Revealed Identities

Miguel A. Ballester Irem Bozbay

1st August 2021

Abstract

We propose a model of identity-based choice behavior, where identity is derived from sharing characteristics with others and choice is the result of prescriptive norms associated to identities. By first modelling data as a choice correspondence that describes the support of alternatives chosen by each individual, we are able to show how to learn which choice is prescribed by which identity. By fully exploiting all available stochastic choice information, we show how to also learn the frequency with which each individual adopts each of her, possibly multiple, identities. We also provide sets of characterizing properties, and the conditions for identification, for our representation.

1 Introduction

Identity is an important determinant of a variety of individual decisions, ranging from those taken in highly socially-charged situations to the ones that may correspond to a more personal sphere.[1] This causal linkage requires a close inspection of, at least, the following questions. First, how identity is formed, an aspect that usually revolves around the fact that individuals share some characteristics with others and are aware of the social relevance of such characteristics. Ultimately, people end up dividing themselves and others into social categories, or identities, on the basis of shared physical or psychological characteristics.[2] Second, how different identities prescribe different behaviors, usually in some form of identity-based social norms or recommendations. Third, as it is evident that people are born with or end up having not one but a variety of identity-relevant characteristics, the frequency with which each of the possible identities of an individual is expressed.[3]

[1] For instance, it has been shown that food consumption is significantly related to religious identity (Atkin et al., 2019), that gender identity impacts marriage market outcomes as well as labor force participation and the income of women in a family (Bertrand et al., 2015), and that ethnic or racial identities play a role in shaping time or risk preferences (Benjamin et al., 2010), as well as labor force participation (Constant and Zimmermann, 2008). In their seminal work, Akerlof and Kranton (2000) provide the first theoretical analysis of *identity economics* and show that, by incorporating identity into a general model of preferences, some puzzling economic decisions can be explained.

[2] This idea is solidly rooted in social psychology. For some references in the psychological construction of identity, see Tajfel (1978) and Tajfel et al. (1979).

[3] We refrain from an extensive discussion on the exact determinants of these intensities, but several factors may intuitively affect it. We adopt the view that strong societal and kin educational structures may make the identities defined by some characteristics more salient than others. See Wichardt (2008)

In this paper, we propose a model of identity-based behavior entertaining all the above-mentioned aspects, and use revealed preference techniques to better learn how identity shapes individual choices. More concretely, we start by considering a finite set of relevant, dichotomous characteristics. In line with the idea of identities as social categories formed by those individuals sharing attributes, an identity is simply defined by considering the set of all individuals in the population that either possess or lack any given characteristic. For instance, if having college education is a characteristic that can be of interest for the study of identity, all individuals with college education form one identity, while all the other individuals form the opposite identity. To incorporate the role of identity-based social norms, we assume that each identity prescribes the choice of a unique alternative or course of action. These prescriptions are the main object of analysis of Section 2, where we abstract from intensity considerations and consider a qualitative, simplified version of our model. We initially model data as a choice correspondence that describes which alternatives have been chosen by each individual and use revealed preference techniques to offer a battery of useful results.[4] First, we show how to learn, from the observation of choices that have been generated by the model, the prescription map connecting identities to choices. Second, we construct two intuitive testable properties that are necessary and sufficient to claim that some data is generated by the model.

In Section 3, we briefly analyze the uniqueness of the proposed representation, and a clearcut picture emerges. A number of identities are *revealed*, in the sense that we can exactly learn which is the alternative prescribed by them. The main feature of most of these identities and alternatives is that at least one individual in society fails to choose the alternative at stake. However, if two or more alternatives are chosen by all individuals in the population, some identities remain *unrevealed*. Consequently, in Section 4, we offer an intuitive restriction of the model that guarantees that at most one such common alternative exists. This restricted model is in line with most social environments, where opposite identities (such as college educated vs non-college educated individuals) are usually conflicted, and prescribe behaviors that are somewhat opposite.[5]

We then move to the quantitative analysis of our model, in which each characteristic becomes salient with certain probability. As a result, each of the identities of an individual is expressed stochastically, ultimately generating stochastic choices.

and Kranton (2016) for a discussion, and Costa-Font and Cowell (2015) for a review of concepts of identity in economics. Understanding and identifying diverse intensities of group associations may be important beyond the context of choice as these may be linked to crucial outcomes via channels of homophily or clustering among others. For instance, homophily may slow down convergence to a consensus within a society (Golub and Jackson, 2012), and may impact social mobility by affecting decisions about investment in education (Calvó-Armengol and Jackson, 2009). See Jackson (2011) for a review of such social network effects.

[4] The choice correspondence can simply be seen as the support of the stochastic choice data of the individual. We later expand the basic model to incorporate fully the standard stochastic/quantitative choice information. We simply think that this two-stage presentation helps to understand better the technical contents of the paper.

[5] To formalise this, we simply model alternatives as elements of a vectorial space and assume that the ideal prescriptions of two opposite identities belong to the same hyperplane, but depart from the origin in opposite directions.

Section 5 provides a theoretical analysis which, to the best of our knowledge, constitutes the first foundational exercise combining identity economics and stochastic choice.[6] More concretely, we extend our revealed preference techniques to this quantitative setting, learning not only the map from identities to choices, but also the frequency with which each identity is expressed. By providing stochastic analogues of our initial properties together with a third property related to the social salience of characteristics, we characterize this fully-fledged version of our model. We conclude this analysis with some remarks on uniqueness. Section 6 provides a brief discussion of the contents of the paper, and proofs are contained in an Appendix.

1.1 Related Literature

This paper is evidently connected to a body of literature stemming from Akerlof and Kranton (2000), who incorporate identity into a general model of preferences and provide the first theoretical approach to the study of identity economics.[7] In this literature, identity is usually based on social categories but the causal influence is modelled through changes in utility.[8] While we adopt social categorization theory, our model allows to reveal the causal influence and the intensity with which each of the possible identities of an individual is expressed.

The possibility of multiple identities is discussed in Wichardt (2008), where it is suggested as a result of association with many different groups whose interests may not always be aligned. This aspect is also illustrated by a series of experiments where specific social categories are made more salient, by 'priming' them within the treatment, in order to identify their marginal effect on behavior.[9] Naturally, the priming method allows for inferring the normative prescription of different identities and these papers can be seen as critical advances in the identification of norm prescriptions and the marginal effects of changes in the social conditions. Some other empirical works, such as Sandberg (2018) and Atkin et al. (2019), provide evidence for the existence, and influence, of competing identities. Using data from the Olympic sport of dressage, the former paper shows that national identities are more salient than gender identities in the judges' biases. The latter uses data on consumption

[6] We review some contributions within these two bodies of literature in the next subsection.

[7] In subsequent projects, the authors demonstrate how identity may affect relevant economic outcomes by adapting their model to gender discrimination in the workplace, poverty and social exclusion, the household division of labor, schooling (Akerlof and Kranton, 2002) or contract theory (Akerlof and Kranton, 2005).

[8] A different but complementary approach to identity economics is by Bénabou and Tirole (2011) and Fang and Loury (2005). In the former, identity is modelled through beliefs and the latter proposes a two-stage game where individuals choose their identity in the first stage and interact in the second stage. Identity is then endogenous, and evolves as the outcome of social interaction and utility maximization as an equilibrium outcome, and eventually helps with collective action problems.

[9] For instance, a seminal work by Benjamin et al. (2010) focuses on the effect of ethnicity, race and gender on risk and time preferences. See also Benjamin et al. (2016) for the effect of religion identity on risk, time and social preferences, Cohn et al. (2015) for the impact of criminal identity on cheating behavior, and Chen et al. (2014) for the interplay of identity and coordination and cooperation behaviors. Even arbitrary social categorizations rather than natural ones may affect behavior (see, for instance, Kranton et al., 2013, 2020), and the marginal effects of the priming method have been proven to be relevant (see e.g. Eckel and Grossman, 2005, Charness et al., 2007).

of food that is identity-relevant (such as the avoidance of beef, pork and alcohol, and vegetarianism) to uncover ethnic and religious identity choices in India and to understand the determinants of identity choice. Our paper addresses these and other aspects of identity economics from a theoretical point of view.

Our paper also connects to a theoretical body of literature concerned with understanding preferences from revealed choice data, particularly in stochastic environments.[10] In most of these works, the stochastic choice data of a single individual is analyzed.[11] Like ours, some recent works focus on multiple decision makers. For instance, Chambers et al. (2020) extend the well-known logit model to investigate how a set of interacting agents influence each other. We significantly depart from all these papers by explicitly considering the role of identity in shaping individual choices. Also, we do not require data to be available from multiple menus, as it is common in all this literature. Alternatively, and more in line with panel data exercises, we simply analyze the behavior of multiple individuals over the same choice problem.

2 Identities and choices

We start by describing the basic components of our model of behavior.

1. An identity is formed by the collection of all individuals, or types, that either possess or lack a given characteristic. Formally, let C be a finite set of **characteristics** and let 2^C be the set of **types**. Given characteristic $c \in C$, **identity** $[c, 1]$ is formed by all types possessing characteristic c, i.e., by the collection of types $\{T \in 2^C : c \in T\}$, while identity $[c, 0]$ is formed by all types lacking characteristic c, i.e., by the collection of types $\{T \in 2^C : c \notin T\}$.

2. Each identity prescribes the choice of a (possibly different) alternative. Formally, let A be a set of **alternatives**. A **prescription map** $p : C \times \{0, 1\} \to A$ details which alternative is prescribed to be chosen by, or is particular of, each identity. For expositional purposes, the prescription of identity $[c, k]$ will be simply written as $p[c, k]$.

3. The choices of a type are those alternatives that are prescribed by at least one of the identities of this type, and only these alternatives. Formally, and denoting the standard indicator function by $\mathbb{1}$, the set of alternatives chosen by type T under **identity-based choice** with prescription map p are $\bigcup_{c\in C}\{p[c, \mathbb{1}_{c\in T}]\}$, which can be also written as $\bigcup_{c\in T}\{p[c, 1]\} \cup \bigcup_{c\notin T}\{p[c, 0]\}$.

To briefly exemplify the model, consider the set of 3 characteristics $\{e, h, w\}$, with e meaning *to have college education*, h meaning *to have high-income* and w

[10] See Chambers and Echenique (2016) for a review of the revealed preference approach as a general tool.

[11] These papers aim to understand the interplay between preferences and either the structure of alternatives (such as in Gul and Pesendorfer, 2006, Apesteguia et al., 2017) or the role of some cognitive factors such as reference dependence (as in Ok et al., 2015) and limited attention (as in Masatlioglu et al., 2012, Manzini and Mariotti, 2014, Cattaneo et al., 2020).

meaning *to be a woman*. This defines $2^3 = 8$ possible types in the society where, e.g., type $T = \{h\}$ is the type formed by high-income men lacking college education. It also defines $3 \times 2 = 6$ identities, where, e.g., identity $[w, 0]$ is the one formed by the collection of all men. Each identity prescribes the choice of an alternative, e.g., $p[w, 0] \in A$ is the alternative prescribed by the man identity. Under identity-based choice, the alternatives chosen by high-income men lacking college education must therefore be $\{p[h, 1], p[e, 0], p[w, 0]\}$, i.e., the alternatives prescribed by being of high-income, lacking college education and being a man, respectively. Figure 1 shows how to visually represent the model on a hypercube. Part a) highlights the identity part of the story. Each characteristic corresponds to a dimension. Each vertex of the hypercube is therefore a type, where e.g., type $\{h\}$ corresponds to the vertex $(1, 0, 1)$.[12] Each face of the hypercube represents an identity, where e.g., identity $[w, 0]$ corresponds to the base of the cube in our three-dimensional example. Part b) highlights the prescription part of the story. Each identity prescribes one alternative where, e.g., $p[w, 0] = a_2$. Each type is then choosing the alternatives prescribed by the corresponding identities where, e.g., the choices of type $\{h\}$ are $\{p[h, 1], p[e, 0], p[w, 0]\} = \{a_2, a_3, a_1\}$.

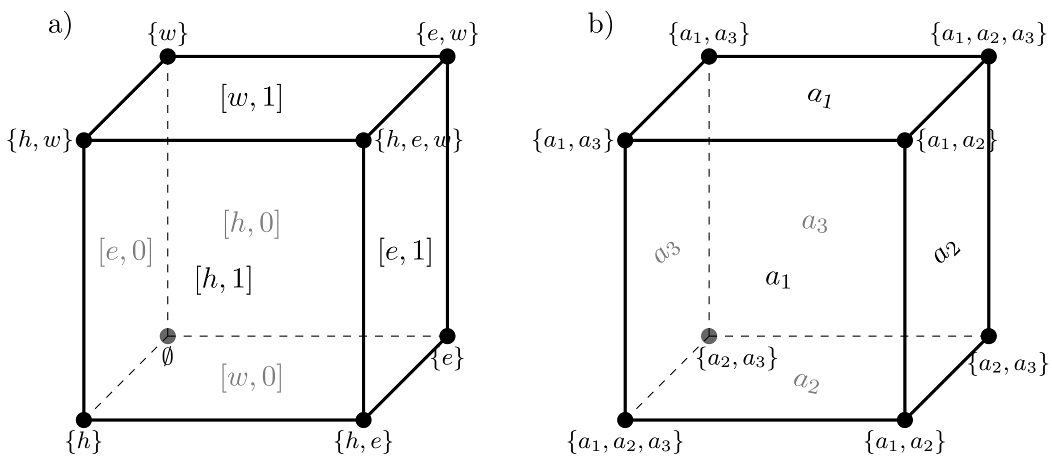

Figure 1: A graphical representation of the identity choice

2.1 Revealed Identities

We start the analysis of the model by studying the following simple question. If choices are the result of identity-based choice, is it possible to recover the prescription map from observed choices? In our first result, we show that we can learn a substantial part of this map by merely analyzing the disparity of observed actions between specific pairs of types.[13]

[12] The size of the hypercube is obviously irrelevant at this stage, but will be used later in the quantitative version of the model.

[13] Proofs are provided in the Appendix.

Lemma 1 *Suppose that the choice correspondence $\mathcal{A} : 2^C \rightrightarrows A$ is the result of identity choice, with prescription map p. Then: (i) for every $T \in 2^C$ and every $c \in C \setminus T$, $a \in \mathcal{A}_{T\cup\{c\}} \setminus \mathcal{A}_T$ implies $p[c,1] = a$, and (ii) for every $T \in 2^C$ and every $c \in C \setminus T$, $a \in \mathcal{A}_T \setminus \mathcal{A}_{T\cup\{c\}}$ implies $p[c,0] = a$.*

Consider two types that differ only in one characteristic, e.g., consider a high-income man with college education, $T = \{h, e\}$, and a high-income woman with college education, $T \cup \{w\} = \{h, e, w\}$. Suppose that a particular alternative a is used only by one of the two types, say the woman in the pair. Under identity choice and noticing that both types share all identities except that of gender, it must be the case that a is prescribed by the woman identity, $[w, 1]$. This technique is used frequently, so we now formalize the idea. Given $T \in 2^C$ and $c \in C \setminus T$, if $a \in \mathcal{A}_{T\cup\{c\}} \setminus \mathcal{A}_T$ (resp. if $a \in \mathcal{A}_T \setminus \mathcal{A}_{T\cup\{c\}}$), we say that the pair of types $(T, T \cup \{c\})$ **reveals the prescription** of alternative a by identity $[c, 1]$ (resp. identity $[c, 0]$). We also say that the identity has been revealed, and denote by $\mathcal{R}$ the set of all revealed identities. If an identity has not been revealed throughout this technique, we call it **unrevealed**, and denote by $\mathcal{U}$ the set of all unrevealed identities.

Importantly, Lemma 1 is not only helpful to identify (part of) the prescription map when data has been generated by the identity-based choice model. The revelation technique also suggests a property that can be used to test whether any choice correspondence $\mathcal{A}$ is indeed the result of identity-based choice for some prescription map.[14]

Consistency of Revealed Identities (CRI): Let $c \notin T \cup S$. If $(T, T \cup \{c\})$ reveals the prescription of alternative a by identity $[c, k]$, then: (i) if $(S, S \cup \{c\})$ reveals the prescription of alternative b by identity $[c, k]$, it must be $b = a$, and (ii) the type in $(S, S \cup \{c\})$ that shares identity $[c, k]$ must also choose alternative a.

The structure of the identity choice model requires that: (i) two different prescriptions cannot be revealed for the same identity, and (ii) a specific prescription revealed for one identity must be accompanied by the fact that all types sharing the identity also choose this alternative.

2.2 Unrevealed Identities

As of now, nothing guarantees that all identities will be revealed using the above-mentioned technique. For instance, it may well be the case that the prescription $p[c, k]$ remains unrevealed simply because all types having identity $[c, 1 - k]$ happen to also choose this alternative. This may be the result of either identity $[c, 1 - k]$, or some other combination of identities, prescribing alternatively $p[c, k]$. Importantly for our purposes, this type of situation contains vital information for the analyst.

Lemma 2 *Suppose that the choice correspondence $\mathcal{A} : 2^C \rightrightarrows A$ is the result of identity-based choice, with prescription map p. Then: (i) If $[c, k] \in \mathcal{U}$, then $p[c, k] \in$*

[14] We omit the immediate proof that identity-based choice must satisfy the property, as well as other necessity proofs proposed later in the paper.

$\bigcap_{T\in 2^C} \mathcal{A}_T$, *and (ii) for every* $a \in \bigcap_{T\in 2^C} \mathcal{A}_T$, *there exists* $c \in C$ *such that* $p[c,k] = p[c,1-k] = a$.

According to Lemma 2, unrevealed identities must prescribe the choice of alternatives that are common to all types, as otherwise the revelation method would produce results and reveal the identity.[15] Conversely, every alternative that is common to all types must indeed be prescribed by a pair of opposite, and thus unrevealed, identities, $[c,k]$ and $[c,1-k]$. The reason for this is that, if this was not the case, we would be able to construct a type such that, for each characteristic, has the identity that fails to prescribe the alternative at hand. This type would obviously fail to choose this alternative, contradicting its commonality.

Again, the ideas contained in Lemma 2 are also helpful to design a second test of identity-based choice. The property derives from the fact that Lemma 2 suggests intuitive bounds on the number of actions common to all types. Denote by $C_{\mathcal{U}}$ the set of characteristics c for which both $[c,k]$ and $[c,1-k]$ belong to $\mathcal{U}$. Then, if $\mathcal{A}$ is identity-based choice for some prescription map, the following property must also be satisfied.

Consistency of Unrevealed Identities (CUI):

$$\mathbb{1}_{C_{\mathcal{U}}\neq\emptyset} \leq |\bigcap_{T\in 2^C} \mathcal{A}_T| \leq |C_{\mathcal{U}}|.$$

The lower bound is given by the fact that, if any characteristic has both identities unrevealed, the prescription of these identities must be an alternative common to all types. Thus, at least one such alternative must exist. The upper bound is given by the fact that each alternative that is common to all types must be prescribed by two unrevealed identities of the form $[c,k]$ and $[c,1-k]$. Given the structure of identity choice, different alternatives require to be prescribed by different pairs.

2.3 Characterization Result

Importantly, the two simple tests defined by CRI and CUI are not only necessary but also sufficient for identity-based choice. That is, any correspondence $\mathcal{A}$ that passes these two tests can be seen as the result of identity-based choice for some prescription map p. Formally,

Theorem 1 *The choice correspondence* $\mathcal{A}$ *satisfies CRI and CUI if and only if there exists a prescription map* p *such that* $\mathcal{A}_T = \bigcup_{c\in C}\{p[c,\mathbb{1}_{c\in T}]\}$ *for every* $T \in 2^C$.

The proof of Theorem 1 is constructive. We start by using the revelation technique, whenever possible. Notice that CRI guarantees that each revealed identity will be assigned a unique prescribed alternative. We then complete the map by assigning

[15] Notice also that the revelation method makes apparent that revealed identities cannot prescribe common alternatives.

common alternatives to unrevealed identities, respecting the fact that each of the common alternatives must be assigned to at least two unrevealed identities of the form $[c, k]$ and $[c, 1-k]$. Notice that CUI guarantees that this is possible. In a second step, we prove that each type uses all its prescribed alternatives. This must be so because either the prescribed alternative had been revealed and hence CRI applies, or is common to all types, and thus chosen. In a third step, we prove that no type can choose a non-prescribed alternative. The reason is that any alternative a used by type T must be either common to all types (and thus prescribed by a pair of opposing identities, with one of them obviously possessed by T), or not common. In the latter case, let S be a type that is not using alternative a. We can construct a minimal sequence of types that starts at T and finishes at S, with two consecutive types only differing in a single characteristic. By considering the first pair of consecutive types that reveals the prescription of a, we learn that T has the prescribing identity, concluding the proof.

3 Identification of the Model

In this section, we briefly discuss a series of questions regarding the uniqueness of the identity-based choice representation and, as a consequence, the extent of identification of the model. Given the structure of identity-based choice, this discussion revolves around which prescription maps can produce the same choice correspondence. To formalize our result, suppose that a choice correspondence $\mathcal{A}$ has been generated by identity-based choice. Then, we denote by $P_{\mathcal{A}}$ the collection of all prescription maps p satisfying (i) and (ii) in the statement of Lemma 1 and (i) and (ii) in the statement of Lemma 2. We then have:

Theorem 2 *$P_{\mathcal{A}}$ corresponds to the set of prescription maps that generate $\mathcal{A}$. Moreover, $P_{\mathcal{A}}$ is a singleton set if and only if $|\bigcap_{T \in 2^C} \mathcal{A}_T| \leq 1$.*

Theorem 2 describes all maps that generate the same given identity-based choice correspondence. Importantly, the basic Lemmas 1 and 2 are again extremely useful in this respect because they lay down exactly the list of properties that a prescription map must satisfy in order to generate the identity-based choice correspondence $\mathcal{A}$. The structure becomes apparent. First, all these maps coincide when attention is restricted to $\mathcal{R}$, with a prescription of a non-common alternative that the revelation technique fully identifies. Hence, the reader may want to notice that the model is always fully identified when attention is restricted to $\mathcal{R}$. The prescription maps generating $\mathcal{A}$ can however differ when attention is restricted to $\mathcal{U}$. Any map assigning common actions to unrevealed identities and with a restriction to $C_{\mathcal{U}}$ that is an onto map to the set of common alternatives produces exactly the same choice correspondence. Hence, if more than one common alternative exists, full identification is not possible. To see how identification fails, suppose that alternatives a_1 and a_2 are common to all types. We know that two pairs of opposite unrevealed identities, $[c_1, k], [c_1, 1-k]$ and $[c_2, k], [c_2, 1-k]$, with each of these pairs prescribing one of the alternatives, must exist. However, the choice correspondence generated by assigning

a_1 to c_1 identities and a_2 to c_2 identities is indistinguishable from the one generated by assigning a_1 to c_2 identities and a_2 to c_1 identities.

Whenever no alternative is common to all types, all identities are revealed and the entire prescription map is fully identified. Importantly, there is also a second case in which the model is also fully identified, despite not all identities being revealed throughout the technique described in Lemma 1. Namely, when at most one alternative is common to all types. In this case, it is immediate that any unrevealed identity must prescribe exactly this unique common alternative and hence, we can also learn exactly this part of the map. Full identification is again obtained. We devote the next section to show that this feature of the model may not correspond to a rare degenerate case, but the direct consequence of imposing an intuitive extra structure to the model.

4 Identities in Conflict

In this section, we provide intuitive restrictions of the identity-based choice model in which at most one common alternative always exists. For example, suppose that A is a vector space and every pair of identities $[c, k]$ and $[c, 1-k]$ prescribe alternatives that lie in the same hyperplane, but depart from the origin in opposite directions. Formally, we assume that there exists $k_c \geq 0$ such that $p[c, k] + k_c p[c, 1-k] = 0$.[16] In this trivial model, a common alternative can only be the result of an irrelevant characteristic c for which $k_c = 0$ and, when this happens, the common alternative must correspond to the origin of the space A. Interestingly, this conclusion can be extended to a more realistic environment, that we simply called **conflicted identity-based choice** model. Formally, A is *any* subset of a normed vectorial space V and the prescription of identity $[c, k]$ is the closest alternative to the (possibly unfeasible) ideal $v[c, k] \in V$, i.e., $p[c, k] = \arg\min_{a \in A} ||a - v[c, k]||$, with ideals being conflicted.[17] We now show that any conflicted identity-based choice model generates at most one common action.

Theorem 3 *If $\mathcal{A}$ is conflicted identity-based choice,* $|\bigcap_{T \in 2^C} \mathcal{A}_T| \leq 1$.

We illustrate the intuition of Theorem 3 with a two-dimensional vector space in Figure 2, in which the set of available alternatives $A = \{a_1, a_2, a_3\}$ is considered. Alternative a_1 is common. We know from Lemma 2 that, for a_1 to be common, there must exist a characteristic c_1 such that both identities $[c_1, k]$ and $[c_1, 1-k]$ prescribe it. This is the case illustrated in part a) of the figure, where the inherent conflict involving ideals $v[c_1, k]$ and $v[c_1, 1-k]$ is such that both of these ideals are

[16] Notice that this property can indeed be the result of a more foundational principle. For example, the prescription of an identity may simply correspond to the average choice of all individuals in the population sharing this identity. In that case, it is evident that the ideals of identities $[c, k]$ and $[c, 1-k]$ must lie in the same hyperplane, but depart from the origin (in this case, the average choice of the entire society) in different directions. We abstract from these general considerations from now on.

[17] That is, for every $c \in C$, there exists $k_c \geq 0$ such that $v[c, k] + k_c v[c, 1-k] = 0$. We furthermore assume that there is always a unique minimizer.

yet sufficiently close to a_1 so as to make it the optimal choice for both identities. For a second alternative, say a_2, to be common, we should be able to find a second characteristic c_2 such that both identities $[c_2, k]$ and $[c_2, 1-k]$ prescribe it. This is impossible. The red affine hyperplane separates alternatives closer to a_1 than to a_2 from the opposite ones. Notice how, as the origin lies to the right of this affine hyperplane, the conflicted ideal related to c_2 that is located in the same half-space than the origin, say $v[c_2, 1]$, cannot lead to alternative a_2, because a_1 is closer.[18]

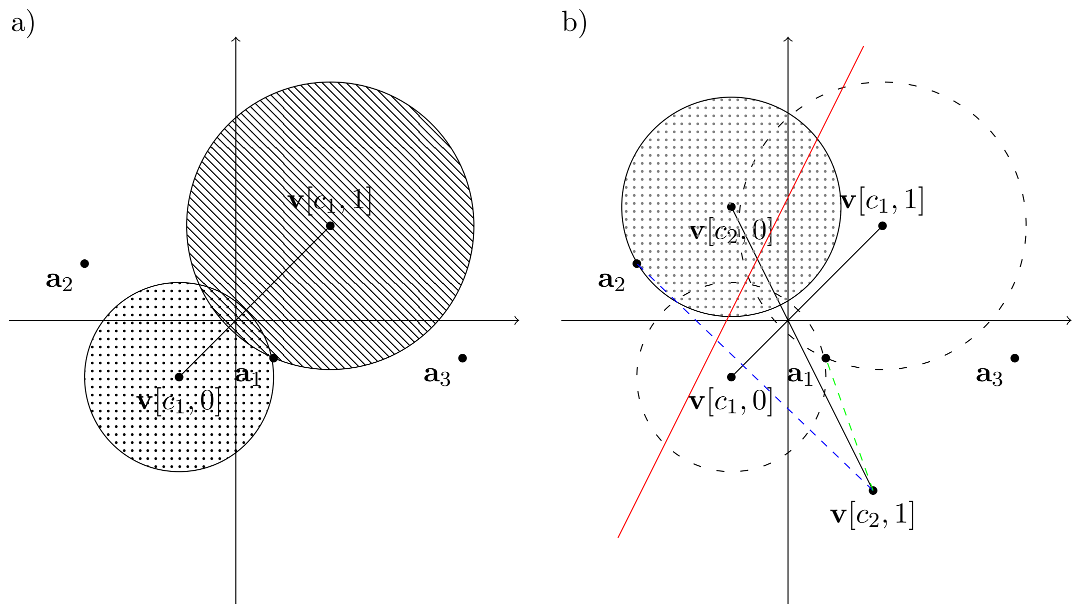

Figure 2: Conflicted Identity Choice.

5 Identities and Expression

In Section 2, we have paid special attention to the prescription map. This map provides information on how different types may present qualitatively different patterns of choice. In many occasions, there is another aspect of identity that is relevant. Namely, some of the characteristics are more salient than others and hence, the identities associated to these characteristics are expressed more often. As a result, some alternatives are chosen more often than others. We consequently enrich the data and now discuss stochastic choices.

We can think of a type as the group of people with the exact same characteristics, each possibly expressing a different identity while making a choice. The resulting data then exhibits interpersonal variation. Under this interpretation, our model focuses on the prevalence of each identity within groups. Alternatively, our framework can model a single individual of each type choosing repeatedly from the same set of options and

[18]Though this is not the case in the figure, it may of course lead to a third different alternative, say a_3.

possibly expressing a different identity each time.[19] The resulting data can then be used to study the degree of expression of an identity by an individual.

We build upon the basic identity-based choice model discussed in Section 2 by means of items (1) to (3). To model the prevalence of identities, we enrich the model by assuming the existence of a probability map $q : C \to (0, 1)$, where $q_c > 0$ captures the probability with which the identity associated to characteristic c is expressed. Then, under **identity-based stochastic choice**, the probability that type T chooses alternative a is:

$$\sum_{c \in C} q_c \mathbb{1}_{a=p[c,\mathbb{1}_{c\in T}]} = \sum_{c \in T} q_c \mathbb{1}_{a=p[c,1]} + \sum_{c \notin T} q_c \mathbb{1}_{a=p[c,0]}.$$

That is, under identity-based stochastic choice, type T chooses alternative a with the cumulated probability of expressing any identity that prescribes alternative a.

Let's recall the example in Section 2, where we considered the set of 3 characteristics $\{h, e, w\}$, with h meaning *to have high-income*, e meaning *to be educated* and w meaning *to be a woman.* In Figure 3, we illustrate the identity stochastic choice using this simple example. We represent each characteristic in $\{h, e, w\}$ with a dimension. Each edge of the rectangular cuboid has the length equal to (or proportional to) the probability expression of the characteristic in that dimension. Each vertex of the rectangular cuboid is a type, and each face represents the identity that is common to every vertex (hence, type) that intersects the face (just as shown in Figure 1 for the case of identity choice). The varying shades on the faces show the relative salience of each identity, representing how strongly it is expressed. Suppose that the probability map is given as $q_h = 0.3$, $q_e = 0.5$ and $q_w = 0.2$ and the prescribed alternative by each identity is as shown on the corresponding face of the cuboid. We show the generated choice maps for two types, namely, for $\{h, w\}$ and $\{h, e\}$.

[19]Stochastic choice is consistently observed in experiments when subjects are asked to choose from the same set of options multiple times. See, for instance, Agranov and Ortoleva (2017) where the majority of subjects select different options for the same question even when it is repeated in a row, and especially when none of the options is clearly better than the other. Deliberate randomization (Cerreia-Vioglio et al., 2019) predicts the stochastic choice behavior of subjects in this experiment.

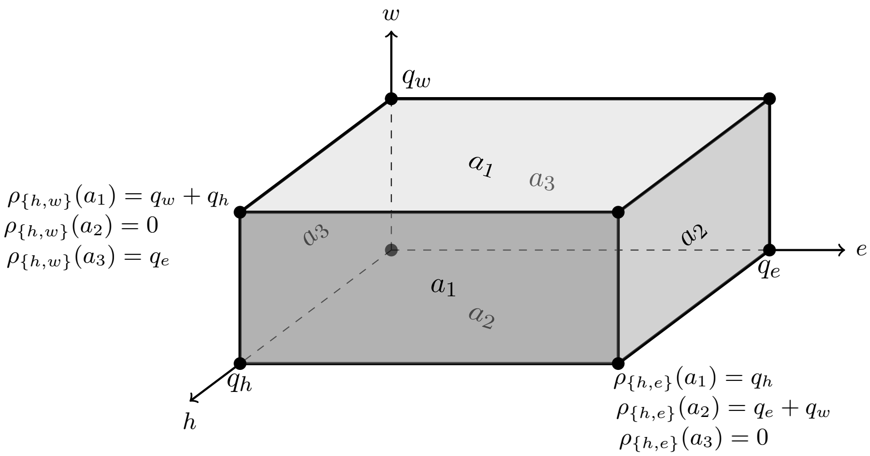

Figure 3: A graphical representation of the identity stochastic choice

Before moving on, we want to emphasize the importance of Section 2. Notice that, under identity stochastic choice, the support of any type defines a choice correspondence, that we now perfectly understand because of the analysis there performed. Naturally, choice probabilities bring extra information, not only on the degrees of expression q but, sometimes, also on the prescription map p. We start with this latter point.

5.1 Stochastic Revelation of Identities

In Section 2, the revelation technique was informative on the prescriptions made by revealed identities, bringing full identification to the restriction of p to $\mathcal{R}$. We now show that the stochastic enrichment of the model and data helps to identify some of the unrevealed identities, even when multiplicity of common actions exists. The following result is an extension of Lemma 1.

Lemma 3 *Suppose that the choice map $\rho : 2^C \times A \to [0, 1]$ is identity stochastic choice with prescription p (and expression q). Then: (i) for every $T \in 2^C$ and $c \in C \setminus T$, $\rho_{T\cup\{c\}}(a) > \rho_T(a)$, implies $p[c, 1] = a$, and (ii) for every $T \in 2^C$ and $c \in C \setminus T$, $\rho_T(a) > \rho_{T\cup\{c\}}(a)$, implies $p[c, 0] = a$.*

Suppose that the probability of choice of a particular alternative is strictly larger for a high-income educated woman than for a high-income educated man. Under identity stochastic choice, this increase in probability can only be the result of expressing the *woman* identity and hence, the woman identity must prescribe that particular alternative. As we announced, this result is simply an extension of Lemma 1. If an action belongs to the support of one type in a pair but not the other, it must be obviously the case that its choice probability increases, and hence the qualitative revelation takes place. Fortunately, stochastic revelation brings extra information. More concretely, it allows to learn the prescription of any identity $[c, k]$ such that

$p[c,k] \neq p[c,1-k]$, which includes two interesting cases not covered before: (i) unrevealed identities $[c,k]$ for which the prescription of $[c,1-k]$ had been revealed and, (ii) a pair of unrevealed identities $[c,k]$ and $[c,1-k]$ prescribing different alternatives. In consonance with our previous analysis, we say that the pair of types $(T, T \cup \{c\})$ **stochastically reveals the prescription** of alternative a by identity $[c,1]$ (resp. identity $[c,0]$) whenever $\rho_{T\cup\{c\}}(a) > \rho_T(a)$ (resp. whenever $\rho_T(a) > \rho_{T\cup\{c\}}(a)$). We similarly use the terms **stochastically revealed** and **stochastically unrevealed** and denote the corresponding sets as $\mathcal{SR}$ and $\mathcal{SU}$.[20]

5.2 Stochastic Revelation of Expression

We now briefly discuss how to learn the expression probabilities q from the data. The following result applies to any stochastic data ρ generated by identity stochastic choice.

Lemma 4 *Suppose that the choice map $\rho : 2^C \times A \to [0,1]$ is identity stochastic choice with prescription p and expression q. For every $T \in 2^C$ and $c \in C \setminus T$, $\rho_T(a) \neq \rho_{T\cup\{c\}}(a)$ implies $q_c = |\rho_T(a) - \rho_{T\cup\{c\}}(a)|$.*

Lemma 4 describes a method to learn the degree of expression of characteristics for which identities are stochastically revealed. When we compare the data from two types sharing all but one characteristic c, any change in their choice probabilities must come from the prescriptions of identities $[c,k]$ and $[c,1-k]$, respectively. As long as these opposing identities are stochastically revealed, i.e., prescribe different alternatives, we can find out the degree of expression of characteristic c by simply observing the changes in the choice probabilities. Say that we observe a disparity in the choice probability of a specific alternative when moving from the high-income educated woman to the high-income educated man. It must be the case that this action is the prescription of either the man or woman identity, and the degree of its expression must correspond to the difference in the observed choice probability.

5.3 A Characterization of Identity Stochastic Choice

We now present stochastic versions of the axioms considered in Section 2, together with a new property that guarantees the consistency of the degrees of expression of identities.[21] As in our first characterization result, all these properties follow immediately from the stochastic revelation method. Accordingly, we denote by $C_{\mathcal{SU}}$ the collection of characteristics for which the two identities are stochastically unrevealed, i.e., for which the stochastic choices of both types are always the same.

Consistency of Stochastically Revealed Identities (CSRI). Let $c \notin T \cup S$. Suppose that $(T, T \cup \{c\})$ and $(S, S \cup \{c\})$ stochastically reveal the prescription of alternatives a and b, respectively, for identity $[c,k]$. Then, $a = b$.

[20]Hence, any revealed identity is stochastically revealed, i.e., $\mathcal{R} \subseteq \mathcal{SR}$ and any stochastically unrevealed identity is unrevealed, i.e., $\mathcal{SU} \subseteq \mathcal{U}$.

[21]Alternatively, we may impose CRI and CUI over the support of ρ and slightly strengthen this new property. The results are available upon request.

Consistency of Stochastically Unrevealed Identities (CSUI).

$$\mathbb{1}_{C_{\mathcal{SU}}\neq\emptyset} \leq |\bigcap_{T\in 2^C} \mathcal{A}_T| \leq |C_{\mathcal{SU}}|.$$

Consistency of Expression (CE). Let $c \notin T \cup S$. Then, for every $a \in A$:

$$\rho_T(a) - \rho_{T\cup\{c\}}(a) = \rho_S(a) - \rho_{S\cup\{c\}}(a).$$

Theorem 4 *The stochastic map ρ satisfies CSRI, CSUI and CE if and only if there exists a prescription map p and a probability expression q such that $\rho_T(a) = \sum_{c\in C} q_c \mathbb{1}_{a=f[c,\mathbb{1}_{c\in T}]}$ for every $T \in 2^C$ and $a \in A$.*

Theorem 4 provides a characterization of identity stochastic choice. We have already discussed that CSRI, CSUI and CE are necessary properties for identity stochastic choice. To show that they are also sufficient, we proceed in three steps. First, we construct a prescription map p using data ρ, using the stochastic revelation method. By CSRI, every such identity must be assigned a unique alternative. For stochastically unrevealed identities, we assign an alternative common to all types, exhausting all common alternatives and guaranteeing that unrevealed identities $[c, k]$ and $[c, 1-k]$ prescribe the same alternative. CSUI guarantees that this can be done. In a second step, we construct a probability distribution q over the set of characteristics. For stochastically revealed identities, this is simply the difference of choice probabilities that prompted the revelation. Notice that CE guarantees that q_c is consistently defined this way. For a stochastically unrevealed identity assigned to a, we need to consider the difference between the choice probability of a in a type and the cumulated probability of stochastically revealed identities leading to a. This difference must be the result of stochastically unrevealed identities, so we can allocate it to all the ones prescribing a proportionally. In the final step, we show that ρ is identity stochastic choice with a map p and a probability distribution q.

Given data with the identity stochastic choice representation, can we identify the prescription map p and the probability distribution q that gives rise to the data? We answer this question in the next theorem (whose proof we omit as it follows from the discussions in the proof of Theorem 4 and Section 3.)

Theorem 5 *Let $\rho : 2^C \times A \to [0, 1]$ be identity stochastic choice, with map p and probability distribution q. Then, p is uniquely identified if and only if $|\bigcap_{T\in 2^C} \mathcal{A}_T| \leq 1$, and q is uniquely identified if and only if $|C_{SU}| \leq 1$ (which also guarantees $|\bigcap_{T\in 2^C} \mathcal{A}_T| \leq 1$).*

Just as in the identity choice model, if we restrict attention to the set $\mathcal{SR}$, the model is fully identified. When there is at most one common alternative to all types, the prescription map p is identified uniquely following the discussion in Section 3. The full identification of q is obtained when there is at most one characteristic in $C_{\mathcal{SU}}$ which also guarantees that there is a unique prescription map giving rise to the data.

6 Final Remarks

In this paper, we have proposed and studied a model of choice based on identity, where each individual adopts each of her identities with some probability, leading to a choice that is normatively prescribed by this identity. We have first studied the choice correspondence of each type, paying attention to the prescription map and how to identify it. We have then extended our analysis to incorporate the salience of characteristics and their strength of expression. Both the reduced and the fully-fledged models apply to a variety of settings where the choices of decision makers can be expressed as a function of their characteristics. We have shown that both models can be tested using simple properties and can be fully identified in intuitive settings.

An obvious extension of the model can be obtained by linking each identity to a rational preference or utility function over the set of alternatives and expanding the analysis to choices from different menus. Fixing a menu, we can apply our analysis and, under the conditions specified in the paper, we can identify the top ranked alternative in the menu for the preference prescribed by each identity. Notice that, for any type, we essentially observe the collection of menu maximal elements of a collection of preferences. Such choice correspondence is naturally not rationalizable by a single preference, but is known to be pseudo-rationalizable[22].

We have defined the expression of each identity as the salience of the characteristic giving rise to the identity. Alternatively, we can think of type dependent salience and define the map q_T for each type T, mapping each identity in T to the identity's strength of expression by that type. This would allow us to study settings where the salience of being, say, educated, might be high or low depending on which other identities there are in the same type. Indeed, the relative importance of being educated may vary depending whether someone has high income or not, or is an immigrant or not. Although this approach would provide important insights, this degree of generality comes with the cost of restricted identification of the model (except for data where identities are very diverse)[23].

A Appendix

Proof of Lemma 1: Suppose that $\mathcal{A}$ is identity choice, with prescription map p. Let $T \in 2^C$ and $c \in C \setminus T$. Types T and $T \cup \{c\}$ share all identities except that relative to characteristic c. Under identity choice, if alternative a is used by one of these two types but not by the other, it can only be prescribed by the identity, relative to characteristic c, that the first type has. By noticing that T has identity $[c, 0]$ and $T \cup \{c\}$ has identity $[c, 1]$, the result follows. $\square$

Proof of Lemma 2: Suppose that $\mathcal{A}$ is identity choice, with prescription map p. To prove claim (i), suppose by contradiction that identity $[c, k] \in C \times \{0, 1\}$ is unrevealed

[22] The characterization of pseudo-rationalizable choice was first introduced in Aizerman and Malishevski (1981). A discussion of this result as well as a proof of the characterization can be found in Moulin (1985).

[23] The results for this model are available upon request.

but there exists T such that $p[c,k] \notin \mathcal{A}_T$. By definition of identity choice, it must be that T has identity $[c, 1-k]$. Let T' be the type sharing all identities with T, except $[c,k]$. Then, $p[c,k] \in \mathcal{A}_{T'} \setminus \mathcal{A}_T$, which contradicts the fact that $[c,k]$ is unrevealed. To prove claim (ii), let $a \in \bigcap_{T \in 2^C} \mathcal{A}_T$. Suppose, by way of contradiction, that no characteristic exists with both associated identities prescribing a. Then, for every $c \in C$, we can find $k_c \in \{0,1\}$ such that identity $[c, k_c]$ does not prescribe a. It is then immediate to see that the type T defined by $c \in T \Leftrightarrow k_c = 1$ cannot choose a, contradicting the fact that a is common to all types. This concludes the proof. □

Proof of Theorem 1: The necessity of both CRI and CUI has been discussed. We now prove sufficiency. To do so, we proceed in three steps. First, we use the revelation technique to define a prescription map p. Second, we show that, for every type T, and for every identity $[c,k]$ of T, $p[c,k]$ belongs to $\mathcal{A}_T$. Third, we complete the proof by showing that $\mathcal{A}_T$ contains no other action.

Step 1. We first define part of the prescription map $p : C \times [0,1] \to A$, throughout the revelation method. That is, we define $p[c,k] = a$ whenever a has been revealed to be the prescription of $[c,k]$ for some pair of types. Notice that CRI, part (i), guarantees that at most one alternative will be assigned to each identity. Moreover, if $\mathcal{U} = \emptyset$, the map is completely defined and we are done. Suppose then that $\mathcal{U} \neq \emptyset$. Consider $C_{\mathcal{U}} = \{c_1, c_2, \ldots, c_l\}$ and $\bigcap_{T \in 2^C} \mathcal{A}(T) = \{a_1, a_2, \ldots, a_m\}$. From CUI, we know that it must be $1 \leq m \leq l$. We can then set $p[c_i, k] = p[c_i, 1-k] = a_i$ whenever $i \leq m$, and complete the map p by simply setting $p[c,k] = a_m$ for any other unrevealed identity. This concludes step 1.

Step 2. We now show that, given the prescription map p, all alternatives prescribed by the identities of any type T are chosen by this type, i.e., $\bigcup_{c \in C} \{p[c, \mathbb{1}_{c \in T}]\} \subseteq \mathcal{A}_T$. To see this, consider any identity $[c,k]$ of type T. If this identity is unrevealed, the construction of p guarantees that its prescribed alternative is common to all types and, hence, belongs to $\mathcal{A}_T$. If identity $[c,k]$ has been revealed, part (ii) in CRI guarantees that every type with identity $[c,k]$, such as type T, must choose the prescribed alternative. This concludes the proof of step 2.

Step 3. We finally show that, given the prescription map p, only the alternatives prescribed by the identities of any type T can be chosen by this type, i.e., $\bigcup_{c \in C} \{p[c, \mathbb{1}_{c \in T}]\} \supseteq \mathcal{A}_T$. Suppose, by way of contradiction, that $a \in \mathcal{A}_T$, but $a \notin \bigcup_{c \in C} \{p[c, \mathbb{1}_{c \in T}]\}$. Consider a linear order $<$ over the set of characteristics and denote alternatives accordingly as $1 < 2 < \cdots < |C|$. Given type T, consider the unique sequence of types $T^0, T^1, \ldots, T^{|C|}$ with the property that T^i has identity $[j, \mathbb{1}_{j \in T}]$ if and only if $j > i$. That is, $T^0 = T$, T^1 only differs from T in the identities defined by the first characteristic, T^2 differs from T exactly in the first two identities, and so on and so forth, with $T^{|C|}$ sharing no identity whatsoever with T. We now prove recursively that alternative a is chosen by all the types in this sequence. By assumption, $a \in T^0$. For the inductive part, suppose that we have proved a to be chosen by all types up to T^n and suppose by contradiction that $a \notin \mathcal{A}_{T^{n+1}}$. Then, the pair (T^n, T^{n+1}) reveals that a is prescribed by identity $[n+1, \mathbb{1}_{(n+1) \in T}]$ and we know that $a = p[n+1, \mathbb{1}_{(n+1) \in T}]$ and $a \in \bigcup_{c \in C} \{p[c, \mathbb{1}_{c \in T}]\}$, a contradiction. This proves the induction argument and shows that a is chosen by all types in the sequence. We now show that a must indeed be chosen by all types in 2^C. To see this, notice that the previous reasoning applies to any linear order over the set of characteristics and every type S must belong to at least one sequence induced by one such linear order. Formally, S belongs to the sequence induced by every order $<_{TS}$ in which all characteristics belonging to $S \cap T$ come after all characteristics not belonging to $S \cap T$. Hence, a belongs to $\bigcap_{T \in 2^C} \mathcal{A}(T)$. From our construction of p, we know that a has been assigned to at least one pair of identities of the form $[c, k]$ and $[c, 1-k]$. But then, it must belong to $\bigcup_{c \in C}\{p[c, \mathbb{1}_{c \in T}]\}$, which is a contradiction. This concludes step 3 and the proof. □

Proof of Theorem 2: Suppose that $\mathcal{A}$ is identity choice. Let p and p' be two maps in $P_{\mathcal{A}}$. We claim that, for every type T, the set of prescribed alternatives for T must be equal in the two maps. Suppose, by way of contradiction, that this is not true. Let $a = p[c, k]$, with type T having identity $[c, k]$, but assume that p' does not prescribe a for type T. From part (ii) in the statement of Lemma 2, we know that a cannot be chosen by all types, as otherwise there should exist $[c, k]$ and $[c, 1-k]$ with $p'[c, k] = p'[c, 1-k] = a$, contradicting the assumption that p' is not prescribing a for type T. From part (i) in the statement of Lemma 2, we know that $[c, k]$ must be revealed, as otherwise $p[c, k]$ would be chosen by all types. Since $[c, k]$ is revealed, it is evident from the statement of Lemma 1 that $p[c, k] = p'[c, k]$, a contradiction. This shows that the claim is true. Moreover, since any map generating $P_{\mathcal{A}}$ must belong to $P_{\mathcal{A}}$, as proved in Lemmas 1 and 2, we are done.

As a consequence, if there is no common chosen alternative, all identities must be revealed and since $p = p'$ over the revealed identities, uniqueness of the prescription map follows. If there is only one common alternative, notice that all unrevealed identities must prescribe it, and the uniqueness result follows again. When there are at least two common alternatives to all types, $a_1 \neq a_2$, the set C_U is not a singleton and any permutation of p over the characteristics in C_U represents a different prescription map with the same observed identity choice correspondence. This concludes the proof. □

Proof of Theorem 3: We proceed by contradiction. Suppose that $\mathcal{A}$ is conflicted identity choice, but $\{a_1, a_2\} \subseteq \bigcap_{T \in 2^C} \mathcal{A}_T$, with $a_1 \neq a_2$. Since $\mathcal{A}$ is identity choice, we know that there exist two characteristics $c_1, c_2 \in C$ satisfying $p[c_1, 0] = p[c_1, 1] = a_1$ and $p[c_2, 0] = p[c_2, 1] = a_2$. Moreover, we know that the ideals $v[c_i, 0]$ and $v[c_i, 1]$ are conflicted, and must lie in opposite directions of the same hyperplane H_i. Consider the half-space V_{12} (resp. V_{21}) of vectors that are strictly closer to a_1 than to a_2 (resp. strictly closer to a_2 than to a_1), and the affine hyperplane $\hat{H}_{12}$ that separates these two half-spaces. Suppose first that $0 \in V_{12} \cup V_{21}$ and let, w.l.o.g., $0 \in V_{12}$. If H_2 is an affine translation of $\hat{H}_{12}$, both ideals $v[c_2, 0]$ and $v[c_2, 1]$ will belong to V_{12}, and hence $p[c_2, 0]$ and $p[c_2, 1]$ cannot be equal to a_2, a contradiction. If H_2 is not one such translation of $\hat{H}_{12}$, then the hyperplanes H_2 and $\hat{H}_{12}$ must intersect. Given the conflicted ideals, at least one of the two ideal alternatives $v[c_2, 0]$ or $v[c_2, 1]$ must belong to the same half-space to which 0 belongs, and hence must be strictly closer to a_1 than to a_2. Its prescribed alternative cannot be a_2, which is a contradiction.

Then, it must be $0 \notin V_1 \cup V_2$, i.e., $0 \in \hat{H}_{12}$. It must be either $H_2 = \hat{H}_{12}$ or not. The former case is impossible because of the assumption of unique minimizers. The latter case guarantees that exactly one of the ideal alternatives, $v[c_2, 0]$ or $v[c_2, 1]$, is closer to a_1 than to a_2, which produces a contradiction that concludes the proof. $\square$

Proof of Lemma 3: Suppose that ρ is identity stochastic choice, for p and q. Let $T \in 2^C$ and $c \in C \setminus T$. Types T and $T \cup \{c\}$ share all identities except the ones relative to characteristic c. If the choice probability of alternative a is strictly larger for one of the two types, it can only be as a result of expressing the identity relative to characteristic c, that must therefore prescribe alternative a. By noticing that T has identity $[c, 0]$ and $T \cup \{c\}$ has identity $[c, 1]$, the result follows. $\square$

Proof of Lemma 4: Suppose that ρ is (f,q)-identity-based behavior. Let $T \in 2^C$ and $c \in C \setminus T$. Whenever $\rho_T(a) \neq \rho_{T\cup\{c\}}(a)$, Lemma 3 guarantees that $f[c, k] = a$ for some $k \in \{0, 1\}$. Under (f,q)-identity-based behavior, as types T and $T \cup \{c\}$ share all identities except that relative to characteristic c, it follows that $q_c = |\rho_T(a) - \rho_{T\cup\{c\}}(a)|$, as desired. $\square$

Proof of Theorem 4: The necessity of the axioms has been discussed. We now prove sufficiency. To do so, we proceed in three steps. In the first two steps, we define a prescription map p and a probability distribution q using the revelation methods described in the text. In the final step, we show that ρ corresponds to identity stochastic choice with map p and probability distribution q.

Step 1. We first define a prescription map $p : C \times [0, 1] \to A$ throughout the stochastic revelation method described in the text. That is, we define $p[c, k] = a$ whenever a has been stochastically revealed to be the prescription of $[c, k]$ for some pair of types. Notice that CSRI guarantees that at most one alternative will be assigned to each identity throughout this revelation method. If $C_{\mathcal{SU}} = \emptyset$, the map is completely defined and we are done. Otherwise, let $C_{\mathcal{SU}} = \{c_1, c_2, \ldots, c_l\} \neq \emptyset$ and $\bigcap_{T\in 2^C} \mathcal{A}(T) = \{a_1, a_2, \ldots, a_m\}$. From CSUI, we know that it must be $1 \leq m \leq l$. We can then set $p[c_i, k] = p[c_i, 1-k] = a_i$ whenever $i < m$, and $p[c_i, k] = p[c_i, 1-k] = a_m$ otherwise. This completes step 1.

Step 2. We now define the probability distribution q. For every $c \notin C_{\mathcal{SU}}$, we know that $[c, 1]$ and $[c, 0]$ are stochastically revealed. Suppose that the corresponding prescriptions are a and $b \neq a$, respectively. We can simply select any T such that $c \notin T$ and define $q_c = \rho_{T\cup\{c\}}(a) - \rho_T(a)$. CE guarantees that this is a single-valued map on $C \setminus C_{\mathcal{SU}}$. For any $c_i \in C_{\mathcal{SU}}$, $i < m$, we can set $q_{c_i} = \rho_C(a_i) - \sum_{c \notin C_{SU}} q_c \mathbb{1}_{a_i = p[c,1]}$. Finally, for any $c_j \in C_{\mathcal{SU}}$, $j \geq m$, we can set $q_{c_j} = \frac{\rho_C(a_m) - \sum_{c\notin C_{\mathcal{SU}}} q_c \mathbb{1}_{a_m = p[c,1]}}{l-m+1}$. This concludes step 2.

Step 3. Denote by $\hat{\rho}_T(a) = \sum_{c\in C} q_c \mathbb{1}_{a=f[c,\mathbb{1}_{c\in T}]}$ the identity stochastic choice generated by p and q. To conclude the proof, we just need to prove that $\rho_T(a) = \hat{\rho}_T(a)$ for every $T \in 2^C$ and $a \in A$. We show this throughout a series of claims.

Claim 1: For every $T \in 2^C$ and $a \in A$: $\rho_T(a) > 0 \Leftrightarrow a = p[c, \mathbb{1}_{c\in T}]$ for some $c \in C$.

Proof of *Claim 1:* This is immediate from the fact that the support of ρ is identity choice, and $q > 0$. $\square$

Define:

$$\delta_c(a) = \begin{cases} -q_c & \text{if } a = p[c,0] \neq p[c,1] \\ q_c & \text{if } a = p[c,1] \neq p[c,0] \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Claim 2: For every $T \in 2^C$ and $a \in A$, $\rho_T(a) = \rho_\emptyset(a) + \sum_{c\in T} \delta_c(a)$.

Proof of *Claim 2:* Consider an arbitrary ordering of the characteristics and label characteristics accordingly as $c_1, c_2, \ldots c_{|C|}$. Define the sequence of types $T^0 = \emptyset$, $T^1 = \{c_1\}$, $T^2 = \{c_1, c_2\}$, $\ldots$, $T^{|C|} = C$. From the construction of q, we know that we can write:

$$\begin{aligned}
\rho_{T^1}(a) &= \rho_\emptyset(a) + \delta_{c_1}(a) \\
\rho_{T^2}(a) &= \rho_{T^1} + \delta_{c_2}(a) = \rho_\emptyset(a) + \delta_{c_1}(a) + \delta_{c_2}(a) \\
&\vdots \\
\rho_{T^n}(a) &= \rho_{T^{n-1}}(a) + \delta_{c_n}(a) = \rho_\emptyset(a) + \sum_{c\in T^n} \delta_c(a)
\end{aligned}$$

where $n \in \{1, \ldots |C|\}$. Given that this reasoning applies to any linear order of characteristics, we can simply consider an order that places first the characteristics of T and type T will belong to the sequence. The result follows. $\square$

Claim 3: If $a \notin \bigcap_{T\in 2^C} \mathcal{A}(T)$, then $\rho_\emptyset(a) = \hat{\rho}_\emptyset(a)$.

Proof of *Claim 3*: Given that a is not common, there must exist a type T for which $\rho_T(a) = 0$. By claim 2, it must be $\rho_T(a) = 0 = \rho_\emptyset(a) + \sum_{c\in T} \delta_c(a)$ and, hence, $\rho_\emptyset(a) = -\sum_{c\in T} \delta_c(a) = \sum_{c\in T} q_c \mathbb{1}_{a=p[c,0]\neq p[c,1]} - \sum_{c\in T} q_c \mathbb{1}_{a=p[c,1]\neq p[c,0]}$. Given that $\rho_T(a) = 0$, Claim 1 guarantees that there does not exist $c \in T$ with $p[c,1] = a$ and hence, $\rho_\emptyset(a) = \sum_{c\in T} q_c \mathbb{1}_{a=f[c,0]}$. Claim 1 also guarantees that there does not exist $c \in C \setminus T$ with $p[c,0] = a$ and hence, we can rewrite the expression summing over all characteristics, leading to $\rho_\emptyset(a) = \sum_{c\in C} q_c \mathbb{1}_{a=p[c,0]}$, which is the desired result. $\square$

Claim 4: If $a \notin \bigcap_{T\in 2^C} \mathcal{A}(T)$, then $\rho_T(a) = \hat{\rho}_T(a)$ for every $T \in 2^C$.

Proof of *Claim 4*: Let $a \notin \bigcap_{T\in 2^C} \mathcal{A}(T)$. Then,

$$\begin{aligned}
\rho_T(a) &= \rho_\emptyset(a) + \sum_{c\in T} \delta_c(a) \text{ (by Claim 2)} \\
&= \sum_{c\in C} q_c \mathbb{1}_{a=p[c,0]} + \sum_{c\in T} \delta_c(a) \text{ (by Claim 3)} \\
&= \sum_{c\in C} q_c \mathbb{1}_{a=p[c,0]} + \sum_{c\in T} q_c \mathbb{1}_{a=p[c,1]\neq p[c,0]} - \sum_{c\in T} q_c \mathbb{1}_{a=p[c,0]\neq p[c,1]} \text{ (by construction)} \\
&= \sum_{c\in C} q_c \mathbb{1}_{a=p[c,0]} + \sum_{c\in T} q_c \mathbb{1}_{a=p[c,1]} - \sum_{c\in T} q_c \mathbb{1}_{a=p[c,0]},
\end{aligned}$$

where the last step is true because there is no $c \in C$ with $p[c,k] = p[c,1-k] = a$ as $a \notin \bigcap_{T \in 2^C} \mathcal{A}(T)$. Notice that the last line can be rewritten as

$$\begin{aligned}\rho_T(a) &= \sum_{c \in C \setminus T} q_c \mathbb{1}_{a=p[c,0]} + \sum_{c \in T} q_c \mathbb{1}_{a=p[c,1]} \\ &= \sum_{c \in T} q_c \mathbb{1}_{a=p[c,\mathbb{1}_{c\in T}]},\end{aligned}$$

which is the desired result. $\square$

To complete the proof, it remains to show the following.

Claim 5: For each $T \in 2^C$ and $a \in \bigcap_{T \in 2^C} \mathcal{A}(T)$, $\rho_T(a) = \hat{\rho}_T(a)$.

Proof of *Claim 5*: By Claim 2, we can write

$$\begin{aligned}\rho_C(a_i) &= \rho_\emptyset(a_i) + \sum_{c \in C} \delta_c(a_i) \\ &= \rho_\emptyset(a_i) + \sum_{c \in C} q_c \mathbb{1}_{a_i=p[c,1]\neq p[c,0]} - \sum_{c \in C} q_c \mathbb{1}_{a_i=p[c,0]\neq p[c,1]}. \end{aligned} \tag{2}$$

Notice that the characteristics that are not in $C_{\mathcal{SU}}$ are exactly those with $p[c,1] \neq p[c,0]$. So, $\sum_{c \in C} q_c \mathbb{1}_{a_i=p[c,1]\neq p[c,0]} = \sum_{c \notin C_{\mathcal{SU}}} q_c \mathbb{1}_{a_i=p[c,1]}$. In Step 2, we have set $q_{c_i} = \rho_C(a_i) - \sum_{c \notin C_{\mathcal{SU}}} q_c \mathbb{1}_{a_i=p[c,1]}$ for any $c_i \in C_{\mathcal{SU}}$, $i < m$. Combining this with (2), we obtain $q_{c_i} = \rho_\emptyset(a_i) - \sum_{c \in C} q_c \mathbb{1}_{a_i=p[c,0]\neq p[c,1]}$ which can be rearranged to

$$\rho_\emptyset(a_i) = q_{c_i} + \sum_{c \in C} q_c \mathbb{1}_{a_i=p[c,0]\neq p[c,1]}. \tag{3}$$

By Claim 2, we can write $\rho_T(a_i)$ for any $T \in 2^C$ as $\rho_T(a_i) = \rho_\emptyset(a_i) + \sum_{c \in T} \delta_c(a_i)$. Combining with (3),

$$\rho_T(a_i) = q_{c_i} + \sum_{c \in C} q_c \mathbb{1}_{a_i=p[c,0]\neq p[c,1]} + \sum_{c \in T} q_c \mathbb{1}_{a_i=p[c,1]\neq p[c,0]} - \sum_{c \in T} q_c \mathbb{1}_{a_i=p[c,0]\neq p[c,1]}. \tag{4}$$

Notice that $\sum_{c \in C} q_c \mathbb{1}_{a_i=p[c,0]\neq p[c,1]} - \sum_{c \in T} q_c \mathbb{1}_{a_i=p[c,0]\neq p[c,1]} = \sum_{c \in C \setminus T} q_c \mathbb{1}_{a_i=p[c,0]\neq p[c,1]}$, so we can simplify (4) to obtain

$$\begin{aligned}\rho_T(a_i) &= q_{c_i} + \sum_{c \in C \setminus T} q_c \mathbb{1}_{a_i=p[c,0]\neq p[c,1]} + \sum_{c \in T} q_c \mathbb{1}_{a_i=p[c,1]\neq f[c,0]} \\ &= q_{c_i} + \sum_{c \in C} q_c \mathbb{1}_{a_i=p[c,k]\neq p[c,1-k]} \\ &= \sum_{c \in C} q_c \mathbb{1}_{a_i=p[c,\mathbb{1}_{c\in T}]}\end{aligned}$$

which is the desired result. The case of $c_j \in C_{\mathcal{SU}}$, $j \geq m$ can be derived analogously by letting $\rho_C(a_j) = (l-m+1)q_{c_j} + \sum_{c \notin C_{\mathcal{SU}}} q_c \mathbb{1}_{a_j=p[c,1]}$ and rewriting (2), (3) and (4) above, which gives

$$\begin{aligned}\rho_T(a_j) &= (l-m+1)q_{c_j} + \sum_{c\in C\setminus T} q_c \mathbb{1}_{a_j=p[c,0]\neq p[c,1]} + \sum_{c\in T} q_c \mathbb{1}_{a_j=p[c,1]\neq p[c,0]} \\ &= (l-m+1)q_{c_j} + \sum_{c\in C} q_c \mathbb{1}_{a_j=p[c,k]\neq p[c,1-k]}.\end{aligned}$$

Since $(l-m+1)q_{c_j} = \sum_{c_j\in C_{\mathcal{SU}}:j\geq m} q_{c_j}$, it follows that $\rho_T(a_j) = \sum_{c\in C} q_c \mathbb{1}_{a_j=p[c,\mathbb{1}_{c\in T}]}$ for all $T \in 2^C$, proving the claim. □

This concludes step 3 and the proof. □

References

Agranov, M. and P. Ortoleva (2017): "Stochastic choice and preferences for randomization," *Journal of Political Economy*, 125, 40–68.

Aizerman, M. and A. Malishevski (1981): "General theory of best variants choice: Some aspects," *IEEE Transactions on Automatic Control*, 26, 1030–1040.

Akerlof, G. A. and R. E. Kranton (2000): "Economics and identity," *The Quarterly Journal of Economics*, 115, 715–753.

——— (2002): "Identity and schooling: Some lessons for the economics of education," *Journal of Economic Literature*, 40, 1167–1201.

——— (2005): "Identity and the Economics of Organizations," *Journal of Economic Perspectives*, 19, 9–32.

Apesteguia, J., M. A. Ballester, and J. Lu (2017): "Single-Crossing Random Utility Models," *Econometrica*, 85, 661–674.

Atkin, D., E. Colson-Sihra, and M. Shayo (2019): "How do we choose our identity? A revealed preference approach using food consumption," *Journal of Political Economy, forthcoming.*

Bénabou, R. and J. Tirole (2011): "Identity, morals, and taboos: Beliefs as assets," *The Quarterly Journal of Economics*, 126, 805–855.

Benjamin, D. J., J. J. Choi, and G. Fisher (2016): "Religious identity and economic behavior," *Review of Economics and Statistics*, 98, 617–637.

Benjamin, D. J., J. J. Choi, and A. J. Strickland (2010): "Social identity and preferences," *American Economic Review*, 100, 1913–28.

Bertrand, M., E. Kamenica, and J. Pan (2015): "Gender identity and relative income within households," *The Quarterly Journal of Economics*, 130, 571–614.

Calvó-Armengol, A. and M. O. Jackson (2009): "Like father, like son: social network externalities and parent-child correlation in behavior," *American Economic Journal: Microeconomics*, 1, 124–50.

Cattaneo, M. D., X. Ma, Y. Masatlioglu, and E. Suleymanov (2020): "A random attention model," *Journal of Political Economy*, 128, 2796–2836.

Cerreia-Vioglio, S., D. Dillenberger, P. Ortoleva, and G. Riella (2019): "Deliberately stochastic," *American Economic Review*, 109, 2425–45.

Chambers, C. P., T. Cuhadaroglu, and Y. Masatlioglu (2020): "Behavioral Influence," *Working paper*.

Chambers, C. P. and F. Echenique (2016): *Revealed preference theory*, vol. 56, Cambridge University Press.

Charness, G., L. Rigotti, and A. Rustichini (2007): "Individual behavior and group membership," *American Economic Review*, 97, 1340–1352.

Chen, Y., S. X. Li, T. X. Liu, and M. Shih (2014): "Which hat to wear? Impact of natural identities on coordination and cooperation," *Games and Economic Behavior*, 84, 58–86.

Cohn, A., M. A. Maréchal, and T. Noll (2015): "Bad boys: How criminal identity salience affects rule violation," *The Review of Economic Studies*, 82, 1289–1308.

Constant, A. F. and K. F. Zimmermann (2008): "Measuring ethnic identity and its impact on economic behavior," *Journal of the European Economic Association*, 6, 424–433.

Costa-Font, J. and F. Cowell (2015): "Social identity and redistributive preferences: a survey," *Journal of Economic Surveys*, 29, 357–374.

Eckel, C. C. and P. J. Grossman (2005): "Managing diversity by creating team identity," *Journal of Economic Behavior & Organization*, 58, 371–392.

Fang, H. and G. C. Loury (2005): ""Dysfunctional Identities" Can Be Rational," *American Economic Review*, 95, 104–111.

Golub, B. and M. O. Jackson (2012): "How homophily affects the speed of learning and best-response dynamics," *The Quarterly Journal of Economics*, 127, 1287–1338.

Gul, F. and W. Pesendorfer (2006): "Random expected utility," *Econometrica*, 74, 121–146.

Jackson, M. O. (2011): "An overview of social networks and economic applications," *Handbook of social economics*, 1, 511–585.

KRANTON, R., M. PEASE, S. SANDERS, AND S. HUETTEL (2013): "Identity, Groups, and Social Preferences," *Working paper*.

——— (2020): "Deconstructing bias in social preferences reveals groupy and not-groupy behavior," *Proceedings of the National Academy of Sciences*, 117, 21185–21193.

KRANTON, R. E. (2016): "Identity economics 2016: Where do social distinctions and norms come from?" *American Economic Review*, 106, 405–09.

MANZINI, P. AND M. MARIOTTI (2014): "Stochastic choice and consideration sets," *Econometrica*, 82, 1153–1176.

MASATLIOGLU, Y., D. NAKAJIMA, AND E. Y. OZBAY (2012): "Revealed attention," *American Economic Review*, 102, 2183–2205.

MOULIN, H. (1985): "Choice functions over a finite set: a summary," *Social Choice and Welfare*, 2, 147–160.

OK, E. A., P. ORTOLEVA, AND G. RIELLA (2015): "Revealed (p) reference theory," *American Economic Review*, 105, 299–321.

SANDBERG, A. (2018): "Competing identities: a field study of in-group bias among professional evaluators," *The Economic Journal*, 128, 2131–2159.

TAJFEL, H., J. C. TURNER, W. G. AUSTIN, AND S. WORCHEL (1979): "An integrative theory of intergroup conflict," *Organizational identity: A reader*, 56, 65.

TAJFEL, H. E. (1978): *Differentiation between social groups: Studies in the social psychology of intergroup relations.*, Academic Press.

WICHARDT, P. C. (2008): "Identity and why we cooperate with those we do," *Journal of Economic Psychology*, 29, 127–139.