



**Universitat
Pompeu Fabra**
Barcelona

Department
of Economics and Business

Economics Working Paper Series

Working Paper No. 1673

**Dual decision processes: retrieving
preferences when some choices are
automatics**

Francesco Cerigioni

October 2019

DUAL DECISION PROCESSES:

RETRIEVING PREFERENCES WHEN SOME CHOICES ARE AUTOMATIC

Francesco Cerigioni*
Universitat Pompeu Fabra and Barcelona GSE

October 11, 2019

ABSTRACT

Evidence from the cognitive sciences suggests that some choices are conscious and reflect individual volition while others tend to be automatic, being driven by analogies with past experiences. Under these circumstances, standard economic modeling might not always be applicable because not all choices are the result of individual tastes. We propose a behavioral model that can be used in standard economic analysis that formalizes the way in which conscious and automatic choices arise by presenting a decision maker comprised of two selves. One self compares past decision problems with the one the decision maker faces and, when the problems are similar enough, it replicates past behavior (Automatic choices). Otherwise, a second self is activated and preferences are maximized (Conscious choices). We then present a novel method capable of identifying a set of conscious choices from observed behavior and discuss its usefulness as a framework for studying asymmetric pricing and empirical puzzles in different settings.

(JEL D01, D03, D60)

KEYWORDS: Dual Processes, Similarity, Revealed Preferences, Fluency, Automatic Choice

*Department of Economics and Business, Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, 08005, Barcelona, Spain. E-mail: francesco.cerigioni@upf.edu. The paper has previously circulated under the title “Dual Decision Processes: Retrieving Preferences when some Choices are Intuitive”. I am grateful to the Ministerio de Ciencia y Innovacion for the grant PGC2018-098949-B-I00 and financial support through the Severo Ochoa Programme for Centers of Excellence in R&D SEV-2015-0563. I thank Miguel A. Ballester for his insightful comments that were essential for the development of this paper and the editor Joel Sobel and four anonymous referees for their critiques and suggestions that greatly improved this work. 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, Larbi Alaoui, Jose Apesteguia, B. Douglas Bernheim, Caterina Calsamiglia, Andrew Caplin, Laurens Cherchye, Vincent Crawford, Tugce Cuhadaroglu, Jan Eeckhout, 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, Sandro Shelegia, John Tsoukalis, Peio Zuazo-Garin and participants at SITE, BRIC and D-TEA for their comments and suggestions. All mistakes are mine.

1 Introduction

Behavioral economics has posed a serious challenge to standard economic theory by documenting and analysing numerous behaviors inconsistent with preference maximization. Inconsistencies tend to emerge when individuals do not exert the effort to consciously analyze the problem at hand, e.g., Carroll et al. (2009). A key question then arises. When are people choosing consciously?

This paper provides a first theoretical framework in which to consider the co-existence of conscious and automatic, i.e., unconscious, behavior and a means to understand the nature of individual decisions in different situations once this duality is taken into account. Studying this dichotomy is crucial to the understanding of market outcomes. In fact, as highlighted by Simon (1987) and Kahneman (2002), *experts* such as managers, doctors, traders or policy makers often make automatic decisions. Nevertheless, the sources of automatic choices remain underexplored in economics.¹

Evidence from the cognitive sciences suggests that the *familiarity* of the choice environment is the key determinant of the relationship between conscious and automatic choices. Unconscious, automatic choices are made in familiar environments. Thus, if we wish to understand the role of automatic decisions in markets, we need (i) to create a model of automatic choices that takes into account the role played by the familiarity of the decision environment and (ii) to analyze whether it is possible from observed behavior to distinguish between conscious and automatic choices in order to understand the underlying preferences. These research questions will be addressed in this paper.

To answer the first question, we propose a simple formalization of evidence drawn from the cognitive sciences, the main contribution of which is to provide a first model describing when and how choices should be conscious or automatic.² In Section 2, we present a decision maker described by a simple procedure. Whenever the decision

¹Although there is a growing attention in economics to conscious and intuitive or automatic reasoning. See for example Rubinstein (2007) and Rubinstein (2016) for a distinction between conscious and intuitive strategic choices by players of a game or the distinction in Cunningham and de Quidt (2015) between implicit and explicit attitudes.

²See section 6.2 for the cognitive foundations of the model.

environment is familiar, i.e., whenever its similarity with past experience, measured by a similarity function, passes a certain threshold, past behavior is replicated. This is the source of *automatic* choices. Otherwise, the best option is chosen by maximizing a rational preference relation. This is the source of *conscious* choices. Consider, for example, a consumer who buys a bundle of products from the shelves of a supermarket. The first time he faces the shelves, he tries to find the bundle that best fits his preferences. Subsequently, if the prices and arrangement of products do not change *too much*, he will perceive the decision environment as familiar and will automatically stick to the bundle chosen previously. If, on the other hand, the change in prices and arrangement of the products is evident to him, his choice will again be based on preference maximization.³

Even in such a simple framework, there is no trivial way to distinguish which choices are made automatically and which are made consciously. Following the example, suppose our consumer is faced once more with the same problem, but this time, although a new bundle is available, he sticks with the original choice. Is it because he prefers the original bundle to the new one? Or is it because he is choosing automatically?

We show how to find conscious choices and hence restore the standard revealed preference analysis by being able to tell which environments were not familiar. Section 3 assumes (i) that the decision maker behaves according to our model and (ii) that the similarity function is known, while the threshold is not.⁴ We then show that, for every sequence of decision problems, it is possible by means of an algorithm to identify a set of conscious observations and an interval within which the similarity threshold should lie. That is, we provide a novel method for restoring revealed preference analysis.

First notice that *new* observations, i.e., those in which the choice is for an alternative never chosen before, must be conscious. No past behavior could have been replicated. Starting from these observations, the algorithm iterates the following idea. If an observation is consciously generated, any other *less familiar* observation, that

³This is in line with the ideas presented in Woodford (2019) where it is stated that “... *people should not be modeled as behaving differently in situations that they do not recognize as different.*...”

⁴See Sections 3 and 5 for a justification of the latter hypothesis. See the appendix for an empirical strategy for estimating the similarity function when shared in a population.

is any decision problem which is less similar to those that preceded it, must be also consciously generated. 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 also have done so on all those occasions where the change was even more evident.

The algorithm identifies a set of automatic decisions in a similar fashion, that is, it first highlights some decisions that must be automatic and then finds *more familiar* observations to reveal other automatic decisions. Notice that knowing whether some decisions were made automatically is very important for understanding the way in which familiarity of environments is determined. Even if automatic choices do not reveal individual preferences, they tell us which problems are considered familiar, i.e., similar enough, by the decision maker, hence enabling the identification of the interval in which the similarity threshold should lie.

The algorithm assumes that the decision maker behaves according to our model, hence the falsifiability of the model becomes a central concern. In Section 4 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. Section 5 contains a generalization of the identification algorithm and of the characterization, assuming only partial knowledge of similarity comparisons. Section 6 clarifies the connection between the model and methods introduced here and the economics and cognitive sciences literatures. Section 7 uses the model as a framework for understanding asymmetric pricing in markets. Section 8 concludes. The appendix describes a procedure for empirically estimating the similarity function when shared by a population of otherwise heterogeneous agents.

2 Dual Decision Processes

2.1 The Model

Let X and E be two sets. The decision maker (DM) faces, at time t , a *decision problem* (A_t, e_t) with $A_t \subseteq X$ and $e_t \in E$. The set of available alternatives A_t at time t , from which the DM has to make a choice, is called the *menu*. An alternative is any choice element, such as a consumption bundle, a lottery or a consumption

stream. The *environment* e_t is a description of the possible characteristics of the problem faced by the DM at time t . As highlighted below, an environment can be the menu, the set of attributes of the alternatives contained in the menu, the framing, etc. We denote by $a_t \in A_t$ the chosen alternative at time t . With a little abuse of the notation, we refer to the pair 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. $D = \{1, \dots, T\}$. Notice that the same menu or environment can appear more than once in the sequence.

A chosen alternative is the outcome of a two-stage choice procedure which describes the DM and formalizes the duality of automatic and conscious choices. Formally, let $\sigma : E \times E \rightarrow [0, 1]$ be the *similarity function*. The value $\sigma(e, e')$ measures the degree of similarity between environment e and environment e' . The *automatic self* 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 faced with the decision problem (A_t, e_t) , the automatic self 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. The *maximizing self* is endowed with a *preference relation* \succ over the set of alternatives. For ease of exposition, we assume that \succ is a strict order, i.e. an asymmetric, transitive and complete binary relation, defined over X . At time t , if the maximizing self is activated, it chooses the alternative a_t that maximizes \succ in A_t .⁵ Summarizing:

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

Three remarks are useful here. First, notice that automatic and conscious decisions are separated by the behavioral parameter α . In some sense, α is summarizing the cost of using the cognitively demanding system. The higher the cost, the lower

⁵The idea that conscious behavior arises from the maximization of a preference relation is a simplification which 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 8.

the threshold. Thus, parameter α captures individual heterogeneity as preferences do. Notice that, while the similarity function has been widely studied in cognitive sciences, e.g. Tversky (1977), Medin et al. (1993) and Hahn (2014), the cognitive costs of activating conscious decision processes remain an unknown quantity. Therefore, the method we propose in section 3 can be seen as a first attempt to identify, among observed behavior, the interval in which such costs should lie, given the similarity function.⁶

As a second remark, notice that we are describing a class of models because we do not impose any particular structure on replicating behavior. We do not specify which alternative will be chosen when more than one past choice can be replicated. This class can accommodate many different behaviors, e.g., choosing the alternative that was chosen in the most similar past environment, or choosing the alternative that maximizes the preference relation over those chosen in similar enough past environments, etc. All of the following analysis is valid for the class as a whole.

As a final remark, below, for illustrative purposes, we propose some examples of relevant environments for economic applications and a possible similarity function for use in such cases.

Environments as Menus: *In many economic applications it seems sensible to view the entire 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 to be as similar as their menus are. In this framework, $E = 2^X \setminus \{\emptyset\}$.*

Environments as Attributes: *The alternatives faced by decision makers are often bundles of attributes. In such contexts, it is reasonable to assume that the attributes of the available alternatives determine the decision environment. If \mathcal{A} is the set containing all possible attributes, then $E = 2^{\mathcal{A}} \setminus \{\emptyset\}$.*

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, the way the products are displayed on a shelf.*

⁶There is growing attention in decision theory to the empirical estimation of cognitive parameters for richer theoretical models, e.g., Cerreia-Vioglio et al. (2012), Rustichini et al. (2016), and Dardanoni et al. (2018).

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 $E = 2^F \setminus \{\emptyset\}$.

The cognitive sciences provide different ways to model the similarity function based on empirical evidence. In all the previous examples it is natural to assume that the similarity function relates different environments according to their commonalities and differences. For example, $\sigma(e, e') = \frac{|e \cap e'|}{|e \cup e'|}$, that is, the similarity of two environments increases with the proportion of shared characteristics. This function is just a symmetric specification of the more general class considered in Tversky (1977). Although, it is not always possible to obtain all the information relating to the similarity function, a case analyzed in section 5, henceforth, we take E and σ as given.

2.2 An example

This section provides an example to illustrate the behavioral process we are modeling. Although the example is abstract, it shows very directly how the model works. For a more concrete application of the model to an interesting economic setting, please refer to section 7.

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

$t = 1$	$t = 2$	$t = 3$	$t = 4$	$t = 5$	$t = 6$	$t = 7$	$t = 8$
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, the DM has no prior experiences, therefore the maximizing self is active. Thus, the choice results from the maximization of preferences, that is, $a_1 = 3$. Then, in the following period, given that $\sigma(A_2, A_1) = \frac{4}{6} > .55$, the automatic self is active and therefore we will observe a replication of past choices, that is, $a_2 = 3$. Period 1's environment makes period 2's one fluent. Now, in period 3, notice that the similarity between A_3 and A_2 or A_1 is always below the similarity threshold and

that this makes the maximizing self active. The preference relation is maximized and therefore $a_3 = 1$. By similar reasoning applied for the fourth and fifth periods, it can be seen that $a_4 = 2$ and $a_5 = 1$. Then, in period six, the automatic self is active given that $\sigma(A_6, A_3) = \frac{3}{5} > .55$, leading to $a_6 = 1$. In period seven, given no sufficiently similar past environment, the maximizing self is active and therefore $a_7 = 2$. Finally, in the last period, the automatic self again becomes active, given that $\sigma(A_8, A_7) = \frac{3}{5} > .55$ and behavior will therefore be replicated, i.e. $a_8 = 2$.

One may wonder what an external observer would understand from this choice behavior. Section 3 addresses this point.

3 The Revealed Preference Analysis of Dual Decisions

In this section, we discuss how to recognize which observations were automatically or consciously generated 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. In fact, we can think of the similarity function as a cultural component common to most individuals, while the similarity threshold is the source of individual heterogeneity in automatic processes. Thus, the similarity function is taken as given, while the similarity threshold α is elicited from observed behavior.⁷ Notice that 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 relating to individual cognitive costs without too many degrees of freedom hampering the technical analysis.

⁷This assumption is in line with the stance taken in the cognitive sciences. In fact, similarity comparisons are generally taken as a more interpersonal component while cognitive costs are specific to the individual. See for example Tversky (1977), Medin et al. (1993) or Hahn (2014) and subsequent related literature, which also provides a variety of methods for identifying the similarity function. The idea here is that, even if, everyone in a certain society implicitly agrees that an Italian restaurant is more similar to a French one than to a Chinese one, they might not all perceive the Italian and the French ones, for example, to be so close as to be indistinguishable, i.e., to be similar enough. The appendix uses this assumption to estimate the similarity function. In the online appendix, the function is identified with rich enough data.

Although the exercise is very different from standard analysis, assuming knowledge of the similarity function is actually analogous to the usual assumptions made for the parametric estimation of decision theoretical models, where the functional form of interest is usually taken as given. Here we assume knowledge of the similarity function in order to enable the estimation of individual preferences and the similarity threshold. Section 5 shows that the hypothesis can be relaxed in order to generalize the analysis presented in this section.

It is easy to recognize a set of observations that are undoubtedly generated by the maximizing self. Notice that all those observations in which the chosen alternative was not a previous choice must belong to this category. This is so because, as no past behavior has been replicated, the automatic self could not be active. We call these observations *new* observations.

In order to identify a first set of automatically-generated observations, notice that, being rational, the maximizing self cannot generate cycles of revealed preference. As is standard, a set of observations t_1, t_2, \dots, t_k forms a *cycle* if $a_{t_{i+1}} \in A_{t_i}$, $i = 1, \dots, k-1$ and $a_{t_1} \in A_{t_k}$, where all chosen alternatives are different. Given the above reasoning, for every cycle there must be at least one observation that is automatically generated. Intuition suggests that the one corresponding to the most familiar environment should be a decision mimicking past behavior. The following definition helps in formalizing this idea.

The *unconditional familiarity* of observation t is

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

Whenever there is no $s < t$ such that $a_s \in A_t$, we say $f(t) = 0$.

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 *most familiar in a cycle* if it is part of a cycle of observations within which it maximizes the unconditional familiarity value. Given the above reasoning, these observations must be generated by the automatic self.

The major challenge is to relate pairs of observations in such a way as to allow

the knowledge of which self generated one of them to be transferred to the other. In order to do so, we introduce a second measure of familiarity of an observation t , which we call *conditional familiarity*. Formally,

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

Whenever there is no $s < t$ such that $a_s = a_t$, we say $f(t|a_t) = 0$.

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 at t . The main difference between $f(t)$ and $f(t|a_t)$ is that the former is an ex-ante concept, i.e., before considering the choice, while the latter 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, \dots, t_k such that $t = t_1$, $t_k = s$ and $t_i \in L(t_{i+1})$ for every $i = 1, 2, \dots, k - 1$.*

The definition formalizes the key idea behind the algorithm. Observation t is linked to observation s if its conditional familiarity is lower than the unconditional familiarity of s . As explained below, once an observation is categorized as consciously or automatically generated, it is through its link to other observations that such knowledge can be extended.

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 most familiar observations in a cycle are indirectly linked. The binary relation determined by the concept of linked observations is clearly reflexive; thus, new observations and most familiar observations in a cycle are contained in D^N and D^C respectively.

We are now ready to present the main result of this section, which establishes that observations in D^N are generated by the maximizing self, while observations in D^C are generated by the automatic self. As a consequence, it guarantees that the revealed preference of observations in D^N is informative with respect to the preferences of the

individual and it moreover identifies an interval within which the similarity threshold must lie. Before stating the proposition, it is useful to highlight that x is revealed preferred to y for a set of observations O , i.e., $xR(O)y$, if there is a sequence of different alternatives x_1, x_2, \dots, x_k such that $x_1 = x, x_k = y$ and for every $i = 1, 2, \dots, k - 1 \in O$, 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 the maximizing self while all observations in D^C are generated by the automatic self,*
2. *if x is revealed preferred to y for the set of observations D^N , then $x \succ y$,*
3. $\max_{t \in D^N} f(t) \leq \alpha < \min_{t \in D^C} f(t|a_t)$.

Proof. We start by proving the statement regarding conscious observations. Trivially, new observations must be generated by the maximizing self since they cannot be replicating any past behavior. Consider an observation $t \in D^N$. By definition, there exists a sequence of observations t_1, t_2, \dots, t_n with $t_1 = t, f(t_i|a_{t_i}) \leq f(t_{i+1})$ for all $i = 1, 2, \dots, n - 1$ and t_n being new. We prove that t is generated by the maximizing self recursively. We know that t_n is consciously generated. Now assume that t_k is generated by the maximizing self and suppose by contradiction that t_{k-1} is generated by the automatic one. 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 the maximizing self, and the recursive analysis proves that observation t is also consciously generated.

We now prove the statement regarding automatic observations. First, consider an observation t which is a most familiar in a cycle and assume by contradiction that it is generated by the maximizing self. Then, $f(t) \leq \alpha$. By definition of most familiar 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 the maximizing self. This is a contradiction with the maximization of a preference relation. Now consider an observation $t \in D^C$. By definition, there exists a sequence of observations t_1, t_2, \dots, t_n with $t_n = t, f(t_i|a_{t_i}) \leq f(t_{i+1})$ for

all $i = 1, 2, \dots, n - 1$ and t_1 being a most familiar in a cycle. We again proceed recursively. Since t_1 is generated by the automatic self, we have $f(t_1|a_{t_1}) > \alpha$. Now assume that t_k is generated by the automatic self and suppose by contradiction that t_{k+1} is generated by the maximizing one. 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 the maximizing self, it is straightforward to see that the revealed information from such a set must refer to the preferences of the DM. Regarding α , notice that since observations in D^N are generated by the maximizing self, we know that $\max_{t \in D^N} f(t) \leq \alpha$ and also that, since observations in D^C are generated by the automatic self, $\alpha < \min_{t \in D^C} f(t|a_t)$, which concludes the proof. ■

To understand the reasoning behind Proposition 1, first consider an observation t which is known to be new, and hence consciously generated. This implies 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 also be generated by the maximizing self. 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 the maximizing self. It is easy to see that this reasoning can be iterated. In fact, any observation linked with an observation generated by the maximizing self must also be generated by the maximizing self.

Similarly, consider a most familiar observation in a cycle t , that is known to be automatically generated. Any observation s for which the unconditional familiarity is greater than the conditional familiarity of t must also be generated by the automatic self. In fact, we know that $\alpha < f(t|a_t)$ because t is generated by the automatic self. Then, any observation s to which t is linked has unconditional familiarity greater than α , which implies that some past behavior *could be* replicated by the automatic self, and therefore that such an observation must also be generated by the automatic self. Again, the reasoning can be iterated. Thus, we can start from a small subset of

observations undoubtedly generated either automatically or consciously, and thence infer which other observations are of the same type. We use the example in section 2.2 to show how the algorithm works.

Algorithm: Example of Section 2.2

Suppose that we observe the decisions made by the DM in the example in section 2.2, without any knowledge regarding 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; hence, we can directly infer that the corresponding choices were conscious.

We can now go one step further and consider observation 5. From the observed behavior we cannot tell whether the choice comes from maximization of preferences or the replication of past behavior in period 3. Nevertheless, the choice was conscious in period 3 and one can easily see that $f(5|a_5) = \frac{2}{5} \leq \frac{3}{7} = f(3)$, which means that observation 5 is linked with observation 3 and, by Proposition 1, this reveals that it, too, was conscious.

Now consider observation 7. We cannot directly link observation 7 either to observation 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 . 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 automatic choices from D^N .

This example presents inconsistencies in the revealed preference. Observations 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 the automatic self thanks to Proposition 1, given they are most familiar in a cycle.

Notice then, however, that observation 6 is linked to observation 8 given that

$f(6|a_6) = \frac{3}{5} \leq f(8) = \frac{3}{5}$, thus revealing that the latter must also have been automatically generated. Thus, we obtain that $D^C = \{2, 6, 8\}$, which were the observations rightfully excluded from D^N . No decision made by the maximizing self has been cataloged as automatic.

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 is indeed the case. The value of the similarity threshold α can, by Proposition 1, be correctly determined to be in the interval $[0.5, 0.6)$ thanks to the information retrieved from observations 7 and 8 respectively.

D^N contains new observations and those indirectly linked to them.⁸ It may be the case that some consciously-generated observations are not linked to any observation in D^N , thus making D^N a proper subset of the set of all consciously-generated observations. For this reason, nothing guarantees that $D \setminus D^N$ are automatically-generated observations and hence, Proposition 1 must show how to dually construct a set of automatic decisions D^C . Nonetheless, if the observations are *rich enough*, it is possible to guarantee that D^N and D^C contain all conscious and automatic decisions, respectively.⁹ More importantly, notice that Proposition 1 relies on one important assumption, which is that the collection of observations is generated by a DD process. The following section addresses this issue.

4 A Characterization of Dual Decision Processes

In Section 3 we showed how to elicit the preferences and the similarity threshold of an individual who follows a DD process. Building upon the results of that section, we here 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.

From the construction of the set D^N , we understand that a necessary condition

⁸ D^N is never empty because it always contains the first observation.

⁹Material available in the online appendix.

for a dataset to be generated by a DD process is that the indirect revealed preference obtained 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. A simple postulate of choice characterizes the whole class of DD processes.

Axiom 1 (Link-Consistency) *A sequence of observations $\{(A_t, e_t, a_t)\}_{t=1}^T$ satisfies Link-Consistency if, in every cycle, there is at least one observation that is not indirectly linked to a new observation.*

This is a weakening of the Strong Axiom of Revealed Preference. In fact, it allows for cycles but only of a particular kind. Preferential information gathered from observations in D^N cannot be inconsistent. The following theorem shows that this 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.*

Proof. 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. Notice that Link-Consistency implies that the revealed preference relation defined over D^N , i.e. $R(D^N)$, is asymmetric. In fact, asymmetry of $R(D^N)$ means that it is not possible to construct cycles comprised of observations in D^N . Using 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 $s \notin D^N$, $f(s|a_s) > \alpha$; thus, it is possible to find, for each one, a preceding observation which it would appear 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 . ■

The theorem is saying that Link-Consistency makes it possible to determine whether the DM is following a DD process or not. In particular, when the property is satisfied, we can characterize the DM's preferences 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 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 $f(t|a_t)$. The next section shows how the algorithm would still work even if with extremely partial knowledge of the similarity comparisons. Notice, however, that Link-Consistency implies that there must be no cycles among new observations, a concept which is independent of the particular choice of similarity function, thus making the model always falsifiable. Finally, notice that we do not assume any particular structure for the sequence of observations we use as data, and hence that the characterization of preferences does not need to be unique, even when the similarity is known.¹⁰

5 Partial Information on the Similarity

In this section we show that our algorithmic analysis is robust to very weak assumptions concerning knowledge of the similarity function. In particular, we study the case in which only a partial preorder, i.e., a reflexive, transitive but incomplete binary relation, over pairs of environments is known, and denoted by S . For example, the Symmetric Difference between sets satisfies this assumption. Such an extension is very general. It is potentially relevant in many contexts where it is not possible to estimate the cardinal similarity function but only the relative order it represents. It is also robust to even weaker assumptions, when, for example, only some of the binary

¹⁰Material providing a full characterization of the procedure with rich enough data is available in the online appendix. Richness is obtained either by assuming that the data come from a homogeneous population, as is standard in the literature when dealing with path dependent models, or by assuming that the implicit memory of the DM is bounded and that the bound is known.

comparisons between pairs of environments are known. Returning to the example used in the introduction, we might not know how the DM compares different prices and dispositions of products on the shelf, but we might know that, for any combination of prices, a minor change in just one price results in a more similar environment than a major 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 conscious and automatic observations respectively, and that one consistency requirement characterizes all DD processes. In order to do this, we assume that, if the individual follows a DD process, the similarity σ cardinally represents a completion of such a partial order. Thus, for any $e, e', g, g' \in E$, $(e, e')S(g, g')$ implies that $\sigma(e, e') \geq \sigma(g, g')$ and we say that (e, e') dominates (g, g') . We first adapt the key concepts on which the algorithmic analysis is based in order to run the more general analysis.

The two concepts of familiarity need to be changed. 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:

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

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

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 in which e_t is compared with past observations, the choice of which *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 $(e_s, e')S(e_t, e)$, 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. Next, notice that, this time, we define the link

between an observation t and a *set* of observations O . This helps us to tell whether an observation is generated by the maximizing self once we know that another observation is. If all observations in O are generated consciously, and for each one of them there exists a pair of environments that dominates a pair in $F(t|a_t)$, then t must also have been generated by the maximizing self. 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 *Consciously-indirectly linked* to the set of observations O if there exists a sequence of observations t_1, \dots, t_k such that $t = t_1$, t_k is linked to O and t_i is linked to $\{t_{i+1}, t_{i+2}, \dots, t_k\} \cup O$ for every $i = 1, 2, \dots, k-1$. Define $D^{\hat{N}}$ as the set containing all new observations and all those observations indirectly linked to the *set* of new observations. Proposition 2 shows that $D^{\hat{N}}$ contains only observations generated by the maximizing self.

What about automatic choices? As in the main analysis, whenever a cycle is present in the data, we know that at least one of the observations in the cycle must be generated by the automatic self. This time, given that we assume only partial knowledge of the similarity comparisons, it is not always possible to define a most familiar observation in a cycle.¹¹ Notice, however, that, whenever an observation is inconsistent with the revealed preference constructed from $D^{\hat{N}}$, such an observation must be automatically generated.¹² Thus, say that observation t is *cloned* if it is either a most familiar in a cycle or $xR(t)y$ while $yR(D^{\hat{N}})x$.

Say that observation t is *Automatically-indirectly linked* to observation s if there exists a sequence of observations t_1, \dots, t_k such that $t = t_1$, $t_k = s$ and t_i is linked to t_{i+1} for every $i = 1, 2, \dots, k-1$. Whenever we know that observation t is automatically generated, we can infer that observation s is generated by the automatic self 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 *some* pairs of environments

¹¹Obviously, in this context, a most familiar 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 $(e_t, e')S(e_s, e)$.

¹²Notice that this implies that even with no knowledge of any similarity comparisons, the knowledge that new observations give us regarding the preferences of the individual highlights some observations as automatic given they are in contradiction with those preferences.

contained in $F(t|a_t)$ have similarity above the threshold. As before, let $D^{\hat{C}}$ be the set containing all cloned observations and the observations to which they are indirectly linked. Proposition 2 below shows that $D^{\hat{C}}$ contains only observations generated by the automatic self.

Proposition 2 *For every collection of observations D generated by a DD process where only a partial preorder over pairs of environments is known:*

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

Proof. First, we show that any observation t linked to a set O of conscious observations must also be generated by the maximizing self. 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 that 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 $(e_s, e')S(e_t, e)$ 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 in the paper, we know that new observations are consciously generated and, by applying the previous reasoning iteratively, it is shown that $D^{\hat{N}}$ must contain only observations generated by the maximizing self.

As a second step, we show that any observation s to which an observation t generated by the automatic self is linked, must be also automatically generated. Given that t is generated by the automatic self, 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 that t is linked to s , there exists a pair $(e_s, e') \in F(s)$ such that $(e_s, e')S(e_t, e_w)$. This implies that $\sigma(e_s, e') \geq \sigma(e_t, e_w) > \alpha$, and the result follows.
- Let $(e_t, e_w) \notin F(t|a_t)$. Then, there exists a $w' < t$ such that $(e_t, e_{w'})S(e_t, e_w)$ 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 that 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 $(e_s, e')S(e_t, e)$. In particular, there exists a $(e_s, e') \in F(s)$ such that $(e_s, e')S(e_t, e_{w'})$. This implies that $\sigma(e_s, e') \geq \sigma(e_t, e_{w'}) > \alpha$, and the result follows.

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

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

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 only a single condition, that is, $D^{\hat{N}}$ -Consistency.

Axiom 2 ($D^{\hat{N}}$ -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 that preorder. This is what is stated in the following theorem. Notice that S is assumed to be known.

Theorem 2 *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 S and a similarity threshold α that characterize a DD process.*

Proof. 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 2, $D^{\hat{N}}$ contains only conscious 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, using 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 S . 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 $(e_s, e')S(e_t, e)$ for some $(e_s, e') \in F(s)$. That is, for all observations not in $D^{\hat{N}}$, there exists a pair of environments which is not dominated by any pair of environments of observations in $D^{\hat{N}}$. We call this pair *undominated*. Then, let S' 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 $(e_s, e') \in F(s)$, $(e_t, e)S'(e_s, e')$ and not $(e_s, e')S'(e_t, e)$. Let S'' be the transitive closure of $S \cup S'$. Notice that S'' is an extension of S that preserves its reflexivity and transitivity. Thus, we can find a completion S^* of S'' and a similarity function $\sigma : E \times E \rightarrow [0, 1]$ that represents S^* .

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 for it to replicate.

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

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

6 The Model in Context

In this section, we show how the paper relates to the literature and what evidence from cognitive sciences it is formalizing.

6.1 Related Literature

In our model the presence of similarity comparisons makes behavior automatic. 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, which was proposed in Gilboa and Schmeidler (1995) and summarized in Gilboa and Schmeidler (2001). In case-based decision theory, as in our model, the 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 selves interacting during the decision-making process. Second, from a technical point of view, our model uses 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 8 suggests, case-based decision theory may 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.

We would like to stress that, although our proposed behavioral model is new, the idea that observed behavior may be the outcome of interaction between two different selves is not novel and dates back, at least, to Strotz (1955). Strotz-type models, such as Laibson (1997), Gul and Pesendorfer (2001) or Fudenberg and Levine (2006), are different from the behavioral model introduced here, since they represent the two selves as two agents with different and conflicting preferences, usually *long-run* vs *short-run* preferences.¹³ In our approach, however, the two selves are *inherently*

¹³In some models, the difference between the two selves stems from the fact that they have different

different one from the other. One uses analogies to deal with the environment in which the decision maker acts, while the other uses a preference relation to consciously choose among the available options. Furthermore, the issue of which self drives a particular decision problem depends on problems experienced in the past and their degree of similarity with the current one, and not on whether the decision affects the present or the future. Nevertheless, we do not exclude the possibility that analogy formation may be influenced to some extent by whether the decision affects the present or the future.

Finally, it is important to notice that the preference revelation strategy used in this paper agrees with the one used in Bernheim and Rangel (2009), who analyze the same problem of eliciting individual preferences from behavioral datasets, which they do 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.¹⁴ Albeit similar, our approach differs in two important aspects. Firstly, we model conscious and automatic choices, and, in section 3, we propose a particular method for finding the welfare-relevant domain, i.e. the algorithm highlighting a set of conscious choices. Secondly, we propose a specific choice procedure, and perform 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 et al. (2012) and Manzini and Mariotti (2014), who make the case for welfare analysis based on the understanding of the behavioral process that generated the data.

information. See, for example, Cunningham (2015), which proposes a model of decision making where the two selves hierarchically aggregate information before choosing an alternative. Other papers use more than two selves to rationalize individual choices, but all selves are still represented by different preference relations; see, for example, Kalai et al. (2002), Manzini and Mariotti (2007) and Ambrus and Rozen (2015).

¹⁴Notice that Apesteguia and Ballester (2015) also propose a choice-procedure free approach for measuring the welfare of an individual from a given dataset. They do this by providing a model-free method to measure how close actual behavior is to the preference that best reflects the choices in the dataset.

6.2 Cognitive Foundations of the Model

The cognitive sciences have always distinguished between conscious and unconscious processes. More recently, this duality has been incorporated into Dual Process Theory, as described in Evans and Frankish (2009) and Kahneman (2011), which divides mental processes into two categories, analytical (or conscious) and analogical (or automatic).¹⁵ Using these findings, the objective here is twofold, (i) to propose a simple initial formalization of automatic processes and (ii) to model the interaction between conscious and automatic processes in a tractable way. We abstract from more in-depth modeling of conscious behavior, which is represented here by the rational model. Nevertheless, the framework developed in this paper is flexible enough to allow for alternative models of conscious processes, as explained in more detail in section 8.

Unconscious or automatic processes are extremely context dependent. In particular, the main determinant for the activation of non-deliberative processes is the subjective experience of ease associated with a particular situation, i.e., *fluency*, due to the characteristics of the environment. Studies surveyed by Oppenheimer (2008), one of the leading scholars in fluency research, show that in fluent, familiar, situations, individuals' decisions are less conscious, that is, more automatic. Disfluent situations, on the other hand, lead to conscious behavior.¹⁶ Hence, it becomes crucial for the purposes of this paper to be able to formalize what fluency of an environment is.

Paraphrasing Oppenheimer (2008), a situation becomes more fluent as it becomes more familiar. That is, an experience is more fluent the easier is the unconscious perception of its similarity with past experiences.¹⁷ Thus, we model familiarity here

¹⁵Many papers helped in developing the theory, including Schneider and Shiffrin (1977), Evans (1977), McAndrews et al. (1987), Evans (1989), Reber (1989), Epstein (1994) and Evans and Over (1996).

¹⁶This literature is closely related to that on limited attention in cognitive sciences, where agents consciously analyze a situation only if it is *novel enough*, see Woodford (2019).

¹⁷This is in line with research on *priming*, which studies the influence of unconscious or implicit memory on behavior through environmental cues, which appears to be one of the main sources of fluency. As defined in Tulving and Schacter (1990), priming is an unconscious change in the ability to identify or produce an item as a result of a specific prior encounter with that item. Priming creates a sense of familiarity and ease, which make the environment fluent and behavior automatic

through two main channels. First, an environment is more familiar the greater its similarity with environments experienced in the past.¹⁸ This is measured by the similarity function. Second, problems requiring new solutions due a lack of previous solutions, cannot be considered familiar. That is, a decision problem is familiar if past behavior can be replicated. Clearly, these are simplifications; but they are in line with evidence in cognitive sciences showing that automatic processes are based on analogical reasoning, i.e., similarity judgments with the past, which activate automatic responses.¹⁹ The main idea we are depicting is that environments that are similar with the past prompt automatic responses that are drawn from past experiences. This is also in line with the evidence in Woodford (2019) which calls for the introduction of these ideas in economics.

Then, the second important step is to model the dichotomy between conscious and automatic processes. The majority of the evidence points to the fact that analogical reasoning is parallel.²⁰ That is, automatic processes operate continuously; unconsciously assessing the nature and familiarity of the environment. On the other hand, conscious processes are costly and are fully activated only when the environment is disfluent or novel, i.e., unfamiliar. In the model, the parallel nature of analogical reasoning is represented by the fact that similarity comparisons are always drawn and the costly nature of the conscious system is captured by the similarity threshold. The higher the threshold, the lower the cost; that is, fewer environments will be perceived as fluent or familiar. This assumption enriches the model and aligns it more closely with the evidence. In fact, the fluency of an environment is a concept that intertwines similarity with past experiences with the cost of conscious processing.²¹ Conscious processes are needed when new responses are required due to the novelty of the environment, either because it is not similar enough with past experiences or

and coherent with the context. See Bargh (2005) for a survey on priming.

¹⁸Notice that , given that we are describing implicit memory, we assume perfect memory, in line with research in neurosciences as already summarized in Graf (1990). For a more recent survey, see Dew and Cabeza (2011).

¹⁹See Bargh (2005), Evans and Frankish (2009), Kahneman (2011) and Lisman and Sternberg (2013).

²⁰See Evans and Frankish (2009) for a survey of the evidence.

²¹Notice that the model abstracts from the notion that the cost of conscious processing can be a function of the environment. This is clearly a simplification to avoid having too many degrees of freedom, but it is certainly an interesting avenue for future research.

because no behavior can be replicated.²² This is in line with research in neuroscience, see Lisman and Sternberg (2013), which sustains that unconscious processes generate habits while non-habits are the outcome of conscious responses.

Finally, notice that, in this model, the relationship between automatic and conscious processes does not have to be completely dichotomous. As already stated, here, we present and analyze a class of models, some specifications of which make the relationship between automatic and conscious processes more complementary. For example, choosing the alternative that maximizes preferences over those chosen in similar-enough past environments. In this sense, the automatic system would be simplifying the problem by creating a *fluent* consideration set on which the conscious system would be maximizing preferences.

7 An Application: Asymmetric Pricing

In this section, the model is used to provide a simple explanation for the phenomenon of *asymmetric pricing*, that is, the asymmetric response of firms to changes in costs. Firms increase prices when costs go up, while they tend not to decrease prices when costs go down. This well- documented fact is at odds with standard economic theory, despite the fact that it affects two thirds of markets, as highlighted in the seminal paper Peltzman (2000). The evidence in Peltzman (2000) is quite overwhelming and excludes more standard explanations, such as collusion or menu costs. In fact, the phenomenon is present in all kind of markets, regardless of their degree of competitiveness. Peltzman (2000) calls for a new theory to deal with this phenomenon. Various models have been put forward, as explained below, but, first, we show that DD processes provide a fairly natural and alternative explanation of the phenomenon.

Consider two firms competing over a market comprised of consumers described by a DD process. Consumers' automatic systems make analogies over price vectors, i.e., given constant wealth, over budget sets. Meanwhile, firms face cost shocks. The key idea is that the presence of DD consumers creates a demand in the second period that depends on prices in the first and that has asymmetric responses to changes

²²We do not exclude the possibility that, in such instances, the conscious response would be to choose a *similar enough* alternative in a *similar enough* environment.

in prices. If prices rise in the second period, consumers react rationally because they are unable to replicate past behavior, given that their budget sets have shrunk. On the other hand, if prices fall in the second period, not all consumers will adapt their behavior, given that they have heterogeneous similarity thresholds and are able to replicate past behavior. Thus, if the decrease in prices is not big enough, some consumers will not perceive the change, thus making the demand much less elastic. Under certain circumstances, this particular demand structure gives firms incentives to maintain prices when costs go down but adjust them when costs go up. The following subsections clarify this reasoning.

7.1 Setting

The setting is very simple. There are two firms, A and B , which compete à la Bertrand over two periods.²³ The structure of the game is common knowledge among the firms. First period costs are symmetric, marginally constant and normalized to 1, that is, $c_A^1 = c_B^1 = 1$, where c_i^t represent the costs of firm i in period t . In period 2, costs are independently drawn for each firm $i = A, B$ as follows:

$$c_i^2 = \begin{cases} (1 + \beta) & \text{with } \pi \text{ probability} \\ (1 - \beta) & \text{with } 1 - \pi \text{ probability} \end{cases}$$

with $\beta \in (0, 1)$.

Consumers are described by a DD process and have V units of wealth in each period. The automatic system makes comparisons between price vectors. In the first period, consumers are rational, given that there is no past to replicate. In the second period, consumers' behavior depends on firms' pricing decisions and consumers' similarity thresholds, which are distributed in the population with density function f . This means that heterogeneity in the population is modeled through different costs of maximizing-self activation.

²³Competition à la Bertrand assures that asymmetric pricing is not the outcome of firms' market power while also making the assumption of having just two firms less relevant. In fact, the whole analysis can be generalized to n firms quite easily.

7.2 Analysis

In this section, we look for the subgame perfect equilibria of the game. We focus on the existence of an equilibrium in pure strategies. First of all, notice that, given the symmetry of the environment, we must have that firms choose the same action in period one. For now, let $p_A^1 = p_B^1 = p^1$, where p_i^t is the price firm i charges in period t . We now look at the second period decision. W.l.o.g., the analysis concerns firm A's decision in the second period.

The demand faced by firm A in period two depends not only on prices in period one, but also on the pricing decisions of firm B in period two. Let:

$$\mu(\alpha_n) = \int_0^{\alpha_n} f(\alpha) d\alpha$$

with

$$\alpha_n = \sigma(p_A^1, p_B^1; p_A^2, p_B^2)$$

That is μ represents the fraction of consumers who use the automatic system to purchase in period two. Thus we have the following.

If $p_B^2 = p^1$:

$$d_A^2(p_A^1, p_B^1, p_A^2, p_B^2) = \begin{cases} 0 & \text{if } p_A^2 > p_B^2, \\ \frac{1}{2} \frac{V}{p^1} & \text{if } p_A^2 = p_B^2, \\ \mu(\alpha_1) \frac{1}{2} \frac{V}{p^1} + (1 - \mu(\alpha_1)) \frac{1}{2} \frac{V}{p_A^2} & \text{if } p_A^2 < p_B^2. \end{cases}$$

While if $p_B^2 \neq p^1$ we get:

$$d_1^2(p_A^1, p_B^1, p_A^2, p_B^2) = \begin{cases} 0 & \text{if } p_A^2 > p_B^2 \geq p^1, \\ \frac{1}{2} \frac{V}{p_A^2} & \text{if } p^1 < p_A^2 = p_B^2, \\ \frac{1}{2} \frac{V}{p^1} & \text{if } p^1 = p_A^2 = p_B^2, \\ \frac{1}{2} \frac{V}{p^1} & \text{if } p^1 = p_A^2 < p_B^2, \\ \mu(\alpha_2) \frac{1}{2} \frac{V}{p^1} + (1 - \mu(\alpha_2)) 0 & \text{if } p_B^2 < p_A^2 \leq p^1, \\ \mu(\alpha_2) \frac{1}{2} \frac{V}{p^1} + (1 - \mu(\alpha_2)) \frac{1}{2} \frac{V}{p_A^2} & \text{if } p_A^2 < p_B^2 \leq p^1. \end{cases}$$

Notice that $\alpha_1 \geq \alpha_2$, given that, in scenario 2, both firms might have changed prices so fewer consumers will perceive the environment as being similar enough.

First of all, when costs are high, i.e., $c_A^2 = 1 + \beta$, firm A will adapt prices such that $p^1 \neq p_A^2 = 1 + \beta$. Suppose firm A does not adapt prices, that is, $p_A^2 = p^1$. Suppose, then, that $p_A^2 = p^1 \geq 1 + \beta$. In this case, firms would be making positive profits in both periods, under any circumstance, and hence would have incentives to undercut each other in the first period, thus leading to $p^1 < 1 + \beta$. Clearly, it cannot be that $p_A^2 = p^1 < 1 + \beta$, because firm A would have incentives to deviate in the second period to avoid incurring in negative profits. Thus, it must be that $p^1 \neq p_A^2 = 1 + \beta$.

As a second point, notice that, when firm A has low costs in period two, i.e., $c_A^2 = 1 - \beta$, while firm B has high costs, $c_B^2 = 1 + \beta$, we will have that $p_A^2 = 1 + \beta - \epsilon$, given that $p_B^2 = 1 + \beta$, as per the previous reasoning. For simplicity, we assume that, in this case, $p_A^2 = 1 + \beta$ and that firm A wins the entire market.

Now, to examine what happens when costs are low for both firms, that is, $c_A^2 = c_B^2 = 1 - \beta$, let us assume that firm A believes that firm B does not change prices, that is, $p_B^2 = p^1$. Then, if firm A lowers prices it gets:

$$(p_A^2 - (1 - \beta))(\mu(\alpha_1)\frac{1}{2}\frac{V}{p^1} + (1 - \mu(\alpha_1))1\frac{V}{p_A^2})$$

On the other hand, if it does not lower prices, it gets:

$$(p^1 - (1 - \beta))\frac{1}{2}\frac{V}{p^1}$$

That is, firm A does not charge p_A^2 if:

$$(p^1 - (1 - \beta))\frac{1}{2}\frac{V}{p^1} \geq (p_A^2 - (1 - \beta))(\mu(\alpha_1)\frac{1}{2}\frac{V}{p^1} + (1 - \mu(\alpha_1))1\frac{V}{p_A^2})$$

with $p^1 > p_A^2 > (1 - \beta)$. This condition can be rewritten as:

$$\mu(\alpha_1) \geq \frac{2(p_A^2 - (1 - \beta))p^1 - (p^1 - (1 - \beta))p_A^2}{(p_A^2 - (1 - \beta))(2p^1 - p_A^2)}$$

That is:

$$F(\alpha_1) \geq \frac{2(p_A^2 - (1 - \beta))p^1 - (p^1 - (1 - \beta))p_A^2}{(p_A^2 - (1 - \beta))(2p^1 - p_A^2)}$$

If we let $p_A^2 = p^1 - \epsilon$, we can rewrite the condition as follows:

$$F(\alpha_1) \geq \frac{p^1}{p^1 + \epsilon} - \frac{\epsilon}{(p^1 + \epsilon)} \frac{1 - \beta}{(p^1 - \epsilon - (1 - \beta))} \quad (1)$$

The following subsection will show that this condition can hold under many general circumstances. For now, let us assume that the condition is satisfied, to see whether it is possible to obtain an equilibrium with asymmetric pricing in pure strategies.

We need to solve for p^1 . Clearly, price competition among firms requires p^1 to be such that expected profits in period 1 are zero. That is:

$$(p^1 - 1)\frac{1}{2}\frac{V}{p^1} + \pi^2(p^1 - (1 - \beta))\frac{1}{2}\frac{V}{p^1} + \pi(1 - \pi)2\beta\frac{V}{1 + \beta} = 0$$

So we get:

$$p^1 = \frac{(1 + \beta)(1 + (1 - \beta)\pi^2)}{(1 + \pi^2)(1 + \beta) + \pi(1 - \pi)4\beta}$$

First notice that $p^1 < 1$. In fact:

$$(1 + \beta)(1 + \pi^2(1 - \beta)) < (1 + \beta)(1 + \pi^2) + \pi(1 - \pi)4\beta$$

Given that $(1 + \beta)(1 + \pi^2(1 - \beta)) < (1 + \beta)(1 + \pi^2)$ and $\pi(1 - \pi)4\beta > 0$. Moreover, clearly, $p^1 > 1 - \beta$. In fact:

$$p^1 - (1 - \beta) = \frac{\beta((1 + \beta) - \pi(1 - \pi)4(1 - \beta))}{(1 + \pi^2)(1 + \beta) + \pi(1 - \pi)4\beta} > 0$$

given that

$$\frac{1 + \beta}{1 - \beta} > 4\pi(1 - \pi),$$

where the left hand side is greater than 1, while the maximum of the right hand side is 1.

To be sure that p^1 can be an equilibrium price when Condition (1) is satisfied, firms must have no incentives to deviate. In fact, it might be the case that firms still

want to deviate unilaterally in the first period. That is, firms might decide to incur in greater short-term losses in order to win the entire market in the second period, and thereby make greater long-term profits. Thus, for p^1 to be an equilibrium price, it must be that:

$$\begin{aligned} & (p^1 - 1)\frac{1}{2}\frac{V}{p^1} + \pi^2(p^1 - (1 - \beta))\frac{1}{2}\frac{V}{p^1} + \pi(1 - \pi)2\beta\frac{V}{1 + \beta} > \\ & > (p^1 - \epsilon - 1)\frac{V}{p^1 - \epsilon} + \pi^2(p^1 - \epsilon - (1 - \beta))\frac{V}{p^1 - \epsilon} + \pi(1 - \pi)2\beta\frac{V}{1 + \beta} \end{aligned}$$

Or:

$$0 > (p^1 - \epsilon - 1)\frac{V}{p^1 - \epsilon} + \pi^2(p^1 - \epsilon - (1 - \beta))\frac{V}{p^1 - \epsilon} + \pi(1 - \pi)2\beta\frac{V}{1 + \beta}.$$

which can be rewritten as:

$$\frac{(1 + \beta)(1 + (1 - \beta)\pi^2)}{(1 + \pi^2)(1 + \beta) + \pi(1 - \pi)2\beta} > p^1 - \epsilon.$$

A sufficient condition is:

$$\frac{(1 + \beta)(1 + (1 - \beta)\pi^2)}{(1 + \pi^2)(1 + \beta) + \pi(1 - \pi)2\beta} > p^1 = \frac{(1 + \beta)(1 + (1 - \beta)\pi^2)}{(1 + \pi^2)(1 + \beta) + \pi(1 - \pi)4\beta},$$

which is always true. Firms have no incentive to deviate. Thus, it is possible to obtain asymmetry of pricing in pure strategies in this setting, as highlighted in the following proposition.

Proposition 3 *There is an equilibrium with asymmetric pricing only if Condition (1) holds.*

Notice that proposition 3 is actually saying two things. Firstly, that, if Condition (1) does not hold, then firms have incentives to deviate unilaterally and therefore, by standard arguments, prices will be equal to costs, thus leading to no asymmetry in pricing. Secondly, that, if Condition (1) holds, there is an equilibrium with asymmetric pricing but it might be the case that there are more equilibria. In fact, the whole analysis is based on firm A's belief that firm B does not change prices. If firm A

believes otherwise, then its incentives to change prices might be stronger and an additional standard equilibrium, with prices equal to costs, would be possible.²⁴

Finally, it is interesting to highlight the following.

Remark 1 *Whenever Condition (1) holds, there is an equilibrium in which firms enjoy a mark up in period 2 when costs are low.*

In fact, if Condition (1) is satisfied, $p^1 \in (1 - \beta, 1)$ and therefore firms have negative profits in period 1 that are compensated by expected positive profits in period 2. This implies that prices align with costs only when costs are high. While not central to understanding asymmetry in pricing, this remark is interesting because it means that, in this simple setting, firms might enjoy mark-ups in pure strategies, even in a fairly competitive environment as occurs under Bertrand competition.

7.3 Analysis of Condition (1)

In this section, we analyze whether there are general settings under which Condition (1) holds or fails.

To understand when Condition (1) is satisfied, we need to make some assumption regarding the automatic system and the distribution of the similarity threshold in the population. As a first step, let the similarity function be:

$$\sigma(p_A^1, p_B^1; p_A^2, p_B^2) = \frac{1}{1 + d((p_A^1, p_A^2), (p_B^1, p_B^2))}$$

where $d((p_A^1, p_A^2), (p_B^1, p_B^2)) = ((p_A^1 - p_A^2)^2 + (p_B^1 - p_B^2)^2)^{\frac{1}{2}}$ is the Euclidean distance between price vectors in the two periods. Under the assumptions of Condition (1) we have that $((p_A^1 - p_B^1)^2 + (p_A^2 - p_B^2)^2)^{\frac{1}{2}} = \epsilon$. Thus, we have:

$$\mu(\alpha_1) = \int_0^{\frac{1}{1+\epsilon}} f(\alpha) d\alpha = F\left(\frac{1}{1+\epsilon}\right)$$

²⁴Notice that, if Condition (1) holds, such an equilibrium would not be trembling-hand perfect, given that the belief that firm B will change prices can be sustained only by allowing mistakes by firm B.

Thus, Condition (1) can be rewritten as follows:

$$F\left(\frac{1}{1+\epsilon}\right) \geq \frac{p^1}{p^1+\epsilon} - \frac{\epsilon}{(p^1+\epsilon)} \frac{1-\beta}{(p^1-\epsilon-(1-\beta))},$$

which leads to remark 2.

Remark 2 *If $F\left(\frac{1}{1+\epsilon}\right) \geq \frac{1}{1+\epsilon}$ for any ϵ , Condition (1) is satisfied.*

This stems from the fact that $p^1 < 1$ and so:

$$F\left(\frac{1}{1+\epsilon}\right) \geq \frac{1}{1+\epsilon} > \frac{p^1}{p^1+\epsilon} > \frac{p^1}{p^1+\epsilon} - \frac{\epsilon}{(p^1+\epsilon)} \frac{1-\beta}{(p^1-\epsilon-(1-\beta))},$$

where the last inequality stems from the fact that the last term is always positive. Notice that $F\left(\frac{1}{1+\epsilon}\right) \geq \frac{1}{1+\epsilon}$ for any ϵ for many distributions, e.g. the uniform distribution or any positively-skewed Beta distribution, such as $Beta(1, n)$ with $n > 1$. However, this is only a sufficient condition. Whenever $F\left(\frac{1}{1+\epsilon}\right) < \frac{1}{1+\epsilon}$ for some ϵ , Condition (1) must be checked and no general conclusions can be drawn without making further assumptions. Nevertheless, an interesting implication of the model can be highlighted.

Remark 3 *If $F\left(\frac{1}{1+\epsilon}\right) < \frac{p^1}{p^1+\epsilon}$ for some ϵ , then there exists a β big enough such that Condition (1) fails.*

The remark is a direct consequence of the fact that

$$\lim_{\beta \rightarrow 1} \frac{1-\beta}{(p^1-\epsilon-(1-\beta))} = 0$$

Remark 3 is of particular interest given that it goes in the direction of empirical findings. In fact, Peltzman (2000) finds that asymmetric pricing is present in those markets in which cost shock are not particularly big. That is, empirically, the bigger the cost shocks, the less asymmetric the behavior of firms in adjusting prices to costs. Similar results are found in Chen et al. (2008). The intuition behind Remark 3 in our setting is quite straightforward. *Ceteris paribus*, the higher the β , the more room there is for a firm to decrease the price to a degree that is perceived by enough consumers to make it profitable. In fact, in general, given that the second term of the

right hand side of Condition (1) is decreasing in β , the greater the β , the more likely it is for Condition (1) to fail. Thus, under certain circumstances, if the cost shock is big enough, the asymmetric response of demand to changes in prices disappears.

This is a crucial distinction between the theoretical explanation provided here and the models that have been presented in the literature to reconcile asymmetric pricing with economic theory. In fact, since Peltzman (2000), there have been many attempts to explain the phenomenon of asymmetric pricing by abstracting from menu costs or market power, which have been rejected by the data. All models, in a similar fashion to our framework, include some kind of inelasticity of demand in response to changes in prices.

One strand of the literature (Yang and Ye (2008), Tappata (2009), Lewis (2011), Cabral and Fishman (2012)) has used search models to explain such inelasticities. The main idea is that consumers form expectations on the distribution of prices based on past realization of prices or costs and thus tend to search less intensely after high prior realization of prices or costs, while searching more intensely after low realizations. This gives incentives to firms to react asymmetrically to cost changes. Another strand, Levy et al. (2004), has used rational inattention to create inelastic response of demand to price changes. According to this model, consumers decide not to allocate attention to small price changes because it is costly for them to do so. This means that demand is symmetrically inelastic for small price changes, irrespective of their direction. Clearly, firms facing this kind of demand have incentives to raise prices for small changes in costs given the demand does not change but have no incentives to reduce prices for minor cost changes because they would not increase demand.

The explanation provided here, despite its simplicity, has some crucial differences that reconcile the two strands of literature with the empirical evidence. First of all, as Remark 3 shows, it can connect the presence of asymmetric pricing with the magnitude of the cost shock an industry suffers like the models of rational inattention. All the papers using search models are unable to reconcile this fact, because the friction created by searching would make asymmetric pricing always optimal for firms, regardless of the magnitude of the shock. Secondly, like search models, the model maintains the asymmetric response of demand to price changes, which is present in

the data, but which rational inattention models cannot explain in this context.²⁵

Finally, it is useful to underline one last feature of the general framework presented here which sets it apart from other settings. The model allows for each consumer's choice to be conscious or automatic, i.e. rational or not, depending on the features of the environment, hence making the overall proportion of agents that are rational or not in the population endogenous to the problem at hand. That is, heterogeneity of behavior is easily obtained through the distribution of α , thus making the model tractable.

8 Final Remarks

The cognitive sciences have highlighted the fact that choices can be divided into two categories; conscious and automatic. This greatly hinders the use of standard economic models, which do not generally take into account the existence of automatic choices.

In this paper, we proposed a possible formalization of this duality, which allows us to restore standard revealed preference analysis. For this reason, 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 also arise when choosing consciously. Fortunately, the framework and analysis developed here do not depend on which particular type of conscious behavior is assumed. In fact, 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 containing those that are not. Such a partition is based on only one, simple assumption; that automatic choices must stem from the replication of past behavior, in line with research in cognitive and neuro sciences, as explained in Lisman and Sternberg (2013). Alternative conscious behaviors are possible. The only element of the formal analysis that must be changed is the consistency requirement, which

²⁵The asymmetric response of demand to price changes has received a great deal of attention in marketing (Krishnamurthi et al. (1992), Kalyanaram and Winer (1995), Mazumdar et al. (2005), Karle et al. (2015), Cornelsen et al. (2018)) and many other fields, such as finance, telecommunications, energy and transportation (Bidwell et al. (1995), Chen et al. (2004), Adeyemi and Hunt (2007), Wadud (2014), Hymel and Small (2015)).

needs to be tested on those problems in D^N . Obviously, dually, to find automatic choices, we should analyze violations of such requirements. Similarly, the assumption that automatic behavior stems from the replication of past behavior is much less restrictive than would appear at a first glance. Automatic choices must be *familiar* choices. For example, we can consider cases where the DM stores in his memory not only his own experiences, but also those of his parents, siblings or friends, thus making the concept more general than a literal interpretation of the model would suggest.

Section 7 showed that the model is very tractable and can be used in standard economic analysis. It formalizes a novel way of understanding the coexistence of sticky and adaptive behavior that generates predictions different from those of other theories. 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. On the other hand, if past behavior cannot be replicated, individual choices are perfectly responsive to changes in the environment. This asymmetry is one of the key differences that distinguish this model from others in the literature and enable it to encompass various phenomena in different fields that usually require specific explanatory models. In fact, there is a large and diverse body of empirical evidence showing that individual behavior does not immediately adapt to changes in the economic environment and that its responsiveness depends on the nature of the change. Consumption tends to be sticky, as shown by Carroll et al. (2011) among others. However, the response is asymmetric; that is, consumption is less responsive as prices go down, than as they go up, as explained in section 7. Traders tend to show under-reaction to news, and trading behavior is less responsive to favorable market conditions; see for example Braun et al. (1995) and Chan et al. (1996). Finally, doctors, for no rationally explicable reason, sometimes tend to stick to suboptimal treatments see, for example, Hellerstein (1998), whereas, in other cases, they are quick adopters of tailored treatments for certain patients, as found in Frank and Zeckhauser (2007). The framework presented here allows for a formal analysis of these different findings in a simple and tractable environment, where it is possible to model individual heterogeneity in behavior through the distribution of a simple

parameter such as α .²⁶

This model may also be useful for addressing welfare issues. Obviously, restoring a revealed preferences approach is crucial for welfare analysis. Nevertheless, the approach proposed here highlights an issue that is usually neglected, i.e., cognitive costs. The upper and lower bounds of the interval in which the similarity threshold lies are defining, respectively, a lower and upper bound of the cognitive costs of activation of the maximizing self. This may be of help in the analysis of welfare, but it raises the question of how to relate such costs to utility, which is a problem that we leave for future research. A first possibility would be the following. Suppose the similarity of two environments depends on how close the maximum utility obtainable in both of them is. Clearly, in such a specification α summarizes the cost of activating the maximizing self in *utils*, because similarity comparisons are also determined by utility. Notice, however, that even when all primitives are related to individual preferences, suboptimal choices are possible, thus maintaining the importance of the revealed preference analysis in section 3. Imagine, for example, a sequence of decision problems where behavior can be replicated in each step of the sequence and the environments are close enough to activate the automatic self. Suppose that the utility attainable in each problem is increasing throughout the sequence. Obviously, even in this specification of the model, a decision maker replicating the initial choice would eventually make a suboptimal decision. This is because the environment of a given decision problem is automatically compared with all past environments, not only those associated with a problem that was solved by the maximizing self.²⁷

Finally, another point worth further attention is that some simplifying assumptions were made in order to develop a new framework encompassing automatic choices. In particular, we assumed that the similarity function is known and the environments are given. Although the first of these assumptions can be relaxed, as shown in sec-

²⁶See the online appendix for an analysis of these economic environments through the lenses of DD processes.

²⁷Although this might appear an odd feature of the model, it closely maps cognitive science ideas, e.g., Chugh and Bazerman (2007) and references therein, by enabling the depiction of the *boiling frog* syndrome. See Offerman and Van Der Veen (2015) for evidence of the phenomenon in economics. Notice that this feature also distinguishes this specification of the model from satisficing and rational inattention models.

tion 5, it is nonetheless essential to know something about similarity comparisons.²⁸ Automatic behavior can be consistent with the maximization of a given preference relation, thus making the observation of inconsistent behavior insufficient for the correct categorization of decisions. Analogies add a layer of complexity to the problem, thus calling for richer data. A first step in this direction is to address the second of the above-mentioned assumptions, i.e., that environments are known. An understanding of the key elements in a decision problem for similarity comparisons, is crucial for the study of individual decision making as noted by Woodford (2019). The model proposed here provides a setting for the structured consideration of this issue but leaves the question open for future research.

A Appendix

A.1 Estimation of the Similarity Function

The similarity function is a key component of a DD process and, for the sake of exposition, in this paper, it is assumed to be at least partially known. Nonetheless, we discuss in what follows how to estimate it by studying the choice behavior of a group of individuals sharing the same similarity function.²⁹ In other words, we are assuming that the population follows a DD process. Notice, however, that, once the similarity is estimated, this assumption can be falsified, as explained in the main text.

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. Suppose, for now, that the preference distribution is known. This assumption is quite common in empirical analysis and can be relaxed in some cases, as explained below. It implies that, for every $A \subseteq X$, the proportion of agents in the population that would choose alternative $x \in A$, for any alternative $x \in A$, is known. Let $\pi(x|A)$ represent such a proportion. Before proceeding to the analysis, notice that, even

²⁸Alternatively, one needs a large enough population of individuals sharing the same similarity function, as shown in the appendices.

²⁹The idea that similarity comparisons are affected by culture has been studied at least since Whorf (1941).

in this circumstance, the revelation of individual preferences would still be relevant for policy purposes. In fact, knowing the aggregate preference distribution is not equivalent to knowing the preferences of each individual.

Now, assume that, for every pair of environments $e, e' \in E$, there exists a representative subpopulation of agents such that, for each agent:

- there exists some t' such that t' is new, $a_{t'} = x$ for some $x \in X$ and $e_{t'} = e'$.
- there exists some $t > t'$ such that $x \in A_t$, no alternative in $A_t \setminus \{x\}$ has been chosen before t and $e_t = e$.
- there is no $s \in (t', t)$ such that $a_s = x$.

The main result of this section shows that we can compare the similarity of two different pairs of environments by considering the aforementioned respective subpopulations and sampling them. Formally, denote by $\nu(x|e, e', A_t)$ the average relative number of randomly-sampled individuals sticking to x at t , as defined in the previous richness condition. That is, for any pair of environments (e, e') , we take a sample of finite magnitude n from the aforementioned subpopulations and compute the average ratio of individuals in that sample who stick to y . This average is $\nu(x|e, e', A_t)$. Notice that the considered alternatives and sequences of decision problems that satisfy the richness condition might differ across different pairs of environments.

Proposition 4 (Eliciting the Similarity) *For every two pairs of environments (e, e') and (g, g') , $Pr(|\nu(x|e, e', A_t) - \pi(x|A_t)| \geq |\nu(x'|g, g', A_{t'}) - \pi(x'|A_{t'})| | \sigma(g, g') > \sigma(e, e')) \xrightarrow{p} 0$. That is, the probability of having $|\nu(x|e, e', A_t) - \pi(x|A_t)| \geq |\nu(x'|g, g', A_{t'}) - \pi(x'|A_{t'})|$ when $\sigma(g, g') > \sigma(e, e')$ probabilistically converges to zero.*

Proof. First, notice that $|\nu(x|e, e', A_t) - \pi(x|A_t)| \neq 0$ for only two reasons, (i) sampling noise, (ii) automatic decisions. Given that the samples to estimate $\nu(x|e, e', A_t)$ are independent, as the dimension of the sample grows, the law of large numbers applies; therefore the first concern disappears in the limit. This leaves us with the second. Given that α is continuously and independently distributed in the population, if $|\nu(x|e, e', A_t) - \pi(x|A_t)| \geq |\nu(x'|g, g', A_{t'}) - \pi(x'|A_{t'})|$ in the limit, it must be because more people are replicating behavior in sequence (e, e') than in sequence

(g, g') . That is, there always exists a non-negligible part of the whole population with similarity threshold α in the interval $[\sigma(g, g'), \sigma(e, e')$ and, since thresholds are independent of preferences, and the only behavior that can be replicated is that in which the choice of x or x' was new, the result follows. ■

The main intuition of Proposition 4 is the following. $|\nu(x|e, e', A_t) - \pi(x|A_t)|$ measures how different from the underlying primitives the behavior in the representative sample is with respect to the preferences in the whole population. There are two reasons why $|\nu(x|e, e', A_t) - \pi(x|A_t)|$ might not be zero. The sample either has some agents making automatic choices or some noise. The law of large numbers causes the second concern to disappear, thus enabling us to reveal the similarity between different pairs of environments simply by comparing the previous differences for the different pairs.

The previous result is general and holds true for any underlying relationship between environments and menus. Nevertheless, whenever $A_t \cap e_t = \emptyset$, e.g., decision environments are frames, we can relax the assumption on the knowledge of preferences. In fact, we would only need to know $\pi(x|A)$ for all $A \subseteq X$ for just one alternative x and the analysis would follow as before, except that, this time, the same alternative can be used to estimate the similarity of different pairs of environments.

References

- Adeyemi, O. I. and L. C. Hunt (2007). Modelling oecd industrial energy demand: Asymmetric price responses and energy-saving technical change. *Energy Economics* 29(4), 693–709.
- Ambrus, A. and K. Rozen (2015). Rationalising choice with multi-self models. *The Economic Journal* 125(585), 1136–1156.
- Apesteguia, J. and M. A. Ballester (2015). A measure of rationality and welfare. *Journal of Political Economy* 123(6), 1278–1310.
- Bargh, J. A. (2005). Bypassing the will: Toward demystifying the nonconscious

- control of social behavior. In R. R. Hassin, J. S. Uleman, and J. A. Bargh (Eds.), *The New Unconscious*, pp. 37–58. Oxford: Oxford University Press New York.
- Bernheim, B. and A. Rangel (2009). Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics. *The Quarterly Journal of Economics* 124(1), 51–104.
- Bidwell, M. O., B. X. Wang, and J. D. Zona (1995). An analysis of asymmetric demand response to price changes: The case of local telephone calls. *Journal of Regulatory Economics* 8(3), 285–298.
- Braun, P., D. B. Nelson, and A. M. Sunier (1995). Good news, bad news, volatility, and betas. *The Journal of Finance* 50(5), 1575–1603.
- Cabral, L. and A. Fishman (2012). Business as usual: A consumer search theory of sticky prices and asymmetric price adjustment. *International Journal of Industrial Organization* 30(4), 371–376.
- Carroll, C. D., J. Slacalek, and M. Sommer (2011). International evidence on sticky consumption growth. *Review of Economics and Statistics* 93(4), 1135–1145.
- Carroll, G. D., J. J. Choi, D. Laibson, B. C. Madrian, and A. Metrick (2009). Optimal defaults and active decisions. *The Quarterly Journal of Economics* 124(4), 1639–1674.
- Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and A. Rustichini (2012). A theory of conscientiousness. *Working paper*.
- Chan, L. K., N. Jegadeesh, and J. Lakonishok (1996). Momentum strategies. *Journal of Finance* 51(5), 1681–1713.
- Chen, H., G. Noronha, and V. Singal (2004). The price response to s&p 500 index additions and deletions: Evidence of asymmetry and a new explanation. *The Journal of Finance* 59(4), 1901–1930.
- Chen, H. A., D. Levy, S. Ray, and M. Bergen (2008). Asymmetric price adjustment in the small. *Journal of Monetary Economics* 55(4), 728–737.

- Chugh, D. and M. H. Bazerman (2007). Bounded awareness: What you fail to see can hurt you. *Mind & Society* 6(1), 1–18.
- Cornelsen, L., M. Mazzocchi, and R. Smith (2018). Between preferences and references: Evidence from great britain on asymmetric price elasticities. *Working Paper*.
- Cunningham, T. (2015). Hierarchical aggregation of information and decision-making. https://dl.dropboxusercontent.com/u/13046545/paper_heuristics.pdf.
- Cunningham, T. and J. de Quidt (2015). Implicit preferences inferred from choice. *Unpublished*.
- Dardanoni, V., P. Manzini, M. Mariotti, and C. J. Tyson (2018). Inferring cognitive heterogeneity from aggregate choices. *Working Paper*.
- Dew, I. T. and R. Cabeza (2011). The porous boundaries between explicit and implicit memory: Behavioral and neural evidence. *Annals of the New York Academy of Sciences* 1224(1), 174–190.
- Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist* 49(8), 709.
- Evans, J. S. (1977). Toward a statistical theory of reasoning. *Quarterly Journal of Experimental Psychology* 29, 297–306.
- Evans, J. S. (1989). *Bias in Human Reasoning: Causes and Consequences*. Hove and London: Lawrence Erlbaum Associates, Inc.
- Evans, J. S. and K. E. Frankish (2009). *In Two Minds: Dual Processes and Beyond*. Oxford: Oxford University Press.
- Evans, J. S. and D. E. Over (1996). *Reasoning and Rationality*. New York: Psychology Press.
- Frank, R. G. and R. J. Zeckhauser (2007). Custom-Made versus Ready-to-Wear Treatments: Behavioral Propensities in Physicians’ Choices. *Journal of Health Economics* 26(6), 1101–1127.

- Fudenberg, D. and D. K. Levine (2006). A dual-self model of impulse control. *American Economic Review* 96(5), 1449–1476.
- Gilboa, I. and D. Schmeidler (1995). Case-based decision theory. *The Quarterly Journal of Economics* 110(3), 605–639.
- Gilboa, I. and D. Schmeidler (2001). *A Theory of Case-Based Decisions*. Cambridge: Cambridge University Press.
- Graf, P. (1990). Life-span changes in implicit and explicit memory. *Bulletin of the Psychonomic Society* 28(4), 353–358.
- Gul, F. and W. Pesendorfer (2001). Temptation and self-control. *Econometrica* 69(6), 1403–1435.
- Hahn, U. (2014). Similarity. *Wiley Interdisciplinary Reviews: Cognitive Science* 5(3), 271–280.
- Hellerstein, J. K. (1998). The importance of the physician in the generic versus trade-name prescription decision. *RAND Journal of Economics* 29(1), 108–136.
- Hymel, K. M. and K. A. Small (2015). The rebound effect for automobile travel: Asymmetric response to price changes and novel features of the 2000s. *Energy Economics* 49, 93–103.
- Kahneman, D. (2002). Maps of bounded rationality: A perspective on intuitive judgment and choice. *Nobel Prize Lecture* 8, 351–401.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
- Kalai, G., A. Rubinstein, and R. Spiegel (2002). Rationalizing choice functions by multiple rationales. *Econometrica* 70(6), 2481–2488.
- Kalyanaram, G. and R. S. Winer (1995). Empirical generalizations from reference price research. *Marketing Science* 14(3 supplement), G161–G169.

- Karle, H., G. Kirchsteiger, and M. Peitz (2015). Loss aversion and consumption choice: Theory and experimental evidence. *American Economic Journal: Microeconomics* 7(2), 101–20.
- Krishnamurthi, L., T. Mazumdar, and S. Raj (1992). Asymmetric response to price in consumer brand choice and purchase quantity decisions. *Journal of consumer research* 19(3), 387–400.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics* 112(2), 443–478.
- Levy, D., H. A. Chen, S. Ray, and M. E. Bergen (2004). Asymmetric price adjustment in the small: An implication of rational inattention. *TC Koopmans Research Institute Working Paper*.
- Lewis, M. S. (2011). Asymmetric price adjustment and consumer search: An examination of the retail gasoline market. *Journal of Economics & Management Strategy* 20(2), 409–449.
- Lisman, J. and E. J. Sternberg (2013). Habit and nonhabit systems for unconscious and conscious behavior: Implications for multitasking. *Journal of Cognitive Neuroscience* 25(2), 273–283.
- Manzini, P. and M. Mariotti (2007). Sequentially rationalizable choice. *American Economic Review* 97(5), 1824–1839.
- Manzini, P. and M. Mariotti (2014). Welfare economics and bounded rationality: The case for model-based approaches. *Journal of Economic Methodology* 21(4), 343–360.
- Masatlioglu, Y., D. Nakajima, and E. Y. Ozbay (2012). Revealed attention. *American Economic Review* 102(5), 2183–2205.
- Mazumdar, T., S. P. Raj, and I. Sinha (2005). Reference price research: Review and propositions. *Journal of Marketing* 69(4), 84–102.

- McAndrews, M. P., E. L. Glisky, and D. L. Schacter (1987). When priming persists: Long-lasting implicit memory for a single episode in amnesic patients. *Neuropsychologia* 25(3), 497–506.
- Medin, D. L., R. L. Goldstone, and D. Gentner (1993). Respects for similarity. *Psychological Review* 100(2), 254.
- Offerman, T. and A. Van Der Veen (2015). How to subsidize contributions to public goods: Does the frog jump out of the boiling water? *European Economic Review* 74, 96–108.
- Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences* 12(6), 237–241.
- Peltzman, S. (2000). Prices rise faster than they fall. *Journal of Political Economy* 108(3), 466–502.
- Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General* 118(3), 219.
- Rubinstein, A. (2007). Instinctive and cognitive reasoning: A study of response times. *The Economic Journal* 117(523), 1243–1259.
- Rubinstein, A. (2016). A typology of players: Between instinctive and contemplative. *The Quarterly Journal of Economics* 131(2), 859–890.
- Rubinstein, A. and Y. Salant (2012). Eliciting welfare preferences from behavioral dataset. *Review of Economic Studies* 79, 375–387.
- Rustichini, A., C. G. DeYoung, J. E. Anderson, and S. V. Burks (2016). Toward the integration of personality theory and decision theory in explaining economic behavior: An experimental investigation. *Journal of Behavioral and Experimental Economics* 64, 122–137.
- Salant, Y. and A. Rubinstein (2008). (a, f): Choice with frames. *The Review of Economic Studies* 75(4), 1287–1296.

- Schneider, W. and R. M. Shiffrin (1977). Controlled and automatic human information processing i: Detection, search and attention. *Psychological Review* 84(1), 1–66.
- Simon, H. A. (1987). Making management decisions: The role of intuition and emotion. *The Academy of Management Executive (1987-1989)*, 57–64.
- Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies* 23(3), 165–180.
- Tappata, M. (2009). Rockets and feathers: Understanding asymmetric pricing. *The RAND Journal of Economics* 40(4), 673–687.
- Tulving, E. and D. L. Schacter (1990). Priming and human memory systems. *Science* 247(4940), 301–306.
- Tversky, A. (1977). Features of similarity. *Psychological Review* 84(4), 327.
- Wadud, Z. (2014). The asymmetric effects of income and fuel price on air transport demand. *Transportation Research Part A: Policy and Practice* 65, 92–102.
- Whorf, B. L. (1941). Languages and logic. *Foundations of Cognitive Psychology*, 707.
- Woodford, M. (2019). Modeling imprecision in perception, valuation and choice. *Working Paper*.
- Yang, H. and L. Ye (2008). Search with learning: Understanding asymmetric price adjustments. *The RAND Journal of Economics* 39(2), 547–564.