new observation. The third requirement imposes that for any pair of environments and any pair of alternatives there is a sequence of observations such that

[29] It is possible to argue that individuals that share the same socioeconomic characteristics and have similar cognitive capabilities are likely to be described by the same DD process.

[30] Notice that the first two conditions play the same role of the Universal Domain assumption in standard choice theory.

the environments are part of two consecutive decision problems in which the two alternatives have never been chosen before. Thus, the choice in t can be either new or the same as in $t-1$. These conditions allow us to perfectly identify the preferences of the homogeneous population and how analogies are made whenever the observed choices satisfy some consistency requirements.

Before stating such requirements, it is useful to define a set for any environment $e \in E$ that contains all those environments that would be considered similar enough to e by a DM following a DD process. Suppose, without loss of generality, that for some $D \in \mathcal{D}$ there exists $t \in D$ such that $A_t = \{x, y\}$ and $a_t = x$ with $x, y \neq a_s$ for $s < t$. If the observations are generated by a DM following a DD process, then t would be a new observation and x would be revealed preferred to y. Then let $S(e)$ be defined as follows:

$$S(e) = \{e' \in E |\ \exists D \in \mathcal{D} \text{ such that, for some } t \in D, t-1 = (\{y\}, e', y) \text{ is new and } t = (\{x, y\}, e, y)\}.$$

For the same reasoning developed before, if the observations are generated by individuals following a DD process, $S(e)$ would contain only environments considered similar enough to e because the observed inversion of preferences in t is possible only when past behavior is replicated. In fact, if y is chosen over x in t, it must be because of replication of behavior in $t-1$. This is a concept similar to the one of revealed preferences. Environment e' is revealed similar to e whenever such inversion of preferences occurs. Notice that by richness, $S(e)$ would contain all those environments that are considered similar enough to e.

Finally, for any collection $D \in \mathcal{D}$ define for all observations $t \in D$ the following set:

$$I(t) = \{x \in X |\ x = a_s, \text{ for some } s < t \text{ such that } a_s \in A_t \text{ and } e_s \in S(e_t)\}.$$

If the observations are generated by DD processes, $I(t)$ would contain all those past choices that could be replicated in t. Now we have all the ingredients to state the consistency requirements that characterize the whole class of DD processes for a rich dataset $\mathcal{D}$. The following axioms are intended for $D, D' \in \mathcal{D}$.

The first axiom requires that conscious choices are consistent. That is, there do not exist two observations in which x is consciously chosen over y in one of them and y over x in the other. This is a weakening of the Weak Axiom of Revealed Preference.

Axiom 2 (Conscious Consistency (CC)) *For any $t \in D$ and $t' \in D'$ such that $I(t) = I(t') = \emptyset$, if $x, y \in A_t \cap A_{t'}$ and $x = a_t$ then $y \neq a_{t'}$.*

The second axiom requires that intuitive choices come from replication of past behavior.

Axiom 3 (Intuitive Consistency (IC)) *For any $t \in D$ such that $I(t) \neq \emptyset$, $a_t \in I(t)$.*

Then we can state the following theorem.

Theorem 2 *A rich dataset $\mathcal{D}$ satisfies CC and IC if and only if there exist a preference relation $\succ$, a similarity function σ and a similarity threshold α that characterize a DD process. Moreover, the preference relation $\succ$ and the binary similarity function defined by σ and α are uniquely identified.*

Intuitively, the first two requirements of a rich dataset plus CC assure that the revealed preference relation constructed from the observations that would have to be explained as conscious choices is complete and transitive. Then, IC assures that those choices that should be explained as intuitive, replicate past behavior. Uniqueness comes from the fact that every pairwise comparison between alternatives and between environments is observable thanks to richness.

5.1 On Conscious and Intuitive Behavior

Throughout all the paper we have assumed that conscious behavior is the maximization of a given preference relation. In some cases this assumption can be too strong. Inconsistent choices might arise also when choosing consciously. The framework and analysis we develop here do not depend on the particular conscious behavior that is assumed. In fact, as the previous section should clarify, for any given decision environment analogies determine a partition with two components, one containing those problems that are similar enough to the reference environment and another one containing those that are not. Such partition is based only on one and simple assumption, intuitive behavior must come from the replication of past behavior.

Alternative conscious behaviors are possible. The only element of the formal analysis we conducted that has to be changed is what kind of consistency requirement to test on those problems that fall in the set that has to be assigned to conscious behavior.[31] Thus we can think about preferences that depend on the decision environment or preferences that satisfy only quasi-transitivity or less consistent choice behaviors that are falsifiable and we would still be able to run the same kind of analysis.

The purpose of this kind of flexible framework is to provide a theoretical benchmark to analyze in a structured way what conscious and intuitive behavior are. The model we propose here is just a first step into this new direction and hence it is simplified to keep the analysis focused on the novel aspect of the problem we address, that is, (i) the relationship between intuitive and conscious choices and (ii) the structure observed behavior should have once we consider these two sources of individual behavior.

Furthermore, notice that the assumptions that intuitive behavior comes from the replication of past behavior is much less restrictive than what would appear at a first glance. The idea we are trying to capture is that intuitive choices have to be *familiar* choices. Intuitive choices come from replication or combination of past experiences. Obviously, familiar choices can come from other sources different than one's memory. We have considered only the case of replicating past choices made by *the same* DM because we are considering a fictitious and quite restrictive decision theoretic environment. It would be easy to incorporate social considerations by changing what

[31] Obviously, dually, to find intuitive choices we should analyze violations of such requirements.

experiences and decision problems form part of the DM's memory. For example, we can consider cases where the DM stores in his memory not only his experiences, but also his parents', siblings' or friends' ones thus making the concept more general than what the literal interpretation of the model would suggest.

Finally, it is important to underline that we are not assuming that conscious choices are slower or less automatic than intuitive ones. The idea here is that conscious and intuitive choices differ because the first ones are generated by rules and analytical thinking while the second ones are generated by similarities and analogical thinking, that is, we are underlining the different cognitive sources of the two types of choices.[32]

6 Extensions

The analysis developed in the previous sections is based on two assumptions that we relax here. First, we have assumed that the DM has perfect memory. In this section we show that such an assumption is not needed to perform the algorithmic analysis. Moreover, if some richness conditions are satisfied by the sequence of decision problems under study, we show that by considering imperfect memory we can completely identify the preferences of the DM and determine which system generated every single observation by studying just one sequence of observations. The second assumption we relax concerns the similarity function. We show that the analysis of the previous sections is perfectly valid, even if we have only partial information regarding the similarity function.

6.1 A Forgetful Decision Maker

So far, we have assumed that the DM has perfect memory, i.e. intuitive decisions can come from the replication of *any* past choice. In this section, we depart from such assumption and analyze the possibility of a DM that *forgets* older choices.

Suppose the DM can remember up until $m \geq 1$ periods of past choices. In a DD-m process, the chosen action in period t is:

$$a_t = \begin{cases} a_s \text{ for some } t-m \leq s < t \text{ such that } \sigma(e_t, e_s) > \alpha \text{ and } a_s \in A_t, \\ \text{the maximal element in } A_t \text{ with respect to } \succ, \text{ otherwise.} \end{cases}$$

Notice that we have just changed the periods that are considered by S1 for the replication of behavior, the structure of the process is otherwise unchanged. Thus, if we take this new assumption into account, we should be able to directly apply the logic behind the algorithm to this new framework. This is indeed the case.

[32] On average though, we expect conscious choices to be slower than intuitive ones.

We say that an observation is *new with imperfect memory* whenever $a_t \neq a_s$ for all $t-m \leq s < t$. That is, the choice in t was never chosen in the previous m periods. Such observations must be generated by S2 for the same logic explained in section 4. In a similar fashion, let *unconditional* and *conditional familiarity with imperfect memory* be as follows:

$$\overline{f}(t) = \max_{t-m \leq s < t, a_s \in A_t} \sigma(e_t, e_s).$$

$$\overline{f}(t|a_t) = \max_{t-m \leq s < t, a_s = a_t} \sigma(e_t, e_s).$$

Again, the only change is that now, for any observation t, only the preceding m periods are important for the replication of behavior, and so they are the only ones considered when defining the two concepts of familiarity. Then, we say that an observation is *the least novel in a cycle with imperfect memory* whenever it maximizes the unconditional familiarity with imperfect memory among those observations in the cycle. For the same logic we used before, any least novel observation in a cycle must be generated by S1.

Finally, we say that observation t is linked to observation s whenever $\overline{f}(t|a_t) \leq \overline{f}(s)$, and indirectly linked observations are defined in an analogous way. Thus, considering imperfect memory only changes the key definitions on which the algorithm is based, not the logic behind it. Again, any observation linked to an observation generated by S2 must be generated by S2 too. Symmetrically, any observation to which an observation generated by S1 is linked, must be generated by S1 too. Denote with $D^{\overline{N}}$ the set containing all observations that are indirectly linked to new observations with imperfect memory and with $D^{\overline{C}}$ the set containing all observations to which a least novel in the cycle with imperfect memory is linked. Then we can state the parallel version of Proposition 1. The proof is omitted.

Proposition 2 *For every collection of observations D generated by a DD-m process:*

1. *all observations in $D^{\overline{N}}$ are generated by S2 while all observations in $D^{\overline{C}}$ are generated by S1,*
2. *if x is revealed preferred to y for the set of observations $D^{\overline{N}}$, then $x \succ y$,*
3. $\max_{t \in D^{\overline{N}}} \overline{f}(t) \leq \alpha < \min_{t \in D^{\overline{C}}} \overline{f}(t|a_t)$.

Similarly, let Link-Consistency* be the parallel version of Link-Consistency defined over $D^{\overline{N}}$. We can state the analogous version of Theorem 1. Again, the proof is omitted.

Theorem 3 *A sequence of observations $\{(A_t, e_t, a_t)\}_{t=1}^T$ satisfies Link-Consistency* if and only if there exist a preference relation $\succ$ and a similarity threshold α that characterize a DD-m process.*

The analysis of a forgetful individual can provide additional insights if we impose some *richness* conditions on the sequence of decision problems the DM faces. A sequence of observations $\{(A_t, e_t, a_t)\}_{t=1}^T$ is *m-rich* for an individual following a $DD-m$ process whenever:

- for any $x, y \in X$ there exists a t such that $x, y \neq a_s$ for $t - m \leq s < t$ and $A_t = \{x, y\}$.
- for any $x, y, z \in X$ there exists a t such that $x, y, z \neq a_s$ for $t - m \leq s < t$ and $A_t = \{x, y, z\}$.
- for any $e, e' \in E$ and any $x, y \in X$, there exists a t such that $x, y \neq a_s$ for $t-m-1 \leq s < t-1$ and $A_{t-1} = \{y\}$, $e_{t-1} = e'$, $A_t = \{x, y\}$ and $e_t = e$.

These conditions are similar to the ones we studied in section 5 so we do not analyze them further. One important thing to notice is that in this case richness is imposed on *one* sequence of observations not on a *collection* of sequences. Then, we can state the following:

Proposition 3 *For every $m-$rich sequence of observations $\{(A_t, e_t, a_t)\}_{t=1}^{T}$ generated by a $DD-m$ process:*

1. *$D^{\overline{N}}$ contains all the decisions generated by S2 and $D^{\overline{C}}$ contains all the decisions generated by S1, that is $D^{\overline{N}} \cup D^{\overline{C}} = D$,*
2. *x is revealed preferred to y for the set of observations $D^{\overline{N}}$ if and only if $x \succ y$.*

Thus, Proposition 3 highlights the fact that preferences and similarity comparisons of a forgetful decision maker can be recovered entirely without the need of observing social data whenever the sequence of observations is rich enough.

Furthermore, for the same reasoning developed in Section 5, an $m-$rich sequence of observations allows for the characterization of a DD$-m$ process and the identification of the preferences and analogies of the DM. Given the analysis would be almost identical, in fact only the definitions of $S(e)$ and $I(t)$ would slightly change, we omit it to avoid repetitions.

6.2 Revealing S1 and S2 with Partial Information on the Similarity

In this section we show that our algorithmic analysis is robust to weaker assumptions concerning the knowledge of the similarity function. In particular, we study the case in which only a partial preorder over pairs of environments is known, denoted by $\succeq$.[33] Such extension can be relevant in many contexts where it is not possible to estimate the similarity function. In such cases, it is sensible to assume that at least some binary comparisons between pairs of environments are known. Coming back to the example of the introduction, we might not know how the DM compares different prices and dispositions of the products on the shelf, but we might know that for any combination of prices, a small change in just one price, results in a more similar environment than a big change in all prices.

We show here that, even if the information regarding similarity comparisons is partial, it is still possible to construct two sets that contain only S1 and S2 observations respectively, and that one consistency requirement of the data characterizes all DD processes. In order to do so, we assume

[33] A partial preorder is a reflexive and transitive binary relation. The Symmetric Difference between sets satisfies this assumption.

that, if the individual follows a DD process, the similarity σ cardinally represents a completion of such partial order. Thus, for any $e, e', g, g' \in E$, $(e, e') \succeq (g, g')$ implies $\sigma(e, e') \geq \sigma(g, g')$ and we say that (e, e') dominates (g, g'). As with the analysis of a forgetful DM, we first adapt the key concepts on which the algorithmic analysis is based in order to encompass this new assumption.

The two concepts of familiarity need to be adapted. In particular, given that it is not always possible to define the *most* familiar past environment, the new familiarity definitions will be sets containing undominated pairs of environments. Let $F(t)$ and $F(t|a_t)$ be defined as follows:

$$F(t) = \{(e_t, e_s) | s < t, a_s \in A_t \text{ and there is no } w < t \text{ such that } (e_t, e_w) \succeq (e_t, e_s) \text{ and } a_w \in A_t\},$$

$$F(t|a_t) = \{(e_t, e_s) | s < t, a_s = a_t \text{ and there is no } w < t \text{ such that } (e_t, e_w) \succeq (e_t, e_s) \text{ and } a_w \in A_t\}.$$

That is, $F(t)$ and $F(t|a_t)$ generalize the idea behind $f(t)$ and $f(t|a_t)$, respectively. In fact, $F(t)$ contains all those undominated pairs of environments where e_t is compared with past observations which choice *could be* replicated. Similarly, $F(t|a_t)$ contains all those undominated pairs of environments where e_t is compared with past observations which choice *could have been* replicated. We can easily redefine the concept of link. We say that observation t is *linked to the set* of observations O whenever either $F(t|a_t) = \emptyset$ or for all $(e_t, e) \in F(t|a_t)$ there exists $s \in O$ such that $(e_s, e') \succeq (e_t, e)$, for some $(e_s, e') \in F(s)$. Two things are worth underlining. First, notice that $F(t|a_t) = \emptyset$ only if t is new, thus, as in the main analysis, new observations are linked with any other observation. Second notice that this time we defined the link between an observation t and a *set* of observations O. This helps understand whether an observation is generated by S2 once we know that another observation is. If all observations in O are generated by S2 and for each one of them there exists a pair of environments that dominates a pair in $F(t|a_t)$ then it must be that S2 generated t too. This is because for all observation s in O, the similarity of *all* pairs of environments contained in $F(s)$ must be below the similarity threshold.

Then, we say that observation t is *S2-indirectly linked to* the set of observations O if there exists a sequence of observations $t_1, \ldots, t_k$ such that $t = t_1$, t_k is linked to O and t_i is linked to $\{t_{i+1}, t_{i+2}, ..., t_k\} \cup O$ for every $i = 1, 2, \ldots, k-1$. Define $D^{\hat{N}}$ as the set containing all new observations and all those observations indirectly linked to the *set* of new observations. Proposition 4 shows that $D^{\hat{N}}$ contains only S2 observations.

What about S1? As in section 4, whenever a cycle is present in the data, we know that at least one of the observations in the cycle must be generated by S1. This time, given that we assume only a partial knowledge of the similarity comparisons, it is not always possible to define a least novel observation in a cycle.[34] Nevertheless, notice that whenever an observation is inconsistent with the revealed preference constructed from $D^{\hat{N}}$, it must be that such observation is generated by S1. Thus, say that observation t is *cloned* if it is either a least novel in a cycle or $xR(t)y$ while

[34]Obviously, in this context, a least novel observation in a cycle would be an observation t belonging to a cycle such that for any other observation s in the cycle, $F(t)$ dominates $F(s)$. That is, for any $(e_s, e) \in F(s)$ there exists $(e_t, e') \in F(t)$ such that $(e_t, e') \succeq (e_s, e)$.

$yR(D^{\hat{N}})x$.

Say that observation t is *S1-indirectly linked to* observation s if there exists a sequence of observations $t_1, \ldots, t_k$ such that $t = t_1$, $t_k = s$ and t_i is linked to t_{i+1} for every $i = 1, 2, \ldots, k-1$. Whenever we know that observation t is generated by S1, we can infer that observation s is generated by S1 too, only if for all pairs of environments in $F(t|A_t)$ there exists a pair of environments in $F(s)$ that dominates it. In fact, in general, only the similarity of *some* pairs of environments contained in $F(t|a_t)$ is above the similarity threshold. As before, let $D^{\hat{C}}$ be the set containing all cloned observations and the observations to which they are indirectly linked. Proposition 4 below shows that $D^{\hat{C}}$ contains only S1 observations.

Proposition 4 *For every collection of observations D generated by a dual decision process where only a partial preorder over pairs of environments is known:*

1. *all decisions in $D^{\hat{N}}$ are generated by S2 and all decisions in $D^{\hat{C}}$ are generated by S1,*
2. *if x is revealed preferred to y for the set of observations $D^{\hat{N}}$, then $x \succ y$.*

Thus, we see that knowing only a partial preorder does not heavily affect the structure of the algorithm and the main logical steps behind it. What is of interest is that even with this assumption it is possible to characterize a DD process with just one single condition, that is $D^{\hat{N}}$-Consistency.

Axiom 4 ($D^{\hat{N}}$-Consistency) *A sequence of observations $\{(A_t, e_t, a_t)\}_{t=1}^{T}$ satisfies $D^{\hat{N}}$-Consistency whenever $xR(D^{\hat{N}})y$ implies not $yR(D^{\hat{N}})x$.*

$D^{\hat{N}}$-Consistency imposes asymmetry on the revealed preference obtained from $D^{\hat{N}}$. If a sequence of decision problems satisfies $D^{\hat{N}}$-Consistency when only a partial preorder is known, then we are able to characterize the preferences of the individual, the similarity threshold and, more importantly, a similarity function that respects such preorder. This is what the next theorem states. Notice that $\succeq$ is assumed to be known.

Theorem 4 *A sequence of observations $\{(A_t, e_t, a_t)\}_{t=1}^{T}$ satisfies $D^{\hat{N}}$-Consistency if and only if there exist a preference relation $\succ$, a similarity function σ representing $\succeq$ and a similarity threshold α that characterize a DD process.*

Intuitively, the observations in $D^{\hat{N}}$ are used to construct the preference relation of the individual. The similarity function represents a possible extension of the partial preorder that respects the absence of links between observations in $D^{\hat{N}}$ and the ones outside that set. This is possible thanks to how $D^{\hat{N}}$ has been constructed and it allows for the definition of the similarity threshold in a similar fashion as before.

7 Final Remarks

Cognitive sciences have highlighted the fact that choices can be divided into two categories. Conscious choices and intuitive ones. This hinders greatly the use of standard economic models that generally do not take into account automatic or intuitive choices.

In this paper, we make two main contributions. First, we propose a new behavioral model that incorporates the ideas coming from cognitive sciences where individual behavior is seen as the result of the interaction of two systems, one conscious and rule based, the other one unconscious and analogy based, being the latter the source of more intuitive choices. Intuitive choices are the result of replication of past behavior in those decision problems that are perceived as *familiar*, that is those problems that are similar enough with other ones that have been already experienced. Second, we study the implications of the model for observed behavior and we propose a general algorithm that allows to understand which choices are truly informative regarding individual preferences and to identify an interval in which the cognitive costs of activating the conscious system should lie. That is, we propose a novel way of restoring revealed preference analysis with behavioral datasets. Finally, we provide an axiomatic characterization of the model and we analyze some possible extensions.

The model of decision making we propose can be seen as a possible explanation of different phenomena and puzzles that are observed in different contexts. In particular it can be seen as a source of *stickiness*, i.e. inertia, in individual behavior. Whenever past behavior can be replicated, analogies between different decision problems make individual choices less responsive to changes in the quality and number of available options.[35] There is a lot of empirical evidence showing that individual behavior does not adapt immediately to changes in the economic environment. Consumption tends to be sticky as Carroll, Slacalek and Sommer (2011) show. Traders tend to show under-reaction to news and trading behavior is less responsive to market conditions, see for example Chan, Jegadeesh and Lakonishok (1996). Finally, doctors tend to stick to suboptimal treatments even when no other rational explanation can explain this behavior, e.g. see Hellerstein (1998). The model we propose allows for a formal analysis of these different frameworks in a simple and tractable environment. Take a doctor for example. If analogies are made between patients' symptoms, different patients can be treated with the same drug because they have similar enough symptoms for the doctor, even if such a treatment is suboptimal for the patients. If this is the case, older doctors, by having more experience, should be more prone to make this kind of mistakes, in line with the evidence in Hellerstein (1998).[36]

The model provides a novel way of understanding sticky behavior that does not depart too much from the standard framework used in economic modeling and that provides a natural bench-

[35] Notice in fact that whenever past choice can be replicated a necessary condition for behavior to adapt to changes in the menu is that the environment changes sufficiently enough. Clearly this decreases the responsiveness of choices to changes in menus.

[36] See King et al. (2013) and Norman and Eva (2010) for a general discussion of the problems in health care and the need of understanding how dual processes influence doctors behavior.

mark to formalize heterogeneity in behavior. In fact, the model allows for adaptive or conscious behavior and for sticky or intuitive one within the same population. Given a population, the aggregate choices will be less responsive to changes in the environment because a part of the population will make intuitive decisions but still there will be a reaction because of the individuals choosing consciously. Individuals with a low enough similarity threshold will stick to past behavior disregarding new possibilities. On the other hand, the rest of the population will react to the change in the environment adapting their choices. As a result, the aggregate choices adapt in a slower manner than what the standard framework would imply while individual choices vary from individual to individual. This kind of dynamics can be crucial for the efficacy of the implementation of a policy, e.g. change in the interest rate, and can lead to wrong predictions if they are not taken into account.

A Appendix

A.1 Figures

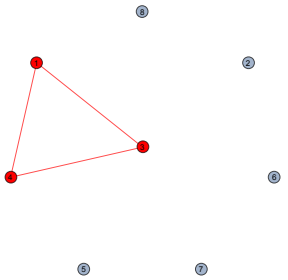

Figure 1: New observations are linked with each other.

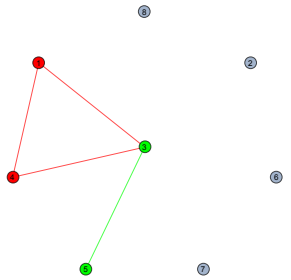

Figure 2: 5 is linked with 3

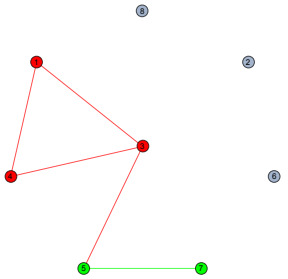

Figure 3: 7 is (indirectly) linked with (3) 5

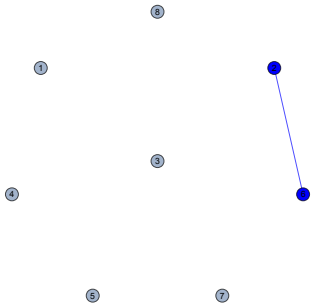

Figure 4: Least novel observations in a cycle

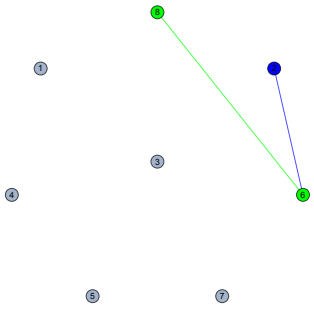

Figure 5: 6 is linked with 8

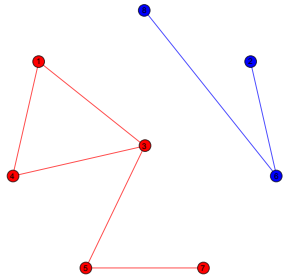

Figure 6: D^N and D^C.

A.2 Proofs

Proof of Proposition 1. We start by proving the statement regarding conscious observations. Trivially, new observations must be generated by S2 since they cannot replicate any past behavior. Consider an observation $t \in D^N$. By definition, there exists a sequence of observations $t_1, t_2, \ldots, t_n$ with $t_1 = t$, $f(t_i|a_{t_i}) \leq f(t_{i+1})$ for all $i = 1, 2, ..., n-1$ and t_n being new. We prove that t is generated by S2 recursively. We know that t_n is generated by S2. Now assume that t_k is generated by S2 and suppose by contradiction that t_{k-1} is generated by S1. From the assumption on t_k, we know that $f(t_k) \leq \alpha$. From the assumption on t_{k-1}, we know that $f(t_{k-1}|a_{t_{k-1}}) > \alpha$, which implies $f(t_{k-1}|a_{t_{k-1}}) > f(t_k)$, a contradiction with the hypothesis. Hence, t_{k-1} is also generated by S2, and the recursive analysis proves that observation t is generated by S2.

We now prove the statement regarding intuitive observations. Consider first an observation t which is a least novel in a cycle and assume by contradiction that it is generated by $S2$. Then, $f(t) \leq \alpha$. By definition of least novel in a cycle, it must be $f(s) \leq \alpha$ for every s in the cycle, making all decisions in the cycle being generated by $S2$. This is a contradiction with the maximization of a preference relation. Consider now an observation $t \in D^C$. By definition, there exists a sequence of observations $t_1, t_2, \ldots, t_n$ with $t_n = t$, $f(t_i|a_{t_i}) \leq f(t_{i+1})$ for all $i = 1, 2, ..., n-1$ and t_1 being a least novel in a cycle. We proceed recursively again. Since t_1 is generated by S1, we have $f(t_1|a_{t_1}) > \alpha$. Now assume that t_k is generated by S1 and suppose by contradiction that t_{k+1} is generated by S2. We then know that $f(t_k|a_{t_k}) > \alpha \geq f(t_{k+1})$, which is a contradiction concluding the recursive argument.

For the revelation of preferences part, since D^N can only contain observations generated by S2, it is straightforward to see that the revealed information from such a set must respond to the preferences of the DM. Regarding α, notice that since observations in D^N are generated by S2, we know that $\max_{t \in D^N} f(t) \leq \alpha$ and also that, since observations in D^C are generated by S1, we know that $\alpha < \min_{t \in D^C} f(t|a_t)$, which concludes the proof. ■

Proof of Theorem 1. Necessity: If D is generated by a DD process, then it satisfies Link-Consistency as explained in the text.

Sufficiency: Now suppose that D satisfies Link-Consistency. We need to show that it can be explained as if generated by a DD process. We first show that Link-Consistency implies that the revealed preference relation defined over D^N, i.e. $R(D^N)$, is asymmetric. Asymmetry of $R(D^N)$ means that it is not possible to construct cycles composed by observations in D^N. Suppose by contradiction that we have a cycle in D^N. That is, there is a set of observations $C = \{t_1, t_2, \ldots, t_k\} \subseteq D^N$ such that $a_{t_{i+1}} \in A_{t_i}$, $i = 1, \ldots, k-1$ and $a_{t_1} \in A_{t_k}$. Take the observation in the cycle with the highest unconditional familiarity. Denote it with t_{i^*}. Then all the other observations in the cycle are linked to t_{i^*}, that is, $C \subseteq L(t_{i^*})$, contradicting Link-Consistency. Thus, $R(D^N)$ must be asymmetric. By standard mathematical results, we can find a transitive completion of $R(D^N)$, call it $\succ$. By construction, all decisions in D^N can be seen as the result of

maximizing $\succ$ over the corresponding menu.

Define $\alpha = \max_{t \in D^N} f(t)$. Notice that by definition of D^N, there is no observation $s \notin D^N$ such that $f(s|a_s) \leq f(t)$ for some $t \in D^N$. This implies that for all $s \notin D^N$, $f(s|a_s) > \alpha$, so, for all of them, it is possible to find a preceding observation they would seem to replicate. In particular, the one defining $f(s|a_s)$.

Thus, we can represent the choices as if generated by an individual with preference relation $\succ$ and similarity threshold α. ■

Proof of Theorem 2. Necessity: As said in the text, if all DM in the population follow a DD process with common preferences, similarity function and similarity threshold, then $I(t)$ would contain only those choices that can be replicated at t because they come from some preceding period which environment is similar enough. Then, by definition of a DD process CC and IC must be satisfied.

Sufficiency: Suppose that $\mathcal{D}$ is rich and satisfies CC and IC. We prove that $\mathcal{D}$ can be represented as if generated by a DD process by steps. First we characterize the preference relation, then we characterize the similarity function and similarity threshold and finally we show that the preference relation and the binary similarity function defined by the combination of similarity function and similarity threshold are unique.

As a first step, let P be a revealed preference relation defined as follows. For any $D \in \mathcal{D}$, let xPy if and only if, for some $t \in D$ such that $I(t) = \emptyset$ and $A_t = \{x, y\}$, $x = a_t$. It is easy to see that CC implies that P is irreflexive and asymmetric. Furthermore, richness of $\mathcal{D}$ implies that the relation is also complete. To see that P is transitive, suppose that xPy, yPz but zPx for some $x, y, z \in X$. By $\mathcal{D}$ being rich, for some $D \in \mathcal{D}$ there exists an observation $t \in D$ such that $x, y, z \neq a_s$ for $s < t$ and $A_t = \{x, y, z\}$. Clearly, given that x, y and z have never been chosen before, $I(t) = \emptyset$. W.l.o.g. suppose that $a_t = x$. By zPx we know that $z = a_{t'}$ for some $t' \in D' \in \mathcal{D}$ such that $A_{t'} = \{x, z\}$ and $I(t') = \emptyset$. But then, by CC we cannot have $a_t = x$ and so we reached a contradiction. Hence P must be transitive. Let $P = \succ$.

As a second step we need to characterize the similarity threshold α and the similarity function σ. Let $\alpha = 0$ and the similarity function be as follows. For any $e \in E$:

- $\sigma(e, e') = 1$ whenever $e' \in S(e)$.
- $\sigma(e, e') = 0$ whenever $e' \notin S(e)$.

The third requirement of a rich dataset assures that if $e' \notin S(e)$ then no inversion of preferences is observed due to replication of behavior associated with e' when e is the reference environment.

We now show that constructing the preference relation and the similarity function in this way allows to explain all choices. In fact, the definition of the similarity function and threshold implies that any observation t such that $I(t) = \emptyset$ must be explained as the outcome of maximization of preferences and any other observation t' such that $I(t') \neq \emptyset$ must be the outcome of replication of past similar behavior. To analyze the first point, take any observation t such that $I(t) = \emptyset$.

Suppose that x maximizes $\succ$ in the menu A_t, but $y = a_t$ for some $t \in D \in \mathcal{D}$ and $y \neq x$. By definition of P, x being the maximal element in A_t implies that for any $y \in A_t \setminus \{x\}$, there exists $t' \in D' \in \mathcal{D}$ such that $A_{t'} = \{x, y\}$ and $a_{t'} = x$. Thus, by CC, we cannot have $a_t = y$ and so we reached a contradiction. Thus, we can identify the preference relation with P. Regarding the second point, take any observation t' such that $I(t') \neq \emptyset$. Then, there is some $s < t'$ such that $a_s \in A_{t'}$ and $\sigma(e_{t'}, e_s) = 1 > \alpha$ and then IC assures that the choice in t' is the same than the choice in one of such past problems as a DD process would require.

Finally, we need to show that the preference relation and the binary similarity function are uniquely identified. This is equivalent to show that P and $S(e)$ are uniquely determined. First, suppose that we can determine two different relations P and P'. This implies that there exist some $x, y \in X$ such that xPy and $yP'x$. Given that by richness P andP' must be complete, the previous binary relations imply that there are two observations t and t' such that $I(t) = \emptyset$ and $A_t = \{x, y\}$, $x = a_t$ and $I(t') = \emptyset$ and $A_{t'} = \{x, y\}$, $y = a_{t'}$ which contradicts CC, so P is unique. Second, suppose that for some $e \in E$, $S(e)$ is not uniquely determined. That is, suppose we can determine two different sets $S(e)$ and $S'(e)$. This implies that there is an $e' \in E$ such that $e' \in S(e)$ and $e' \notin S'(e)$. By definition of $S(e)$ this is not possible. In fact, $e' \in S(e)$, given richness, can only imply that there are two alternatives $x, y \in X$ such that xPy and that there is some observation t where $x, y \neq a_s$ for $s < t - 1$ and $t - 1 = \{\{y\}, e', y\}$ and $t = \{\{x, y\}, e, y\}$, then e' must be in $S'(e)$ too. ■

Proof of Proposition 3. Given Proposition 2, we just need to show the following:

1. If t is generated by S2 then it is contained in $D^{\overline{N}}$.
2. If t is generated by S1 then it is contained in $D^{\overline{C}}$.
3. If $x \succ y$ then $xR(D^{\overline{N}})y$.

We start with point 3. Notice that, by m-richness, for any pair of alternatives $x, y \in X$ there exists a t such that $x, y \neq a_s$ for $t - m \leq s < t$ and $A_t = \{x, y\}$. Thus, for any pair of alternative there is a new observation that reveals the preferences of the DM and the result follows.

Now we analyze point 1. First notice that for any DD-m process there exists a pair of environments (e^*, e^{**}) for which there is no $(e, e') \in E \times E$ such that $\sigma(e^*, e^{**}) < \sigma(e, e') \leq \alpha$. That is, e^* and e^{**} maximize the value of the similarity function among those pairs of environments that that are considered *dissimilar enough* by the DM. Take $x, y \in X$. By the proof of point 3, we know that for any pair of alternatives we can determine the preference of the DM, thus we can assume w.l.o.g. that $x \succ y$. Then, by m-richness there exists a t such that $x, y \neq a_s$ for $t - 1 - m \leq s < t - 1$ and $A_{t-1} = \{y\}$, $e_{t-1} = e^{**}$, $A_t = \{x, y\}$ and $e_t = e^*$. Given $\sigma(e^*, e^{**}) \leq \alpha$, it must be $a_t = x$ making t a new observation and so $t \in D^{\overline{N}}$. Hence, given that only behavior in $t - 1$ could be replicated in t, we must have $\overline{f}(t) = \sigma(e^*, e^{**})$ and so, by definition, for any observation s generated by S2 we must have $\overline{f}(s|a_s) \leq \overline{f}(t) = \sigma(e^*, e^{**})$. Thus, all S2 observations are linked to t, and hence any S2

observation must be in $D^{\overline{N}}$. By a dual argument point 2 can be shown to be valid. This concludes the proof. ■

Proof of Proposition 4. First we show that any observation t linked to a set O of S2 observations must be generated by S2 too. In fact, notice that for any $s \in O$ we know that $\sigma(e_s, e') \leq \alpha$ for all $(e_s, e') \in F(s)$. Then, given t is linked to O we know that for any $(e_t, e) \in F(t|a_t)$ there exists $s \in O$ such that $(e_s, e') \succeq (e_t, e)$, for some $(e_s, e') \in F(s)$. Now, given the definition of $F(t|a_t)$ this implies that $\sigma(e_t, e_w) \leq \alpha$ for all $w < t$ such that $a_w = a_t$ and the result follows. Then, by Proposition 1 we know that new observations are generated by S2 and applying the previous reasoning iteratively it is shown that $D^{\hat{N}}$ must contain only S2 observations.

As a second step, we show that any observation s to which an observation t generated by S1 is linked, must be generated by S1 too. Given t is generated by S1 it means that there exists a $w < t$ such that $\sigma(e_t, e_w) > \alpha$ and $a_w = a_t$. Then, either $(e_t, e_w) \in F(t|a_t)$ or $(e_t, e_w) \notin F(t|a_t)$.

- Let $(e_t, e_w) \in F(t|a_t)$. Then, given t is linked to s, there exists a pair $(e_s, e') \in F(s)$ such that $(e_s, e') \succeq (e_t, e_w)$. This implies $\sigma(e_s, e') \geq \sigma(e_t, e_w) > \alpha$, and the result follows.

- Let $(e_t, e_w) \notin F(t|a_t)$. Then, there exists a $w' < t$ such that $(e_t, e_{w'}) \succeq (e_t, e_w)$ and $(e_t, e_{w'}) \in F(t|a_t)$. This implies that $\sigma(e_t, e_{w'}) > \sigma(e_t, e_w) > \alpha$. Then, given t is linked to s we know that for all $(e_t, e) \in F(t|a_t)$, there exists a $(e_s, e') \in F(s)$ such that $(e_s, e') \succeq (e_t, e)$. In particular, there exists a $(e_s, e') \in F(s)$ such that $(e_s, e') \succeq (e_t, e_{w'})$. This implies $\sigma(e_s, e') \geq \sigma(e_t, e_{w'}) > \alpha$, and the result follows.

Then, given that cloned observations are generated by S1, we can apply the previous reasoning iteratively to show that $D^{\hat{C}}$ must contain only S1 observations.

Finally, by a reasoning similar to the one developed in the proof of Proposition 1, given all observations in $D^{\hat{N}}$ must be generated by S2, $R(D^{\hat{N}})$ reveals the preference of the DM. ■

Proof of Theorem 4. Necessity: Suppose that the sequence $\{(A_t, e_t, a_t)\}_{t=1}^{T}$ is generated by a DD process. Then it satisfies $D^{\hat{N}}$-Consistency given that, according to Proposition 4, $D^{\hat{N}}$ contains only S2 observations and $\succ$ is a linear order.

Sufficiency: Suppose that the sequence $\{(A_t, e_t, a_t)\}_{t=1}^{T}$ satisfies $D^{\hat{N}}$-Consistency. We need to show that it can be explained as if generated by a DD process. Notice that $D^{\hat{N}}$-Consistency implies that the revealed preference relation defined over $D^{\hat{N}}$, i.e. $R(D^{\hat{N}})$, is asymmetric. Thus, by standard mathematical results, we can find a transitive completion of $R(D^{\hat{N}})$, call it $\succ$. By construction, all decisions in $D^{\hat{N}}$ can be seen as the result of maximizing $\succ$ over the corresponding menu.

We now define σ. We first complete $\succeq$. Notice that by construction of $D^{\hat{N}}$, for all $t \notin D^{\hat{N}}$ there exists a pair $(e_t, e) \in F(t|a_t)$ such that there is no $s \in D^{\hat{N}}$ for which $(e_s, e') \succeq (e_t, e)$, for some $(e_s, e') \in F(s)$. That is, for all observations not in $D^{\hat{N}}$ there exists a pair of environments that is not dominated by any pair of environments of observations in $D^{\hat{N}}$, a pair that we call

undominated. Then, let $\succeq'$ be the following reflexive binary relation. For any undominated pair $(e_t, e) \in F(t|a_t)$ with $t \notin D^{\hat{N}}$, let for all $s \in D^{\hat{N}}$ and for all $(e_s, e') \in F(s)$, $(e_t, e) \succeq' (e_s, e')$ and not $(e_s, e') \succeq' (e_t, e)$. Let $\succeq''$ be the transitive closure of $\succeq \cup \succeq'$. Notice that $\succeq''$ is an extension of $\succeq$ that preserves its reflexivity and transitivity. Thus we can find a completion $\succeq^*$ of $\succeq''$ and a similarity function $\sigma : E \times E \to [0, 1]$ that represents $\succeq^*$.

Finally, we can define α. For any observation t, let $f^*(t)$ be as follows:

$$f^*(t) = \max_{s<t, a_s \in A_t} \sigma(e_t, e_s),$$

Then let $\alpha = \max_{t \in D^{\hat{N}}} f^*(t)$. Notice that by construction of σ for all $t \notin D^{\hat{N}}$ there exists a pair of environments $(e_t, e) \in F(t|a_t)$ such that for all $s \in D^{\hat{N}}$, $\sigma(e_t, e) > f^*(s)$, hence $\sigma(e_t, e) > \alpha$. So, for every observation not in $D^{\hat{N}}$ we can find a preceding observation to imitate.

Thus, we can represent the choices as if generated by an individual with preference relation $\succ$, similarity function σ and similarity threshold α. ■

A.3 Estimation of the Similarity Function

The similarity function is a key component of a DD process and for the sake of exposition it is assumed to be known in the main part of the paper. Nonetheless, we discuss here how to identify it by studying the choice behavior of a group of individuals sharing it.

Consider a continuous population of individuals sharing the similarity function σ, with a continuous and independent distribution of the similarity threshold over $[0, 1]$. Sequences of decision problems and preferences are independently distributed.

Consider a pair of alternatives $x, y \in X$ such that each of them is considered better than the other for a non-negligible part of the population. For every pair of environments $e, e' \in E$, assume there exists a non-negligible subpopulation for which there is an observation t as follows:

- $x \notin A_t$, $A_{t+1} = A_t \cup \{x\}$ and no alternative in A_{t+1} was chosen before t,
- $e_t = e'$ and $e_{t+1} = e$ and
- $a_t = y$.

The main result of this section shows that we can compare the similarity of two different pairs of environments by considering the corresponding aforementioned subpopulations and sampling them. Formally, denote by $\nu(e, e')$ the average relative number of randomly sampled individuals sticking to y at $t + 1$. That is, for any pair of environments (e, e') we take a sample of finite magnitude n from the aforementioned subpopulations and we compute the average of the relative number of individuals that stick to y in such sample. Such average is $\nu(e, e')$.

Proposition A1 (Eliciting the Similarity) *For every two pairs of environments (e, e') and (g, g'), $Pr(\nu(e, e') \geq \nu(g, g') | \sigma(g, g') > \sigma(e, e')) \xrightarrow{p} 0$. That is, the probability of having $\nu(e, e') \geq \nu(g, g')$ when $\sigma(g, g') > \sigma(e, e')$ probabilistically converges to zero.*

Proof. Let $\mu(e, e')$ be the relative number of individuals that would choose y *in the whole non-negligible subpopulation* from which the sample defining $\nu(e, e')$ is taken. First, notice that whenever $\mu(e, e') \geq \mu(g, g')$ it implies that environments (e, e') are more similar than environments (g, g'). In fact, given that preferences and decision problems are independently distributed in the population, any of the subpopulations we consider is representative of the whole population. Then, given that in the whole population no alternative is preferred over the other by every individual, if the actual relative numbers are different it implies that one pair of environments is more similar than the other. Moreover, given α is continuously and independently distributed in the population we can compare any pair of environments with any other pair and find how they are related by the similarity function. That is, for any $e, e', g, g' \in E$ such that $\sigma(e, e') < \sigma(g, g')$ it must be $\mu(e, e') < \mu(g, g')$ given that there always exists a non-negligible part of the whole population with similarity threshold α in the interval $[\sigma(e, e'), \sigma(g, g'))$ and again, the subpopulations are representative. Then, notice that given that every subpopulation is infinite, any sample that defines $\nu(e, e')$ for any $e, e' \in E$ can be considered independent. Thus, the law of large numbers applies in this context and we get,

$$Pr(|\nu(e, e') - \mu(e, e')| > \epsilon) \xrightarrow{p} 0,$$

and the result follows. ■

Given our assumptions, each sample we take to calculate $\nu(e, e')$ is an independent estimation of the relative number of individuals that sticks to y. Thus, we are getting a consistent estimate of the relative number of individuals that would stick to y in the whole subpopulation. Then, comparing the average relative number of individuals that stick to alternative y in $t + 1$ for different pairs of environments, gives the required information on the similarity function. The main intuition of Proposition A1 is the following. There are two reasons that force the average relative number of people sticking to y with the pair of environments (e, e') to be bigger than the average relative number of people sticking to y with another pair of environments (g, g'). First, the pair of environments (e, e') is more similar than the pair of environments (g, g'). This implies that S1 is active in $t+1$ for a larger number of individuals on average, which leads to replication of the choice in t, i.e. alternative y.[37] Second, y is preferred to x by a larger number of sampled people among those using S2 with the pair of environments (e, e'). The law of large numbers makes the second concern disappear as the number of samples taken to measure ν grows, revealing the similarity of different pairs of environments.

[37] Notice that our assumptions imply that only the choice in t can be replicated if S1 is active.

References

Alter, Adam L, Daniel M Oppenheimer, Nicholas Epley, and Rebecca N Eyre. 2007. "Overcoming Intuition: Metacognitive Difficulty Activates Analytic Reasoning." *Journal of Experimental Psychology: General*, 136(4): 569.

Ambrus, Attila, and Kareen Rozen. 2015. "Rationalising Choice with Multi-self Models." *The Economic Journal*, 125(585): 1136–1156.

Apesteguia, Jose, and Miguel A Ballester. 2015. "A Measure of Rationality and Welfare." *Journal of Political Economy*, 123(6): 1278–1310.

Bernheim, B.Douglas, and Antonio Rangel. 2009. "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics." *The Quarterly Journal of Economics*, 124(1): 51–104.

Carroll, Christopher D, Jiri Slacalek, and Martin Sommer. 2011. "International Evidence on Sticky Consumption Growth." *Review of Economics and Statistics*, 93(4): 1135–1145.

Chan, Louis KC, Narasimhan Jegadeesh, and Josef Lakonishok. 1996. "Momentum Strategies." *Journal of Finance*, 51(5): 1681–1713.

Chetty, Raj. 2015. "Behavioral Economics and Public Policy: A Pragmatic Perspective." *American Economic Review P&P*, 105(5): 1–33.

Cunningham, Tom. 2015. "Hierarchical Aggregation of Information and Decision-Making." *https://dl.dropboxusercontent.com/u/13046545/paper_heuristics.pdf*.

Cunningham, Tom, and Jonathan de Quidt. 2015. "Implicit Preferences Inferred from Choice." *Unpublished*.

Diemand-Yauman, Connor, Daniel M Oppenheimer, and Erikka B Vaughan. 2011. "Fortune Favors the **Bold** (and *Italicized*): Effects of Disfluency on Educational Outcomes." *Cognition*, 118(1): 111–115.

Duhigg, Charles. 2012. *The Power of Habit: Why We Do What We Do in Life and Business.* New York:Random House.

Epstein, Seymour. 1994. "Integration of the Cognitive and the Psychodynamic Unconscious." *American Psychologist*, 49(8): 709.

Evans, Jonathan S.B.T. 1977. "Toward a Statistical Theory of Reasoning." *Quarterly Journal of Experimental Psychology*, 29: 297–306.

Evans, Jonathan S.B.T. 1989. *Bias in Human Reasoning: Causes and Consequences.* Hove and London:Lawrence Erlbaum Associates, Inc.

Evans, Jonathan S.B.T., and David E Over. 1996. *Reasoning and Rationality.* New York:Psychology Press.

Evans, Jonathan S.B.T., and Keith E. Frankish. 2009. *In Two Minds: Dual Processes and Beyond.* Oxford:Oxford University Press.

Fudenberg, Drew, and David K Levine. 2006. "A Dual-Self Model of Impulse Control." *American Economic Review*, 96(5): 1449–1476.

Gilboa, Itzhak, and David Schmeidler. 1995. "Case-Based Decision Theory." *The Quarterly Journal of Economics*, 110(3): 605–639.

Gilboa, Itzhak, and David Schmeidler. 2001. *A Theory of Case-Based Decisions.* Cambridge:Cambridge University Press.

Gul, Faruk, and Wolfgang Pesendorfer. 2001. "Temptation and Self-Control." *Econometrica*, 69(6): 1403–1435.

Hahn, Ulrike. 2014. "Similarity." *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(3): 271–280.

Hellerstein, Judith K. 1998. "The Importance of the Physician in the Generic Versus Trade-Name Prescription Decision." *RAND Journal of Economics*, 29(1): 108–136.

Kahneman, Daniel. 2002. "Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice." *Nobel Prize Lecture*, 8: 351–401.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow.* New York:Farrar, Straus and Giroux.

Kalai, Gil, Ariel Rubinstein, and Ran Spiegler. 2002. "Rationalizing Choice Functions by Multiple Rationales." *Econometrica*, 70(6): 2481–2488.

King, Dominic, Felix Greaves, Ivo Vlaev, and Ara Darzi. 2013. "Approaches Based on Behavioral Economics Could Help Nudge Patients and Providers Toward Lower Health Spending Growth." *Health Affairs*, 32(4): 661–668.

Manzini, Paola, and Marco Mariotti. 2007. "Sequentially Rationalizable Choice." *American Economic Review*, 97(5): 1824–1839.

Manzini, Paola, and Marco Mariotti. 2014. "Welfare Economics and Bounded Rationality: The Case for Model-Based Approaches." *Journal of Economic Methodology*, 21(4): 343–360.

Masatlioglu, Yusufcan, Daisuke Nakajima, and Erkut Y Ozbay. 2012. "Revealed Attention." *American Economic Review*, 102(5): 2183–2205.

McAndrews, Mary Pat, Elizabeth L Glisky, and Daniel L Schacter. 1987. "When Priming Persists: Long-lasting Implicit Memory for a Single Episode in Amnesic Patients." *Neuropsychologia*, 25(3): 497–506.

Medin, Douglas L, Robert L Goldstone, and Dedre Gentner. 1993. "Respects for Similarity." *Psychological Review*, 100(2): 254.

Norman, Geoffrey R, and Kevin W Eva. 2010. "Diagnostic Error and Clinical Reasoning." *Medical education*, 44(1): 94–100.

Oppenheimer, Daniel M. 2008. "The Secret Life of Fluency." *Trends in Cognitive Sciences*, 12(6): 237–241.

Reber, Arthur S. 1989. "Implicit Learning and Tacit Knowledge." *Journal of Experimental Psychology: General*, 118(3): 219.

Rubinstein, Ariel. 2007. "Instinctive and Cognitive Reasoning: A Study of Response Times." *The Economic Journal*, 117(523): 1243–1259.

Rubinstein, Ariel. 2015. "A Typology of Players: Between Instinctive and Contemplative." $http://arielrubinstein.tau.ac.il/papers/Typology.pdf$.

Rubinstein, Ariel, and Yuval Salant. 2012. "Eliciting Welfare Preferences from Bheavioral Dataset." *Review of Economic Studies*, 79: 375–387.

Salant, Yuval, and Ariel Rubinstein. 2008. "(A, f): Choice with Frames." *The Review of Economic Studies*, 75(4): 1287–1296.

Schneider, Walter, and Richard M. Shiffrin. 1977. "Controlled and Automatic Human Information Processing I: Detection, Search and Attention." *Psychological Review*, 84(1): 1–66.

Simon, Herbert A. 1987. "Making Management Decisions: The Role of Intuition and Emotion." *The Academy of Management Executive (1987-1989)*, 57–64.

Stanovich, Keith E. 1999. *Who is Rational?: Studies of Individual Differences in Reasoning.* New York:Psychology Press.

Strotz, Robert H. 1955. "Myopia and Inconsistency in Dynamic Utility Maximization." *The Review of Economic Studies*, 23(3): 165–180.

Tversky, Amos. 1977. "Features of Similarity." *Psychological Review*, 84(4): 327.