CentER

Discussion Paper

No. 2010–21

DECISIONS WITH ENDOGENEOUS PREFERENCE PARAMETERS

By Patricio S. Dalton, Sayantan Ghosal

February 2010

ISSN 0924-7815

TILBURG UNIVERSITY

Decisions with Endogeneous Preference Parameters*

Patricio S. Dalton†

CentER and Economics Department, Tilburg University

Sayantan Ghosal‡

Economics Department, University of Warwick

February 2009

Abstract

We relate the normative implications of a model of decision-making with endogenous preference parameters to choice theoretic models (Bernheim and Rangel 2007, 2009; Rubinstein and Salant, 2008) in which observed choices are determined by frames or ancillary conditions.

JEL classification numbers: D01, D62, C61, I30.

Keywords: Decisions, choice, frames, standard, behavioral, welfare.

*This paper was prepared for a special issue of *Social Choice and Welfare* on Reconciling Normative and Behavioural Economics, edited by Ben McQuillin, Robert Sugden and Marc Fleurbaey. It constitutes a short and selective version of the paper presented at the University of East Anglia Conference on Reconciling Normative and Behavioural Economics on April 2008. Both authors acknowledge support from ESRC-DFID grant RES-167-25-0364. Patricio Dalton gratefully acknowledges financial support from a EST Marie Curie Fellowship, Royal Economic Society Junior Fellowship and Warwick Economics Department.

†e-mail: P.S.Dalton@uvt.nl

‡e-mail: S.Ghosal@warwick.ac.uk

1 Introduction

This paper relates decision problems introduced in Dalton and Ghosal (2008) where preference parameters are endogenous to choice theoretic models (Bernheim and Rangel, 2007, 2009), Rubinstein and Salant, 2008) in which observed choices are determined by frames or ancillary conditions.

In Dalton and Ghosal (2008), preference parameters are broadly interpreted to include psychological states (e.g. reference points, beliefs, emotions, feelings, self-esteem, will-power, aspirations, etc.) or moral states (e.g. personal commitments, individual values, etc.). There is considerable work from social psychology and economics that such psychological and moral factors affect preferences[1]. In addition, there is also a great deal of evidence which suggests that the relationship between preferences and behavior may also go in the opposite direction: what a person does (or expect to do) may define her psychological and moral states. William James (1890/1981) used the term "self-esteem" to refer to the way individuals feel about themselves which, in turn, depends on the success they have in accomplishing those things that they wish to accomplish (in Pajares and Schunk, 2001,2002). In a similar vein, Baron (2008, pp. 68) argues that emotions are partly under our control: individuals can "induce or suppress emotions in themselves almost on cue." Some people may reshape their character, so that their emotional responses change. There is extensive work in social cognitive theory by Albert Bandura[2], *reciprocal determinism* (see Bandura 1986), that views human functioning as the product of a dynamic interplay of personal, behavioral, and environmental influence: the way in which people interpret the results of their own behavior informs and alters their environments and personal factors which, in turn, inform and alter subsequent behavior through an "environmental feedback effect."

With this motivation in mind, Dalton and Ghosal (2008) study decision problems where preference parameters are potentially endogenous. A decision state is a profile consisting of both actions and preference parameters. A consistent decision state

[1] Elster (1998) provides a review on how individual preferences are affected by emotions; Sen (1977) discusses how personal values shape preferences; Appadurai (2004) studies the relationship between aspirations and behaviour; Benabou and Tirole (2002) study how self-confidence interact with preferences.

[2] See Bandura (1997, 2001) for a survey.

is a profile of actions and preference parameters where the preference parameter is generated by the action profile via the feedback effect. A standard decision problem is one where the chosen action is required to be optimal when the feedback effect from actions to preference parameters is fully internalized. A behavioral decision problem is one where the chosen actions are optimal taking preference parameters as given[3] although chosen actions and preference parameters are required to be mutually consistent[4]. In general, the outcomes of the two decision problems do not coincide.

Our model is related to dual-self problems proposed by Shefrin and Thaler (1988) and Bernheim and Rangel (2004), Benhabib and Bisin (2004) and Loewenstein and O'Donoghue (2005) among others. We interpret a decision problem with endogenous frames as a dual self intra-personal game where one self chooses actions and the other self chooses preference parameters. Consistent with this interpretation, a standard decision problem corresponds to a Stackelberg equilibrium and a behavioral decision problem corresponds to a Nash equilibrium of the dual self intra-personal game. Therefore, the outcomes of a standard decision problem correspond to the choices of a fully rational decision maker. In contrast, in a behavioral decision problem, factors that are irrelevant to a rational decision-maker will impact on choices.

Bernheim and Rangel (2007, 2009) (hereafter BR) and Rubinstein and Salant (2008) (hereafter RS) model choice problems where observed choices are determined by frames (RS) or ancillary conditions (BR). Examples of frames or ancillary conditions include the order in which candidates are listed on a ballot, default alternatives, salience of the alternative, deadline for making a choice (RS) or the point in time at which a choice is made, the manner in which information or alternatives are presented, the labeling of a particular option as the "status-quo," the salience of a default option, or exposure to an anchor (BR). For both BR and RS, frames or ancillary conditions are exogenous to the individual at the point when choices are made. Formally, in both papers, binary relations are constructed solely on observed

[3]Note that in a standard decision problem the individual fully internalizes all the possible consequences of her actions. However, in a behavioural decision problem, the individual doesn't internalize all the possible consequences of her actions and as such imposes an externality on herself.

[4]We argue that the decision framework with (potentially) endogenous preference parameters studied here can be obtained as reduced form representation of seemingly disconnected types of decision models studied in the literature on behavioral economics (see section 2 below).

choice and it is shown that such derived binary relations can be used to rank actions available to the decision-maker from a welfare viewpoint[5].

A key dilemma raised by behavioral economics is whether the construction of welfare metrics (and therefore, a benevolent social planner) can rely on observed choice alone. When choice is affected by frames or ancillary conditions, the issue is whether such frames or ancillary conditions matter from a welfare viewpoint. There are at least two diametrically opposite views: (i) the choices of a fully rational decision maker will not depend on frames or ancillary conditions and therefore when frames and ancillary conditions do affect choices they could, in principle, matter for welfare, and (ii) frames or ancillary conditions are exogenous to choices (if they appear endogenous this only because the decision problem isn't correctly specified) and are welfare irrelevant from a social planner's viewpoint[6].

We relate our decision model to the choice frameworks with frames or ancillary conditions as follows. If the starting point is that, in principle, the choices of a fully rational decision-maker do not depend on frames or ancillary conditions, we argue that the choice mappings studied in BR and RS correspond to a decision model where each frame or ancillary condition is consistent with any action chosen by the decision-maker. In a behavioral decision problem, the decision-maker treats the frame or ancillary condition as given while in a standard decision problem frames or ancillary conditions are considered to be characteristics of actions and are fully taken into account. We provide examples that illustrate the difference (details in

[5] There are some technical differences between the two frameworks (see section 4 below). Further, BR focus on the welfare implications of such framework while RS relate their framework to classical model of choice.

[6] RS define a frame as "observable information that is irrelevant in the rational assessment of the alternatives, but nonetheless affects choice" (RS, RES 2008, abstract). For BR an ancillary condition is "an exogenous feature of the choice environment that may affect behavior, but is not taken as relevant to a social planner's evaluation" (BR, NBER 2008, pp. 4). Further quotations: "(...) if the individual's behavior appears to determine the ancillary condition endogenously, the decision problem has been defined incorrectly. For example, if he can choose to make his selection .... under one of two conditions, A or B, then it is inappropriate to describe A and B as endogenous ancillary conditions. Rather, the correct ancillary condition describes the two-stage decision process." (Bernheim 2008 pp. 34); "treating a condition of choice as a welfare-relevant characteristic of available objects would seem to defy common-sense...classifying it as an ancillary condition should be relatively uncontroversial" (pp. 35).

section 3) between the two decision problems. In this case, we show that (i) the set of outcomes of a standard decision problem is a subset of the set of outcomes of a behavioral decision problem and (ii) the set of welfare optima derived solely from observed choice contains all the actions corresponding to the solution of a standard decision problem. However, for a general feedback map from actions to preference parameters, we show that the weak welfare optima derived from observed choice may have an empty intersection with, or exclude elements of, the set of actions that correspond to the outcomes of a standard decision problem.

If, on the other hand, both frames and ancillary conditions are truly exogenous to choices and welfare irrelevant for a social planner, then we argue that the choice mappings studied in BR and RS are generated in a family of distinct decision problems each indexed by a different preference parameter. In this case, the feedback from actions to preference parameters is a constant, the outcomes of both standard and behavioral decision problems coincide and the set of welfare optima derived solely from observed choice is exactly the actions corresponding to the solution of a standard decision problem.

In section 5, we analyze some policy implications of our argument. Clearly, when the outcomes of both standard and behavioral decision problems coincide, there is no case for any sort of intervention by a social planner. In contrast, in scenarios where the set of outcomes of a standard decision problem is contained in the set of outcomes of a behavioral decision problem, we argue that the "libertarian paternalism" approach to policy interventions, advocated by Thaler and Sunstein (2003), that only seeks to alter the frames or ancillary conditions (reference points) of individuals might work. In general, we argue that policy interventions should aim to ensure that decision-makers internalize the feedback from actions to preference parameters.

The remainder of the paper is organized as follows. Section 2 introduces the general model and gives some examples. Section 3 summarizes the choice-theoretic framework in BR and RS and clarifies the relationship between their frameworks and our model. Section 4 is devoted to the analysis of welfare. Section 5 discusses policy implications. The last section concludes and discusses directions for further research.

2 Decisions with Endogenous Preference Parameters

There are two sets, a set $A \subset \Re^k$ of actions, and a set $Q \subset \Re^n$ of preference parameters, where $\Re^k$ and $\Re^n$ are finite dimensional Euclidian spaces. A decision state is a pair of actions and preference parameters (a, q) where $a \in A$ and $q \in Q$.

The preferences of the decision-maker are denoted by $\succeq$, a binary relation ranking pairs of decision states in $(A \times Q) \times (A \times Q)$. The expression $\{(a, q), (a', q')\} \in \succeq$ is written as $(a, q) \succeq (a', q')$ and is to be read as "(a, q) is weakly preferred to (a', q') by the decision-maker".

There is a map $\pi : A \rightarrow Q$ modelling the feedback effect from actions to preference parameters and it is assumed that $\pi(a)$ is non-empty for each $a \in A$.

A consistent state is a decision state (a, q) such that $q \in \pi(a)$.

There are two types of decision problems studied here:

1. A *standard decision problem* (S) is one where the decision-maker chooses a pair (a, q) within the set of consistent decision states. A consistent pair (a, q) is the outcome of a standard decision problem if and only if

$$(a, q) \succeq \left(a', q'\right) \; for \; all \; \left(a', q'\right) \; with \; q' \in \pi(a').$$

Let M denote the set of all outcomes of a standard decision problem.

2. A *behavioral decision problem* (B) is one where the decision maker takes as given the preference parameter q when choosing a. A consistent pair (a, q) is the outcome of a behavioral decision problem if and only if

$$(a, q) \succeq \left(a', q\right) \; for \; all \; a' \in A.$$

Let E denote the set of outcomes of a behavioral decision problem.

Remarks[7]*:*

1. Suppose $Q = A$ and $a \in \pi(a)$. In this case, the model of decision-making studied here corresponds to situation where "the reference state usually corresponds to the decision maker's current state." (Tversky and Kahneman, 1991, pp. 1046). More generally, it turns out, we show that the decision framework with endogenous preferences is a reduced form representation of a number of positive models of

[7] The following remarks rely on results contained in Dalton and Ghosal (2008) where the interested reader can find more detailed arguments.

decision-making studied in behavioral economics such as cognitive dissonance (Akerlof and Dickens 1982), psychological games with a single active player (Geanakoplos, Pearce and Stacchetti, 1989), loss aversion games with a single player (Shalev, 2000), reference dependent consumption and personal equilibrium (Koszegi, 2005; Koszegi and Rabin, 2006, 2007), aspiration traps (Ray, 2006 and Heifetz and Minelli, 2007) and moral hazard with state-dependent utility (Dreze and Rustichini, 1999). It is also possible to show that the outcome of a behavioral decision corresponds to long-run preferences in a decision-process with changing preferences and a myopic decision-maker (von Weizsacker 1971, Hammond 1976, Pollack 1978).

2. A decision problem with endogenous preference parameters can be viewed as a dual self intra-personal game where one self (the "extrovert") chooses actions and the other self (the "introvert") chooses preference parameters: in this case, the feedback from actions to preference parameters $\pi(a)$ is simply the best-response of the "introvert" to the choices made by the "extrovert". Thus, the outcome of a behavioral decision problem corresponds to a pure strategy Nash equilibrium of the same intra-personal game while the outcome of a standard decision problem corresponds to a Stackelberg equilibrium. This captures the intuition that in a behavioral decision problem, the individual imposes an externality on herself that she doesn't fully internalize.

3. Even this special case, where $Q = \{q_1, q_2\}$ and $A = \{a_1, a_2\}$ and $q_i = \pi(a_i)$, it is possible to show that, in general, $M \neq E$ and in this case, behavioral decision outcomes have properties normally associated with two-person normal form games. Assume that the preferences $\succeq$ are represented by an utility function $u : A \times Q \to \Re$. Let $\alpha(q) = \arg\max_{a \in A} u(a, q)$. A pure action outcome of a behavioral decision is an action profile $\tilde{a}$ such that $\tilde{a} \in \alpha(\tilde{q})$ and $\tilde{q} = \pi(\tilde{a})$. A pure action outcome of a standard decision problem is one where $\hat{a} \in \arg\max_{a \in A} u(a, \pi(a))$. In each example, the decision problem is represented by a payoff table where rows are actions and columns are the utility parameters. Under the assumptions made so far, the payoffs associated with consistent decision states are on the diagonal of these payoff tables.

Example 1. *A unique inefficient behavioral decision in dominant actions: addiction*

Consider the following payoff table:

	q_1	q_2
a_1	1	-1
a_2	2	0

(Table 1)

We interpret these payoffs as an example of addiction where a_2 corresponds to *smoking* and a_1 corresponds to *not smoking* and q_i to different health states of the individual (q_2 is less healthy than q_1). In this case, in a behavioral decision problem, the decision maker always chooses a_2 as a_2 is the dominant action for each value of q: if the individual takes her health state q as given she always prefers to smoke. The unique behavioral decision outcome is (a_2, q_2) with a payoff of 0. However, note that the consistent decision state (a_1, q_1) with a payoff of 1 is the only element of M: once the individual takes the feedback from actions to health states into account, she always chooses not to smoke.

Example 2. *No pure action behavioral decision: the grass is always greener on the other side*

	q_1	q_2
a_1	0	1
a_2	1	0

(Table 2)

We interpret these payoffs as an example of a situation where the individual makes a choice between two different lifestyle so that q_i denotes a specific lifestyle and a_i denotes the action that chooses location q_i. Starting from q_1, the decision-maker prefers a_2 to a_1 while starting from q_2, the decision-maker prefers a_1 to a_2: the individual always believes that the grass is greener on the other side. There is no behavioral decision in pure strategies. The decision-maker is, however, indifferent between both the two consistent decision-states (a_1, q_1) and (a_2, q_2).

Example 3. *Multiple welfare ranked equilibria: aspirations*

	q_1	q_2
a_1	1	0
a_2	0	2

(Table 3)

We interpret these payoffs as an example of an aspiration failure. Let a_1=*as undertaking an action that perpetuates the status quo* and a_2=*undertaking that changes* *the status quo*, with q_2 = *"high aspirations"* and q_1 = *"low aspirations"* being the consistent psychological states associated with a_1 and a_2 respectively. In this example, there are two strict behavioral decision outcomes (a_1, q_1) and (a_2, q_2). Note that the pure action equilibrium (a_1, q_1) is dominated by the pure action equilibrium (a_2, q_2). When decision-maker's aspirations are high, $(a_2, q_2) \succ (a_1, q_2)$, while when her aspirations are low, $(a_2, q_1) \succ (a_1, q_1)$. Thus, the behavioral decision outcome (a_1, q_1) is an instance of an aspirations failure.

4. Example 2 demonstrates that, in general, E may be empty even when M isn't. However, given the discussion so far, a behavioral decision outcome can be interpreted as a Nash equilibrium of a two person game so that as long as A and Q are finite, a mixed strategy behavioral decision outcome always exists.

3 Choice with Ancillary Conditions or Frames

In this section and the following one, we assume that both A and Q are non-empty finite sets containing at least two elements each.

Both BR and RS study generalized (or extended) choice problems (A, q) where q is a frame or an ancillary condition. An individual's choices[8] are described by a correspondence $c(A', q) \subseteq A'$ where $A' \in \tilde{\mathbf{A}}$ the set of all non-empty subsets of A and further, that $c(A', q)$ is non-empty for all pairs (A', q). Define aP^*b[9] iff for all admissible (A', q) with $a, b \in A'$, $b \notin c(A', q)$: when aP^*b, following BR, a is strictly unambiguously chosen over b[10]. Define aR^*b iff $\sim aP^*b$: there is some generalized choice problem where both a and b are present and a is chosen. Define xI^*y iff xR^*y and yR^*x: there is some generalized choice problem where both a and b are present and a is chosen and some other generalized choice problem where both a and b are present and b is chosen. BR show that R^* is necessarily complete: for any a and b the individual must necessarily choose either a or b from any $(\{a, b\}, q)$. Moreover, they also show that none of the binary relations need be transitive although they do show that P^* is acyclic i.e. for any a_1,...,a_K, if $a_k P^* a_{k+1}$ then $\sim a_K P^* a_1$. BR,

[8]RS study choice functions while BR allow for choice correspondences.

[9]For ease of exposition, since we allow for choice correspondences and focus on normative implications of choice, we follow BR although we note that RS also derive a preference relation similar to P^*.

[10]In words, the statement "aP^*b" means that whenever a and b are available, b is never chosen.

then, go on to make the following definition:

Definition (*Weak welfare optimum, Bernheim and Rangel, 2009*): It is possible to strictly improve on a choice $a \in A$ if there is $b \in A$ such that bP^*a. When a strict improvement is impossible, a is defined as a weak welfare optimum.

BR show the following result that underpins their welfare analysis:

Result (*FACT 1, Bernheim and Rangel, 2007*): If $a \in c(A', q)$ for some (A', q), then a is a weak welfare optimum.

3.1 Characteristics of the choice object

Both BR and RS make the point that, in practice, it is difficult to draw a distinction between characteristics of elements in A and variables in Q which could also be viewed as characteristics of elements in A. In any actual decision problem studied in their papers, an individual takes the frame or ancillary condition as given when choosing an action. As already pointed out in the introduction, there are two mutually exclusive interpretations of choice with frames or ancillary conditions.

If the starting point is that the choices of a fully rational decision-maker would never depend on frames or ancillary conditions, it is possible to relate the choice scenarios in BR and RS to those studied here by assuming that $\pi(a) = Q$ for all $a \in A$ as each frame or ancillary condition is consistent with any action chosen by the decision-maker. Therefore, the outcomes of a standard decision problem corresponds the choices of a fully rational decision-maker: in this case, choices do not depend on frames or ancillary conditions as all characteristics of actions are taken into account and the decision-maker chooses a pair $(a, q) \in A \times Q$. In contrast, in a behavioral decision problem the objects of choice are $a \in A$ taking as given $q \in Q$ and therefore frames or ancillary conditions impact on choices and may matter from a welfare viewpoint.

The following two examples distinguish between the outcomes of a standard and behavioral decision problem behavioral decision problems when $\pi(a) = Q$ for all $a \in A$.

Example 4 A possible interpretation of q emphasized in both BR and RS is that of a label attached to objects of choice (such as "default options"). Consider $A = \{a, a'\}$ and $Q = \{q = \text{"}a \textit{ is the default option}\text{"}, q' = \text{"}a' \textit{is the default option}\text{"}\}$. In a standard decision problem, the individual will consider the label as a characteristic of the available objects and choose the optimal pair $(a, q) \in A \times Q$. In a behavioral decision problem, however, she will take the label as given when making a choice without considering it as a characteristic pertaining to the object, and may choose a over a' at q and a' over a at q'.

Example 5 If $Q = \{t = 1, t = 2, t = 3\}$ and $A = \{a_1, a_2, a_3\}$ where $a_t =$"*complete task at* t, *do nothing at* $t' \neq t$", $t = 1, 2, 3$. In a standard decision problem, the individual will choose both $(a_t, t) \in A \times Q$, while in a behavioral decision-problem the individual will take t as given so that, for example, at $t = 1$, a_2 will be chosen while at $t = 2$, a_3 will be chosen, thus being dynamically inconsistent.

Alternatively, if both frames and ancillary conditions are truly exogenous to choices and welfare irrelevant for a social planner, it is possible to relate the choice scenarios in BR and RS to those studied here by assuming that for all $a \in A$, $\pi(a) = q_c$ for some fixed $q_c \in Q$. In this case, each $q \in Q$ indexes a distinct decision problem (with a different π) and in each of these distinct decision problems, $M = E$.

4 Choice and Welfare

The framework studied in this paper implies that welfare ranking should take place over consistent decision states using the preference relation $\succeq$. The approach of RS and BR, on the other hand, is based solely on choice data. In this section, we compare and contrast the normative implications of our decision model with choice with frames or ancillary conditions.

We begin by assuming that $\pi(a) = Q$ for all $a \in Q$. Clearly for RS and BR, what matters for welfare is the binary relation P^* constructed solely from choices. In contrast, what matters for welfare purposes in our model is the ranking of consistent decision states. The important question then is whether the ranking over actions using the binary relation P^* constructed solely on the basis of observed choices coincides with the fixed underlying preference relation $\succsim$ *over the set of consistent decision states.*

In what follows, we assume that $\pi(a) = Q$ and introduce the following notation. For any $A' \in \tilde{\mathbf{A}}$, let $\hat{c}(A') \subseteq A'$ be defined as the choice correspondence of a standard decision maker:

$$\hat{c}\left(A'\right) = \left\{a \in A' : (a,q) \succeq \left(a',q'\right) \forall a' \in A' \textit{ and } q' \in \pi(a')\right\}$$

and let $\tilde{c}\left(A'\right) \subseteq A'$ be defined as the choice correspondence of a behavioral decision maker:

$$\tilde{c}\left(A'\right) = \left\{a \in A' : (a,q) \succeq \left(a',q\right) \forall a' \in A' \textit{ and some } q \in \pi(a)\right\}.$$

Define the *standard* binary relation $a\hat{P}^{*}b$ iff for all admissible (A',q) with $a,b \in A'$, $b \notin \hat{c}(A')$ and define the *behavioral* binary relation $a\tilde{P}^{*}b$ iff for all admissible (A',q) with $a,b \in A'$, $b \notin \tilde{c}(A')$ with $\hat{R}^{*}, \hat{I}^{*}$ and $\tilde{R}^{*}, \tilde{I}^{*}$ defined analogously.

We begin by stating the following result as an immediate consequence of the analysis presented so far:

Proposition 1. Suppose $\pi(a) = Q$ for all $a \in A$. Then, $\tilde{c}\left(A'\right) = \cup_{q \in Q} c(A',q)$ and $\tilde{P}^{*} = P^{*}$.

Proof: Since $\pi(a) = Q$ for all $a \in A$, any pair (a,q) is a consistent decision state. Thus, for each q if the decision maker solves the decision problem in a behavioral way, for each $A' \subseteq A$, each chosen action will be an element of $c(A',q)$ and therefore, every choice will be a weak individual welfare optimum in A. ■

In what follows, we assume that the decision maker always solves a behavioral decision problem so that observed choice is described by $\tilde{c}(A')$ for each $A' \subseteq A$. Let $\hat{W}$ denote the weak welfare optima corresponding to $\hat{P}^{*}$ and let $\tilde{W}$ denote the weak welfare optima corresponding to P^{*} and $\tilde{P}^{*}$. The following proposition examines the link between $\hat{W}$ and $\tilde{W}$.

Proposition 2. Suppose $\pi(a) = Q$ for all $a \in A$. Then, $M \subseteq E$, $\hat{c}\left(A'\right) \subseteq \tilde{c}\left(A'\right)$ and $\hat{W} \subseteq \tilde{W}$.[11]

Proof: If $\pi(a) = Q$, then $(a,q) \in M$ implies that $(a,q) \succeq (a',q)$ for all $a' \in A$ and therefore, $(a,q) \in E$. That $\hat{c}\left(A'\right) \subseteq \tilde{c}\left(A'\right)$ is a direct consequence of $M \subseteq E$. Finally, as $\hat{c}\left(A'\right) \subseteq \tilde{c}\left(A'\right)$, $b \notin \tilde{c}\left(A'\right)$ implies $b \notin \hat{c}\left(A'\right)$. It follows that if $a \in \hat{W}$ there is no $b \in A$ such that $b\hat{P}^{*}a$. Therefore, for every $b \in A$, there exists a non-empty $A' \subseteq A$ with $a,b \in A'$ such that $a \in \hat{c}\left(A'\right)$. As $\hat{c}\left(A'\right) \subseteq \tilde{c}\left(A'\right)$, it also follows that there is no $b \in A$ such that $b\tilde{P}^{*}a$.Therefore, $a \in \tilde{W}$ and $\hat{W} \subseteq \tilde{W}$. ■

[11]Recall that M is the set of all outcomes (a,q) of a standard decision problem and E is the set of all outcomes (a,q) of a behavioral decision problem.

The importance of proposition 2 is that in situations with exogenous frames, the set of weak welfare optima derived solely from observed choice contain all the actions corresponding to elements in M. However, not all the weak welfare optima derived from observed choices will correspond to elements in M.

Example 6. Modify example *3* so that $\pi(a) = Q$ for all $a \in A$ with no other changes. In this case, $\hat{W} = \{a_1\} \subset \tilde{W} = \{a_1, a_2\}$. Therefore, $a_2 \in \tilde{W}$ but $a_2 \notin \hat{W}$.

Proposition 2 has the following further implication: as long as M is non-empty, so is E.

Proposition 3. Suppose $\pi(a) = Q$ for all $a \in A$. Suppose $\succ$ (the strict preference relation over $A \times Q$ corresponding to $\succeq$) is acyclic. Then, M, and hence E, is non-empty.

Proof: As both A and Q are finite, and $\succ$ is acyclic, M (which corresponds to the maximal elements of $\succ$) is non-empty. As $M \subseteq E$, so is E. ■

Next, we examine under what conditions, the conclusion derived in Proposition 1 extends to the case when $\pi(a) \subset Q$ for some $a \in A$. In this case, it is possible that the weak welfare optima derived from observed choices may have an empty intersection with actions corresponding to elements in M.

In example *1*, $\hat{W} = \{a_1\} \cap \tilde{W} = \{a_2\}$ is empty whereas in example *2*, $\hat{W} = \{a_1, a_2\}$ but $\tilde{W}$ is the empty set. In both these examples, the weak welfare optima derived from observed choice have no connection with the actions corresponding to maximal consistent decision states. However, if both examples *1* and *2* are modified so that $\pi(a) = Q$ for all $a \in A$ with no other changes, then, in the now modified example *1*, $\hat{W} = \tilde{W} = \{a_2\}$ while in the modified example *2*, $\hat{W} = \tilde{W} = \{a_1, a_2\}$.

The following proposition states that a necessary and sufficient condition for obtaining this conclusion for a general $\pi(.)$ is that $M \subseteq E$.

Proposition 4. Suppose $\pi(a) \subset Q$ for some $a \in A$. Then, $\hat{W} \subseteq \tilde{W}$ iff $M \subseteq E$.

Proof: Step 1: We begin by showing that $M \subseteq E$ iff the following condition (condition C) holds: for $(a, q), (a', q') \in A \times Q$ such that $(a, q) \succeq (a', q')$, then $(a, q) \succeq (a', q)$ for some $q \in \pi(a)$. Suppose $(a, q) \in M$. As $(a, q) \succeq (a', q')$ for all $(a', q') \in A \times Q$ such that $q' \in \pi(a')$, by (C), $(a, q) \succeq (a', q)$ for some $q \in \pi(a)$. It follows that $(a, q) \in E$. Next, suppose, by contradiction, $M \subseteq E$ but (C) doesn't hold. As $(a, q) \in M$, $(a, q) \succeq (a', q')$ for all $(a', q') \in A \times P$ such that $q' \in \pi(a')$. As, by assumption, (C) doesn't hold, there exists $a' \in A$ such that $(a', q) \succ (a, q)$ for some $q \in \pi(a)$. But, then, $(a,q) \notin E$, a contradiction.

Step 2: Clearly, if $M \subseteq E$, $\hat{c}(A') \subseteq \tilde{c}(A')$ and by using argument similar to those in Proposition 1, $\hat{W} \subseteq \tilde{W}$. Next, suppose that $\hat{W} \subseteq \tilde{W}$ but $M \subsetneq E$. As $M \subsetneq E$, by step 1 of the argument, it is equivalent to assume that condition (C) doesn't hold. As (C) doesn't hold, there exists $a \in A$ such that $(a,q) \succeq (a',q')$ for some $q \in \pi(a)$ and $q' \in \pi(a')$ but $(a,q) \prec (a',q)$ for all $q \in \pi(a)$. It follows that there exists some non-empty $A' \subseteq A$, such that $a \in \hat{c}(A')$ but $a \notin \tilde{c}(A')$ and therefore, $\hat{c}(A') \subsetneq \tilde{c}(A')$. But, then, there exists $b \in A$ (take $b = a'$) and a non-empty subset of A, A', such that $a \in \hat{c}(A')$ but $a \notin \tilde{c}(A')$. Therefore, $a \in \hat{W}$ but $a \notin \tilde{W}$, a contradiction. ■

Finally, when for all $a \in A$, $\pi(a) = q_c$ for some fixed $q_c \in Q$. In this case, each $q \in Q$ indexes a distinct decision problem and in each of distinct decision problems, $M = E$. For each $q \in Q$, and any $A' \in \tilde{\mathbf{A}}$, let $\hat{c}(A',q) \subseteq A'$ be defined as the choice correspondence of a standard decision maker and let $\tilde{c}(A',q) \subseteq A'$ be defined as the choice correspondence of a behavioral decision maker. Clearly $c(A',q) = \hat{c}(A',q) = \tilde{c}(A',q)$ and the corresponding set of weak welfare optima all coincide.

5 Policy Implications

If for all $a \in A$, $\pi(a) = q_c$ for some fixed $q_c \in Q$ and each $q \in Q$ indexes a distinct decision problem with a different π, the preference ranking derived from choice data alone P^* is the only one that is appropriate for constructing any welfare metric.

What are the policy implications of our model when $\pi(a) = Q$ for all $a \in A$? As we have shown, in this case necessarily, $M \subseteq E$. With complete information about the individual preferences, if individuals choose a taking q as given, policy interventions that aim to shift the reference points of individuals (the $q's$) will ensure that individuals eventually choose actions that correspond to choices in M. So, our model can provide a set of conditions (complete information about preferences and $\pi(a) = Q$ for all $a \in A$) that justifies the "soft paternalism" approach to policy interventions recently advocated by Thaler and Sunstein (2003).

However, even in this case, there is a different form of intervention, namely relaxing the internal constraint that the individual takes q as given instead of choosing both a and q, that would achieve the same objective but wouldn't rely on any information about the preferences of the individual.

In general, however, with incomplete information about individual preferences and feedback effects, direct policy intervention along the lines of "soft" paternalism (changing frames) or "hard" paternalism (making choices directly for the individual) could make matters worse.

One possible policy recommendation in scenarios with incomplete information about an individual's preferences is to directly act on the way in which a person internalizes the feedback effect from actions to frames[12].

To fix ideas, consider example *1.* In this example, if the individual doesn't take the feedback effect from actions to psychological states into account, she always chooses a_2 (*smoking)* over a_1 *(not smoking)*; however, the reverse would be true, if she took the feedback into account. Let α, $0 \leq \alpha \leq 1$, denote the probability with which the individual does take the feedback effect into account. A straightforward computation shows that as long as $\alpha > \frac{1}{2}$, the individual will choose a_1 over a_2: as long as the individual takes the feedback effect into account with a high enough probability, she will choose not to smoke.

Likewise, in example *3,* a policy intervention that might work may be an "empowerment" policy, that would help the individual to become aware of her "internal constraints" and thus "gaining control over her own life[13]."

6 Final Remarks

The results reported here have some empirical caveats. Both, the endogenous frames and the feedback-map are key variables for policy considerations, though they are not directly observable or even inferred from choice behavior. One possible approach to identify these "unobservable" may be to use evidence from neuroscience and psychology on the neural processes driving decision making.

Extending the one-person model studied here to $n-$players, dynamic and sequential decision scenarios are topics for future research.

[12] An example of such policies could be psychotherapy sessions, projects aiming to foster people's emotional intelligence and empowerment, etc.

[13] See for instance Appadurai (2004) on the "capacity to aspire" or World Bank (2002) and Stern (2004) on Empowerment.

References

[1] Akerlof G. A. and W. T. Dickens (1982) "The Economic Consequences of Cognitive Dissonance," American Economic Review, 72, 3, 307-319.

[2] Appadurai (2004), "The capacity to aspire: Culture and the Terms of Recognition", in V.Rao and M. Walton (Eds.), Culture and Public Action, IBRD-World Bank, Washington DC.

[3] Bandura, A. (1986), Social foundations of thought and action: A social cognitive theory, Englewood Cliffs, NJ: Prentice Hall

[4] Bandura, A. (1997), Self-efficacy:The Exercise of Control, New York: Freeman.

[5] Bandura A. (2001) "Social cognitive theory: an agentic perspective," Annual Review of Psychology, 52, 1-26.

[6] Baron J. (2008) Thinking and Deciding, Cambridge University Press, 4th edition, N.Y.

[7] Benabou, R. and J. Tirole (2002) "Self-Confidence and Personal Motivation," Quarterly Journal of Economics, 117, 3, 871-915.

[8] Benhabib J. and A. Bisin (2004) "Modeling Internal Commitment Mechanisms and Self-Control: A Neuroeconomics Approach to Consumption-Saving Decisions," Games and Economic Behavior, Special Issue on Neuroeconomics, 52(2), 460-92.

[9] Bernheim, B. D. (2008), "Behavioral Welfare Economics", Schumpeter Lecture, EEA-ESEM 2008 Meetings, Milan.

[10] Bernheim D. and A. Rangel (2004) "Addiction and Cue-Triggered Decision Processes," American Economic Review, Vol. 94, 5, pp. 1558-1590.

[11] Bernheim, B.D. and A. Rangel (2007) "Toward Choice-Theoretic Foundations for Behavioral Welfare Economics", American Economic Review Papers and Proceedings 97(2), 464-470.

[12] Bernheim, B. D. and A. Rangel (2008) “Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics,” NBER WP 13737,

[13] Bernheim, B. D. and A. Rangel (2009) “Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics,” Quarterly Journal of Economics, forthcoming.

[14] Dalton P. and S. Ghosal (2008) "Behavioral Decisions and Welfare," mimeo, University of Warwick.

[15] Dreze J.H. and A. Rustichini (1999) "State-dependent Utility and Decision Theory," Ch. 16, Handbook of Utility Theory, edited by Salvador Barbera, Peter J. Hammond and Christian Seidl.

[16] Elster, J. (1998) "Emotions and Economic Theory," Journal of Economic Literature, Vol 36, No 1, 47-74.

[17] Geanakoplos, J., D. Pearce and E. Stacchetti (1989), "Psychological games and sequential rationality", Games and Economic Behavior, 1, 60-79.

[18] Hammond, P.J. (1976) "Endogenous Tastes and Stable Long-Run Choice," Journal Economic Theory, 13, 329-340.

[19] Heifetz, A. and E. Minelli (2006), "Aspiration traps," mimeo.

[20] James, W. (1890/1981), The Principles of Psychology, Cambridge, MA: Harvard University Press.

[21] Koszegi, B. (2005), "Utility from anticipation and personal equilibrium", mimeo, University of California, Berkeley.

[22] Koszegi, B. and M. Rabin (2006), "A model of reference-dependent preferences," forthcoming, Quarterly Journal of Economics.

[23] _______________________ (2007), "Reference-Dependent Risk Attitudes," American Economic Review,. 97(4), 1047-1073.

[24] Loewenstein G. and T. O’Donoghue (2005) "Animal Spirits: Affective and Deliberative Processes in Economic Behavior," mimeo, Carnegie Mellon University.

[25] Pajares F. and D.H. Schunk (2001), "Self-beliefs and school success: self-efficacy, self-concept and school achievement", chapter in R. Riding and S. Rayner (Eds.), (2001), Perception (pp. 239-266), London: Ablex Publishing.

[26] Pollak R. A. (1978) "Endogenous Tastes in Demand and Welfare Analysis," American Economic Review, 68, 2, Papers and Proceedings, 374-379.

[27] Ray, D. (2006) "Aspirations, Poverty and Economic Change," in A. Banerjee, R. Benabou and D. Mookherjee (eds) Understanding Poverty, Oxford University Press

[28] Rubinstein A. and Salant Y. (2008) "(A,f) Choice with Frames," Review of Economic Studies, 75, 1287-1296.

[29] Sen A. (1977) "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory," Philosophy and Public Affairs, 6, 4, 317-344.

[30] Shalev, J. (2000), "Loss aversion equilibrium," International Journal of Game Theory, 29, 269-87.

[31] Shefrin, H. M. and R. H. Thaler (1988) "The Behavioral Life-Cycle Hypothesis," Economic Inquiry, 26, 609-643.

[32] Stern N., J-J. Dethier and F. H. Rogers (2005) Growth and Empowerment. Making Development Happen, Munich Lectures in Economics, The MIT Press, Cambridge, Massachusetts, and London, England

[33] Thaler R. and C. R. Sunstein (2003) "Libertarian Paternalism," American Economic Review, Papers and Proceedings, 93, 2, 175-179

[34] Tversky, A. and D. Kahneman (1991), "Loss aversion in riskless choice: A reference-dependent model", Quarterly Journal of Economics, Vol. 106(4), 1039-1061.

[35] von Weizsacker C., (1971) "Notes on Endogenous Change of Tastes," Journal of Economic Theory, 3, 345-72.

[36] World Bank (2002), Empowerment and Poverty Reduction: A Sourcebook, preliminary draft, pp. 280.