Norms and Emotions*

Hiroki Saruya† and Masayuki Yagasaki‡

November 19, 2025

Abstract

Social norms are an important determinant of behavior, but the behavioral and welfare effects of norms are not well understood. We propose and axiomatize a decision-theoretic model in which a reference point is formed by the decision maker's perceptions of which actions are admired (prescriptive norms) and which are prevalent (descriptive norms), and utility depends on the *pride* of exceeding the reference point or the *shame* of falling below it. The model is simple, yet provides a unified explanation for previous empirical findings, and is useful for behavioral and welfare analysis of norm-evoking policies with a revealed preference approach.

Keywords: norms, reference dependence, pride, shame, public recognition, norm nudge

JEL classification: D80, D81, D90, D91

*This paper substantially extends Yagasaki's previously circulated paper "Pride, Shame and Social Comparisons." We thank Leonardo Bursztyn, Takaaki Hamada, Yosuke Hashidate, Hidehiko Ichimura, Ryota Iijima, Michihiro Kandori, Akihiko Matsui, Ran Spiegler, Norio Takeoka, and seminar participants at the University of Tokyo and Osaka University for their invaluable feedback. Any remaining errors are the sole responsibility of the authors.

†Economic and Social Research Institute, Cabinet Office, Government of Japan and Tohoku University, `hiroki.saruya@aya.yale.edu`

‡Tohoku University, `masayuki.yagasaki.c4@tohoku.ac.jp`

1 Introduction

Social norms are receiving increasing attention as a key determinant of behavior in various contexts. Norms can take effect through simple interventions such as making decisions or outcomes publicly observable[1] or providing social information.[2] As a result, policymakers have become increasingly interested in social norms as a cost-effective policy lever to induce desired behavioral changes.

Despite the growing interest, the behavioral and welfare effects of such policies are not well understood. Norm-evoking policies may produce desired behavioral outcomes in some cases, but they may fail to do so or even backfire in others.[3] The lack of understanding is partly due to the lack of theoretical foundations on how norms affect the decision maker's payoffs and how they are revealed from choice data. This gap also makes it unclear how revealed preferences are useful for welfare analysis in norm-conscious decision-making.

This paper presents a novel decision-theoretic model to describe the behavior of a decision maker (DM) who is concerned with social norms. We consider a two-stage choice problem (Gul and Pesendorfer 2001; Noor and Takeoka 2015) adapted to decisions under social image concerns (e.g. Dillenberger and Sadowski 2012; Saito 2015; Evren and Minardi 2017; Hashidate 2021). The DM first *privately* chooses a menu (i.e., choice set) and then *publicly* chooses an alternative from the menu. This setting naturally expresses the behavioral effect of norms by the discrepancy between preferences in the private (norm-free) and public (norm-conscious) stages, and is also suitable to study the avoidance of choice opportunities (e.g., Dana et al. 2006) or the welfare effects of norms. We axiomatize a utility representation called a *pride-shame representation*, in which utility depends on an endogenously derived reference point (cf. Ok et al. 2015; Lleras et al. 2019; Kıbrıs et al. 2023).

In our model, the reference point is determined by an interaction of two types of *subjective* norms, referred to as *descriptive norms* and *prescriptive norms*. Economists typically emphasize descriptive norms, which express the DM's perception of what behavior is prevalent or common, i.e., what others *choose* to do. In contrast, social psychologists also emphasize prescriptive norms (e.g., Cialdini et al. 1991; Bicchieri 2005; Bicchieri and Dimant 2022),

[1] Researchers have studied the effects of publicity on educational investment (Bursztyn and Jensen 2015), career choice (Bursztyn et al. 2017), tax compliance (Perez-Truglia and Troiano 2018), charitable giving (Butera et al. 2022, and see also DellaVigna et al. 2012), blood donations (Lacetera and Macis 2010), and voting (Gerber et al. 2008). See Bursztyn and Jensen (2017) for a review.

[2] Information about other individuals' behavior or normative opinions affects charitable donation (Frey and Meier 2004), tax compliance (Frey and Torgler 2007; Hallsworth et al. 2017), energy conservation (Schultz et al. 2007; Allcott 2011; Allcott and Rogers 2014), and female labor participation (Bursztyn et al. 2020).

[3] Publicity of decisions may increase or decrease target behavior (Bursztyn and Jensen 2015). Providing information about the behavior of others may lead to the avoidance of a choice opportunity (Klinowski 2021) or an undesirable choice (Schultz et al. 2007).

which express the DM's perception of what behavior is approved of or admired, i.e., what others think one *should* do.[4] Although economists have also studied prescriptive norms (e.g., Akerlof and Kranton 2000), they have not extensively studied how the two notions of norms interact. We show, by an application to prosocial behavior, that interactions between these norms can explain a variety of previously documented behavioral patterns. Crucially, the two types of norms are the subjective beliefs of the DM and are allowed to be biased.

An essential determinant of behavior is social emotions, such as pride and shame, which arise from comparing one's own behavior with the typical behavior of others as reference behavior. To illustrate, consider a DM who expects a donation solicitor to arrive at her home shortly (DellaVigna et al. 2012). The DM's satisfaction with donating an amount, say \$10, depends on how she perceives the behavior of others. If she believes that her neighbors donate \$0, then she gains a positive sense of pride from the \$10 donation because her behavior is perceived as normatively superior to that of her neighbors. The degree of pride depends on the perceived desirability of each action: if donating \$10 is considered much (barely) more desirable than donating \$0, then the payoff gain from pride is large (small). In contrast, if she believes that her neighbors donate \$100, she suffers a negative sense of shame from donating \$10 because her behavior is considered normatively inferior. The payoff loss from shame, in turn, depends on the perceived admirability of each action. As this example shows, descriptive norms determine which behavior the DM focuses on as reference behavior (donating \$0 or \$100) to which to compare her own choice (donating \$10), and prescriptive norms determine the payoff from the comparison. The norms then affect the DM's behavior. Suppose she initially plans to donate \$10, but then thinks that her neighbors are donating \$100. If a solicitor is already at her door, she may increase her planned donation to avoid shame. Alternatively, if the solicitor has not yet arrived, she may leave the house, thereby avoiding the opportunity to donate.

Using a simple example of prosocial behavior, we illustrate that our model provides useful insights for the behavioral and welfare effects of norms. First, it clarifies how the choice of an action depends on descriptive and prescriptive norms, and when policies such as providing social information or making decisions public may be (in)effective. For example, if information about others' behavior (normative opinions) mainly affects the descriptive (prescriptive) norm, then changing this perceived norm is the main mechanism behind the effect of providing information. The effectiveness of the policy then depends on how sensitive the perceived norms are to the policy and how the DM evaluates the resulting pride or shame.

Second, the two-stage modeling allows us to study choice avoidance and the welfare effects of policies directly. For example, if a DM strictly prefers a menu {\$0} over another menu

[4]Prescriptive norms are also known as injunctive norms (see Section 5 for terminology).

$\{\$0, \$10\}$, this indicates her avoidance of an opportunity to donate \$10, and a negative welfare effect of making the choice public. Without using the preference over the menus, we might draw a false welfare conclusion: e.g., if we only observe that the DM chooses a \$10 donation from the menu $\{\$0, \$10\}$, we might mistakenly infer that adding the option to donate \$10 is beneficial, even if she chooses it simply to avoid the shame from not donating.[5] In addition, our model illustrates how policies to influence perceived norms exert differential impacts on the participation in a donation opportunity and on the donation decision conditional on participation. For example, it can explain the laboratory findings of Klinowski (2021) that informing individuals about others' high level of donation *after* participation increases the amount donated, but doing so *before* participation discourages participation.

Third, our model can account for behavioral regularities that are well documented in psychology but have received limited attention in economics. For example, it can rationalize previous findings that providing information about descriptive or prescriptive norms is more effective at inducing prosocial behavior when they are aligned than when they are misaligned (Cialdini 2003), and that the descriptive norm has a greater influence in the latter case (e.g., Bicchieri and Xiao 2009). An individual is more likely to make a donation when others say that one should donate *and* they do donate, than when others say one should donate *but* they do not. Intuitively, when both norms point to prosocial behavior, failure to follow them generates shame. In contrast, if the prescriptive norm points to prosocial behavior but the descriptive norm points to the opposite, acting prosocially generates pride. If avoiding shame is a stronger motivator than seeking pride, which is empirically supported (DellaVigna et al. 2017; Butera et al. 2022), then aligned norms are more likely to induce prosocial behavior.

The first step toward axiomatically deriving our representation is to characterize the DM's subjective reference. Our approach is similar in spirit to that of Masatlioglu et al. (2012) and Kıbrıs et al. (2023), who elicit the DM's consideration and reference, respectively, by observing a "choice reversal," whereby removing an unchosen alternative from a menu affects the choice from the menu. Instead of requiring a choice reversal, we exploit observations such that removing an unchosen alternative affects the preference over menus. Suppose we observe that the DM donates \$10 whether or not she has the option to decline donation ($\mathcal{C}(\{\$0, \$10\}) = \mathcal{C}(\{\$10\}) = \{\$10\}$), but that she strictly prefers to donate with the option to decline ($\{\$0, \$10\} \succ \{\$10\}$). This suggests that the option to decline donation improves the DM's utility from donating by generating pride, which then implies that \$0 is the reference choice at the menu $\{\$0, \$10\}$. We generalize this observation to elicit a subjective reference set, i.e., the set of reference alternatives, at each menu.

The second key step is to describe how preferences for smaller or larger menus emerge

[5] This situation is similar to the situation of "product market traps" (Bursztyn et al. 2025).

depending on the reference set. Consider first a DM who perceives that her neighbors do not donate. Then, answering the door to meet a solicitor will never hurt ($\{\$0, \$10\} \succeq \{\$0\}$), because she can decline the donation without shame, or she can even feel pride by choosing to donate. Lemma 2(i), derived from our axioms, formalizes this idea: the DM will exhibit a preference for larger menus (cf. Evren and Minardi 2017) when the additional alternatives do not enter the reference set. Next, consider a DM who privately does not want to donate ($\{\$0\} \succeq \{\$10\}$). Suppose she notices that some neighbors are donating \$10, so \$10 enters her reference set when she faces the menu $\{\$0, \$10\}$. Then, the option to donate \$10 will not improve the DM's feelings about not donating, because of the shame of falling below her neighbors' standard. She then prefers to avoid the donation option ($\{\$0\} \succeq \{\$0, \$10\}$). Lemma 2(ii) characterizes such a preference: the DM will exhibit a preference for smaller menus (cf. Gul and Pesendorfer 2001; Dillenberger and Sadowski 2012) when the extra alternatives enter the reference set.

Our contribution is to propose a simple, tractable, and theoretically and axiomatically founded model of norm-conscious decisions that facilitates applied analysis. (1) Our model of norms using a descriptive norm and a prescriptive norm is simple, yet it can explain various previous empirical findings. It also clarifies mechanisms behind policy effects, facilitating policy analysis. (2) The model is tractable in that it does not require an equilibrium assumption; instead, it is directly disciplined by observed choices. Thus, the DM's perceived norms are allowed to be biased and are revealed from choice data. (3) The model is closely aligned to the social psychological theory of norms and it also has an axiomatic foundation. The transparent link between choice data and utility representation facilitates a revealed preference approach to studying behavioral and welfare effects of norms and norm-evoking policies, e.g., to infer who feel pride and who feel shame in a given situation (cf. Toussaert 2018), and welfare impacts of public recognition programs (Butera et al. 2022).

The paper is organized as follows. In Section 2, we illustrate our model and its implications by a simple example of prosocial behavior. Section 3 presents our axioms and the representation result. Section 4 discusses how our model can be useful for empirical research. Section 5 reviews the literature. Section 6 concludes. Proofs and additional results are presented in the Appendices.

2 Illustrative Model

Denote a typical menu of alternatives by A. With a simplified version of our utility representation, the preference $\succeq$ over menus is represented by

$$V_{PS}(A) = \max_{x \in A} U(x; A), \tag{1}$$

and the ex-post choice from each menu coincides with $\mathcal{C}_{PS}(A) = \arg\max_{x \in A} U(x; A)$, where $U(x; A)$, the utility of choosing alternative x from A, is expressed as

$$U(x; A) = \underbrace{u(x)}_{\text{intrinsic}} - \underbrace{\max\{w(\varphi_r(A)) - w(x), 0\}}_{\text{“shame”} \geq 0} + \beta \underbrace{\max\{w(x) - w(\varphi_r(A)), 0\}}_{\text{“pride”} \geq 0}. \tag{2}$$

In Eq. (2), $\beta > 0$, $\varphi_r(A) = \arg\max_{a \in A} r(a)$, which is assumed to be a singleton for illustrative purposes, and u, w, and r are expected utility (EU) functions.[6] The function u represents the DM's intrinsic utility function, which describes her private preference ranking.[7] The term $w(\varphi_r(A))$ represents a social reference point, which consists of two distinct components. First, the function r is called the *descriptive norm function*, which expresses the DM's perception of the prevalence of each alternative. $\varphi_r(A)$ is then interpreted as the alternative that the DM thinks is typically chosen by other people in her society. Second, the function w is called the *prescriptive norm function*, which expresses the DM's perception of the admirability of each alternative. Together, $w(\varphi_r(A))$ represents the normative desirability of the socially prevalent choice, as perceived by the DM. We allow the DM to have biased beliefs about others' behavior or normative opinions.

The last two terms in Eq. (2) denote the utility from social emotions. If the DM chooses an alternative x that is normatively inferior to the reference alternative $\varphi_r(A)$, she feels *shame*, which reduces her utility by $w(\varphi_r(A)) - w(x)$. Conversely, if she chooses x that is normatively superior to $\varphi_r(A)$, she feels *pride*, which increases her utility by $\beta\,[w(x) - w(\varphi_r(A))]$. This modeling is closely aligned to the social psychological theory of norms (see Section 5). Also, by letting $\beta \neq 1$, we allow the DM to care about a downward

[6]The EU may not well suit some contexts of social decision-making under uncertainty (e.g., Saito 2013). However, it remains appropriate in contexts where the social consequences of actions involve probabilistic uncertainty and can be meaningfully aggregated through expected values. For example, offering aid to a politically corrupt and impoverished country may probabilistically result in either alleviating suffering among the poor or entrenching the authority of a corrupt regime. Such cases are well suited to expected utility analysis (Rabin 1995). While we acknowledge the limitations of the EU in capturing more complex forms of moral uncertainty, we adopt it for tractability and as a first step. Extending the framework to broader decision-making contexts is a promising direction for future research. Also, the applied analysis in Section 2.1 and Section 4 does not rely on the properties of the EU.

[7]u may capture not only her self-interest, but also other concerns such as altruism, warm glow, and moral concerns that are not influenced by social image concerns.

deviation from the reference point (shame) differently from an upward deviation (pride).[8]

2.1 A Simple Example of Prosocial Behavior

We illustrate the implications of our model by the following simple example. Let $x \in A = \{0,1\}$ denote the DM's choice of an alternative, where $x = 1$ indicates the DM engaging in prosocial behavior, and $x = 0$ indicates non-engagement. Let $u(0) = \bar{u} > 0 = u(1)$, $w(0) = 0 < \bar{w} = w(1)$, and $\beta\bar{w} < \bar{u} < \bar{w}$. Thus, the DM privately prefers the non-prosocial choice but believes that the prosocial choice is more admired. Also, $\beta < 1$ means that the DM is more sensitive to shame than she is to pride.

Benchmark behavior. The DM chooses $x = 0$ or $x = 1$ by comparing the utility from each alternative:

$$\begin{aligned} U(0;\{0,1\}) &= \underbrace{\bar{u}}_{\text{intrinsic}} - \underbrace{[w(\varphi_r(\{0,1\})) - 0]}_{\text{shame}} = \bar{u} - w(\varphi_r(\{0,1\})) \\ U(1;\{0,1\}) &= \underbrace{0}_{\text{intrinsic}} + \beta \underbrace{[\bar{w} - w(\varphi_r(\{0,1\}))]}_{\text{pride}} = \beta\,[\bar{w} - w(\varphi_r(\{0,1\}))] \end{aligned} \tag{3}$$

The expressions are simpler than Eq. (2) because $x = 0$ never causes pride and $x = 1$ never causes shame, regardless of the reference alternative $\varphi_r(\{0,1\})$.

As a benchmark, suppose $r(1) < r(0)$, i.e., the DM believes that other people in her society do not typically engage in prosocial behavior. The reference alternative is $\varphi_r(\{0,1\}) = 0$ and the reference point is $w(0) = 0$. Choosing $x = 0$ gives the DM the intrinsic utility $\bar{u}$ and no utility from social emotion, because she chooses the action dictated by the norm. On the other hand, choosing $x = 1$ gives the DM zero intrinsic utility but gives a positive utility from pride. Since $\beta\bar{w} < \bar{u}$, the DM chooses $x = 0$.

Perceived norms and behavior. The model predicts how the DM's behavior depends on the descriptive and prescriptive norms. Consider the following analysis, where each type of norms shifts toward prosocial behavior relative to the above benchmark.

(i) *Higher descriptive norm.* Suppose the descriptive norm function becomes r' such that $r'(0) < r'(1)$, shifting the reference point to $w(\varphi_{r'}(\{0,1\})) = \bar{w}$. Now, choosing $x = 0$ gives the DM utility $\bar{u} - \bar{w} < 0$, whereas choosing $x = 1$ yields zero utility. Thus, the DM chooses $x = 1$.

[8] We emphasize the case with $\beta \in (0,1)$, which expresses shame aversion (cf. Butera et al. 2022), though our theory allows for $\beta \geq 1$. Also, it accommodates $\beta = 0$ as a limit case.

(ii) *Higher prescriptive norm.* Suppose that the prescriptive norm function becomes w' such that $w'(0) = 0$ and $w'(1) = \bar{w}' > \frac{\bar{u}}{\beta}$. Then, choosing $x = 1$ provides a pride benefit of $\beta\bar{w}'$, which exceeds the utility $\bar{u}$ from $x = 0$. Thus, the DM chooses $x = 1$.

The DM switches to prosocial behavior $x = 1$ in both cases, but for different reasons. In case (i), she chooses $x = 1$ because she would feel shame if she stuck to the less admirable choice $x = 0$ while perceiving that others choose $x = 1$. In contrast, in case (ii), she chooses $x = 1$ because she feels greater pride from $x = 1$ perceiving that others choose $x = 0$.

This analysis is insightful for analyzing the effect of "norm nudges," which guide people's decisions by providing social information. Many economic studies have explored the effect on decisions of information about how others behave (e.g., Frey and Meier 2004; Allcott 2011) or what others think is the appropriate behavior (e,g, Hallsworth et al. 2017; Bursztyn et al. 2020). Our model helps clarify the mechanisms underlying such a norm-nudging. For example, if information about others' behavior (normative opinions) mainly affects an individual's perceived descriptive (prescriptive) norm,[9] then the main mechanism of the effect of such information is to alter r (w) in favor of prosocial behavior, thereby generating shame of non-prosociality (increasing pride of prosociality). Thus, the relative effectiveness of each type of information depends on the DM's sensitivity to each social emotion, expressed by β (more on this below). Of course, effectiveness also depends on the quality of the information.

Public recognition and prosociality. The model illustrates how public observability affects the DM's prosociality. Her choice of action under a private decision environment is expressed as a choice between two menus $\{0\}$ and $\{1\}$, with the utility from each option $V_{PS}(\{x\}) = u(x)$ (note the absence of social emotions). By contrast, her choice under a public environment is expressed as a choice between two actions 0 and 1 from the menu $\{0, 1\}$, with the utility from each option $U(x; \{0, 1\})$ in Eq. (3).[10] Because $U(x; \{0, 1\})$ is strictly increasing in $w(x)$, the DM becomes more prosocial in the public environment than in the private environment. The above analysis also suggests when policies such as public recognition programs are ineffective for inducing prosocial behavior: they are ineffective when the descriptive and prescriptive norms do not sufficiently favor prosocial behavior, or when the DM is relatively insensitive to pride.

Perceived norms and avoidance. The two-stage model enables us to study how norms

[9]In reality, information about one norm type may also affect the perception of the other.

[10]The private and public preferences represent norm-free and norm-conscious preferences, respectively. Thus, our model applies to more general settings where some environmental cue (including publicity as an example) triggers the DM to focus on norms.

affect the DM's decision to participate in the opportunity for prosocial behavior, as well as her decision on prosocial behavior itself. Analysis of the participation decision is important for two reasons. First, laboratory and field experiments have documented that a large fraction of individuals avoid opportunities to engage in prosocial behavior, even if they can choose non-engagement after participation and even if avoidance is costly (e.g., Dana et al. 2006; Broberg et al. 2007; Lazear et al. 2012; DellaVigna et al. 2012; Andreoni et al. 2017; Klinowski 2021). Our model clarifies how such avoidance depends on the perceived descriptive and prescriptive norms. Second, the participation decision is informative of the DM's willingness-to-pay (WTP) for publicity and can be used to study the welfare impacts of policies such as public recognition programs, assuming that pride and shame are welfare-relevant. For example, her valuation of the menu $\{0, 1\}$ relative to that of the singleton menu $\{0\}$ is informative of her WTP for public recognition.[11]

Our model illustrates how perceived norms can differentially impact participation and choice of action. Suppose that the DM first chooses whether to participate in the opportunity for prosocial behavior. If she decides to participate, she proceeds to the binary-choice stage described above. Alternatively, she can decide not to participate and be given a singleton menu $\{0\}$, which gives her utility $V_{PS}(\{0\}) = \bar{u}$. In the benchmark case, participation gives utility $V_{PS}(\{0, 1\}) = \max\{U(0; \{0, 1\}), U(1; \{0, 1\})\} = \bar{u}$, and it is indifferent to non-participation. Therefore, the DM can optimally participate in the opportunity and then choose not to engage in prosocial behavior.

Now, suppose that the descriptive norm shifts toward prosocial behavior (i.e., r changes to r'). A possible interpretation is that the DM updates her perception of the norm after she is given information about others' actions. By the above analysis, the DM switches to prosocial behavior *conditional on participation*. On the other hand, with the descriptive norm r', we have $V_{PS}(\{0, 1\}) = \max\{\bar{u} - \bar{w}, 0\} = 0 < V_{PS}(\{0\})$, so the DM avoids the opportunity for prosocial behavior. Thus, the more prosocial descriptive norm induces the DM to take a prosocial action if she has no option to avoid the choice occasion, but it induces her to avoid the occasion if possible.

The theoretical predictions match empirical evidence quite well. In a laboratory experiment, Klinowski (2021) demonstrates that (1) when individuals receive information that others have made a large donation after they participate in the opportunity, they increase the amount of donation relative to the no-information benchmark, whereas (2) when they receive the same information prior to the decision to participate, the participation rate drops

[11] Butera et al. (2022) use an incentive-compatible mechanism to elicit individuals' WTP for public recognition, in the context of charitable behavior. Our framework infers the WTP using preferences over menus, which might be useful when WTP-elicitation surveys are unavailable (e.g., naturally occurring data).

relative to the benchmark. Our model can rationalize these findings by the shift of the descriptive norm caused by the information treatment.

Aligned vs. misaligned norms. The simple model also explains why the descriptive and prescriptive norms induce larger behavioral changes when they are aligned than when they are misaligned (Cialdini 2003), and why the descriptive norm tends to trump the prescriptive norm in the latter case (Tyran and Feld 2006; Bicchieri and Xiao 2009).[12] When the prescriptive norm dictates prosocial behavior ($w(0) < w(1)$) but the descriptive norm dictates otherwise ($r(0) > r(1)$),[13] the DM behaves prosocially if the pride benefit $\beta\bar{w}$ outweighs the intrinsic benefit of non-engagement, $\bar{u}$. By contrast, when both norms point to the prosocial behavior ($w(0) < w(1)$ and $r'(0) < r'(1)$), she behaves prosocially if the shame cost $\bar{w}$ of non-engagement outweighs its intrinsic benefit. If the DM is shame-averse ($\beta < 1$), which finds some empirical support,[14] aligned norms induce prosocial behavior more effectively than misaligned ones. In words, when both norms point to prosocial behavior, failing to follow them causes shame for falling below social expectations, which is a strong motivator for prosocial behavior. In contrast, when the prescriptive norms point to prosocial behavior but the descriptive norms point to the opposite, the social motivation for prosocial behavior is pride from exceeding social expectations, which may not be a strong motivator. Thus, the DM will not feel pressure to respect the prescriptive norms and she will behave non-prosocially, as the descriptive norms dictate.

Other results. In Supplemental Appendix S.B, we show that the model can explain other previous empirical findings, e.g., why providing information on others' prosocial behavior can *reduce* the amount of prosocial behavior (e.g., Schultz et al. 2007).

3 Model

We adopt the framework of Gul and Pesendorfer (2001) (henceforth GP). Let (Z, ρ) be a compact metric space, where Z is a finite set of prizes, and let $\Delta \equiv \Delta(Z)$ denote the

[12] Allcott (2011) and Hallsworth et al. (2017) find evidence that descriptive norm messages are more effective than prescriptive norm messages for inducing electricity saving and tax compliance, respectively. Heinicke et al. (2022) find that descriptive norms exhibit a stronger correlation with behavior than prescriptive norms in the context of mini-dictator games.

[13] The opposite misalignment ($\tilde{w}(0) > \tilde{w}(1)$ and $r'(0) < r'(1)$) is irrelevant; the DM chooses $x = 0$ because it is better both intrinsically and socially than $x = 1$, making $r'(0) < r'(1)$ implausible and irrelevant. Put differently, a relevant case should have conflicting u and w, and r may be aligned to either.

[14] Butera et al. (2022) find evidence for shame aversion in prosocial behavior. DellaVigna et al. (2017) find that non-voters in an election sort out of a survey due to the negative feeling from admitting non-voting or lying about it, while voters do not sort in to enjoy the positive feeling from saying that they voted.

set of probability measures on the Borel σ-algebra of Z endowed with the weak topology. Denote by $\mathcal{A}$ a set of all closed subsets of Δ, and endow $\mathcal{A}$ with the topology generated by the Hausdorff metric.[15] A typical lottery $a \in \Delta$ is called an alternative (or choice), and a typical element $A \in \mathcal{A}$, a set of alternatives, is called a menu (or choice set). Define $\alpha A + (1-\alpha)B \equiv \{z \in \Delta : z = \alpha a + (1-\alpha)b, a \in A, b \in B\}$ for $A, B \in \mathcal{A}$ and $\alpha \in [0,1]$.

We consider a DM who has a preference $\succeq$ over menus and also makes a choice from a menu by a choice rule $\mathcal{C}$. Specifically, $\succeq$ is a binary relation over $\mathcal{A}$, and $\mathcal{C} : \mathcal{A} \twoheadrightarrow \Delta$ satisfies $\emptyset \neq \mathcal{C}(A) \subseteq A$ for all $A \in \mathcal{A}$. We assume that both $\succeq$ and $\mathcal{C}$ are observed.

We consider a DM whose choice from a menu depends on a reference point that consists of her subjective beliefs. Specifically, she references the alternative in the menu that she believes is most commonly chosen by others. If multiple alternatives are perceived to be most common, she references the one that she believes is the most admirable. She then derives a positive (negative) emotion from her choice if she believes that it is more (less) admirable than the reference alternative. The beliefs on prevalence/commonality are expressed as descriptive norms, and the beliefs on admirability are expressed as prescriptive norms. The preferences over menus reflect the anticipated payoffs from such emotions, although the choices between menus per se do not generate such emotions. We capture this situation by assuming a two-stage process, where the DM *privately* chooses a menu in the first stage and then *publicly* chooses an alternative from the menu in the second stage.

3.1 Axioms

We first introduce some basic axioms.

Axiom 1. (Order) $\succeq$ *is complete and transitive.*

Axiom 2.

(i) (Lower Semi-Continuity) *For any* $A \in \mathcal{A}$, $\{B \in \mathcal{A} : A \succeq B\}$ *is closed.*

(ii) (Upper von Neumann-Morgenstern Continuity) $A \succ B \succ C$ *implies* $B \succ \alpha A + (1-\alpha) C$ *for some* $\alpha \in (0,1)$.

(iii) (Upper Singleton Continuity) $\{\{b\} \in \mathcal{A} : \{b\} \succeq \{a\}\}$ *is closed.*

Axiom 1 is standard. Axioms 2(i)-(iii), similar to axioms in GP to characterize preferences without self-control, weaken standard continuity. They yield a reference point that is of a "Strotz representation" (Strotz 1955), which may change discontinuously. Such a specification seems attractive given that social preferences often feature discontinuities.[16]

[15] That is, $d_H(A,B) = \max\left\{\max_{a \in A} \min_{b \in B} d(a,b), \max_{b \in B} \min_{a \in A} d(a,b)\right\}$, where d is a metric that metrizes the weak topology.

[16] E.g., an equal split in dictator games (cf. Andreoni and Bernheim 2009) may be a discontinuity point.

We proceed by introducing a "revealed descriptive (norm) ranking" $\succeq_r$, which elicits the DM's subjective beliefs about the prevalence of each alternative from observed behavior.

Definition 1. (Revealed Descriptive Ranking)

(i) $a \succ^* b$ *if there exists* $A \ni b$ *such that* $A \cup \{a\} \succ A$ *and* $a \notin \mathcal{C}(A \cup \{a\})$.

(ii) $a \succ_r b$ *if either of the following conditions holds:*

- **a.** $a \succ^* b$.
- **b.** *There exists some* $c \in \operatorname{int}(\Delta)$ *such that* $c \not\succ^* a$ *and* $c \succ^* b$.

(iii) $a \sim_r b$ *if neither* $a \succ_r b$ *nor* $b \succ_r a$*, and* $a \succeq_r b$ *if either* $a \succ_r b$ *or* $a \sim_r b$.

$a \succ^* b$ means that adding a to a menu $A \ni b$ makes the menu more attractive ($A \cup \{a\} \succ A$) though a is unchosen ($a \notin \mathcal{C}(A \cup \{a\})$). This suggests that a improves the menu A by lowering its reference point. We then infer that the DM references a at $A \cup \{a\}$, which we interpret as her perceiving a to be more prevalent than the other alternatives in A, including b.[17] Case (ii-b) addresses technical difficulties when a is a boundary element in Δ (see Supplemental Appendix S.D).

Next, we define a "revealed prescriptive (norm) ranking" $\succeq_w$, which partially elicits the DM's subjective beliefs about the normative desirability/admirability of each alternative. It is "partial" in that it elicits the true prescriptive norms w only among alternatives with the same descriptive ranking (see Theorem 2).

Definition 2. (Revealed Prescriptive Ranking)

(i) $a \succ_w b$ *if one of the following conditions holds.*

- **a.** $\{b\} \succ \{a, b\}$
- **b.** $\{b\} \sim \{a, b\}$ *and* $\mathcal{C}(\{a, b\}) = \{a\}$.
- **c.** $a \sim_r b$ *and* $\{a\} \sim \{a, b\} \succ \{b\}$.

(ii) $a \sim_w b$ *if neither* $a \succ_w b$ *nor* $b \succ_w a$. $a \succeq_w b$ *if either* $a \succ_w b$ *or* $a \sim_w b$.

The elicitation of the prescriptive ranking is similar to that of the temptation ranking (GP; Noor and Takeoka 2015). In case (i-a), adding a makes the menu $\{b\}$ less attractive. This suggests that a raises the reference point, which we interpret as the DM perceiving a to be more admirable than b. In case (i-b), if b sets the reference point at $\{a, b\}$, then the DM faces the same reference point as $\{b\}$ but does not choose b, suggesting that she is strictly

[17] We do not infer strict descriptive rankings from preferences for smaller menus ($A \succ A \cup \{a\}$). Theorem 2 shows that $\succeq_r$ alone fully elicits the true descriptive norms r if the data are generated by our representation, as long as $\succeq_r$ is not constant. We use preferences for smaller menus to infer the prescriptive rankings (see Definition 2), just as GP use them to infer the temptation rankings.

better off at $\{a,b\}$, which contradicts $\{b\} \sim \{a,b\}$. Thus, the reference point must be higher at $\{a,b\}$ than at $\{b\}$. Finally, in case (i-c), because both a and b enter the reference set at $\{a,b\}$, the reference point is weakly higher at $\{a,b\}$ than at $\{b\}$. Thus, $\{a,b\} \succ \{b\}$ implies that a is the unique choice at $\{a,b\}$. Then, $\{a\} \sim \{a,b\}$ implies that a must be weakly higher in the prescriptive ranking than b (otherwise, the DM prefers to exclude b, so $\{a\} \succ \{a,b\}$). Below, we focus on PS preferences such that the descriptive and prescriptive rankings are observationally distinct;[18] thus, in case (i-c), $a \sim_r b$ implies that a must be strictly higher than b in the prescriptive ranking.

$a \succ_w b$ also implies that a is at least as high as b in the descriptive ranking; otherwise, the DM does not reference a at $\{a,b\}$, so nothing about its normative desirability is revealed. However, we do not use $a \succ_w b$ to infer the descriptive ranking, because it does not tell us whether a is strictly higher than b or just as high as b in the descriptive ranking.[19]

A natural way for the DM to form her perception about socially prevalent actions is that she imagines a "typical person" and uses that person's behavior as a reference. Thus, we impose axioms to rationalize the descriptive ranking as an expected utility (EU) of someone. For simplicity, we directly impose axioms on $\succ^*$, $\succeq_r$, and $\succeq_w$, though we can rewrite them as properties of $(\succeq, \mathcal{C})$.

Axiom 3. (r-EU)

(i) *If $a \succ_r b$ or $a \succ_w b$, then neither $b \succ_r a$ nor $b \succ_w a$.*

(ii) *$\succ^*$ is transitive. Also, if $a \sim_r b \sim_r c$, $a \succeq_w b$, and $b \succeq_w c$, then $a \succeq_w c$.*

(iii) **a.** *$\{\alpha \in [0,1] : \alpha A + (1-\alpha)C \succeq B\}$ is closed in [0,1].*

b. *If there exists $a^* \in A$ such that $a^* \succ_r a$ for all $a \in A \setminus \{a^*\}$, then for any $\{A_n\}_n$ and $\{a_n\}_n$ such that $A_n \to A$, $a_n \in \mathcal{C}(A_n)$ and $a_n \to a$, we have $a \in \mathcal{C}(A)$.*

(iv) *For any $\alpha \in (0,1)$, $\alpha a + (1-\alpha)c \succ^* \alpha b + (1-\alpha)c$ and $a \in \text{int}(\Delta)$ imply $a \succ^* b$.*

Axiom 3(i) imposes consistency of descriptive rankings and prescriptive rankings revealed by different observations.[20] Recall $a \succ_r b$ reveals that a is perceived to be more prevalent than b. Also, $a \succ_w b$ reveals that a is perceived to be at least as prevalent as b and more admirable than b. Then, to consistently rank alternatives in prevalence and admirability, the choice data should not reveal the opposite relations. Axiom 3(ii) states that the directly revealed descriptive ranking $\succ^*$ is transitive and $\succeq_w$ is transitive on the indifference set of

[18]Distinguishing descriptive and prescriptive norms is pointless if they coincide. In fact, our PS model with identical descriptive and prescriptive norms is observationally equivalent to a PS model with no descriptive norms, i.e., constant r. See also Example 1.

[19]Moreover, the true descriptive norms are fully revealed by $\succeq_r$ alone, without using information of $\succeq_w$, for nondegenerate cases (see Theorem 2).

[20]Similar axioms appear in Dillenberger and Sadowski (2012) and Kıbrıs et al. (2023).

$\succeq_r$. Axiom 3(iii) expresses Archimedeanity of $\succeq_r$. Axiom 3(iv) imposes some linearity on the descriptive ranking. Axiom 3(iv) is only required to deal with boundary elements.

Next, we introduce a weak version of the linearity of $(\succeq, \mathcal{C})$. Due to the potential asymmetry between positive and negative social emotions, the standard independence axiom may be violated. For example, suppose $\$0 \succ_r \$100 \succ_r \$10$, i.e., the DM believes that donation is uncommon but a large donation is common conditional on donating.[21] Consider two menus $\{\$10, \$100\}$ and $\{\$100\}$, both of which have a \$100 donation as the reference alternative. If the shame from donating a small amount despite the social expectation of a large donation is strong, the DM will conform to the expectation and donate \$100 at both menus, so $\{\$10, \$100\} \sim \{\$100\}$. Now, consider two menus, $A = 0.5\{\$10, \$100\} + 0.5\{\$0, \$10\}$ and $B = 0.5\{\$100\} + 0.5\{\$0, \$10\}$, both having $0.5\$100 + 0.5\$0$ as the reference (recall r is linear). If the reference point is sufficiently low (due to the possibility of \$0), then the DM may choose \$10 from A, because it may strike a balance between self-interest and pride. In contrast, such an option is unavailable at B. Thus, we may have $A \succ B$. This phenomenon occurs because the relative attractiveness of two alternatives depends on the social emotions they generate: \$100 is preferred to \$10 if both generate (possibly zero) shame, whereas the converse is true if both generate (possibly zero) pride.

This discussion suggests that we should relax the linearity if mixing two menus alters the types of social emotions generated by each alternative. However, we may keep the linearity if the mixture preserves the types of social emotions.[22] To formalize the idea, let $L_r(a) = \{b \in \Delta : a \succ_r b\}$ denote the set of alternatives which is strictly below a in the descriptive ranking $\succeq_r$. For an arbitrary $a \in \Delta$, any $b \preceq_r a$ belongs to exactly one of the following sets:

$$\mathcal{P}(a) = \{b \in L_r(a) : \{a, b\} \succ \{b\} \text{ and } \mathcal{C}(\{a, b\}) = \{b\}\} \tag{4}$$

$$\mathcal{S}(a) = \{b \in L_r(a) : \{a, b\} \prec \{b\} \text{ and } \mathcal{C}(\{a, b\}) = \{b\}\} \tag{5}$$

$$\mathcal{N}_1(a) = \{b \in L_r(a) : \{a, b\} \sim \{b\} \text{ and } \mathcal{C}(\{a, b\}) = \{b\}\} \tag{6}$$

$$\mathcal{N}_2(a) = \{b \in L_r(a) : a \in \mathcal{C}(\{a, b\})\} \tag{7}$$

$$\mathcal{I}(a) = \{b \in \Delta : b \sim_r a\} \tag{8}$$

Eq.(4)-(7) partition the set of alternatives b that are below a in the descriptive ranking,

[21] Such a non-monotonic ranking can be generated by an EU function r such that the Bernoulli utility function is non-monotonic in the amount of money donated. The non-monotonicity can arise if the Bernoulli function is the sum of two functions, with one decreasing in the donated amount (e.g., representing selfish monetary payoffs) and the other increasing in it (e.g., representing altruism or image concerns).

[22] This assumption leads to a parsimonious model that generates menu effects by the asymmetry between pride and shame, which still provides useful insights, as illustrated in Section 2. A more general model would allow the linearity to hold only conditional on the reference point, which would complicate the analysis.

based on what social emotion the DM feels at the menu $\{a,b\}$. With $b \in \mathcal{P}(a)$, the DM feels pride by choosing b at $\{a,b\}$: Because the DM chooses b at both $\{a,b\}$ and $\{b\}$ but strictly prefers the former, we infer that the unchosen alternative a gives pride by lowering the reference point. Similarly, with $b \in \mathcal{S}(a)$, the DM feels shame by choosing b at $\{a,b\}$, because a sets a higher reference point than b. With $b \in \mathcal{N}_1(a)$, the DM feels no social emotion at $\{a,b\}$ because the chosen alternative b is socially as desirable as the reference alternative a. In these three cases, the DM would feel no social emotion if she deviated to choosing a. With $b \in \mathcal{N}_2(a)$, the DM feels no social emotion at $\{a,b\}$ because she chooses the reference alternative, although she may feel pride or shame by deviating to b. Finally, $\mathcal{I}(a)$ is the set of alternatives that are indifferent to a in the descriptive ranking. In this case, she never feels pride, because the reference point is set by the most admirable alternative.

Now, let $\mathcal{B}_P = \{\{a,b\} \in \mathcal{A} : a = b \text{ or } b \in \mathcal{P}(a) \cup \mathcal{N}_1(a)\}$ collect the binary or singleton menus where the DM never feels shame from any alternative.[23] Similarly, let $\mathcal{B}_S = \{\{a,b\} \in \mathcal{A} : a = b \text{ or } b \in \mathcal{S}(a) \cup \mathcal{N}_1(a) \cup \mathcal{I}(a)\}$ collect the binary or singleton menus where the DM never feels pride from any alternative. Finally, let $\mathcal{B}_N = \{\{a,b\} \in \mathcal{A} : a = b \text{ or } b \in \mathcal{N}_2(a)\}$ collect the singletons and binary menus where we cannot exclude any social emotion. Note we have $\cup_{j=P,S,N}\mathcal{B}_j = \{\{a,b\} : a,b \in \Delta\}$ and $\cap_{j=P,S,N}\mathcal{B}_j = \{\{a\} : a \in \Delta\}$. We now state our linearity axioms on $(\succeq, \mathcal{C})$.

Axiom 4. (Weak Independence) *For any $\alpha \in (0,1)$,*

- **(i)** *$A,B,C \in \mathcal{B}_P$ and $A \succ (\succeq) B$ imply $\alpha A + (1-\alpha)C \succ (\succeq) \alpha B + (1-\alpha)C$.*
- **(ii)** *$A,B,C \in \mathcal{B}_S$ and $A \succ (\succeq) B$ imply $\alpha A + (1-\alpha)C \succ (\succeq) \alpha B + (1-\alpha)C$.*
- **(iii)** *$A,B \in \mathcal{A}$, $c \in \Delta$ and $A \succ (\succeq) B$ imply $\alpha A + (1-\alpha)\{c\} \succ (\succeq) \alpha B + (1-\alpha)\{c\}$.*

Axiom 5. (Weak Linearity) *For any $a,b,c,d \in \Delta$ and $\alpha \in (0,1)$, the following properties hold.*

- **(i)** *Suppose $\{a,b\}, \{c,d\} \in \mathcal{B}_P$ or $\{a,b\}, \{c,d\} \in \mathcal{B}_S$. Then $\mathcal{C}(\alpha\{a,b\} + (1-\alpha)\{c,d\}) = \alpha\mathcal{C}(\{a,b\}) + (1-\alpha)\mathcal{C}(\{c,d\})$.*
- **(ii)** *Let $A = \alpha\{a,b\} + (1-\alpha)\{a,c\}$ and $b \in \mathcal{N}_2(a)$.*
  - **a.** *If $c \in \mathcal{P}(a)$, $\{a, \alpha b + (1-\alpha)c\} \succeq \alpha\{b\} + (1-\alpha)\{a,c\}$, and $\mathcal{C}(\{a, \alpha b + (1-\alpha)c\}) = \{\alpha b + (1-\alpha)c\}$, then $\mathcal{C}(A) = \alpha\mathcal{C}(\{a,b\}) + (1-\alpha)\mathcal{C}(\{a,c\})$.*
  - **b.** *If $c \in \mathcal{S}(a)$, $\alpha\{b\} + (1-\alpha)\{a,c\} \succeq \{a, \alpha b + (1-\alpha)c\}$, and $\mathcal{C}(\{a, \alpha b + (1-\alpha)c\}) = \{\alpha b + (1-\alpha)c\}$, then $\mathcal{C}(A) = \alpha\mathcal{C}(\{a,b\}) + (1-\alpha)\mathcal{C}(\{a,c\})$.*
- **(iii)** *For any $A \in \mathcal{A}$, $\mathcal{C}(\alpha A + (1-\alpha)\{a\}) = \alpha\mathcal{C}(A) + (1-\alpha)\{a\}$.*

[23] We do not distinguish $\{a,b\}$ from $\{b,a\}$: e.g., $a \in \mathcal{P}(b)$ implies $\{a,b\} \in \mathcal{B}_P$.

Axiom 4 states that the independence of $\succeq$ holds within the "pride domain" $\mathcal{B}_P$ and "shame domain" $\mathcal{B}_S$, and that it holds for mixtures with a singleton. Similarly, Axiom 5(i) states that domain-wise linearity of choice holds. Axiom 5(ii) is interpreted similarly, but requires additional conditions to exclude the possibility that one of the mixed menus generates pride and the other generates shame. For (ii-a), note that from $b \in \mathcal{P}(a)$ and $c \in \mathcal{N}_2(a)$, we know that a is superior to b and c in the descriptive ranking. Suppose we also know $\{a, \alpha b + (1-\alpha)c\} \succeq \alpha\{b\} + (1-\alpha)\{a, c\}$ and $\mathcal{C}(\{a, \alpha b + (1-\alpha)c\}) = \{\alpha b + (1-\alpha)c\}$. This suggests that moving a towards b makes the menu less desirable, and this is not because a is a preferred choice. We can then infer that a sets a reference point lower than b. Thus, $\{a, b\}$ never generates shame, so the linearity of choice holds if it is mixed with a menu that never generates shame. Similarly, conditions in (ii-b) suggest that a sets the reference point higher than b, so the linearity holds if $\{a, b\}$ is mixed with a menu that never generates pride. Finally, Axiom 5(iii) states that the linearity holds for to mixtures with a singleton.

The next axiom relates preferences $\succeq$ to ex-post choice $\mathcal{C}$.

Axiom 6. (Sophistication) *Suppose there exists $a^* \in A$ such that $a^* \succeq_r c$ for all $c \in A \cup B$ and $a^* \succeq_w a$ for all $a \in A$.*

- **(i)** *Suppose $a^* \succeq_w b$ for all $b \in B$. Then, $A \cup B \succeq A$. Moreover, $A \cup B \succ A$ if and only if $\mathcal{C}(A \cup B) \cap A = \emptyset$.*
- **(ii)** *Suppose $b^* \succ_w a^*$ for some $b^* \in B$. Then, $A \cup B \succeq A$ implies $\mathcal{C}(A \cup B) \cap A = \emptyset$.*

Axiom 6(i) concerns situations in which some $a^* \in A$ sets the reference point at $A \cup B$. Then, the DM weakly prefers the larger menu $A \cup B$ to A because it expands options without changing the reference point. Moreover, the larger menu is strictly preferred if and only if the added menu contains an option strictly better than all alternatives in A. Axiom 6(ii) concerns situations in which an added alternative $b^* \in B$ sets a higher reference point than the reference point at A. The DM then weakly prefers the larger menu only if the added menu B contains a strictly better alternative to be chosen than alternatives in A.

The next axiom captures the DM's *shame attitude*, i.e., how her social payoff depends on the size and direction of the deviation of her choice from the reference point. We consider a DM whose marginal utility from pride and that from shame are constant, respectively, but who may care about pride and shame differently. The following axiom captures such an attitude toward pride and shame. Let $e^{a,b}$ denote an alternative such that $\{e^{a,b}\} \sim \{a, b\}$. For $\{a, b\} \in \mathcal{B}_S$, such $e^{a,b}$ exists by Lemma S27 in Supplemental Appendix S.C.

Axiom 7. (Constant Shame Attitude) *There exists a unique $\alpha \in (0, 1)$ such that, for any $a, b, c, d \in \Delta$ with $c \in \mathcal{P}(a) \cap \mathcal{P}(b)$ and $d \in \mathcal{S}(a) \cap \mathcal{S}(b)$, we have $\alpha\{a, c\} + (1-\alpha)\{e^{b,d}\} \sim \alpha\{b, c\} + (1-\alpha)\{e^{a,d}\}$.*

To interpret Axiom 7, suppose $c \in \mathcal{P}(a) \cap \mathcal{P}(b)$ and $\{a, c\} \succ \{b, c\}$. At each menu, a or b sets the reference point and the DM feels pride by choosing c. Then, a must be considered normatively inferior to b and generates higher pride for choosing the same alternative c. Suppose also $d \in \mathcal{S}(a) \cap \mathcal{S}(b)$. Then, similarly, the DM prefers $\{a, d\}$ to $\{b, d\}$ because a generates lower shame than b does for choosing the same alternative d. Now, consider the choice between two lotteries: lottery 1 yields the payoff from the high-pride menu $\{a, c\}$ or that from the high-shame menu $\{b, d\}$ with probability α and $1-\alpha$, respectively, and lottery 2 yields the payoff from the low-pride menu $\{b, c\}$ or that from the low-shame menu $\{a, d\}$ with the same mixing rate. As α increases, lottery 1 becomes more desirable, and the DM will be indifferent between the lotteries at some α. Such α indicates the rate at which the DM trades off the gain from more pride with the loss from more shame. Axiom 7 states that this trade-off rate is constant. Moreover, the trade-off rate measures the degree of shame aversion: the higher α, the more shame-averse the DM is, because she demands a higher pride gain to compensate for the loss from shame.

Definition 3. (Shame attitudes) *(i) The DM is α-sensitive to shame if her preference $(\succeq, \mathcal{C})$ satisfies Axiom 7 with $\alpha \in (0, 1)$. (ii) The DM who is α-sensitive to shame is shame-averse if $\alpha > \frac{1}{2}$; shame-neutral if $\alpha = \frac{1}{2}$; and shame-loving if $\alpha < \frac{1}{2}$.*

Our final axiom imposes some consistency of choices across menus. Consider the donation example above which features the violation of independence. There, the DM believes that donation is uncommon but a large donation is common conditional on donating. Then, she may choose \$100 from {\$10, \$100} to avoid the shame of falling behind the social expectation of a large donation, whereas she may choose \$10 from {\$0, \$10, \$100} because a small donation nicely balances self-interest with pride from exceeding the social expectation of zero donation. This choice pattern violates the WARP. This pattern emerges because the relative attractiveness of \$10 and \$100 changes as the reference point changes. This example suggests the following axiom.

Axiom 8. (Weak WARP) *For any $A, B \in \mathcal{A}$, suppose there exists $a^* \in A \cap B$ such that $a^* \succeq_r c$ and $a^* \succeq_w c$ for all $c \in A \cup B$. Then, $a, b \in A \cap B$, $a \in \mathcal{C}(A)$ and $b \in \mathcal{C}(B)$ imply $a \in \mathcal{C}(B)$.*

$a^* \succeq_r a$ and $a^* \succeq_w a$ for all $a \in A$ suggests that a sets the reference point at A. Thus, Axiom 8 says that the standard WARP property applies to menus A and B which share a common reference-setting alternative a^*.

3.2 Representation Theorem

We show that our axioms characterize the following utility representation.

Definition 4. $(\succeq, \mathcal{C})$ *is a* pride-shame (PS) preference *if there are continuous linear functions u,w and r and a constant $\beta > 0$ such that $\succeq$ is represented by*

$$V_{PS}(A) = \max_{x \in A} U(x; A) \tag{9}$$

and $\mathcal{C}$ coincides with $\mathcal{C}_{PS}(A) = \arg\max_{x \in A} U(x; A)$, where $U(x; A)$ is written as

$$U(x; A) = u(x) - \underbrace{\max\left\{\max_{y \in \varphi_r(A)} w(y) - w(x), 0\right\}}_{\text{“shame”}} + \beta \underbrace{\max\left\{w(x) - \max_{y \in \varphi_r(A)} w(y), 0\right\}}_{\text{“pride”}} \tag{10}$$

and $\varphi_r(A) = \arg\max_{a \in A} r(a)$. This representation is called a PS representation*.*

$\max_{y \in \varphi_r(A)} w(y)$ is interpreted as the normative desirability that the DM perceives is expected to achieve, which we simply call the reference point. It consists of two distinct components. First, the *descriptive norm function* r represents the DM's belief about the prevalence/commonality of each alternative. The DM's reference set $\varphi_r(A)$ consists of alternatives which she believes is the most prevalent in A. Second, the *prescriptive norm function* w represents the DM's belief about the social desirability/admirability of each alternative. When $\varphi_r(A)$ contains multiple alternatives, the DM adopts the highest value of w in $\varphi_r(A)$ as the reference point.

We say that a DM with a PS preference *feels pride (shame) by choosing $a \in A$ at A* if $w(a) - \max_{y \in \varphi_r(A)} w(y) > (<) 0$. In words, the DM feels pride (shame) if she chooses an alternative that she perceives is normatively superior (inferior) to the reference alternative. Pride (shame) gives the DM a positive (negative) payoff. Because the DM may care about shame differently than pride (Butera et al. 2022), we allow the DM to be more or less sensitive to shame than to pride, by allowing $\beta \neq 1$.

The PS representation includes GP's model as a degenerate case. For axiomatization, however, we focus on a class of nondegenerate PS preferences.

Definition 5. (Nondegeneracy) $(\succeq, \mathcal{C})$ *is* nondegenerate *if there exist $x, y, y' \in \Delta$ such that $y \in \mathcal{P}(x)$ and $y' \in \mathcal{S}(x)$.*

Nondegeneracy ensures that the DM feels pride at some binary menu and shame at another. The temptation preference of GP is degenerate. In Section 3.4, we show that such a degenerate preference may accommodate non-unique PS representations. Thus, our main theorem focuses on preferences that accommodate a unique PS representation up to positive affine transformation (see Section 3.4). We can show that nondegeneracy holds generically if $\dim(Z) \geq 4$. See Supplemental Appendix S.D for further discussions of nondegeneracy. Note also that nondegeneracy is testable.

We now state our main theorem.

Theorem 1. *A nondegenerate preference* $(\succeq, \mathcal{C})$ *satisfies Axioms 1-8 if and only if it admits a PS representation. Moreover, the decision maker is shame-averse if* $\beta < 1$*, shame-neutral if* $\beta = 1$*, and shame-loving if* $\beta > 1$*.*

3.3 Sketch of the Proof of Theorem 1

Our proof, formally presented in Appendix A, begins by verifying that the descriptive ranking $\succeq_r$ admits a linear representation r.

Lemma 1. *If Axioms 1-5 hold, then* $\succeq_r$ *admits a linear representation* r*. The representation is unique up to positive affine transformation.*

Define the reference correspondence as $\varphi_r(A) = \{a \in A : r(a) \geq r(b)\ \forall b \in A\}$. We can then show the following important properties of φ_r.

Lemma 2. *If Axioms 1-6 hold, then the following conditions hold for any finite* $A, B \in \mathcal{A}$*.*
(i) $\varphi_r(A \cup B) = \varphi_r(A)$ *implies* $A \cup B \succeq A$*.*
(ii) $\varphi_r(A \cup B) = \varphi_r(A) \cup \varphi_r(B)$ *and* $A \succeq B$ *imply* $A \succeq A \cup B \succeq B$*.*

Lemma 2(i) states that if augmenting menu A by menu B does not affect the reference set, then the DM weakly prefers the larger menu. In this case, the DM exhibits a preference for larger menus (cf. Evren and Minardi 2017) because the addition will never worsen her social emotion. In contrast, Lemma 2(ii) states that if the addition of alternatives expands the reference set, then the set betweenness property (Gul and Pesendorfer 2001) holds. In particular, the DM exhibits a preference for smaller menus (cf. Dillenberger and Sadowski 2012) because the addition will never improve her social emotion.

Lemma 2 implies that the preferences over finite menus can be characterized by at most two elements in each menu.

Lemma 3. *If Axioms 1-6 hold, then, for any finite menu* $A \in \mathcal{A}$*, there exist* $a^* \in A$ *and* $b^* \in \varphi_r(A)$ *such that* $A \sim \{a^*, b^*\}$*.*

The remaining components u and w can be constructed by an approach similar to Gul and Pesendorfer 2001, although we address several technical difficulties due to the violation of independence and WARP. We first show that there exists a function V_{PS} that represents $\succeq$ and satisfies some linearity. Let $\mathcal{A}_f$ denote the set of all finite menus in $\mathcal{A}$.

Lemma 4. *If Axioms 1-6 hold, then there exists a function* V_{PS} *that represents* $\succeq$ *on* $\mathcal{A}_f$ *and satisfies the following property:* $A, B \in \mathcal{B}_P$ *or* $A, B \in \mathcal{B}_S$ *implies* $V_{PS}(\alpha A + (1-\alpha)B) = \alpha V_{PS}(A) + (1-\alpha)V_{PS}(B)$*.*

Lemma 4 and nondegeneracy allow us to construct two linear functions, w_P on the pride domain and w_S on the shame domain. Then, Axiom 7 implies that the two functions are proportional: $(1/\beta)w_P(x) = w_S(x) \equiv w(x)$ for some $\beta > 0$.[24]

We can show that the PS representation holds for all binary menus. Then, Lemma 3 allows us to extend the representation to all finite menus, and Axiom 2 further extends the result to all menus in $\mathcal{A}$. For choice $\mathcal{C}$, Axiom 8 extends the representation to all menus.

3.4 Uniqueness of PS Representation

3.4.1 Uniqueness of Descriptive Norm Function

$\succeq_r$ is unique given choice data $(\succeq, \mathcal{C})$, and the representation r of $\succeq_r$ is unique up to positive affine transformation. However, $\succeq_r$ is merely a specific way to reveal the underlying descriptive norms. In general, there can be two PS representations with different reference sets (i.e., different descriptive norm functions), such that both represent the same choice data.

Example 1. *Suppose that the choice data* $(\succeq, \mathcal{C})$ *are generated by the temptation preference of GP:* $V_{GP}(A) = \max_{x\in A}\{u(x) + w(x) - \max_{y\in A} w(y)\}$. *Then, a strict descriptive ranking* $a \succ_r b$ *never occurs, so the descriptive norm function which rationalizes* $\succeq_r$ *is a constant. However, the data can also be represented by another PS preference with* $r = w$*, because* $\max_{y\in\varphi_w(A)} w(y) = \max_{y\in A} w(y)$. *Thus, the data cannot distinguish the two models.*

However, we show below that when the data are generated by a PS preference that satisfies some nondegeneracy property, $\succeq_r$ correctly elicits the true descriptive norm r. Thus, among the PS preferences that satisfy the property, the descriptive norm function r which is consistent with observed data is unique up to positive affine transformation. It turns out that the following weak version of nondegeneracy is sufficient to ensure that the descriptive norm function revealed by $\succeq_r$ is the only one that is consistent with data.

Definition 6. (Weak Nondegeneracy) $(\succeq, \mathcal{C})$ *is* weakly nondegenerate *if there exist some* $\bar{a}, \bar{b} \in \Delta$ *such that* $\bar{a} \succ^* \bar{b}$.

Theorem 2. *Suppose the data are generated by a weakly nondegenerate PS preference represented by* (u, w, r, β). *Then, the following statements hold. (i)* $r(a) > r(b)$ *if and only if* $a \succ_r b$. *(ii) For* $a, b \in \Delta$ *with* $r(a) = r(b)$, $w(a) > w(b)$ *if and only if* $a \succ_w b$.

[24] Another possible way is to construct w directly from $\succeq_w$. However, $\succeq_w$ elicits the true prescriptive ranking w only between alternatives with the same descriptive ranking, so such a proof will involve an incomplete binary relation. Instead, we construct w from V_{PS} and then show in Theorem 2 that w indeed represents $\succeq_w$ on the indifference sets of $\sim_r$.

Under weak nondegeneracy, the only descriptive norm r consistent with the data is the one revealed by $\succeq_r$: if r and r' are both consistent with the data, then $r(a) \geq r(b) \Leftrightarrow a \succeq_r b \Leftrightarrow r'(a) \geq r'(b)$ for any $a, b \in \Delta$. Weak nondegeneracy is implied by nondegeneracy, so Theorem 1 focuses on the case where this uniqueness holds. Also, a similar uniqueness result holds for the prescriptive norm w, among alternatives indifferent in $\sim_r$. Note that the preference in Example 1 is excluded by weak nondegeneracy. [25] See Supplemental Appendix S.D for a graphical illustration of weak nondegeneracy and Theorem 2(i).

3.4.2 Uniqueness of (u, w, r, β)

We can also show that u and w are unique up to affine transformation and that β is unique, when $(\succeq, \mathcal{C})$ satisfies the above axioms and nondegeneracy.

Proposition 1. *Suppose a nondegenerate $(\succeq, \mathcal{C})$ satisfies Axioms 1-8. Then the following statements are equivalent.*

(i) *If a PS representation (u, w, r, β) represents $\succeq$, then (u', w', r', β') also represents $\succeq$.*

(ii) *The following properties hold.*

- **a.** $u' = \theta u + \gamma_u$ *and* $w' = \theta w + \gamma_w$ *for some* $\theta > 0$ *and* $\gamma_u, \gamma_w \in \mathbb{R}$.
- **b.** $r' = \theta_r r + \gamma_r$ *for some* $\theta_r > 0$ *and* $\gamma_r \in \mathbb{R}$.
- **c.** $\beta = \beta'$

3.5 Comparing Shame Aversion

Definition 3 yields a notion of a DM being "more shame-averse" than another DM. For two DMs $i = 1, 2$, let $(\succeq_i, \mathcal{C}_i)$ denote the preference of DM i.

Definition 7. *Suppose DM $i \in \{1, 2\}$ is α_i-sensitive to shame. Then DM 1 is (weakly) more shame-averse than DM 2 if $\alpha_1 > (\geq) \alpha_2$.*

We now state how the PS representation and observed behavior are linked in terms of (relative) shame aversion. For $i = 1, 2$, let $\mathcal{P}_i$ and $\mathcal{S}_i$ denote the set of pride-generating binary menus $\mathcal{P}$ and the set of shame-generating binary menus $\mathcal{S}$ defined in Eq. (4) and (5), respectively, for DM i. Also, let $e_i^{x,y} \in \Delta$ be such that $\{e_i^{x,y}\} \sim_i \{x, y\}$.

Proposition 2. *Suppose DM 1 and DM 2 have a PS preference, with parameters β_1 and β_2, respectively. Then, the following statements are equivalent.*

[25] Conversely, if $(\succeq, \mathcal{C})$ violates weak nondegeneracy, then r is constant. Then, Lemma 2 reduces to Axiom 4 (Set Betweenness) of GP, and V_{PS} in Lemma 4 satisfies linearity for all binary menus. Although we can possibly pursue an axiomatization of the GP representation for this case, we omit it.

(i) $\beta_1 < (\leq)\beta_2$.

(ii) *DM 1 is (weakly) more shame-averse than DM 2.*

(iii) *Take any* $\alpha \in (0,1)$ *and any* a, b, c, d *such that* $c \in \mathcal{P}_i(a) \cap \mathcal{P}_i(b)$, $d \in \mathcal{S}_i(a) \cap \mathcal{S}_i(b)$, $\{a, c\} \succ_i \{b, c\}$ *and* $\{a, d\} \succ_i \{b, d\}$ *for* $i = 1, 2$. *Then,* $\alpha \{b, c\} + (1-\alpha) \left\{e_2^{a,d}\right\} \succeq_2 \alpha \{a, c\} + (1-\alpha) \left\{e_2^{b,d}\right\}$ *implies* $\alpha \{b, c\} + (1-\alpha) \left\{e_1^{a,d}\right\} \succ_1 (\succeq_1) \alpha \{a, c\} + (1-\alpha) \left\{e_1^{b,d}\right\}$.

By the equivalence of (i) and (ii), β in the PS representation characterizes the DM's shame aversion. Also, the equivalence of (ii) and (iii) implies that we can compare the shame aversion of two DMs by the following experiment: Consider two lotteries, lottery 1 giving the payoff of a high-pride menu $\{a, c\}$ and that of a high-shame menu $\{b, d\}$ with probability α and $1-\alpha$ respectively, and lottery 2 giving the payoff of a low-pride menu $\{b, c\}$ and that of a low-shame menu $\{a, d\}$ with probability α and $1-\alpha$ respectively. Ask the DMs to choose between the two lotteries at various α. Then, DM1 is more shame-averse than DM2 if and only if DM1 chooses the "safer" lottery 2 whenever DM2 does.

4 Empirical Perspective

Our PS model can provide a basis for a revealed preference approach to studying the behavioral and welfare effects of social norms and norm-evoking policies. Instead of requiring highly rich data to verify all of the above axioms, researchers may assume that the choice data are generated by a PS peference and then derive stronger conclusions about behavior and welfare. We illustrate this point below. The proofs for the results in this section are straightforward given the representation and hence relegated to Supplemental Appendix S.C.

Some of the most important empirical questions are about whether the descriptive and prescriptive norms are different, and how they affect welfare. Economists typically emphasize either descriptive norms alone or prescriptive norms alone. A small number of studies that feature both (e.g., Allcott 2011; Hallsworth et al. 2017; Heinicke et al. 2022) do not provide detailed mechanisms for how they interact to affect behavior and welfare. Our model proposes that the descriptive and prescriptive norms influence behavior and welfare by shaping social emotions, namely pride and shame, depending on how they are aligned. Thus, it is crucial to ensure that empirical researchers can test whether descriptive and prescriptive norms are different, and how pride and shame are distributed, in a given context of interest.

The following result suggests how we can test whether the descriptive and prescriptive norms are distinct, and also how they are distinct from the intrinsic (i.e., private) preference.

Claim 1. *Suppose the DM has a PS preference. Then, the following statements hold, where* $\|$ *represents equality up to positive affine transformation.*

(i) *$r \not\parallel w$ if weak nondegeneracy holds.*

(ii) *$u \not\parallel w, r$ if there exist some $A \in \mathcal{A}$ and $a \in A$ such that $\{a\} \succ A$.*

Claim 1 provides a basis for empirically distinguishing descriptive norms, prescriptive norms, and intrinsic preferences. To illustrate, suppose a denotes making no donation and b denotes making a donation. If an individual who chooses b anyway ($\mathcal{C}(\{a, b\}) = \{b\} = \mathcal{C}(\{b\})$) is willing to make her decision publicly observed rather than leave it private ($\{a, b\} \succ \{b\}$), then this indicates that r and w are different (in particular, $r(a) > r(b)$ and $w(a) < w(b)$). This is empirically testable, e.g., by asking individuals to choose whether they donate and asking their WTP for publicly recognizing their decision (cf. Butera et al. 2022). Next, if an individual avoids the opportunity to donate ($\{a\} \succ \{a, b\}$), this indicates that u is different from w and r. This is again empirically testable, e.g., by asking the individual to choose whether they want to participate in the choice opportunity, possibly with a cost of avoidance (e.g., Dana et al. 2006; Broberg et al. 2007; Lazear et al. 2012; DellaVigna et al. 2012; Andreoni et al. 2017; Klinowski 2021).

Furthermore, the PS model allows us to test whether individuals feel pride or shame. Recall that a DM feels pride (shame) by choosing a at A if $w(a) - \max_{y \in \varphi_r(A)} w(y) > (<) 0$.

Claim 2. *Suppose $(\succeq, \mathcal{C})$ is a PS preference such that $\mathcal{C}(\{a, b\}) = \{b\}$.[26] Then,*

(i) *the DM feels pride by choosing b at $\{a, b\}$ if and only if $\{b\} \prec \{a, b\} \succ \{a\}$.*

(ii) *the DM feels shame by choosing b at $\{a, b\}$ if and only if $\{b\} \succ \{a, b\} \succ \{a\}$.*

Importantly, Claim 2 allows us to investigate the heterogeneity in social emotions. In many contexts, some people will feel pride while others feel shame (e.g., Butera et al. 2022). Understanding such heterogeneity is important for welfare analysis as well as behavioral predictions, because the welfare effects of policies can crucially depend on the distribution of behavioral responses (Allcott et al. 2025). Our framework allows researchers to investigate how behavioral responses to a policy and associated social emotions differ, e.g., between female and male workers or between black and white students.

Finally, our framework facilitates an appropriate welfare analysis even when individuals do not feel any social emotion. Recall the example of prosocial behavior in Section 2.1 with norms (r', w) such that $r'(0) < r'(1)$ and $w(0) < w(1)$. In this case, both norms dictate prosocial behavior $x = 1$, and the DM follows the norms to avoid the shame from deviating to $x = 0$, so she does not feel pride or shame from her choice. However, a public recognition policy that forces the DM to make a choice at $\{0, 1\}$ is welfare reducing relative to the outside option of avoiding prosocial behavior in private ($\{0\} \succ \{0, 1\}$). Incorporating preferences

[26] If $\mathcal{C}(\{a, b\}) = \{a, b\}$, then the social emotion depends on the realization of the ex-post choice.

over menus is useful to study the welfare effects of norms and norm-evoking policies with a revealed preference approach, when social pressure forces the DM to choose an action that she nonetheless wants to have removed (cf. Bursztyn et al. 2025).

5 Related Literature

Social Norms and Image Concerns. To the best of our knowledge, this paper is the first to formalize and axiomatize the notion of descriptive and prescriptive norms in a decision-theoretic model. In social psychology (Cialdini et al. 1991; Schultz et al. 2007), *descriptive norms* refer to norms that dictate individuals to do what is typically done by others, and *prescriptive* (or *injunctive*) *norms* refer to norms which dictate them to do what people approve of. Bicchieri and Dimant (2022) define a social norm as a behavioral rule that individuals prefer to follow because they believe that (i) others follow it *and* (ii) others think it should be followed. In our model, social norms are shaped by two functions r and w, where r represents the former belief (perception) and w represents the latter.[27] Also, the DM feels payoff-relevant social emotions by comparing her own choice with others' choice (if r reflects others' choice), which closely follows the literature on social comparisons (Festinger 1954).

We also contribute to the growing literature on social pressure or image concerns (e.g., Bénabou and Tirole 2006, and see also footnote 1). Our contribution is to propose a model with image concerns which is useful for applied analysis. By distinguishing the two types of norms in a simple manner, the model explains various behavioral patterns parsimoniously, and it clarifies the mechanisms behind the behavioral and welfare impacts of norms and norm-evoking policies (see Section 2). Our model is also tractable in that it does not impose an equilibrium assumption (unlike social signaling models); instead, it is directly disciplined by observable choice data. This facilitates analysis of misperceived norms. Finally, our utility representation is transparently linked to choice data, and it facilitates empirical analysis based on revealed preferences. Section 4 discusses possible applied analyses.

Axiomatic Decision Theory. Our model relates to the axiomatic two-stage models of choices with temptation (Gul and Pesendorfer 2001; Noor and Takeoka 2015) or social emotions (Dillenberger and Sadowski 2012; Saito 2015; Evren and Minardi 2017; Hashidate 2021), and axiomatic models of endogenous reference points (Ok et al. 2015; Lleras et al. 2019; Kıbrıs et al. 2023). The former consider decision problems where the DM chooses a menu of alternatives in the first stage and then chooses an alternative from the menu in the

[27]Bicchieri and Dimant (2022) refer to the former belief as *empirical expectations* and the latter as *normative expectations*, and they define a social norm as a behavioral rule governed by these expectations.

second stage. The models with social emotions, including ours, also assume that the first stage is private whereas the second stage is publicly observed, and that the DM anticipates social emotions due to public observability when making the first-stage choice. However, our model differs from the previous models in important ways. The latter models of endogenous reference do not include preferences over menus, so our way to elicit reference is new. Below, we discuss each paper in more detail.

Gul and Pesendorfer (2001) and Noor and Takeoka (2015) consider preferences over menus of lotteries. Noor and Takeoka (2015) also use choices from each menu, as we do, to derive a *Menu-Dependent Self-Control (MDSC)* representation. In the MDSC representation, the self-control cost (similar to the social payoffs in our model) of choosing x from A is specified by $\psi\left(\max_{y\in A} w(y)\right)\left(\max_{y\in A} w(y) - w(x)\right)$, where $\psi(\cdot) \geq 0$. Both their and our representations generalize the GP representation by relaxing the independence axiom (and WARP). While their self-control cost function is more flexible than our piecewise linear "cost" of social emotions, our emotional cost can be negative, which is essential for generating pride.

Dillenberger and Sadowski (2012) and Evren and Minardi (2017) study preferences over menus consisting of social allocations of non-stochastic objects (e.g., dictator games). Dillenberger and Sadowski (2012) characterize shame, which involves a preference for smaller menus. In contrast, Evren and Minardi (2017) characterize warm-glow, which involves a preference for larger menus. Our axioms capture both types of preferences, depending on whether adding alternatives to a menu expands the reference set (see Lemma 2).

Saito (2015) and Hashidate (2021) study preferences over menus consisting of social allocations $p = (p_i)_{i\in\{1\}\cup S}$ of lotteries, where 1 denotes the DM and S denotes the set of other agents. Saito (2015) derives the *generalized utilitarian* (GU) representation, which generates the pride $\beta_1 \max_{q\in A} \alpha_1(u(q_1) - u(p_1)) > 0$ of acting altruistically if the DM compromises her private payoff, and the shame $-\beta_S \max_{r\in A}(\sum_{i\in S} \alpha_i u(r_i) - \sum_{i\in S} \alpha_i u(p_i)) < 0$ of acting selfishly if the DM compromises other agents' private payoffs.[28] Hashidate (2021) generalizes the GU representation, allowing for various social emotions. Their models generate social emotions by comparing the DM's or other agents' private payoffs to reference points. In contrast, pride and shame in our model arise from the comparison of the perceived normative desirability of the own choice against the reference alternative's desirability. We believe separating these emotions from private payoffs is vital for two reasons. First, our model formalizes the concept of social norms discussed above closely. Moreover, we argue that pride and shame should depend on the degree to which the DM can live up to social expectations

[28]Saito (2015) allows for $\beta_1 < 0$, expressing the case of the temptation to act selfishly.

(norms), not the degree to which the DM's or other agents' private payoffs are sacrificed.[29] Second, our formulation explains empirical findings discussed in Section 2 straightforwardly. For example, when information about others' behavior alters the DM's choice, it is plausibly due to changes in perceived norms, rather than changes in private payoffs.

Our model also has differences from previous axiomatic models of endogenous reference dependence,[30] besides how to elicit the reference. Ok et al. (2015) characterize choice behavior which exhibits the "attraction effect." In their model, a dominated (hence unchosen) alternative serves as a reference alternative and restricts the choice set to alternatives that dominate it. In contrast, in our model, the reference alternative may be chosen, and it affects the preferences (beliefs) but not choice sets. Lleras et al. (2019) consider a preference over state-contingent contracts (acts) and derive a representation that evaluates an act based on its expected value and expected gain/loss relative to the expected value. Their representation allows the DM to derive payoffs from either expected gain or loss, but not both. In contrast, the DM in our model may feel pride from an alternative and shame from another. Kıbrıs et al. (2023) consider choice problems generated from a finite set of alternatives, and derive a model where the reference point is determined by an endogenously derived conspicuity ranking, just as an endogenously derived r defines reference in our model. Their representation is quite general, but they do not characterize a specific structure of the reference point in our model, so our axiomatization result is not implied by their work.

6 Conclusion

Despite the growing interest in using social norms to influence behavior, their behavioral and welfare effects are not well understood. We propose and axiomatize a model of reference-dependent decision-making in which the decision maker's perception of others' choice (descriptive norm) and her perception of others' normative opinions (prescriptive norm) together shape a reference point. The key drivers of behavior are social emotions, specifically a positive payoff from pride, which she enjoys if her choice exceeds the reference point, and a negative payoff from shame, which she suffers if her choice falls short of it. The simple model provides useful implications, e.g., when public recognition programs or norm nudges likely induce prosocial behavior, how policies can incentivize more prosocial behavior but also induce choice avoidance. It also explains why aligned descriptive and prescriptive norms are more effective for inducing prosocial behavior compared to misaligned ones, and why the

[29] Scheff (1988) discusses how perceived social expectations set a system in which conformity to norms is sustained by the reward of pride and punishment of shame.

[30] Köszegi and Rabin (2006) develop a non-axiomatic model of endogenous reference formation. In their model, the agent's reference point is constrained by rational expectations.

descriptive norms may have a larger impact in the latter case. Moreover, the model is simple yet useful to study the mechanisms behind policy effects, tractable because it does not impose an equilibrium assumption, and transparent in its relation to observed choice, which may usefully guide empirical analysis based on a revealed preference approach.

This paper has several limitations. First, our model relies on expected utility functions, which may not fully capture the complexity of social decision-making particularly when individuals care not only about outcomes but also about how those outcomes are generated (cf. Saito 2013). Second, we do not model how individuals' perceptions are shaped. For example, they may arise as equilibrium objects of a game (cf. Bénabou and Tirole 2006), and an equilibrium restriction may be necessary to study how norms evolve over time. Alternatively, individuals may form perceptions in a self-serving manner (Heinicke et al. 2022; Bicchieri et al. 2023), and this process may be crucial for predicting the effects of norm nudges. Third, norms may be specified more flexibly. For example, an individual may compare her behavior with the behavior of a group of individuals rather than that of a "typical person." The reference point may then depend on the distribution of the descriptive norms of others, yielding a "random Strotz" representation (Dekel and Lipman 2012). Eliciting information on individuals' reference groups from their choice is an interesting topic for future research.

References

Akerlof, George A. and Rachel E. Kranton, "Economics and Identity," *The Quarterly Journal of Economics*, 08 2000, *115* (3), 715–753.

Allcott, Hunt, "Social Norms and Energy Conservation," *Journal of Public Economics*, 2011, *95* (9-10), 1082–1095.

______ **and Todd Rogers**, "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation," *American Economic Review*, 2014, *104* (10), 3003–3037.

______**, Daniel Cohen, William Morrison, and Dmitry Taubinsky**, "When Do "Nudges" Increase Welfare?," *American Economic Review*, May 2025, *115* (5), 155596.

Andreoni, James and B. Douglas Bernheim, "Social Image and the 5050 Norm: A Theoretical and Experimental Analysis of Audience Effects," *Econometrica*, 2009, *77* (5), 1607–1636.

______**, Justin M Rao, and Hannah Trachtman**, "Avoiding the Ask: A Field Experiment on Altruism, Empathy, and Charitable giving," *Journal of Political Economy*, 2017, *125* (3), 625–653.

Bénabou, Roland and Jean Tirole, "Incentives and Prosocial Behavior," *American Economic Review*, December 2006, *96* (5), 1652–1678.

Bicchieri, Cristina, *The Grammar of Society: The Nature and Dynamics of Social norms*, Cambridge University Press, 2005.

______ **and Erte Xiao**, "Do the Right Thing: But Only if Others Do So," *Journal of Behavioral Decision Making*, 2009, *22* (2), 191–208.

______ **and Eugen Dimant**, "Nudging with Care: The Risks and Benefits of Social Information," *Public Choice*, 2022, *191* (3-4), 443–464.

______ **,** ______ **, and Silvia Sonderegger**, "It's not a lie if you believe the norm does not apply: Conditional norm-following and belief distortion," *Games and Economic Behavior*, 2023, *138*, 321–354.

Broberg, Tomas, Tore Ellingsen, and Magnus Johannesson, "Is Generosity Involuntary?," *Economics Letters*, 2007, *94* (1), 32–37.

Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott, "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia," *American Economic Review*, 2020, *110* (10), 2997–3029.

______ **and Robert Jensen**, " How Does Peer Pressure Affect Educational Investments? ," *The Quarterly Journal of Economics*, 05 2015, *130* (3), 1329–1367.

______ **and** ______ , "Social Image and Economic Behavior in the Field: Identifying, Understanding, and Shaping Social Pressure," *Annual Review of Economics*, 2017, *9* (1), 131–153.

______ **, Benjamin R Handel, Rafael Jimenez-Duran, and Christopher Roth**, "When Product Markets Become Collective Traps: The Case of Social Media," Working Paper 2025.

______ **, Thomas Fujiwara, and Amanda Pallais**, "'Acting Wife': Marriage Market Incentives and Labor Market Investments," *American Economic Review*, November 2017, *107* (11), 3288–3319.

Butera, Luigi, Robert Metcalfe, William Morrison, and Dmitry Taubinsky, "Measuring the Welfare Effects of Shame and Pride," *American Economic Review*, 2022, *112* (1), 122–168.

Cialdini, Robert B., "Crafting Normative Messages to Protect the Environment," *Current Directions in Psychological Science*, 2003, *12* (4), 105–109.

Cialdini, Robert B, Carl A Kallgren, and Raymond R Reno, "A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior," in "Advances in Experimental Social Psychology," Vol. 24, Elsevier, 1991, pp. 201–234.

Dana, Jason, Daylian M Cain, and Robyn M Dawes, "What You Dont Know Wont Hurt Me: Costly (but Quiet) Exit in Dictator Games," *Organizational Behavior and Human Decision Processes*, 2006, *100* (2), 193–201.

Dekel, Eddie and Barton L. Lipman, "Costly Self-Control and Random Self-Indulgence," *Econometrica*, 2012, *80* (3), 1271–1302.

DellaVigna, Stefano, John A List, and Ulrike Malmendier, "Testing for Altruism and Social Pressure in Charitable Giving," *The Quarterly Journal of Economics*, 2012, *127* (1), 1–56.

______ **, John A. List, Ulrike Malmendier, and Gautam Rao**, "Voting to Tell Others," *The Review of Economic Studies*, 10 2017, *84* (1), 143–181.

Dillenberger, David and Philipp Sadowski, "Ashamed to be selfish," *Theoretical Economics*, 2012, *7* (1), 99–124.

Evren, Özgür and Stefania Minardi, "WarmGlow Giving and Freedom to be Selfish," *The Economic Journal*, 02 2017, *127* (603), 1381–1409.

Festinger, Leon, "A Theory of Social Comparison Processes," *Human Relations*, 1954, *7* (2), 117–140.

Frey, Bruno S and Benno Torgler, "Tax Morale and Conditional Cooperation," *Journal of Comparative Economics*, 2007, *35* (1), 136–159.

______ **and Stephan Meier**, "Social Comparisons and Pro-social Behavior: Testing Conditional Cooperation in a Field Experiment," *American Economic Review*, 2004, *94* (5), 1717–1722.

Gerber, Alan S., Donald P. Green, and Christopher W. Larimer, "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment," *American Political Science Review*, 2008, *102* (1), 3348.

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

Hallsworth, Michael, John A List, Robert D Metcalfe, and Ivo Vlaev, "The Behavioralist as Tax Collector: Using Natural Field Experiments to enhance tax compliance," *Journal of Public Economics*, 2017, *148*, 14–31.

Hashidate, Yosuke, "Social Image Concern and Reference Point Formation," *The B.E. Journal of Theoretical Economics*, 2021, *21* (1), 19–84.

Heinicke, Franziska, Christian König-Kersting, and Robert Schmidt, "Injunctive vs. descriptive social norms and reference group dependence," *Journal of Economic Behavior Organization*, 2022, *195*, 199–218.

Kıbrıs, Özgür, Yusufcan Masatlioglu, and Elchin Suleymanov, "A theory of reference point formation," *Economic Theory*, 2023, *75* (1), 137–166.

Klinowski, David, "Reluctant Donors and Their Reactions to Social Information," *Experimental Economics*, 2021, *24* (2), 515–535.

Köszegi, Botond and Matthew Rabin, "A Model of Reference-Dependent Preferences," *The Quarterly Journal of Economics*, 11 2006, *121* (4), 1133–1165.

Lacetera, Nicola and Mario Macis, "Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme," *Journal of Economic Behavior & Organization*, 2010, *76* (2), 225 – 237.

Lazear, Edward P, Ulrike Malmendier, and Roberto A Weber, "Sorting in experiments with application to social preferences," *American Economic Journal: Applied Economics*, 2012, *4* (1), 136–163.

Lleras, Juan Sebastián, Evan Piermont, and Richard Svoboda, "Asymmetric gain–loss reference dependence and attitudes toward uncertainty," *Economic Theory*, 2019, *68* (3), 669–699.

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

Noor, Jawwad and Norio Takeoka, "Menu-dependent self-control," *Journal of Mathematical Economics*, 2015, *61*, 1–20.

Ok, Efe A., Pietro Ortoleva, and Gil Riella, "Revealed (P)Reference Theory," *American Economic Review*, January 2015, *105* (1), 299–321.

Perez-Truglia, Ricardo and Ugo Troiano, "Shaming tax delinquents," *Journal of Public Economics*, 2018, *167*, 120–137.

Rabin, Matthew, "Moral Preferences, Moral Constraints, and Self-serving Biases," 1995.

Saito, Kota, "Social preferences under risk: Equality of opportunity versus equality of outcome," *American Economic Review*, 2013, *103* (7), 3084–3101.

______ , "Impure altruism and impure selfishness," *Journal of Economic Theory*, 2015, *158*, 336–370.

Scheff, Thomas J., "Shame and Conformity: The Deference-Emotion System," *American Sociological Review*, 1988, *53* (3), 395–406.

Schultz, P. Wesley, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldstein, and Vladas Griskevicius, "The Constructive, Destructive, and Reconstructive Power of Social Norms," *Psychological Science*, 2007, *18* (5), 429–434.

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

Toussaert, Séverine, "Eliciting Temptation and Self-control through Menu Choices: A Lab Experiment," *Econometrica*, 2018, *86* (3), 859–889.

Tyran, Jean-Robert and Lars P Feld, "Achieving Compliance when Legal Sanctions are Non-deterrent," *Scandinavian Journal of Economics*, 2006, *108* (1), 135–156.

A Proofs

We first prove the sufficiency part of Theorem 1, i.e., that Axioms 1-8 are sufficient for a nondegenerate preference to have a PS representation. We next prove Theorem 2, which is used to show the necessity part of Theorem 1, i.e., that the representation implies the axioms. Finally, we prove the remaining results. Straightforward results and proofs, including the proof of the necessity part, are relegated to Supplemental Appendix S.C.

We denote a mixed lottery $\alpha a+(1-\alpha)b \in \Delta$ by $a\alpha b$ and a mixed menu $\alpha A+(1-\alpha)B \in \mathcal{A}$ by $A\alpha B$, for any $a, b \in \Delta$, $A, B \in \mathcal{A}$, and $\alpha \in [0,1]$.

We note that under Axioms 1-5, the definition of $\succ_r$ can be simplified as follows: $a \succ_r b$ if and only if there exists some $c \in \text{int}(\Delta)$ such that $c \not\succ^* a$ and $c \succ^* b$. See Lemma S21. Below, we use this simplification without mention.

A.1 Proof of Theorem 1 (Sufficiency of Axioms)

A.1.1 Proof of Lemma 1

We prove that $\succeq_r$ satisfies the following properties.

Completeness. Immediate from the definition of $\succeq_r$.

Transitivity. It is straightforward to show that $\succ_r$ is transitive (see Lemma S22). Now, suppose $a \succeq_r b \succeq_r c$. If $a \succ_r b \succ_r c$, then $a \succ_r c$. Next, suppose $a \succ_r b \sim_r c$. Then, we have $d \not\succ^* a$ and $d \succ^* b$ for some $d \in \text{int}(\Delta)$. By, $b \sim_r c$, we have $d \succ^* c$, hence $a \succ_r c$. We can similarly show $a \succ_r c$ if $a \sim_r b \succ_r c$. Finally, suppose $a \sim_r b \sim_r c$. If $a \not\succeq_r c$, then we have $c \succ_r a \sim_r b$, so the above argument yields $c \succ_r b$, a contradiction. Thus, $a \succeq_r c$.

Independence. Suppose $a \succ_r b$. Then $d \not\succ^* a$ and $d \succ^* b$ for some $d \in \text{int}(\Delta)$. We then have $d\alpha c \succ^* b\alpha c$ (see Lemma S19), and Axiom 3(iv) implies $d\alpha c \not\succ^* a\alpha c$, so $a\alpha c \succ_r b\alpha c$.

Archimedeanity. Let $a \succ_r b \succ_r c$. Then, there exist $d, e \in \text{int}(\Delta)$ such that $d \not\succ^* a$, $d \succ^* b$, $e \not\succ^* b$, and $e \succ^* c$. By Lemma S20 and Axiom 3(iv), we have $d\alpha c \not\succ^* a\alpha c$, $d\alpha c \succ^* b$, and $e \succ^* a\beta c$ for some $\alpha, \beta \in (0,1)$. Thus, we have $a\alpha c \succ_r b \succ_r a\beta c$.

Thus, by the Mixture Space Theorem, $\succeq_r$ admits a linear representation r, and the representation is unique up to positive affine transformation. $\square$

A.1.2 Proof of Lemma 2

(i) If $A \succ A \cup B$, then there exists $b \in B \setminus A$ such that $b \succ_r a$ or $b \succ_w a$ for all $a \in A$ (see Lemma S24(i)). By Axiom 3(i), we have $b \succeq_r a$, hence $r(b) \geq r(a)$, for all $a \in A$. Thus, $\varphi_r(A \cup B) \neq \varphi_r(A)$.

(ii) Suppose $\varphi_r(A \cup B) = \varphi_r(A) \cup \varphi_r(B)$ and $A \succeq B$. If $A \succeq B \succ A \cup B$, then by Lemma S24(i) there exist some $a^* \in A \setminus B$ such that $a^* \succ_r b$ or $a^* \succ_w b$ for all $b \in B$ and some $b^* \in B \setminus A$ such that $b^* \succ_r a$ or $b^* \succ_w a$ for all $a \in A$, which contradicts Axiom 3(i). Next, suppose $A \cup B \succ A \succeq B$. Note that at least one of $\mathcal{C}(A \cup B) \cap A$ or $\mathcal{C}(A \cup B) \cap B$ is nonempty. If $\mathcal{C}(A \cup B) \cap A \neq \emptyset$, then there exists $b \in B \setminus A$ such that $r(b) > r(a)$ for all $a \in A$ (see Lemma S24(ii)). Then, $\varphi_r(A \cup B) = \varphi_r(B) \neq \varphi_r(A) \cup \varphi_r(B)$, a contradiction. A similar contradiction results if $\mathcal{C}(A \cup B) \cap B \neq \emptyset$. Thus, $A \succeq A \cup B \succeq B$. $\square$

A.1.3 Proof of Lemma 3.

For any $a \in A$, pick $b_a \in \varphi_r(A)$ such that $\{a, b\} \succeq \{a, b_a\}$ for all $b \in \varphi_r(A)$. Let $a^* \in A$ be such that $\{a^*, b_{a^*}\} \succeq \{a, b_a\}$ for all $a \in A$ and let $b^* \equiv b_{a^*}$. Then, iteratively applying Lemma 2(ii), $\{a^*\} \cup \varphi_r(A) = \cup_{b \in \varphi_r(A)} \{a^*, b\} \succeq \{a^*, b^*\}$. Further, because $A = (\{a^*\} \cup \varphi_r(A)) \cup (A \setminus (\{a^*\} \cup \varphi_r(A)))$ and $\varphi_r(A) = \varphi_r(\{a^*\} \cup \varphi_r(A))$, applying Lemma 2(i) yields $A \succeq \{a^*\} \cup \varphi_r(A) \succeq \{a^*, b^*\}$. Next, by construction, $\{a^*, b^*\} \succeq \{a, b_a\}$ for all $a \in A$. Then iteratively applying Lemma 2(ii) yields $\{a^*, b^*\} \succeq \cup_{a \in A} \{a, b_a\} = A$. $\square$

A.1.4 Proof of Lemma 4

For $j \in \{P, S\}$, let $\mathcal{A}_j = \left\{A \in \mathcal{A} : A = \sum_{m=1}^{M} \alpha_m A_m, A_m \in \mathcal{B}_j, M < \infty\right\}$ denote the set of finite mixtures over $\mathcal{B}_j$. Also, for notational simplicity, define $\mathcal{A}_N = \mathcal{B}_N$. By Lemma 1 of GP, for each $j \in \{P, S\}$, $\succeq$ restricted to $\mathcal{A}_j$ has a linear representation V^j. In addition, by Axioms 3 and 6, $b \in \mathcal{N}_2(a)$ implies $\{a, b\} \sim \{a\}$ (see Lemma S25). Therefore, for each $A \in \mathcal{A}_N$, we have $A \sim \{a^A\}$ for some known $a^A \in A$. Thus, $V^N(A) \equiv V^P(\{a^A\})$ represents $\succeq$ on $\mathcal{A}_N$. By Lemma 1 of GP, each representation's restriction to singleton sets is continuous.

By construction, for any $j, k \in \{P, S, N\}$, $V^j(\{a\}) \geq V^j(\{b\}) \Leftrightarrow \{a\} \succeq \{b\} \Leftrightarrow V^k(\{a\}) \geq V^k(\{b\})$. The linear representation of $\succeq$ on singletons is unique up to positive affine transformation, so we can normalize V^j so that $V^j(\{a\}) = V^k(\{a\}) \equiv V^{\text{singleton}}(\{a\})$ for all a and all j, k.

To link representations of $\succeq$ across different domains, we note that if $A \in \mathcal{A}_S \cup \mathcal{A}_N$, then there exists some $e \in \Delta$ such that $A \sim \{e\}$ (see Lemma S27). Intuitively, such A is either indifferent to some $a \in A$, or we can find $a, a' \in A$ such that $\{a'\} \succ A \succ \{a\}$, in which case A is indifferent to $a' \alpha a$ for some α.

We now obtain the desired representation of $\succeq$ on $\mathcal{A}_P \cup \mathcal{A}_S \cup \mathcal{A}_N$, which contains all binary menus. For notational simplicity, in Lemma 5, we eliminate from $\mathcal{A}_P$ menus which are contained in $\mathcal{A}_S$ (i.e., $\mathcal{A}_P$ denotes $\mathcal{A}_P \setminus \mathcal{A}_S$).[31]

[31] Such duplicates arise from menus $\{a, b\}$ such that $b \in \mathcal{N}_1(a)$, which belongs to $\mathcal{B}_P \cap \mathcal{B}_S$.

Lemma 5. *Suppose Axioms 1-6 hold. Define* $V : \mathcal{A}_P \cup \mathcal{A}_S \cup \mathcal{A}_N \to \mathbb{R}$ *by*

$$V(A) = \sum_{j \in \{P,S,N\}} V^j(A) \times I\{A \in \mathcal{A}_j \text{ and } |A| > 1\} + V^{\text{singleton}}(A) \times I\{|A| = 1\}.$$

Then, V *represents* $\succeq$ *on* $\mathcal{A}_P \cup \mathcal{A}_S \cup \mathcal{A}_N$*. Moreover,* $A, B \in \mathcal{B}_P$ *or* $A, B \in \mathcal{B}_S$ *implies* $V(A\alpha B) = \alpha V(A) + (1-\alpha)V(B)$.

Proof. Note any $A \in \mathcal{A}_P \cup \mathcal{A}_S \cup \mathcal{A}_N$ belongs to exactly one of $\{B \in \mathcal{A}_j : |B| > 1\}$, $j = P, S, N$, or $\{B : |B| = 1\}$ (by the re-definition of $\mathcal{A}_P$). Suppose $A \succ B$ for $A \in \mathcal{A}_j$ and $B \in \mathcal{A}_k$. If $|A| = 1$, then $V(A) = V^{\text{singleton}}(A) = V^k(A) > V^k(B) = V(B)$. A similar result obtains if $|B| = 1$. Now, suppose $|A| > 1$ and $|B| > 1$. If $j = k = P$, then $V(A) = V^P(A) > V^P(B) = V(B)$. Otherwise, by Lemma S27, there exists some $e \in \Delta$ such that $A \sim \{e\}$ or $B \sim \{e\}$. For the former case, $V(A) = V^j(A) = V^j(\{e\}) = V^k(\{e\}) > V^k(B) = V(B)$, thus $V(A) > V(B)$. Proof for the latter case is analogous. Thus, $A \succ B \Rightarrow V(A) > V(B)$ holds. Similarly, $B \succeq A \Rightarrow V(B) \geq V(A)$ holds.

To prove the last statement, note that for $A \in \mathcal{A}_j$, $j \in \{P, S\}$, we have $V(A) = V^j(A)$. Because $A\alpha B \in \mathcal{A}_j$ for any $A, B \in \mathcal{B}_j$, we obtain $V(A\alpha B) = V^j(A\alpha B) = \alpha V^j(A) + (1-\alpha)V^j(B) = \alpha V(A) + (1-\alpha)V(B)$. □

Finally, we obtain the desired representation as follows. Take V from Lemma 5 and define, for each finite $A \in \mathcal{A}$, $V_{PS}(A) = V\left(\{a^{A*}, b^{A*}\}\right)$ where $a^{A*} \in A$ and $b^{A*} \in \varphi_r(A)$ are constructed as in Lemma 3. Then $A \succeq B \Leftrightarrow \{a^{A*}, b^{A*}\} \succeq \{a^{B*}, b^{B*}\} \Leftrightarrow V\left(\{a^{A*}, b^{A*}\}\right) \geq V\left(\{a^{B*}, b^{B*}\}\right) \Leftrightarrow V_{PS}(A) \geq V_{PS}(A)$. Thus, we have obtained the desired function V_{PS}. □

We additionally have the following result, analogous to GP.

$$V_{PS}(A) = \max_{a \in A} \min_{b \in \varphi_r(A)} V_{PS}(\{a, b\}) = \min_{b \in \varphi_r(A)} \max_{a \in A} V_{PS}(\{a, b\}). \tag{11}$$

Proof is straightforward given the proof of Lemma 3, hence omitted.

A.1.5 Proof of Theorem 1 (Sufficiency), Continued

For some two-component mixtures (denoted A) of binary menus, Axioms 5 and 6 identify a binary subset of A to which A is indifferent. (Proof is in Supplemental Appendix S.C.)

Lemma 6. *Suppose Axioms 1-6 hold. Let* $A = \{a, b\}\,\alpha\,\{c, d\} \in \mathcal{A}$*. Then, the following statements hold.*

(i) If $b \in \mathcal{P}(a) \cup \mathcal{N}_1(a)$ *and* $d \in \mathcal{P}(c) \cup \mathcal{N}_1(c)$*, then* $A \sim \{a\alpha c, b\alpha d\}$*.*

(ii) If $b \in \mathcal{S}(a) \cup \mathcal{N}_1(a)$ *and* $d \in \mathcal{S}(c) \cup \mathcal{N}_1(c)$*, then* $A \sim \{a\alpha c, b\alpha d\}$*.*

(iii) *If $b \in \mathcal{I}(a)$, $\{a,b\} \succ \{a\}$, and $d \in \mathcal{S}(c)$, then, $\{a\alpha c, b\alpha d\} \succeq A$. If, in addition, $\{b\} \succ \{a,b\}$, then $\{a\alpha c, b\alpha d\} \sim A$.*

(iv) *Suppose $b \in \mathcal{I}(a)$, $\{b\} \succ \{a,b\} \sim \{a\}$, $a \in \mathcal{C}(\{a,b\})$, and $d \in \mathcal{S}(c)$. Then, $A \sim \{a\}\,\alpha\,\{c,d\} \succeq \{a\alpha c, b\alpha d\}$, and the latter relation is strict if and only if $\mathcal{C}(\{a,b\}) = \{a\}$.*

(v) *Suppose $b \in \mathcal{I}(a)$, $\{a\} \sim \{a,b\} \sim \{b\}$, and $d \in \mathcal{S}(c)$. Then:*

(v-a) *$\mathcal{C}(\{a,b\}) = \{a\}$ implies $\{b\alpha c, a\alpha d\} \succ A \sim \{a\}\,\alpha\,\{c,d\} \succ \{a\alpha c, b\alpha d\}$.*

(v-b) *$\mathcal{C}(\{a,b\}) = \{a,b\}$ implies $A \sim \{a\}\,\alpha\,\{c,d\} \sim \{a\alpha c, b\alpha d\}$.*

We now construct u and w. For any $a \in \Delta$, let $u(a) = V_{PS}(\{a\})$. By nondegeneracy (Definition 5), there exist $x, y, y' \in \Delta$ such that $y \in \mathcal{P}(x)$ and $y' \in \mathcal{S}(x)$. Below, we fix such x, y, y'. Note there exists $\delta \in (0,1)$ such that, for all $c \in \Delta$, we have $y(1-\delta)c \in \mathcal{P}(x)$ and $y'(1-\delta)c \in \mathcal{S}(x)$ (see Lemma S28). By Axioms 3 and 6, we have $\{x, y(1-\delta)c\} \succ \{x\}$ and $\{x, y'(1-\delta)c\} \succ \{x\}$ (see Lemma S25). Now, define w_P and w_S by

$$w_P(c; x, y, \delta) = \frac{1}{\delta} V_{PS}(\{x, y(1-\delta)c\}) - \frac{1-\delta}{\delta} V_{PS}(\{x, y\}) - V_{PS}(\{c\})$$
$$w_S(c; x, y', \delta) = \frac{1}{\delta} V_{PS}(\{x, y'(1-\delta)c\}) - \frac{1-\delta}{\delta} V_{PS}(\{x, y'\}) - V_{PS}(\{c\}).$$

$w_P(c; x, y, \delta)$ measures how the utility changes as the reference alternative x is moved slightly toward c, keeping the ex post choice constant.$w_S(c; x, y', \delta)$ is interpreted analogously. The next two lemmas show some properties of w_P and w_S, including its linearity and independence of the specific choice of $x, y, y' \in \Delta$. Proof is in Supplemental Appendix S.C.

Lemma 7. *Suppose Axioms 1-6 hold. If $y(1-\delta)c \in \mathcal{P}(x)$ for all $c \in \Delta$, then the following statements hold.*

(i) *If $c \in \mathcal{P}(x)$, then $w_P(c; x, y, \delta) = V_{PS}(\{x, c\}) - V_{PS}(\{c\})$.*

(ii) *$w_P(x; x, y, \delta) = 0$.*

(iii) *$w_P(c\alpha c'; x, y, \delta) = \alpha w_P(c; x, y, \delta) + (1-\alpha) w_P(c'; x, y, \delta)$ for any $\alpha \in (0,1)$.*

(iv) *$w_P(c; x, y, \delta') = w_P(c; x, y, \delta)$ for any $\delta' \in (0, \delta)$.*

(v) *Suppose $b(1-\delta)c \in \mathcal{P}(a)$ for all $c \in \Delta$. Then $w_P(c; x, y, \delta) = w_P(c; a, b, \delta) + w_P(a; x, y, \delta)$.*

Similarly, we have the following result. The proof is analogous, hence omitted.

Lemma 8. *Suppose Axioms 1-6 hold. If $y'(1-\delta)c \in \mathcal{S}(x)$ for all $c \in \Delta$, then the following properties hold.*

(i) *If $c \in \mathcal{S}(x)$, then $w_S(c; x, y', \delta) = V_{PS}(\{x, c\}) - V_{PS}(\{c\})$.*

(ii) *$w_S(x; x, y', \delta) = 0$.*

(iii) $w_S(c\alpha c'; x, y', \delta) = \alpha w_S(c; x, y', \delta) + (1-\alpha) w_S(c'; x, y', \delta)$ *for any* $\alpha \in (0,1)$.

(iv) $w_S(c; x, y', \delta') = w_S(c; x, y', \delta)$ *for all* $\delta' \in (0, \delta)$.

(v) Suppose $b' \in \mathcal{S}(a')$ *for all* $c \in \Delta$. *Then* $w_S(c; x, y', \delta) = w_S(c; a', b', \delta) + w_S(a'; x, y', \delta)$.

Now, suppose Axiom 7 holds with some $\alpha \in (0,1)$. Then, a simple algebra shows that $\beta \equiv \frac{1}{\alpha} - 1 > 0$ satisfies $\frac{1}{\beta} w_P(c; x, y, \delta) = w_S(c; x, y', \delta) \equiv w(c; x, y, y', \delta)$ for all $c \in \Delta$ (see Lemma S29). We first show that the representation holds for binary menus that include x.

Lemma 9. *Suppose Axioms 1-8 hold. Consider* $x, b \in \Delta$ *such that* $r(x) \geq r(b)$ *and* $V_{PS}(\{x,b\}) \geq V_{PS}(\{x\})$. *Suppose* $y(1-\delta)c \in \mathcal{P}(x)$ *and* $y'(1-\delta)c \in \mathcal{S}(x)$ *for all* $c \in \Delta$. *Then* V_{PS} *is expressed as*

$$V_{PS}(\{x,b\}) = \max_{c \in \{x,b\}} g\left(c, \max_{c' \in \varphi_r(\{x,b\})} w(c'; x, y, y', \delta)\right)$$

and $\mathcal{C}(\{x,b\})$ *coincides with* $\mathcal{C}_{PS}(\{x,b\}) = \arg\max_{c \in \{x,b\}} g\left(c, \max_{c' \in \varphi_r(\{x,b\})} w(c'; x, y, y', \delta)\right)$, *where* $g(c, R) = u(c) - \max\{R - w(c; x, y, y', \delta), 0\} + \beta \max\{w(c; x, y, y', \delta) - R, 0\}$ *and* $\varphi_r(A) = \arg\max_A r$.

Proof. Note $V_{PS}(\{x, y(1-\delta)c\}) > V_{PS}(\{x\})$ and $V_{PS}(\{x, y'(1-\delta)c\}) > V_{PS}(\{x\})$ by Axioms 3 and 6 (recall Lemma S25). Consider the following exhaustive cases.

Case 1. Suppose $b \in \mathcal{P}(x)$, so that $\mathcal{C}(\{x,b\}) = \{b\}$ by definition. Note $V_{PS}(\{x,b\}) > V_{PS}(\{x\})$ by Lemma S25. By Lemma 7(i)(ii) and $b \in \mathcal{P}(x)$, we have $w(b; x, y, y', \delta) - w(x; x, y, y', \delta) = \frac{1}{\beta} w_P(b; x, y, \delta) = \frac{1}{\beta}[V_{PS}(\{x,b\}) - V_{PS}(\{b\})] > 0$. Therefore,

$$\begin{aligned} g\left(b, \max_{c' \in \varphi_r(\{x,b\})} w(c'; x, y, y', \delta)\right) &= u(b) + \beta[w(b; x, y, y', \delta) - w(x; x, y, y', \delta)] \\ &= V_{PS}(\{x,b\}) \\ &> V_{PS}(\{x\}) = g\left(x, \max_{c' \in \varphi_r(\{x,b\})} w(c'; x, y, y', \delta)\right). \end{aligned}$$

Thus, the conclusion holds.

Case 2. Suppose $b \in \mathcal{S}(x)$, so that $\mathcal{C}(\{x,b\}) = \{b\}$. By Lemma 8(i)(ii) and $b \in \mathcal{S}(x)$, $w(b; x, y, y', \delta) - w(x; x, y, y', \delta) = w_S(b; x, y', \delta) = V_{PS}(\{x,b\}) - V_{PS}(\{b\}) < 0$. Therefore,

$$\begin{aligned} g\left(b, \max_{c' \in \varphi_r(\{x,b\})} w(c'; x, y, y', \delta)\right) &= u(b) + w(b; x, y, y', \delta) - w(x; x, y, y', \delta) \\ &= V_{PS}(\{x,b\}) \\ &> V_{PS}(\{x\}) = g\left(x, \max_{c' \in \varphi_r(\{x,b\})} w(c'; x, y, y', \delta)\right). \end{aligned}$$

Case 3. Suppose $b \in \mathcal{N}_1,(x)$, so that $\mathcal{C}(\{x,b\}) = \{b\}$. Letting $A' = \{x, y'\}(1-\delta)\{x,b\}$, Lemma 6(ii) and $\{x,b\} \sim \{b\}$ imply $w(b;x,y,y',\delta) - w(x;x,y,y',\delta) = w_S(b;x,y',\delta) = \frac{1}{\delta}\left[V_{PS}(A') - (1-\delta)V_{PS}(\{x,y'\}) - \delta V_{PS}(\{b\})\right] = 0$. Therefore,

$$g\left(b, \max_{c' \in \varphi_r(\{x,b\})} w(c';x,y,y',\delta)\right) = V_{PS}(\{b\}) = V_{PS}(\{x,b\})$$
$$> V_{PS}(\{x\}) = g\left(x, \max_{c' \in \varphi_r(\{x,b\})} w(c';x,y,y',\delta)\right).$$

Case 4. Suppose $b \in \mathcal{N}_2(x)$, so that $x \in \mathcal{C}(\{x,b\})$. By Lemma S25, $V_{PS}(\{x,b\}) = V_{PS}(\{x\})$. Consider first the case where $w(b;x,y,y',\delta) > 0$. By Axiom 4(iii), $\frac{1}{\delta}V_{PS}(\{x, y(1-\delta)b\}) - \frac{1-\delta}{\delta}V_{PS}(\{x,y\}) - V_{PS}(\{b\}) = w_P(b;x,y,\delta) > 0 = \frac{1}{\delta}V_{PS}(\{x,y\}(1-\delta)\{b\}) - \frac{1-\delta}{\delta}V_{PS}(\{x,y\}) - V_{PS}(\{b\})$, so $V_{PS}(\{x,y(1-\delta)b\}) > V_{PS}(\{x,y\}(1-\delta)\{b\})$, which together with Axiom 5(ii-a) implies $\mathcal{C}(A) = \{y\}(1-\delta)\mathcal{C}(\{x,b\})$ where $A = \{x,y\}(1-\delta)\{x,b\}$.[32] Then, by Axiom 6(i), $\{x, y(1-\delta)x\} \sim A \succeq \{x, y(1-\delta)b\}$ and the latter relation is strict if and only if $\mathcal{C}(\{x,b\}) = \{x\}$. Therefore,

$$g\left(b, \max_{c' \in \varphi_r(\{x,b\})} w(c';x,y,y',\delta)\right) = u(b) + \beta \cdot \frac{1}{\beta} w_P(b;x,y,\delta)$$
$$\leq \frac{1}{\delta}V_{PS}(\{x, y(1-\delta)x\}) - \frac{1-\delta}{\delta}V_{PS}(\{x,y\})$$
$$= V_{PS}(\{x\}) = V_{PS}(\{x,b\}) = g\left(x, \max_{c' \in \varphi_r(\{x,b\})} w(c';x,y,y',\delta)\right)$$

where the inequality is strict if and only if $\mathcal{C}(\{x,b\}) = \{x\}$. Next, consider the case with $w(b;x,y,y',\delta) \leq 0$. By Axiom 4(iii), $\frac{1}{\delta}V_{PS}(\{x, y'(1-\delta)b\}) - \frac{1-\delta}{\delta}V_{PS}(\{x,y'\}) - V_{PS}(\{b\}) = w_S(b;x,y',\delta) \leq 0 = \frac{1}{\delta}V_{PS}(\{x,y'\}(1-\delta)\{b\}) - \frac{1-\delta}{\delta}V_{PS}(\{x,y'\}) - V_{PS}(\{b\})$, so we have $V_{PS}(\{x, y'(1-\delta)b\}) \leq V_{PS}(\{x,y'\}(1-\delta)\{b\})$, which together with Axiom 5(ii-b) implies $\mathcal{C}(A') = \{y'\}(1-\delta)\mathcal{C}(\{x,b\})$ where $A' = \{x,y'\}(1-\delta)\{x,b\}$. By Axiom 6(i), $\{x, y'(1-\delta)x\} \sim A' \succeq \{x, y'(1-\delta)b\}$ and the latter relation is strict if and only if $\mathcal{C}(\{x,b\}) = \{x\}$. Therefore,

$$g\left(b, \max_{c' \in \varphi_r(\{x,b\})} w(c';x,y,y',\delta)\right) = u(b) + w_S(b;x,y',\delta)$$
$$\leq \frac{1}{\delta}V_{PS}(\{x, y'(1-\delta)x\}) - \frac{1-\delta}{\delta}V_{PS}(\{x,y'\})$$
$$= V_{PS}(\{x\}) = V_{PS}(\{x,b\}) = g\left(x, \max_{c' \in \varphi_r(\{x,b\})} w(c';x,y,y',\delta)\right)$$

where the inequality is strict if and only if $\mathcal{C}(\{x,b\}) = \{x\}$.

Case 5. Suppose $b \in \mathcal{I}(x)$ and $V_{PS}(\{x,b\}) > V_{PS}(\{x\})$. Note $\mathcal{C}(\{x,b\}) = \{b\}$ by Axiom

[32] To apply Axiom 5(ii), recall $\mathcal{C}(\{x, y(1-\delta)c\}) = \{y(1-\delta)c\}$ for all $c \in \Delta$.

6, and $V_{PS}(\{x, y'(1-\delta)b\}) \geq V_{PS}(A')$ by Lemma 6(iii), where $A' = \{x, y'\}(1-\delta)\{x, b\}$. Now, by Lemma 2, $V_{PS}(\{b\}) \geq V_{PS}(\{x, b\})$. Consider first the case $V_{PS}(\{b\}) = V_{PS}(\{x, b\})$. Then, $V_{PS}(\{x, y'(1-\delta)b\}) \geq (1-\delta)V_{PS}(\{x, y'\}) + \delta V_{PS}(\{x, b\}) = (1-\delta)V_{PS}(\{x, y'\}) + \delta V_{PS}(\{b\})$, so $w(b; x, y, y', \delta) \geq 0 = w(x; x, y, y', \delta)$. Thus, noting $\varphi_r(\{x, b\}) = \{x, b\}$,

$$\begin{aligned} g\left(b, \max_{c' \in \varphi_r(\{x,b\})} w(c'; x, y, y', \delta)\right) &= V_{PS}(\{b\}) \\ &= V_{PS}(\{x, b\}) > V_{PS}(\{x\}) \geq g\left(x, \max_{c' \in \varphi_r(\{x,b\})} w(c'; x, y, y', \delta)\right). \end{aligned}$$

Next, consider the case $V_{PS}(b) > V_{PS}(\{x, b\})$. By Lemma 6(iii), $V_{PS}(\{x, y'(1-\delta)b\}) = V_{PS}(A')$. Therefore, $w(b; x, y, y', \delta) = \frac{1-\delta}{\delta} V_{PS}(\{x, y'\}) + V_{PS}(\{x, b\}) - \frac{1-\delta}{\delta} V_{PS}(\{x, y'\}) - V_{PS}(\{b\}) < 0 = w(x; x, y, y', \delta)$. Thus,

$$\begin{aligned} g\left(b, \max_{c' \in \varphi_r(\{x,b\})} w(c'; x, y, y', \delta)\right) &= u(b) + w(b; x, y, y', \delta) \\ &= V_{PS}(\{x, b\}) > V_{PS}(\{x\}) = g\left(x, \max_{c' \in \varphi_r(\{x,b\})} w(c'; x, y, y', \delta)\right). \end{aligned}$$

Case 6. Suppose $b \in \mathcal{I}(x)$ and $V_{PS}(\{b\}) > V_{PS}(\{x, b\}) = V_{PS}(\{x\})$. Note $x \in \mathcal{C}(\{x, b\})$.[33] Then, by Lemma 6(iv), $V_{PS}(\{x, y'\}(1-\delta)\{x, b\}) \geq V_{PS}(\{x, y'(1-\delta)b\})$. Therefore, we have $w(b; x, y, y', \delta) \leq \frac{1}{\delta} V_{PS}(\{x, y'\}(1-\delta)\{x, b\}) - \frac{1-\delta}{\delta} V_{PS}(\{x, y'\}) - V_{PS}(\{b\}) < 0 = w(x; x, y, y', \delta)$, so

$$\begin{aligned} g\left(b, \max_{c' \in \varphi_r(\{x,b\})} w(c'; x, y, y', \delta)\right) &= u(b) + w(b; x, y, y', \delta) \\ &\leq \frac{1}{\delta} V_{PS}(\{x, y'\}(1-\delta)\{x, b\}) - \frac{1-\delta}{\delta} V_{PS}(\{x, y'\}) \\ &= V_{PS}(\{x, b\}) = V_{PS}(\{x\}) = g\left(x, \max_{c' \in \varphi_r(\{x,b\})} w(c'; x, y, y', \delta)\right). \end{aligned}$$

By Lemma 6(iv) the inequality is strict if and only if $\mathcal{C}(\{x, b\}) = \{x\}$.

Case 7. Suppose $b \in \mathcal{I}(x)$ and $V_{PS}(\{b\}) = V_{PS}(\{x, b\}) = V_{PS}(\{x\})$. If $\mathcal{C}(\{x, b\}) = \{x\}$, then by Lemma 6(v-a), we have $w(b; x, y, y', \delta) < \frac{1}{\delta} V_{PS}(\{x, y'\}(1-\delta)\{x\}) - \frac{1-\delta}{\delta} V_{PS}(\{x, y'\}) - V_{PS}(\{b\}) = 0 = w(x; x, y, y', \delta)$. Therefore,

$$\begin{aligned} g\left(b, \max_{c' \in \varphi_r(\{x,b\})} w(c'; x, y, y', \delta)\right) &= V_{PS}(\{b\}) + w(b; x, y, y', \delta) \\ &< V_{PS}(\{x, b\}) = V_{PS}(\{x\}) = g\left(x, \max_{c' \in \varphi_r(\{x,b\})} w(c'; x, y, y', \delta)\right). \end{aligned}$$

[33] $\{b\} \succ \{x, b\}$ implies $x \succ_w b$. Then, Axiom 6 and $\{x\} \sim \{x, b\}$ imply $x \in \mathcal{C}(\{x, b\})$.

If $\mathcal{C}(\{x,b\}) = \{b\}$, then Lemma 6(v-a) implies $w(b;x,y,y',\delta) > \frac{1}{\delta}V_{PS}(\{x,y'\}(1-\delta)\{b\}) - \frac{1-\delta}{\delta}V_{PS}(\{x,y'\}) - V_{PS}(\{b\}) = 0 = w(x;x,y,y',\delta)$, so

$$\begin{aligned} g\left(b, \max_{c'\in\varphi_r(\{x,b\})} w(c';x,y,y',\delta)\right) &= V_{PS}(\{b\}) \\ &= V_{PS}(\{x,b\}) \\ &> V_{PS}(\{x\}) - w(b;x,y,y',\delta) = g\left(x, \max_{c'\in\varphi_r(\{x,b\})} w(c';x,y,y',\delta)\right). \end{aligned}$$

Finally, if $\mathcal{C}(\{x,b\}) = \{x,b\}$, then Lemma 6(v-b) implies $w(b;x,y,y',\delta) = \frac{1}{\delta}V_{PS}(\{x,y'\}(1-\delta)\{x\}) - \frac{1-\delta}{\delta}V_{PS}(\{x,y'\}) - V_{PS}(\{b\}) = 0$, so the desired representation holds.

Case 8. Suppose $b \in \mathcal{I}(x)$ and $V_{PS}(\{x\}) = V_{PS}(\{x,b\}) > V_{PS}(\{b\})$. Note that Axiom 6 implies $\mathcal{C}(\{x,b\}) = \{x\}$.[34] By Axiom 4, $\{x\} \sim \{x,b\}\,\delta\,\{x\} \succ \{b\}\,\delta\,\{x\}$, so $x \succ_w x(1-\delta)b$ by definition. Thus, Axioms 3(i) and 6(i) imply $\{x, x(1-\delta)b, y'(1-\delta)b\} \succeq \{x, y'(1-\delta)b\}$. Also, $\mathcal{C}(\{x, y'(1-\delta)b\}) = \{y'(1-\delta)b\}$ by construction, and Axiom 8 imply $x \notin \mathcal{C}(\{x, x(1-\delta)b, y'(1-\delta)b\})$. Therefore, by Axiom 6(ii), $\{x(1-\delta)b, y'(1-\delta)b\} \succ \{x, x(1-\delta)b, y'(1-\delta)b\}$. Combining these results, $\{x,y'\}(1-\delta)\{b\} \succ \{x, y'(1-\delta)b\}$. Then, $w(b;x,y,y',\delta) < \frac{1}{\delta}V_{PS}(\{x,y'\}(1-\delta)\{b\}) - \frac{1-\delta}{\delta}V_{PS}(\{x,y'\}) - V_{PS}(\{b\}) = 0 = w(x;x,y,y',\delta)$, so

$$\begin{aligned} g\left(b, \max_{c'\in\varphi_r(\{x,b\})} w(c';x,y,y',\delta)\right) &= V_{PS}(\{b\}) + w(b;x,y,y',\delta) \\ &< V_{PS}(\{x,b\}) \\ &= V_{PS}(\{x\}) = g\left(x, \max_{c'\in\varphi_r(\{x,b\})} w(c';x,y,y',\delta)\right). \quad \square \end{aligned}$$

Next, we prove that the representation holds for an arbitrary binary menu $\{a,b\}$.

Lemma 10. *Suppose Axioms 1-8 and nondegeneracy hold. Then there exist continuous and linear functions u, w and r such that V_{PS} is expressed as*

$$V_{PS}(\{a,b\}) = \max_{c\in\{a,b\}} g\left(c, \max_{c'\in\varphi_r(\{a,b\})} w(c')\right)$$

and $\mathcal{C}(\{a,b\})$ coincides with $\mathcal{C}_{PS}(\{a,b\}) = \arg\max_{c\in\{a,b\}} g\left(c, \max_{c'\in\varphi_r(\{a,b\})} w(c')\right)$, where $g(c,R) = u(c) - \max\{R - w(c), 0\} + \beta\max\{w(c) - R, 0\}$ and $\varphi_r(A) = \arg\max_A r$.

Proof. As above, let $u(a) = V_{PS}(\{a\})$ and let r be the linear function that represents $\succeq_r$. Take $x, y, y' \in \Delta$ such that $y \in \mathcal{P}(x)$ and $y' \in \mathcal{S}(x)$, which exist by nondegeneracy. Note that we can take some $\delta \in (0,1)$ such that $y(1-\delta)c \in \mathcal{P}(x)$ and $y'(1-\delta)c \in \mathcal{S}(x)$ for all

[34] If $b \succeq_w x$, use Axiom 6(i) and $\{x,b\} \succ \{b\}$. Otherwise, use Axiom 6(ii) and $\{x,b\} \succ \{b\}$.

$c \in \Delta$ (see Lemma S28(i)). Now, consider an arbitrary set $\{a, b\}$. First, consider the case $a, b \in \text{int}(\Delta)$. Without loss of generality, suppose $r(a) \geq r(b)$ and $V_{PS}(\{a, b\}) \geq V_{PS}(\{a\})$.[35] Because $a \in \text{int}(\Delta)$, there exist $\bar{a} \in \Delta$ and $\alpha \in (0, 1)$ such that $a = \bar{a}\alpha x$. Define $z = \bar{a}\alpha y$ and $z' = \bar{a}\alpha y'$. By Lemma S28(ii), we have $z \in \mathcal{P}(a)$ and $z' \in \mathcal{S}(a)$. Then, by Lemma S28(i) there exists $\delta' \in (0, 1)$ such that $z(1 - \delta')c \in \mathcal{P}(a)$ and $z'(1 - \delta')c \in \mathcal{S}(a)$ for all $c \in \Delta$. Then, by Lemma 9, we have $V_{PS}(\{a, b\}) = \max_{c \in \{a,b\}} g\left(c, \max_{c' \in \varphi_r(\{a,b\})} w(c'; a, z, z', \delta')\right)$ and $\mathcal{C}(\{a, b\}) = \arg\max_{c \in \{a,b\}} g\left(c, \max_{c' \in \varphi_r(\{a,b\})} w(c'; a, z, z', \delta')\right)$. Now, let $\delta^* = \min\{\delta, \delta'\}$. Then, by Lemmas 7 and 8, we have $w(\cdot; x, y, y', \delta) = w(\cdot; x, y, y', \delta^*)$, $w(\cdot; a, z, z', \delta') = w(\cdot; a, z, z', \delta^*)$ and $w(\cdot; a, z, z', \delta^*) = w(\cdot; x, y, y', \delta^*) + k$ for some constant k. Therefore, defining $w(\cdot) = w(\cdot; x, y, y', \delta)$ yields the conclusion.[36]

Next, suppose $a \in \Delta$ and $b \in \text{int}(\Delta)$. Because $a\alpha b \in \text{int}(\Delta)$ for $\alpha \in (0, 1)$,

$$\begin{aligned}
&V_{PS}(\{a\alpha b, b\}) = \max_{c \in \{a\alpha b, b\}} g\left(c, \max_{c' \in \varphi_r(\{a\alpha b, b\})} w(c')\right) \\
\Leftrightarrow\ &\alpha V_{PS}(\{a, b\}) + (1 - \alpha)V_{PS}(\{b\}) = \alpha \max_{c \in \{a,b\}} g\left(c, \max_{c' \in \varphi_r(\{a,b\})} w(c')\right) + (1 - \alpha)u(b).
\end{aligned}$$

where the right-hand side follows from linearity (cf. the proof of the necessity of Axiom 4(iii)). Letting $\alpha \to 1$ yields the conclusion for V_{PS}. The conclusion for $\mathcal{C}_{PS}$ is obtained analogously using Axiom 5. Proof for the general case with $a, b \in \Delta$ is now straightforward. □

We now extend the representation to any finite menus.

Lemma 11. *Suppose Axioms 1-8 and nondegeneracy hold. Then there exist continuous and linear functions u, w and r such that, over finite menus, V_{PS} is expressed as*

$$V_{PS}(A) = \max_{c \in A} g\left(c, \max_{c' \in \varphi_r(A)} w(c')\right)$$

and $\mathcal{C}(A)$ coincides with $\mathcal{C}_{PS}(A) = \arg\max_{c \in A} g\left(c, \max_{c' \in \varphi_r(A)} w(c')\right)$, where g and φ are defined in Lemma 10.

Proof. Take any finite set A. By Eq.(11), Lemma 10, and the property that $R > R'$ implies $g(c, R) < g(c, R')$ at each $c \in \Delta$, we have

$$\begin{aligned}
V_{PS}(A) &= \min_{b \in \varphi_r(A)} \max_{a \in A} V_{PS}(\{a, b\}) \\
&= \min_{b \in \varphi_r(A)} \max_{a \in A} \max_{c \in \{a,b\}} g\left(c, \max_{c' \in \varphi_r(\{a,b\})} w(c')\right)
\end{aligned}$$

[35]If $r(a) > r(b)$, then Lemma 2(i) implies $V_{PS}(\{a, b\}) \geq V_{PS}(\{a\})$. If $r(a) = r(b)$, then Lemma 2(ii) implies $V_{PS}(\{a, b\}) \geq V_{PS}(\{a\})$ or $V_{PS}(\{a, b\}) \geq V_{PS}(\{b\})$, so the assumption is without loss of generality.

[36]Note that k is differenced out in g, and that (x, y, y', δ) does not depend on (a, b).

$$
\begin{aligned}
&= \min_{b\in\varphi_r(A)} \max_{a\in A} \max\left\{ g\left(a, \max_{c'\in\varphi_r(\{a,b\})} w(c')\right), g\left(b, \max_{c'\in\varphi_r(\{a,b\})} w(c')\right)\right\} \\
&= \min_{b\in\varphi_r(A)} \max_{a\in A} g\left(a, \max_{c'\in\varphi_r(\{a,b\})} w(c')\right) \\
&= \max_{a\in A} g\left(a, \max_{c'\in\varphi_r(A)} w(c')\right).
\end{aligned}
$$

where the fourth equality holds because $b \in \varphi_r(A)$ implies $\max_{a\in A} g\left(b, \max_{c'\in\varphi_r(\{a,b\})} w(c')\right) \leq g(b, w(b)) \leq \max_{a\in A} g\left(a, \max_{c'\in\varphi_r(\{a,b\})} w(c')\right)$.

To prove the result on $\mathcal{C}$, we first introduce some lemmas. Lemmas 12 and 14 establish Lemma 15, which then establishes Lemma 11.

Lemma 12. *Suppose Axioms 1-8 hold. If $r(a) = r(b)$ and $w(a) > w(b)$, then $a \succ_w b$.*

Proof. Consider the following exhaustive cases. (Recall that, by Lemma 10, the representation holds for binary menus.)

Case 1. Suppose $u(a) + w(a) \leq u(b) + w(b)$. Then $V_{PS}(\{a, b\}) = u(b) + w(b) - w(a) < V_{PS}(\{b\})$, so $a \succ_w b$.

Case 2. Suppose $u(a) + w(a) > u(b) + w(b)$ and $u(a) < u(b)$. Then $V_{PS}(\{a, b\}) = u(a) < V_{PS}(\{b\})$, so $a \succ_w b$.

Case 3. Suppose $u(a) + w(a) > u(b) + w(b)$ and $u(a) = u(b)$. Then $V_{PS}(\{a, b\}) = u(a) = V_{PS}(\{b\})$ and $\mathcal{C}(\{a, b\}) = \{a\}$, so $a \succ_w b$.

Case 4. Suppose $u(a) + w(a) > u(b) + w(b)$ and $u(a) > u(b)$. Then $V_{PS}(\{a\}) = V_{PS}(\{a, b\}) > V_{PS}(\{b\})$. This, together with $a \sim_r b$, implies $a \succ_w b$. □

Lemma 13. *Suppose Axioms 1-8 hold. If $y \in \mathcal{P}(x)$, then $r(x) > r(y)$ and $w(x) < w(y)$.*

Proof. By assumption, we have $r(x) > r(y)$ and $g(y, w(x)) = V_{PS}(\{x, y\}) > u(y) = g(y, w(y))$, the latter of which implies $w(x) < w(y)$. □

Lemma 14. *Suppose Axioms 1-8 and nondegeneracy hold. Then, for any $a, b \in \Delta$, $r(a) = r(b)$ and $a \neq b$ imply $w(a) \neq w(b)$.*

Proof. By Lemma 13, $r(x) > r(y)$ and $w(x) < w(y)$ for some $x, y \in \Delta$. If $r(a) = r(b)$ and $w(a) = w(b)$ for some $a \neq b$, then the indifference curves for r and w are parallel straight lines, which contradicts $r(x) > r(y)$ and $w(x) < w(y)$. □

Lemma 15. *Suppose Axioms 1-8 and nondegeneracy hold. Then, for any $A \in \mathcal{A}$ and any $b \in \arg\max_{c'\in\varphi_r(A)} w(c')$, we have $b \succeq_r a$ and $b \succeq_w a$ for all $a \in A$.*

Proof. Take any $a \in A \setminus \{b\}$. Because $b \in \varphi_r(A)$, $b \succeq_r a$. If $a \notin \varphi_r(A)$, then $b \succ_r a$, so Axiom 3(i) implies $b \succeq_w a$. If $a \in \varphi_r(A)$, then $r(a) = r(b)$ and $w(a) \leq w(b)$ by definition. By Lemma 14, $w(a) < w(b)$. Thus, $b \succ_w a$ by Lemma 12. □

Proof of Lemma 11, Continued. We show $\mathcal{C}(A) = \arg\max_{c\in A} g\left(c, \max_{c'\in\varphi_r(A)} w(c')\right)$ by showing each inclusion.

Step 1. Take $a \in \arg\max_{c\in A} g\left(c, \max_{c'\in\varphi_r(A)} w(c')\right)$, $b \in \arg\max_{c'\in\varphi_r(A)} w(c')$, and $d \in \mathcal{C}(A)$. By the first half of Lemma 11 and $\max_{c'\in\varphi_r(A)} w(c') = \max_{c'\in\varphi_r(\{a,b,d\})} w(c') = w(b)$, we have $\{a,b,d\} \sim A$. Also, Lemma 15 implies $b \succeq_r a'$ and $b \succeq_w a'$ for all $a' \in A$. Thus, by Axiom 8, $d \in \mathcal{C}(\{a,b,d\})$. Also, by the representation and Lemma 10, we have $\{a,b\} \sim \{a,b,d\}$ and $a \in \mathcal{C}(\{a,b\})$. By Axiom 6(i), we have $a \in \mathcal{C}(\{a,b,d\})$ or $b \in \mathcal{C}(\{a,b,d\})$, so Axiom 8 implies $a \in \mathcal{C}(\{a,b,d\})$. Thus, again by Axiom 8, $a \in \mathcal{C}(A)$.

Step 2. Suppose $d \in \mathcal{C}(A)$. Take $a \in \arg\max_{c\in A} g\left(c, \max_{c'\in\varphi_r(A)} w(c')\right)$ and $b \in \arg\max_{c'\in\varphi_r(A)} w(c')$. By Lemma 15, we have $b \succeq_r a'$ and $b \succeq_w a'$ for all $a' \in A$. Also, by *Step 1*, we have $a \in \mathcal{C}(\{a,b,d\})$. Thus, Axiom 8 implies $d \in \mathcal{C}(\{a,b,d\})$. By Axiom 6(i), $\{b,d\} \sim \{a,b,d\}$. By the representation and the definition of a and b, we have $\max_{c\in\{b,d\}} g(c, w(b)) = \max_{c\in\{a,b,d\}} g(c, w(b)) = \max_{c\in A} g(c, w(b))$. By Axiom 8, we have $d \in \mathcal{C}(\{b,d\})$, so Lemma 10 implies $d \in \arg\max_{c\in\{b,d\}} g(c, w(b))$. Thus $d \in \arg\max_{c\in A} g\left(c, \max_{c'\in\varphi_r(A)} w(c')\right)$. □

To complete the proof of the sufficiency part of Theorem 1, we use the following result.

Lemma 16. *Suppose Axioms 1-8 hold. Suppose that for any $A \in \mathcal{A}$, there exists a finite subset A' of A such that (i) $\max_A r = \max_{A'} r$ and (ii) for any finite A'' such that $A' \subset A'' \subset A$, $A'' \sim A'$. Then we have $A \sim A'$.*

Proof. Note that there exists a sequence of finite subsets $\{A_n\}_{n=1}^{\infty}$ of A such that $d_H(A_n, A) \to 0$ as $n \to \infty$ by Lemma 0 in GP. By (ii), $A' \sim A_n \cup A'$ for all n, so $A' \succsim A \cup A' = A$ by Axiom 2(i). To show the opposite relation, note that since A is compact, for every $\epsilon > 0$, there are finite $x_1, \cdots, x_n \in A$ such that $A \subset \cup_{i=1}^{n} N(x_i, \epsilon)$. If $(A' \cup \overline{N(x_i,\epsilon)}) \cap A \succ A$ for all $i = 1, \cdots, n$, then iteratively applying Lemma 2(ii) yields $A = \cup_{i=1}^{n}\{(A' \cup \overline{N(x_i,\epsilon)}) \cap A\} \succ A$, a contradiction. Therefore, $A \succsim (A' \cup \overline{N(x',\epsilon)}) \cap A$ for some $x' \in A$. Thus, we can take a sequence $\{x_n\}_{n=1}^{\infty}$ in A such that $A \succsim (A' \cup \overline{N(x_n, \frac{1}{n})}) \cap A$ for all $n = 1, 2, \cdots$. Since A is compact, there exists a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ such that $x_{n_k} \to x^* \in A$. Then letting $k \to \infty$ yields $A \succsim A' \cup \{x^*\} \sim A'$ by Axiom 2(i). □

Proof of Theorem 1 (Sufficiency), Continued.

Take any closed set $A \in \mathcal{A}$, and take $a^* \in \arg\max_{c\in A} g\left(c, \max_{c'\in\varphi_r(A)} w(c')\right)$ and $b^* \in \arg\max_{c'\in\varphi_r(A)} w(c')$. By construction, $\max_{\{a^*,b^*\}} r = \max_A r$. By Lemma 11, $\{a^*, b^*\} \sim A''$ for any finite A'' such that $\{a^*, b^*\} \subset A'' \subset A$. Therefore, by Lemma 16, we have $A \sim \{a^*, b^*\}$.

Thus, defining

$$V_{PS}(A) \equiv V_{PS}(\{a^*, b^*\}) = \max_{c \in \{a^*, b^*\}} g\left(c, \max_{c' \in \varphi_r(\{a^*, b^*\})} w(c')\right) = \max_{c \in A} g\left(c, \max_{c' \in \varphi_r(A)} w(c')\right),$$

V_{PS} represents $\succeq$ on $\mathcal{A}$. Also, following the argument for the proof of Lemma 11, we obtain $\mathcal{C}(A) = \mathcal{C}_{PS}(A) = \arg\max_{c \in A} g\left(c, \max_{c' \in \varphi_r(A)} w(c')\right)$.

Finally, by construction, $\beta = \frac{1}{\alpha} - 1$ if the DM is α-sensitive to shame. Thus, the DM is shame-averse ($\alpha > \frac{1}{2}$) if and only if $\beta < 1$, shame-neutral ($\alpha = \frac{1}{2}$) if and only if $\beta = 1$ and shame-loving ($\alpha < \frac{1}{2}$) if and only if $\beta > 1$. $\square$

A.2 Proof of Theorem 2

We first introduce two lemmas. Lemma 17 is used to prove Lemma 18, which in turn helps us establish $r(a) > r(b) \Leftrightarrow a \succ_r b$.

Lemma 17. *Suppose the data are generated by a weakly nondegenerate PS preference. Then, the following sets are nonempty for all $a \in \text{int}(\Delta)$:*

$$P_1(a) = \{c \in \Delta : r(a) > r(c) \text{ and } g(a, w(a)) < g(c, w(a))\} \quad (12)$$

$$P_2(a) = \{c \in \Delta : r(a) > r(c) \text{ and } w(a) < w(c)\} \quad (13)$$

Proof. Take $\bar{a}, \bar{b} \in \Delta$ such that $\bar{a} \succ^* \bar{b}$. Then, there exist $A \ni \bar{b}$ and $c \neq \bar{a}$ such that $A \cup \{\bar{a}\} \succ A$, $\bar{a} \notin \mathcal{C}(A \cup \{\bar{a}\})$, and $c \in \mathcal{C}(A \cup \{\bar{a}\})$. The first two conditions imply $\varphi_r(A \cup \{\bar{a}\}) = \{\bar{a}\}$ and $w(\bar{a}) = \max_{y \in \varphi_r(A \cup \{\bar{a}\})} w(y) < \max_{y \in \varphi_r(A)} w(y)$ (see Lemma S30). Thus, there is $d \in A$ such that $w(d) = \max_{y \in \varphi_r(A)} w(y) > w(\bar{a})$ and $r(d) < r(\bar{a})$, implying $d \in P_2(\bar{a})$. We additionally have $c \in \mathcal{C}(A \cup \{\bar{a}\})$, so $g(c, w(\bar{a})) > g(\bar{a}, w(\bar{a}))$, hence $c \in P_1(\bar{a})$.

Now, take any $a \in \text{int}(\Delta)$. There exist $\alpha \in (0, 1)$ and $e \in \Delta$ such that $a = \bar{a}\alpha e$. By the linearity of u, w, and r, we have $r(a) > \max\{r(c\alpha e), r(d\alpha e)\}$, $g(c\alpha e, w(a)) > g(a, w(a))$, and $w(a) < w(d\alpha e)$. Therefore, $c\alpha e \in P_1(a)$ and $d\alpha e \in P_2(a)$. $\square$

$c \in P_1(a)$ is an alternative which is below a in the descriptive norm ranking r but which is a choice preferred to a. $d \in P_2(a)$ is an alternatives which is below a in the descriptive norm ranking but above a in the prescriptive norm ranking w. Therefore, if $d \in A \cap P_2(a)$ sets the reference point at a menu A and a sets the reference point at $A \cup \{a\}$, then the latter reference point is lower than the former. Moreover, if $c \in A \cap P_1(a)$, then a is not chosen from $A \cup \{a\}$. Such c, d are key to establishing $r(a) > r(b) \Rightarrow a \succ^* b$ for "generic" cases. Lemma 18 formalizes the idea. See also the graphical illustration in Figure S1 and discussions in Appendix S.D.

Lemma 18. *Suppose the data are generated by a weakly nondegenerate PS preference. Then, for any $a, b \in \Delta$, $a \succ^* b$ implies $r(a) > r(b)$. Moreover, if $a \in \text{int}(\Delta)$, then $r(a) > r(b)$ implies $a \succ^* b$.*

Proof. Suppose $a \succ^* b$, so that we have $A \ni b$ such that $A \cup \{a\} \succ A$ and $a \notin \mathcal{C}(A \cup \{a\})$. Then, we must have $\varphi_r(A \cup \{a\}) = \{a\}$ (see Lemma S30), hence $r(a) > r(b)$. Next, suppose $a \in \text{int}(\Delta)$ and $r(a) > r(b)$. By Lemma 17, there exist $c \in P_1(a)$ and $d \in P_2(a)$. Note that $d \in P_2(a)$ implies $d\alpha a \in P_2(a)$ for all $\alpha \in (0, 1)$. Therefore, we can assume without loss of generality that $r(a) > r(d) > \max\{r(b), r(c)\}$. Then, defining $G(A, R) = \max_{c \in A} g(c, R)$,

$$V_{PS}(\{a,b,c,d\}) = G\left(\{a,b,c,d\}, \max_{y \in \varphi_r(\{a,b,c,d\})} w(y)\right) = G(\{b,c,d\}, w(a))$$
$$> G(\{b,c,d\}, w(d)) = V_{PS}(\{b,c,d\})$$

where the inequality follows from $G(A, \cdot)$ being strictly decreasing, and the second equality follows from $g(a, w(a)) < g(c, w(a))$, which also implies $a \notin \mathcal{C}(\{a,b,c,d\})$. Thus, $a \succ^* b$. $\square$

Proof of Theorem 2, Continued.
(i) Suppose first $r(a) > r(b)$. Take some $c \in \text{int}(\Delta)$ such that $r(a) > r(c) > r(b)$.[37] By Lemma 18, $c \succ^* b$. Also, if $c \succ^* a$, then $r(c) > r(a)$ by Lemma 18, a contradiction. Therefore, $c \not\succ^* a$, hence $a \succ_r b$. Next, suppose $a \succ_r b$, so that we have $c \not\succ^* a$ and $c \succ^* b$ for some $c \in \text{int}(\Delta)$. By Lemma 18, $r(c) > r(b)$. Also, if $r(c) > r(a)$, then $c \succ^* a$ by Lemma 18, a contradiction. Thus, $r(a) \geq r(c) > r(b)$.

(ii) By inspection, $a \succ_w b$ implies $w(a) > w(b)$ (see Lemma S31). The converse can be established by following the proof of Lemma 12.[38] $\square$

A.3 Other Proofs

A.3.1 Proof of Proposition 1

It is easy to show that (ii) implies (i), so we only prove that (i) implies (ii). Let V_{PS} and V'_{PS} denote the PS representations of $\succeq$ using (u, w, r, β) and (u', w', r', β'), respectively. Since V_{PS} is unique up to positive affine transformation, $u'(x) = V'_{PS}(\{x\}) = \theta V_{PS}(\{x\}) + \gamma_u = \theta u(x) + \gamma_u$ for some $\theta > 0$ and $\gamma_u \in \mathbb{R}$. Now, by nondegeneracy and Lemma S28, there exist $x, y \in \Delta$ and $\delta \in (0, 1)$ such that $y(1-\delta)z \in \mathcal{P}(x)$ for all $z \in \Delta$. By the representation, we

[37] Take $c' = a\alpha b$ for some $\alpha \in (0, 1)$. If $c' \in \text{int}(\Delta)$, let $c = c'$. Otherwise, take some $d \in \text{int}(\Delta)$ and let $c = c'\beta d$ where $\beta < 1$ is sufficiently close to 1.

[38] Note that the proof depends on the representation and not on any axiom.

have

$$\begin{aligned}
\frac{1}{\delta}V_{PS}(\{x,y(1-\delta)z\}) - \frac{1-\delta}{\delta}V_{PS}(\{x,y\}) &= \frac{1}{\delta}\{u(y(1-\delta)z) + \beta[w(y(1-\delta)z) - w(x)]\} \\
&\quad - \frac{1-\delta}{\delta}\{u(y) + \beta[w(y) - w(x)]\} \\
&= u(z) + \beta w(z) - \beta w(x).
\end{aligned}$$

Therefore, for any $z \in \Delta$,

$$\begin{aligned}
u'(z) + \beta[w'(z) - w'(x)] &= \frac{1}{\delta}V'_{PS}(\{x,y(1-\delta)z\}) - \frac{1-\delta}{\delta}V'_{PS}(\{x,y\}) \\
&= \frac{1}{\delta}\left[\theta V_{PS}(\{x,y(1-\delta)z\}) + \gamma_u\right] - \frac{1-\delta}{\delta}\left[\theta V_{PS}(\{x,y\}) + \gamma_u\right] \\
&= \theta\left[\frac{1}{\delta}V_{PS}(\{x,y(1-\delta)z\}) - \frac{1-\delta}{\delta}V_{PS}(\{x,y\})\right] + \gamma_u \\
&= \theta\{u(z) + \beta[w(z) - w(x)]\} + \gamma_u.
\end{aligned}$$

Since $u'(z) = \theta u(z) + \gamma_u$, we have $w'(z) = \theta w(z) - \theta w(x) + w'(x) \equiv \theta w(z) + \gamma_w$. Next, by Theorem 2, both r and r' represent $\succeq_r$, so Lemma 1 implies $r' = \theta_r r + \gamma_r$ for some $\theta_r > 0$ and $\gamma_r \in \mathbb{R}$. Finally, by Axiom 7, there exists a unique $\alpha \in (0,1)$ such that, for any $a,b,c,d \in \Delta$ with $c \in \mathcal{P}(a) \cap \mathcal{P}(b)$ and $d \in \mathcal{S}(a) \cap \mathcal{S}(b)$, we have $\{a,c\}\,\alpha\{e^{b,d}\} \sim \{b,c\}\,\alpha\{e^{a,d}\}$, where $\{e^{b,d}\} \sim \{b,d\}$ and $\{e^{a,d}\} \sim \{a,d\}$. By the representation, we have

$$\begin{aligned}
V_{PS}\left(\{a,c\}\,\alpha\left\{e^{b,d}\right\}\right) &= \alpha V_{PS}(\{a,c\}) + (1-\alpha)V_{PS}\left(\left\{e^{b,d}\right\}\right) \\
&= \alpha V_{PS}(\{a,c\}) + (1-\alpha)V_{PS}(\{b,d\}) \\
&= \alpha\{u(c) + \beta[w(c) - w(a)]\} + (1-\alpha)\{u(d) + w(d) - w(b)\}
\end{aligned}$$

and similarly $V_{PS}\left(\{b,c\}\,\alpha\left\{e^{a,d}\right\}\right) = \alpha\{u(c)+\beta[w(c)-w(b)]\} + (1-\alpha)\{u(d)+w(d)-w(a)\}$. Since these values are equal for any $a,b \in \Delta$, we must have $1-\alpha-\alpha\beta = 0$, i.e., $\alpha = \frac{1}{1+\beta}$. Because the two representations must represent the same $(\succeq, \mathcal{C})$, we have $\beta = \beta'$.

A.3.2 Proof of Proposition 2.

Equivalence of (i) and (ii) follows from $\alpha_i = \frac{1}{1+\beta_i}$ (see the proof of Proposition 1). To show the equivalence of (ii) and (iii), note that if DM i is α_i-sensitive to shame, then the representation implies

$$\alpha_i = \frac{V^i_{PS}(\{a,d\}) - V^i_{PS}(\{b,d\})}{V^i_{PS}(\{a,c\}) - V^i_{PS}(\{b,c\}) + V^i_{PS}(\{a,d\}) - V^i_{PS}(\{b,d\})}.$$

Then

$$
\begin{aligned}
&\{b,c\}\,\alpha\left\{e_i^{a,d}\right\} \succeq_i \{a,c\}\,\alpha\left\{e_i^{b,d}\right\} \\
\Leftrightarrow\ &\alpha V_{PS}^i\left(\{a,c\}\right)+(1-\alpha)V_{PS}^i\left(\{b,d\}\right)\leq \alpha V_{PS}^i\left(\{b,c\}\right)+(1-\alpha)V_{PS}^i\left(\{a,d\}\right) \\
\Leftrightarrow\ &\alpha\leq\alpha_i.
\end{aligned}
$$

Therefore, $\alpha_1 > (\geq)\alpha_2$ if and only if $\{b,c\}\,\alpha\left\{e_2^{a,d}\right\} \succeq_2 \{a,c\}\,\alpha\left\{e_2^{b,d}\right\}$ implies $\{b,c\}\,\alpha\left\{e_1^{a,d}\right\} \succ_1$ $(\succeq_1)\,\{a,c\}\,\alpha\left\{e_1^{b,d}\right\}$. $\square$

Supplemental Appendix for "Norms and Emotions"

S.B Other Predictions from the Simple Model

In Section 2.1, we present a simple PS model of prosocial behavior and show its insightfulness. In this section, we discuss additional implications of the model, such as conformative versus pride-seeking behavior, and boomerang effects.

Recall that $x = 1$ denotes engaging in prosocial behavior, and $x = 0$ denotes non-engagement. Private and social payoffs are in conflict: $u(0) = \bar{u} > 0 = u(1)$ and $w(0) = 0 < \bar{w} = w(1)$, with $\beta\bar{w} < \bar{u} < \bar{w}$. At menu $\{0, 1\}$, the DM chooses an action by comparing the ex post utility of action 0, $U(0; \{0, 1\}) = \bar{u} - w(\varphi_r(\{0, 1\}))$, with that of action 1, $U(1; \{0, 1\}) = \beta [\bar{w} - w(\varphi_r(\{0, 1\}))]$. We compare decisions in the benchmark case $r(0) > r(1)$ (prosocial behavior is perceived as uncommon) with those in the post-intervention case $r'(0) < r'(1)$ (prosocial behavior is perceived as common).

Conformity and pride seeking. The DM conforms to the reference alternative both in the benchmark case ($x = \varphi_r(\{0, 1\}) = 0$) and post-intervention case ($x = \varphi_{r'}(\{0, 1\}) = 1$). By contrast, if we modify the benchmark assumption so that $\beta\bar{w} > \bar{u}$, then the DM engages in prosocial behavior under both scenarios. In the modified benchmark, the DM deviates from the reference to seek pride. Thus, our model can produce conformative or pride-seeking behavior depending on β. Typical empirical findings suggest β is small (see footnote 14); still, in some contexts, individuals may seek to perform better than a natural reference point.

Boomerang effect. In a field experiment on electricity consumption, Schultz et al. (2007) find that providing descriptive information on neighbors' electricity usage led to desired electricity saving by high-consuming households but increased consumption by low-consuming households. To explain the latter result (which Schultz et al. (2007) call a "boomerang effect") without complicating the model, let $x = 0$ and $x = 1$ denote high consumption and low consumption of electricity, respectively, and suppose that the low-consuming households originally perceive norms (w, r') but the intervention updates the perceptions to (w, r). By the analysis in Section 2.1, the low-consuming households originally choose $x = 1$ but the intervention causes them to switch to $x = 0$. Thus, our model can explain the boomerang effect by a shift of the perceived descriptive norm toward higher consumption.

The purpose of the above example is to illustrate the importance of considering the perceived norms of individuals when introducing a policy, rather than develop a more thorough model. For example, the reduction in the electricity consumption by high-consuming households can be explained by the opposite shift in the perceived descriptive norm. Instead of developing a model which accommodates both types of households (possibly requiring more than two options), we note that even the direction of a policy effect, as well as its magnitude, crucially depends on what norms the households perceive prior to the intervention.

S.C Supplemental Proofs

S.C.1 Supplemental Proofs for Theorem 1 (Sufficiency Part)

S.C.1.1 Supplemental Results for Lemma 1

Lemma S19. *Suppose Axioms 4 and 5 hold. Then, for any $a, b, c \in \Delta$ and $\alpha \in (0,1)$, $a \succ^* b$ implies $a\alpha c \succ^* b\alpha c$.*

Proof. By definition, we have $A \cup \{a\} \succ A$ and $a \notin \mathcal{C}(A \cup \{a\})$ for some $A \ni b$. Then, mixing each term with $\{c\}$ with mixing rate α yields the result. □

Lemma S20. *Suppose Axioms 1-3 hold. Then, the following statements hold.*

(i) If $a \succ^* b$, then for any $c \in \Delta$, there exists $\alpha^* \in (0,1)$ such that $a\alpha c \succ^* b$ for all $\alpha \in (\alpha^*, 1)$.

(ii) If $b \succ^* c$, then for any $a \in \Delta$, there exists $\beta^* \in (0,1)$ such that $b \succ^* c\beta a$ for all $\beta \in (\beta^*, 1)$.

Proof. (i) By $a \succ^* b$, we have $A \cup \{a\} \succ A$ and $a \notin \mathcal{C}(A \cup \{a\})$ for some $A \ni b$. By Axiom 2(i), there exists $\alpha_1 \in (0,1)$ such that $A \cup \{a\alpha c\} \succ A$ for all $\alpha \in (\alpha_1, 1)$. By Axiom 3(iii-b), we have $\alpha_2 \in (0,1)$ such that $a\alpha c \notin \mathcal{C}(A \cup \{a\alpha c\})$ for all $\alpha \in (\alpha_2, 1)$. Thus, $\alpha^* \equiv \max\{\alpha_1, \alpha_2\}$ has the desired property.

(ii) By $b \succ^* c$, we have $B \cup \{b\} \succ B$ and $b \notin \mathcal{C}(B \cup \{b\})$ for some $B \ni c$. By Axioms 2(i) and 3(iii), there exist $\beta_1, \beta_2 \in (0,1)$ such that $[B\beta\{a\}] \cup \{b\} \succ B\beta\{a\}$ for all $\beta \in (\beta_1, 1)$ and $b \notin \mathcal{C}([B\beta\{a\}] \cup \{b\})$ for all $\beta \in (\beta_2, 1)$.[39] Thus, $b \succ^* c\beta a$ for all $\beta > \beta^* \equiv \max\{\beta_1, \beta_2\}$. □

Lemma S21. *Suppose Axioms 1-5 hold. If $a \succ^* b$ holds, then there exists $c \in \text{int}(\Delta)$ such that $c \not\succ^* a$ and $c \succ^* b$.*

Proof. Suppose $a \succ^* b$. By Lemma S19, we have $a \succ^* a.5b \succ^* b$. If $a.5b \in \text{int}(\Delta)$, then Axiom 3(i) implies $c = a.5b$ has the desired property. Otherwise, take any $\alpha \in (0,1)$ and $d \in \text{int}(\Delta)$, and let $c = (a.5b)\alpha d \in \text{int}(\Delta)$. Then, by Lemma S20, we have $a \succ^* c \succ^* b$ for α sufficiently close to one. □

[39] To show that the former property holds for all sufficiently large $\beta < 1$, note first that Axioms 2(i) and 3(iii-a) ensure $B \cup \{b\} \succ \tilde{B} \succ B$ where $\tilde{B} = \tilde{B}(\gamma) = [B \cup \{b\}]\gamma B$ for some $\gamma \in (0,1)$. (Otherwise, $\Gamma^L = \{\gamma \in [0,1] : B \succeq \tilde{B}(\gamma)\}$ and $\Gamma^U = \{\gamma \in [0,1] : \tilde{B}(\gamma) \succeq B \cup \{b\}\}$ are nonempty closed sets such that $\Gamma^L \cup \Gamma^U = [0,1]$, so $\tilde{B}(\gamma) \succ \tilde{B}(\gamma)$ for $\gamma \in \Gamma^L \cap \Gamma^U$, a contradiction.) Then, by Axioms 2(i) and 3(iii-a), for all sufficiently large $\beta < 1$, we must have $[B\beta\{a\}] \cup \{b\} \succ \tilde{B} \succ B\beta\{a\}$.

Lemma S22. *If Axioms 1-5 hold, then $\succ_r$ is transitive.*[40]

Proof. Suppose $a \succ_r b \succ_r c$. By Lemma S21, we have some $d \in \text{int}(\Delta)$ such that $d \not\succ^* a$ and $d \succ^* b$. If $d \not\succ^* c$, then $c \succ_r b$, contradicting Axiom 3(i). Therefore, $d \succ^* c$, hence $a \succ_r c$. $\square$

S.C.1.2 Supplemental Results for Lemma 2

Below, we impose Axioms 1-5, so $\succeq_r$ admits a linear representation r (Lemma 1).

Lemma S23. *Suppose Axioms 1-5 hold. Then, for any finite $A \in \mathcal{A}$, there exists $a^* \in A$ such that $a^* \succeq_r a$ and $a^* \succeq_w a$ for all $a \in A$.*

Proof. Because $\succeq_w$ is complete by definition and transitive on $\varphi_r(A)$ by Axiom 3(ii), there exists a^* which maximizes $\succeq_w$ on $\varphi_r(A)$. By $a^* \in \varphi_r(A)$, we must have $a^* \succeq_r a$ for all $a \in A$. Also, for any $a \in A \setminus \varphi_r(A)$, we have $a^* \succ_r a$, so Axiom 3(i) implies $a^* \succeq_w a$. Thus, $a^* \succeq_r a$ and $a^* \succeq_w a$ for all $a \in A$. $\square$

Lemma S24. *Suppose that Axioms 1-6 hold and that A and B are finite.*

(i) If $A \succ A \cup B$, there is $b \in B \setminus A$ such that $b \succ_r a$ or $b \succ_w a$ for all $a \in A$.

(ii) If $A \cup B \succ A$ and $\mathcal{C}\,(A \cup B) \cap A \neq \emptyset$, there is $b \in B \setminus A$ such that $b \succ_r a$ for all $a \in A$.

Proof. Note that by the hypotheses, $A \not\supset B$ holds for (i)(ii).

(i) By Lemma S23, there exists $a^* \in A$ such that $a^* \succeq_r a$ and $a^* \succeq_w a$ for all $a \in A$. To prove the contrapositive, suppose that for any $b \in B \setminus A$, there exists $a \in A$ such that $a \succeq_r b$ and $a \succeq_w b$. By transitivity and Axiom 3(i)(ii), $a^* \succeq_r c$ and $a^* \succeq_w c$ for all $c \in A \cup B$. Thus, Axiom 6(i) yields $A \cup B \succeq A$.

(ii) If the conclusion is false, then we have $a^* \in A$ such that $a^* \succeq_r c$ for all $c \in A \cup B$ and $a^* \succeq_w c$ for all $c \in A$. If $a^* \succeq_w b$ for all $b \in B$, then by Axiom 6(i), it is impossible to have $A \cup B \succ A$ and $\mathcal{C}\,(A \cup B) \cap A \neq \emptyset$ simultaneously. If $b \succ_w a^*$ for some $b \in B$, then Axiom 6(ii) yields the same conclusion. $\square$

S.C.1.3 Supplemental Results for Lemma 4.

Lemma S25. *Suppose Axioms 3(i) and 6(i) hold. Then, $b \in \mathcal{P}(a) \cup \mathcal{S}(a) \cup \mathcal{N}_1(a)$ implies $\{a, b\} \succ \{a\}$, and $b \in \mathcal{N}_2(a)$ implies $\{a, b\} \sim \{a\}$.*

Proof. If $a \succ_r b$, Axiom 3(i) implies $a \succeq_w b$. Then, Axiom 6(i) yields the conclusion. $\square$

Lemma S26. *Suppose Axioms 1-6 hold. If $\{c, d\} \in \mathcal{B}_S$, then $\{c\} \succeq \{c, d\}$ or $\{d\} \succeq \{c, d\}$.*

[40] In fact, Axioms 3(i)(ii) are enough to show the transitivity of $\succ_r$, with a longer proof.

Proof. The conclusion trivially holds if $c = d$, so we assume $c \neq d$. Without loss of generality, let $c \succ_r d$. If $d \in \mathcal{S}(c) \cup \mathcal{N}_1(c)$, then $\{d\} \succeq \{c, d\}$ by definition. If $d \in \mathcal{I}(c)$, then Lemma 2(ii) implies either $\{c\} \succeq \{c, d\}$ or $\{d\} \succeq \{c, d\}$. □

Lemma S27. *Suppose Axioms 1-6 hold. If $A \in \mathcal{A}_S \cup \mathcal{A}_N$, then $A \sim \{e\}$ for some $e \in \Delta$.*

Proof. If $A \in \mathcal{A}_N$, then the conclusion follows from Lemma S25. Suppose $A \in \mathcal{A}_S$. Note that iteratively applying Lemma 2 yields $A \succeq \{a\}$ for some $a \in A$.[41] If $A \sim \{a\}$ for some $a \in A$, the conclusion holds. Next, suppose $\{a'\} \succ A \succ \{a\}$ for some $a, a' \in A$. Then, because $\{a\}, \{a'\}, A \in \mathcal{A}_S$, we have $V^S(\{a'\}) > V^S(A) > V^S(\{a\})$. By linearity, there exists $\alpha \in (0,1)$ such that $V^S(\{a'\alpha a\}) = \alpha V^S(\{a'\}) + (1-\alpha)V^S(\{a\}) = V^S(A)$. Thus, $\{a'\alpha a\} \sim A$. Finally, to see that $A \succ \{a\}$ for all $a \in A$ does not occur, recall we can write $A = \sum_{m=1}^{M_A} \alpha_m \{a_{1m}, a_{2m}\}$ where $\{a_{1m}, a_{2m}\} \in \mathcal{B}_S$ and $\sum_{m=1}^{M_A} \alpha_m = 1$. By Lemma S26, there exist $(e_m)_{m=1}^{M_A}$, with $e_m \in \{a_{1m}, a_{2m}\}$ for each m, such that $\{e_m\} \succeq \{a_{1m}, a_{2m}\}$. By Axiom 4, we have $\{\sum_{m=1}^{M_A} \alpha_m e_m\} \succeq A$. □

S.C.1.4 Supplemental Results for Theorem 1 (Sufficiency), Continued

Proof of Lemma 6. (i) By Axiom 5(i), $\mathcal{C}(\{a,b\}\alpha\{c,d\}) = \{b\alpha d\}$. Also, by the linearity of r and Axiom 3(i), we have $r(a\alpha c) > r(z)$, hence $a\alpha c \succeq_w z$, for all $z \in A \setminus \{a\alpha c\}$. Therefore, Axiom 6(i) implies $A \sim \{a\alpha c, b\alpha d\}$.

(ii) The same argument as (i) yields the result.

(iii) Let V_{PS} be a function that represents $\succeq$ over finite menus in $\mathcal{A}$. By Eq.(11), there exists $z \in A$ such that $V_{PS}(A) = \min_{z' \in \varphi_r(A)} V_{PS}(\{z, z'\})$. If $z = a\alpha c$, then by $a\alpha c \in \varphi_r(A)$ and Axioms 3(i) and 6(i), we have $V_{PS}(A) \leq V_{PS}(\{a\alpha c\}) < \alpha V_{PS}(\{a,b\}) + (1-\alpha)V_{PS}(\{c,d\}) = V_{PS}(A)$, a contradiction. If $z = a\alpha d$, then $V_{PS}(A) \leq V_{PS}(\{a\}\alpha\{c,d\}) < V_{PS}(A)$, a contradiction. A similar contradiction results if $z = b\alpha c$. Thus, $V_{PS}(A) = \min_{z' \in \varphi_r(A)} V_{PS}(\{b\alpha d, z'\}) \leq V_{PS}(\{a\alpha c, b\alpha d\})$. Now, suppose $V_{PS}(b) > V_{PS}(\{a,b\})$. Note we have $V_{PS}(A) = V_{PS}(\{b\alpha d, a\alpha c\})$ or $V_{PS}(A) = V_{PS}(\{b\alpha d, b\alpha c\})$. In the latter case, $V_{PS}(A) = \alpha V_{PS}(\{b\}) + (1-\alpha)V_{PS}(\{c,d\}) > V_{PS}(A)$, a contradiction. Thus, $A \sim \{a\alpha c, b\alpha d\}$.

(iv) Note we have $\varphi_r(A) = \{a\alpha c, b\alpha c\}$. Also, by Axiom 4(iii), we have $\{b\alpha c\} \succ \{a\alpha c, b\alpha c\}$, so $a\alpha c \succ_w b\alpha c$. By Axiom 3(i), $a\alpha c \succeq_w z$ for all $z \in A$. Also, by Axiom 5(i), $\mathcal{C}(A) = \mathcal{C}(\{a,b\})\alpha\{d\}$. Therefore, Axiom 6(i) yields the desired conclusion.

(v) Let $\mathcal{C}(\{a,b\}) = \{a\}$. We first prove the last two relations in (v-a). By Axiom 5(i), $\mathcal{C}(A) = \{a\alpha d\}$ and $\mathcal{C}(\{a,b\}\alpha\{c\}) = \{a\alpha c\}$. Also, Axiom 4 implies $\{b\alpha c\} \sim \{a\alpha c, b\alpha c\}$,

[41] Denote $A = \{a_1, \cdots, a_{|A|}\}$ where $\{a_1\} \succeq \{a_2\} \succeq \cdots \succeq \{a_{|A|}\}$. If $a_{|A|-1} \succ_r a_{|A|}$, then Lemma 2(i) implies $\{a_{|A|-1}, a_{|A|}\} \succeq \{a_{|A|-1}\} \succeq \{a_{|A|}\}$. If $a_{|A|} \succ_r a_{|A|-1}$, then Lemma 2(i) implies $\{a_{|A|-1}, a_{|A|}\} \succeq \{a_{|A|}\}$. If $a_{|A|-1} \sim_r a_{|A|}$, then Lemma 2(ii) implies $\{a_{|A|-1}, a_{|A|}\} \succeq \{a_{|A|}\}$. Repeating similar arguments yields $A \succeq \{a_{|A|}\}$.

so $a\alpha c \succ_w b\alpha c$. Also, we have $a\alpha c \succ_r a\alpha d, b\alpha d$, so Axiom 3(i) implies $a\alpha c \succeq_r z$ and $a\alpha c \succeq_w z$ for all $z \in A$. By Axiom 6(i), $A \sim \{a\alpha c, a\alpha d\} \succ \{a\alpha c, b\alpha d\}$. Next, to show the first relation in (v-a), note that $b\alpha c \succeq_r z$ for all $z \in A$, $b\alpha c \succeq_w a\alpha d$ (by Axiom 3(i)), and $a\alpha c \succ_w b\alpha c$. Thus, applying Axiom 6(ii) to $\tilde{A} = \{b\alpha c, a\alpha d\}$ and $\tilde{B} = \{a\alpha c, b\alpha d\}$, we obtain $\{b\alpha c, a\alpha d\} \succ \tilde{A} \cup \tilde{B} = A$. Finally, to show (v-b), suppose $\mathcal{C}(\{a, b\}) = \{a, b\}$. Then we have $\mathcal{C}(A) = \{a\alpha d, b\alpha d\}$, and $a\alpha c, b\alpha c \succeq_r z$ and $a\alpha c, b\alpha c \succeq_w z$ for all $z \in A$. Thus, applying Axiom 6(i) to $\tilde{A} = \{a\alpha c, b\alpha d\}$ and $\tilde{B} = \{b\alpha c, a\alpha d\}$ yields $A \sim \{a\alpha c, b\alpha d\}$ and applying it to $\tilde{A} = \{a\alpha c, a\alpha d\}$ and $\tilde{B} = \{b\alpha c, b\alpha d\}$ yields $A \sim \{a\alpha c, a\alpha d\}$. □

Lemma S28. *Suppose Axioms 1-5 hold, $y \in \mathcal{P}(x)$, and $y' \in \mathcal{S}(x)$.*

(i) There exists $\delta \in (0, 1)$ such that $y(1-\delta)c \in \mathcal{P}(x)$ and $y'(1-\delta)c \in \mathcal{S}(x)$ for all $c \in \Delta$.

(ii) $y(1-\delta)c \in \mathcal{P}(x(1-\delta)c)$ and $y'(1-\delta)c \in \mathcal{S}(x(1-\delta)c)$ for all $c \in \Delta$ and all $\delta \in (0, 1)$.

Proof. (i) By definition, $\{x, y\} \succ \{y\}$ and $\mathcal{C}(\{x, y\}) = \{y\}$. Because the restriction of $\succeq$ to singleton sets is continuous, and because Δ is compact, there exists $\delta_1 \in (0, 1)$ such that $\{x, y(1-\delta)c\} \succ \{y(1-\delta)c\}$ for all $c \in \Delta$ and $\delta \in (0, \delta_1)$.[42] Also, by Axiom 3(iii-b) and compactness, we have some $\delta_2 \in (0, 1)$ such that $\mathcal{C}(\{x, y(1-\delta)c\}) = \{y(1-\delta)c\}$ for all $c \in \Delta$ and $\delta \in (0, \delta_2)$. Therefore, by taking $\underline{\delta}^P = \min\{\delta_1, \delta_2\}$, the first half of the statement holds for all $\delta < \underline{\delta}^P$. An analogous argument yields $\underline{\delta}^S$ such that the second half of the statement holds for all $\delta < \underline{\delta}^S$. Thus, $\delta < \min\left\{\underline{\delta}^P, \underline{\delta}^S\right\}$ satisfies the desired property.

(ii) The conclusion is an immediate consequence of Axioms 4(iii) and 5(iii). □

Proof of Lemma 7. (i) Because $y, c \in \mathcal{P}(x)$, $V_{PS}(\{x, y(1-\delta)c\}) = V_{PS}(\{x, y\}(1-\delta)\{x, c\})$ by Lemma 6(i). Therefore, using $\{x, y\}, \{x, c\} \in \mathcal{B}_P$,

$$\begin{aligned} w_P(c; x, y, \delta) &= \frac{1}{\delta}\left[(1-\delta)V_{PS}(\{x, y\}) + \delta V_{PS}(\{x, c\}) - (1-\delta)V_{PS}(\{x, y\}) - \delta V_{PS}(\{c\})\right] \\ &= V_{PS}(\{x, c\}) - V_{PS}(\{c\}). \end{aligned}$$

(ii) The result follows from $V_{PS}(\{x, y(1-\delta)x\}) = (1-\delta)V_{PS}(\{x, y\}) + \delta V_{PS}(\{x\})$.

(iii) By Lemma 6(i), $V_{PS}(\{x, [y(1-\delta)c]\,\alpha\,[y(1-\delta)c']\}) = V_{PS}(\{x, y(1-\delta)c\}\,\alpha\,\{x, y(1-\delta)c'\})$. Therefore,

$$\begin{aligned} w_P(c\alpha c'; x, y, \delta) =& \frac{1}{\delta}V_{PS}(\{x, [y(1-\delta)c]\,\alpha\,[y(1-\delta)c']\}) - \frac{1-\delta}{\delta}V_{PS}(\{x, y\}) - V_{PS}(\{c\alpha c'\}) \\ =& \frac{\alpha}{\delta}V_{PS}(\{x, y(1-\delta)c\}) + \frac{1-\alpha}{\delta}V_{PS}(\{x, y(1-\delta)c'\}) \\ & - \frac{1-\delta}{\delta}V_{PS}(\{x, y\}) - \alpha V_{PS}(\{c\}) - (1-\alpha)V_{PS}(\{c'\}) \end{aligned}$$

[42] Let $A = \{x, y\}\,.5\,\{y\}$. By Axiom 4, we have $\{x, y\} \succ A \succ \{y\}$. By Axiom 2(i) and the continuity of V_{PS} on singletons, we have $\{x, y(1-\delta)c\} \succ A$ and $A \succ \{y(1-\delta)c\}$ for all sufficiently small δ.

$$=\alpha w_P(c;x,y,\delta)+(1-\alpha)w_P(c';x,y,\delta).$$

(iv) Let $\delta'\in(0,\delta)$. Note that $y(1-\delta')c=y\frac{\delta-\delta'}{\delta}[y(1-\delta)c]$. Because $y,y(1-\delta)c\in\mathcal{P}(x)$, Lemma 6(i) implies $V_{PS}\left(\left\{x,y\frac{\delta-\delta'}{\delta}[y(1-\delta)c]\right\}\right)=V_{PS}\left(\{x,y\}\frac{\delta-\delta'}{\delta}\{x,y(1-\delta)c\}\right)$. Therefore, $V_{PS}(\{x,y(1-\delta)c\})=\frac{\delta}{\delta'}V_{PS}(\{x,y(1-\delta')c\})-\frac{\delta-\delta'}{\delta'}V_{PS}(\{x,y\})$. Substituting this into the definition, we have

$$\begin{aligned}w_P(c;x,y,\delta)=&\frac{1}{\delta'}V_{PS}(\{x,y(1-\delta')c\})-\frac{\delta-\delta'}{\delta\delta'}V_{PS}(\{x,y\})-\frac{1-\delta}{\delta}V_{PS}(\{x,y\})-V_{PS}(\{c\})\\=&w_P(c;x,y,\delta').\end{aligned}$$

(v) Our goal is to show $w_P(c;x,y,\delta)=w_P(c;a,b,\delta)+w_P(a;x,y,\delta)$ or, equivalently,

$$\begin{aligned}&\frac{1}{\delta}V_{PS}(\{x,y(1-\delta)c\})=\frac{1}{\delta}V_{PS}(\{a,b(1-\delta)c\})-\frac{1-\delta}{\delta}V_{PS}(\{a,b\})\\&-V_{PS}(\{a\})+\frac{1}{\delta}V_{PS}(\{x,y(1-\delta)a\}).\end{aligned}$$

By (ii), we have $V_{PS}(\{a\})=\frac{1}{\delta}V_{PS}(\{a,b(1-\delta)a\})-\frac{1-\delta}{\delta}V_{PS}(\{a,b\})$. Substituting this into the above expression, our goal is to show

$$V_{PS}(\{x,y(1-\delta)c\}.5\{a,b(1-\delta)a\})=V_{PS}(\{a,b(1-\delta)c\}.5\{x,y(1-\delta)a\}).$$

Because $y(1-\delta)c,y(1-\delta)a\in\mathcal{P}(x)$ and $b(1-\delta)a,b(1-\delta)c\in\mathcal{P}(a)$, Lemma 6(i) implies that both sides of this equation equal $V_{PS}(\{x.5a,[(1-\delta)(y+b)].5[\delta(a+c)]\})$. $\square$

Lemma S29. *Suppose Axioms 1-7 hold, $y(1-\delta)c\in\mathcal{P}(x)$, and $y'(1-\delta)c\in\mathcal{S}(x)$ for all $c\in\Delta$. Then, $\beta\equiv\frac{1}{\alpha}-1>0$, where $\alpha\in(0,1)$ is as defined in Axiom 7, satisfies the following condition: $w_P(c;x,y,\delta)=\beta w_S(c;x,y',\delta)$ for all $c\in\Delta$.*

Proof. We have

$$\begin{aligned}&\delta\left[w_P(c;x,y,\delta)-\beta w_S(c;x,y',\delta)\right]\\=&V_{PS}(\{x,y(1-\delta)c\})-(1-\delta)V_{PS}(\{x,y\})-\delta V_{PS}(\{c\})\\&-\beta\left[V_{PS}(\{x,y'(1-\delta)c\})-(1-\delta)V_{PS}(\{x,y'\})-\delta V_{PS}(\{c\})\right]\\=&\frac{1}{\alpha}\left[\alpha V_{PS}(\{x,y(1-\delta)c\})+(1-\alpha)V_{PS}(\{x(1-\delta)c,y'(1-\delta)c\})\right]\\&-\frac{1}{\alpha}\left[\alpha V_{PS}(\{x(1-\delta)c,y(1-\delta)c\})+(1-\alpha)V_{PS}(\{x,y'(1-\delta)c\})\right]\\=&\frac{1}{\alpha}\left[\alpha V_{PS}(\{x,y(1-\delta)c\})+(1-\alpha)V_{PS}\left(\left\{e^{x(1-\delta)c,y'(1-\delta)c}\right\}\right)\right]\\&-\frac{1}{\alpha}\left[\alpha V_{PS}(\{x(1-\delta)c,y(1-\delta)c\})+(1-\alpha)V_{PS}\left(\left\{e^{x,y'(1-\delta)c}\right\}\right)\right]\end{aligned}$$

$= 0$

where the last equality holds because $y(1-\delta)c \in \mathcal{P}(x) \cap \mathcal{P}(x(1-\delta)c)$ and $y'(1-\delta)c \in \mathcal{S}(x) \cap \mathcal{S}(x(1-\delta)c)$ hold by Lemma S28, so that Axiom 7 applies. Thus, $w_P(c; x, y, \delta) = \beta w_S(c; x, y', \delta)$ where $\beta > 0$. □

S.C.2 Supplemental Proofs for Theorem 2

Lemma S30. *Suppose the data are generated by a PS preference. If $A \cup \{a\} \succ A$ and $a \notin \mathcal{C}(A \cup \{a\})$, then $\varphi_r(A \cup \{a\}) = \{a\}$ and $w(a) = \max_{c' \in \varphi_r(A \cup \{a\})} w(c') < \max_{c' \in \varphi_r(A)} w(c')$.*

Proof. Suppose $A \cup \{a\} \succ A$ and $a \notin \mathcal{C}(A \cup \{a\})$. If $\max_{c' \in \varphi_r(A \cup \{a\})} w(c') \geq \max_{c' \in \varphi_r(A)} w(c')$, then $V_{PS}(A \cup \{a\}) = G\left(A \cup \{a\}, \max_{c' \in \varphi_r(A \cup \{a\})} w(c')\right) = G\left(A, \max_{c' \in \varphi_r(A \cup \{a\})} w(c')\right) \leq G\left(A, \max_{c' \in \varphi_r(A)} w(c')\right) = V_{PS}(A)$ where the second equality follows from $a \notin \mathcal{C}(A \cup \{a\})$. This is a contradiction. Thus, $\max_{c' \in \varphi_r(A \cup \{a\})} w(c') < \max_{c' \in \varphi_r(A)} w(c')$, and we must have $\varphi_r(A \cup \{a\}) = \{a\}$ and $\max_{c' \in \varphi_r(A \cup \{a\})} w(c') = w(a)$.[43] □

Lemma S31. *Suppose the data are generated by a weakly nondegenerate PS preference. If $a \succ_w b$, then $r(a) \geq r(b)$ and $w(a) > w(b)$.*

Proof. Consider the following exhaustive cases.

Case 1. If $\{b\} \succ \{a, b\}$, then $g(b, w(b)) = u(b) > \max_{c \in \{a,b\}} g\left(c, \max_{c' \in \varphi_r(\{a,b\})} w(c')\right) \geq g\left(b, \max_{c' \in \varphi_r(\{a,b\})} w(c')\right)$, so $\max_{c' \in \varphi_r(\{a,b\})} w(c') > w(b)$, yielding the conclusion.

Case 2. If $\{b\} \sim \{a, b\}$ and $\mathcal{C}(\{a, b\}) = \{a\}$, then $g(b, w(b)) = u(b) = g\left(a, \max_{c' \in \varphi_r(\{a,b\})} w(c')\right) > g\left(b, \max_{c' \in \varphi_r(\{a,b\})} w(c')\right)$, so the conclusion holds as in *Case 1*.

Case 3. If $a \sim_r b$ and $\{a\} \sim \{a, b\} \succ \{b\}$, then we have $a \in \mathcal{C}(\{a, b\})$ (otherwise, $a \succ_r b$, a contradiction). Thus, $u(a) = g\left(a, \max_{c' \in \{a,b\}} w(c')\right)$. This in turn implies $\max_{c' \in \{a,b\}} w(c') = w(a) \geq w(b)$. By weak nondegeneracy, the straight indifference curves of r and w cross each other (see Lemma 17). Because $r(a) = r(b)$ and $a \neq b$, we have $w(a) > w(b)$. □

S.C.3 Proof of Theorem 1 (Necessity Part)

Below, we show that a nondegenerate preference that has a PS representation satisfies each axiom. Proofs of Axiom 1 and Axiom 2(iii) are straightforward and omitted.

To proceed to other axioms, note first that the functions $g(c, R) = u(c) - \max\{R - w(c), 0\} + \beta \max\{w(c) - R, 0\}$ and $G(A, R) = \max_{c \in A} g(c, R)$ are strictly decreasing and continuous in R and that G is continuous in A.

[43] Note that for any $A, B \in \mathcal{A}$, $\varphi_r(A \cup B) \in \{\varphi_r(A), \varphi_r(B), \varphi_r(A) \cup \varphi_r(B)\}$.

Axiom 2(i). Suppose $A \succeq B_n$ for all n and $B_n \to B$. Then, because $\max_{c' \in \varphi_r(B)} w(c') \geq \lim_{n\to\infty} \max_{c' \in \varphi_r(B_n)} w(c')$, we have

$$\begin{aligned} V_{PS}(A) = G\left(A, \max_{c' \in \varphi_r(A)} w(c')\right) &\geq \lim_{n\to\infty} G\left(B_n, \max_{c' \in \varphi_r(B_n)} w(c')\right) \\ &= G\left(B, \lim_{n\to\infty} \max_{c' \in \varphi_r(B_n)} w(c')\right) \\ &\geq G\left(B, \max_{c' \in \varphi_r(B)} w(c')\right) = V_{PS}(B). \end{aligned}$$

Axiom 2(ii). Suppose $A \succ B \succ C$. Note

$$\begin{aligned} V_{PS}(A\alpha C) &= G\left(A\alpha C, \max_{c' \in \varphi_r(A\alpha C)} w(c')\right) \\ &= G\left(A\alpha C, \alpha \max_{c' \in \varphi_r(A)} w(c') + (1-\alpha) \max_{c' \in \varphi_r(C)} w(c')\right). \end{aligned}$$

By the continuity of G, $V_{PS}(A\alpha C) \approx V_{PS}(C) < V_{PS}(B)$ for sufficiently small $\alpha \in (0,1)$.

Axiom 3(i). If $a \succ_r b$, then Theorem 2 implies $r(a) > r(b)$. Thus, Theorem 2 implies $b \not\succ_r a$, and Lemma S31 implies $b \not\succ_w a$. Similarly, if $a \succ_w b$, then Lemma S31 implies $r(a) \geq r(b)$ and $w(a) > w(b)$, so we cannot have $b \succ_r a$ or $b \succ_w a$.

Axiom 3(ii). Suppose $a \succ^* b \succ^* c$. By definition, there exist $A \ni b$ and $B \ni c$ such that $A \cup \{a\} \succ A$, $a \notin \mathcal{C}(A \cup \{a\})$, $B \cup \{b\} \succ B$, and $b \notin \mathcal{C}(B \cup \{b\})$. Now, let $C = A \cup B$. By Lemma S30, we have $\varphi_r(C \cup \{a\}) = \{a\}$ and $w(a) < \max_{c' \in \varphi_r(A)} w(c') = \max_{c' \in \varphi_r(C)} w(c')$.[44] The representation implies $a \notin \arg\max_{d \in C \cup \{a\}} g(d, w(a)) = \mathcal{C}(C \cup \{a\})$ and $C \cup \{a\} \succ C$. Thus, $a \succ^* c$. Next, if $a \sim_r b \sim_r c$, Theorem 2 implies $[a \succeq_w b] \wedge [b \succeq_w c] \Leftrightarrow [w(a) \geq w(b)] \wedge [w(b) \geq w(c)] \Rightarrow w(a) \geq w(c) \Leftrightarrow a \succeq_w c$.

Axiom 3(iii-a). Suppose $A\alpha_n C \succeq B$ for all n and $\alpha_n \to \alpha$. Because

$$\lim_{n\to\infty} \max_{c' \in \varphi_r(A\alpha_n C)} w(c') = \lim_{n\to\infty} \left[\alpha_n \max_{c' \in \varphi_r(A)} w(c') + (1-\alpha_n) \max_{c' \in \varphi_r(C)} w(c')\right] = \max_{c' \in \varphi_r(A\alpha C)} w(c'),$$

we have $A\alpha C \succeq B$ as follows:

$$\begin{aligned} V_{PS}(A\alpha C) &= G\left(A\alpha C, \max_{c' \in \varphi_r(A\alpha C)} w(c')\right) \\ &= \lim_{n\to\infty} G\left(A\alpha_n C, \max_{c' \in \varphi_r(A\alpha_n C)} w(c')\right) = \lim_{n\to\infty} V_{PS}(A\alpha_n C) \geq V_{PS}(B). \end{aligned}$$

Axiom 3(iii-b). Suppose $a^* \in A$ is such that $a^* \succ_r a$ for all $a \in A \setminus \{a^*\}$. Take

[44] For any $a' \in \varphi_r(A)$ and $b' \in \varphi_r(B)$, we have $r(a) > r(a') \geq r(b) > r(b')$.

any $(A_n)_n$ and any $(a_n)_n$ such that $A_n \to A$, $a_n \in \mathcal{C}(A_n)$, and $a_n \to a$. By Theorem 2, $r(a^*) > r(a)$ for all $a \in A \setminus \{a^*\}$, so $\lim_{n\to\infty} \max_{c' \in \varphi_r(A_n)} w(c') = w(a^*)$. By continuity, $g\left(c, \max_{c' \in \varphi_r(A_n)} w(c')\right) \to g\left(c, w(a^*)\right)$ for all c. By $a_n \in \mathcal{C}(A_n)$, we have $g\left(a_n, \max_{c' \in \varphi_r(A_n)} w(c')\right) = G(A_n, \max_{c' \in \varphi_r(A_n)} w(c'))$, so letting $n \to \infty$ yields $g\left(a, w(a^*)\right) = G(A, w(a^*))$. Thus, $a \in \mathcal{C}(A)$.

Axiom 3(iv). By Lemma 18, $a\alpha c \succ^* b\alpha c \Rightarrow r(a) > r(b) \Rightarrow a \succ^* b$.

To prove some of the remaining axioms, we use the following result.

Lemma S32. *Suppose the choice data are generated by a PS preference. (i) If $b \in \mathcal{P}(a)$, then $w(a) < w(b)$. (ii) If $b \in \mathcal{S}(a)$, then $w(a) > w(b)$. (iii) If $b \in \mathcal{N}_1(a)$, then $w(a) = w(b)$.*

Proof. (i) By the representation and the definition of $\mathcal{P}(a)$, $g(b, w(a)) > u(b) = g(b, w(b))$, so $w(a) < w(b)$. (ii) If $b \in \mathcal{S}(a)$, we have $g(b, w(a)) < g(b, w(b))$, so $w(a) > w(b)$. (iii) If $b \in \mathcal{N}_1(a)$, we have $g(b, w(a)) = g(b, w(b))$, so $w(a) = w(b)$. □

Proof of Theorem 1 (Necessity), Continued.

Axiom 4(i). We consider the cases where mixed menus are binary; mixtures with a singleton are considered in Axiom 4(iii). Suppose $b \in \mathcal{P}(a) \cup \mathcal{N}_1(a)$, $d \in \mathcal{P}(c) \cup \mathcal{N}_1(c)$, and $f \in \mathcal{P}(e) \cup \mathcal{N}_1(e)$. By Lemma S32, $r(a) > r(b)$, $w(a) \leq w(b)$, $r(e) > r(f)$, and $w(e) \leq w(f)$. Therefore, for any $\alpha \in (0, 1)$,

$$
\begin{aligned}
&V_{PS}(\{a,b\}\,\alpha\,\{e,f\}) \\
&= \max_{x \in \{a,b\}\alpha\{e,f\}} \left[u(x) + \beta\left(w(x) - w(a\alpha e)\right)\right] \\
&= \alpha \max_{x \in \{a,b\}} \left[u(x) + \beta\left(w(x) - w(a)\right)\right] + (1-\alpha) \max_{x \in \{e,f\}} \left[u(x) + \beta\left(w(x) - w(e)\right)\right)] \\
&= \alpha V_{PS}(\{a,b\}) + (1-\alpha) V_{PS}(\{e,f\}).
\end{aligned}
$$

Similarly, $V_{PS}(\{c,d\}\,\alpha\,\{e,f\}) = \alpha V_{PS}(\{c,d\}) + (1-\alpha) V_{PS}(\{e,f\})$. Thus, $\{a,b\} \succ (\succeq) \{c,d\}$ implies $\{a,b\}\,\alpha\,\{e,f\} \succ (\succeq) \{c,d\}\,\alpha\,\{e,f\}$.

Axiom 4(ii). Again, consider the cases where mixed menus are binary. Take any $\{a,b\}, \{e,f\} \in \mathcal{B}_S$ such that $a \neq b$ and $e \neq f$. By Lemma S32, we can assume without loss of generality that $r(a) \geq r(b)$, $w(a) \geq w(b)$, $r(e) \geq r(f)$, and $w(e) \geq w(f)$. Then,

$$
\begin{aligned}
V_{PS}(\{a,b\}\,\alpha\,\{e,f\}) &= \max_{x \in \{a,b\}\alpha\{e,f\}} \left[u(x) + w(x) - w(a\alpha e)\right] \\
&= \alpha V_{PS}(\{a,b\}) + (1-\alpha) V_{PS}(\{e,f\}).
\end{aligned}
$$

Therefore, the conclusion of Axiom 4(ii) holds.

Axiom 4(iii). Note that for any $x \in A$, we have

$$
\begin{aligned}
& g\left(x\alpha c, \max_{c' \in \varphi_r(A\alpha\{c\})} w(c')\right) \\
&= \alpha u(x) + (1-\alpha)u(c) - \alpha \max\left\{\max_{c' \in \varphi_r(A)} w(c') - w(x), 0\right\} + \alpha\beta \max\left\{w(x) - \max_{c' \in \varphi_r(A)} w(c'), 0\right\} \\
&= \alpha g\left(x, \max_{c' \in \varphi_r(A)} w(c')\right) + (1-\alpha)u(c).
\end{aligned}
$$

Thus, the conclusion follows from

$$
\begin{aligned}
V_{PS}\left(A\alpha\left\{c\right\}\right) &= \max_{x \in A} g\left(x\alpha c, \max_{c' \in \varphi_r(A\alpha\{c\})} w(c')\right) \\
&= \alpha \max_{x \in A} g\left(x, \max_{c' \in \varphi_r(A)} w(c')\right) + (1-\alpha)u(c) = \alpha V_{PS}\left(A\right) + (1-\alpha)V_{PS}\left(\left\{c\right\}\right).
\end{aligned}
$$

Axiom 5(i). Consider first the case where $b \in \mathcal{P}(a) \cup \mathcal{N}_1(a)$ and $d \in \mathcal{P}(c) \cup \mathcal{N}_1(c)$. Following the proof of Axiom 4(i),

$$
\begin{aligned}
& \mathcal{C}(\{a,b\}\,\alpha\,\{c,d\}) \\
&= \arg\max_{x \in \{a,b\}\alpha\{c,d\}} \left[u(x) + \beta\left(w(x) - w(a\alpha c)\right)\right] \\
&= \alpha \arg\max_{x \in \{a,b\}} \left[u(x) + \beta\left(w(x) - w(a)\right)\right] + (1-\alpha) \arg\max_{x \in \{c,d\}} \left[u(x) + \beta\left(w(x) - w(c)\right)\right] \\
&= \mathcal{C}(\{a,b\})\alpha\mathcal{C}(\{c,d\}).
\end{aligned}
$$

Proof for the case $b \in \mathcal{S}(a) \cup \mathcal{N}_1(a) \cup \mathcal{I}(a)$ and $d \in \mathcal{S}(c) \cup \mathcal{N}_1(c) \cup \mathcal{I}(a)$ is analogous: letting $w(a) \geq w(b)$ and $w(c) \geq w(d)$ without loss of generality,

$$
\mathcal{C}(\{a,b\}\,\alpha\,\{c,d\}) = \arg\max_{x \in \{a,b\}\alpha\{c,d\}} \left[u(x) + w(x) - w(a\alpha c)\right] = \mathcal{C}(\{a,b\})\alpha\mathcal{C}(\{c,d\}).
$$

Before proving Axiom 5(ii), we note that Axiom 5(iii) can be shown by following the proof of Axiom 4(iii). Thus, the proof of Axiom 5(iii) is omitted.

Axiom 5(ii). For (ii-a), suppose $A = \{a,b\}\,\alpha\,\{a,c\}$, $b \in \mathcal{N}_2(a)$, $c \in \mathcal{P}(a)$, $\{a, b\alpha c\} \succeq \{b\}\,\alpha\,\{a,c\}$, and $\mathcal{C}\left(\{a, b\alpha c\}\right) = \{b\alpha c\}$. By Axiom 5(iii), we have $\mathcal{C}(\{b\}\alpha\{a,c\}) = \{b\alpha c\}$. By the representation, $g(b\alpha c, w(a)) = V_{PS}(\{a, b\alpha c\}) \geq V_{PS}(\{b\}\,\alpha\,\{a,c\}) = g(b\alpha c, w(b\alpha a))$. Therefore, we have $w(b\alpha a) \geq w(a)$, so $w(b) \geq w(a)$. Because $r(a) > r(b), r(c)$ and $w(a) \leq w(b), w(c)$, following the proof of Axiom 5(i) yields $\mathcal{C}\left(A\right) = \mathcal{C}\left(\{a,b\}\right)\alpha\mathcal{C}\left(\{a,c\}\right)$. Proof of (ii-b) is analogous, once we note that the assumptions imply $g(b\alpha c, w(b\alpha a)) \geq g(b\alpha c, w(a))$

so that we have $w(a) \geq w(b)$, as well as $r(a) > r(b), r(c)$ and $w(a) > w(c)$.

Axiom 6(i). Suppose there exists $a^* \in A$ such that $a^* \succeq_r c$ and $a^* \succeq_w c$ for all $c \in A \cup B$. Then, by Theorem 2, $\max_{y \in \varphi_r(A \cup B)} w(y) = \max_{y \in \varphi_r(A)} w(y) = w(a^*)$. Therefore, $V_{PS}(A \cup B) = \max_{x \in A \cup B} g(x, w(a^*)) \geq \max_{x \in A} g(x, w(a^*)) = V_{PS}(A)$ and the inequality is strict if and only if $\mathcal{C}(A \cup B) \cap A = \arg\max_{x \in A \cup B} g(x, w(a^*)) \cap A = \emptyset$.

Axiom 6(ii). Suppose there exists $a^* \in A$ such that $a^* \succeq_r c$ for all $c \in A \cup B$ and $a^* \succeq_w a$ for all $a \in A$, and there exists $b^* \in B$ such that $b^* \succ_w a^*$. By Lemma S31 and Theorem 2, $r(b^*) = r(a^*) \geq r(b)$ for all $b \in B$ and $w(b^*) > w(a^*)$. Without loss of generality, let b^* be a maximizer of $\succeq_w$ on $\varphi_r(B)$. Then, by Theorem 2, $\max_{y \in \varphi_r(A \cup B)} w(y) = w(b^*) > w(a^*) = \max_{y \in \varphi_r(A)} w(y)$. Therefore, if there exists $c \in \mathcal{C}(A \cup B) \cap A$, then $V_{PS}(A \cup B) = g(c, w(b^*)) < g(c, w(a^*)) \leq V_{PS}(A)$.

Axiom 7. Take any $a, b, c, d \in \Delta$ such that $c \in \mathcal{P}(a) \cap \mathcal{P}(b)$ and $d \in \mathcal{S}(a) \cap \mathcal{S}(b)$. Then,

$$\begin{aligned} V_{PS}(\{a,c\}) + \beta V_{PS}(\{b,d\}) &= u(c) + \beta\left(w(c) - w(a)\right) + \beta\left[u(d) + w(d) - w(b)\right] \\ &= u(c) + \beta\left(w(c) - w(b)\right) + \beta\left[u(d) + w(d) - w(a)\right] \\ &= V_{PS}(\{b,c\}) + \beta V_{PS}(\{a,d\}). \end{aligned}$$

Therefore, by letting $\alpha = \frac{1}{1+\beta} \in (0,1)$, and using $\{e^{b,d}\} \sim \{b,d\}$ and $\{e^{a,d}\} \sim \{a,d\}$,

$$\begin{aligned} V_{PS}\left(\{a,c\}\, \alpha \left\{e^{b,d}\right\}\right) &= \alpha V_{PS}\left(\{a,c\}\right) + (1-\alpha) V_{PS}\left(\left\{e^{b,d}\right\}\right) \\ &= \alpha V_{PS}\left(\{b,c\}\right) + (1-\alpha) V_{PS}\left(\left\{e^{a,d}\right\}\right) = V_{PS}\left(\{b,c\}\, \alpha \left\{e^{a,d}\right\}\right). \end{aligned}$$

Axiom 8. Suppose there exists $a^* \in A \cap B$ such that $a^* \succeq_r c$ and $a^* \succeq_w c$ for all $c \in A \cup B$. By Theorem 2, $\max_{y \in \varphi_r(A)} w(y) = \max_{y \in \varphi_r(B)} w(y) = w(a^*)$. Now, suppose $a, b \in A \cap B$, $a \in \mathcal{C}(A)$, and $b \in \mathcal{C}(B)$. By the representation, $g(a, w(a^*)) = g(b, w(a^*)) \geq g(c, w(a^*))$ for all $c \in A \cup B$. Therefore, $a \in \mathcal{C}(B)$.

S.C.4 Other Proofs

S.C.4.1 Proof of Claim 1

(i) By assumption, there exist $A \in \mathcal{A}$ and $a \in \Delta$ such that $A \cup \{a\} \succ A$ and $a \notin \mathcal{C}(A \cup \{a\})$. By Lemma S30, $\varphi_r(A \cup \{a\}) = \{a\}$ and $w(a) < \max_{c' \in \varphi_r(A)} w(c')$. Thus, for some $b \in A$, we have $r(a) > r(b)$ and $w(a) < w(b)$. (ii) Without loss of generality, assume $\{a\} \succeq \{b\}$ for all $b \in A$. By assumption, $\max_{y \in \varphi_r(A)} w(y) = w(c) > w(a)$ for some $c \in A$ (otherwise, we would have $A \succeq \{a\}$). Therefore, we have $u(a) \geq u(c)$, $w(a) < w(c)$, and $r(a) \leq r(c)$. If we further had $u(a) = u(c)$, then $A \succeq \{c\} \sim \{a\}$, a contradiction. Thus, $u(a) > u(c)$. □

S.C.4.2 Proof of Claim 2

(i) If $\{a,b\} \succ \{b\}$ and $\mathcal{C}(\{a,b\}) = \{b\}$, then we must have $\max_{y\in\varphi(\{a,b\})} w(y) = w(a) < w(b)$, so the DM feels pride by choosing b at $\{a,b\}$. Conversely, pride immediately implies $\{a,b\} \succ \{b\}$, and $\{a,b\} \succ \{a\}$ follows from the representation and $a \notin \mathcal{C}(\{a,b\})$. (ii) Similar to (i). $\square$

S.D Graphical Illustrations of Nondegeneracy and $\succ_r$

In this section, we provide graphical illustrations of the nondegeneracy concepts and the elicitation of the reference ranking, with $\dim(Z) = 3$. Figure S1 illustrates the concepts of nondegeneracy and weak nondegeneracy, providing an example to distinguish the two. It also shows why, in Definition 1, $a \succ^* b$ should be defined using a general menu $A \ni b$ and not just $A = \{b\}$, and it presents a graphical illustration of Theorem 2. Figure S2 then demonstrates how Definition 1(ii-b) helps establish $a \succ_r b$ when we cannot establish the relation via condition (ii-a) (i.e., $a \succ^* b$), which occurs when a is on the boundary of Δ.

Figure S1a illustrates nondegeneracy, which requires that there exist $x, y, y' \in \Delta$ such that $y \in \mathcal{P}(x)$ and $y' \in \mathcal{S}(x)$. For $\mathcal{P}(x)$ to be nonempty, we must have some $y \in \Delta$ such that $r(x) > r(y)$, $w(x) < w(y)$, and $g(x, w(x)) < g(y, w(x))$. The first two conditions ensure that the reference point at $\{x, y\}$ is lower than that at $\{y\}$, and adding the third condition ensures that x is not chosen from $\{x, y\}$. Similarly, for $\mathcal{S}(x)$ to be nonempty, we must have $y' \in \Delta$ such that $r(x) > r(y')$, $w(x) > w(y')$, and $g(x, w(x)) < g(y', w(x))$, ensuring that the reference point at $\{x, y'\}$ is higher than that at $\{y'\}$ and that x is not chosen from $\{x, y'\}$.

Figure S1b provides an example in which the nondegeneracy property is violated. To see this, note that for any $\bar{a}, y \in \Delta$ such that $r(\bar{a}) > r(y)$ and $w(\bar{a}) < w(y)$, we have $g(\bar{a}, w(\bar{a})) > g(y, w(\bar{a}))$, so $\mathcal{P}(\bar{a})$ is empty. In this case, the reference-lowering alternative $\bar{a}$ is also the chosen one, so observing $\{\bar{a}, y\} \succ \{y\}$ does not allow us to tell if the larger menu is preferred because $\bar{a}$ lowers the reference point or because $\bar{a}$ is the preferred choice.

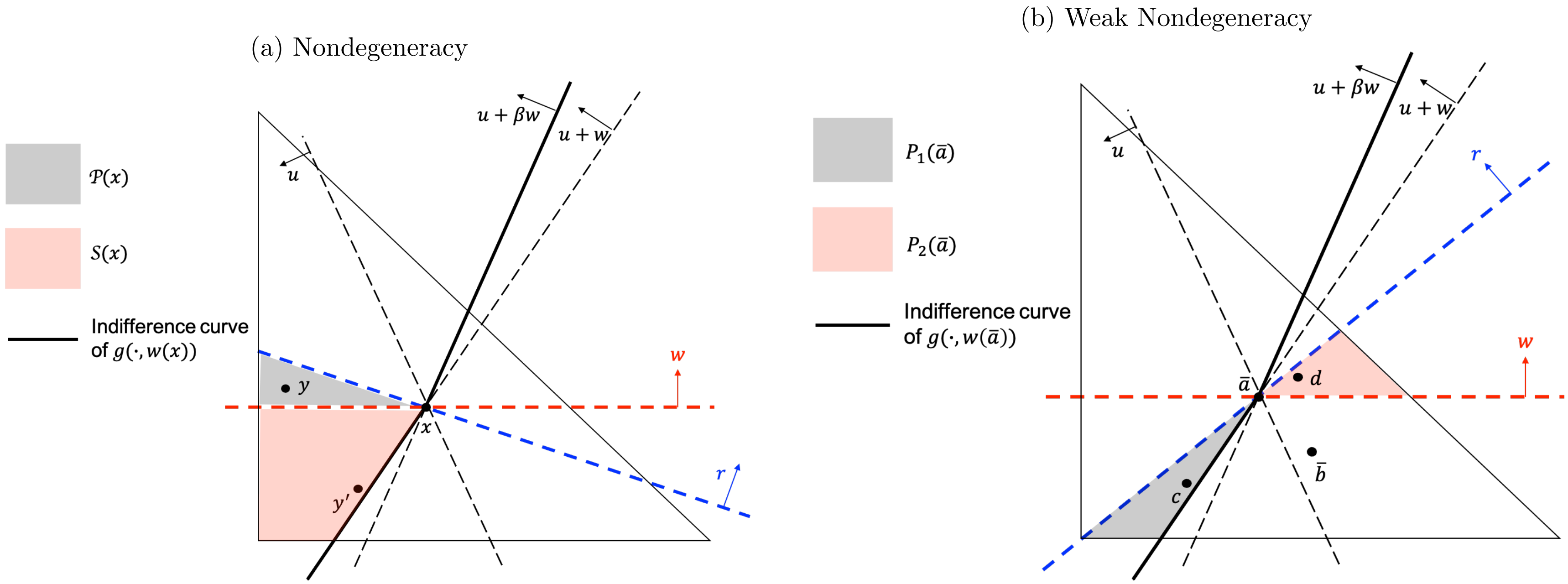
However, the preference illustrated in Figure S1b satisfies the weak nondegeneracy axiom. To see this, note that (i) the reference point at $\{\bar{a}, \bar{b}, c, d\}$, i.e., $w(\bar{a})$, is lower than the reference point at $\{\bar{b}, c, d\}$, i.e., $w(d)$, and (ii) $\mathcal{C}(\{\bar{b}, c, d\}) = \mathcal{C}(\{\bar{a}, \bar{b}, c, d\}) = \{c\}$. In this case, $\bar{a}$ makes the larger menu more desirable even though it is not chosen there, by setting the reference point lower than the reference point at $\{\bar{b}, c, d\}$. Therefore, we have $\bar{a} \succ^* \bar{b}$. This example shows why we cannot confine Definition 1(i) to $A = \{\bar{b}\}$: Even if $r(\bar{a}) > r(\bar{b})$, we may have $\bar{a} \in \mathcal{C}(\{\bar{a}, \bar{b}\})$ ($\bar{a}$ is chosen) or $w(\bar{a}) \geq w(\bar{b})$ ($\bar{a}$ sets a weakly higher reference point), preventing us from concluding $\bar{a} \succ^* \bar{b}$ with $A = \{\bar{b}\}$. Thus, to conclude $\bar{a} \succ^* \bar{b}$, we may need a larger menu A that contains a "choice fixer" $c \in P_1(\bar{a})$ and a "higher reference setter" $d \in P_2(\bar{a})$. The figure also graphically illustrates Theorem 2, in particular that we observe $\bar{a} \succ^* \bar{b}$ whenever $r(\bar{a}) > r(\bar{b})$, as long as $P_1(\bar{a})$ and $P_2(\bar{a})$ are nonempty, which is a quite weak condition.

The nondegeneracy condition is also a quite weak condition in general, because it holds generically if $\dim(Z) \geq 4$. To see the intuition, note that $\mathcal{P}(x)$ is characterized by three linear inequalities (involving r, w, and $u + \beta w$) and that $\mathcal{S}(x)$ is also characterized by three linear inequalities (involving r, w, and $u+w$), as the above discussion of Figure S1a suggests. Thus, as long as the coefficient matrices (of dimension $3 \times (\dim(Z) - 1)$, because the probabilities must sum to one) have a rank of three, which holds generically, they are nonempty.

Figure S2 illustrates how we can elicit $\bar{a} \succ_r \bar{b}$ when $\bar{a}$ cannot satisfy $\bar{a} \succ^* \bar{b}$ even though data are generated by a PS preference with $r(\bar{a}) > r(\bar{b})$. Figure S2a depicts the indifference curves of the same PS preference as in Figure S1b. However, because $\bar{a}$ is on the boundary of Δ, $P_2(\bar{a})$ is empty, and we cannot establish $\bar{a} \succ^* \bar{b}$ with any $A \ni \bar{b}$. In words, when $\bar{a}$ is a unique reference alternative at $A \cup \{\bar{a}\}$ (i.e., $r(\bar{a}) > r(y)$ for all $y \in A$), the reference point is necessarily higher than that at A, so $A \cup \{\bar{a}\} \succ A$ does not occur as long as $\bar{a}$ is unchosen there.

However, we can still conclude $r(\bar{a}) > r(\bar{b})$ by using some $c \in \mathrm{int}(\Delta)$ such that $r(\bar{a}) > r(c) > r(\bar{b})$, as Figure S2b demonstrates. First, we can elicit $c \succ^* \bar{b}$ by Lemma 18 (note also that $P_1(c)$ and $P_2(c)$ are nonempty). Moreover, we cannot have $c \succ^* \bar{a}$, as c cannot set a reference point at A whenever $\bar{a} \in A$. Thus, we can conclude $\bar{a} \succ_r \bar{b}$ via Definition 1(ii-b).

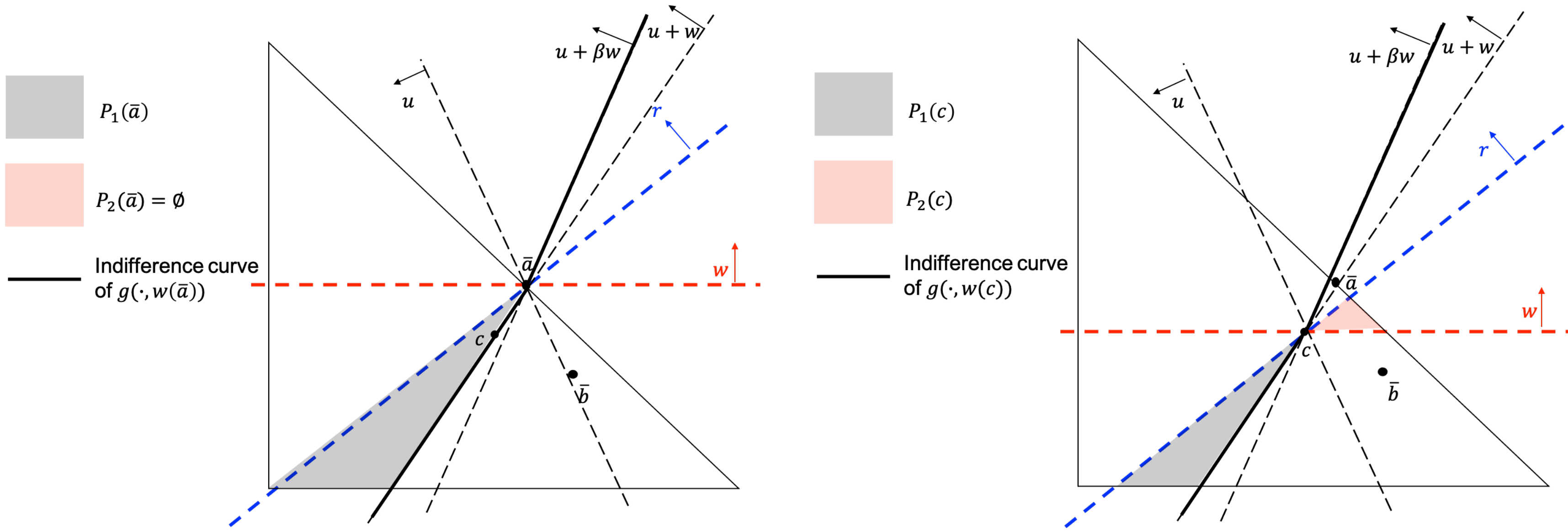
Figure S1: Nondegeneracy and Weak Nondegeneracy



Notes: Panel (a) presents an example of a PS preference that satisfies nondegeneracy. Panel (b) presents an example of a PS preference that satisfies weak nondegeneracy but not nondegeneracy, because $\mathcal{P}(\bar{a})$ is empty for all $\bar{a} \in \Delta$. Each dashed or solid straight line represents an indifference curve of u, w, $u + w$, $u + \beta w$ or r, with an arrow indicating the increasing direction of the utility function. The bold solid line kinked at x in Panel (a) (at $\bar{a}$ in Panel (b)) denotes the indifference curve of the function $g(\cdot, w(x))$ ($g(\cdot, w(\bar{a}))$) defined in Lemma 10. In Panel (a), the black and red shaded areas depict $\mathcal{P}(x)$ and $\mathcal{S}(x)$, respectively, defined in Eq. (4) and (5) in Section 3.1. In Panel (b), the black and red shaded area depicts $P_1(\bar{a})$ and $P_2(\bar{a})$, respectively, defined in Lemma 17. See the text in Appendix S.D for details.

Figure S2: Reference Elicitation on the Boundary

(a) Non-existence of $d \in P_2(\bar{a})$ (b) Mediating alternative $c \in \text{int}(\Delta)$



Notes: Panel (a) presents an example of alternatives $\bar{a}, \bar{b} \in \Delta$ such that $r(\bar{a}) > r(\bar{b})$ but $\bar{a} \not\succ^* \bar{b}$. Panel (b) illustrates how we can establish $\bar{a} \succ_r \bar{b}$ via Definition 1(ii-b) using a mediating alternative $c \in \text{int}(\Delta)$. Each dashed or solid straight line represents an indifference curve of u, w, $u + w$, $u + \beta w$ or r, with an arrow indicating the increasing direction of the utility function. The bold solid line kinked at $\bar{a}$ in Panel (a) (at c in Panel (b)) denotes the indifference curve of the function $g(\cdot, w(\bar{a}))$ $(g(\cdot, w(c)))$ defined in Lemma 10. In Panel (a), the black shaded area depicts $P_1(\bar{a})$ defined in Lemma 17. In Panel (b), the black and red shaded areas depict $P_1(c)$ and $P_2(c)$, respectively. See the text in Appendix S.D for details.