A Theory of Pretense in Public Goods Provision

Huy Nguyen*

November 17, 2017

Abstract

A player decides to help, bystand, or pretend to help in providing a public good in two games: the volunteer's dilemma and the public goods game. Pretending does not contribute, but it costs less than helping and can confer prestige. If actual contribution is less than claimed contribution, some claimants may be doubted as fakes and shamed. When pretense is possible, both the individual's chance to help and the expected level of good provision are weakly less than when pretense is not possible. Whether pretense occurs does not depend on group size. Pretenders dilute the prestige from helping and discourage actual helpers. If pretense causes negative externalities, an organization would actually benefit from anonymizing contributors. Introducing authenticated help at a premium can eliminate pretense. Extensions on asymmetry and incomplete information reveal that equilibria can exist where help, bystand, and pretend are all played.

Keywords: pretense, volunteer's dilemma, public goods game, information, signaling, social image, altruism

JEL Classifications: D71, D74, D83, H41

*University of California, Santa Barbara; huy@umail.ucsb.edu. I express gratitude to my advisor Theodore Bergstrom and committee members Gary Charness and Peter Kuhn for their exemplary mentoring. I thank UCSB for the Graduate Opportunity Fellowship 2015-2016, which funded this research. I appreciate Rodney Garratt, Zachary Grossman, Lones Smith, Charles Stuart, Lissette Wilensky, and my colleagues at UCSB for helpful comments.

1 Introduction

Recent developments in economic theory (Bénabou and Tirole 2006; Andreoni and Bernheim 2009) have emphasized the importance of image motivation.[1] People derive satisfaction from appearing virtuous to others, as a positive image confers real benefits such as influence and access to resources.[2] An overlooked aspect of image motivation is that it incentivizes pretense. If we value what others think, then appearing to be helpful can be just as good as actually helping. In fact, while pretense has been studied in some contexts like quality uncertainty or entry deterrence, it has not received much attention in public economics.[3] Motivated by these considerations, my paper explores the effects of social image concerns in a fundamental economic context—public good provision—when pretense is possible.

I introduce a player's decision to help, bystand, or pretend to help in providing a public good in two classic games: the volunteer's dilemma and the public goods game. Pretending does not contribute, but it costs significantly less than helping and can confer prestige. Conversely, a player suspected to be pretending or outright exposed suffers a shame cost. This pretense and prestige dynamic occurs in many real-life scenarios: U.N. states endorsing noble goals but taking little or no action (e.g. 2015 Millennial Development Goals, Paris Agreement); charities spending money primarily on advertisement; people registering as a blood/organ/marrow donor but later reneging; social media users verbalizing support as substitute for real aid (e.g. via retweets or Facebook profile frames).

Such pretense may hinder or crowd out real help, so it is important to understand the role of pretense in public goods provision. In each of these examples, time, distance, or bureaucracy can make distinguishing a helper from a pretender difficult, if not impossible. Indeed, anyone in modern society has almost certainly witnessed an instance where a CEO, politician, coworker, or neighbor extols a virtue but does not follow their words. A formal model can help us identify the conditions conducive to pretense and the equilibrium amount of pretending. Consequently, we might design countermeasures to reduce pretense and improve good provision.

Thus I ask: How do potential contributors behave when pretending is an option? How do the equilibrium proportions of helpers, bystanders, and pretenders vary with group size, prestige, or asymmetric preferences? What are the consequences of pretense, and how might an organization mitigate them?

In Section 2, I explore pretense in the context of a volunteer's dilemma game. In addition to help or bystand, players can instead pretend to help in order to gain a *prestige* benefit.

[1] In the dictator game, one player of two decides how to split a sum of money. A 'rational' player takes all, yet many split 50-50 not only out of fairness but also because they wish to be *perceived* as fair by the observers.

[2] Indeed, Maslow's (1943) hierarchy of needs lists self-esteem and self-respect as major motivations after physiological needs.

[3] A person may disclose or mimic strength to deter conflict. In an I.Q. contest, players signaled over-confidence to deter entry and under-confidence provoke entry (Charness et al. 2013).

Pretending is costly but still much cheaper than helping.[4] Initially, the audience is naïve and, if the good is provided, confers prestige to everyone who appeared helpful. If the good is not provided, all claimants are exposed as fakes and suffer shame. Helpers also experience *warm glow*, a psychological benefit, which bystanders and pretenders do not experience.[5]

In Section 2.1 (existence and uniqueness), I note key assumptions and solve for pure- and mixed-strategy equilibria. In pure strategy equilibria, one player helps while all others pretend. In mixed-strategy equilibria, players mix either on help/pretend or help/bystand; which of the two they select depends on costs and benefits, but surprisingly not on group size. Broadly two types of societies exist, one that sufficiently values image and one that does not. In other words, helping is always an option in both societies, but the alternative action is pretend in the image-oriented society and bystand in the other. When pretending is an option, the individual's chance to help and the group's chance of provision are both *weakly less than* when pretending is not possible.

In Section 2.2 (comparative statics), I show that an individual's chance to help increases with benefits and decreases with costs; as group size increases, it converges to zero. This is consistent with the original game. Three new dimensions of warm glow, prestige, and shame add novel results. Warm glow encourages only helping. Prestige incentivizes helping but also incentivizes pretending. Shame deters pretending. When pretense causes negative externalities, an organization may—surprisingly—prefer to reduce or even eliminate the prestige associated with helping. While an organization may not be able to directly choose prestige, they could instead control information structures like keeping donors anonymous.

In Section 2.3 (endogenous prestige), I consider a variant where players are sophisticated and aware that some who appear helpful may be pretenders. Consequently, prestige is now endogenous rather than exogenous. A key finding here is that the presence of pretenders discourages players from helping because their prestige is diluted by fakes, much like a lemons market. Discouraged helpers are thus willing to pay a premium to distinguish their help as authentic, provided the premium is not too costly. If an organization could introduce an authentication option, this would also minimize pretense.

In Section 3, I explore pretense in the context of a simplified public goods game where players can help by donating 1 unit toward the pot, bystand, or pretend to help. As before, pretending costs less than helping, appears indistinguishable, and can confer prestige. Prestige in the public goods game is always endogenous, and therefore discounted, because everyone can see the difference between the pot size and the number of help claims.

In Section 3.1 (existence and uniqueness), I prove that there exists up to two symmetric,

[4] Mimicry costs some effort to convince others. Further, most people are averse to lying because it exacts some cognitive or emotional toll. This is true even when lying is undetectable, absent strategic motives, or leads to improved monetary outcomes for everyone (Gneezy et al. 2013; Abeler et al. 2014; Erat and Gneezy 2011).

[5] Egoistic utility from giving, as opposed to *pure* altruism for the recipient's welfare. The colloquial term was first coined by Andreoni in 1989.

pure-strategy Nash equilibria—all help and/or all bystand—and zero asymmetric, pure-strategy equilibria. All-pretend cannot be an equilibiurm. Essentially, players treat the game environment as a shop to 'buy' prestige and warm glow. There also exists a symmetric, mixed-strategy equilibrium, which takes one of three cases: players mix on help/bystand and no one pretends; players mix on help/pretend and no one bystands; or players mix on all three actions. A three-way equilibrium exists in the public goods game because a marginal helper increases the belief that a claimant is a helper while a marginal pretender decreases this belief.

In Section 3.2 (mixed strategies in two- & three-player games), I calculate the mixed-strategy provision rates for small partnerships. This rate varies depending on whose good opinion players care to impress: only other contributors, all others equally, or no one at all? For some combinations of group size $n \in \{2, 3\}$ and weighting schemas, a mixed strategy does not exist. For $n = 2$, pretending cannot be an action because pretenders face immediate exposure. For $n = 3$, if the value of prestige is sufficiently large, the expected good size is greater when players care to impress only other contributors. On the other hand, a culture that weakly values prestige contributes more from weighting everyone's opinion equally.

In Section 3.3 (incomplete information), I treat players as having (i.i.d.) random but private valuations of both prestige and warm glow.[6] In equilibrium, players select one of three actions based on their preferences. Generally, players who care little about looking good bystand while those who care much about feeling good help; those who value looking good, but not feeling good, pretend. An equilibrium consisting of all three actions can thus be sustained this way. Two graphical examples illustrate this partition, one for the volunteer's dilemma and one for the public goods game.

In Section 4, I discuss my findings in the context of prominent literature and suggest subsequent avenues of exploration. To my best knowledge, my paper is the first to model pretense in public goods provision.

2 Pretending Volunteer's Dilemma

The volunteer's dilemma (Diekmann 1985) is an n-player game in which each player i benefits b_i from the provision of a public good if at least one player volunteers to pay the cost c_i, where $0 < c_i < b_i$. Each gains from the good's provision but prefers to let someone else pay for the good, a phenomenon known among social psychologists as *diffusion of responsibility* or the *bystander effect* (Darley and Latane 1968). Examples include helping an injured victim, taking out the trash, or enforcing a social norm such as confronting a smoker at a restaurant.[7] Various authors have since covered extensions like asymmetric costs (Diekmann

[6]One might argue that players should vary in the cost of helping, but warm glow serves the same purpose.

[7]Animals like penguins and marmots are known to volunteer serving as a lookout for predators (Dawkins 1976).

1993), incomplete information (Weesie 1994), or cost sharing (Weesie and Franzen 1998). The studies closest to my paper incorporate behavioral elements like warm glow (Andreoni 1990; Bergstrom et al. 2015) and prestige from helping (Harbaugh 1998; Andreoni and Petrie 2004) but do not model the strategic *pretense* of helping.

A problem occurs, and n witnesses simultaneously decide whether to help (H), bystand (B), or a third option: pretend to help (P). Helping costs c and, if at least one person helps, confers three benefits: first, everyone consumes a material benefit $b > 0$; secondly, helpers and pretenders earn a prestige $g > 0$; thirdly, helpers enjoy a warm glow $w > 0$. Thus, I decompose the helper's benefit into material, prestige, and warm glow components.

Bystanding costs zero and contributes nothing; a bystander receives the material benefit b only if someone else helped. Pretending costs less than helping but more than zero, $0 < c_P < c$, and confers prestige g only if someone else helped.[8] If no one helped, the public good is not provided so all pretenders are exposed; exposed pretenders gain no prestige and suffer a shame cost s. A player's payoff from each action, conditional on whether or not someone else offers to help, is summarized below:

	u_i (if someone else helps)	u_i (if no one else helps)
help (H)	$b + g + w - c$	$b + g + w - c$
bystand (B)	b	0
pretend (P)	$b + g - c_P$	$-c_P - s$

where

$b =$ material benefit	$g =$ prestige	$w =$ warm glow
$c =$ cost	$c_P =$ pretend cost	$s =$ shame

Table 1. Actions and Payoffs

All parameters have positive value. H is the best response if no one else helps, so it cannot be strictly dominated. An assumption is needed to ensure that neither B nor P is weakly dominated.

Assumption 1. $0 < c_P < g < g + w < c < b$.

The assumption that $c < b$ is the original volunteer's dilemma condition. That $g + w < c$ means prestige and warm glow alone, without any material benefit, do not incentivize a bystander to help; else, H weakly dominates B. $c_P < g$ means that pretending is profitable if someone else helps; else, B weakly dominates P.

Initially, I model prestige g as fixed and exogenous. In later sections, I explore heterogeneity, where $g_i \sim U[0, \bar{g}]\ \forall\ i$, and endogenous prestige, where g is discounted by the Bayesian probability an observed helper is actually helping.

[8] Prestige here is an abstract, non-rival benefit that makes more sense duplicated than divided equally. All else equal, three volunteers generate more prestige than one volunteer.

2.1 Existence and Uniqueness

Proposition 1. *Given Assumption 1, there exist n pure-strategy Nash equilibria where exactly one player helps and everyone else pretends.*

Proof. Given that all others pretend—and therefore do not help—the one helper receives positive utility $u_i = b+g+w-c > 0$. Deviating from action H would cause $u_i \leq 0$ because the public good would not be provided. Given that someone else helps, action P yields the greatest utility. □

Pretenders not only free ride the material benefit, but also share in the prestige.[9]

Proposition 2.
a. *Given Assumption 1, exactly one of two symmetric, mixed-strategy Nash equilibria exists: 1) players mix on help/bystand and no one pretends; or 2) players mix on help/pretend and no one bystands.*
b. *When* $\frac{c-g-w}{b} < \frac{c-c_P-w}{b+g+s}$*, only the help/bystand equilibrium exists. When* $\frac{c-g-w}{b} > \frac{c-c_P-w}{b+g+s}$*, only the help/pretend equilibrium exists. When* $\frac{c-g-w}{b} = \frac{c-c_P-w}{b+g+s}$*, either the help/pretend equilibrium exists or the help/bystand equilibrium exists.*
c. *A three-way, symmetric, mixed-strategy Nash equilibrium where players mix on all three actions–help, bystand, and pretend–with positive probability does not exist.*

Proof. See Appendix. □

The three-way equilibrium does not exist because in equilibrium, $E(u_i|H) = E(u_i|B) = E(u_i|P)$ occurs with measure zero in parameter space. In a later subsection with endogenous prestige, a three-way mixing is possible because endogenous prestige depends on the composition of helpers and pretenders. The individual's probability of helping and the group's probability of provision in each equilibrium are summarized below:

	individual helps (p_H^*)	group provides ($\hat{P}^*$)
help/bystand (H/B)	$1-\left(\frac{c-g-w}{b}\right)^{1/(n-1)}$	$1-\left(\frac{c-g-w}{b}\right)^{n/(n-1)}$
help/pretend (H/P)	$1-\left(\frac{c-c_P-w}{b+g+s}\right)^{1/(n-1)}$	$1-\left(\frac{c-c_P-w}{b+g+s}\right)^{n/(n-1)}$

Table 2. Equilibrium Probabilities

While the provision chance decreases in group size n, whether the help/bystand or help/pretend equilibrium exists does not depend on n. This condition also appears in Table

[9]Whether this is socially wasteful is a matter of perspective, as the resources expended to pretend are compensated by the utility gained from shared prestige.

2; equivalently, only the equilibrium with the smaller probability to help exists. Proposition 2.b implies that when pretending is possible, both the individual's probability of helping and the group's probability of provision are *weakly less than* if pretending were not an option.

2.2 Comparative Statics

Proposition 3
a. *Given Assumption 1, within the help/bystand and within the help/pretend equilibria,* $\partial p_H^*/\partial b$, $\partial p_H^*/\partial g$, $\partial p_H^*/\partial w > 0$ *and* $\partial p_H^*/\partial c < 0$. *Within only the help/pretend equilibrium,* $\partial p_H^*/\partial c_P$, $\partial p_H^*/\partial s > 0$.
b. *As* $n \longrightarrow \infty$, $p_H^* \longrightarrow 0$ *and* $\hat{P}^* \longrightarrow 1 - max\left\{\frac{c-g-w}{b}, \frac{c-c_P-w}{b+g+s}\right\}^{n/(n-1)}$.

Proof. Refer to Table 2. Marginal effects are deduced from each parameter's sign (+ or −) and position (numerator or denominator). $c_P < g < g + w < c$ implies the numerators are positive, $c-g-w > 0$ and $c-c_P-w > 0$. Then, $c < b$ implies the denominators are greater, so $p_H^* \in (0, 1)$. Asymptotics are also deduced by each exponent's convergence (either 0 or 1) as $n \longrightarrow \infty$. □

Asymptotic probabilities p_H^* and $\hat{P}^*$ are consistent with the original volunteer's dilemma. Notice that the expected utility is always $E(u_i) = E(u_i|H) = b + g + w - c$, which is unaffected by c_P or by whether one equilibrium exists or the other.

Proposition 4. *Given Assumption 1, when the prestige parameter g is sufficiently large and shame parameter s is sufficiently small, only the help/pretend equilibrium exists.*

Proof. Proposition 2.b states that only the help/pretend equilibrium exists when $\frac{c-g-w}{b} > \frac{c-c_P-w}{b+g+s}$. Rearranging this expression in terms of prestige g yields:

$$0 < g^2 + (b + w - c + s) \cdot g + s(w - c) - b \cdot c_P \quad (1)$$

The right side of Inequality 1 is a convex parabola $f(g)$ with a negative vertical intercept at $f(0) = s(w - c) - b \cdot c_P < 0$. This means somewhere in the domain $g > 0$, there exists a unique threshold g^T where $f(g^T) = 0$ crosses the horizontal axis. Specifically, this crossing point occurs at $g^T = (-B + \sqrt{B^2 - 4C})/2$ where $B = b+w-c+s$ and $C = s(w-c)-b\cdot c_P$. When prestige exceeds this threshold, the help/pretend equilibrium exists.

Rearranging Inequality 1 in terms of shame s shows that the help/pretend equilibrium exists when:

$$0 > (c - g - w)s + [-g^2 + g(c - b - w) + b \cdot c_P] \quad (2)$$

This is a linear function of s. Assumption 1 implies the slope is positive and the intercept sign is ambiguous. If the intercept is non-negative, then s does not affect equilibrium selection and all players prefer mixing H/N. If the intercept is negative, then there exists a horizontal-axis crossing in $s > 0$ that determines the threshold between mixing H/P versus H/N. The intercept is negative when g is sufficiently large. □

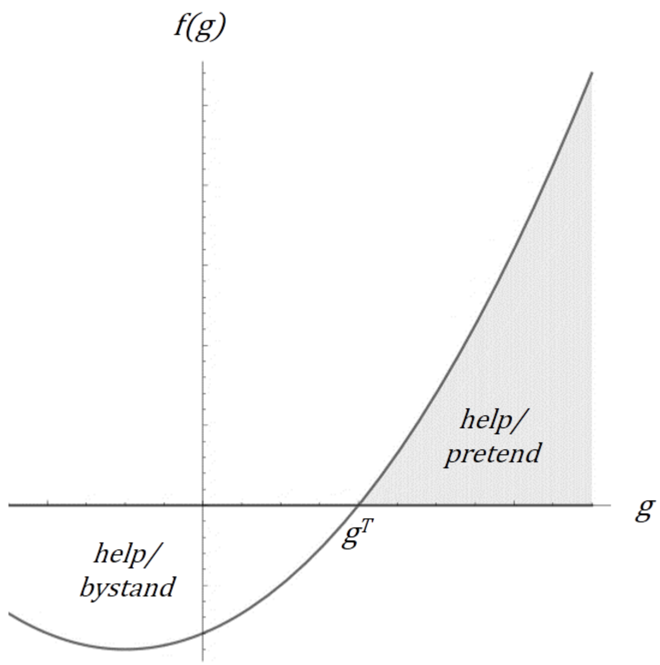

Fig. 1. Large Values of Prestige Induce a Help/Pretend Equilibrium

Prestige encourages contribution but also pretense. Societies that strongly value prestige develop a norm to appear helpful, even if that means pretending. Societies that care little for prestige develop a norm to bystand if one does not help. In some cases, pretense may waste social resources. For example, registering as an organ or marrow donor but reneging when asked can disrupt a clinic's plan. When pretense imposes negative externalities, society may fare better if people value prestige less!

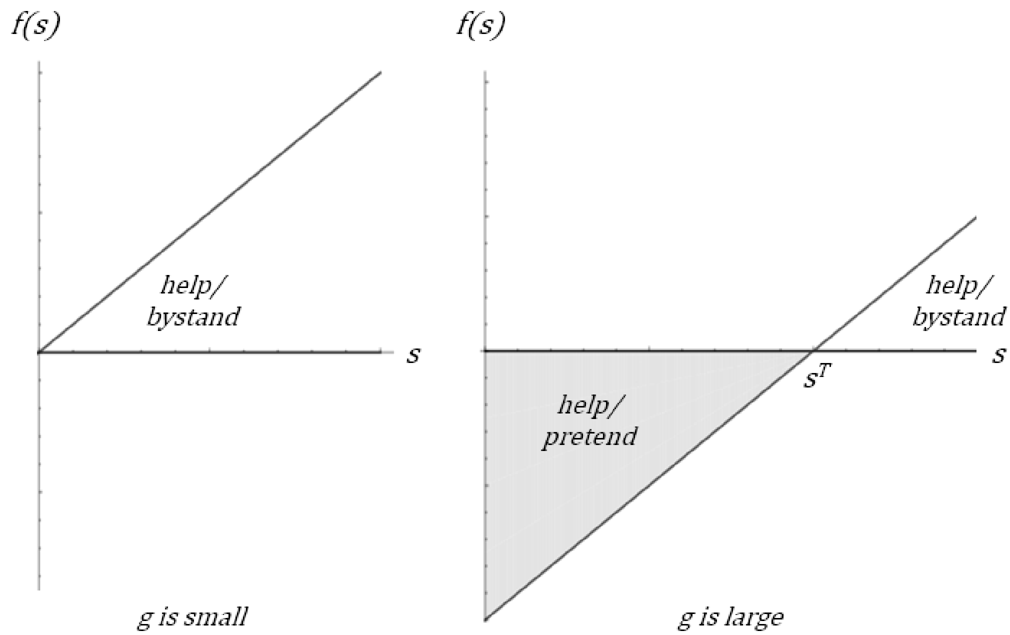

Fig. 2. Large Values of Shame Deter a Help/Pretend Equilibrium

When prestige is small, shame does not determine which type of equilibrium exists. When prestige is large, a corresponding large value of shame deters pretending.

2.3 Endogenous Prestige

Assumption 2. g is discounted by $p_H^*/(p_H^* + p_P^*)$.
I now relax the assumption that all claimants earn prestige when the public good is provided. If audiences are sophisticated, they would discount the prestige g by the Bayesian probability an observed helper is actually helping. This discount factor is $p_H^*/(p_H^* + p_P^*)$, where p_H^* and p_P^* are the equilibrium probabilities that a player helps or pretends. This makes prestige endogenous, affected by the composition of helpers versus pretenders in equilibrium.

Proposition 5. *Given Assumptions 1-2, exactly one of three symmetric, mixed-strategy Nash equilibria exists: 1) players mix on help/bystand and no one pretends; 2) players mix on help/pretend and no one bystands; or 3) players mix on all three actions—help, bystand, and pretend–with prositive probability.*

Proof. See Appendix. □

Proposition 6. *Given Assumptions 1-2, both a player's probability of helping and the group's probability of provision in the help/pretend equilibrium are strictly less than under Assumption 1 only.*

Proof. When the equilibrium mixes only on actions H/B, any helpers are clearly helpers. When the equilibrium mixes only on H/P, prestige g is discounted. In this case, $p_P^* = 1 - p_H^*$ so the discount factor simplifies to p_H^*. Equating $E(u_i|H) = E(u_i|P)$ simplifies to:

$$\frac{c - c_P - w}{b + g \cdot p_H + s} = (1 - p_H)^{n-1} \in (0, 1) \tag{3}$$

A discounted $g \cdot p_H < g$ makes the denominator smaller, the fraction larger, and thus p_H smaller on the right side. □

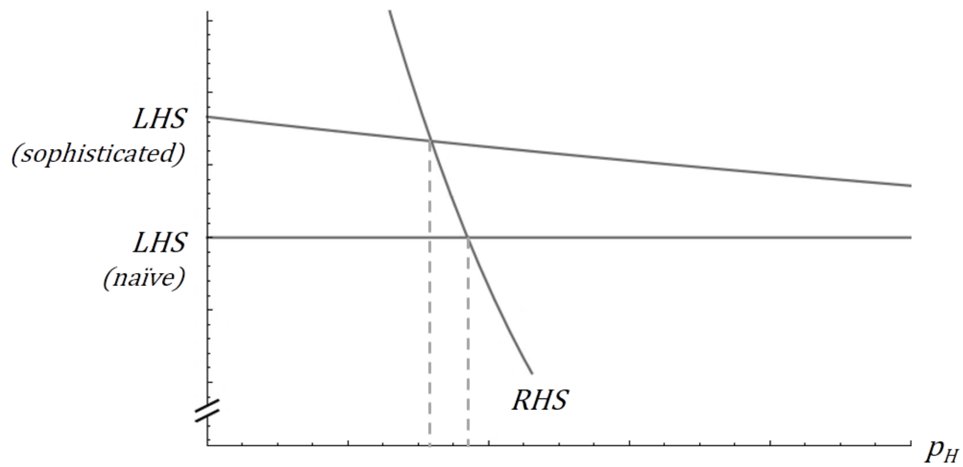

Fig. 3. Audience Sophistication Reduces Provision

Fig. 3 illustrates how Equation 3 determines equilibrium p_H^* depending on whether the left side uses endogenous prestige $g \cdot p_H$ (sophisticated audience) or exogenous prestige g (naïve audience). The presence of pretenders has a double-negative effect on provision. First, pretenders do not contribute. Second, pretenders dilute prestige, which discourages helpers who realize their help may be doubted as fake. This discouragement shares a structural similarity to the market for lemons (Akerlof et al. 1970).[10]

In an environment where audiences are sophisticated and some players pretend, helpers feel frustrated to be pooled with pretenders. Authentication is relevant only in a help/pretend equilibrium. In a help/bystand equilibrium, any help is already believed. Suppose now those helpers have the option to authenticate (A) their help. Authenticating costs more than helping ($c_A > c$), but it appears distinct from any other action and guarantees that the helper earns an undiluted prestige $g > g \cdot p_H$.

[10]Akerlof, Spence, and Stiglitz shared the 2001 Nobel Prize in Economics for their work in asymmetric information, of which the market for lemons was a central idea.

Proposition 7.
a. *Given Assumptions 1-2, when authenticating is an option, exactly one of two symmetric, mixed-strategy Nash equilibria exists: 1) players mix on authenticate/bystand and no one pretends or unauthenticated-helps; or 2) players mix on unauthenticated-help/pretend and no one bystands or authenticates.*
b. *If the premium is less than the prestige loss, the authenticate/bystand equilibrium exists. If the premium is greater than the prestige loss, the help/pretend equilibrium exists.*

Proof. Helpers who are pooled with pretenders authenticate if $E(u_i|A) > E(u_i|H)$. This occurs when:

$$(c_A - c) < g(1 - p_H^*) \tag{4}$$

The left side of Inequality 4 is the extra cost to authenticate while the right side is the extra benefit from restoring full prestige. If marginal benefit outweighs marginal cost, players pay this premium. If so, players mix on actions A/B (i.e. they play H and P with probability zero). If the premium is too costly, no one plays A and the equilibrium reverts to mixing on H/P. Mixing on A/P cannot be an equilibrium because players can deduce any claimant who does not authenticate help must be a pretender. The case where both sides of Inequality 4 are exactly equal occurs on a set of measure zero in parameter space. □

When pretense is widespread and problematic, an organization may consider introducing an authentication option at a premium to minimize pretending behavior. This is an alternative solution to minimizing prestige itself via, for example, anonymizing donors.

3 Pretending Public Goods Game

The public goods game (Samuelson 1954) is a scenario where each of n members in a group can voluntarily contribute some or all of their private resources toward a common pool. This sum contribution is then multiplied by a given fraction $m \in (1/n, 1)$ and awarded identically to each member. The constraint $m > 1/n$ means that, should one contribute, the gain in group welfare outweighs the loss in individual welfare. That $m < 1$ means players face a strategic tension between a lesser, private welfare versus a greater, public welfare. Else, donating 1 unit would immediately return $m \geq 1$ to the donor. The dominant strategy in this game is to donate zero, which shares some similarity to an n-player prisoner's dilemma. Regardless what others do, donating always reduces one's own material payoff.

Numerous experimental studies, however, show that people do donate for a variety of reasons: altruism, in that helpers feel a warm glow, but only when others are also helping; efficiency, in that helpers seek to increase the economic pie; or confusion, in that helpers choose arbitrarily (Andreoni 1995; Houser and Kurzban 2002). Extensions reveal that the threat of punishment (Fehr and Gächter 2000) or exclusion (Cinyabuguma et al. 2005; Charness and Yang 2014) can pressure players to donate at the social norm.

There are two important differences between the pretending volunteer's dilemma and the pretending public goods game: marginal returns and beliefs. In the first game, a single helper provides a fixed-size public good. If several players claim to help and the group sees that good provision succeeded, it is unclear how many claimants helped. This merited the discussion on naïve versus sophisticated audiences. In the second game, each marginal helper increases the size of the public good. The group sees both the number of claimants and the actual contribution. Players immediately know the fraction of helpers among claimants, especially when each has private information about their own action.

A group of n players convenes over a public good project. Each player simultaneously chooses one of three actions: 1) help (H), which costs one unit and contributes one unit ($c_i = x_i = 1$) toward the project; 2) bystand (B), which costs nothing and contributes nothing ($c_i = x_i = 0$), or pretend (P) to help, which costs less than helping ($0 < c_i = c_P < 1$) but contributes nothing ($x_i = 0$). The sum contribution $X = \sum_{i=1} x_i$ in the project, colloquially called the pot, is then scaled by a known multiplier $m \in (1/n, 1)$ and enjoyed identically by each player.[11] This composes the material benefit. Helping costs 1 but returns some material benefit via m. Let $c_H = 1 - m$ then denote the (discounted) cost of helping.

Players also value the good opinion of others, though not necessarily with equal weights. People prefer the company of altruistic and productive others, so naturally they value most the opinion of helpers. To a lesser degree, they also value the opinion of bystanders. They value least, or perhaps not at all, the opinion of vile pretenders. Thus, the opinion weight of another player j is $\alpha_j \in \{\alpha_H, \alpha_B, \alpha_P\}$, where $1 \geq \alpha_H \geq \alpha_B \geq \alpha_P \geq 0$. A player i who claims to help receives an expected prestige benefit $g \cdot E(\alpha_j \cdot \beta_{ij})$ and suffers an expected shame cost $s \cdot E[\alpha_j(1 - \beta_{ij})]$, where $\beta_{ij} \in \{\beta_{HH}, \beta_{HB}, \beta_{HP}, \beta_{PH}, \beta_{PB}, \beta_{PP}\}$ is the conditional probability that claimant i is a helper from player j's perspective, which is a function of their respective actions. The belief that a bystander is a helper from anyone's perspective is zero. That is, $\beta_{BH} = \beta_{BB} = \beta_{BP} = 0$.

For example, suppose four players each possessing one coin are asked to donate. One player abstains while three others walk up to a box and gesture a donation. After all players have (simultaneously) acted, the donation box is revealed to contain only two coins! From a helper's perspective, each other claimant has a $\beta_{HH} = \beta_{PH} = 1/2$ chance to be a helper. From a bystander's perspective, each claimant has a $\beta_{HB} = \beta_{PB} = 2/3$ chance to be a helper. From the pretender's perspective, the other claimants were certainly helpers, so $\beta_{HP} = 1$. These probabilities are not independent. If some claimants are believed to be pretending, the other claimants must be helping. These inferences are clear enough that we can treat all players as sophisticated.[12]

[11] I normalize cost and contribution to 1 because m already acts as a scalar.

[12] It would be inconceivably naïve to see two coins and believe all three were helpers.

A helper also enjoys a warm glow $w > 0$. A bystander receives neither prestige nor warm glow. If each player helps, pretends, or bystands with respective probabilities p_H, p_P, and p_B, player i's expected payoffs from each action is summarized below:

	material profit	expected prestige	warm glow
help (H)	$(n-1)p_H \cdot m - c_H$	$+(g+s)E(\alpha_j \cdot \beta_{Hj}) - s \cdot E(\alpha_j)$	$+w$
bystand (B)	$(n-1)p_H \cdot m - 0$	0	0
pretend (P)	$(n-1)p_H \cdot m - c_P$	$+(g+s)E(\alpha_j \cdot \beta_{Pj}) - s \cdot E(\alpha_j)$	0

where

m = multiplier	g = prestige	w = warm glow
x_j = player j's contribution	c_P = pretend cost	s = shame
α_j = j's belief weight	β_{Hj} = j's belief if i helps	β_{Pj} = j's belief if i pretends

Table 3. Actions and Payoffs

Assumption 3. $0 < c_P < c_H$; $1/n < m < 1$; $0 \leq \alpha_P \leq \alpha_B \leq \alpha_H \leq 1$.

3.1 Existence and Uniqueness

Proposition 8.

a. *Given Assumption 3, if a symmetric, pure-strategy Nash equilibrium exists, then it is one of three cases: 1) all-help is the only equilibrium; 2) all-bystand is the only equilibrium; or 3) all-help and all-bystand are both equilibria. All-pretend cannot be an equilibrium.*

b. *The conditions that determine which equilibria exist are summarized by the truth table:*

$(g \cdot \alpha_H + w \geq c_H) \wedge$ $[(g+s)\alpha_H/(n-1) + w \geq c_H - c_P]$	true	true	false	false
$c_H \geq g \cdot \alpha_B + w$	true	false	true	false
$PSNE$	H and B	H only	B only	none ($MSNE$)

Proof. An all-help equilibrium exists if $H \succeq B$ and $H \succeq P$ given $p_H = 1$. When all others are helping, $\alpha_j = \alpha_H \ \forall\ j$, $\beta_{iH} = \beta_{HH} = 1$ if i helps, and $\beta_{iH} = \beta_{PH} = (n-2)/(n-1)$ if i pretends. $E(u_i|H) \geq E(u_i|B)$ implies:

$$(g+s)E(\alpha_j \cdot \beta_{Hj}) - s \cdot E(a_j) + w \geq c_H \tag{5}$$

$$\implies g \cdot \alpha_H + w \geq c_H \tag{6}$$

$E(u_i|H) \geq E(u_i|P)$ implies:

$$(g+s)E[\alpha_j(\beta_{Hj} - \beta_{Pj})] + w \geq c_H - c_P \tag{7}$$

$$\implies (g+s)\alpha_H/(n-1) + w \geq c_H - c_P \tag{8}$$

An all-bystand equilibrium exists if $B \succeq H$ and $B \succeq P$ given $p_B = 1$. When all others are bystanding, $\alpha_j = \alpha_B \ \forall \ j$, $\beta_{iB} = \beta_{HB} = 1$ if i helps, and $\beta_{iB} = \beta_{PB} = 0$ if i pretends. $B \succeq P$ here is always true because a pretender would be immediately exposed when everyone sees the pot is empty. The pretender suffers $-c_P - s \cdot \alpha_B < 0$ and would be better off bystanding.[13] $E(u_i|B) \geq E(u_i|H)$ implies:

$$c_H \geq (g+s)E(\alpha_j \cdot \beta_{Hj}) - s \cdot E(a_j) + w \tag{9}$$

$$\implies c_H \geq g \cdot \alpha_B + w \tag{10}$$

An all-pretend equilibrium cannot exist because everyone's payoff is $-c_P - s \cdot \alpha_P < 0$. Even if everyone were shameless ($\alpha_P = 0$), $-c_P < 0$ is still negative. In this case, each player would do better to bystand and receive zero. □

If $w \geq c_H - c_P$, Inequality 8 is true independent of n; else, n must be sufficiently small such that $H \succeq P$. This threshold is:

$$n \leq \frac{(g+s) \cdot \alpha_H}{c_H - c_P - w} + 1 \tag{11}$$

Essentially, the option of guaranteed prestige and warm glow collapses this game into a shop where players can 'buy' looking and feeling good via their contribution. This forms the pure-strategy equilibrium where everyone plays H.

Proposition 9. *Given Assumption 3, an asymmetric, pure-strategy Nash equilibrium cannot exist.*

Proof. See Appendix. □

Proposition 10.
a. *Given Assumption 3, in a two-player game, there exists a unique, mixed-strategy Nash equilibrium where players mix on help/bystand and no one pretends.*
b. *In a three-player game, exactly one of three symmetric, mixed-strategy Nash equilibria exists: 1) players mix on help/bystand and no one pretends; 2) players mix on help/pretend and no one bystands; or 3) players mix on all three actions, help, bystand, and pretend.*

Proof. See Appendix. □

[13] Because any deviation from B is to H, the off-equilibrium-path belief of anyone not bystanding must be $\beta_{iB} = \beta_{HB} = 1$. I assert not only that $\beta_{iB} = 1$ sustains the equilibrium, but also the stronger claim that the unique $\beta_{iB} = 1$ must be deduced if players are rational and know others are rational.

3.2 Mixed Strategies in Two- & Three-Player Games

Whose good opinion do players value? In this subsection, I compare equilibrium probabilities for different α_j weighting schemas that reflect, broadly, three distinct cultures. On one extreme, people care only about the opinion of helpers, so $(\alpha_H, \alpha_B, \alpha_P) = (1, 0, 0)$. In a Confucian society, for example, one's virtue is linked with benevolence or altruism toward one's community; this is known as *'ren'*. On the other extreme, they may value all opinions equally, so $(\alpha_H, \alpha_B, \alpha_P) = (1, 1, 1)$. This reflects a forgiving, egalitarian, or non-discriminatory culture; the Scandinavian societies, for example, feature the lowest Gini coefficients in the world and allow prisoners to live like regular citizens. For completeness, I examine also the case where prestige is meaningless and everyone is shameless, so $(\alpha_H, \alpha_B, \alpha_P) = (0, 0, 0)$; equivalently, $(g, s) = (0, 0)$. An image-less society is arguably a distinct 'culture', too.

culture	virtuous	egalitarian	image-less
$(\alpha_H, \alpha_B, \alpha_P)$	(1,0,0)	(1,1,1)	(0,0,0)
p_H^* (H/B)	$\frac{c_H - w}{g}$	*D.N.E.*	*D.N.E.*
p_H^* (H/P)	$\frac{2(c_H - c_P - w)}{g + s}$	*D.N.E.*	*D.N.E.*
p_H^* $(H/P/B)$	$\frac{s + \sqrt{s^2 + 2(g+s)c_P}}{g + s}$	$\frac{2(c_P + s)}{g + s}$	*D.N.E.*

Table 4. Equilibrium Provision Rates By Cultural Schema

Assumption 4. $(\alpha_H, \alpha_B, \alpha_P) = (1, 0, 0)$
Assumption 5. $(\alpha_H, \alpha_B, \alpha_P) = (1, 1, 1)$

I calculate these provision rates by inputting the α_j parameters into the system of equations in Proof 10 and solving for p_H^*. For the 'virtuous' schema, the provision rate in a help/bystand equilibrium is the same whether in a two-, three-, or even n-player game. The help/pretend and three-way equilibria exist for three, but not two, players. This is because pretending is immediately exposed with only two players. For the 'egalitarian' schema, only a three-way equilibrium exists. An 'image-less' schema produces no mixed strategies. Such a schema reverts to the original public goods game where the dominant strategy is to bystand.

Proposition 11. *The equilibrium provision rate in a three-way game is greater under Assumptions 3 & 4 than under Assumptions 3 & 5 if* $(g - s)/2 > c_P$*, lesser if* $(g - s)/2 < c_P$*, and equal if* $(g - s)/2 = c_P$*.*

Proof. $\frac{s+\sqrt{s^2+2(g+s)c_P}}{g+s} \geq\leq \frac{2(c_P+s)}{g+s} \implies (g-s)/2 \geq\leq c_P$ □

Fixing s and c_P, increasing g makes p_H^* from the 'virtuous' schema greater than from p_H^* from the 'egalitarian' schema. This suggests that when the value of prestige is sufficiently large, caring only about the opinion of other contributors improves the expected contribution level. On the other hand, a society that weakly values prestige contributes more from weighting everyone's opinion equally.

3.3 Incomplete Information

The base model assumes every player is identical. A more realistic assumption would be that players are similar in preferences, but each varies in some individual way. Because prestige and warm glow are behavioral terms introduced to the model, it is most natural to introduce heterogeneity via g and w. Let each player have fixed benefits of prestige $g_i \sim U[0, \bar{g}]$ and warm glow $w_i \sim U[0, \bar{w}]$ that are identically and independently distributed uniformly. Individuals know their own preferences but not those of others; they know only the distribution from which these preferences are drawn.

To derive an equilibrium, I use the Bayesian approach for games of incomplete information (Harsanyi 1967-1968). A strategy here maps a player i's type (g_i, w_i) to an optimal action (H, B, or P).[14] In a Bayesian equilibrium, each player best-responds to maximize expected utility given the stochastic preferences (g_j, w_j, where $j \neq i$) of other players.

Proposition 12. *Given Assumption 3, under heterogeneity and incomplete information, a unique equilibrium exists where players choose exactly one of three actions—help, pretend, bystand—based on their type.*

Proof. See Appendix. □

While the strategy functions are complex and parameter-dependent, we can visualize a mapping of types to optimal actions. A rectangle of dimensions $\bar{g} \times \bar{w}$ represents the set space for $g_i \sim U[0, \bar{g}] \times w_i \sim U[0, \bar{w}]$. This type space is divided by three concentric rays into distinct regions corresponding to the three actions—help, bystand, or pretend. Every player uses this same mapping to play a pure strategy corresponding to their given type.[15]

[14] As shown in Weesie 1994, randomization among alternatives is unnecessary for an equilibrium to exist and is thus discarded.

[15] In fact, I conjecture that a valid partition always exists for any ratio of $H : N : P$. That is, given three distinct angles of rays originating from a central point P and given a desired ratio of $H : N : P$, I claim that there always exists at least one placement of P, either inside or outside the rectangle, that partitions the type space into the desired shares. However, the geometric proof for this is beyond the scope of this paper. I simply mention this for interested mathematicians.

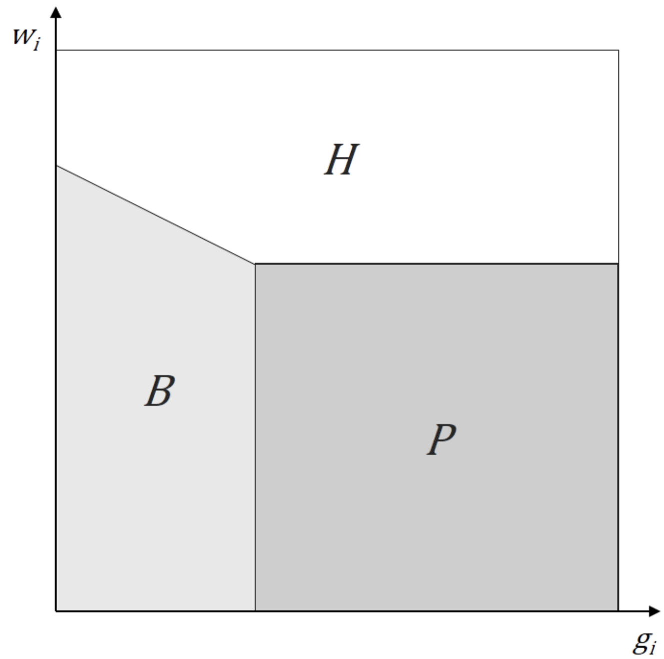

Fig. 4. Mapping from Preferences to Action (PGG)

At the indifference between B/P, any increase in prestige g_i would tip a marginal player to P. This is because warm glow does not exist in either B or P, but prestige matters in P. Likewise, at the indifference between H/B, any increase in g_i or w_i tips a marginal player to H. In fact, this threshold's slope is flatter than 45° because $\partial[E(u_i|H) - E(u_i|P)]/\partial g_i < \partial[E(u_i|H) - E(u_i|P)]/\partial w_i$. A helper receives warm glow with certainty and prestige only with some probability $\beta < 1$. The margin of H/P is separated by only w_i, with warm glow encouraging helping. Both helpers and pretenders are subject to the same judgment β so any difference must lie in w_i.

In this hypothetical illustration, the areas of H, B, and P are respectively 35%, 25%, and 40%.[16] A player's chance to help is 1 if their type falls in H and 0 otherwise. However, from an observer's perspective, it appears as if $p^* = 35\%$ per individual. For comparison, the equilibrium in the volunteer's dilemma, when audiences are sophisticated, resembles Fig. 4. When audiences are naïve, however, the equilibrium resembles Fig. 5. The H/B threshold's slope is exactly 45° because $\partial[E(u_i|H) - E(u_i|B)]/\partial g_i = \partial[E(u_i|H) - E(u_i|B)]/\partial w_i$. The H/P threshold's slope is flatter than 45° because $\partial[E(u_i|H) - E(u_i|P)]/\partial g_i < \partial[E(u_i|H) - E(u_i|P)]/\partial w_i$. Pretenders already receive g_i if the public good is provided, so it takes greater prestige than warm glow to tip a marginal pretender into helping.

[16]These examples approximate the mean populations of altruistic, selfish, and 'reluctant' types across previous studies: Lazear et al. (2012), Dana et al. (2006), Dana et al. (2007).

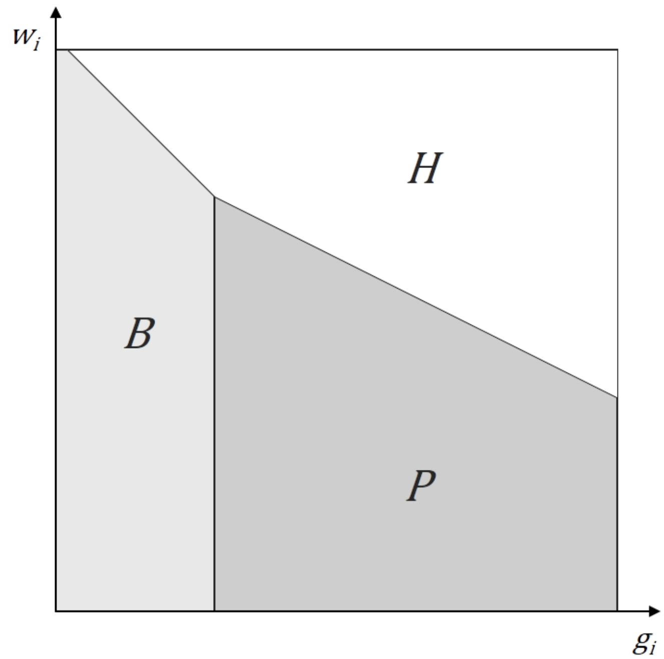

Fig. 5. Mapping from Preferences to Action (VOD, naïve)

4 Discussion

The key insight of this paper is that in many public goods contexts—economic, political, social—a 'help' signal may not correspond to a 'help' action. That is, people sometimes pretend to contribute, potentially free riding a positive image without paying the full help cost. How do decision makers behave when pretending is an option? How do group size, prestiges, or asymmetry affect outcomes? What are consequences to pretense and how might we mitigate them? Until now, this has been a gap in the public goods literature.

I fill this gap by modeling two games, the volunteer's dilemma and the public goods game, with a third alternative: pretending to help. Pretending appears identical to helping, contributes zero toward the public good, and costs significantly less than helping ($0 < c_P < c_H$). Secondly, I decompose the benefits from helping (or pretending) into material (b), prestige (g), and warm glow (w). These three factors comprise the primary motivations for contributing toward a public good. Helping confers a material benefit via a stronger public good and also via prestige and warm glow. Pretending confers prestige if one is believed to be a helper and shame if not. Bystanding costs nothing and confers nothing.

In Diekmann's (1985) original, n-player volunteer's dilemma, there are n pure-strategy Nash equilibria where exactly one player helps and everyone else bystands. When I introduce the pretend option, these same equilibria instead become exactly one player helping while everyone else pretends.[17] The intuition is that if provision succeeds, pretenders can also share in prestige (Result 1). Whether this outcome is socially wasteful is a matter of perspective, as the resources expended to pretend are compensated by the utility gained from prestige. Like the original model, the pretense model has a unique, symmetric, mixed-strategy Nash equilibrium where everyone mixes on help/bystand. Unlike the original model, it additionally has one where everyone mixes on help/pretend. Exactly one of these two equilibria must exist (Result 2a).

Regarding the effect of group size, in both equilibria as n grows toward infinity, the probability an individual helps shrinks toward zero and the likelihood the good is provided approaches a constant between zero and one. This is consistent with Diekmann's model (Result 3b). Whether the mixed-strategy equilibrium is help/bystand or help/pretend depends on costs and benefits, but surprisingly not on group size. Specifically, the equilibrium with the smaller individual probability to help exists. A somber logic implies that when the pretending occurs played, the public good is even *less* likely to be provided than when pretending is not possible (Result 2b).

For sufficiently large values of prestige and small values of shame, the help/pretend mix dominates. From an anthropological view, this means that societies which value image heavily develop a norm to appear helpful, which motivates pretending (Result 4). The rise of social media, for instance, has created audience effects and norms of philanthropy. At the same time, the ease of forwarding solicitations can make pretense an attractive option. Telecommunications enables people to not only *not* help, but furthermore *pretend* to help. If pretense becomes widespread and problematic, an organization may actually prefer to reduce the prestige accorded to helpful behavior, perhaps by anonymizing donors.

The pretending volunteer's dilemma shares some structural features to Akerlof's (1970) famous market for lemons. Akerlof argued that when buyers cannot distinguish between a high-quality car ('peach') and a low-quality car ('lemon'), they will in expectation pay only the average price of a peach and lemon. Sellers, on the other hand, know the quality of their car. Given the lower, average price at which buyers would buy, 'lemons' sell while 'peaches' leave the market. Similarly, when the audience is sophisticated, they discount prestige by the probability an observed helper is actually helping. Thus, the presence of pretenders has a double-negative effect. First, pretenders do not help. Second, pretenders dilute prestige and discourage potential helpers from helping. Discouraged helpers may fear that their help will be doubted as fake (Result 6).

Interestingly, discouraged helpers facing a diluted prestige are willing to pay extra costs, up to the prestige loss, to authenticate their help. When an authentication option is introduced, either all mix on authenticate/bystand or all mix on help/pretend (Result 7). While it may seem strange, people sometimes do pay a premium to authenticate their help. A

[17] With endogenous prestige, the cost of pretending must be cheaper than the benefit of diluted prestige; else, the rest would bystand.

worker might work in an inconvenient but public space for visibility. A philanthropist might increase donation to a higher bracket to avoid being listed among lower-tier donors who may be donating the bare minimum (e.g. $100+ group instead of $1 to $99 group). This result is supported by data from Harbaugh's (1998a) work on philanthropy and prestige, which shows that when donations are both tiered and publicized, most donations bunch at the minimum required for inclusion in each tier.

I extend pretense to a simplified public goods game. In Samuelson's (1954) original, n-player public goods game, the dominant strategy is to donate zero. In my pretender's public goods game, this is also a Nash equilibrium, but there is a second outcome where everyone donates. This occurs if prestige and warm glow are sufficiently high to offset material costs.[18] In essence, the game simplifies into a shop where individuals may purchase looking and feeling good via their donation (Result 8). There are no asymmetric, pure-strategy Nash equilibria (Result 9). A second, striking difference is that, for some weighting schemas ('virtuous'), the pretender's public goods game supports mixed-strategy Nash equilibria where none had originally existed (Result 10). For three-player games, both the 'virtuous' and 'egalitarian' schemas support a three-way equilibrium that mixes on help, pretend, and bystand. When prestige is high, the expected contribution level is greater when players care to impress only other contributors, relative to impressing everyone (Result 11).

In an incomplete information equilibrium, a two-dimensional space exists that maps types to actions (Result 12). Recent work on 'reluctant' helpers by Lazear et al. (2012), Dana et al. (2006, 2007) suggest that people can be grouped into altruistic, selfish, or 'reluctant' types. Reluctant helpers are motivated by prestige but would rather not help if unobserved. These types correspond closely to helpers, bystanders, and pretenders in my heterogeneity model. The work of these authors estimate that about 35% of the population are altruistic, 25% are selfish, and 40% are reluctant. In crafting Figs. 4 and 5, I took special effort to equate the areas of the three regions to these percentages.

My theory on pretense in public goods games enables several extensions. To construct a comprehensive theory, I made the simplifying assumption that players in the three public goods games (constant, increasing, and decreasing returns) contribute either 1 unit or 0 units. An extension might consider how the pretender's public goods model changes when players can contribute any continuous fraction of their endowment. Within this fraction, perhaps players can further decide how much is real and how much is fake. For example, a player endowed with 10 coins could submit 3 real coins and 4 (cheaper) fake coins.

Experimentally, when pretending is an option, should we expect an increase in *observed* donation levels, but a decrease in *actual* donation levels? If so, this would imply that some helpers were relucant and also that some bystanders would 'buy' prestige if only prestige were cheaper. Relative to donation levels found by Isaac and Walker (1988a, 1998b), does

[18]In some cases, it may also depend on n being sufficiently small.

the pretender's public goods game decay more rapidly over time? In an environment where pretending is possible, to what extent would introducing an authenticate option improve donation levels?

References

[1] **Abeler, J., Becker, A., Falk, A.** 2014. Representative evidence on lying costs. Journal of Public Economics. 113: 96-104.

[2] **Akerlof, G. A.** 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. The Quarterly Journal of Economics. 84: 488-500.

[3] **Andreoni, J.** 1989. Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. Journal of Political Economy. 97: 1447-1458.

[4] **Andreoni, J.** 1990. Impure altruism and donations to public goods: A theory of warm-glow giving. The Economic Journal. 100: 464-477.

[5] **Andreoni, J.** 1995. Cooperation in Public-Goods Experiments: Kindness or Confusion? The American Economic Review. 84: 891-904.

[6] **Andreoni, J., Bernheim, B. D.** 2009. Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects. Econometrica. 77: 1607-1636.

[7] **Andreoni, J., Petrie, R.** 2004. Public goods experiments without condentiality: a glimpse into fund-raising. Journal of public Economics. 88: 1605-1623.

[8] **Bénabou, R., Tirole, J.** 2006. Incentives and Prosocial Behavior. American Economic Review. 96: 1652-1678.

[9] **Bergstrom, T. C., Garratt, R., Leo, G.** 2015. Let me, or let George? Motives of Competing Altruists. University of California, Santa Barbara working paper.

[10] **Charness, G., Yang, C.** 2014. Starting small toward voluntary formation of efficient large groups in public goods provision. Journal of Economic Behavior & Organization. 102: 119-132.

[11] **Charness, G., Rustichini, A., van de Ven, J.** Self-Confidence and Strategic Behavior. University of California, Santa Barbara, University of Minnesota, University of Amsterdam working paper.

[12] **Cinyabuguma, M., Page, T., Putterman, L.** 2005. Cooperation under the threat of expulsion in a public goods experiment. Journal of Public Economics. 89: 1421-1435.

[13] **Dana, J., Cain, D. M., Dawes R.** 2006. What you don't know won't hurt me: Costly (but quiet) exit in a dictator game. Organizational Behavior and Human Decision Processes. 100: 193-201.

[14] **Dana, J., Weber R. A., Kuang J. X.** 2007. Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness. Economic Theory. 33: 67-80.

[15] **Darley, J., Latane, B.** 1968. Bystander "Apathy". American Scientist. 57: 244-268.

[16] **Dawkins, R.** 1976. The selfish gene. New York: Oxford University Press.

[17] **Dieckmann, A.** 1985. Volunteer's Dilemma. Journal of Conflict Resolution. 29: 605-610.

[18] **Diekmann, A.** 1993. Cooperation in Asymmetric Volunteer's Dilemma Game: Theory and Experimental Evidence. International Journal of Game Theory. 22: 75-85.

[19] **Erat and Gneezy** 2011. White Lies. Management Science. 58: 723-733.

[20] **Fehr, E., Gächter, S.** 2000. Cooperation and Punishment in Public Goods Experiments. The American Economic Review. 90: 980-994.

[21] **Franzen, A.** 1995. Group Size and One Shot Collective Action. Rationality and Society. 7: 183-200.

[22] **Glazer, A., Konrad, K.** 1996. A Signaling Explanation for Charity. The American Economic Review. 86: 1019-1028.

[23] **Gneezy, U., Rockenbach, B., Serra-Garcia, M.** 2013. Measuring lying aversion. Journal of Economic Behavior & Organization. 93: 293-300.

[24] **Harbaugh, W. T.** 1998a. What do donations buy? A model of philanthropy based on prestige and warm glow. Journal of Public Economics. 67: 269-284.

[25] **Harbaugh, W. T.** 1998b. The prestige motive for making charitable transfers. American Economic Review. 88: 277-282.

[26] **Harsanyi, J. C.** 1967-1968. Games with incomplete information played by "Bayesian" players. Management Science. 14: 159-182, 320-334, 486-502.

[27] **Houser, D., Kurzban, R.** 2002. Revisiting Kindness and Confusion in Public Goods Experiments. The American Economic Review. 92: 1062-1069.

[28] **Lazear, E. P., Malmendier, U., Weber, R. A.** 2012. Sorting in Experiments with Application to Social Preferences. American Economic Journal: Applied Economics. 4: 136-163.

[29] **Maslow, A. H.** 1943. A theory of human motivation. Psychological Review. 50: 370-396.

[30] **Murnighan, J. K., Kim, J., Metzger, A. R.** 1993. The volunteer dilemma. Administrative Science Quarterly. 38: 515-538.

[31] **Samuelson, P. A.** 1954. The pure theory of public expenditure. The Review of Economics and Statistics. 36: 387-389.

[32] **Weesie, J.** 1994. Incomplete Information and Timing in the Volunteer's Dilemma: A Comparison of Four Models. Journal of Conflict Resolution. 38: 557-585.

[33] **Weesie, J., Franzen, A.** 1998. Cost sharing in a volunteers dilemma. Journal of Conflict Resolution. 42: 600-618.

Appendix

A1. Proofs

Proof of Proposition 2. Let p_H be the probability an individual helps, p_B the probability of bystanding, and p_P that of pretending. With three actions, there potentially exist up to four symmetric, mixed-strategy equilibria: 1) mixing only on actions B/P (i.e. $p_B + p_P = 1, p_H = 0$); 2) mixing only on H/B; 3) mixing only on H/P; or 4) mixing on all three actions $H/P/B$.[19]

Mixing only on B/P is not an equilibrium because $p_H = 0$ means the public good is not provided and thus each player prefers action H to either B or P. Mixing only on H/B is an equilibrium if and only if there exists a p_B^* (and implicitly $p_H^* = 1 - p_B^*$) that makes a player indifferent between H and B but still weakly prefer either to P. Given that each other player helps with probability $1 - p_B^*$, a helper guarantees a payoff of $b + g + w - c$. A bystander receives b only if someone else helps, which occurs with probability $1 - (p_B)^{n-1}$. A pretender receives $b + g - c_P$ if someone else helps and $-c_P - s$ if no one else helps. In terms of expected utility, this translates to:

$$E(u_i|H) = E(u_i|B) \geq E(u_i|P) \tag{12}$$

$$b + g + w - c = b[1 - (p_B)^{n-1}] \geq (b + g)[1 - (p_B)^{n-1}] - s(p_B)^{n-1} - c_P \tag{13}$$

$$\implies p_B^* = \left(\frac{c - g - w}{b}\right)^{1/(n-1)} \geq \left(\frac{c - c_P - w}{b + g + s}\right)^{1/(n-1)} \tag{14}$$

If instead mixing only on H/P were an equilibrium, then:

$$E(u_i|H) = E(u_i|P) \geq E(u_i|B) \tag{15}$$

$$b + g + w - c = (b + g)[1 - (p_P)^{n-1}] - s(p_B)^{n-1} - c_P \geq b[1 - (p_P)^{n-1}] \tag{16}$$

$$\implies p_P^* = \left(\frac{c - c_P - w}{b + g + s}\right)^{1/(n-1)} \geq \left(\frac{c - g - w}{b}\right)^{1/(n-1)} \tag{17}$$

The right side of Inequality 14 is p_P^* in the H/P equilibrium; conversely, the right side of Inequality 17 is p_B^* in the H/B equilibrium. At least one of these inequalities must be true, but generally not both at once. The condition that decides which equilibrium exists is:

$$\frac{c - g - w}{b} \geq\leq \frac{c - c_P - w}{b + g + s} \tag{18}$$

If the left side is strictly greater, then the H/P equilibrium exists. If the right side is strictly greater, then the H/B equilibrium exists. In the coincidence that both sides are

[19] A game with k actions has up to $2^k - k - 1$ symmetric, (non-pure) mixed-strategy equilibria.

equal, the equilibrium mixes on either H/P (and $p_B^* = 0$) or H/B (and $p_P^* = 0$). However, this happens on a set of measure zero in parameter space.

Suppose a three-way, symmetric, mixed strategy exists. Given that each other player helps, pretends, or bystands with respective probabilities p_H, p_P, p_B, the expected utility from each action is:

$$E(u_i|H) = b + g + w - c \tag{19}$$
$$E(u_i|P) = (b + g)[1 - (1 - p_H)^{n-1}] - s(1 - p_H)^{n-1} - c_P \tag{20}$$
$$E(u_i|B) = b[1 - (1 - p_H)^{n-1}] \tag{21}$$

Equating $E(u_i|B) = E(u_i|P)$ yields $p_H^* = 1 - \left(\frac{g - c_P}{g + s}\right)^{1/(n-1)}$. However, $E(u_i|P)$ evaluated at p_H^* equals $E(u_i|H)$ on a set of measure zero in parameter space. So, a three-way, symmetric, mixed strategy does not exist. □

Proof of Proposition 5. An equilibrium that mixes only on H/B exists when:

$$E(u_i|H) = E(u_i|B) \geq E(u_i|P) \tag{22}$$
$$b + g + w - c = b[1 - (p_B)^{n-1}] \geq (b + g)[1 - (p_B)^{n-1}] - s(p_B)^{n-1} - c_P \tag{23}$$
$$\implies p_B^* = \left(\frac{c - g - w}{b}\right)^{1/(n-1)} \geq \left(\frac{c - c_P - w}{b + g + s}\right)^{1/(n-1)} \tag{24}$$

An equilibrium that mixes only on H/P exists when:

$$E(u_i|H) = E(u_i|P) \geq E(u_i|B) \tag{25}$$
$$b + g \cdot p_H + w - c = (b + g \cdot p_H)[1 - (p_P)^{n-1}] - s(p_P)^{n-1} - c_P \geq b[1 - (p_P)^{n-1}] \tag{26}$$
$$\implies p_P^* = \left(\frac{c - c_P - w}{b + g \cdot p_H + s}\right)^{1/(n-1)} \geq \left(\frac{c - g \cdot p_H - w}{b}\right)^{1/(n-1)} \tag{27}$$

When Inequalities 24 and 27 are both false, then the three-way mixed strategy equilibrium exists. Given that each other player helps, pretends, or bystands with respective probabilities p_H, p_P, p_B, the expected utility from each action is:

$$E(u_i|H) = b + g\frac{p_H}{p_H + p_P} + w - c \tag{28}$$
$$E(u_i|P) = (b + g\frac{p_H}{p_H + p_P})[1 - (1 - p_H)^{n-1}] - s(1 - p_H)^{n-1} - c_P \tag{29}$$
$$E(u_i|B) = b[1 - (1 - p_H)^{n-1}] \tag{30}$$

Subtracting $E(u_i|B)$ from each and equating $E(u_i|H) = E(u_i|P) = E(u_i|B)$ yields:

$$0 = b(1-p_H)^{n-1} + g\frac{p_H}{p_H+p_P} + w - c \tag{31}$$

$$0 = g\frac{p_H}{p_H+p_P}[1-(1-p_H)^{n-1}] - s(1-p_H)^{n-1} - c_P \tag{32}$$

Equilibrium (p_H^*, p_P^*) must satisfy this system of two equations. There is no general, closed-form solution due to the interaction between exponents and fractions. By the Nash Existence Theorem, however, every n-player game with finite actions has at least one equilibrium. If there are no pure-strategy equilibria, then there must be a unique, mixed-strategy equilibrium. If the mixed-strategy equilibrium does not mix on only H/B, H/P, or B/P, then it must mix on $H/P/B$. □

Proof of Proposition 9. An asymmetric, pure-strategy Nash equilibrium, if it exists, must take one of four forms: H/B, H/P, B/P, or $H/P/B$.

Suppose one exists in only H/B such that X players help and the rest $(n-X)$ bystand.[20] Given this composition, a helper must prefer $H \succeq B$ and a bystander must prefer $H \preceq B$.[21] $E(u_H|H) \geq E(u_H|B)$ and $E(u_B|H) \leq E(u_B|B)$ respectively simplify to:

$$\alpha_H(X-1) + \alpha_B(n-X) \geq (c_H - w)(n-1)/g \tag{33}$$

$$\alpha_H(X) + \alpha_B(n-1-X) \leq (c_H - w)(n-1)/g \tag{34}$$

$$\implies \alpha_H \leq \alpha_B \tag{35}$$

However, $\alpha_H \geq \alpha_B$. Even if $\alpha_H = \alpha_B$, then $g \cdot \alpha_H + w = c_H$, which occurs on a set of measure zero in parameter space. Likewise, $E(u_H|H) \geq E(u_H|P)$ and $E(u_P|H) \leq E(u_P|P)$ respectively simplify to:

$$\alpha_H(X-1) + \alpha_P(n-X) \geq (c_H - c_P - w)(n-1)^2/(g+s) \tag{36}$$

$$\alpha_H \cdot X + \alpha_P(n-1-X) \leq (c_H - c_P - w)(n-1)^2/(g+s) \tag{37}$$

$$\implies \alpha_H \leq \alpha_P \tag{38}$$

However, $\alpha_H \geq \alpha_P$. Even if $\alpha_H = \alpha_P$, then $(g+s)\alpha_H/(n-1) + w = c_H - c_P$, which again occurs on a set of measure zero.

An asymmetric equilibrium in only B/P cannot exist because P requires H.

Lastly, consider an asymmetric equilibrium in $H/P/B$. Given the composition that X players help, k pretend, and $(n-X-k)$ bystand, no type should switch to any other type.

[20] This notation matches our previous X being the pot size, or total contribution.

[21] $H \succeq P$ and $B \succeq P$ must also be true, but this is unnecessary. In general, a subset of preference relations is sufficient to disprove the existence of asymmetric, pure-strategy equilibria.

Then, $E(u_H|H) \geq E(u_H|B)$ and $E(u_B|B) \geq E(u_B|H)$), which simplify respectively to:

$$\alpha_H \cdot \frac{X-1}{n-1} \cdot \frac{X-1}{X+k-1} + \alpha_P \cdot \frac{k}{n-1} \cdot \frac{X}{X+k-1} + \alpha_B \cdot \frac{n-X-k+1}{n-1} \cdot \frac{X}{X+k} \geq (c_H - w)/(g+s) \quad (39)$$

$$\alpha_H \cdot \frac{X}{n-1} \cdot \frac{X}{X+k} + \alpha_P \cdot \frac{k}{n-1} \cdot \frac{X+1}{X+k} + \alpha_B \cdot \frac{n-X-k}{n-1} \cdot \frac{X+1}{X+k+1} \leq (c_H - w)/(g+s) \quad (40)$$

Inequalities 39-40 are simultaneously true only if $k = 0$ and $\alpha_H \leq \alpha_B$. $k = 0$ means that there are no pretenders, which is sufficient to refute a three-way asymmetry. So, an asymmetric, pure-strategy equilibrium in $H/P/B$ cannot exist and thus no asymmetric, pure-strategy equilibrium can exist at all. □

Proof of Proposition 10. A symmetric, mixed-strategy Nash equilibrium, if it exists, must take one of four forms: mixing only on H/B, only on H/P, only on B/P, or on $H/P/B$.

In a two-player game, a player who pretends is immediately exposed because the other player sees their own action and the pot size. Among two, no mixed strategy can include pretending so the only possibility mixes on H/B. Given the expectation that the other player j is a helper p_H of the time and a bystander $1-p_H$ of the time, player i is indifferent between H and B. All claimants are helpers, so $\beta_{Hj} = 1$. $E(u_i|H) = E(u_i|B)$ simplifies to:

$$\alpha_H \cdot p_H + \alpha_B(1 - p_H) = (c_H - w)/g \quad (41)$$

$$\implies p_H^* = \frac{c_H - w - \alpha_B \cdot g}{(\alpha_H - \alpha_B)g} \quad (42)$$

This relies on $\alpha_H > \alpha_B$. Recall that β_{ij} is the probability that player i is believed to be a helper by another player j. In a three-player game with mixed-strategy (p_H, p_B, p_P), this belief depends on the realization of random actions.

		i helps	i bystands	i pretends
probability	actions of j, k	β_{Hj}, β_{Hk}	β_{Bj}, β_{Bk}	β_{Pj}, β_{Pk}
p_H^2	H, H	1, 1	0, 0	0.5, 0.5
$2 \cdot p_H \cdot p_B$	H, B	1, 1	0, 0	0, 0.5
$2 \cdot p_H \cdot p_P$	H, P	0.5, 1	0, 0	0, 0.5
p_B^2	B, B	1, 1	0, 0	0, 0
$2 \cdot p_B \cdot p_P$	B, P	0.5, 1	0, 0	0, 0
p_P^2	P, P	0.5, 0.5	0, 0	0, 0

Table 5. Belief i Helped Based on Actions of i, j, k

Suppose an equilibrium exists that mixes only on H/B. Then $p_H + p_B = 1$ and $p_P = 0$. $E(u_i|H) = E(u_i|B) \implies p_H^2 \cdot \alpha_H \cdot g + 2 \cdot p_H(1-p_H)(\alpha_H + \alpha_B) \cdot g/2 + (1-p_H)^2 \alpha_B \cdot g = c_H - w$, which simplifies to:

$$p_H^* = \frac{c_H - w - \alpha_B \cdot g}{(\alpha_H - \alpha_B)g} \tag{43}$$

This relies on $\alpha_H > \alpha_B$. That $p_P = 0$ requires $E(u_i|B) \geq E(u_i|P)$, which simplifies to:

$$0 \geq p_H^2 \cdot \frac{\alpha_H(g+s) - 2 \cdot \alpha_B \cdot g}{2} + p_H \cdot \frac{\alpha_B(g+3s) - 2 \cdot \alpha_H \cdot s}{2} + \alpha_B(-s) - c_P \tag{44}$$

This is a parabola with solution $p_H^* \leq (-B + \sqrt{B^2 - 4AC})/2A$, where A, B, C are the quadratic coefficients.[22] An equilibrium that mixes only on H/P implies $p_H + p_P = 1$ and $p_B = 0$. $E(u_i|H) = E(u_i|P) \implies p_H^2 \cdot \alpha_H(g+s)/2 + 2 \cdot p_H(1-p_H)(\alpha_H + \alpha_P)(g+s)/4 + (1-p_H^2)\alpha_P(g+s)/2 = c_H - c_P - w$, which simplifies to:

$$p_H^* = \frac{2(c_H - c_P - w) - \alpha_P(g+s)}{(\alpha_H - \alpha_P)(g+s)} \tag{45}$$

This relies on $\alpha_H > \alpha_P$. That $p_B = 0$ requires $E(u_i|B) \leq E(u_i|P)$, which simplifies to:

$$0 \leq p_H^2 \cdot \frac{\alpha_H(g+s) - 2 \cdot \alpha_P \cdot g}{2} + p_H \cdot \frac{\alpha_P(g+3s) - 2 \cdot \alpha_H \cdot s}{2} + \alpha_P(-s) - c_P \tag{46}$$

This is a parabola with solution $p_H^* \geq (-B + \sqrt{B^2 - 4AC})/2A$, where A, B, C are the quadratic coefficients. When Inequalities 44 and 46 are both false, then a three-way mixed-strategy Nash equilibrium exists. An equilibrium that mixes on $H/P/B$ exists if, given $p_H + p_B + p_P = 1$, $E(u_i|H) = E(u_i|B) = E(u_i|P)$.

$$\begin{aligned} 0 = -c_H + w + p_H(\alpha_H - \alpha_B)g - p_H \cdot p_P[\alpha_H(3g-s)/2 + (\alpha_P - \alpha_B)g] \\ + p_P^2[\alpha_P(g-s)/2 + \alpha_B \cdot g] - p_P(2 \cdot \alpha_B \cdot g) + \alpha_B \cdot g \end{aligned} \tag{47}$$

$$\begin{aligned} 0 = -c_P + p_H^2(\alpha_H - \alpha_B)(g+s)/2 + p_H[\alpha_H(-s) + \alpha_B(g+3s)/2] \\ + p_H \cdot p_P(\alpha_P - \alpha_B)(g+s)/2 + p_P(\alpha_B - \alpha_P)s + \alpha_B(-s) \end{aligned} \tag{48}$$

Equilibrium $(p_H^*, p_P^*, 1 - p_H^* - p_P^*)$ solves this system of two equations, two variables. □

Proof of Proposition 12. Let the binary function $S_i(g_i, w_i) \in \{0, 1\}$ determine the probability a player i helps (H) or not ($\neg H$) given their type. Further, if $S_i = 0$ for a player, then a second binary function $T_i(g_i, w_i) \in \{0, 1\}$ determines the probability i pretends (P), as opposed to bystand (B). Then, i's expected payoff is a function of their type and the

[22]In the following subsection, I compare equilibrium probabilities for several different α_j weighting schemas.

vectors $S = (S_1, S_2, ..., S_n)$, $T = (T_1, T_2, ..., T_n)$ that map types to actions.

$$E(u_i|H) = (n-1)E(S_j) \cdot m - c_H + w_i + (g_i + s)E(\alpha_j \cdot \beta_{Hj}) - s \cdot E(\alpha_j) \tag{49}$$

$$E(u_i|B) = (n-1)E(S_j) \cdot m - 0 \tag{50}$$

$$E(u_i|P) = (n-1)E(S_j) \cdot m - c_P + (g_i + s)E(\alpha_j \cdot \beta_{Pj}) - s \cdot E(\alpha_j) \tag{51}$$

where $E(\beta_{ij})$ is the *expected* belief that claimant i is a helper, from j's perspective. In this case, it is the expected population of helpers out of the expected population of helpers and pretenders:[23]

$$E(H_{pop}) = \int_0^{\bar{w}} \int_0^{\bar{g}} S_i(g_i, w_i) dg_i dw_i \tag{52}$$

$$E(P_{pop}) = \int_0^{\bar{w}} \int_0^{\bar{g}} (1 - S_i(g_i, w_i)) T_i(g_i, w_i) dg_i dw_i \tag{53}$$

$$E(\beta) = \frac{E(H_{pop})}{E(H_{pop}) + E(P_{pop})} \tag{54}$$

A player i selects the action which yields the greatest expected utility. In equilibrium, the functions S_i, T_i must hold true given all other S_j, T_j where $j \neq i$. The equilibrium is unique because S, being a vector of functions, inputs a set of parameters and outputs exactly one threshold between helping (H) and not ($\neg H$) and thus one expected likelihood of provision $E(P^*)$. Vector T in turn inputs this $E(P^*)$ to output one threshold between pretending (P) and bystanding (B). Thus, the incomplete information equilibrium is unique. □

[23] β_{ij} in the base model is determined in equilibrium. Here $E(\beta_{ij})$ is an expected value because preferences are private and random so no one can ex-ante deduce the realized value of β_{ij}.