Identity in Public Goods Contribution*

Fuhai Hong† Larry Karp‡ Tat How Teh§

February 2, 2018

Abstract

Agents' decision whether to join a group, and their subsequent contribution to a public good, depend on the group's ideals. Agents have different preference for this public good, e.g. reductions in greenhouse gas emissions. People who become "climate insiders" obtain identity utility, but suffer disutility if they deviate from the group ideal. That ideal might create a wide but shallow group, having many members but little effect on behavior, or a narrow but deep group. Greater heterogeneity of preferences causes the contribution-maximizing ideal to create narrow but deep groups. The contribution-maximizing ideal maximizes welfare if the population is large.

Keywords: Ideals, Identity, Public Goods, Social Categories.

JEL Classification: D03, H41, Q54.

*We thank Roland Bénabou, Parkash Chander, In-Koo Cho and conference and seminar participants in the 2016 EAERE Conference (Zurich), the 2017 AERE Conference (Pittsburgh), Arne Ryde Workshop on Identity, Image and Economic Behavior (Lund), NTU Workshop on Environmental Economics, 4th Hanyang-Kobe-NTU Symposium in Economics, Hong Kong Baptist University, Hong Kong University of Science and Technology, Kobe University, National University of Singapore, Peking University HSBC Business School, Renmin University of China, University of Groningen, and Xiamen University for helpful comments.

†Department of Economics, Lingnan University, Hong Kong, and Division of Economics, Nanyang Technological University, Singapore. Email: fuhaihong@ln.edu.hk.

‡Department of Agricultural and Resource Economics, University of California, Berkeley. Email: karp@berkeley.edu.

§Department of Economics, National University of Singapore. Email: tehtathow@gmail.com.

1 Introduction

A group's norms and ideals influence a person's decision whether to join the group, and potentially influence their subsequent behavior. Connections between group ideals and agent behavior arise for many types of groups, so the results we obtain are widely applicable. Because of its importance, we consider groups with different views on climate change. Environmentalists accept the existence of anthropogenic climate change, support public policy to reduce carbon emissions, and may make lifestyle changes to reduce their carbon footprint. Climate skeptics refuse to take personal action or to support public policy to reduce carbon emissions. An agent may obtain utility from identifying with a particular group, but membership can also create costs, pressuring members to increase their contribution to a public good.

In this "identity game", based on Akerlof and Kranton (2000), the population has a distribution of preferences for a socially beneficial action, such as reducing carbon emissions. Agents self-select into the green or non-green group. Those who join the green group receive a warm glow or status- or commercial-related benefit; they decide how closely to match the group ideal, and bear costs associated with deviations from the ideal. Those who do not join the group, do not change their behavior. Akerlof and Kranton (2000) and extensions use the identity model to analyze behavior in the workplace, the school and the family, focusing on discrimination, poverty, labor division, and corporate culture (Akerlof and Kranton, 2002, 2005, 2008, 2010, Hiller and Verdier, 2014). Most of these models take the ideal prescribed by the group as fixed; we determine the optimal level of ideal.[1]

In our setting, the green ideal is a recommended level of contribution to the public good, such as a level of abatement. The rigor of this ideal influences the self-selection and the behavior of those who identify as green. A demanding ideal might lead to significant changes in behavior, but only amongst the small group that self-identifies as green: this group is deep but narrow. A relaxed ideal can lead to large green membership but only modest changes in behavior: a wide but shallow group.

The ideal that maximizes aggregate contributions to the public good depends on the distribution of preferences. Under uniformly distributed preferences, width trumps depth when agent heterogeneity is small: the contribution-maximizing ideal makes the agent with weakest green preferences indifferent between joining the green group and remaining outside. When preference heterogeneity is large, it is too expensive to attract all agents into the group, but many of those who join significantly increase their contribution to the public good; here, depth trumps width. For more

[1] In a school setting, Akerlof and Kranton (2002) consider a tradeoff in choosing the school's ideal: A higher ideal raises the effort choice of those in the right tail of the distribution but causes other students to reject the school and exert less effort. They do not solve for the optimal ideal though. Other approaches of modelling identity include oppositional identities (Bisin et al 2011), social identity as collective reputation (Carvalho, 2016) and identity investment when a player has incomplete information on her own type (Bénabou and Tirole, 2011). In Shayo (2009), identity utility is derived from status while identity cost comes from deviation from other group members' behavior rather than an ideal. Almudi and Chóliz (2011) introduce an environmentally friendly identity which only depends on the individual's consumption level, while there is no ideal behavior or social categorization in their model. Costa-i-Font and Cowell (2013) review the related literature. Our approach is also related to Hsiaw (2013) who considers an individual setting a goal to deal with self-control problems.

general preference distributions, the contribution-maximizing ideal depends on skewness. If the distribution is symmetric, greater preference heterogeneity still leads to a narrow but deep group, as in the case of uniform distribution.

Agents' individually rational behavior constrains the possibility of increasing contributions by means of manipulating the group ideal: beyond some level, the greater depth arising from a higher ideal does not make up for the resulting loss in width. In important cases, this constraint is binding for the welfare-maximization problem; here, the contribution-maximizing ideal also maximizes welfare. However, for some distributions of preferences, the welfare-maximizing ideal exceeds the level that maximizes contributions. In this situation it is feasible but not optimal to increase contributions by lowering the ideal. Here, an ideal above the contribution-maximizing level increases welfare by improving the match between types and levels of contribution.

The characterization of the optimal ideal is analogous to the result in the mechanism design problem with adverse selection (Laffont and Martimort, 2001). There, a principal delegates production decisions to agents who have private information about their efficiency. The optimal menu of offers induces both types of agents to produce if the difference in efficiency is small; in contrast, when the difference in efficiency is large, it is not worthwhile attracting the inefficient agent, and the optimal menu attracts the efficient agent only. The trade-offs in the two problems are similar: In the mechanism design problem, the principal trades off between the rent extracted from the efficient agent and the participation of the inefficient agent; in the identity model, the trade-off is between the contribution by strong-preference individuals and the participation of the weak-preference individuals.

Agents might be individuals or actors such as companies, cities, or states. Individuals can contribute to the public good by reducing their energy consumption below their individually rational level. California's steps to reduce carbon emissions perhaps exceed efforts that would maximize its (narrowly construed) welfare. Some explanations for this behavior are unrelated to identity: Californians might expect benefits from early adoption or they might believe that their demonstration will encourage others to follow their lead. However, identity benefits can be an important motivator. For individuals, these benefits might be psychological, but for companies or states, there may be brand-related commercial benefits or political prestige.

Individuals' self-identification and behavior are correlated. Kotchen and Moore (2008) find that environmentalists are more likely to voluntarily restrain their consumption of goods and services that generate negative externalities. Kahn (2007) finds that those who vote for green policies and register for liberal/environmental political parties live a greener lifestyle, commuting by public transit more often, favoring hybrid vehicles, and consuming less gasoline than non-environmentalists.

Psychologists and management scientists show that persuasion strategies or nudges can change behavior (Thaler and Sunstein 2008, Schultz et al 2007, Goldstein, Cialdini, and Griskevicius, 2008). Costa and Kahn (2013) find that an electricity conservation nudge that provides feedback to households on their own and peers' electricity usage is much more effective with liberals/environmentalists than with conservatives. These authors ascribe the asymmetry to self-identification: the ideologies that people accept, influence their behavior. Here, nudges appear useful only for those who identify with the ideology embedded in the nudges; the goals or norms provided by one group have little impact on people who do not belong to that group.

An extensive behavioral economics literature studies public goods. A moral imperative, arising from introspection and associated with Kantian absolute laws, can enhance public good provision (Brekke, Kverndokk and Nyborg, 2003). In the Akerlof and Kranton (2000) framework, the ideal is a social but perhaps not moral norm; people's acceptance of the social norm affects their self-selection into social categories. Fehr and Schmidt (1999) examine the role of inequality aversion in voluntary public good contribution. Andreoni (1990) and Holländer (1990) consider warm glow and social approval as by-products of contributing to a public good. These studies do not consider self-selection into social groups, the focus of our paper. Rege (2004) endogenizes the strength of social approvals, emphasizing interactions among contributors and non-contributors, in a model without an ideal public good contribution.

Group identity has significant effects on interpersonal interactions even in laboratory settings (Chen and Li, 2009). Empirical evidence demonstrates the role of identity in public good provision outside the laboratory. Burlando and Hey (1997), Benjamin, Choi and Fisher (2010), Solow and Kirkwood (2002), and Croson, Marks and Snyder (2003) estimate the effects of national, religious, social and gender identities on public good contribution.

Section 2 develops the model and Section 3 discusses the ideal that maximizes the expected level of public good contribution or social welfare. Section 4 concludes. Short proofs appear in footnotes, and longer proofs in the Appendix.

2 The Model

The population contains N agents, each of whom makes a voluntary contribution to a public good. In the environmental context, the contribution equals pollution abatement. Agent i contributes a_i, for $i \in \{1, 2, ..., N\}$, incurring the private cost $\frac{1}{2}a_i^2$. Agent i's constant marginal utility of the public good is the realization of an identically and independently distributed random variable β_i, having continuous probability density function $f(\beta_i)$ and cumulative distribution function $F(\beta_i)$ defined on $[\underline{\beta}, \overline{\beta}]$, with $\overline{\beta} > \underline{\beta} > 0$. Many public economics models use these functional assumptions, which lead to a simple equilibrium in dominant strategies (e.g. Barrett 1994, Goeschl and Perino 2015 and Ali and Bénabou, 2016).

The heterogeneity of β_i may be due to differences in tastes, information, business opportunities, or social preferences. An individual's preference for air quality may depend on income, which affects their opportunities for adaptation (e.g. air conditioning or filters). A state's preference might depend on population density. Environmentalists and climate skeptics may have different information or beliefs about the consequences of the accumulation of Greenhouse gas (GHG), and therefore about the benefit of abatement.[2]

[2]Altruism/spite can also create heterogeneity of β_i. Hammond (1987) characterizes altruism and discusses its relevance in public good provision. Cason and Saijo (2002), and Cason et al. (2004) provide empirical evidence of spite in public good games. If β_i incorporates altruism, rather than merely tastes and beliefs, the socially optimal level of abatement must take that altruism into account in evaluating the costs as well as the benefits of abatement (Bergstrom 2006). To avoid this complication, we assume that heterogeneity of β_i arises from reasons unrelated to altruism.

Agent i's utility associated with the public good (ignoring identity-related utility) is $\beta_i \left(a_i + \sum_{j \neq i} a_j \right) - \frac{1}{2} a_i^2$. The agent takes as given other agents' contributions. Without identity-related utility (our baseline), agent i chooses a_i to maximize $\beta_i a_i - \frac{1}{2} a_i^2$, resulting in the baseline level of abatement

$$a_i^b = \beta_i > 0. \tag{1}$$

2.1 Identity and Utility

The group ideal level of contribution to the public good, a^*, affects the sorting of agents and the public good contribution of those who join the group. Section 3 discusses the choice of the ideal. "Insiders" self-select into the green group, and identify with the ideal; "outsiders" ignore the ideal and choose a_i^b. Individuals who join the green group obtain utility from being an insider, $V > 0$. Insiders have a sense of belonging and a feeling of pride. Companies, cities, or states may obtain commercial or political benefits associated with green membership, V. V is a club good.

Insiders who deviate from the ideal suffer a utility loss. An insider who contributes not more than the ideal (weakly) under-contributes, and one who contributes more than the ideal over-contributes. Their losses are

$$\text{under-contributing insider's loss (if } a_i \leq a^*\text{): } \frac{\theta}{2} \left(a_i - a^* \right)^2$$

$$\text{over-contributing insider's loss (if } a_i > a^*\text{): } \frac{\gamma}{2} \left(a_i - a^* \right)^2.$$

Under-contributors may feel guilty about contributing less than the group ideal. Companies or states that strictly under-contribute may be vulnerable to bad publicity. Over-contributors might also incur disutility from exceeding the ideal. Monin, Sawyer, and Marquez (2008) provide experimental evidence showing that people's positive self-image may be threatened by those who "do the right thing", leading to resentment against them, and a utility loss for over-contributors. Bénabou and Tirole (2011) endogenize the ostracism towards the virtuous "do-gooders". We adopt Akerlof and Kranton's (2002) assumption that the utility loss is a quadratic function of the gap between the insider's action and the ideal level, but we relax their assumption of symmetric loss. With $\theta \geq 0$ and $\gamma \geq 0$, the model includes both the symmetric loss case ($\gamma = \theta$) and the case where there is no loss from over-contribution ($\gamma = 0$).

The game's timeline is:

- At stage 0, an influential entity chooses (or adjusts) the ideal a^*.
- At stage 1, agents learn their types and individually decide whether to identify with the ideal a^* (and become an insider) or remain an outsider.
- At stage 2, agents individually choose their public good contribution.

2.2 Public Good Contributions (Stage 2)

Because identity does not affect the outsider's preference, an outsider contributes a_i^b. An insider with β_j solves:

$$\max_{a_j} U_j\left(a_j|a^*\right) = \begin{cases} \beta_j a_j - \frac{1}{2}a_j^2 - \frac{\theta}{2}\left(a_j - a^*\right)^2 + V \text{ if } a_j \leq a^* \\ \beta_j a_j - \frac{1}{2}a_j^2 - \frac{\gamma}{2}\left(a_j - a^*\right)^2 + V \text{ if } a_j > a^*. \end{cases} \quad (2)$$

Lemma 1 *An insider (superscript i) with β_j contributes*

$$a_j^i = \begin{cases} \frac{\theta(a^*-\beta_j)}{1+\theta} + \beta_j \textit{ if } \beta_j \leq a^* \\ \frac{\gamma(a^*-\beta_j)}{1+\gamma} + \beta_j \textit{ if } \beta_j > a^*. \end{cases} \quad (3)$$

Define $\Delta\left(\beta_j\right) = a_j^i - a_j^b$, the change of public good contribution due to membership in the group:

$$\Delta\left(\beta_j\right) = \begin{cases} \frac{\theta(a^*-\beta_j)}{1+\theta} \geq 0 \text{ if } \beta_j \leq a^* \\ \frac{\gamma(a^*-\beta_j)}{1+\gamma} < 0 \text{ if } \beta_j > a^*. \end{cases} \quad (4)$$

Equation (4) implies:[3]

Remark 1 *The effect of a^* on an insider's contribution depends on whether the agent is an under- or over-contributing insider, and is proportional to the gap $a^* - \beta_j$. Membership increases an under-contributing insider's action and decreases an over-contributing insider's action.*

Figure 1 illustrates the relation between a^*, a_j^b and a_j^i, when $a^* \in \left(\underline{\beta}, \overline{\beta}\right)$. A larger utility loss from deviating from the ideal (larger θ or γ) induces the insider to move toward the ideal.[4]

Remark 2 *Membership in the green group decreases the difference between the contributions of insiders with different preferences, but might increase or decrease the difference in contributions between insiders and outsiders: For any a^* and any $\beta_j > \beta_i$, $a^b\left(\beta_j\right) - a^b\left(\beta_i\right) \geq a^i\left(\beta_j\right) - a^i\left(\beta_i\right)$.*

Figure 1 shows that $\left|a^i\left(\beta_j\right) - a^i\left(\beta_i\right)\right| < \left|a^b\left(\beta_j\right) - a^b\left(\beta_i\right)\right|$ for any $\beta_j \neq \beta_i$. Suppose that an agent with β_i becomes an insider and an agent with β_j remains an outsider. With either $\beta_i < a^* < \beta_j$ or $\beta_j < a^* < \beta_i$, there is convergence in contribution between the insider and the outsider: $\left|a^b\left(\beta_j\right) - a^b\left(\beta_i\right)\right| > \left|a^b\left(\beta_j\right) - a^i\left(\beta_i\right)\right|$. In contrast, with either $\beta_j < \beta_i < a^*$ or $\beta_j > \beta_i > a^*$, there is divergence in contribution: $\left|a^b\left(\beta_j\right) - a^b\left(\beta_i\right)\right| < \left|a^b\left(\beta_j\right) - a^i\left(\beta_i\right)\right|$.

[3] We confirm Remark 1 using Lemma 1 and Equations (1) and (4); these imply that when $\beta_j \leq a^*$, $\beta_j \leq a_j^i \leq a^*$ and when $\beta_j > a^*$, $\beta_j > a_j^i > a^*$.

[4] For $\beta_j \leq a^*$, $\frac{da_j^i}{d\theta} = \frac{a^*-\beta_j}{(1+\theta)^2} \geq 0$ and for $\beta_j > a^*$, $\frac{da_j^i}{d\gamma} = \frac{a^*-\beta_j}{(1+\gamma)^2} < 0$.

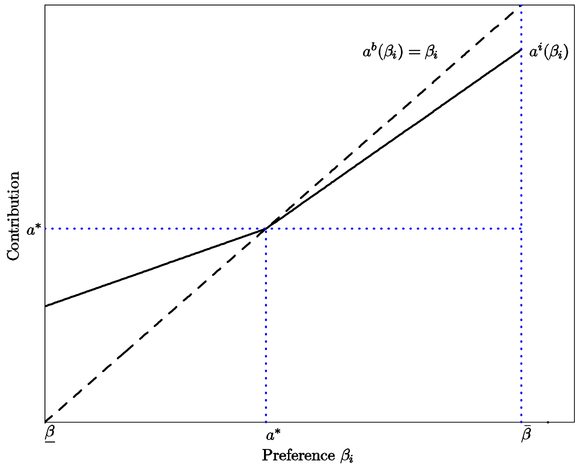

Figure 1: The solid line shows a_i^i, and the dashed line shows $a_i^b = \beta_i$. $\left[\underline{\beta}, \overline{\beta}\right] = [0, 9/5]$, $a^* = 7/8$, $\theta = 3/2$, $\gamma = 1/4$.

2.3 Self-selection (Stage 1)

At stage 1, agents compare their utility as insiders and outsiders and decide whether to join the green group. Strategies are dominant, so the agent's choice does not depend on other agents' action. Using Equation (1) and suppressing the payoff due to other agents' actions, agent i's utility of remaining an outsider equals

$$\beta_i^2 - \frac{1}{2}\beta_i^2 = \frac{1}{2}\beta_i^2. \tag{5}$$

Using Equation (3) in Expression (2), the insider's utility equals

$$\begin{array}{l} \frac{1}{2}\frac{\beta_i(\beta_i+2\theta a^*)-\theta a^{*2}}{1+\theta} + V \text{ if } \beta_i \leq a^* \\ \frac{1}{2}\frac{\beta_i(\beta_i+2\gamma a^*)-\gamma a^{*2}}{1+\gamma} + V \text{ if } \beta_i > a^*. \end{array} \tag{6}$$

Define

$$B \equiv \left(\frac{2V(1+\theta)}{\theta}\right)^{1/2} \text{ and } D \equiv \left(\frac{2V(1+\gamma)}{\gamma}\right)^{1/2}.$$

These measures increase with identity utility (V) and decrease with the insiders' cost of departing from the ideal (θ and γ). B and D thus provide measures of attractiveness of joining the group to under- and over- contributors. Under the tie-breaking assumption that an agent who is indifferent between the choices decides to join the group, we have:

Lemma 2 *Agents with*

$$\beta_i \in \left[\max\left(a^* - B, \underline{\beta}\right), \min\left(a^* + D, \overline{\beta}\right)\right]$$

join the group, and other agents remain outsiders. Among the insiders, agents with $\beta_i \leq a^*$ are under-contributors, while agents with $\beta_i > a^*$ are over-contributors.

An increase in the attractiveness of joining the group, B or D, (weakly) increases the domain of insiders. Define

$$I^- = \max\left(a^* - B, \underline{\beta}\right);\ I^+ = \min\left(a^* + D, \overline{\beta}\right);\ I^* = \begin{cases} \overline{\beta} \text{ if } a^* > \overline{\beta} \\ a^* \text{ if } a^* \in \left[\underline{\beta}, \overline{\beta}\right] \\ \underline{\beta} \text{ if } a^* < \underline{\beta} \end{cases}.$$

We divide the domain of insiders, $[I^-, I^+]$, into the domain of under-contributors, $[I^-, I^*]$, and of over-contributors, $(I^*, I^+]$. We have[5]

Remark 3 *The largest increase in individual contribution that can be implemented by a group ideal is $\frac{\theta}{1+\theta}B$.*

3 Choice of the Group Ideal (Stage 0)

An influential person or entity nudges the ideal, prior to the membership decisions, in order to increase provision of the public good. Influence-molders such as Al Gore or James Hansen, or those who support them, may be able to use the media, schools, and churches to alter the green ideal. Political entities may be able to use non-binding contracts or international agreements to adjust the green ideal. We identify the choice of a^* that maximizes provision of the public good, and show when this level also maximizes aggregate welfare.

Denote by $g(a^*)$ the expectation, at stage 0, of the effect of a^* on a random agent's contribution to the public good. Using Equation (4), an agent's expected increase in contribution is:

$$g(a^*) = \frac{\gamma}{1+\gamma}\int_{I^*}^{I^+} (a^* - \beta_i) f(\beta_i)\, d\beta_i + \frac{\theta}{1+\theta}\int_{I^-}^{I^*} (a^* - \beta_i) f(\beta_i)\, d\beta_i. \tag{7}$$

Lemma 3 provides bounds on the contribution-maximizing ideal:[6]

Lemma 3 *The contribution-maximizing ideal lies in the interval $\left(\underline{\beta}, \overline{\beta} + B\right)$.*

Equation (7) shows that a higher a^* increases insiders' contributions, leading to a deeper group. The ideal also affects the ranges of the under- and over-contributing insiders, $[I^-, I^*]$ and $(I^*, I^+]$, altering the group's width. The contribution-maximizing a^* typically involves a trade-off between depth and width.

[5] By Equation (4), the largest increase in contribution implemented by the ideal a^* is $\frac{\theta\left(a^* - I^-\right)}{1+\theta}$. Using Lemma 2, this increase equals $\frac{\theta\left(a^* - \max\left(a^* - B, \underline{\beta}\right)\right)}{1+\theta}$. This function increases in a^* for $a^* < B + \underline{\beta}$ and is constant at $\frac{\theta B}{1+\theta}$ for $a^* \geq B + \underline{\beta}$.

[6] An ideal $a^* \leq \underline{\beta}$ weakly decreases contributions. For this ideal, $I^* = \underline{\beta}$, so all insiders (if any) are over-contributors: membership decreases their contribution. An ideal $a^* \geq \overline{\beta} + B$ implies $I^- = \overline{\beta}$, meaning that a^* is too high to attract any insiders, so it does not affect the public good contribution. Therefore, $g(a^*) \leq 0$ if $a^* \leq \underline{\beta}$, and $g(a^*) = 0$ if $a^* \geq \overline{\beta} + B$. Note that Remark 3 provides the upper bound on the increase in contribution, conditional on being an insider. Lemma 3, in contrast, is a statement about the unconditional expected increase in contribution.

3.1 Uniformly Distributed Preferences

Here we assume that β_i is uniformly distributed over $\left[\underline{\beta}, \overline{\beta}\right]$, as in Tabarrok (1998), Barbieri and Malueg (2008), and Kotchen (2009). We first consider the relation between the ideal and expected contributions, and then turn to welfare effects. We let $\bar{a}$ denote the ideal that maximizes the expected contribution.

Proposition 1 *Under uniformly distributed β_i, (i) $\bar{a} = \max\left\{\underline{\beta} + B, \overline{\beta}\right\}$ maximizes $g(a^*)$; it is the unique maximizer if $\gamma > 0$; (ii) when $\overline{\beta} - \underline{\beta} < B$ (so $\bar{a} = \underline{\beta} + B$) the effect of this ideal on expected individual contributions is $g\left(\underline{\beta} + B\right) = \frac{\theta}{1+\theta}\left(B - \frac{\overline{\beta}-\underline{\beta}}{2}\right)$, and when $\overline{\beta} - \underline{\beta} \geq B$ (so $\bar{a} = \overline{\beta}$) the effect of the ideal is $g\left(\overline{\beta}\right) = \frac{\theta}{1+\theta}\frac{B^2}{2\left(\overline{\beta}-\underline{\beta}\right)} = \frac{V}{\overline{\beta}-\underline{\beta}}$.*

3.1.1 Intuition

The effect of the ideal depends on its influence on the depth and width of the group. The width depends on the triplet (I^-, I^*, I^+), which gives the ranges of β_i for agents who join either as under- or over-contributing insiders. By Remark 1 and Equation (4), the increase in insider i's contribution is proportional to $a^* - \beta_i$. Figures 2 and 3 illustrate the contribution-maximizing $\bar{a}$ when $B \geq \overline{\beta} - \underline{\beta}$ and when $B < \overline{\beta} - \underline{\beta}$, respectively.

First suppose $B \geq \overline{\beta} - \underline{\beta}$, where the contribution-maximizing ideal is $\bar{a} = \underline{\beta} + B \geq \overline{\beta}$. Here, every agent becomes an under-contributing insider (Lemma 2), so $(I^-, I^*, I^+) = \left(\underline{\beta}, \overline{\beta}, \overline{\beta}\right)$. The ideal increases contributions of the lowest-preference agent by an amount proportional to B and of the highest-preference agent by an amount proportional to $\underline{\beta} - \overline{\beta} + B$. The trapezoid *abcd* in Figure 2 represents the total effect of the ideal $a^* = \underline{\beta} + B$ on expected contribution.[7] Slightly increasing or decreasing a^* from $\underline{\beta} + B$ are both counterproductive. A small decrease of a^* to $\underline{\beta} + B - \epsilon$, where ϵ is a small positive number, does not alter the triple (I^-, I^*, I^+). The perturbation makes the group shallower but no wider. The trapezoid *hbci* (which is smaller than the original trapezoid *abcd*) represents the contribution under this ideal. Next consider a small increase of a^* to $\underline{\beta} + B + \epsilon$, implying $(I^-, I^*, I^+) = \left(\underline{\beta} + \epsilon, \overline{\beta}, \overline{\beta}\right)$. This perturbation induces the insiders with $\beta_i \in \left[\underline{\beta}, \underline{\beta} + \epsilon\right]$ to drop out, making the group narrower. The trapezoid *fgce* represents the effect of the ideal $a^* = \underline{\beta} + B + \epsilon$. The area of the trapezoid *abcd* is larger than that of the trapezoids *hbci* and *fgce*.

Next suppose $B < \overline{\beta} - \underline{\beta}$, where the contribution-maximizing ideal is $\bar{a} = \overline{\beta}$ leading to $(I^-, I^*, I^+) = \left(\overline{\beta} - B, \overline{\beta}, \overline{\beta}\right)$ by Lemma 2. Under this ideal, agents with $\beta_i \in \left[\overline{\beta} - B, \overline{\beta}\right]$ are under-contributing insiders, while lower-preference agents stay out. The area of triangle *abc* in Figure 3 represents the ideal-induced increase in contributions. Slightly increasing or decreasing a^* from $\overline{\beta}$ are both counterproductive. The perturbation that reduces a^* to $\overline{\beta} - \epsilon$ causes $(I^-, I^*, I^+) = \left(\overline{\beta} - \epsilon - B, \overline{\beta} - \epsilon, \overline{\beta}\right)$; this perturbation encourages agents with $\beta_i \in \left[\overline{\beta} - \epsilon - B, \overline{\beta} - \epsilon\right]$ to contribute, but

[7] We compare areas of polygons to show the effect of perturbations on expected contributions. In every case, the actual effects on contributions are proportional to these areas, with the factors of proportionality equal to either $\frac{\theta}{1+\theta}$ or $\frac{\gamma}{1+\gamma}$. As shorthand, we say that the area "represents" the effect on contributions.

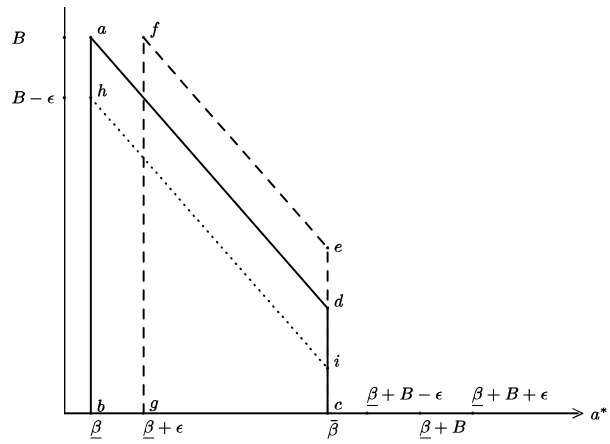

Figure 2: The expected effect of a^* on contribution: $B \geq \overline{\beta} - \underline{\beta}$. The downward sloping lines are all 45 degree lines.

discourages agents with $\beta_i \in (\overline{\beta}-\epsilon, \overline{\beta}]$ from contributing. It does not change contributions from under-contributors: the areas of triangles ghi and abc are equal. However, the perturbation creates a set of over-contributors; the area of triangle ijc represents the decreased contributions from these agents. A perturbation that increases the ideal to $\overline{\beta}+\epsilon$ causes $(I^-, I^*, I^+) = \left(\overline{\beta}+\epsilon-B, \overline{\beta}, \overline{\beta}\right)$. The trapezoid $cdef$, which is smaller than the area of the triangle abc, represents contributions under this perturbation.

3.1.2 Implications

The contribution-maximizing ideal is high enough to encourage all insiders to weakly increase their contribution to the public good ($a_j^i \leq \bar{a}$):

Corollary 1 *Under uniform distribution of β_i, an ideal that maximizes the expected contribution to the public good never elicits over-contribution.*

Proposition 1 implies that when preference heterogeneity is small ($\overline{\beta} - \underline{\beta} \leq B$), the contribution-maximizing ideal attracts all agents to become insiders, and weakly increases their public good contribution; the resulting group is wide but shallow. Here, the agent with the lowest demand for the public good ($\beta_i = \underline{\beta}$) is indifferent between becoming an insider and staying out. This conclusion is consistent with Knack and Keefer's (1997) and Hardin's (2005) empirical finding that in more homogeneous societies, there is typically a higher degree of acceptance to social norms. For non-controversial campaigns, such as those that encourage people to use public transportation to reduce traffic congestion, most people accept the ideal promoted. When preference heterogeneity is large ($\overline{\beta} - \underline{\beta} > B$), the contribution-maximizing ideal equals the baseline contribution of the agent with the highest demand for the public good $\bar{a} = \overline{\beta} = a^b\left(\overline{\beta}\right)$, leading to a narrow but deep group; only agents with sufficient demand for the public good are attracted to be under-contributing insiders; the others remain outsiders.

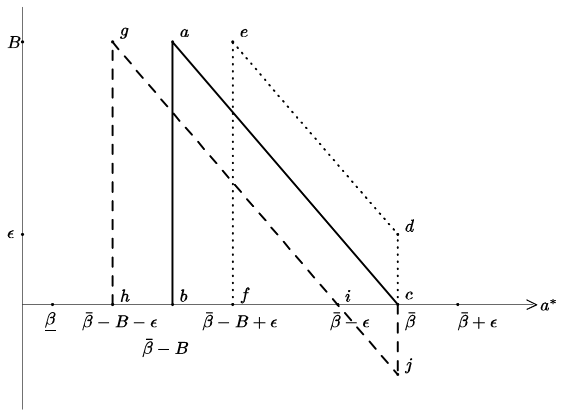

Figure 3: The expected effect of a^* on contribution: $B < \overline{\beta} - \underline{\beta}$. The downward sloping lines are all 45 degree lines.

People have diverse and even contradictory views about climate change, leading to considerable public disagreement about the value of GHG abatement: there is large preference heterogeneity for this public good. Al Gore's receipt of the Nobel prize arguably increased the value of membership, V, and the level of the ideal. A stricter ideal might have changed insiders' behavior; it could also have persuaded some people to remain climate skeptics: a stricter low-carbon ideal/target might discourage some people from identifying with climate insiders, leading to higher emissions.

Proposition 1.ii identifies the expected change of an agent's public good contribution under the contribution-maximizing ideal. The following remark collects additional comparative statics.

Remark 4 *(i) If $\overline{\beta} - \underline{\beta} < B$, the contribution-maximizing ideal, $\bar{a} = \underline{\beta} + B = \underline{\beta} + \left(\frac{2V(1+\theta)}{\theta}\right)^{1/2}$, is increasing in V and decreasing in θ. (ii) If $\overline{\beta} - \underline{\beta} \geq B$, under the contribution-maximizing ideal, $\bar{a} = \overline{\beta}$, agents with $\beta_i \geq \overline{\beta} - B$ become insiders. Here, the expected proportion of insiders is $\frac{B}{\overline{\beta}-\underline{\beta}} = \frac{1}{\overline{\beta}-\underline{\beta}}\left(\frac{2V(1+\theta)}{\theta}\right)^{1/2}$, which increases in V, and decreases with both θ and with agent heterogeneity, $\overline{\beta} - \underline{\beta}$. (iii) In all cases, the expected increase in public good contribution, $\underline{g}(\bar{a})$ increases in V and decreases in $\overline{\beta} - \underline{\beta}$. (iv) The contribution-maximizing ideal, its associated proportion of insiders, and $\underline{g}\left(\overline{a}\right)$ are independent of γ.*

These claims follow from inspection. Note that $B = \left(\frac{2V(1+\theta)}{\theta}\right)^{1/2}$ increases in V, the identity utility of membership, and decreases in θ, which determines under-contributors' cost of departing from the ideal. Therefore, the contribution-maximizing ideal increases in V and decreases in θ when agent heterogeneity is small; moreover, the expected proportion of agents accepting the ideal increases in V and decreases in θ when the agent heterogeneity is large. Because there are no over-contributors under the contribution-maximizing ideal (Corollary 1), the over-contributors' cost

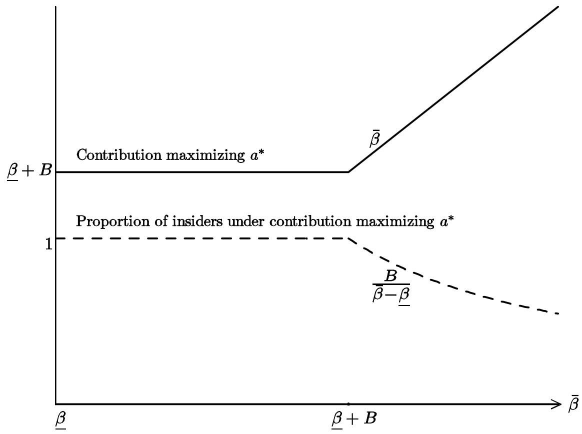

Figure 4: The horizontal line represents $\overline{\beta}$, with the origin set at $\overline{\beta} = \underline{\beta}$. The solid curve shows the contribution-maximizing ideal while the dashed curve shows the proportion of insiders under this ideal, assuming $\underline{\beta} + B > 1$.

of departing from the ideal (determined by γ) has no effect on: the contribution-maximizing ideal; the proportion of insiders; or the level of public good contribution.

Preference heterogeneity, $\overline{\beta} - \underline{\beta}$, affects the contribution-maximizing ideal, the resulting size of the insider group, and the efficacy of identity. In Figure 4 (for $\underline{\beta} + B > 1$), the solid curve shows the contribution-maximizing ideal and the dashed curve shows the proportion of insiders under this ideal. Fixing $\underline{\beta}$, when $\overline{\beta} - \underline{\beta} \leq B$, the contribution maximizing ideal equals $\underline{\beta} + B$, where all agents join the group. When $\overline{\beta} - \underline{\beta} > B$, the contribution-maximizing ideal increases in (and is equal to) $\overline{\beta}$, and the proportion of insiders decreases in $\overline{\beta}$. Social identity is less effective in enhancing public good contribution in a more heterogenous population (Remark 4.iii).

Remark 2 notes that identity may lead to convergence or divergence in contributions across insiders and outsiders . We revisit this issue under the contribution-maximizing ideal, $\bar{a}$. When $\overline{\beta} - \underline{\beta} < B$, where all agents are insiders, Remark 2 implies that identity leads to convergence in contribution among any two agents. When $\overline{\beta} - \underline{\beta} \geq B$, with $a^* = \bar{a}$, the gap in public good contribution between insiders and outsiders increases; here, agents with $\beta_j \geq \overline{\beta} - B$ become insiders and increase their contribution, while outsiders do not change their contribution. This result is consistent with the asymmetric effects, across groups, of energy conservation nudges (Costa and Kahn, 2013). The nudges influence insiders (political liberals/environmentalists) but not outsiders (political conservatives). The outsiders do not identify with the ideology embedded in the nudges, so the nudges widen the gap in energy use between the two groups of people.

3.1.3 Welfare

Welfare assessments for behavioral models can be controversial (Bernheim and Rangel, 2005). We adopt a welfare criterion, $M(a^*)$, that considers only the direct costs and benefits arising from actions, not the psychological effect of identity on utility:

$$M(a^*) = E_{\beta_1,\ldots,\beta_N}\left[\beta_i\left(a_i + \sum_{j\neq i} a_j\right) - \frac{1}{2}a_i^2\right]. \tag{8}$$

The function $M(a^*)$ represents materialistic welfare, permitting an apples-to-apples comparison between the cases with and without identity. The identity-related utility, V, might be manipulated by opinion-makers promoting the ideal or it might include commercial or political advantages that have offsetting costs. For example, a stronger green brand for insiders amounts to a relatively weaker brand for outsiders. By ignoring identity benefits in the welfare criterion, we avoid taking a stand on the extent to which they are manipulated or impose offsetting costs.

The difference between the first-best level of expected contribution, $NE(\beta_i)$, and the expected baseline level, $E(\beta_i)$, is

$$n \equiv (N-1)\,E(\beta_i);$$

n is the unconstrained first-best increase in expected contribution.

Proposition 2 *If*

$$n \geq B\left(\frac{\theta}{1+\theta}\right), \tag{9}$$

then the contribution-maximizing ideal, $a^* = \bar{a}$, maximizes expected welfare, $M(a^*)$, subject to the contribution constraints in Equation (3) and the participation decisions stated in Lemma 2.

Inequality (9) states that the unconstrained first-best increase in expected contribution (n) weakly exceeds the maximum ideal-induced increase in an insider's contribution (the right side of the inequality, by Remark 3).[8] This condition is sufficient but not necessary; it holds if the population (N) is large.

Identity increases agents' expected contribution. Each agent benefits from other agents' higher contribution, but insiders incur a cost from deviating from their baseline, β_i. The contribution-maximizing ideal, $\bar{a}$, is independent of N, so the insiders' expected cost due to contributing more than their baseline level is also independent of N. However, the benefit due to other agents' higher contribution is proportional to N. For sufficiently large N, the agent's benefit resulting from other agents' increased contribution exceeds the cost due to its own increased contribution, over the domain of contributions that can be supported by a group ideal.

[8] A calculation shows that, for the uniform distribution, Inequality (9) implies that $n \geq g(\bar{a})$. If it were possible to write welfare of the representative agent, $M(a^*)$, as a function of only the expected increase in abatement, relative to the baseline, and if in addition that function were quasi-concave, then $n \geq g(\bar{a})$ would imply that the contribution-maximizing ideal also maximizes welfare. However, the proof of Proposition 2 is more complicated, because $M(a^*)$ depends on the distribution of the expected increase in abatement, not only on the expected increase in abatement.

3.2 General Distribution of Preferences

Here we consider the set of preference densities that nest increasing and decreasing density functions, the uniform density, and the single-peaked density.

Assumption 1 *The density function for preferences,* $f(\beta)$*, is continuously differentiable over the support* $\left[\underline{\beta}, \overline{\beta}\right]$*. There exists* $\hat{\beta} \in \left[\underline{\beta}, \overline{\beta}\right]$ *such that* $f'(\beta_i) \geq 0$ *for* $\beta_i \in \left(\underline{\beta}, \hat{\beta}\right)$*, while* $f'(\beta_i) \leq 0$ *for* $\beta_i \in \left(\hat{\beta}, \overline{\beta}\right)$*.*

The density function is weakly decreasing if $\hat{\beta} = \underline{\beta}$, weakly increasing if $\hat{\beta} = \overline{\beta}$, single-peaked if $\hat{\beta}$ is unique, and uniform if $f'(\beta) = 0$ for all β. If $\hat{\beta}$ has multiple values, let $\min \hat{\beta}$ and $\max \hat{\beta}$ denote its minimum and maximum values respectively.

We focus on a most important case, where $\gamma = 0$. Agents with $\beta_i \geq a^*$ identify with the ideal but contribute their baseline, and incur no disutility. The influence-molder does not have to worry about discouraging high-demand agents from contributing, and can focus on the trade-off between attracting more under-contributing members and increasing their contributions.

The contribution-maximizing ideal depends on the distribution's shape, and especially on its skewness. For a symmetric distribution, the contribution-maximizing ideal again depends on the preference heterogeneity, as under the uniform distribution. With $\gamma = 0$, Equation (7) simplifies to

$$g(a^*) = \frac{\theta}{1+\theta} \int_{I^-}^{I^*} (a^* - \beta_i) f(\beta_i) \, d\beta_i. \tag{10}$$

The expected contribution-maximizing ideal maximizes $g(a^*)$. Proposition 3 identifies the trade-off between depth and width in setting the ideal.

Proposition 3 *Under Assumption 1 with* $\gamma = 0$*, if*

$$Bf\left(\underline{\beta}\right) \geq F\left(\underline{\beta} + B\right), \tag{11}$$

then the ideal $\bar{a} = \underline{\beta} + B$ *maximizes* $g(a^*)$*. If inequality (11) fails,* $\bar{a} > \underline{\beta} + B$*, where* $\bar{a}$ *is a solution to*

$$Bf(\bar{a} - B) = F(\bar{a}) - F(\bar{a} - B). \tag{12}$$

Moreover, $\bar{a} \in \left[\max\left\{\underline{\beta} + B, \min \hat{\beta}\right\}, \max \hat{\beta} + B\right]$*.*

When Inequality (11) holds, the contribution-maximizing ideal makes the agent with the lowest demand indifferent between joining; all agents join. If Inequality (11) fails, the contribution-maximizing ideal is larger than $\underline{\beta} + B$, attracting only agents with sufficiently high preference. The shape of the distribution affects whether Inequality (11) holds.

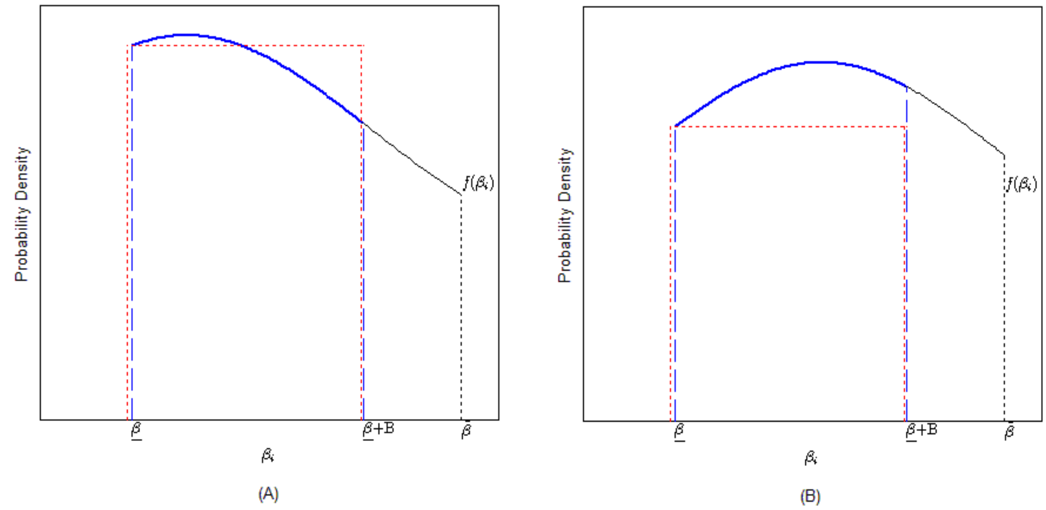

Figure 5: The left figure shows the situation where $Bf\left(\underline{\beta}\right) > F\left(\underline{\beta}+B\right)$, while the right figure shows the situation where $Bf\left(\underline{\beta}\right) \leq F\left(\underline{\beta}+B\right)$.

3.2.1 Intuition

The contribution-maximizing ideal is at least as large as $\underline{\beta}+B$. A lower ideal implies that $I^- = \underline{\beta}$. Increasing that ideal raises both the mass of under-contributors and each of their contributions; there are no effects on over-contributors because $\gamma = 0$. A trade-off exists only when $a^* \geq \underline{\beta}+B$. Inequality (11) means that the cost of a marginal increase of a^* at $\underline{\beta}+B$ outweighs its benefit. With the marginal increase in a^*, agents with $\beta_i = \underline{\beta}$ drop out, reducing the group's width. The density of these agents is $f\left(\underline{\beta}\right)$. Membership causes those agents to increase their contribution by an amount proportional to $a^* - \beta_i = B$. (See Remark 1; the factor of proportionality is $\frac{\theta}{1+\theta}$.) The loss in contributions when those agents drop out is therefore proportional to $Bf\left(\underline{\beta}\right)$. Membership increases contributions by an amount proportional to $a^* - \beta_i$ so a marginal increase in the ideal (leading to an increase in the group's depth) increases contributions by an amount proportional to $F\left(\underline{\beta}+B\right)$, the measure of under-contributing insiders. Inequality (11) states that the loss due to a marginal increase in the ideal, at $a^* = \underline{\beta}+B$, is at least as great as the gain.

This situation arises when the distribution of preference is skewed towards the lower end, as in Figure 5(A). In both Figures 5(A) and 5(B), $Bf\left(\underline{\beta}\right)$ equals the area enclosed in the dotted red lines (a rectangle), and $F\left(\underline{\beta}+B\right)$ is the area below the solid blue line (and enclosed by the blue dashed lines). Therefore Inequality (11) holds if the distribution is sufficiently skewed towards the lower end, as in Figure 5(A). Figure 5(B) illustrates a situation where Inequality (11) does not hold. Here, it is worth excluding the agents with the lowest preference to encourage remaining insiders to contribute more.

Remark 1 also helps in understanding the optimality condition when $\bar{a} > \underline{\beta}+B$, Equation (12). At the solution to this equation, a marginal increase in the ideal causes the lowest preference members, with density $f\left(\bar{a}-B\right)$, to drop out. Membership increases each of these agents' contribution by an amount proportion to B, so their defection reduces contributions by an amount proportional to $Bf(\bar{a}-B)$, the left side of the equation. The higher ideal causes a proportional marginal increase in each of the under-contributing members' contribution; their measure is $F(\bar{a})-F(\bar{a}-B)$, the right side of the equation. At the contribution-maximizing ideal, the marginal loss, due to an increase in the ideal, equals the marginal gain.

For a single-peaked distribution ($\hat{\beta}$ is unique and therefore $\min\hat{\beta}=\max\hat{\beta}$), the last line of Proposition 3 implies that the contribution-maximizing ideal is (weakly) higher than the mode, $\hat{\beta}$, but small enough to induce the modal agents to join. A marginal increase in the ideal from a^* to $a^*+\varepsilon$ ($\varepsilon>0$) causes agents in the interval of $[a^*, a^*+\varepsilon]$ to become under-contributing insiders, and agents in the interval of $[a^*-B, a^*-B+\varepsilon]$ to drop out. Thus, a small increase in a^* changes the measure of insiders by

$$\delta(a^*,\varepsilon)\equiv F(a^*+\varepsilon)-F(a^*)-[F(a^*-B+\varepsilon)-F(a^*-B)].$$

This perturbation also increases the existing members' contributions by an amount proportional to their measure, $F(a^*)-F(a^*-B)$. If $a^*<\widehat{\beta}$ (the single peak), then the perturbation increases both the measure of members, and their contributions; therefore, the contribution-maximizing ideal is no less than $\widehat{\beta}$. If $a^*>\widehat{\beta}+B$, a reduction in the ideal ($\varepsilon<0$) increases the measure of under-contributing members, but decreases members' contributions. The proof of the proposition shows that the net change is positive, so the contribution-maximizing ideal is no greater than $\widehat{\beta}+B$.

3.2.2 Extensions

The following result provides the comparative statics with respect to V.

Remark 5 *Under Assumption 1 with $\gamma=0$, the contribution-maximizing ideal, $\bar{a}$, and the maximum feasible increase in expected contribution, $g(\bar{a})$, are weakly increasing in V.*

The contribution-maximizing ideal may depend on the shape, particularly the skewness, of the distribution. The following corollary discusses the special cases of symmetric ("the least skew"), increasing ("extremely right-skewed") and decreasing ("extremely left-skewed") density functions, and the uniform density.

Corollary 2 *Under Assumption 1 with $\gamma=0$,*

- *(i) if f is symmetric, when $\bar{\beta}-\underline{\beta}>B$, then $\bar{a}$ as determined by (12), maximizes $g(a^*)$; when $\bar{\beta}-\underline{\beta}\le B$ and $f(\underline{\beta})$ is sufficiently close to $f\left(\hat{\beta}\right)$, then $\bar{a}=\underline{\beta}+B$ maximizes $g(a^*)$;*
- *(ii) if $f'\ge 0$ for all β_i, then $\bar{a}$, determined by (12), maximizes $g(a^*)$;*
- *(iii) if $f'\le 0$ for all β_i, then $\bar{a}=\underline{\beta}+B$ maximizes $g(a^*)$; and*
- *(iv) if β_i has a uniform distribution, when $\bar{\beta}-\underline{\beta}\ge B$, then any ideal in $[\underline{\beta}+B,\bar{\beta}]$ maximizes $g(a^*)$, and when $\bar{\beta}-\underline{\beta}<B$, $\bar{a}=\underline{\beta}+B$ maximizes $g(a^*)$.*

Corollary 2.i shows that under symmetric densities, the contribution-maximizing ideal depends on the size of preference heterogeneity, as with the uniform distribution analyzed in Section 3.1. When preference heterogeneity is large, under the contribution-maximizing ideal $\bar{a} > \underline{\beta} + B$, agents with weak preferences remain outsiders; however, when preference heterogeneity is small, the contribution-maximizing ideal induces everyone to be insiders, under the additional condition that $f\left(\underline{\beta}\right)$ is sufficiently close to $f\left(\hat{\beta}\right)$. When this additional condition fails, the benefit of inducing the lowest-preference agents to join does not merit the cost of being unable to induce higher contributions from other members.

Corollary 2.ii and iii state that the contribution-maximizing ideal takes the interior solution of Proposition 3 under increasing density functions, and the corner solution under decreasing density functions. For an increasing density function, $\widehat{\beta} = \overline{\beta}$; the last part of Proposition 3 thus implies that the contribution-maximizing ideal $\bar{a} \geq \overline{\beta}$, i.e. the ideal is always higher than the highest type of the population, $\bar{\beta}$. When agents are more densely located towards the higher end of the distribution, the contribution-maximizing ideal should be sufficiently high to motivate these agents. For a decreasing density function, the contribution-maximizing ideal makes the lowest-type agent indifferent, and all agents become insiders. In this case, agents are more densely located in the lower end of the distribution so the contribution-maximizing ideal motivates these agents to contribute.

A uniform distribution is symmetric, and both weakly increasing and weakly decreasing, so it satisfies all the properties in Corollary 2.i-iii. Corollary 2.iv extends the analysis of Section 3.1 by showing the effect of γ. When preference heterogeneity is large enough, with $\gamma > 0$, Proposition 1 states that the contribution-maximizing ideal equals $\bar{\beta}$, so that no insider is induced to contribute less than their baseline level. However, with $\gamma = 0$, the risk of over-contribution vanishes, so any ideal in $\left[\underline{\beta} + B, \bar{\beta}\right]$ does an equally good job in motivating contribution.

Proposition 4 considers the relation between the contribution-maximizing and the welfare-maximizing ideal, extending Proposition 2.

Proposition 4 *Under Assumption 1 with $\gamma = 0$ and Inequality (9):*

- *(i) There exists an ideal in $\left[\max\left\{\underline{\beta} + B, \min\hat{\beta}\right\}, \max\hat{\beta} + B\right]$ that maximizes welfare, $M\left(a^*\right)$, subject to the contribution constraints in Equation (3) and the participation decisions stated in Lemma 2.*

- *(ii) If $f' \leq 0$ for all β_i, then the contribution-maximizing ideal $\underline{\beta} + B$ is also welfare-maximizing.*

- *(iii) If $f' \geq 0$ for all β_i, the welfare-maximizing ideal is weakly higher than the contribution-maximizing ideal $\bar{a}$ identified in Proposition 3.*

- *(iv) As $N \to \infty$, then for any distribution, the contribution-maximizing ideal and the welfare maximizing ideal converge.*

The contribution-maximizing ideal maximizes other agents' contributions, but also causes an agent's contribution to be higher than his baseline. The increase in $M(a^*)$ arising from the first effect is proportional to N, and the decrease in $M(a^*)$ arising from the second effect does not depend on N. When N is sufficiently large, the first effect dominates, so the contribution-maximizing ideal also maximizes welfare (Proposition 4.iv). Propositions 3 and 4.i imply that the welfare-maximizing ideal lies in the same interval as the contribution-maximizing ideal: $\left[\max\left\{\underline{\beta}+B, \min\hat{\beta}\right\}, \max\hat{\beta}+B\right]$.

Agents' individual rationality constrains the ability of the ideal to raise contributions. When the density is weakly decreasing, this constraint binds in the welfare maximization problem: agents would have higher welfare if it were possible to increase the contribution, but that increase is not feasible (Proposition 4.ii). In contrast, for increasing densities, the welfare-maximizing ideal exceeds the contribution-maximizing level (Proposition 4.iii). To understand this ranking, consider the welfare effect of a marginal increase in the ideal, beginning with the interior contribution-maximizing level. This change creates only a second order effect on the expected contribution (and thus an agent's expected external effects from others' actions), but it has a first order effect on an agent's welfare from own action. The higher ideal causes the lowest-type insiders to leave the group, reducing their contribution cost substantially, and increasing aggregate welfare. The higher ideal raises the remaining insiders' contribution costs, lowering aggregate welfare. However, increasing density functions assign more weight to higher-type insiders; their welfare loss from own actions is small.[9] Therefore the increase in welfare due to the exit of marginal insiders exceeds the reduction in welfare due to higher costs for remaining insiders. Because an increase in the ideal beginning with the contribution-maximizing level increases aggregate welfare, the welfare-maximizing ideal exceeds the contribution-maximizing ideal.

4 Conclusion

Free-riding typically leads to under-provision of a public good. The global and persistent nature of GHG pollution exacerbates the under-provision of climate services. National sovereignty and differing views about the severity of climate change make it difficult to reach an effective international agreement on GHG regulation. Voluntary contribution to the public good of emission abatement might nevertheless be important in curbing climate change. We adopt a behavioral perspective, showing how identity-related benefits can influence voluntary public good contributions. The agents in our setting might be individuals, in which case the identity-related benefits are primarily psychological. Agents might also be companies or cities or states, in which case the identity-related benefits may include both status and commercial or political benefits. A member whose public good contribution differs from the group ideal has a loss in identity benefit.

We examine the effect of the ideal, without attempting to explain the mechanism that determines it. Opinion-molders, including politicians, educators, and religious leaders, might influence the ideal through public policies, media, school, and church.

[9] A type-β_i insider's welfare from own action, denoted by ϖ, is $\varpi = \beta_i a_i - \frac{a_i^2}{2} = \beta_i\left(\beta_i + \frac{\theta(a^*-\beta_i)}{1+\theta}\right) - \frac{1}{2}\left(\beta_i + \frac{\theta(a^*-\beta_i)}{1+\theta}\right)^2$. We have $\frac{d\varpi}{da^*} = -\left(\frac{\theta}{1+\theta}\right)^2(a^*-\beta_i)$.

Nonenforceable international agreements such as the Kyoto Protocol or the 2015 Paris climate agreement can also alter the group ideal. Manipulation of the group ideal can help alleviate the free-riding problem associated with voluntary public good contributions. A change in the ideal can alter both the depth and the width of the group associated with it. A higher ideal encourages insiders to contribute more, but reduces the range of the agents who become insiders. For uniformly distributed preferences, homogeneity tends to make the contribution-maximizing group wide but shallow, and heterogeneity tends to make that group narrow but deep. The same comparison holds for symmetric general distributions if members do not obtain disutility from over-contributing.

The contribution-maximizing ideal also maximizes welfare if the number of agents is large. If members face no utility loss from contributions above the group ideal, then the welfare-maximizing ideal equals the contribution-maximizing ideal if the density of types is decreasing; the welfare-maximizing ideal is higher than the contribution-maximizing ideal if the density of types is increasing.

A Proofs

Proof. (Lemma 1) Because both alternative forms of $U_j(a_j|a^*)$ in the right side of (2) are concave, by first order conditions, we have

$$\arg\max_{a_j} \beta_j a_j - \frac{1}{2}a_j^2 - \frac{\theta}{2}(a_j - a^*)^2 + V = \beta_j + \frac{\theta(a^* - \beta_j)}{1+\theta}, \tag{13}$$

and

$$\arg\max_{a_j} \beta_j a_j - \frac{1}{2}a_j^2 - \frac{\gamma}{2}(a_j - a^*)^2 + V = \beta_j + \frac{\gamma(a^* - \beta_j)}{1+\gamma}. \tag{14}$$

If $\theta = \gamma$, then trivially the solution is given by either of these two expressions, which are the same. Now suppose $\theta \neq \gamma$. Observe that

$$\beta_j + \frac{\eta(a^* - \beta_j)}{1+\eta} \leq a^* \iff \beta_j \leq a^* \tag{15}$$

for both $\eta \in \{\theta, \gamma\}$. Suppose $\beta_j \leq a^*$. Then the right side of (13) is no higher than a^* and thus optimal given $a_j \leq a^*$. We then claim that $\frac{dU_j(a_j|a^*)}{da_j} \leq 0$ for $\forall a_j \geq a^*$: Concavity of $U_j(a_j|a_j \geq a^*)$ implies that $\frac{dU_j(a_j|a^*)}{da_j} \leq 0$ for all $a_j \geq \beta_j + \frac{\gamma(\beta_j - a^*)}{1+\gamma}$ by (14), and meanwhile $a_j \geq a^*$ and $\beta_j \leq a^*$ imply $a_j \geq a^* \geq \beta_j + \frac{\gamma(\beta_j - a^*)}{1+\gamma}$ by (15). Therefore, $\frac{dU_j(a_j|a^*)}{da_j} \leq 0$ for all $a_j \geq a^*$. Consequently the right side of (13) is optimal if $\beta_j \leq a^*$.

Suppose $\beta_j > a^*$, then the right side of (14) is no lower than a^* and thus optimal for $a_j \in [a^*, \infty)$. A logic similar to the above shows that $\frac{dU_j(a_j|a^*)}{da_j} \geq 0$ for $\forall a_j \leq a^*$ and consequently the right side of (14) is optimal if $\beta_j > a^*$. ∎

Proof. (Remark 2) Using Equations (1) and (3), we have the following. If $a^* \geq \beta_j > \beta_i$, meaning that both agents with β_i and β_j will be under-contributing insiders by Footnote 3, then

$$a^b(\beta_j) - a^b(\beta_i) = \beta_j - \beta_i > \frac{\beta_j - \beta_i}{1+\theta} = a^i(\beta_j) - a^i(\beta_i).$$

If $a^* < \beta_i < \beta_j$, meaning that both agents with β_i and β_j will be over-contributing insiders by Footnote 3, then

$$a^b(\beta_j) - a^b(\beta_i) = \beta_j - \beta_i \geq \frac{\beta_j - \beta_i}{1+\gamma} = a^i(\beta_j) - a^i(\beta_i).$$

If $\beta_j > a^* \geq \beta_i$, meaning that the agent with β_i will be an under-contributor while the other an over-contributor as insiders, then

$$\begin{aligned}
&a^b(\beta_j) - a^b(\beta_i) \\
&= \beta_j - \beta_i \\
&> \left[\beta_j + \frac{\gamma(a^* - \beta_j)}{1+\gamma}\right] - \left[\beta_i + \frac{\theta(a^* - \beta_i)}{1+\theta}\right] \\
&= a^i(\beta_j) - a^i(\beta_i),
\end{aligned}$$

because $\frac{\gamma(a^*-\beta_j)}{1+\gamma} - \frac{\theta(a^*-\beta_i)}{1+\theta} < 0$. ■

Proof. (Lemma 2) Comparing (5) and (6) while using the definitions of B and D, we have the following.

$$\begin{aligned}
&\frac{1}{2}\frac{\beta_i(\beta_i+2\theta a^*)-\theta a^{*2}}{1+\theta}+V\geq\frac{\beta_i^2}{2} \qquad (16)\\
&\Leftrightarrow 2V(1+\theta)\geq\theta\beta_i^2-2\theta a^*\beta_i+\theta a^{*2}\\
&\Leftrightarrow \frac{2V(1+\theta)}{\theta}\geq(\beta_i-a^*)^2\\
&\Leftrightarrow a^*-B\leq\beta_i\leq a^*+B.
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\frac{1}{2}\frac{\beta_i(\beta_i+2\gamma a^*)-\gamma a^{*2}}{1+\gamma}+V\geq\frac{1}{2}\beta_i^2\\
&\Leftrightarrow a^*-D\leq\beta_i\leq a^*+D.
\end{aligned}$$

If $\beta_i \leq a^*$, the agent will be an under-contributor as an insider, and if Inequality (16) is satisfied, the utility of being an (under-contributing) insider will be not lower than that of being an outsider. Combining $\beta_i \leq a^*$ and Inequality (16), an agent will identify with the ideal and become an under-contributing insider if and only if $a^* - B \leq \beta_i \leq a^*$. Similar logic with $\beta_i \geq a^*$ shows that an agent will identify with the ideal and become an over-contributing insider if and only if $a^* < \beta_i \leq a^* + D$. Combining the above analysis and taking into account the constraint $\underline{\beta} \leq \beta_i \leq \overline{\beta}$, the lemma is proved. ■

The proofs of the propositions make use of the following technical lemma.

Lemma 4 *For any function defined by*

$$Y(x) = \int_{\underline{\beta}}^{z(x)} y(\beta_i, x)\, d\beta_i,$$

if $y(\beta_i, x)$ is bounded and continuous and $z(x)$ is continuous, then Y is continuous.

Proof. (Proof of Lemma 4) We want to show $\lim_{x\to c} Y(x) = Y(c)$ for any c. Since y is bounded, there exists a finite number W such that $y(\beta_i, x) \leq W$ for $\forall (\beta_i, x)$. By definition,

$$\begin{aligned}
Y(x)-Y(c) &= \int_{\underline{\beta}}^{z(x)} y(\beta_i,x)\,d\beta_i - \int_{\underline{\beta}}^{z(c)} y(\beta_i,c)\,d\beta_i\\
&= \int_{\underline{\beta}}^{z(x)} y(\beta_i,x)\,d\beta_i - \int_{\underline{\beta}}^{z(x)} y(\beta_i,c)\,d\beta_i + \int_{\underline{\beta}}^{z(x)} y(\beta_i,c)\,d\beta_i - \int_{\underline{\beta}}^{z(c)} y(\beta_i,c)\,d\beta_i\\
&= \int_{\underline{\beta}}^{z(x)} \left[y(\beta_i,x)-y(\beta_i,c)\right]d\beta_i + \int_{z(c)}^{z(x)} y(\beta_i,c)\,d\beta_i.
\end{aligned}$$

This implies

$$\begin{aligned}
|Y(x) - Y(c)| &\le \left| \int_{\underline{\beta}}^{z(x)} [y(\beta_i, x) - y(\beta_i, c)] \, d\beta_i \right| + \left| \int_{z(c)}^{z(x)} y(\beta_i, c) \, d\beta_i \right| \\
&\le \int_{\underline{\beta}}^{z(x)} |y(\beta_i, x) - y(\beta_i, c)| \, d\beta_i + \left| \int_{z(c)}^{z(x)} y(\beta_i, c) \, d\beta_i \right| \\
&\le \int_{\underline{\beta}}^{z(x)} |y(\beta_i, x) - y(\beta_i, c)| \, d\beta_i + W |z(x) - z(c)|,
\end{aligned}$$

where the first two inequalities follow from the triangle inequality, and the last inequality follows from $y \le W$. Next, the continuity of z and y implies $\lim_{x \to c} z(x) = z(c)$ and $\lim_{x \to c} y(\beta_i, x) = y(\beta_i, c)$. Consequently,

$$\lim_{x \to c} \int_{\underline{\beta}}^{z(x)} |y(\beta_i, x) - y(\beta_i, c)| \, d\beta_i + W |z(x) - z(c)| = 0,$$

which implies that

$$0 \le \lim_{x \to c} |Y(x) - Y(c)| \le 0.$$

By the Squeeze Theorem, we have $\lim_{x \to c} |Y(x) - Y(c)| = 0$, and therefore $\lim_{x \to c} Y(x) = Y(c)$. ■

Proof. (Proposition 1) Following Lemma 3, we consider only $a^* \in \left(\underline{\beta}, \bar{\beta} + B\right)$. The following exhaustive list shows all the possible triplets for (I^-, I^*, I^+) :

$$\begin{aligned}
&\left\{\text{All possible } \left(I^-, I^*, I^+\right)\right\} \\
&= \left\{ \begin{array}{c} \left(\underline{\beta}, a^*, a^* + D\right), \left(\underline{\beta}, a^*, \overline{\beta}\right), \left(\underline{\beta}, \overline{\beta}, \overline{\beta}\right) \\ \left(a^* - B, a^*, a^* + D\right), \left(a^* - B, a^*, \overline{\beta}\right), \left(a^* - B, \overline{\beta}, \overline{\beta}\right) \end{array} \right\}.
\end{aligned}$$

Recall that $a^* > \overline{\beta} \implies I^* = \overline{\beta}$, $a^* + D > \overline{\beta} \implies I^+ = \overline{\beta}$, and $a^* - B < \underline{\beta} \implies I^- = \underline{\beta}$. Meanwhile the possibilities with $I^* = \underline{\beta}$, $I^+ = \underline{\beta}$ or $I^- = \overline{\beta}$ are eliminated because $a^* \in \left(\underline{\beta}, \bar{\beta} + B\right)$. Under uniform distribution, the density function is $f(\beta_i) = \frac{1}{\overline{\beta} - \underline{\beta}}$. Following (7), for any given triplet (I^-, I^*, I^+) the increase in contribution is given by

$$\begin{aligned}
g(a^*) &= \frac{1}{\overline{\beta} - \underline{\beta}} \left(\frac{\gamma}{1+\gamma} \int_{I^*}^{I^+} (a^* - \beta_i) \, d\beta_i + \frac{\theta}{1+\theta} \int_{I^-}^{I^*} (a^* - \beta_i) \, d\beta_i \right) \qquad (17) \\
&= \frac{1}{\overline{\beta} - \underline{\beta}} \left(\frac{\gamma}{1+\gamma} \underbrace{\left(I^+ - I^*\right) \left(a^* - \frac{1}{2} \left(I^+ + I^*\right) \right)}_{\text{Term 1}} + \frac{\theta}{1+\theta} \underbrace{\left(I^* - I^-\right) \left(a^* - \frac{1}{2} \left(I^* + I^-\right) \right)}_{\text{Term 2}} \right).
\end{aligned}$$

Claim 0: $g(a^*)$ is continuous.
To see this, rewrite $g(a^*)$ as

$$g(a^*) = \frac{1}{\overline{\beta} - \underline{\beta}} \left[\begin{array}{c} \frac{\gamma}{1+\gamma} \left(\int_{\underline{\beta}}^{I^+} (a^* - \beta_i) \, d\beta_i - \int_{\underline{\beta}}^{I^*} (a^* - \beta_i) \, d\beta_i \right) \\ + \frac{\theta}{1+\theta} \left(\int_{\underline{\beta}}^{I^*} (a^* - \beta_i) \, d\beta_i - \int_{\underline{\beta}}^{I^-} (a^* - \beta_i) \, d\beta_i \right) \end{array} \right],$$

where for each integral, the integrands are continuous and bounded (since we focus on $a^* \in \left(\underline{\beta}, \bar{\beta}+B\right)$ by Lemma 3), and the upper supports (I^* or I^+) are also continuous in a^*. By Lemma 4, $g\left(a^*\right)$ is continuous.

Claim 1: $g\left(\underline{\beta}+B\right) > g\left(a^*\right)$ for all $a^* < \underline{\beta}+B$.

Observe that $a^* \leq \underline{\beta}+B \implies a^*-B \leq \underline{\beta}$ implying $I^-=\underline{\beta}$. Therefore the triplet $\left(I^-, I^*, I^+\right)$ can be either $\left(\underline{\beta}, a^*, a^*+D\right)$, $\left(\underline{\beta}, a^*, \bar{\beta}\right)$, or $\left(\underline{\beta}, \bar{\beta}, \bar{\beta}\right)$. For each of these possibilities, Term 1 and Term 2 in the last line of (17) are expressed by the following.

Case	$\left(I^-, I^*, I^+\right)$	Term 1	Term 2
1	$\left(\underline{\beta}, a^*, a^*+D\right)$	$-\frac{1}{2}D^2$	$\frac{1}{2}\left(a^*-\underline{\beta}\right)^2$
2	$\left(\underline{\beta}, a^*, \bar{\beta}\right)$	$-\frac{1}{2}\left(\bar{\beta}-a^*\right)^2$	$\frac{1}{2}\left(a^*-\underline{\beta}\right)^2$
3	$\left(\underline{\beta}, \bar{\beta}, \bar{\beta}\right)$	0	$\left(\bar{\beta}-\underline{\beta}\right)\left(a^*-\frac{1}{2}\left(\bar{\beta}+\underline{\beta}\right)\right)$

Substituting Term 1 and Term 2 in the above table back to (17), we observe that for all the cases, $g\left(a^*\right)$ is increasing, given the boundaries of a^* defined by the corresponding $\left(I^-, I^*, I^+\right)$. When a^* changes, the triplet $\left(I^-, I^*, I^+\right)$ may switch from one case to another. However, given the continuity of $g\left(a^*\right)$ and that $g\left(a^*\right)$ is increasing for each of the three possible triplet cases, we have $g\left(\underline{\beta}+B\right) > g\left(a^*\right)$ for all $a^* < \underline{\beta}+B$.

Claim 2: $g\left(\bar{\beta}\right) \geq g\left(a^*\right)$ for all $a^* < \bar{\beta}$ and strictly so if $\gamma > 0$.

Observe that $a^* \leq \bar{\beta} \implies I^* = a^*$, since we consider only $a^* > \underline{\beta}$. Therefore the triplet $\left(I^-, I^*, I^+\right)$ can be either $\left(\underline{\beta}, a^*, a^*+D\right)$, $\left(\underline{\beta}, a^*, \bar{\beta}\right)$, $\left(a^*-B, a^*, a^*+D\right)$, or $\left(a^*-B, a^*, \bar{\beta}\right)$. For each of these possibilities, Term 1 and Term 2 in the last line of (17) are expressed by the following.

Case	$\left(I^-, I^*, I^+\right)$	Term 1	Term 2
1	$\left(\underline{\beta}, a^*, a^*+D\right)$	$-\frac{1}{2}D^2$	$\frac{1}{2}\left(a^*-\underline{\beta}\right)^2$
2	$\left(\underline{\beta}, a^*, \bar{\beta}\right)$	$-\frac{1}{2}\left(\bar{\beta}-a^*\right)^2$	$\frac{1}{2}\left(a^*-\underline{\beta}\right)^2$
3	$\left(a^*-B, a^*, a^*+D\right)$	$-\frac{1}{2}D^2$	$\frac{1}{2}B^2$
4	$\left(a^*-B, a^*, \bar{\beta}\right)$	$-\frac{1}{2}\left(\bar{\beta}-a^*\right)^2$	$\frac{1}{2}B^2$

Substituting Term 1 and Term 2 in the above table back to (17), we observe that for all the cases, $g\left(a^*\right)$ is (weakly) increasing in a^*. Similarly to the logic above, given the continuity of $g\left(a^*\right)$, we have $g\left(\bar{\beta}\right) \geq g\left(a^*\right)$ for all $a^* < \bar{\beta}$.

Now suppose $\gamma > 0$. If $\bar{\beta}-\underline{\beta} > B+D$, when a^* changes, the triplet $\left(I^-, I^*, I^+\right)$ may switch from Case 1 to Case 3 and then Case 4. If $\bar{\beta}-\underline{\beta} \leq B+D$, while the triplet $\left(I^-, I^*, I^+\right)$ may switch, Case 3 never occurs. Note that $g\left(a^*\right)$ is strictly increasing in a^* for all the cases except Case 3, where $g\left(a^*\right)$ is constant in a^*. Therefore, in both of these two situations, $g\left(\bar{\beta}\right) > g\left(a^*\right)$ for all $a^* < \bar{\beta}$, given the continuity of $g\left(a^*\right)$.

Claim 3: For $a^* \geq \max\left\{\bar{\beta}, \underline{\beta}+B\right\}$, $g\left(a^*\right)$ is uniquely maximized at $\max\left\{\bar{\beta}, \underline{\beta}+B\right\}$.

Consider $a^* \geq \max\left\{\bar{\beta}, \underline{\beta}+B\right\}$. $a^* \geq \bar{\beta}$ implies $I^* = \bar{\beta}$ and $I^+ = \bar{\beta}$, while $\bar{\beta}+B \geq a^* \geq \underline{\beta}+B$ implies $I^- = a^*-B$. Substitute I^-, I^*, and I^+ and observe that in the last line of (17) Term 1 is zero and Term 2 is:

$$
\begin{aligned}
&\left(\bar{\beta}-\left(a^*-B\right)\right)\left(a^*-\frac{1}{2}\left(\bar{\beta}+\left(a^*-B\right)\right)\right)\\
&=\frac{1}{2}\left(B^2-\left(a^*-\bar{\beta}\right)^2\right),
\end{aligned}
$$

which is decreasing in a^* because $a^* \geq \bar{\beta}$ and thus maximized uniquely at the lower bound $a^* = \max\{\bar{\beta}, \underline{\beta} + B\}$.

Combining Claims 1-3 above, Part (i) is thus proved.

For Part (ii), when $\overline{\beta} - \underline{\beta} < B$, $\overline{a} = \max\{\bar{\beta}, \underline{\beta} + B\} = \underline{\beta} + B$, under which $(I^-, I^*, I^+) = (\underline{\beta}, \overline{\beta}, \overline{\beta})$. Substituting these into (17), we have $g(\underline{\beta} + B) = \frac{\theta}{1+\theta}\left(B - \frac{\overline{\beta}-\underline{\beta}}{2}\right)$. When $\overline{\beta} - \underline{\beta} \geq B$, $\overline{a} = \max\{\bar{\beta}, \underline{\beta} + B\} = \bar{\beta}$, under which $(I^-, I^*, I^+) = (\overline{\beta} - B, \overline{\beta}, \overline{\beta})$. Substituting these into (17), we have $g(\overline{\beta}) = \frac{\theta}{1+\theta}\frac{B^2}{2(\overline{\beta}-\underline{\beta})} = \frac{V}{\overline{\beta}-\underline{\beta}}$. ∎

Proof. (Proof of Proposition 2) Suppose that $n \geq B\left(\frac{\theta}{1+\theta}\right)$. Since β_i's are i.i.d, Equation (8) can be written as

$$\begin{aligned} M(a^*) &= E\left[\beta_i a_i - \frac{1}{2}a_i^2\right] + E(\beta_i)(N-1)E(a_i) \\ &= E\left[\beta_i a_i - \frac{1}{2}a_i^2\right] + nE(a_i). \end{aligned} \tag{18}$$

Given an insider's contribution in Equation (3) and an outsider's contribution β_i, after some simple algebra we obtain

$$\begin{aligned} &E(a_i) \\ &= \int_{\underline{\beta}}^{\bar{\beta}} \beta_i dF(\beta_i) + \frac{\theta}{1+\theta}\int_{I^-}^{I^*}(a^* - \beta_i)\,dF(\beta_i) + \frac{\gamma}{1+\gamma}\int_{I^*}^{I^+}(a^* - \beta_i)\,dF(\beta_i), \end{aligned}$$

and

$$\begin{aligned} &E\left[\beta_i a_i - \frac{1}{2}a_i^2\right] \\ &= \frac{1}{2}\int_{\underline{\beta}}^{\bar{\beta}} \beta_i^2 dF(\beta_i) - \frac{1}{2}\int_{I^-}^{I^*}\left(\frac{\theta(a^* - \beta_i)}{1+\theta}\right)^2 dF(\beta_i) - \frac{1}{2}\int_{I^*}^{I^+}\left(\frac{\gamma(a^* - \beta_i)}{1+\gamma}\right)^2 dF(\beta_i). \end{aligned}$$

Substituting both expressions into Equation (18) to obtain:

$$\begin{aligned} M(a^*) = &\int_{\underline{\beta}}^{\bar{\beta}}\left(\frac{1}{2}\beta_i^2 + n\beta_i\right)dF(\beta_i) + \frac{n\theta}{1+\theta}\int_{I^-}^{I^*}(a^* - \beta_i)\,dF(\beta_i) + \frac{n\gamma}{1+\gamma}\int_{I^*}^{I^+}(a^* - \beta_i)\,dF(\beta_i) \\ &-\frac{1}{2}\left(\int_{I^-}^{I^*}\left(\frac{\theta(a^* - \beta_i)}{1+\theta}\right)^2 dF(\beta_i) + \int_{I^*}^{I^+}\left(\frac{\gamma(a^* - \beta_i)}{1+\gamma}\right)^2 dF(\beta_i)\right). \end{aligned} \tag{19}$$

Following the same logic as Claim 0 in the proof of Proposition 1 and using Lemma 4, we obtain the continuity of $M(a^*)$.

Using the density function $f(\beta_i) = \frac{1}{\bar{\beta}-\underline{\beta}}$ and rearranging the above equation yields,

$$\begin{aligned}
\left(\bar{\beta}-\underline{\beta}\right) M(a^*) &= \int_{\underline{\beta}}^{\bar{\beta}} \left(\frac{1}{2}\beta_i^2 + n\beta_i\right) d\beta_i \\
&+ \frac{\gamma}{1+\gamma} \underbrace{\int_{I^*}^{I^+} \left[-\frac{1}{2}\frac{\gamma\left(a^*-\beta_i\right)^2}{1+\gamma} + n\left(a^*-\beta_i\right)\right] d\beta_i}_{\text{Term 1}} \\
&+ \frac{\theta}{1+\theta} \underbrace{\int_{I^-}^{I^*} \left[-\frac{1}{2}\frac{\theta\left(a^*-\beta_i\right)^2}{1+\theta} + n\left(a^*-\beta_i\right)\right] d\beta_i}_{\text{Term 2}}.
\end{aligned} \tag{20}$$

Term 1 in (20), denoted by $T_1(I^*, I^+)$, can be written as

$$T_1\left(I^*, I^+\right) = -\frac{1}{6}\left(\frac{\gamma}{1+\gamma}\right)\left(\left(I^+-a^*\right)^3 + \left(a^*-I^*\right)^3\right) + n\left(I^+-I^*\right)\left(a^* - \frac{1}{2}\left(I^+ + I^*\right)\right), \tag{21}$$

while Term 2 in (20), denoted by $T_2(I^-, I^*)$, can be written as

$$T_2\left(I^-, I^*\right) = -\frac{1}{6}\left(\frac{\theta}{1+\theta}\right)\left(\left(I^*-a^*\right)^3 + \left(a^*-I^-\right)^3\right) + n\left(I^*-I^-\right)\left(a^* - \frac{1}{2}\left(I^* + I^-\right)\right). \tag{22}$$

Similar to the proof of Proposition 1, we shall structure our proof according to triplets (I^-, I^*, I^+) and focus on $a^* \in \left(\underline{\beta}, \bar{\beta}+B\right)$ because a^* outside this domain is counterproductive or ineffective. This rules out the possibilities with $I^* = \underline{\beta}$, $I^+ = \underline{\beta}$ or $I^- = \overline{\beta}$. So we have

$$\begin{aligned}
&\left\{\text{All possible } \left(I^-, I^*, I^+\right)\right\} \\
&= \left\{ \begin{array}{c} \left(\underline{\beta}, a^*, a^*+D\right), \left(\underline{\beta}, a^*, \overline{\beta}\right), \left(\underline{\beta}, \overline{\beta}, \overline{\beta}\right) \\ \left(a^*-B, a^*, a^*+D\right), \left(a^*-B, a^*, \overline{\beta}\right), \left(a^*-B, \overline{\beta}, \overline{\beta}\right) \end{array} \right\}.
\end{aligned}$$

We first establish how T_1 and T_2 change with a^* under each possibility. For $T_1(I^*, I^+)$, the arguments have 3 possibilities: duplets (a^*, a^*+D), $\left(a^*, \overline{\beta}\right)$, and $\left(\overline{\beta}, \overline{\beta}\right)$. By (21),

$$T_1(a^*, a^*+D) = -\frac{1}{6}\left(\frac{\gamma}{1+\gamma}\right) D^3 - \frac{n}{2} D^2,$$

which is a constant. Next

$$T_1\left(a^*, \overline{\beta}\right) = -\frac{1}{6}\left(\frac{\gamma}{1+\gamma}\right)\left(\bar{\beta}-a^*\right)^3 - \frac{n}{2}\left(\bar{\beta}-a^*\right)^2,$$

where

$$\frac{dT_1\left(a^*, \overline{\beta}\right)}{da^*} = \left(\bar{\beta}-a^*\right)\left(\frac{1}{2}\left(\frac{\gamma}{1+\gamma}\right)\left(\bar{\beta}-a^*\right) + n\right) \geq 0,$$

as duplet $(I^*, I^+) = \left(a^*, \overline{\beta}\right)$ implies $\bar{\beta} \geq a^*$. Lastly

$$T_1\left(\overline{\beta}, \overline{\beta}\right) = 0, \tag{23}$$

which is a constant.

For $T_2\left(I^-, I^*\right)$, the arguments have 4 possibilities: duplets $\left(\underline{\beta}, a^*\right), \left(a^* - B, a^*\right), \left(\underline{\beta}, \bar{\beta}\right)$ and $\left(a^* - B, \overline{\beta}\right)$. By (22),

$$T_2\left(\underline{\beta}, a^*\right) = -\frac{1}{6}\left(\frac{\theta}{1+\theta}\right)\left(a^* - \underline{\beta}\right)^3 + \frac{n}{2}\left(a^* - \underline{\beta}\right)^2$$

where

$$\frac{dT_2\left(\underline{\beta}, a^*\right)}{da^*} = \left(a^* - \underline{\beta}\right)\left(n - \frac{1}{2}\left(\frac{\theta}{1+\theta}\right)\left(a^* - \underline{\beta}\right)\right) \geq 0,$$

given $n \geq B\left(\frac{\theta}{1+\theta}\right)$. This is because duplet $\left(I^-, I^*\right) = \left(\underline{\beta}, a^*\right)$ implies that $a^* - B \leq \underline{\beta}$, and thus $n \geq \frac{B}{2}\left(\frac{\theta}{1+\theta}\right) \geq \frac{1}{2}\left(\frac{\theta}{1+\theta}\right)\left(a^* - \underline{\beta}\right)$. Then

$$T_2\left(a^* - B, a^*\right) = -\frac{1}{6}\left(\frac{\theta}{1+\theta}\right) B^3 + \frac{n}{2} B^2,$$

which is a constant. Next

$$T_2\left(\underline{\beta}, \bar{\beta}\right) = -\frac{1}{6}\left(\frac{\theta}{1+\theta}\right)\left(\left(\bar{\beta} - a^*\right)^3 + \left(a^* - \underline{\beta}\right)^3\right) + n\left(\bar{\beta} - \underline{\beta}\right)\left(a^* - \frac{1}{2}\left(\bar{\beta} + \underline{\beta}\right)\right),$$

where

$$\begin{aligned}
\frac{dT_2\left(\underline{\beta}, \bar{\beta}\right)}{da^*} &= \left(\bar{\beta} - \underline{\beta}\right)\left[-\frac{1}{2}\left(\frac{\theta}{1+\theta}\right)\left(2a^* - \left(\bar{\beta} + \underline{\beta}\right)\right) + n\right] \\
&\geq \left(\bar{\beta} - \underline{\beta}\right)\left[-\frac{1}{2}\left(\frac{\theta}{1+\theta}\right) 2B + n\right] \\
&\geq 0.
\end{aligned}$$

where the first inequality makes use of

$$I^- = \underline{\beta} \Rightarrow a^* - B \leq \underline{\beta} \Rightarrow \bar{\beta} + \underline{\beta} > 2\underline{\beta} \geq 2\left(a^* - B\right),$$

and the second inequality uses $n \geq B\left(\frac{\theta}{1+\theta}\right)$. Lastly,

$$T_2\left(a^* - B, \overline{\beta}\right) = -\frac{1}{6}\left(\frac{\theta}{1+\theta}\right)\left(\left(\overline{\beta} - a^*\right)^3 + B^3\right) + \frac{n}{2}\left[B^2 - \left(\overline{\beta} - a^*\right)^2\right],$$

where

$$\begin{aligned}
\frac{dT_2\left(a^* - B, \overline{\beta}\right)}{da^*} &= \left(\overline{\beta} - a^*\right)\left(n + \frac{1}{2}\left(\frac{\theta}{1+\theta}\right)\left(\overline{\beta} - a^*\right)\right) \qquad (24) \\
&\leq 0.
\end{aligned}$$

To show this inequality, note that $a^* \geq \overline{\beta}$ (because $I^* = \bar{\beta}$ here), and also $n \geq \frac{1}{2}\left(\frac{\theta}{1+\theta}\right) B > \frac{1}{2}\left(\frac{\theta}{1+\theta}\right)\left(a^* - \overline{\beta}\right)$ by $a^* < \overline{\beta} + B$; thus in the right side of the first line of (24), the first term is nonpositive while the second term is positive, which implies that $\frac{dT_2\left(a^* - B, \overline{\beta}\right)}{da^*}$ is nonpositive.

Now we are ready to prove the proposition.

Claim 1: $M\left(\underline{\beta}+B\right) \geq M\left(a^*\right)$ for all $a^* \leq \underline{\beta}+B$.

Observe that $a^* \leq \underline{\beta}+B \implies a^*-B \leq \underline{\beta}$ implying $I^- = \underline{\beta}$. Therefore the triplet (I^-, I^*, I^+) can be either $\left(\underline{\beta}, a^*, a^*+D\right)$, $\left(\underline{\beta}, a^*, \bar{\beta}\right)$, or $\left(\underline{\beta}, \bar{\beta}, \bar{\beta}\right)$. For all these possibilities, as discussed above, Term 1 and Term 2 are either constant or increasing in a^*, which are summarized in the following table.

(I^-, I^*, I^+)	Term 1	Term 2
$\left(\underline{\beta}, a^*, a^*+D\right)$	$\frac{dT_1(a^*,a^*+D)}{da^*} = 0$	$\frac{dT_2\left(\underline{\beta},a^*\right)}{da^*} \geq 0$
$\left(\underline{\beta}, a^*, \bar{\beta}\right)$	$\frac{dT_1\left(a^*,\bar{\beta}\right)}{da^*} \geq 0$	$\frac{dT_2\left(\underline{\beta},a^*\right)}{da^*} \geq 0$
$\left(\underline{\beta}, \bar{\beta}, \bar{\beta}\right)$	$\frac{dT_1\left(\bar{\beta},\bar{\beta}\right)}{da^*} = 0$	$\frac{dT_2\left(\underline{\beta},\bar{\beta}\right)}{da^*} \geq 0$

While the triplet (I^-, I^*, I^+) may switch from one to another when a^* changes, $M\left(a^*\right)$ is continuous. Using Equation (20) and the results in this table, we have $M\left(\underline{\beta}+B\right) \geq M\left(a^*\right)$ for $\forall a^* \leq \underline{\beta}+B$.

Claim 2: $M\left(\bar{\beta}\right) \geq M\left(a^*\right)$ for all $a^* \leq \bar{\beta}$.

Observe that $a^* \leq \bar{\beta} \implies I^* = a^*$ as we consider $a^* > \underline{\beta}$ only. Therefore the triplet (I^-, I^*, I^+) can be either $\left(\underline{\beta}, a^*, a^*+D\right)$, $\left(\underline{\beta}, a^*, \bar{\beta}\right)$, (a^*-B, a^*, a^*+D), or $\left(a^*-B, a^*, \bar{\beta}\right)$. For all these possibilities, as discussed above, Term 1 and Term 2 are either constant or increasing in a^*, which are summarized in the following table.

(I^-, I^*, I^+)	Term 1	Term 2
$\left(\underline{\beta}, a^*, a^*+D\right)$	$\frac{dT_1(a^*,a^*+D)}{da^*} = 0$	$\frac{dT_2\left(\underline{\beta},a^*\right)}{da^*} \geq 0$
$\left(\underline{\beta}, a^*, \bar{\beta}\right)$	$\frac{dT_1\left(a^*,\bar{\beta}\right)}{da^*} \geq 0$	$\frac{dT_2\left(\underline{\beta},a^*\right)}{da^*} \geq 0$
(a^*-B, a^*, a^*+D)	$\frac{dT_1(a^*,a^*+D)}{da^*} = 0$	$\frac{dT_2(a^*-B,a^*)}{da^*} = 0$
$\left(a^*-B, a^*, \bar{\beta}\right)$	$\frac{dT_1\left(a^*,\bar{\beta}\right)}{da^*} \geq 0$	$\frac{dT_2(a^*-B,a^*)}{da^*} = 0$

Using Equation (20) and the results in this table as well as the continuity of $M\left(a^*\right)$, we have $M\left(\bar{\beta}\right) \geq M\left(a^*\right)$ for $\forall a^* \leq \bar{\beta}$.

Claim 3: $a^* = \max\left\{\bar{\beta}, \underline{\beta}+B\right\}$ maximizes $M\left(a^*\right)$.

Given Claim 1 and 2, we restrict our attention to $a^* \geq \max\left\{\bar{\beta}, \underline{\beta}+B\right\}$. $a^* \geq \bar{\beta}$ implies $I^* = \bar{\beta}$ and $I^+ = \bar{\beta}$, while $\underline{\beta}+B \leq a^* < \bar{\beta}+B$ implies $I^- = a^*-B$. Using the results in (23) and (24) for Equation (20), we have $M\left(\max\left\{\bar{\beta}, \underline{\beta}+B\right\}\right) \geq M\left(a^*\right)$ for all $a^* \geq \max\left\{\bar{\beta}, \underline{\beta}+B\right\}$, and therefore, Claim 3 and thus the proposition are proved.
■

Proof. (Proof of Proposition 3) Following Lemma 3, we consider only $a^* \in \left(\underline{\beta}, \bar{\beta}+B\right)$. In this domain, the following exhaustive list shows all the possible pairs of (I^-, I^*):

$$\begin{aligned}
&\left\{\text{All possible } \left(I^-, I^*\right)\right\} \\
&= \left\{\left(\underline{\beta}, a^*\right), \left(\underline{\beta}, \bar{\beta}\right), \left(a^*-B, a^*\right), \left(a^*-B, \bar{\beta}\right)\right\}.
\end{aligned}$$

Given (10), we are looking for a^* that maximizes $g(a^*)$ where

$$\begin{aligned} \frac{1+\theta}{\theta} g(a^*) &= \int_{I^-}^{I^*} (a^* - \beta_i) f(\beta_i) \, d\beta_i \\ &= (a^* - \beta_i) F(\beta_i) \big|_{I^-}^{I^*} + \int_{I^-}^{I^*} F(\beta_i) \, d\beta_i \\ &= (a^* - I^*) F(I^*) - \left(a^* - I^-\right) F\left(I^-\right) + \int_{I^-}^{I^*} F(\beta_i) \, d\beta_i \\ &\equiv T\left(I^-, I^*\right). \end{aligned} \tag{25}$$

Our problem is thus equivalent to maximize $T(I^-, I^*)$. We have slightly abused the notation here because $T(I^-, I^*)$ is a function of a^*, but this representation highlights the dependence of the function form on the pair (I^-, I^*). We shall use $T(I^-, I^*)$ and $T(a^*)$ interchangeably.

Claim 1: $\frac{dT}{da^*} \geq 0$ for all $a^* \leq \underline{\beta} + B$.

To show this, observe that $a^* \leq \underline{\beta} + B \implies I^- = \underline{\beta}$. Then the pair (I^-, I^*) will be $\left(\underline{\beta}, a^*\right)$ if $a^* \leq \overline{\beta}$, or $\left(\underline{\beta}, \overline{\beta}\right)$ otherwise. For both possibilities, substitute (I^-, I^*) into T and utilize $F\left(\underline{\beta}\right) = 0$ and $F\left(\bar{\beta}\right) = 1$ to obtain

$$T\left(\underline{\beta}, a^*\right) = \int_{\underline{\beta}}^{a^*} F(\beta_i) \, d\beta_i, \tag{26}$$

and

$$T\left(\underline{\beta}, \overline{\beta}\right) = \left(a^* - \bar{\beta}\right) + \int_{\underline{\beta}}^{\bar{\beta}} F(\beta_i) \, d\beta, \tag{27}$$

where both expressions are increasing in a^*. Moreover, T is continuous at the neighborhood of $a^* = \overline{\beta}$. This implies that $\frac{dT}{da^*} \geq 0$ for all $a^* \leq \underline{\beta} + B$.

Claim 1 allows us to focus on $a^* \in \left[\underline{\beta} + B, \bar{\beta} + B\right)$, where the pair (I^-, I^*) is either $(a^* - B, a^*)$ or $\left(a^* - B, \overline{\beta}\right)$. Substitute them into T to obtain:

$$T(a^* - B, a^*) = -BF(a^* - B) + \int_{a^*-B}^{a^*} F(\beta_i) \, d\beta_i;$$

$$\begin{aligned} T\left(a^* - B, \bar{\beta}\right) &= a^* - \bar{\beta} - BF(a^* - B) + \int_{a^*-B}^{\bar{\beta}} F(\beta_i) \, d\beta_i \\ &= \int_{\bar{\beta}}^{a^*} F(\beta_i) \, d\beta_i - BF(a^* - B) + \int_{a^*-B}^{\bar{\beta}} F(\beta_i) \, d\beta_i \\ &= -BF(a^* - B) + \int_{a^*-B}^{a^*} F(\beta_i) \, d\beta_i, \end{aligned}$$

where the second equality for $T\left(a^* - B, \bar{\beta}\right)$ makes use of $F(\beta_i) = 1$ for $\beta_i \geq \bar{\beta}$. Notice that these two expressions are the same. Together with Claim 1, our optimization problem now boils down to

$$\max_{a^* \in \left[\underline{\beta}+B, \bar{\beta}+B\right)} -BF(a^* - B) + \int_{a^*-B}^{a^*} F(\beta_i) \, d\beta_i. \tag{28}$$

Its first and second order derivatives are:

$$\frac{dT}{da^*} = -Bf\left(a^* - B\right) + F\left(a^*\right) - F\left(a^* - B\right), \tag{29}$$

and

$$\frac{d^2T}{da^{*2}} = -Bf'\left(a^* - B\right) + \left[f\left(a^*\right) - f\left(a^* - B\right)\right]. \tag{30}$$

For the subsequent analysis, whenever $\frac{dT}{da^*}$ and $\frac{d^2T}{da^*}$ are mentioned, we exclusively refer to the two expressions above. The next claim allows us to restrict our attention to $a^* \in \left[\max\left\{\min\hat{\beta}, \underline{\beta} + B\right\}, \max\hat{\beta} + B\right]$.

Claim 2: $\frac{dT}{da^*} \geq 0$ for all $a^* \leq \min\hat{\beta}$, and $\frac{dT}{da^*} \leq 0$ for all $a^* \geq \max\hat{\beta} + B$.

If $\underline{\beta} + B \geq \min\hat{\beta}$, then the first part of Claim 2 is proved given Claim 1. Now suppose $\underline{\beta} + B < \min\hat{\beta}$. We focus on $a^* \in \left[\underline{\beta} + B, \min\hat{\beta}\right]$. Recall that $a^* \leq \min\hat{\beta}$ implies $f'\left(\beta_i\right) \geq 0$ for $\beta_i \leq a^*$, which leads to $F\left(a^*\right) - F\left(a^* - B\right) = \int_{a^*-B}^{a^*} f\left(\beta_i\right) d\beta_i \geq Bf\left(a^* - B\right)$. So by (29), $\frac{dT}{da^*} \geq 0$. Combining this with Claim 1, we have $\frac{dT}{da^*} \geq 0$ for all $a^* \leq \min\hat{\beta}$.

Next consider $a^* \geq \max\hat{\beta} + B$, which implies $a^* - B \geq \max\hat{\beta}$ so that $f'\left(\beta_i\right) \leq 0$ for all $\beta_i \geq a^* - B$. This means that $F\left(a^*\right) - F\left(a^* - B\right) = \int_{a^*-B}^{a^*} f\left(\beta_i\right) d\beta_i \leq Bf\left(a^* - B\right)$, implying $\frac{dT}{da^*} \leq 0$ by (29). The second part of Claim 2 is thus proved.

Next we prove a useful property of $\frac{dT}{da^*}$.

Claim 3: $\frac{dT}{da^*}$ is a single-crossing function for $a^* \in \left[\max\left\{\min\hat{\beta}, \underline{\beta} + B\right\}, \max\hat{\beta} + B\right]$.

For $a^* \in \left[\max\left\{\min\hat{\beta}, \underline{\beta} + B\right\}, \max\hat{\beta} + B\right]$, $\frac{dT}{da^*}$ is single-crossing, if $\frac{dT}{da^*}|_{a^*=a_1} \leq 0$ implies $\frac{dT}{da^*}|_{a^*=a_2} \leq 0$ wherever $a_1 < a_2$. By (29), $\frac{dT}{da^*}|_{a^*=a_1} \leq 0$ implies that $f\left(a_1\right) \leq f(a_1 - B)$ given Assumption 1. The fact that we are restricting to domain $a^* \in \left[\max\left\{\min\hat{\beta}, \underline{\beta} + B\right\}, \max\hat{\beta} + B\right]$ implies $f\left(a^*\right)$ is (weakly) decreasing while $f(a^* - B)$ is (weakly) increasing, and therefore $f\left(a^*\right) \leq f(a^* - B)$ will hold for all $a^* > a_1$. Then this implies that for all $a^* > a_1$, the second term in the right side of (30) is (weakly) negative, while the first term is also (weakly) negative because $a^* - B \leq \max\hat{\beta}$. So we have $\frac{d^2T}{da^{*2}} \leq 0$ for $a^* > a_1$. This implies $\frac{dT}{da^*}|_{a^*=a_2} \leq 0$, given $a_1 < a_2$ and $\frac{dT}{da^*}|_{a^*=a_1} \leq 0$, which proves Claim 3.

We are now ready to prove the proposition. If $Bf\left(\underline{\beta}\right) \geq F\left(\underline{\beta} + B\right)$ as stated in (11), then $\frac{dT}{da^*}|_{a^*=\underline{\beta}+B} \leq 0$ by (29). By the single-crossing property the derivative will remain negative for $a^* > \underline{\beta} + B$, so the corner solution $a^* = \underline{\beta} + B$ is optimal.

If $Bf\left(\underline{\beta}\right) < F\left(\underline{\beta} + B\right)$, then $\frac{dT}{da^*}|_{a^*=\underline{\beta}+B} > 0$ by (29). Meanwhile, $\frac{dT}{da^*}|_{a^*=\min\hat{\beta}} \geq 0$ by Claim 2. Therefore, we have $\frac{dT}{da^*}|_{a^*=\max\left\{\min\hat{\beta}, \underline{\beta}+B\right\}} \geq 0$. Also given $\frac{dT}{da^*}|_{a^*=\max\hat{\beta}+B} \leq 0$ from Claim 2, by continuity and the single-crossing property, there must exist some cutoff-point in $\left[\max\left\{\min\hat{\beta}, \underline{\beta} + B\right\}, \max\hat{\beta} + B\right]$, before which $\frac{dT}{da^*} \geq 0$ and $\frac{dT}{da^*} \leq 0$ otherwise. By the definition of quasi-concavity, this implies that T is quasi-concave in $\left[\max\left\{\min\hat{\beta}, \underline{\beta} + B\right\}, \max\hat{\beta} + B\right]$ and thus the solution to the maximization problem is interior. Therefore, when $Bf\left(\underline{\beta}\right) \leq F\left(\underline{\beta} + B\right)$, the optimal a^* is given by the first order condition $\frac{dT}{da^*} = 0$, where $\frac{dT}{da^*}$ is given in (29). ■

Proof. (Proof of Remark 5) Given $B \equiv \left(\frac{2V(1+\theta)}{\theta}\right)^{1/2}$, which is increasing in V, we want to show that the contribution maximizing $\overline{a}$ is increasing in B. According to Proposition 3, under the general distribution of β_i with $\gamma = 0$, the contribution-maximizing ideal is either $\underline{\beta} + B$, which is increasing in B, or determined by (12).

When $\overline{a}$ is determined by (12), by the theorem of monotone comparative statics (Milgrom and Shannon, 1994), to show that $\overline{a}$ increases in B, it suffices to show that $g(a^*)$ has increasing differences in $(a^*; B)$, i.e. for all $B' \geq B$, $g(a^*; B') - g(a^*; B)$ is non-decreasing in a^*, for $a^* \in \left[\max\left\{\underline{\beta} + B, \min\hat{\beta}\right\}, \max\hat{\beta} + B\right]$. By the analysis between (25) and (28), we have

$$\frac{1+\theta}{\theta}\left[g(a^*; B') - g(a^*; B)\right] = BF(a^* - B) - B'F(a^* - B') + \int_{a^*-B'}^{a^*-B} F(\beta_i)\, d\beta_i.$$

Differentiating w.r.t a^* yields

$$\begin{aligned}
&\frac{d\frac{1+\theta}{\theta}\left[g(a^*; B') - g(a^*; B)\right]}{da^*} = Bf(a^* - B) - B'f(a^* - B') + \int_{a^*-B'}^{a^*-B} f(\beta_i)\, d\beta_i \\
&\geq -(B' - B) f(a^* - B') + \int_{a^*-B'}^{a^*-B} f(\beta_i)\, d\beta_i \\
&\geq 0,
\end{aligned}$$

where the inequalities come from $a^* - B' \leq a^* - B \leq \max\hat{\beta}$, given

$$a^* \in \left[\max\left\{\underline{\beta} + B, \min\hat{\beta}\right\}, \max\hat{\beta} + B\right].$$

Therefore $g(a^*)$ has increasing differences in $(a^*; B)$, so $\bar{a}$ is increasing in B.

Next we will show that for both $\overline{a} = \underline{\beta} + B$ or $\overline{a}$ as determined by (12), we have $\frac{dg(\overline{a})}{dB} \geq 0$. In the proof of Proposition 3 we showed that our maximization problem can be simplified to (28). If the contribution maximizing $\overline{a} = \underline{\beta} + B$, then

$$g\left(\underline{\beta} + B\right) = \frac{\theta}{1+\theta}\left[-BF\left(\underline{\beta}\right) + \int_{\underline{\beta}}^{\underline{\beta}+B} F(\beta_i)\, d\beta_i\right]$$

and therefore

$$\frac{dg\left(\underline{\beta} + B\right)}{dB} = \frac{\theta}{1+\theta}\left[F\left(\underline{\beta} + B\right) - F\left(\underline{\beta}\right)\right] \geq 0.$$

If the contribution maximizing $\overline{a}$ is determined by (12), then

$$g(\overline{a}) = \frac{\theta}{1+\theta}\left[-BF(\overline{a} - B) + \int_{\overline{a}-B}^{\overline{a}} F(\beta_i)\, d\beta_i\right].$$

By the envelope theorem,

$$\frac{dg(\bar{a})}{dB} = \frac{\theta}{1+\theta} Bf(\bar{a} - B) \geq 0.$$

Since B is increasing in V, then under the contribution maximizing ideal, $g(\overline{a})$ is weakly increasing in V. ■

Proof. (Proof of Corollary 2) (i) Given the symmetry and Assumption 1, $f\left(\underline{\beta}\right) = \min_{\beta_i} f(\beta_i)$. When $\bar{\beta} - \underline{\beta} > B$, we have $Bf\left(\underline{\beta}\right) < \left(\bar{\beta} - \underline{\beta}\right) f\left(\underline{\beta}\right) \leq \int_{\underline{\beta}}^{\underline{\beta}+B} f(\beta_i)\, d\beta_i = F\left(\underline{\beta} + B\right)$, therefore condition (11) fails and Proposition 3 shows that the ideal determined by (12) maximizes $g(a^*)$. When $\bar{\beta} - \underline{\beta} \leq B$, then it follows that $F\left(\underline{\beta} + B\right) = 1$. Moreover, the density at the peak satisfies $f\left(\hat{\beta}\right) \geq \frac{1}{\bar{\beta}-\underline{\beta}} \geq \frac{1}{B}$. Therefore if $f\left(\underline{\beta}\right)$ is sufficiently close to $f\left(\hat{\beta}\right)$, we will have $Bf\left(\underline{\beta}\right) \geq 1 = F\left(\underline{\beta} + B\right)$, where condition (11) is satisfied and Proposition 1 shows that $a^* = \underline{\beta} + B$ maximizes $g(a^*)$.

(ii) For upward-sloping density functions, $f' \geq 0$ for all β_i implies that Inequality (11) is not satisfied. By Proposition 3, the contribution-maximizing a^* is the ideal determined by (12).

(iii) $f' \leq 0$ for all β_i implies $Bf\left(\underline{\beta}\right) \geq F\left(\underline{\beta} + B\right)$. Proposition 3 implies that $a^* = \underline{\beta} + B$ is contribution maximizing.

(iv) Under the uniform distribution, when $\bar{\beta} - \underline{\beta} < B$, $Bf\left(\underline{\beta}\right) = \frac{B}{\bar{\beta}-\underline{\beta}} > 1 \geq F\left(\underline{\beta} + B\right)$, by Proposition 3, $a^* = \underline{\beta} + B$ maximizes $g(a^*)$. When $\bar{\beta} - \underline{\beta} \geq B$, $Bf\left(\underline{\beta}\right) = \frac{B}{\bar{\beta}-\underline{\beta}} = F\left(\underline{\beta} + B\right)$, by Proposition 3, $a^* = \underline{\beta} + B$ maximizes $g(a^*)$. Meanwhile, given $\gamma = 0$ and for any $a^* \in \left[\underline{\beta} + B, \bar{\beta}\right]$, Equation (17) degenerates to

$$\begin{aligned} g(a^*) &= \frac{1}{\bar{\beta} - \underline{\beta}} \frac{\theta}{1+\theta} \left(I^* - I^-\right) \left(a^* - \frac{1}{2}\left(I^* + I^-\right)\right) \\ &= \frac{1}{\bar{\beta} - \underline{\beta}} \frac{\theta}{1+\theta} \left[a^* - (a^* - B)\right] \left(a^* - \frac{1}{2}(a^* + a^* - B)\right), \end{aligned}$$

which is a constant. Therefore, any $a^* \in \left[\underline{\beta} + B, \bar{\beta}\right]$ maximizes $g(a^*)$. ■

Proof. (Proof of Proposition 4) (i) The Proof of Proposition 2 shows that Equation (8) can be written as Equation (19). Substituting $\gamma = 0$ into Equation (19) gives:

$$\begin{aligned} M(a^*) = &\int_{\underline{\beta}}^{\bar{\beta}} \left(\frac{1}{2}\beta_i^2 + n\beta_i\right) f(\beta_i)\, d\beta_i + \frac{n\theta}{1+\theta} \int_{I^-}^{I^*} (a^* - \beta_i) f(\beta_i)\, d\beta_i \\ &- \frac{1}{2} \left(\int_{I^-}^{I^*} \left(\frac{\theta(a^* - \beta_i)}{1+\theta}\right)^2 f(\beta_i)\, d\beta_i\right). \end{aligned}$$

Define

$$H\left(I^-, I^*\right) \equiv \int_{I^-}^{I^*} \left[-\frac{1}{2}\frac{\theta(a^* - \beta_i)^2}{1+\theta} + n(a^* - \beta_i)\right] dF(\beta_i).$$

Again we have slightly abused the notation here, since $H(I^-, I^*)$ is a function of a^*. H is continuous in a^* given the continuity of $M(a^*)$. The ideal a^* that maximizes $H(I^-, I^*)$ will maximize $M(a^*)$. Any $a^* \notin \left(\underline{\beta}, \bar{\beta} + B\right)$ will not attract any under-contributing insiders and will be ineffective with $\gamma = 0$. We thus focus on

$a^* \in \left(\underline{\beta}, \bar{\beta} + B\right)$. We have

$$\frac{dH\left(I^-, I^*\right)}{da^*} = -\frac{1}{2}\left(\frac{\theta}{1+\theta}\right)\frac{d}{da^*}\int_{I^-}^{I^*}\left(a^* - \beta_i\right)^2 dF\left(\beta_i\right) + n\frac{d}{da^*}\int_{I^-}^{I^*}\left(a^* - \beta_i\right) dF\left(\beta_i\right)$$
$$= -\frac{1}{2}\left(\frac{\theta}{1+\theta}\right)\frac{d}{da^*}\int_{I^-}^{I^*}\left(a^* - \beta_i\right)^2 dF\left(\beta_i\right) + n\frac{dT\left(a^*\right)}{da^*}, \tag{31}$$

where $T\left(a^*\right)$ is defined in (25). Referring to Equations (26), (27) and (29) in the proof of Proposition 3, for each possible pair $\left(I^-, I^*\right)$, $\frac{dT(a^*)}{da^*}$ is listed in the table below.

$\left(I^-, I^*\right)$	$\frac{dT(a^*)}{da^*}$
$\left(\underline{\beta}, a^*\right)$	$F\left(a^*\right)$
$\left(\underline{\beta}, \overline{\beta}\right)$	1
$\left(a^* - B, a^*\right)$	$-Bf\left(a^* - B\right) + F\left(a^*\right) - F\left(a^* - B\right)$
$\left(a^* - B, \overline{\beta}\right)$	$-Bf\left(a^* - B\right) + F\left(a^*\right) - F\left(a^* - B\right)$ with $a^* \geq \overline{\beta}$

(32)

Meanwhile,

$$\int_{I^-}^{I^*}\left(a^* - \beta_i\right)^2 f\left(\beta_i\right) d\beta_i$$
$$= \left(a^* - I^*\right)^2 F\left(I^*\right) - \left(a^* - I^-\right)^2 F\left(I^-\right) + \int_{I^-}^{I^*} 2\left(a^* - \beta_i\right) F\left(\beta_i\right) d\beta_i$$
$$= \left(a^* - I^*\right)^2 F\left(I^*\right) - \left(a^* - I^-\right)^2 F\left(I^-\right) + \left[2\left(a^* - \beta_i\right) G\left(\beta_i\right)\right]_{I^-}^{I^*} + \int_{I^-}^{I^*} 2G\left(\beta_i\right) d\beta_i, \tag{33}$$

where we let G denote the indefinite integral of F.

Claim 1: For any $a^* \leq \underline{\beta} + B$, (8) is maximized at $a^* = \underline{\beta} + B$.

To show this, observe that $a^* \leq \underline{\beta} + B \implies I^- = \underline{\beta}$. Then the pair $\left(I^-, I^*\right)$ can either be $\left(\underline{\beta}, a^*\right)$ or $\left(\underline{\beta}, \overline{\beta}\right)$. Consider first $\left(I^-, I^*\right) = \left(\underline{\beta}, a^*\right)$. Substituting $\left(I^-, I^*\right)$ into (33) we have

$$\int_{I^-}^{I^*}\left(a^* - \beta_i\right)^2 f\left(\beta_i\right) d\beta_i = -\left(a^* - \underline{\beta}\right)^2 F\left(\underline{\beta}\right) - 2\left(a^* - \underline{\beta}\right) G\left(\underline{\beta}\right) + \int_{\underline{\beta}}^{a^*} 2G\left(\beta_i\right) d\beta_i.$$

Thus

$$\frac{d}{da^*}\int_{I^-}^{I^*}\left(a^* - \beta_i\right)^2 f\left(\beta_i\right) d\beta_i = -2\left(a^* - \underline{\beta}\right) F\left(\underline{\beta}\right) - 2G\left(\underline{\beta}\right) + 2G\left(a^*\right).$$
$$= \int_{\underline{\beta}}^{a^*} 2F\left(\beta_i\right) d\beta_i.$$

Substituting this and $\frac{dT(a^*)}{da^*} = F(a^*)$ (from (32)) into (31) to obtain

$$
\begin{aligned}
\frac{dH\left(\underline{\beta}, a^*\right)}{da^*} &= -\frac{1}{2}\frac{\theta}{1+\theta}\int_{\underline{\beta}}^{a^*} 2F(\beta_i)\,d\beta_i + nF(a^*) \\
&= \frac{1}{a^* - \underline{\beta}}\left(-\frac{\theta\left(a^* - \underline{\beta}\right)}{1+\theta}\int_{\underline{\beta}}^{a^*} F(\beta_i)\,d\beta_i + nF(a^*)\left(a^* - \underline{\beta}\right)\right) \\
&\geq \frac{1}{a^* - \underline{\beta}}\left(-\frac{\theta}{1+\theta}\left(a^* - \underline{\beta}\right)\int_{\underline{\beta}}^{a^*} F(\beta_i)\,d\beta_i + n\int_{\underline{\beta}}^{a^*} F(\beta_i)\,d\beta_i\right) \\
&\geq 0,
\end{aligned}
$$

where the first inequality is implied by increasing $F(\beta_i)$, and the second inequality comes from $n \geq \frac{\theta}{1+\theta}\left(a^* - \underline{\beta}\right)$, which is due to $n \geq \frac{\theta}{1+\theta}B$ and $a^* \leq \underline{\beta} + B$.

Next, consider $(I^-, I^*) = \left(\underline{\beta}, \bar{\beta}\right)$. From (33) we have

$$
\int_{I^-}^{I^*} (a^* - \beta_i)^2 f(\beta_i)\,d\beta_i = \left(a^* - \bar{\beta}\right)^2 + 2\left(a^* - \overline{\beta}\right)G\left(\overline{\beta}\right) - 2\left(a^* - \underline{\beta}\right)G\left(\underline{\beta}\right) + \int_{\underline{\beta}}^{\overline{\beta}} 2G(\beta_i)\,d\beta_i.
$$

Thus

$$
\begin{aligned}
\frac{d}{da^*}\int_{I^-}^{I^*} (a^* - \beta_i)^2 f(\beta_i)\,d\beta_i &= 2\left(a^* - \bar{\beta}\right) + 2\left(G\left(\overline{\beta}\right) - G\left(\underline{\beta}\right)\right) \\
&= 2\left(a^* - \bar{\beta}\right) + 2\int_{\underline{\beta}}^{\bar{\beta}} F(\beta_i)\,d\beta_i.
\end{aligned}
$$

Substituting this and $\frac{dT(a^*)}{da^*} = 1$ (from (32)) into (31) to obtain

$$
\begin{aligned}
\frac{dH\left(\underline{\beta}, \bar{\beta}\right)}{da^*} &= -\frac{\theta}{1+\theta}\left(a^* - \bar{\beta} + \int_{\underline{\beta}}^{\bar{\beta}} F(\beta_i)\,d\beta_i\right) + n \\
&\geq -\frac{\theta}{1+\theta}B + n \\
&\geq 0,
\end{aligned}
$$

where the first inequality comes from

$$
a^* - \bar{\beta} + \int_{\underline{\beta}}^{\bar{\beta}} F(\beta_i)\,d\beta_i \leq B - \left(\bar{\beta} - \underline{\beta}\right) + \int_{\underline{\beta}}^{\bar{\beta}} F(\beta_i)\,d\beta_i \leq B
$$

(due to $a^* \leq \underline{\beta} + B$ and $F(\beta_i) \leq 1$ for any β_i), and the second inequality makes use of $n \geq \frac{\theta}{1+\theta}B$. Therefore for both possible duplets (I^-, I^*), $\frac{dH\left(I^-, I^*\right)}{da^*} \geq 0$ for any $a^* \leq \underline{\beta} + B$. Note that while duplet (I^-, I^*) may switch when a^* changes, H is continuous in a^*. Consequently, we have $H\left(\underline{\beta} + B\right) \geq H(a^*)$ for all $a^* < \underline{\beta} + B$ and Claim 1 is proved.

Claim 1 allows us to focus on $a^* \in \left[\underline{\beta}+B, \bar{\beta}+B\right)$, where the pair (I^-, I^*) is either (a^*-B, a^*) or $\left(a^*-B, \overline{\beta}\right)$. Substitute $(I^-, I^*) = (a^*-B, a^*)$ into (33) to get:

$$\int_{a^*-B}^{a^*} (a^*-\beta_i)^2 f(\beta_i)\, d\beta_i = -B^2 F(a^*-B) - 2BG(a^*-B) + \int_{a^*-B}^{a^*} 2G(\beta_i)\, d\beta_i, \tag{34}$$

and substitute $(I^-, I^*) = \left(a^*-B, \overline{\beta}\right)$ into (33) to get:

$$\begin{aligned}
&\int_{I^-}^{I^*} (a^*-\beta_i)^2 f(\beta_i)\, d\beta_i \qquad (35)\\
&= \left(a^*-\bar{\beta}\right)^2 - B^2 F(a^*-B) + \left[2(a^*-\beta_i) G(\beta_i)\right]_{a^*-B}^{\bar{\beta}} + \int_{a^*-B}^{\bar{\beta}} 2G(\beta_i)\, d\beta_i\\
&= -B^2 F(a^*-B) - 2BG(a^*-B) + \int_{a^*-B}^{\bar{\beta}} 2G(\beta_i)\, d\beta_i + 2\left(a^*-\bar{\beta}\right) G\left(\bar{\beta}\right) + \left(a^*-\bar{\beta}\right)^2.
\end{aligned}$$

Note that $F(\beta_i) = 1$ for $\forall \beta_i \geq \bar{\beta}$ implies $G(\beta_i)$ is linear in this range, implying

$$2\left(a^*-\bar{\beta}\right) G\left(\bar{\beta}\right) + \left(a^*-\bar{\beta}\right)^2 = \left(a^*-\bar{\beta}\right) \times \left(G\left(\bar{\beta}\right) + G\left(\bar{\beta}\right) + \left(a^*-\bar{\beta}\right)\right) = \int_{\bar{\beta}}^{a^*} 2G(\beta_i)\, d\beta_i. \tag{36}$$

Substituting (36) to (35) yields (34). This means that (34) covers both $(I^-, I^*) = \left(\underline{\beta}, a^*\right)$ and $\left(\underline{\beta}, \overline{\beta}\right)$, and therefore it suffices to focus our attention on (34). Its first-order derivative is:

$$\begin{aligned}
&\frac{d}{da^*} \int_{a^*-B}^{a^*} (a^*-\beta_i)^2 f(\beta_i)\, d\beta_i \qquad (37)\\
&= -B^2 f(a^*-B) - 2BF(a^*-B) + 2G(a^*) - 2G(a^*-B)\\
&= B\left[-Bf(a^*-B) - F(a^*-B)\right] - BF(a^*-B) + 2G(a^*) - 2G(a^*-B)\\
&= B\left(-F(a^*) + \frac{dT(a^*)}{da^*}\right) - BF(a^*-B) + 2G(a^*) - 2G(a^*-B)\\
&= B\frac{dT(a^*)}{da^*} - BF(a^*) - BF(a^*-B) + 2G(a^*) - 2G(a^*-B),
\end{aligned}$$

where the third equality uses $\frac{dT(a^*)}{da^*} = -Bf(a^*-B) + F(a^*) - F(a^*-B)$ for both $(I^-, I^*) = (a^*-B, a^*)$ and $\left(a^*-B, \overline{\beta}\right)$, as shown in the table in (32). Define Φ to be the last terms in the last line of (37):

$$\begin{aligned}
\Phi &\equiv -BF(a^*) - BF(a^*-B) + 2G(a^*) - 2G(a^*-B)\\
&= 2\int_{a^*-B}^{a^*} F(\beta_i)\, d\beta_i - B\left(F(a^*) + F(a^*-B)\right). \qquad (38)
\end{aligned}$$

Φ, which is a function of a^*, may be positive or negative depending on the shape of F in $[a^*-B, a^*]$. Specifically, $\Phi = 0$ if F is linear (f constant), $\Phi \geq 0$ if F is concave (f decreasing), and $\Phi \leq 0$ if F is convex (f increasing). Substituting (38) and (37) to (31) yields

$$\frac{dH}{da^*} = \left(n - \frac{B}{2}\frac{\theta}{1+\theta}\right)\frac{dT(a^*)}{da^*} - \frac{1}{2}\left(\frac{\theta}{1+\theta}\right)\Phi. \tag{39}$$

Note that this expression covers both possibilities of (I^-, I^*) being either $(a^* - B, a^*)$ or $\left(a^* - B, \overline{\beta}\right)$ as mentioned before, and $\frac{dT(a^*)}{da^*}$ for both possibilities has the same form. Given $n \geq \frac{\theta}{1+\theta}B$, then the term in the parentheses in the first term of $\frac{dH}{da^*}$ in (39) is positive. Since $\Phi(a^*)$ can be both positive or negative depending on the distribution, $\frac{dT(a^*)}{da^*} = 0$ does not necessarily imply $\frac{dH}{da^*} = 0$. Nevertheless, we are ready to prove Part (i) of the proposition.

To see this, first consider $a^* \geq \max \hat{\beta} + B$, implying $f(\beta_i)$ is decreasing for $\forall \beta_i \geq a^* - B$, which leads to $\Phi(a^*) \geq 0$ and also by (29)

$$F(a^*) - F(a^* - B) \leq Bf(a^* - B) \implies \frac{dT}{da^*} \leq 0.$$

Consequently we have $\frac{dH}{da^*} \leq 0$ for all $a^* \geq \max \hat{\beta} + B$. Next consider $a^* \leq \min \hat{\beta}$. Suppose $\min \hat{\beta} < \underline{\beta} + B$. Then combining Claim 1 and the result above, there exists $a^* \in \left[\max\left\{\underline{\beta} + B, \min \hat{\beta}\right\}, \max \hat{\beta} + B\right]$ that maximizes H. Suppose $\min \hat{\beta} \geq \underline{\beta} + B$. Consider $a^* \in \left[\underline{\beta} + B, \min \hat{\beta}\right]$. Recall that $a^* \leq \min \hat{\beta}$ implies $f(\beta_i)$ is increasing for $\forall \beta_i \leq a^*$, which leads to $\Phi(a^*) \leq 0$ and also by (29)

$$F(a^*) - F(a^* - B) \geq Bf(a^* - B) \implies \frac{dT}{da^*} \geq 0.$$

Consequently we have $\frac{dH}{da^*} \geq 0$ for all $a^* \in \left[\underline{\beta} + B, \min \hat{\beta}\right]$. This implies that H is maximized at $\min \hat{\beta}$ for all $a^* \leq \min \hat{\beta}$, using Claim 1. Combining Claim 1 with the results here, there exists a welfare-maximizing a^* in interval $\left[\max\left\{\underline{\beta} + B, \min \hat{\beta}\right\}, \max \hat{\beta} + B\right]$.

(ii) If f is decreasing, then $\Phi \geq 0$. Meanwhile we have $\frac{dT(a^*)}{da^*} = -Bf(a^* - B) + F(a^*) - F(a^* - B) \leq 0$ for $\forall a^* \in \left[\underline{\beta} + B, \bar{\beta} + B\right)$, which implies that $\frac{dH}{da^*} \leq 0$ in this interval. Consequently, H is maximized at the corner point $\underline{\beta} + B$.

(iii) If f is increasing, then $\Phi \leq 0$. Corollary 2 shows that the contribution maximizing a^* is determined by (12) where $\frac{dT(\overline{a})}{da^*} = 0$. Thus for all $a^* < \bar{a}$, $\frac{dT(a^*)}{da^*} > 0$, and thus $\frac{dH(\overline{a})}{da^*} > 0$ given $\Phi \leq 0$. Therefore, a^* that maximizes H is (weakly) greater than $\overline{a}$.

(iv) Suppose that $Bf\left(\underline{\beta}\right) < F\left(\underline{\beta} + B\right)$ and by Proposition 3 the contribution-maximizing ideal $\overline{a}$ is determined by $\frac{dT(\overline{a})}{da^*} = 0$. Part (i) of Proposition 4 allows us to focus on $a^* \in \left[\underline{\beta} + B, \bar{\beta} + B\right)$ for welfare-maximization. In this domain, both $\frac{dT(a^*)}{da^*}$ and Φ are continuous and so is $\frac{dH}{da^*}$. We have

$$\frac{dH\left(\overline{\beta} + B\right)}{da^*} = -\left(n - \frac{B}{2}\frac{\theta}{1+\theta}\right) Bf\left(\bar{\beta}\right) \leq 0;$$

$$\begin{aligned}
&\frac{dH\left(\underline{\beta} + B\right)}{da^*} \\
&= \left(n - \frac{B}{2}\frac{\theta}{1+\theta}\right)\left[-Bf\left(\underline{\beta}\right) + F\left(\underline{\beta} + B\right)\right] - \frac{1}{2}\left(\frac{\theta}{1+\theta}\right)\left[\int_{\underline{\beta}}^{\underline{\beta}+B} 2F(\beta_i)\, d\beta_i - BF\left(\underline{\beta} + B\right)\right],
\end{aligned}$$

where the first term of the right side is positive because $Bf\left(\underline{\beta}\right) < F\left(\underline{\beta}+B\right)$ and $n \geq \frac{\theta}{1+\theta}B$. Thus $\frac{dH\left(\underline{\beta}+B\right)}{da^*} \geq 0$ when N is large enough. Therefore, for each N that is large enough, by the intermediate value theorem there exists $\widetilde{a}\left(N\right)$ such that $\frac{dH(\widetilde{a}(N))}{da^*} = 0$. Next, the infinite $\left(n - \frac{B}{2}\frac{\theta}{1+\theta}\right)$ and finite $-\frac{1}{2}\left(\frac{\theta}{1+\theta}\right)\Phi$ imply that $\lim_{N\to\infty}\frac{dT(\widetilde{a})}{da^*} = 0$, as otherwise we will have $\frac{dH(\widetilde{a}(N))}{da^*} \neq 0$. This means $\lim_{N\to\infty}\widetilde{a}\left(N\right) = \overline{a}$ as defined above. Finally, when N is large enough, H is quasi-concave because $\frac{dT(a^*)}{da^*}$ is single-crossing and so does $\frac{dH}{da^*}$. $\widetilde{a}\left(N\right)$ thus maximizes H.

Now suppose that $Bf\left(\underline{\beta}\right) \geq F\left(\underline{\beta}+B\right)$ i.e the contribution-maximizing ideal $\overline{a} = \underline{\beta}+B$, the corner solution. This implies that Inequality (11) holds, which further implies that $\min\widehat{\beta} \leq \underline{\beta}+B$, by our distribution assumption Assumption 1. Part (i) of this proposition thus implies that the welfare maximizing ideal lies in $\left[\underline{\beta}+B, \max\hat{\beta}+B\right]$. The proof of Proposition 3 shows under the contribution-maximizing ideal $\overline{a} = \underline{\beta}+B$, $\frac{dT(a^*)}{da^*} \leq 0$ for $\overline{a} \geq \underline{\beta}+B$. When $N \to \infty$, $n \to \infty$, and therefore we have $\frac{dH}{da^*} \leq 0$ for $\overline{a} \geq \underline{\beta}+B$ since $-\frac{1}{2}\left(\frac{\theta}{1+\theta}\right)\Phi$ is finite. This implies that the welfare maximizing ideal is also $\underline{\beta}+B$. ■

References

[1] Akerlof, G. and R. Kranton, 2000, "Economics and identity," *Quarterly Journal of Economics*, CXV(3), 715-751.

[2] Akerlof, G. and R. Kranton, 2002, "Identity and schooling: Some lessons from the Economics of Education," *Journal of Economic Literature*, 40(4), 1167-1201.

[3] Akerlof, G. and R. Kranton, 2005, "Identity and the economics of organizations," *Journal of Economic Perspectives*, 19(1), 9-32.

[4] Akerlof, G. and R. Kranton, 2008, "Identity, supervision, and work groups," *American Economic Review Papers and Proceedings*, 98(2), 212-217.

[5] Akerlof, G. and R. Kranton, 2010, *Identity Economics: How Our Identities Shape Our Work, Wages and Well-being*, Princeton University Press.

[6] Ali, S. N. and R. Bénabou, 2016, "Image Versus Information: Changing Societal Norms and Optimal Privacy," NBER Working Paper 22203.

[7] Almudi, I. and J. S. Chóliz, 2011, "Sustainable use of renewable resources: an identity approach," *Journal of Bioeconomics*, 13: 97-123.

[8] Andreoni, J., 1990, "Impure altruism and donations to public goods: a theory of warm-glow giving," *Economic Journal*, 100, 464–477.

[9] Barbieri, S. and D. Malueg, 2008, "Private provision of a discrete public good: Continuous-strategy equilibria in the private-information subscription game," *Journal of Public Economic Theory*, 10(4), 529–545.

[10] Barrett, S., 1994, "Self-Enforcing International Environmental Agreements," *Oxford Economic Papers*, 46, 878-894.

[11] Bénabou, R. and J. Tirole, 2011, "Identity, dignity and taboos: Beliefs as assets," *Quarterly Journal of Economics*, 126, 805-855.

[12] Benjamin, D., J. Choi, and G. Fisher, 2010, "Religious identity and economic behavior," NBER Working Paper 15925.

[13] Bergstrom, T., 2006, "Benefit-Cost in a Benevolent Society, *American Economic Review*, 96(1), 339-351.

[14] Bernheim, B.D. and A. Rangel, 2005, "Behavioral Public Economics: Welfare and Policy Analysis with Non-Standard Decision-Makers," NBER Working Paper 11518.

[15] Bisin, A., E. Patacchini, T. Verdier, and Y. Zenou, 2011, "Formation and persistence of oppositional identities," *European Economic Review*, 55, 1046-1071.

[16] Brekke, K., S. Kverndokk and K. Nyborg, 2003, "An economic model of moral motivation," *Journal of Public Economics*, 87, 1967-1983.

[17] Burlando, R. and J. Hey, 1997, "Do Anglo-Saxons free-ride more?" *Journal of Public Economics*, 64(1), 41-60.

[18] Carvalho, Jean-Paul, 2016, "Identity-based organizations," *American Economic Review*, 106(5), 410-414.

[19] Cason, T., T. Saijo and T. Yamato, 2002, "Voluntary Participation and Spite in Public Good Provision Experiments: An International Comparison," *Experimental Economics*, 5(2), 133-153.

[20] Cason, T., T. Saijo, T. Yamato and K. Yokotani, 2004, "Non-excludable public good experiments," *Games and Economic Behavior*, 49(1), 81-102.

[21] Chen, Y., and S. X. Li, 2009, "Group identity and social preferences," *American Economic Review*, 99(1), 431-457.

[22] Costa, D. and M. Kahn, 2013, "Energy conservation 'nudges' and environmentalist ideology: Evidence from a randomized residential electricity field experiment," *Journal of the European Economic Association*, 11(3), 680-702.

[23] Costa-i-Font, J. and F. Cowell, 2013, "Social identity and redistributive preferences: A survey," CESifo Working Paper, No. 4440.

[24] Croson, R., M. Marks, and J. Snyder, 2003, "Groups work for women: Gender and group identity in the provision of public goods," *The Negotiation Journal*, 24 (4), 411-427.

[25] Fehr, E. and K. Schmidt, 1999, "A theory of fairness, competition and cooperation," *Quarterly Journal of Economics*, 114(3), 817-868.

[26] Goeschl, T. and G. Perino, 2016, "The Climate Policy Hold-Up: Green Technologies, Intellectual Property Rights, and the Abatement Incentives of International Agreements," *Scandinavian Journal of Economics*, forthcoming.

[27] Goldstein, N., R. Cialdini and V. Griskevicius, 2008, "A room with a viewpoint: Using social norms to motivate environmental conservation in hotels," *Journal of Consumer Research*, 18(5), 429-434.

[28] Hammond, P., 1987, "Altruism," In J. Eatwell, M. Milgate, and P. Newman (Eds.), *The New Palgrave: A Dictionary of Economics*, 85–87, London: Macmillan Press Ltd.

[29] Hardin, R., 2005, "From communal norms to networks," *Chinese Public Affairs Quarterly*, 1(2), 93-102.

[30] Hiller, V. and T. Verdier, 2014, "Corporate culture and identity investment in an industry equilibrium," *Journal of Economic Behavior & Organization*, 103, 93-102.

[31] Holländer, H., 1990, "A social exchange approach to voluntary cooperation," *American Economic Review*, 80(5), 1157-1167.

[32] Hsiaw, A., 2013, "Goal-setting and self-control," *Journal of Economic Theory*, 148, 601-626.

[33] Kahn, M., 2007, "Do greens drive hummers or hybrids? Environmental ideology as a determinant of consumer choice," *Journal of Environmental Economics and Management*, 54(2), 129-145.

[34] Kotchen, M. and M. Moore, 2008, "Conservation: From voluntary restraint to a voluntary pice premium," *Environmental & Resource Economics*, 40(2), 195-215.

[35] Kotchen, M., 2009, "Voluntary provision of public goods for bads: A theory of environmental offsets," *Economic Journal*, 119, 883-899.

[36] Knack, S. and P. Keefer, 1997, "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," *Quarterly Journal of Economics*, 112(4), 1251-1288.

[37] Laffont, J.-J. and D. Martimort, 2001, *The Theory of Incentives: The Principal-Agent Model*, Princeton University Press.

[38] Milgrom, P. and C. Shannon, 1994, "Monotone comparative statics," *Econometrica*, 62(1), 157-180.

[39] Monin, B., P. Sawyer, and M. Marquez, 2008, "The rejection of moral rebels: Resenting those who do the right thing," *Journal of Personality and Social Psychology*, 95(1), 76-93.

[40] Rege, M., 2004, "Social norms and private provision of public goods," *Journal of Public Economic Theory*, 6(1), 65-77.

[41] Schultz, P., J. Nolan, R. Cialdini, N. Goldstein, and V. Griskevicius, 2007, "The constructive, destructive, and reconstructive power of social norms," *Psychological Science*, 18, 429 –34.

[42] Shayo, M., 2009, "A model of social identity with an application to political economy: Nation, class, and redistribution," *American Political Science Review*, 103, 147-174.

[43] Solow, J. and N. Kirkwood, 2002, "Group identity and gender in public goods experiments," *Journal of Economic Behavior & Organization*, 48, 403-412.

[44] Tabarrok, A., 1998, "The private provision of public goods via dominant assurance contracts," *Public Choice*, 96, 345-362.

[45] Thaler, R. and C. Sunstein, 2008, *Nudge: Improving Decisions about Health, Wealth and Happiness*, Yale University Press.