

*This draft is for submission to the World Public Choice Society Meetings, Miami, 2012.
Please do not cite or quote yet.*

Identity, Expression and Act Orientated Choice

by

Nicholas Baigent

Choice Group, London School of Economics and Political Science, Houghton St.,
London, WC2A 2AE, England
n.baigent@lse.ac.uk

November, 2011.

Abstract

It is argued that rational agents may rationally be motivated to choose by properties of acts independently of properties of outcomes. Recent developments in the theory of Expressive Choice and in the Economics of Identity provide prominent examples. In a model of the limiting case in which agents only have preferences over acts, it is shown that a choice function on outcomes may not exist. The main result characterizes act preferences for which a choice function on outcomes exists by a single neutrality property. It is also shown that this result extends to mixed cases in which agents care about both acts and their outcomes. Finally, another property is shown to characterize those cases in which a choice function on outcomes is induced by a choice function on acts. In all cases, respect for some sort of “output equivalence” provides the key characterizing properties.

1 Introduction

The purpose of this paper is to consider an issue in the foundations of two prominent developments in the theory of rational choice, namely expressive choice¹ and choice based on identity². Common to both of these recent developments is that agents may care about acts independently of the outcomes to which they lead. If that is the case, the following issue arises: Under what conditions is this consistent with the usual formulation of rational choice in which agents choose whatever outcomes are highest ranked according to a preference on outcomes? The existing literature implicitly assumes, apparently without exception, that their central concepts of expressive and identity based choice require no additional restrictions on the usual formulation. However, it will be shown that the contrary is the case. If acts matter to agents independently of their outcomes, and whether or not they also care about outcomes, does require strong restrictions on the usual formulation of preference optimizing behavior.

The reason that these strong restrictions have gone unnoticed is probably the following. While the existing literature on expressive and identity based choice extensively notes and indeed emphasizes that these concepts require acts to matter independently of their outcomes, attention is focused on the implications obtained by remaining within a utility maximizing formulations. But this begs the question of whether this requires in general any further restriction on the way that acts matter independently of outcomes. For several formulations of expressive and identity based choice, this paper characterizes the cases in which this is consistent with preference/utility optimization on outcomes.

It should be emphasized that the formulation in this paper is very much more general than those formulations in the literature on expressive and identity based choice in the following sense. Both expressive and identity based formulations of choice also make many additional assumptions motivated by other aspects of the interpretations of their formulations. However, in this paper none of these are imposed. It is only assumed

¹ The earliest contributions are in Buchanan (1954), and Riker and Ordeshook (1968). Hamlin and Jennings (2011) offer a survey. Brennan (2008) offers a particularly eloquent explanation of expressive choice. See also Hillman (2010).

² See Akerlof, G.A. & Rachel E. Kranton (2000 and 2010).

that agents are wholly or partly motivated by acts independently of their outcomes in ways that are made precise.

It is easy to argue in general that there are motives for rationally choosing acts that are independent of outcomes. Though detailed examples will be given later, it is immediate that motives such as duty, honor and integrity, as well as other reasons, may focus largely or even entirely on acts irrespective of outcomes. Apart from such general considerations, two rapidly developing areas offer more detailed motivation, namely Expressive Choice and choice based on identity. The remainder of this section offers a short discussion of expressive and identity based choice that relates them to the role of acts in this paper.

In the case of Expressive Choice, agents may not be motivated to choose an act not by a preference for its outcome over the outcomes of other acts. Rather, they may be motivated by something that the act expresses. Examples that are not mutually exclusive include identity, beliefs, values, commitments and integrity among others. Such motives need not necessarily involve preferences for outcomes and may refer solely to aspects or qualities of acts. Theories of Expressive Choice have been usefully employed in the analysis of voting, though the concept is certainly not limited to such collective choice problems.

In the case of the Economics of Identity, identities prescribe particular acts and although agents may choose counter prescriptively their payoffs are then decreased as a result. Identities may be given as in the cases of gender and race, or they may be socially determined as in the cases of boss, worker, student, parent, military officer and so on. It should be emphasized that it is acts and not their outcomes that are prescribed. Models of choice with payoff sensitive prescribed acts have led to different results in Principal – Agent models, in particular in the area of organization theory, education, race and gender among others. Akerlof and Kranton (2010) provide a primer to this literature.

Though the boundary between Expressive Choice and Identity influenced choice is certainly not clear, what is clear is that acts and not their outcomes play the central role in both cases. The discussion in Hamlin and Jennings (2010) is explicit that for Expressive Choice it is choice in the sense of action that is important and not

consequences of choices and actions. For example: “First, distinguish between direct and indirect accounts of choice/action, where a direct account focuses attention on some property of the choice/act itself as a source of motivation, ... “.

The focus on acts rather than their consequential outcomes is equally clear in the Akerlof-Kranton Identity model. Indeed, Identity in their “prototype model” is an extensive form game of perfect and complete information with one additional feature. This is that at some decision nodes a particular move is prescribed and eventually terminal payoffs depend on whether the prescribed move is taken or not. Moves in an extensive form game are clearly acts and not outcomes since in general the latter are only determined at terminal nodes and not decision nodes. While there is no strategic interaction in the model in the present paper, its focus on acts and the motivation for this can be motivated by exactly the arguments given in Akerlof & Kranton (2000).

Therefore, the present paper abstracts from the typically rich detail that is typical in Expressive Choice theory and the AK Identity model in order to focus sharply on the fundamental concept that they share, a focus on acts. In doing so, the approach in the present paper is not limited to, but certainly covers, these active areas of research.

Both Expressive Choice theory and the Economics of Identity have simply augmented existing rational choice theory by allowing agents to be motivated by properties of acts independently of their outcomes. In both cases there is no apparent awareness that there may be a price to pay for the additional insights to which these extensions lead. The results in this paper show that there is indeed a price to be paid. Roughly, it may be expressed as follows. Firstly, it must be assumed that the number of acts does not exceed the number of outcomes so that there is not “more than one way of skinning a cat”. If this is not the case however, then it must be assumed that preferences for acts have a particular neutrality property. This somewhat restrictive property requires that acts that have the same outcomes must be ranked in the same way relative to any act that has different outcomes. If this property is violated, choice functions on outcomes may not exist even though choices of acts are well determined. In this case, acts must necessarily be distinguished from outcomes.

There are two ways in general in which acts may motivate choices independently of outcomes. In one, agents may have a preference order on acts. An extreme

formulation is offered in section 2 where agents explicitly maximize a preference on acts and only implicitly determine outcomes via an outcome function. Less extreme formulations are offered in section 3 in which agents care about (have a preference order over) both acts and outcomes. In section 4 choice of acts is determined explicitly only by a choice function on acts rather than by a preference on acts, and it is this that implicitly induces choices of outcomes. Finally, some conclusions are offered in section 5.

2 Preferences for acts

Let X denote a set of outcomes where $1 < |X| < \infty$ and let A denote a set of acts where $2 < |A| < \infty$. It will be assumed that $|X| < |A|$, in order for the problem raised to be interesting. If $|X| = |A|$, distinguishing between acts and their outcomes is redundant, while if $|X| > |A|$, this case can be reduced to the case $|X| = |A|$ by letting X denote the set of outcomes that can be obtained by some act in A . The assumption that there is more than one way to obtain an outcome also seems reasonable. Stealing an item and successfully begging for the money to buy it both have the outcome that the item is obtained.

Let binary relations $Q \subseteq A \times A$, assumed to be reflexive, complete and acyclic, be called act preferences. Q should not necessarily be assumed to be a weak order, given the features of acts that may be relevant to ranking them may not be easily or even plausibly traded off against each other, as the examples in the introduction make clear. However, if Q does not even induce a choice function on acts, then it can hardly be expected to induce a choice function on outcomes! Thus, it will be assumed that act preferences Q are reflexive, complete and acyclic.

A function $\mu : A \rightarrow X$ such that $\mu(A) = X$ is an outcome function.

Choice functions on a set in general make non empty selections from all non empty subsets of that set. Since an act preference Q is assumed to be reflexive, complete and acyclic, maximizing it does induce a choice function on A . It follows that

$F(\cdot, Q) : 2^A \setminus \emptyset \rightarrow 2^A \setminus \emptyset$ such that:

$F(B, Q) = \{a \in B : (\forall b \in B) aQb\} \neq \emptyset$, for all $B \in 2^A \setminus \emptyset$,

is a choice function on acts.

Section 2.1

The intuition suggested by the examples in section 1 is that ranking of outcome equivalent acts relative to one that is not outcome equivalent should be the same. That is, if $\mu(a) = \mu(a') \neq \mu(b)$ then aQb if and only if $a'Qb$ and bQa if and only if bQa' , for all $a, a', b \in A$. An act preference that has this property is called Outcome Equivalent Neutral (OEN).

The natural way to induce choices on outcomes from act preference maximization is, for all act preferences Q , consider $C_{\mu Q}(B) = \mu(F(B, Q))$. Since $\mu(F(B, Q)) \neq \emptyset$ for all $B \in 2^X \setminus \emptyset$, the function $C_{\mu Q} : 2^X \setminus \emptyset \rightarrow 2^X \setminus \emptyset$ is well defined if and only if outcome equivalent subsets of acts induce identical subsets of outcomes. That is, $C_{\mu Q} : 2^X \setminus \emptyset \rightarrow 2^X \setminus \emptyset$ is well defined if and only if, for all $B, B' \in 2^A \setminus \emptyset$: $\mu(B) = \mu(B')$ implies $\mu(F(B, Q)) = \mu(F(B', Q))$. Choice functions as defined earlier also require that, for a choice function on outcomes, $\mu(2^A \setminus \emptyset) = 2^X \setminus \emptyset$. However, this follows immediately from $|A| > |X|$ and $\mu(A) = X$. This consideration can therefore be safely ignored, and attention focused completely on OEN. Indeed, theorem 1 shows that OEN is necessary and sufficient for $C_{\mu Q} : 2^X \setminus \emptyset \rightarrow 2^X \setminus \emptyset$ to be well defined.

Theorem 1: *For all act preferences Q : Q is OEN if and only if $C_{\mu Q}$ is well defined.*

Proof: Assume that act preference Q is OEN and consider $B, B' \in 2^A \setminus \emptyset$ such that $\mu(B) = \mu(B')$. That is, B and B' are output equivalent. Let $x \in \mu(F(B, Q))$ and, for some $a \in B$, $a \in F(B, Q)$ and $\mu(a) = x$. Since B and B' are output equivalent, it follows that for some $a' \in B'$, $\mu(a') = x$ and $a'Qb'$ for all $b' \in B'$ such that $\mu(b') \neq x$. Thus, either $a' \in F(B', Q)$ or $a'' \in F(B', Q)$ for some $a'' \in B'$ such that $\mu(a'') = x$. Therefore, $x \in F(B', Q)$ and, since $x \in \mu(F(B, Q))$ is assumed: $x \in \mu(F(B, Q))$ implies $x \in \mu(F(B', Q))$. A similar argument shows that $x \in \mu(F(B', Q))$ implies

$x \in \mu(F(B, Q))$. Thus, $\mu(B) = \mu(B')$ implies that $\mu(F(B, Q)) = \mu(F(B', Q))$, and $C_{\mu Q}$ is well defined.

Let $C_{\mu Q}$ be well defined and assume contrary to hypothesis, that Q is not OEN. Thus, for some $a, a', b \in A$ such that $\mu(a) = \mu(a') \neq \mu(b)$, either aQb and $bP(Q)a'$ or bQa and $a'P(Q)b$. Consider the first case only, since the second may be dealt with in a similar way. So, given aQb and $bP(Q)a'$, it follows that $a \in F(\{a, b\}, Q)$ and $a' \notin F(\{a', b\}, Q)$. Therefore, $\mu(a) \in \mu(F(\{a, b\}, Q))$ and $\mu(a') \notin \mu(F(\{a', b\}, Q))$ which implies that $\mu(F(\{a, b\}, Q)) \neq \mu(F(\{a', b\}, Q))$. But since $\mu(\{a, b\}) = \mu(\{a', b\})$ follows from the previous assumption, $\mu(a) = \mu(a') \neq \mu(b)$, there is a contradiction with the assumption that $C_{\mu Q}$ is well defined. Thus, If $C_{\mu Q}$ is well defined, Q must be OEN.

Turning now to the issue of revealed preference, it will be shown that if Q is OEN so that, from theorem 1, $C_{\mu Q}$ is well defined then it is rationalized by a quasi transitive binary relation on X .

Since $\mu(A) = X$, the subsets $\mu^{-1}(x)$, $x \in X$, of A partition A into iso-outcome subsets relative to μ . Let A/μ denote this partition.

For all $x, y \in X$, all $a, a' \in \mu^{-1}(x)$ and all $b, b' \in \mu^{-1}(y)$, $x \neq y$, OEN ensures that the following binary relation $\tilde{Q} \subseteq (A/\mu) \times (A/\mu)$ is well defined:

$$\mu^{-1}(x) \tilde{Q} \mu^{-1}(y) \text{ if and only if } aQb, a \in \mu^{-1}(x) \text{ and } b \in \mu^{-1}(y)$$

From this, define the binary relation $R_Q \subseteq X \times X$ as follows. For all $x, y \in X$:

$$xR_Q y \text{ if and only if } \mu^{-1}(x) \tilde{Q} \mu^{-1}(y).$$

In general, a binary relation $R \subseteq X \times X$ rationalizes a choice function C on X if and only if, for all $S \in 2^X \setminus \emptyset$: $C(S) = \{x \in S : (\forall y \in S) xRy\}$.

Theorem 2: *If Q is OEN then R_Q rationalizes $C_{\mu Q}$.*

Proof: Consider $S \in 2^X \setminus \emptyset$ Given OEP and theorem 1,

$C_{\mu_Q}(S) = \mu(F(B, Q))$, $\mu(B) = S$. Let $x \in C_{\mu_Q}(S)$ and $y \in S$. Then, for some $a \in \mu^{-1}(x) \cap F(B, Q)$, and for all $b \in B$, aQb . Since Q is OEN, so that \tilde{Q} is well defined, it follows that $\mu^{-1}(x)\tilde{Q}\mu^{-1}(y)$ and xR_Qy as required for C_{μ_Q} to be rationalized by R_Q .

Since R_Q is obtained from Q , given μ , xR_Qy may be interpreted in the following way: outcome x is preferred to outcome y since, given outcome function μ , any act for which x is the outcome is at least as good as any act for which y is the outcome. Roughly, one outcome is preferred to another if what must be done to get the former is preferred to what must be done to get the latter.

3 Preferences on acts and outcomes

Of course, in the previous section, an extreme or limiting case is analyzed. Agents who care about the properties of acts and outcomes are probably more common. One way to formulate this is to consider preferences of the following kind: *Getting alternative x from act a is at least as good as getting alternative y from act b* . Such mixed preferences rankings on act-outcome pairs are now considered.

Let $\Omega(\mu) = \{(a, x) \in A \times X : \mu(a) = x\}$ and let $\tilde{Q} \subseteq \Omega(\mu) \times \Omega(\mu)$ be a mixed preference which is assumed to be a reflexive, complete and acyclic binary relation on act-outcome pairs. For all $B \in 2^A \setminus \emptyset$ and all mixed preferences \tilde{Q} , let $\tilde{G}_A(B, \tilde{Q}) = \{a \in B : (\forall b \in B)(a, \mu(a))\tilde{Q}(b, \mu(b))\}$ and let $\tilde{G}_x(\mu(B), \tilde{Q}) = \mu(G_A(B, \tilde{Q}))$.

$\tilde{G}_x(\cdot, \tilde{Q})$ is a well defined choice function on outcomes if and only if:

$$(\mu(B) = \mu(B')) \Rightarrow (\tilde{G}_x(\mu(B), \tilde{Q}) = \tilde{G}_x(\mu(B'), \tilde{Q})).$$

A mixed preference \tilde{Q} is Outcome Equivalent Neutral (OEN) if and only if, for all $a, a', b \in A$: $(\mu(a) = \mu(a') \neq \mu(b)) \Rightarrow [((a, \mu(a))\tilde{Q}(b, \mu(b)) \Leftrightarrow ((a', \mu(a'))\tilde{Q}(b, \mu(b)))) \wedge ((b, \mu(b))\tilde{Q}(a, \mu(a)) \Leftrightarrow ((b, \mu(b))\tilde{Q}(a', \mu(a')))]$.

Theorem 1a: $\tilde{G}_X(\cdot, \tilde{Q})$ is a well defined choice function on outcomes if and only if \tilde{Q} is OEN.

The proof of this theorem adapts the proof of theorem 1 in an obvious way.

Let the mixed preference \tilde{Q} be OEN. Then, for all $x, y \in X$, $R(\tilde{Q}) \subseteq X \times X$ is the derived outcome preference defined as follows: $xR(\tilde{Q})y$ if and only if $(a, x)\tilde{Q}(b, y)$ where $(a, x), (b, y) \in \Omega(\mu)$.

Theorem 2a: If \tilde{Q} is OEN then $R(\tilde{Q})$ rationalizes $G_X(\cdot, \tilde{Q})$.

Again, the proof is an obvious adaptation of the proof of theorem 2.

A similar variation of theorem 1 may be used for the case in which agents have two preferences, an act preference and an outcome preference. Outcomes may be determined according to dominance as follows. An act may be chosen if and only if there is no other act that is strictly preferred to it or an outcome strictly preferred to that from the act in question. Outcome Equivalent Neutrality as defined in section 2.1 needs no reformulation. As the second example above makes clear, the following “theorem” may be proved:

If outcomes are determined by undominated choice of acts then OEN is necessary and sufficient for the existence of an induced choice function on outcomes.

4 A Generalization: Choice Functions on Acts:

It can certainly be argued that agents may have an imperative rather than a preferential view of acts. That is, instead of having a preference on acts, they may simply have a view as to what may and may not be chosen in a way that is not representable by any binary relation. For example, any act that is consistent with professional integrity may be acceptable and any act that violates professional integrity may be unacceptable. This viewpoint suggests that agents have a choice function on acts as described in the following formal model.

For an arbitrary set Z , a choice function C_Z on Z is a function $C_Z : 2^Z \setminus \emptyset \rightarrow 2^Z \setminus \emptyset$ such that, for all $Y \in 2^Z \setminus \emptyset$, $C_Z(Y) \in 2^Y \setminus \emptyset$. In the following, use is made of choice functions C_A and C_X on Acts and Outcomes respectively.

For all $B \in 2^A \setminus \emptyset$, let B/μ be defined similarly to the definition of A/μ in section 2, and for all $b \in B$, let $[b]$ denote an element of B/μ . Thus, $[b] \in B/\mu$ denotes a largest outcome equivalent subset of acts in B .

Given an outcome function μ , a choice function C_A on acts *induces an choice function* C_X on outcomes if and only if, for all $S \in 2^X \setminus \emptyset$ and all $B \in 2^A \setminus \emptyset$ such that $\mu(B) = S$:

$$(3.1) \quad C_X(S) = \mu(C_A(B))$$

Clearly, for $C_X(S)$ to be *well defined* in (3.1) requires that, for all $B, B' \in 2^A \setminus \emptyset$, if $\mu(B) = \mu(B')$ then:

$$(3.2) \quad \mu(C_A(B)) = \mu(C_A(B'))$$

Thus, in proving that a choice function on acts induces a choice function on outcomes requires showing that (3.2) holds given whatever conditions are in question.

The condition in question will in fact be the following. Given an outcome function μ , C_A has the μ -Outcome Equivalent Contraction (μ -OEC) property if and only if, for all $B_1, B_2 \in 2^A \setminus \emptyset$ such that $\mu(B_1) = \mu(B_2)$, $B_1 \subseteq B_2$ and $B_1 \cap C_A(B_2) \neq \emptyset$:

$$(3.3) \quad C_A(B_1)/\mu = (B_1 \cap C_A(B_2))/\mu$$

This property requires that if some chosen acts are still available after deleting some acts, the subset of outcomes induced by post deletion act choices is the same as those induced by pre deletion acts that are not deleted. It is interesting to note that this property is related to one widely known as Arrow's property, first given in Arrow (1959). Indeed, μ -OEC can be obtained from Arrow's property by replacing the role of alternatives by their outcome equivalence classes. Thus, μ -OEC is Arrow's

property up to outcome equivalence. It is this property that, for given outcome function, characterizes those choice functions on acts that induce choice functions on outcomes, as the following results shows.

Theorem 3: μ and C_A induce a choice function C_X if and only if C_A has the μ -OEC property.

Proof: Assume that C_A has the μ -OEC property and consider subsets of acts,

$B_1, B_2 \in 2^A \setminus \emptyset$, such that $\mu(B_1) = \mu(B_2)$. Let $B = B_1 \cup B_2$, so that

$B/\mu = B_1/\mu = B_2/\mu$. Since $B_1 \subseteq B$ and $B_2 \subseteq B$, $C_A(B_1)/\mu = (B_1 \cap C_A(B))/\mu$,

$C_A(B_2)/\mu = (B_2 \cap C_A(B))/\mu$ and $(B_1 \cap C_A(B))/\mu = (B_2 \cap C_A(B))/\mu$ from μ -OEC

and the assumed output equivalence of B_1, B_2 and B . Therefore:

$$(3.4) \quad C_A(B_1)/\mu = C_A(B_2)/\mu$$

Now let the function $\mu^*: A/\mu \rightarrow X$ be defined as follows: For all $[a] \in A/\mu$, let

$\mu^*([a]) = \mu(a)$. For all $B' \in 2^A \setminus \emptyset$ and all $b' \in B'$, it follows that $\mu(b') = \mu^*([b'])$.

This implies:

$$(3.5) \quad \mu^*(B'/\mu) = \mu(C_A(B'))$$

$\mu(C_A(B_1)) = \mu(C_A(B_2))$ now follows from (3.4) and (3.5). This shows that (3.2) holds and a choice function on outcomes is induced by μ and C_A .

To complete the proof, assume that given μ , C_A induces C_X so that (3.2) holds, and assume that C_A does have the μ -OEC property. An argument similar to that given above easily provides a contradiction with the given assumption. The details are omitted.

4 Conclusions

The results in sections 2 and 3 both show that the usual mainstream formulation of optimizing a preference on outcomes and choices of outcomes revealing a preference on outcomes, are not generally consistent with models in which acts matter

independently of, though possibly as well as, their outcomes. Whether or not the way acts matter independently of outcomes is via a preference or a choice function on acts, strong additional requirements must be met if a choice function on outcomes is to be induced. In both cases, the additional restrictions that are required take account in a particular way, some sort of output equivalence. Such output equivalence is therefore required by the use of utility maximizing models in the literature on expressive and identity based choice.

References

- Akerlof, G.A. & Rachel E. Kranton (2000): “Economics and Identity”, *Quarterly Journal of Economics*, 715-753
- Akerlof, G.A. & Rachel E. Kranton (2010): *Identity Economics: How our identities shape our work, wages, and well being*, Princeton University Press, Princeton and Oxford.
- Arrow, K.J. (1959): “Rational Choice Functions and Orderings”, *Economica* (N.S.) 26: 121-7.
- Baigent, N. (1995): “Behind the Veil of Preference”, *Japanese Economic Review*, Vol. 46, No. 1, March 1995, pp. 88-101.
- Baigent, N. and W.Gaertner (1996): “Never Choose the Uniquely Largest: A Characterization,” *Economic Theory*, Vol.8, pp.239–249
- Binmore, K. (2009): *Rational Decisions*; Princeton University Press.
- Blume, L. (2010): Review of *Rational Decisions* by Ken Binmore, *Economic Journal*, Vol. 120, No. 542, (Feb., 2008), pp. 157-161.
- Brennan, G. (2008): “Psychological Dimensions in Voter Choice”, *Public Choice*, 137: 475–489.
- Buchanan, J.M. (1954): “Individual Choice in Voting and the Market”, *Journal of Political Economy*, 62, 334-343.
- Hamlin, A. & C. Jennings (2010): “Expressive Political Behaviour: Foundations, scope and implications”, *British Journal of Political Science*, forthcoming.
- Hargreaves-Heap, S., Martin Hollis, Bruce Lyons, Robert Sugden and Albert Weale (1992): *The Theory of Choice: A critical guide*; Blackwell, Oxford UK and Cambridge USA.
- Hillman, A.L. (2010): “Expressive Behaviour in Economics and Politics?”, *European Journal of Political Economy*, 26, 403–18.

Riker, W.H. & Peter C. Ordeshook (1968): "A Theory of the Calculus of Voting", *American Political Science Review*, 62, 25-42.

Samuelson, P.A. (1938): "A note on the pure theory of consumer's behavior", *Economica*, New Series, Vol. 5, No. 17 (Feb., 1938), pp. 61-71.

Sen, A.K. (1993): "Internal Consistency of Choice", *Econometrica*, 61, 495-521.