LE
7/24/08
9:30 am

Privacy, Publicity, and Choice*

Andrew F. Daughety
Jennifer F. Reinganum

Department of Economics
and Law School
Vanderbilt University
Nashville, TN 37235

andrew.f.daughety@vanderbilt.edu
jennifer.f.reinganum@vanderbilt.edu

May 2008

* We thank the Division of Humanities and Social Sciences, Caltech; the Berkeley Center for Law, Business and the Economy, Boalt Law School, the University of California, Berkeley; and NYU Law School for providing us with a supportive research environment during the writing of this paper. We thank Rebecca Morton, Joel Sobel, Eric Talley, and Abraham Wickelgren for helpful conversations and suggestions.

Privacy, Publicity, and Choice

Andrew F. Daughety
Jennifer F. Reinganum

ABSTRACT

We develop and explore a new model of the economics of privacy. Previous work has focused on "privacy of type," wherein an agent privately knows an immutable characteristic. We consider "privacy of action," wherein privacy means that an agent's choice of action is unobservable to others. To show how a policy of privacy can be socially optimal, we assume that an agent derives utility from an action he takes, from the aggregate of all agents' actions, and from other agents' perceptions of the agent's type (that are based on his action). If his action is observable, then he distorts it (relative to his full-information optimal action) so as to enhance the perceptions that others have of him. This contributes to aggregate welfare through increasing the public good, but the disutility associated with the distortion of agents' actions is also a social cost. If his action is unobservable, then he can take his full-information optimal action and still be "pooled" with other types. When the disutility of distortion is high relative to the marginal utility of the public good, a policy of privacy is optimal. We also consider a policy of waivable privacy, and find that equilibria exist in which some, but not all, types waive privacy. More significantly, if policies of privacy or publicity are costlessly enforceable, then a policy of waivable privacy is never socially preferred. Finally, we consider a number of examples (some of which involve a public bad and/or social disapproval): open-source software development; charitable giving; recycling; consumption of health services; DNA dragnets; student rankings; constraints on information disclosure at trial; electricity and water usage during periods of voluntary rationing; shaming of speeders; and the use of earmarks by Congress.

1. Introduction

In this paper we develop and explore a new model of the economics of privacy. In contrast with much of the previous work on the topic of privacy, the notion of privacy we explore is "privacy of action;" that is, we consider privacy as a limitation on the public observability of an individual agent's choice of a level of a good or service to obtain, or of an action to take. We especially emphasize individual actions that may have public-good (or public-bad) implications and characterize conditions under which it is socially preferable to provide a policy of privacy. When actions are private, agents with different types (for example, different preferences over what they read, or how much they will contribute to a charity, or over the level of health services they want to consume) need not conform their actions to those of others in order to pool their types with those of others. Unlike the usual pooling notion in information-economics models, here unobservability of action means that pooling can occur even though different types choose their individual full-information levels of the action. This notion of privacy leads to conditions wherein it is individually and socially preferred to the alterative, which we refer to as "publicity," where individual choice is distorted by social pressure. Alternatively put, we develop a model of a demand for privacy without assuming a direct preference (or taste) for privacy on the part of the individual agent. We further consider an intermediate policy of waivable privacy, wherein an agent may choose to make his actions observable. Despite the absence of an exogenously-determined cost of disclosure, we obtain an equilibrium that involves limited waiver of privacy. Waivable privacy is never *ex ante* optimal when it is costless to enforce a policy either of privacy or publicity. However, it turns out to be socially preferred to policies of either pure privacy or pure publicity when those pure policies are sufficiently costly to enforce, and many of the privacy policies in our society are of the waivable form (from the 5th Amendment's right to silence to opt-outs from privacy restrictions in insurance policies). We provide an analysis of the *ex ante* and *interim* socially optimal choices of policy and discuss examples of the application of such policies to open-source software development; charitable giving; recycling; consumption of health services; DNA dragnets; student rankings; constraints on information disclosure at trial; electricity and water usage during periods of voluntary rationing; shaming of speeders; and the use of earmarks by Congress.

The traditional notion of privacy in economics has centered around what might be called "privacy of type."[1] There the basic idea is that individual agents have characteristics which they might wish to hide (keep secret) when engaged in market transactions or social interactions. Thus, for example, a worker might wish to hide a characteristic that might affect his productivity when he negotiates with a prospective employer whose profitability would be adversely influenced by the characteristic involved (for example, the worker's proclivity to use alcohol or drugs or his potential for contracting a debilitating disease which might be predictable via information on the worker's DNA). This form of privacy derives from the standard notion of private information about type from information economics. The basic dilemma in such analyses is that either some or all of the types pool (which demands that all take the same action, thereby suffering a loss due to the requirement to conform to a common action) or some form of inefficient information transfer occurs, either because agents on the other side of the transaction engage in some type of costly screening of the privately informed agents, or because conditions obtain wherein distortionary signaling by the privately informed agents ensues.[2] In other words, private information in such a context has costs and (without positing a taste for privacy) seems to have no social benefits, thereby leading to the classical assertion that privacy is welfare-reducing in the economy. For example, Posner focuses on privacy as secrecy about personal characteristics and likens privacy protection to the protection of possibly fraudulent claims.[3] As we will see,

---

[1] For early contributions, see Posner (1978, 1981) and Stigler (1980). For an extensive list of recent work, and links to a number of recent papers, see the web site maintained by Acquisti: http://www.heinz.cmu.edu/~acquisti/economics-privacy.htm; accessed April 4, 2008.

[2] Disclosure might be possible if it can be achieved credibly and costlessly, but then private information as a notion is rather empty of significance. Costly disclosure means that only some types will bear the cost. In the analysis at hand, we assume that agents cannot costlessly and credibly disclose their types, though they can costlessly disclose their actions.

[3] "The basic point I wish to assert is the symmetry between 'selling' oneself and selling a product. If fraud is bad in the latter context ... it is bad in the former context, and for the same reasons: it reduces the amount of information in the market, and hence the efficiency with which the market – whether the market for labor, or spouses, or friends – allocates resources." (Posner, 1981, p. 406).

if we consider privacy with regard to actions, then although such privacy does afford some protection of private information about type, it can be welfare-enhancing.

It is important to note that we do not assume that privacy of type is not an important and interesting topic. Clearly a number of laws (for example, concerning the release of information about DNA or blood tests) focus on such privacy. Rather, our point is that privacy of action is also a focus of policy and that such a form of privacy actually can effectively (and, more efficiently) act to provide privacy of type. Indeed, privacy of action may be necessary to maintain privacy of type, since privacy of type alone (without privacy of action) can frequently lead to separating outcomes wherein type ends up being revealed, usually accompanied by distortionary behavior on the part of one or more agents.

For example, information on individuals which is contained in governmental records is protected via various privacy laws. The Privacy Act of 1974 applies to a variety of federal agencies, and augments or primarily determines agency privacy protection policies. Passport records maintained by the U.S. Department of State, which include information on international travel by passport holders, are subject to Privacy Act restrictions; choice of where to travel is an action choice, not a personal characteristic. IRS records, which are also the subject of Privacy Act restrictions, contain information on sources of income (for example, what stocks were sold or jobs held) and on the disposition of wealth (for example, what charities received contributions). Neither sources of income nor distributions of wealth made are immutable characteristics of an individual – that is, they are not "types" drawn from a type-space; rather they are actions taken (that is, choices made within a set of available options). Thus, part of the purpose of this Act is to provide privacy of action when actions taken by an individual become part of a record kept by a federal agency.

Plan of the Paper

In Section 2 we provide a brief review of some of the recent papers on privacy and also some related work on conformity and on pro-social activity. Section 3 provides the basic analysis of pure policies of privacy of action or publicity (that is, openness) of action; here actions give rise to intrinsic utility and esteem (or future opportunities to trade), but they may also contribute to the provision of a public good (think of contributing to the local symphony). We characterize how a privacy or publicity policy affects *ex ante* social welfare, and we show how there can be a substantial conflict between the policy that maximizes *ex ante* expected social welfare and the policy that is *interim*-preferred by the median type. In Section 4 we add the policy of waivable privacy and see that such a policy can be *ex ante* preferred only if pure privacy or pure publicity are sufficiently costly to enforce. Section 5 provides a number of examples that illustrate our results while Section 6 provides a summary and conclusions. An Appendix provides primary results and proofs, while a Technical Appendix[4] provides supplementary results and proofs; we indicate which appendix contains what results as appropriate.

2. Related Literature

This paper is related to several strands of the economics literature, including the disclosure literature (in which an agent can reveal his type directly) and the signaling literature (in which an agent cannot reveal his type directly, but can reveal it indirectly through costly signaling).[5] Like the signaling literature, we assume that an agent cannot credibly reveal his type directly, but can signal it if his action is publically observable. In a variation on the disclosure literature, we assume that an agent can costlessly disclose his action (rather than his type). In Section 5 we discuss several areas of application of our model; related literature that is specific to these applications is discussed briefly in the context of the examples.

In terms of the agent's payoff, our model is related to that of Benabou and Tirole (2006), who explore the effect of rewards on individuals' pursuit of pro-social activity. In their model, an individual has an intrinsic utility from engaging in an activity; he also consumes the public good thereby created. In addition

---

[4] Available at: http://www.vanderbilt.edu/Econ/faculty/Daughety/PPCTechAppendix.pdf

[5] Early papers on disclosure include Grossman (1981) and Milgrom (1981); an early paper on signaling is Spence (1973). Daughety and Reinganum (2007) provide a discussion and unification of the disclosure and signaling models in the context of private information about product quality.

he receives monetary and reputational rewards. An agent's type is two-dimensional (the degree of altruism and the degree of responsiveness to monetary rewards), but there is a single activity level to be chosen, leading to a "signal extraction" problem: does a higher level of activity mean that the agent is more altruistic or greedier? They demonstrate how the use of monetary rewards can undermine the value of engaging in the activity as a signal of altruism, and discuss the determination of optimal monetary rewards. The strength of reputational incentives depends upon a parameter that reflects the extent to which the agent's action is observable. They show that the agents' aggregate supply of the activity increases in this parameter.[6]

Our model and goals are different from those of Benabou and Tirole (2006) in several ways. The most significant difference is that our objective is to compare behavior and welfare (both *interim* and *ex ante*) under various privacy policies, including privacy, publicity, and waivable privacy. Our agents' utility functions also include intrinsic utility from the action, utility from consuming the public good, monetary rewards or costs, and utility from the esteem of others. Since our interest is in the comparison of alternative privacy policies rather than the conflicting reputational concerns that result in rewards undermining participation, we assume that agents have private information only about their intrinsic utility, and not about their responsiveness to rewards. In this case, the agent's action (if public) is a clear signal of his (single) type. We also consider anti-social activities, and those that have a mixed effect (for instance, an agent's action may contribute to a public good, but result in an adverse reputational effect for that agent).[7]

---

[6] Linardi and McConnell (2008) conduct an experiment based on Benabou and Tirole's model and find that subjects do volunteer more time when their contributions are public than when they are private; however, they find that monetary incentives have little effect on the extent of volunteering.

[7] Becker's (1974) complete information model of social interactions includes intrinsic utility and a status motive for engaging in an activity. Bernheim's (1994) incomplete information model of conformity involves intrinsic utility and esteem, where esteem is based on the agent's perceived (inferred) type. Since the average type is accorded the most esteem, every type distorts his public action toward this agent's ideal action. By contrast, in our model the highest esteem is accorded to an extreme type, so all types distort their public actions in the same direction (either upward or downward, depending on whether the highest or lowest type is accorded the greatest esteem). In addition, we include a term that reflects the utility (or disutility) associated with the aggregate of agents' actions. A proper comparison of privacy policies requires all three potential sources of utility.

There are several previous papers that address the question of privacy, but in the context of market transactions between a buyer with private information and a sequence of sellers. Although the details of the models differ, they obtain similar results. Taylor (2004) considers a market wherein a buyer with private information buys sequentially from two sellers. The buyer's valuations for the two goods are correlated, so the information that the first seller obtains by screening the buyer would allow the second seller to engage in price discrimination. If buyers fully-anticipate the sale of their information, they modify their purchase behavior so as to reduce the extent of information that is revealed, thus undermining the first seller's direct profits as well as its profits from the sale of information. In this case, the first seller has an incentive to commit not to sell the information he obtains by screening the buyer. Acquisti and Varian (2005) analyze a related model wherein the first and second seller are the same firm. They provide conditions under which the firm would find it optimal to commit not to use the buyer's first-period purchase history in its second-period pricing decisions. Calzolari and Pavan (2006) use a principal-agent model wherein a buyer contracts sequentially with two sellers. Since the first contract may sort buyer types (and this information could be conveyed to the second seller), the first seller and the buyer must also contract over the extent of the buyer's privacy. They provide sufficient conditions on the preferences and correlations between the buyer's values for trade with each seller for the optimal first contract to involve privacy (respectively, disclosure). These papers may be viewed as addressing "privacy of action" in the sense that the buyer's first-period action is rendered unobservable to the second seller when the first seller promises privacy; however, buyers are still engaged in a game of asymmetric information with the first seller (and with the second as well), so their decisions are still distorted away from their full-information optimal levels. By contrast, in our model a policy of privacy allows the agent to retain private information about his type while choosing his full-information optimal action.

We also address the issue of waivable privacy of action. In the context of privacy of type, Hermalin and Katz (2006) consider the effect of waivable privacy in a model wherein knowledge of a buyer's type may allow the seller to price discriminate or, alternatively, may facilitate improved matching between buyers and sellers. They find that privacy rights must be mandated rather than waivable in order to have any effect. This latter result is due to their assumption that the agent, by waiving his right to privacy, reveals his type costlessly. This leads to complete "unraveling" because every agent type is induced to disclose his type in order not to be "pooled" with a collection of even worse types. By contrast, we consider a costlessly-waivable privacy of action. By waiving his right to privacy, the agent can reveal his action – but not his type – costlessly. Although waiving privacy of action is costless, this waiver itself results in a distorted choice of action through which type is revealed. Thus, there is an endogenously-determined cost associated with revealing type in our model and complete unraveling need not occur. Although mandated privacy may be more efficient than waivable privacy in our model, waivable privacy still has some effect in the sense that some types may invoke their (waivable) privacy rights in equilibrium. Moreover, if a policy of pure privacy (respectively, pure publicity) is costly to enforce then waivable privacy can be more efficient.

3. Privacy versus Publicity Policies: Individual and Social Preferences

We initially formulate our model to address the possibility of actions that might generate public goods as well as personal esteem due to the perceptions of others; in a later section we will modify the model slightly to consider actions that might generate public bads and/or social disapproval. We structure the model so that we can address three policies: 1) complete privacy (that is, no individual's action is observable by others); 2) complete publicity (that is, each individual's action is observable by all, though each individual's type is not observable directly); and 3) waivable privacy, wherein each agent can choose to make her action observable or unobservable. We focus on the first two policies in this section and then extend the analysis to the waivable-privacy case in the next section.

In what follows we model an agent's utility as being comprised of three parts: 1) an intrinsic utility term reflecting consumption of a composite commodity and the action of interest; 2) an extrinsic utility term equal to the individual's perceived benefit from a public good arising from the aggregate actions of all agents; and 3) an extrinsic utility term reflecting the esteem an individual is accorded by others.

3.1 Model Setup

There are n agents, with agent i's utility over the consumption of a composite good c_i and an action g_i assumed to be quasilinear in the composite consumption good; thus any income not spent on g_i is devoted to c_i. Let θ_i denote agent i's type; we will assume that $\theta_i \in [0, \bar{\theta}]$ for all i where $\bar{\theta}$ is finite. Furthermore, assume that each agent observes his type privately, but it is common knowledge that each agent's type is an i.i.d. draw from a commonly known distribution, H, with support $[0, \bar{\theta}]$ and a strictly positive and continuous density h on that support. We assume that the agent's intrinsic utility of the action (that is, the agent's utility divorced from any public goods and esteem effects as well as any associated costs) is quadratic in the level of the action, g_i, and in the type, θ_i; thus, the intrinsic utility of the action is given by $\gamma g_i - (g_i - \theta_i)^2/2$, with $\gamma > 0$. The marginal intrinsic utility is $\gamma - g_i + \theta_i$, which is diminishing in the level of the action but increasing in type, so that higher types have higher intrinsically-optimal actions. We have chosen this particular form of the intrinsic utility so as to provide a sufficiently simple and manipulable model when we allow for incomplete information; in certain of the examples below in Section 5, we will modify the model slightly and indicate the effects of the modification on the results we obtain in this section.

Each agent's action will be a function only of his own (privately observed) type, since agents make simultaneous choices; that is, a strategy for each agent maps his type into the non-negative real line.[8] Agent i also derives utility from the aggregate of the agents' actions. From agent i's perspective, the aggregate expected activity by the n agents is $G_i \equiv g_i + G_{-i}$, where $G_{-i} \equiv \sum_{j \neq i} E(g_j(\theta_j))$ and $g_j(\theta_j)$ is the (commonly-held) conjecture of the action agent j will take as a function of his type. The associated value to agent i is given by αG_i: the non-negative parameter α allows us to vary the intensity of G_i in agent i's utility.

We further assume that overall utility also depends upon the esteem (social approval) in which agent

[8] Therefore, any inferences about an agent's type are assumed to depend only on that agent's action. Fudenberg and Tirole refer to this as "no signaling what you don't know" (1991, pp. 332-3).

i is held by others; note that esteem might be a personal consumption value, or it might reflect continuation values from future trading opportunities. We assume that esteem is increasing in the agent's perceived type, denoted $\tilde{\theta}_i$, and we specifically model this term as $\beta\tilde{\theta}_i$, with β a positive parameter.[9] If the action is unobservable, then $\tilde{\theta}_i$ equals the mean type of those whose actions would be unobservable (either due to a policy or a choice). If actions are observable, then others will try to infer agent i's type from his action.[10]

Agent i's income is denoted as I_i, and since c_i is the numeraire good, we take its price to be 1; the price of the action, g_i, is denoted as p, so that agent i's budget constraint is $c_i + pg_i = I_i$. For example, if the action involves giving money to a cause, then p may reflect anticipated administrative costs (and would be greater than 1), while if the action involves a physical activity (e.g., volunteering, reading books, recycling) then p would be the cost of the activity (respectively, lost wages, cost of reading materials, direct costs plus the time value of money spent in recycling activity). Finally, given the quasilinearity assumption, we assume that I_i is large enough to assure that a positive amount of the composite commodity c_i is consumed by each agent i; this thereby assures us that the demand for the action g_i is independent of the agent's income level.

Thus, our model of agent i's choice problem, after substituting for the numeraire composite consumption good and employing the functional forms described above, entails the agent choosing the level of the action of interest (g_i) so as to maximize overall utility, denoted as $V_i(g_i, \theta_i, \tilde{\theta}_i, G_{-i})$, under the specified rule of privacy (that is, complete privacy, complete publicity, or waivable privacy), where:

$$V_i(g_i, \theta_i, \tilde{\theta}_i, G_{-i}) \equiv \gamma g_i - (g_i - \theta_i)^2/2 + \alpha(g_i + G_{-i}) + \beta\tilde{\theta}_i + I_i - pg_i. \quad (1)$$

In what follows, we will contrast agent i's equilibrium choice of g_i under the assumption that it is private (that is, unobservable to other agents) versus public (that is, observable to other agents). In principle,

---

[9] In Section 5 we allow α to be negative, so as to consider public bads; we also allow β to be negative so as to consider social disapproval. The details of the analysis for $\beta < 0$ are in the Appendix.

[10] Notice that, in this formulation, every type receives positive esteem. Alternatively, one might formulate this utility term as $\beta(\tilde{\theta}_i - \mu)$, where μ is the mean type. In this case, perceived types below the mean receive negative esteem while perceived types above the mean receive positive esteem. This turns out to have no effect on our results since it simply subtracts a constant, $\beta\mu$, from every type's payoff.

with a finite number of agents, knowledge of the total amount of the public good and knowledge of his own action (g_j) would allow agent j to infer something about the action of agent i, even when i's action is not observable directly. This inference could be made more tenuous by assuming that the public goods production process is noisy (e.g., G_i is subject to multiplication by a random scale factor with mean 1 and/or the addition of a random factor with mean 0). We neglect this inference in what follows so that when actions are specified as being private, agent i takes her action to be unknown by the other agents. Alternatively, one could use a "large-economy" version of the public good in which agent i's utility depends on the average action of a large number of agents; that is, $G_i \equiv \lim_{n\to\infty} \{g_i + G_{-i}\}/n$. In this case, the model and results proceed as before with two minor changes that will be noted in the text when they arise.

3.2 Privacy of Action

First, assume that the n-person society of agents operates under a policy of privacy of action, so that each agent's perceived type is the average of the group, $\mu = \int_A \theta h(\theta)d\theta$, where $A = [0, \bar{\theta}]$; we use a superscript "P" to denote the actions under this regime. Next, since V_i is quasilinear (and thus the optimal action is independent of agent i's income) and the term G_{-i} is a constant from agent i's perspective, the optimal action function is symmetric over all agents: agent i's problem is to pick an action function (that is, a type-dependent strategy), denoted as $g^P(\theta_i)$, which maximizes $V_i(g_i, \theta_i, \mu, G_{-i})$, which again reflects the symmetry of the agents. Since V_i is strictly concave in g_i, the optimal solution for agent i is $g^P(\theta_i) \equiv \gamma + \alpha - p + \theta_i$. In order to simplify the notation, let $g_{min} \equiv \gamma + \alpha - p$,[11] and assume parameter restrictions so that $g_{min} \geq 0$. Thus, agent i's equilibrium level of action under a policy of privacy is:

$$g^P(\theta_i) = g_{min} + \theta_i \text{ for all } \theta_i \in [0, \bar{\theta}]. \tag{2}$$

For example, if γ exceeds the direct marginal cost of the action, p, then independent of the intensity of the public-goods effect (via α), the agent will choose a positive amount of the action, and that amount is

[11] Note that under the "large-economy" version of the public good discussed above, α would not appear in the expression g_{min} (thus, in this case, $g_{min} = \gamma - p$) since agent i's impact on the public good would be truly negligible.

increasing in type. Finally, note that the agent's optimal value function, $V_i(g^P(\theta_i), \theta_i, \mu, G_{-i})$, is readily shown to be increasing in θ_i.

Notice that in this equilibrium, while the types are pooled in the usual sense that no observer can infer which type characterizes any particular individual (that is, there is insufficient information to identify the type of any agent), the lack of observability of the actions means that, unlike the usual pooling equilibrium, the different types do not need to take the same action in order to pool: *privacy of action is sufficient to allow pooling of types without constraining the actual choice of the level of the action itself*. This fundamentally distinguishes privacy of action (wherein an agent's choice of g is not itself observable) from privacy of type (private information about individual characteristics, such as θ).

3.3 Publicity of Action

Consider the same society, but now operating under a policy of publicity (or, for notational convenience, "O" for "openness"). In what follows we will characterize a separating equilibrium;[12] publicity assures that actions are observable while separation assures that each action allows inference of the type that would take that action in equilibrium. Since agents are identical in all observable aspects (except for income, which does not affect the choice of g_i) beliefs about which type will have chosen any particular observed action are the same for all agents. Let B(g) be the belief function that relates an agent's choice of observable action level to his perceived type; thus, if agent i chooses action level g_i (and this level is publically observable), then he is inferred to have type $B(g_i) \in [0, \bar{\theta}]$. Given openness, agent i choosing observable action level g_i, with true type θ_i and perceived type $\tilde{\theta}_i = B(g_i)$ obtains utility $V_i(g_i, \theta_i, B(g_i), G_{-i})$. As observed earlier, simultaneous choice by all agents means that agent i's strategy is a function of his type alone, as there is no strategic interaction among the agents. While G_{-i} contains conjectures about the (expected) actions of

---

[12] We focus on the separating equilibrium because it is typically selected by belief-based refinements such as the Intuitive Criterion or D1 (see Cho and Kreps, 1987). However, an analysis based on pooling equilibria is conjectured to have similar results regarding preferences over policies because in both the separating and pooling equilibria, types do not choose their full-information optimal actions, and thus there is a utility loss associated with public actions.

the other agents, this term is merely a constant in agent i's objective function. Thus, the separating perfect Bayesian equilibrium can be found by analyzing an individual agent's incentive compatibility conditions. In addition to incentive compatibility constraints that ensure separation, a separating equilibrium requires that observing agents infer correctly the acting agent's type from his publically observable action; that is, the beliefs must be consistent with equilibrium play. This is formalized in the following definition.

Definition 1. A separating perfect Bayesian equilibrium consists of the action function, $g^O(\bullet)$, and beliefs, $B^O(\bullet)$, such that for all $\theta_i \in [0, \bar{\theta}]$:

(a) $V_i(g^O(\theta_i), \theta_i, \theta_i, G_{-i}) \geq \max_g V_i(g, \theta_i, B^O(g), G_{-i})$ $\quad i = 1, ..., n;$ (IC)

(b) $B^O(g^O(\theta_i)) = \theta_i$. (Consistency)

The separating equilibrium described in Proposition 1 below is derived in the Appendix.

Proposition 1. The Separating Equilibrium under a Policy of Publicity. For i = 1, ..., n:

(a) There is a unique separating perfect Bayesian equilibrium with agent i's equilibrium choice of action function, $g^O(\theta_i)$, and common agent beliefs, $B^O(g)$, defined as follows.

i) $g^O(\theta_i)$ is the (implicit) solution to $g^O(\theta_i) = g_{min} + \theta_i + \beta(1 - \exp[- (g^O(\theta_i) - g_{min})/\beta])$, with $g^O(\theta_i) \geq g_{min}$ for all $\theta_i \in [0, \bar{\theta}]$.

ii) This equilibrium is supported by the beliefs $B^O(g)$ given by:

$$B^O(g) = g - g_{min} - \beta(1 - \exp[- (g - g_{min})/\beta])$$

for $g \in [g_{min}, g^O(\bar{\theta})]$, with out-of-equilibrium beliefs:

$$B^O(g) = 0 \text{ for } g < g_{min}, \text{ and } B^O(g) = \bar{\theta} \text{ for } g > g^O(\bar{\theta}).$$

(b) $g^O(0) = g^P(0)$ and $g^O(\theta_i) > g^P(\theta_i)$ for all $\theta_i \in (0, \bar{\theta}]$.

(c) G_i in the publicity equilibrium strictly exceeds G_i in the privacy equilibrium: $G_i^O > G_i^P$.

The full derivation of part (a) is in the Appendix; we provide a brief sketch of the proofs for parts (b) and (c) below. The signaling distortion is $g^O(\theta_i) - g^P(\theta_i)$; to understand the effect of this distortion, let us examine the following implicit relationship for g:

$$g = g_{min} + \theta_i + \beta(1 - \exp[- (g - g_{min})/\beta]). \tag{3}$$

Solutions to this implicit expression for g will be the solutions, for given θ_i, for $g^O(\theta_i)$. For all $\beta > 0$ and for all $\alpha \geq 0$, we can illustrate the solution to equation (3) as the intersection of the 45°-line with the curve described by the right-hand-side of equation (3); this is done in Figure 1 below.

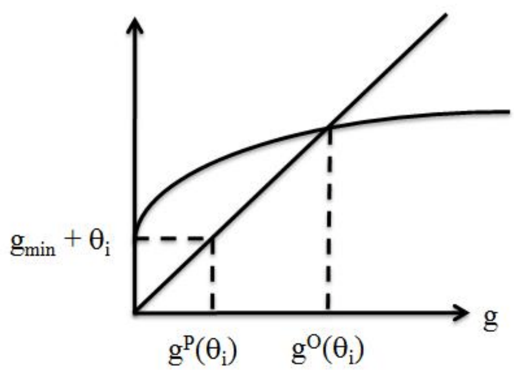

Figure 1: Comparisons of Action Choices for Given θ_i

Note that, for a given agent i and for a given type of that agent $\theta_i > 0$, the curve described by the right-hand-side of equation (3) intersects the vertical axis at $g_{min} + \theta_i$, which is the privacy action choice, $g^P(\theta_i)$, as indicated on the horizontal axis. Since the third term on the right-hand-side of equation (3) is increasing and concave in g, the right-hand-side of equation (3) cuts the 45°-line at a point that implies that $g^O(\theta_i) > g^P(\theta_i)$: under a policy of publicity each agent chooses a higher level of the action (so as to signal type) than he would under a policy of privacy. Notice, this means that G_i under a policy of publicity obtains a value that exceeds its value under privacy: publicity leads to a higher level of the public good.

A second informative illustration of the relationship between $g^O(\theta_i)$ and $g^P(\theta_i)$ is shown in Figure 2, which graphs each of the action strategies with respect to θ_i. The lower, linear function is $g^P(\theta_i)$ using equation (2) above; note that the slope of this line is one and the intercept is g_{min}. The curve for $g^O(\theta_i)$ starts at the same intercept and lies everywhere above the $g^P(\theta_i)$-line; the slope of $g^O(\theta_i)$ is greater than one.[13] Therefore the vertical distance between $g^O(\theta_i)$ and $g^P(\theta_i)$ is increasing as θ_i increases, indicating that (under a policy of publicity) higher types must distort more in order to separate from lower types who would otherwise be tempted to mimic the higher types so as to garner an increase in esteem. That is, the distortion

$$\delta(\theta_i; \beta) \equiv g^O(\theta_i) - g^P(\theta_i) = \beta(1 - \exp[-(g^O(\theta_i) - g_{min})/\beta])$$

is increasing in θ_i.

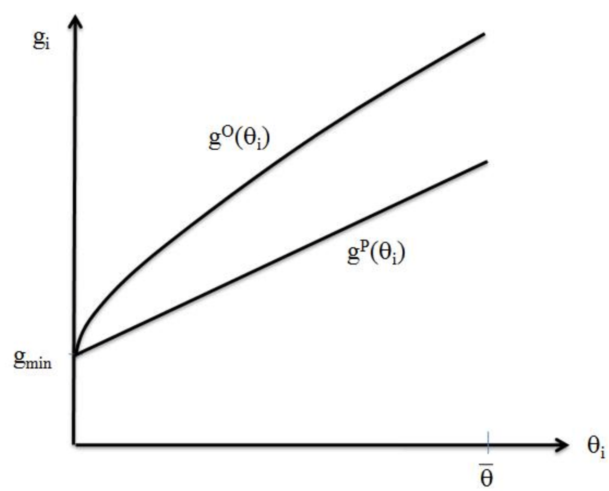

Figure 2: Comparisons of Action Functions

Again, due to the envelope theorem, agent i's optimal value function, $V_i(g^O(\theta_i), \theta_i, B^O(g^O(\theta_i)), G_{-i})$ is readily shown to be increasing in θ_i. With a little more work, one can show that $\partial g^O(\theta_i)/\partial\gamma = \partial g^O(\theta_i)/\partial\alpha = 1$, $\partial g^O(\theta_i)/\partial p = -1$, and $\partial g^O(\theta_i)/\partial\beta > 0$. Finally, for future use, we note that the distortion $\delta(\theta_i; \beta)$ is <u>independent</u> of γ, α, and p, and increasing in β; that is, an increase in the importance to the agent of esteem results in an increased degree of distortion in order to signal type.

3.4 Ex Ante Comparisons of Privacy and Publicity

From an *ex ante* perspective, the decision by a social planner as to which policy, P or O, to pick

---

[13] This and other comparative statics results are proved in the Technical Appendix.

depends upon an interesting trade-off that involves the importance of the public good, α, and the disutility from the distortion $\delta(\theta_i; \beta)$. To be more precise, we define the *ex ante* expected social payoff from a policy of publicity, denoted as W^O, as $E_\Theta[\sum_{i=1}^n V_i(g^O(\theta_i), \theta_i, B^O(g^O(\theta_i)), G^O_{-i})]$, where the expectation is taken over the n-vector of possible types, Θ, for the n agents, and G^O_{-i} is the aggregate level of the public good provided by all agents other than agent i when operating under a policy of publicity. Since these types are i.i.d. draws, this simplifies W^O, making it:

$$W^O = n\gamma E_\theta(g^O(\theta)) - (n/2)E_\theta(g^O(\theta) - \theta)^2 + \alpha n^2 E_\theta(g^O(\theta)) + n\beta\mu + \sum_{i=1}^n I_i - npE_\theta(g^O(\theta))$$

$$= (ng_{min} + \alpha n(n - 1))E_\theta(g^O(\theta)) - (n/2)E_\theta(g_{min} + \delta(\theta; \beta))^2 + n\beta\mu + \sum_{i=1}^n I_i.$$

Similarly, we define the *ex ante* expected social payoff from a policy of privacy, denoted as W^P, as $E_\Theta[\sum_{i=1}^n V_i(g^P(\theta_i), \theta_i, \mu, G^P_{-i})]$, where the expectation is taken over the n-vector of possible types, Θ, for the n agents, and G^P_{-i} is the aggregate level of the public good provided by all agents other than agent i when operating under a policy of privacy. As before, since types are i.i.d. draws, this simplifies W^P, making it:

$$W^P = (ng_{min} + \alpha n(n - 1))E_\theta(g^P(\theta)) - (n/2)E_\theta(g_{min})^2 + n\beta\mu + \sum_{i=1}^n I_i.$$

The question for the social planner concerns the difference between W^O and W^P. Let us denote this difference as Φ^{PO}, which is expressed as a function of α, since that will be a continuing focus of our analysis in this section. That is, let $\Phi^{PO}(\alpha) \equiv W^P - W^O$. It is straightforward to show that:

$$\Phi^{PO}(\alpha) = (n/2)E(\delta^2) - \alpha n(n-1)E(\delta), \quad (4)$$

where, for readability and notational convenience, we let $E(\delta^2)$ denote $E_\theta(\delta^2(\theta; \beta))$, the second moment of $\delta(\theta; \beta)$, and let $E(\delta)$ denote $E_\theta(\delta(\theta; \beta))$, the first moment of $\delta(\theta; \beta)$. Notice that, under either policy, each agent's *ex ante* expected esteem is $\beta\mu$ and therefore this term does not appear in $\Phi^{PO}(\alpha)$.

The sign of the right-hand-side of equation (4) is influenced by two factors: 1) the marginal utility of the public good, α; and 2) the distribution of $\delta(\theta; \beta)$, which is in turn influenced by the distribution of θ, H, and by the esteem parameter, β. First, since $\Phi^{PO}(\alpha)$ is decreasing in α, it is greater than, equal to, or less than zero as the marginal utility of the public good, α, is greater than, equal to, or less than

$E(\delta^2)/[2(n - 1)E(\delta)]$. Let $\alpha^{PO} \equiv E(\delta^2)/[2(n-1)E(\delta)]$ denote the value of α that yields social indifference between P and O.[14] In other words, for values of $\alpha < \alpha^{PO}$, privacy is socially preferred to publicity; this is because the distortionary effects of publicity create greater disutility than the added utility from the increased provision of the public good. On the other hand, if $\alpha > \alpha^{PO}$, publicity is socially preferred to privacy; now the individual valuation of the public good is sufficiently high that each individual would prefer to bear the expected disutility from distortion under publicity, since it will be imposed simultaneously on all other agents and will lead to a sufficiently greater provision of the public good: misery loves company if it contributes to an important public good.[15] In other words, unless the public good is of sufficient importance to the agents, the *ex ante* preferred policy is privacy even though there is a public-goods aspect to the actions by the agents. This reflects the costly effects on the agents of the distortion induced by the esteem term.[16] The foregoing is summarized in the next proposition.

Proposition 2. *Ex Ante* Desirability of Privacy or Publicity.

Publicity is *ex ante* socially preferred to privacy if and only if α is sufficiently large; that is, if and only if $\alpha > \alpha^{PO}$.

---

[14] In this paper we vary α and β independently so as to consider all possible combinations, but one could imagine situations in which β is an increasing function of α (contributing to a more important public good garners greater esteem for a given type). Assuming that $\beta(0) > 0$, a policy of privacy is still preferred for sufficiently low α. If $\beta'(\alpha)$ is sufficiently small, then there is still a unique value of α beyond which a policy of publicity is preferred. If $\beta'(\alpha)$ is sufficiently large, then privacy may always be the optimal policy or there may exist multiple disconnected sets of α-values for which privacy is optimal.

[15] Under the "large-economy" version of the public good, the *ex ante* expected utility of the average agent under a policy of publicity is $W^O = (g_{min} + \alpha)E_\theta(g^O(\theta)) - (1/2)E_\theta(g_{min} + \delta(\theta; \beta))^2 + \beta\mu + I$, where I is the common income of the agents. Similarly, the *ex ante* expected utility of the average agent under a policy of privacy is $W^P = (g_{min} + \alpha)E_\theta(g^P(\theta)) - (1/2)E_\theta(g_{min})^2 + \beta\mu + I$. Finally, the value of α that yields social indifference between P and O in the "large-economy" version is $\alpha^{PO} \equiv E(\delta^2)/[2E(\delta)]$. In all of these expressions, $g_{min} = \gamma - p$.

[16] As discussed in Section 5, this also happens if both α and β are negative (that is, a public bad that involves social disapproval); if the magnitude of α is not very large, privacy will still be preferred.

Second, now consider the distribution of possible δ-values, denoted as H^{δ}, which is induced by the distribution of θ-values, H, and the equilibrium action function, $g^{O}(\theta)$. A simple re-arrangement of equation (4) reveals that a mean-preserving spread of H^{δ} results in an increased social preference for privacy. To see this, observe that equation (4) can be written as:

$$\Phi^{PO}(\alpha) = (n/2)[E(\delta^2) - (E(\delta))^2] + nE(\delta)[E(\delta)/2 - \alpha(n - 1)].$$

A mean preserving-spread in the distribution of $\delta(\theta; \beta)$ results in an increase in the first term in brackets (which is the variance of $\delta(\theta; \beta)$) and no change in the second term in brackets (for given α). Thus, after a mean-preserving spread is applied to H^{δ}, α^{PO} must increase to make $\Phi^{PO}(\alpha^{PO})$ again equal to zero. Our formal result is as follows.

Proposition 3. The Effect of a Mean-Preserving Spread of H^{δ}.

A mean-preserving spread of H^{δ} increases α^{PO}, increasing the set of α-values for which privacy is *ex ante* socially preferred.

Unfortunately, the complexity of the form of $\delta(\theta; \beta)$ has precluded any general characterization of what changes in H create a mean-preserving spread of H^{δ}; however, the discussion below on the effects of β on $\delta(\theta; \beta)$ is at least suggestive of the relationship between the two distributions.

In the Technical Appendix we show that if $\beta' > \beta$, then $H^{\delta}(\delta; \beta')$ first-order stochastic dominates $H^{\delta}(\delta; \beta)$, so that increases in β (the marginal utility of esteem) increase the expectation of any increasing function of δ; two such expectations are $E(\delta^2)$ and $E(\delta)$. Therefore, since α^{PO} is proportional to $E(\delta^2)/E(\delta)$, the parameter β will potentially affect α^{PO}. While we have not been successful in finding a theoretical characterization of the effect of β on α^{PO}, we have used computational techniques as follows.[17] First, under the assumption that H is the uniform distribution, we have computed the necessary and sufficient condition

[17] The calculations were produced using Mathematica 6; the notebook used is available from the authors upon request.

for $\partial(E(\delta^2)/E(\delta))/\partial\beta$ to be positive (and therefore, that $\partial\alpha^{PO}/\partial\beta > 0$) for a broad range of positive values of β and found the condition to be satisfied.

Second, we have computed $E(\delta^2)/E(\delta)$ for a few alternative (but reasonably representative) values of β and for four different distributions on [0, 1]; see the Appendix for details. The computations suggest that, for a given density, increasing β increases α^{PO}, so that $\Phi^{PO}(\alpha^{PO})$ shifts up, associating more values of α with privacy as the *ex ante* socially preferred policy. Moreover, the computations are consistent with the conjecture that a shift in the distribution H of θ-values to a new distribution H', where H' first-order stochastic dominates H, results in higher values of α^{PO} as well.[18] Alternatively put, for fixed α, a group with a greater expected preponderance of high-θ members is more likely, *ex ante*, to prefer privacy over publicity as a policy, than would a reference group with a lower expected preponderance of high-θ-members. This makes sense since publicity is costly to engage in (due to the incentive it provides to distort action) and common knowledge that realized types are more likely to be high-θ substitutes for the need for individuals to distort so as to distinguish themselves from lower types. We summarize the computational results below.

Remark 1. Parametric Influences on the *Ex Ante* Social Desirability of Privacy.

Computational experience supports the notion that increases in β, or first-order stochastic dominance-generating shifts in H, will increase the range of α-values such that privacy is *ex ante* socially preferred to publicity.

3.5 Interim Comparisons of Privacy and Publicity

We now compare *interim* utility payoffs for each individual (that is, type-specific payoffs) under a policy of privacy versus a policy of publicity. Our primary result will be that, whatever is the *ex ante* policy

---

[18] However, the table in the Appendix also suggests that no such nice generalization to mean-preserving spreads of H will be possible, as revealed by the results for the Uniform and Middle Triangle Densities.

adopted, reasonable conditions exist such that the median type will *interim*-prefer the opposite policy; this will then lead to the next section on a mixture of privacy and publicity, namely waivable privacy.

Interim policy preferences for an agent of type θ_i are summarized via the agent's net value of a policy of privacy compared with a policy of publicity; we denote this net value as $\Gamma^{PO}(\theta_i, \alpha)$, for $\theta_i \in [0, \bar{\theta}]$, where:

$$\Gamma^{PO}(\theta_i, \alpha) \equiv V_i(g^P(\theta_i), \theta_i, \mu, G^P_{-i}) - V_i(g^O(\theta_i), \theta_i, \theta_i, G^O_{-i}). \qquad (5)$$

Thus type θ_i strictly *interim*-prefers a policy of privacy to a policy of publicity if and only if $\Gamma^{PO}(\theta_i, \alpha) > 0$. Substituting the action functions $g^P(\theta)$ and $g^O(\theta)$ into equation (5) and simplifying yields:

$$\Gamma^{PO}(\theta_i, \alpha) = \beta(\mu - \theta_i) + (\delta(\theta_i; \beta))^2/2 - \alpha(n - 1)E(\delta). \qquad (6)$$

The first term on the right-hand-side is positive or negative depending upon whether θ_i is less than or greater than μ; the second term (the numerator of which is the signaling distortion squared) is strictly positive. Finally, as before, the term $\alpha(n - 1)E(\delta)$ is the individual's gain in utility from the public good if all other agents' actions are public rather than private.

Taking expectations of both sides of equation (6), we get that $E(\Gamma^{PO}(\theta, \alpha)) = \Phi^{PO}(\alpha)/n$. This allows us to partition the possible α-values into three intervals which correspond to whether or not *interim* preferences reinforce, or potentially conflict with, *ex ante* social choices between the pure privacy and pure publicity policies. This is summarized in the following proposition (details of the interim preferences and of the parameters in Proposition 4 are provided in the Technical Appendix).

Proposition 4. Marginal Utility of the Public Good and the *Interim* Preference for Privacy.

(a) There exist α-values denoted as $\underline{\alpha}^{PO}$ and $\bar{\alpha}^{PO}$, with $0 < \underline{\alpha}^{PO} < \bar{\alpha}^{PO}$, such that:

(i) all types *interim* prefer P to O for any $\alpha \leq \underline{\alpha}^{PO}$;

(ii) all types *interim* prefer O to P for any $\alpha \geq \bar{\alpha}^{PO}$;

(iii) some types prefer P to O; the rest prefer O to P, when α is in the interval $(\underline{\alpha}^{PO}, \bar{\alpha}^{PO})$.

(b) Furthermore, the α-value such that society is *ex ante* indifferent between P and O, α^{PO}, is always in this latter interval: $\alpha^{PO} \in (\underline{\alpha}^{PO}, \bar{\alpha}^{PO})$. Therefore, for α-values in this interval there is always a conflict between *ex ante* and *interim* preferences over privacy and publicity.

In fact, it is possible for the median type to (*interim*) prefer the policy opposite to whichever policy maximizes the *ex ante* social payoff. To see this contrast in a case wherein O is *ex ante* preferred but P is *interim* preferred by the median type, let $\theta^{PO}(\alpha)$ be the marginal type such that $\Gamma^{PO}(\theta^{PO}(\alpha), \alpha) = 0$, for $\alpha \geq 0$. Note that $\theta^{PO}(\alpha)$ is decreasing in α, and that $\theta^{PO}(0) > \mu$, the mean (and median) type if H is symmetric. Thus, there is an α* such that $\theta^{PO}(\alpha^*) = \mu$. It is straightforward to show that $\alpha^* \in (\alpha^{PO}, \bar{\alpha}^{PO})$, so that for any value of α in the interval (α^{PO}, α^*), the *ex ante* social payoff-maximizing choice of policy is O, but on an *interim* basis, the median type would prefer P to O.

To see how the reverse conflict can occur, assume that $\alpha = 0$. Since $\alpha^{PO} > 0$, this means that society *ex ante* prefers P to O. Since $\theta^{PO}(0) > \mu$, then any density h which has a median to the right of $\theta^{PO}(0)$ implies that the median type prefers O to P. This is because signaling type to gain esteem is sufficiently valuable to the median type (but is irrelevant in the case of the *ex ante* decision) for those types to *interim* prefer O to P. This conflict between the *ex ante* and *interim* settings is summarized below.

Remark 2. Conflicting *Ex Ante* and *Interim* Preferences over Policies.

H symmetric: There are values of α such that while a policy of publicity is *ex ante* socially preferred, the alternative policy of privacy is *interim*-preferred by the median type.

H sufficiently right- skewed: There are values of α such that while a policy of privacy is *ex ante* socially preferred, the alternative policy of publicity is *interim*-preferred by the median type.

4. Waivable Privacy

If policies of pure privacy or pure publicity are costly to enforce, the foregoing discussion demonstrating the possibility of conflict between *ex ante* and *interim* decisions as to which policy to implement suggests that we consider a third possible policy, wherein privacy is waivable. That is, we assume that privacy is not enforced *per se*, and that any type who desires to publicize his action choice may elect to do so and does not incur any direct cost of making such a disclosure.

Without a cost of disclosure, one might expect that all types would disclose, in keeping with the unraveling result developed in the literature on disclosure (see the discussion in Section 2 above). That is, one might readily expect that some form of unraveling would occur in equilibrium, so that all types disclose (choose "O") since otherwise a type that does not disclose will be perceived to be worse than his true type. However, as we show, this need not occur: an equilibrium wherein some types choose privacy and some types choose publicity can exist. This is because choosing to waive privacy means making one's action choice observable, and the optimal public action involves distortion due to the social judgment that affects esteem. This, in turn, means that there is an endogenously-determined cost of disclosing type which can imply the existence of an interior marginal type who is indifferent between keeping his action private or making it public. Thus, under waivable privacy an equilibrium wherein only some types waive privacy (and others do not) can exist; in this sub-section we characterize such an equilibrium and then compare it with the pure policies discussed earlier.

4.1 Extension of the Model and the Waivable Privacy Equilibrium

As before, let the two type-dependent strategies under the two policy regimes of privacy and publicity be denoted as $g^P(\theta)$ and $g^O(\theta)$, respectively. We will analyze the problem from agent i's point of view. Consider a strategy wherein, for any $\hat{\theta} \in [0, \bar{\theta}]$, agent j chooses $(P, g^P(\theta_j))$ if $\theta_j \in [0, \hat{\theta})$ and $(O, g^O(\theta_j))$ if $\theta_j \in [\hat{\theta}, \bar{\theta}]$. The marginal type $\hat{\theta}$ thereby characterizes a cutoff rule for every agent $j \neq i$. Thus, denote the *waivable-privacy action function* for agent j, $g^W(\theta_j; \hat{\theta})$, as $g^P(\theta_j)$ if $\theta_j \in [0, \hat{\theta})$ and as $g^O(\theta_j))$ if $\theta_j \in [\hat{\theta}, \bar{\theta}]$, and let the perceived type for any agent be $\mu(\hat{\theta})$ if the agent chooses P (where $\mu(\hat{\theta})$ is the mean conditional[19] on

[19] That is, the conditional mean, $\mu(\hat{\theta}) \equiv \int_T th(t)dt/H(\hat{\theta})$, with $T = [0, \hat{\theta}]$.

$[0, \hat{\theta}))$ and $B^O(g)$ if the agent chooses O and takes observable action g. It is clear that agent i's payoff under P, $V_i(g_i, \theta_i, \mu(\hat{\theta}), G_{-i}(\hat{\theta}))$, is maximized at $g^P(\theta_i)$ while agent i's payoff under O, $V_i(g_i, \theta_i, B^O(g_i), G_{-i}(\hat{\theta}))$, is maximized at $g^O(\theta_i)$, since $\mu(\hat{\theta})$ and $G_{-i}(\hat{\theta})$ do not affect the optimal solutions. Then the net value for agent i of type θ_i of privacy over publicity, given that all others use the strategy and beliefs specified above is:

$$\Delta(\theta_i; \hat{\theta}) \equiv V_i(g^P(\theta_i), \theta_i, \mu(\hat{\theta}), G_{-i}(\hat{\theta})) - V_i(g^O(\theta_i), \theta_i, B^O(g^O(\theta_i)), G_{-i}(\hat{\theta})). \quad (7)$$

It is straightforward to show that this net value is decreasing in θ_i; that is, $\partial\Delta(\theta_i; \hat{\theta})/\partial\theta_i = g^P(\theta_i) - g^O(\theta_i) < 0$.[20] Thus, lower values of θ_i are associated with higher net values of privacy, which rationalizes the hypothesized form of the strategy wherein low-θ types choose P while high-θ types choose O.

Let θ^W denote an equilibrium value of $\hat{\theta}$: that is, a commonly-conjectured cutoff value such that no individual agent of type θ will defect from using the cutoff rule $g^W(\theta; \theta^W)$. The following proposition, which is proved in the Appendix, provides a full characterization of the possible waiver equilibria.

Proposition 5. Equilibria under Waivable Privacy.

(a) There is always a *full-waiver equilibrium*, wherein all types choose to waive privacy and an agent of type θ chooses action level $g^O(\theta)$.

(b) If $\mu \geq \bar{\theta} - (\delta(\bar{\theta}; \beta))^2/2$, then there is also a *no-waiver equilibrium*, wherein all types choose not to waive privacy and an agent of type θ chooses action level $g^P(\theta)$.

(c(i)) If $\mu < \bar{\theta} - (\delta(\bar{\theta}; \beta))^2/2$, then a no-waiver equilibrium does not exist, but at least one *partial-waiver* (that is, interior) equilibrium does exist, where $\theta^W \in (0, \bar{\theta})$ solves $\Delta(\theta^W; \theta^W) = 0$.

(c(ii)) This equilibrium is supported by the beliefs:

- if the agent chooses P, then $\tilde{\theta} = \mu(\theta^W)$;
- if the agent chooses (O, g), then $\tilde{\theta} = B^O(g)$.

(c(iii)) A sufficient condition for this interior equilibrium to be the unique interior equilibrium is that

[20] Indeed, $\Delta(\theta_i; \hat{\theta})$ is decreasing in θ_i regardless of the form of the other agents' strategies.

the conditional mean $\mu(\theta)$ is everywhere concave in θ.

There are always at least two pure-strategy equilibria, one involving waiver by all types so that the publicity outcome is an equilibrium, and the other(s) involving some degree of privacy, including (possibly) a no-waiver equilibrium with privacy chosen by all types. Moreover, under mild assumptions (e.g., that the conditional mean is weakly concave[21]), this second type of equilibrium is unique; therefore, in what follows we assume that $\Delta(\theta; \theta)$ is concave. Figure 3 illustrates $g^W(\theta; \theta^W)$for the partial waiver equilibrium. The base curves are those shown in Figure 2 earlier, but now there is a jump discontinuity at θ^W. This jump arises because the marginal type is indifferent between taking his full information action but being believed to be the mean type in $[0, \theta^W)$, and distorting his action to obtain the esteem associated with his true type. Thus, equilibrium action choice is always made along the solid portion of the curves.

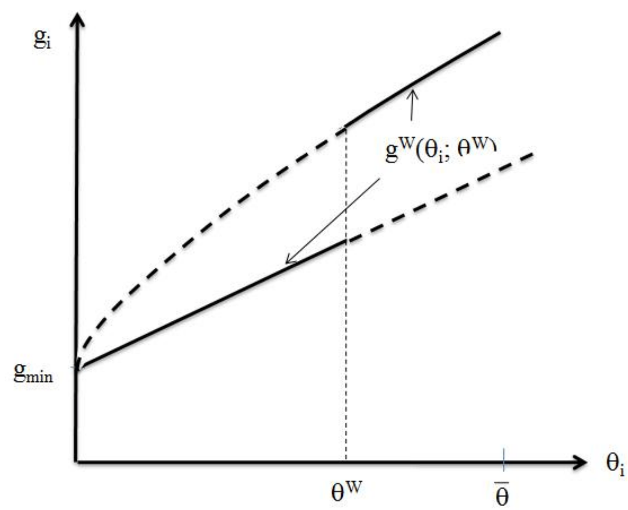

Figure 3: Action Function in Partial Waiver Equilibrium

We have employed the full-publicity beliefs to support the partial-waiver equilibrium. That is, action choices that are disclosed and that are in the interval $[g^O(\theta^W), g^O(\bar{\theta})]$ are taken to be from the

[21] For example, if the distribution H is the uniform distribution, this condition holds. In fact, if H is one of the family of Beta distributions, extensive computational analysis suggests that as long as the density is bounded, $\mu''(\theta) \leq 0$ for all θ. Thus, this sufficient condition appears to hold for a highly relevant class of distributions.

corresponding type in the interval $[\theta^W, \bar{\theta}]$ given by $B^O(g)$. What if, however, one observes an action g which is less than $g^O(\theta^W)$? This is an out-of-equilibrium action choice, which has been made public. This could be because of either of two possible sources of error: 1) perhaps the agent miscalculated θ^W (computed a value below his privately observed type, θ, when the correct value was above θ), and then proceeded to choose the public action choice $g^O(\theta)$; or 2) since there is always a second equilibrium at $\theta^W = 0$, perhaps there was a coordination failure wherein the agent in question played as if he anticipated the full-waiver equilibrium instead of the partial-waiver equilibrium; that is, he played his equilibrium strategy from the other equilibrium. In either case, mapping the g-value via the dotted curve in Figure 3 (i.e., the part of $g^O(\theta)$ not included in $g^W(\theta; \theta^W)$) to the inferred value of θ provides the needed disincentive to keep types from strategically engaging in such behavior.

Notice that although disclosure of action is costless, the jump in the action function reflects the endogenously-determined cost of disclosing type that accompanies choosing to produce a public, rather than private, action and derive any esteem based on this choice (thereby causing the upward distortion).

4.2 Ex Ante Comparisons of Waivable Privacy with Policies of Privacy and of Publicity

Recall from Section 3.4 above the computed versions of the *ex ante* expected social payoff associated with P (W^P) and O (W^O). In a similar fashion, the *ex ante* expected social payoff associated with waivable privacy, denoted as W^W, is:[22]

$$W^W = (ng_{min} + \alpha n(n - 1))E_\theta(g^W(\theta; \theta^W)) - (n/2)E_\theta(g^W(\theta; \theta^W) - \theta)^2 + n\beta\mu + \textstyle\sum_{i=1}^{n} I_i.$$

Thus, the net *ex ante* expected social payoff of P versus W, denoted as $\Phi^{PW}(\alpha)$, is:

$$\Phi^{PW}(\alpha) = W^P - W^W = (n/2)E((\delta^W)^2) - \alpha n(n - 1)E(\delta^W), \tag{10}$$

where, for readability and notational convenience, we let $E((\delta^W)^2)$ denote $\int(\delta(\theta; \beta))^2 h(\theta)d\theta$, the partial second moment of $\delta(\theta; \beta)$, and $E(\delta^W)$ denote $\int\delta(\theta; \beta)h(\theta)d\theta$, the partial first moment of $\delta(\theta; \beta)$, where both integrals are over the interval $[\theta^W, \bar{\theta}]$.

[22] Each agent's *ex ante* expected esteem is $\beta((\mu(\theta^W)H(\theta^W) + (\mu - \mu(\theta^W)H(\theta^W))) = \beta\mu$.

Analogously to before, if $\alpha = \alpha^{PW} \equiv E((\delta^W)^2)/(2(n-1)E(\delta^W))$, then $\Phi^{PW}(\alpha^{PW}) = 0$, so that when comparing the P and W regimes, when $\alpha < \alpha^{PW}$, a policy of privacy is strictly preferred to a policy of waivable privacy, while if $\alpha > \alpha^{PW}$, a policy of waivable privacy is strictly preferred to a policy of privacy.

Finally, the net *ex ante* expected social payoff of W versus O, denoted as $\Phi^{WO}(\alpha)$, is:

$$\Phi^{WO}(\alpha) = W^W - W^O = (n/2)[E((\delta)^2) - E((\delta^W)^2)] - \alpha n(n-1)[E(\delta) - E(\delta^W)]. \tag{11}$$

Analogously to before, if $\alpha = \alpha^{WO} \equiv [E((\delta)^2) - E((\delta^W)^2)]/(2(n-1)[E(\delta) - E(\delta^W)])$, then $\Phi^{WO}(\alpha^{WO}) = 0$, so that when comparing the W and O regimes, if $\alpha < \alpha^{WO}$, a policy of waivable privacy is strictly preferred to a policy of publicity, while if $\alpha > \alpha^{WO}$, a policy of publicity is strictly preferred to a policy of waivable privacy.

Note that the equations providing $\Phi^{PO}(\alpha)$, $\Phi^{PW}(\alpha)$, and $\Phi^{WO}(\alpha)$ (that is, equations (4), (10) and (11), respectively) are all of the same form, except for the terms having to do with the first and second moments of the distortion $\delta(\theta; \beta)$. In the case of $\Phi^{PO}(\alpha)$ and the associated α-value α^{PO}, these are the full moments (that is, the integration in the expectations is over $[0, \bar{\theta}]$) while the other Φ-functions and associated α-values employ partial expectations; for the PW comparison, the partial first and second moments of $\delta(\theta; \beta)$ are taken over $[\theta^W, \bar{\theta}]$, while for the WO comparison, the partial first and second moments of $\delta(\theta; \beta)$ are taken over $[0, \theta^W]$. Therefore, by construction:

$$\Phi^{PO}(\alpha) = \Phi^{PW}(\alpha) + \Phi^{WO}(\alpha).$$

Using these functions and the α-values where each function switches from positive to negative (a reversal of pair-wise preference) leads us to Proposition 6.

Proposition 6. *Ex ante* Social Ordering of Privacy, Publicity, and Waivable Privacy.

(a) $0 < \alpha^{WO} < \alpha^{PO} < \alpha^{PW}$.

(b) Absent enforcement costs, waivable privacy is never an *ex ante* first-best policy. If $\alpha < \alpha^{PO}$, privacy is strictly preferred, while if $\alpha > \alpha^{PO}$, publicity is strictly preferred.

(c) Absent enforcement costs, waivable privacy is an *ex ante* second-best policy when $\alpha < \alpha^{WO}$ or when $\alpha > \alpha^{PW}$.

(d) If the aggregate social cost of enforcing privacy is $k > \Phi^{PW}(\alpha^{WO})$, then there exists a range of α-values such that waivable privacy is the *ex ante* first-best policy.

(e) If the aggregate social cost of enforcing publicity is $k > - \Phi^{PO}(\alpha^{PW})$, then there exists a range of α-values such that waivable privacy is the *ex ante* first-best policy.

A main result of Proposition 6 is that (absent enforcement costs) waivable privacy is always *ex ante* dominated by either a policy of privacy or a policy of publicity. The rankings in Proposition 6(b) and (c) are shown in Figure 4 below, which illustrates the three functions ($\Phi^{WO}(\alpha)$ as "WO", $\Phi^{PO}(\alpha)$ as "PO", and $\Phi^{PW}(\alpha)$ as "PW") for the case wherein H(θ) is the Uniform distribution on [0, 1] and β = 1.

Figure 4 also helps us illustrate the effect of enforcement costs, as discussed in items (d) and (e) of Proposition 6. For example, if there is a cost k of enforcing a policy of privacy, then the Φ-functions PW and PO both shift downward. If that cost exceeds $\Phi^{PW}(\alpha^{WO})$, then the crossing point of the PO and PW Φ-

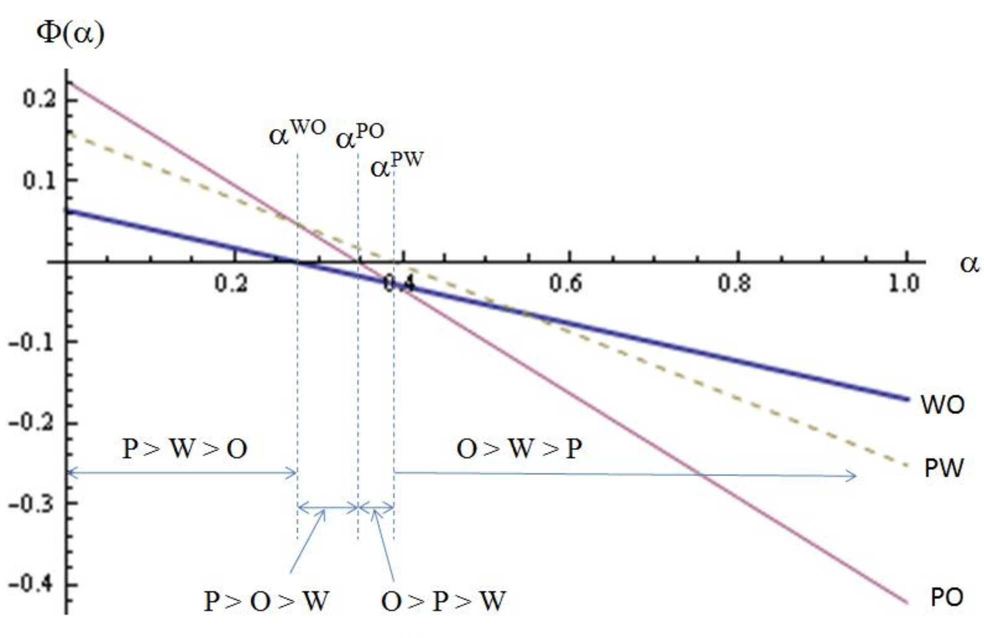

Figure 4: *Ex Ante* Social Ordering over P, O, and W Policies

functions shifts below the WO Φ-function, making W a first-best policy over a range of α-values, which depends upon the magnitude of k. Alternatively, if publicity is costly to enforce (and the cost is k), then the PO and WO Φ-functions shift up, while the PW Φ-function does not move. If that cost is greater than $-\Phi^{PO}(\alpha^{PW})$, then the crossing point for the PO and WO Φ-functions shifts up, making W a first-best policy over a range of α-values, which depends upon the magnitude of k.[23]

The foregoing may explain why, for example, a pure policy of privacy is rarely observed (though we will discuss at least one in Section 5); enforcement may be too costly, especially if the person whose actions are deemed private does not desire privacy, or is willing to trade that privacy for some specified advantage (which in the model is captured by the esteem, or future trading opportunities, term). Instead, most policies that are concerned with privacy are designed to be waivable by the person who is the subject of the policy. Even the 5th Amendment's prohibition that no one can be forced to testify against himself is waivable by the defendant. Companies that are particularly privacy-conscious and run web sites that use "cookies" (which can provide tracking information) to personalize customer visits may request permission from a customer to operate such a cookie, or may warn the customer so that they may disable the cookie. Personal insurance contracts often include a limited privacy waiver that allows the insurance firm limited ability to exchange information with corporate partners (e.g., firms providing marketing services for the company in question).

Some waivable privacy policies are concerned with the communication of information to and with advisors. Classic examples of advisor/advisee waivable privacy policies are priest/penitent, lawyer/client, and doctor/patient. In each case one party (the advisor) cannot unilaterally waive privacy, while the other (the advisee) may unilaterally waive privacy. In the case of the spousal privilege, courts enforce the privilege (a pure privacy policy) in the case of "marital communication" by not allowing one spouse to testify against the other with respect to communications in the marriage that were intended to be confidential. However,

[23] The conditions for *interim* preferences over the policies are provided and explored in the Technical Appendix. Unfortunately, these conditions are very complex. However, there we show that conditions exist such that a subset of the types strictly *interim* prefers W to either P or O even without enforcement costs.

in criminal cases, a spouse can be called as a witness concerning criminal acts by a defendant and the witness-spouse can choose to waive the privilege unilaterally.

5. Examples of Policies of Privacy, Publicity, and Waivable Privacy

In this section we provide some examples that illustrate how our model might shed light on some current (and some more speculative) policies of privacy and publicity. In general, when α and β are of the same sign, then either privacy or publicity may be *ex ante* preferred, depending on the strength of the public-good (or public-bad) effect. On the other hand, when α and β are of opposite signs, then a policy of privacy is *ex ante* socially preferred to a policy of publicity. We provide examples for each of the possible parameter configurations below.

5.1 Examples Wherein α > 0 and β > 0

Open-Source Software Development

One plausible application of our model involves open-source software development, wherein independent programmers contribute towards the improvement of a program that is available freely to end-users. These projects typically follow a policy of publicity wherein programmers are credited with their contributions (though of course contributions can be made anonymously). In discussing what motivates programmers to contribute, Lerner and Tirole (2002, p. 213) discuss costs (such as the opportunity cost of the time spent on the project) and benefits, both immediate (such as utility from working on a "cool" open-source project) and delayed (such as possible job opportunities, future access to venture capital, or ego gratification due to peer recognition, presumably all based on others' inferences about the programmer's talent). They refer to these delayed benefits collectively as the "signaling incentive," and provide empirical evidence regarding the benefits that have accrued subsequently to open-source contributors. In terms of our model, the action g_i represents the extent of the improvement made by agent i, while θ_i represents the programmer's talent or dedication to solving the problem. A programmer enjoys intrinsic utility from advancing the project, and the intrinsically-optimal improvement is higher for more talented or dedicated programmers. Spending time on generating the improvement g_i has an opportunity cost, reflected in $p > 0$. Being inferred to be more talented or dedicated contributes positively to esteem or future trading opportunities, which is reflected in $\beta > 0$. Finally, since the resulting software is available freely, these individual improvements contribute to a public good, which is reflected in $\alpha > 0$. If we take the prevailing policy (that is, publicity) to be the *ex ante* optimal policy, then our model would suggest that α is relatively high in the OSS context.

Charitable Giving

Many papers explore various motives for charitable giving, including intrinsic utility and the utility associated with consuming the aggregate public good (Andreoni, 1989), and the desire to signal some attribute to acquire status (Harbaugh, 1998; Glazer and Konrad, 1996). Glazer and Konrad (1996) discuss the use of charitable giving to signal income, where status is assumed to be increasing in perceived (inferred) income. They observe that the charity benefits from the upward-distortion associated with public giving, and that this can result in too much of the public good. However, their paper does not include intrinsic utility and does not evaluate welfare under alternative policies.[24]

In order to analyze charitable giving (such as contributing to the local symphony), the model can be generalized to include the agent's income in the intrinsic utility term. We assume that agent i derives intrinsic utility ("warm glow") from giving according to the utility function $\gamma g_i - (g_i - \theta_i I_i)^2/2$, where g_i denotes agent i's gift and θ_i reflects agent i's generosity. Thus, more generous agents give more, based on intrinsic utility alone. Suppose that the charitable contributions fund a public good, and that agents who are perceived as more generous receive greater esteem or enhanced future trading opportunities; thus, both α and β are

[24] There is a literature on fund-raising that finds that a charity can benefit by making early donations public when donors move sequentially. This occurs because: an informed donor can signal the "quality" of a charity via an upward-distorted gift (Vesterlund, 2003, 2005); there may be complementarities in giving (Romano and Yildirim, 2001); and this may solve coordination problems when a fixed total contribution is required (Andreoni, 1998; Marx and Matthews, 2000). Our model does not involve any of these effects, so the timing of contributions is irrelevant.

positive. Assuming that income is observable, it is straightforward to show that agent i's gift under a policy of privacy is $g^P(\theta_i; I_i) = g_{min} + \theta_i I_i$ for all $\theta_i \in [0, \bar{\theta}]$, while her gift under a policy of publicity is given implicitly by $g^O(\theta_i; I_i) = g_{min} + \theta_i I_i + (\beta/I_i)(1 - \exp[-I_i(g^O(\theta_i; I_i) - g_{min})/\beta])$. It is clear that the gift under privacy, $g^P(\theta_i; I_i)$, is an increasing function of income for all $\theta_i \in (0, \bar{\theta}]$; it can be shown that the gift under publicity, $g^O(\theta_i; I_i)$, is also an increasing function of income for all $\theta_i \in (0, \bar{\theta}]$. However, the distortion $g^O(\theta_i; I_i) - g^P(\theta_i; I_i)$ is a decreasing function of income for all $\theta_i \in (0, \bar{\theta}]$. Thus, an agent might give more either because she is more generous or because she has more income; since we assume that income is observable, the agent's gift (if public) reveals her generosity. Assume that, *ex ante*, agents' incomes are distributed identically and independently of their types (their generosity). If each agent's marginal utility for the public good is low, then a policy of privacy is *ex ante* optimal. On the other hand, if each agent's marginal utility for the public good is high, then a policy of publicity is *ex ante* optimal; while each agent is induced to give more when contributions are made public, each agent benefits substantially from the upward-distorted gifts of the other agents. Contributions to the symphony are acknowledged publicly (often in discrete classifications, rather than actual dollar amounts), and an agent is likely to benefit substantially from others' contributions to her local symphony (thus, α is likely to be large). On the other hand, contributions to global relief funds, where the public good arguably has a smaller direct impact on the giver, are typically not acknowledged publically.[25]

Recycling

Another possible application is to recycling. The action g_i represents the extent to which agent i engages in recycling, while θ_i represents the agent's taste for recycling (or "greenness"). The agent derives some intrinsic utility from recycling, and greener agents naturally engage in more recycling. Recycling may involve a cost ($p > 0$). Being inferred to be more socially-responsible contributes positively to esteem or future trading opportunities, which is reflected in $\beta > 0$. Finally, since recycling conserves scarce resources

[25] Of course, giving to either a local symphony or a global relief fund may generate a very high intrinsic value for the giver. In both cases the policy is waivable; one can make an anonymous contribution to the symphony and one can post a receipt for a contribution to a global relief fund on the web, if desired.

it contributes to a public good, which is reflected in $\alpha > 0$. In the US, recycling is largely a private matter at present, though as more communities adopt curbside recycling and more schools and businesses provide receptacles, it is becoming more public/observable. Moreover, as resource constraints become tighter, the value of α will arguably increase, which might tip the optimal policy from being one of privacy to one of publicity. In urban South Korea, household recycling has been incentivized since 1995; for instance, individuals must purchase specifically-authorized bags for waste disposal and the disposal of separated recyclables is free (Lee, 2003). There is a fine for the use of unauthorized bags, and even a monetary reward for identifying offenders.[26] Individuals deposit waste and recyclables at public collection points near their apartment buildings. Thus there is likely to be a substantial component of observability and (presumably) social disapproval associated with poor compliance.

5.2 Examples Wherein $\alpha > 0$ and $\beta < 0$

When $\beta < 0$, higher perceived types suffer greater disapproval. In the Appendix we show that in this case, the direction of an agent's distortion is downward: every type (except the highest type, which is now viewed as the "worst" type) chooses a lower level of the action when it is observable than when it remains private. If $\alpha > 0$, then this downward distortion reduces both the agent's intrinsic utility and her contribution to the public good. Thus, the case of $\alpha > 0$ and $\beta < 0$ clearly implies that a policy of privacy is always socially-preferred to a policy of publicity.

Consumption of Health Care

A prime example with this pattern of the parameters is the consumption of health care (perhaps especially the consumption of mental health care). The action g_i represents the amount of health care consumed by agent i, while θ_i represents the agent's need for health care. An agent with greater health care

---

[26] According to Jong (2007), "In the early days of implementing the policy, there were news reports of people rummaging through their neighbours' trash to find out the identifies of offenders so that they could report the culprits and get the cash reward."

needs has a higher marginal intrinsic utility for health care.[27] Health care has a cost, which is reflected in $p > 0$. Obtaining needed health care contributes to the public good (that is, $\alpha > 0$) under a variety of interpretations. For instance, it suppresses communicable diseases, and it improves work productivity (to the extent that there are complementarities among workers, these benefits extend beyond the private benefits reaped by the worker). However, being inferred to require more health care may contribute negatively to esteem or future trading opportunities; thus $\beta < 0$. Hence, a policy of privacy results in an increased consumption of health care (relative to a policy of publicity), which provides both a private and a public good.

DNA Dragnets

Another interesting example is the DNA dragnet, wherein a crime leaving DNA evidence has occurred, and members of an entire community are asked to provide a sample.[28] Although the action space here is binary (provide a sample, or do not provide a sample), many of the features of our model are present. Providing a DNA sample may contribute to a public good, that is, the apprehension of a criminal and at (presumably) lower cost; thus $\alpha > 0$. In terms of intrinsic utility, a guilty agent will be disinclined to provide a sample, but even innocent agents might be concerned about future negative consequences (e.g., the sample might be retained despite the promise that it would be disposed of at the closure of the investigation; it might be mis-handled, resulting in a risk of conviction in the instant, or some subsequent, crime; or the sample may contain other information not pertinent to the instant crime, but which might leave the agent open to blackmail). Thus, one might think of θ_i as reflecting an innocent agent's degree of suspicion about the police and the use that might be made of the DNA. People who are perceived to be more suspicious (or less trusting) may incur more disapproval (for unwillingness to cooperate in apprehending a perpetrator) or lower future

[27] The form of the intrinsic utility function implies that maximized intrinsic utility is increasing in θ, which does not seem plausible for the application to health care (since it implies that agents with greater health care needs, if they receive intrinsically-optimal care, have higher utility). However, it is straightforward to modify the intrinsic utility function (e.g., by subtracting a term that depends on θ_i but not g_i) to reverse this implication without changing the optimal action g_i or any of the model's other implications.

[28] DNA dragnets first appeared in the UK in 1987 in a serial rape and kidnaping case, and are now employed in the UK, Europe, and more recently, the US; see Drobner (2000).

trading opportunities, as well as being pooled with the guilty party, so $\beta < 0$. Thus, one would predict that the set of types would be partitioned into one set that volunteers a sample (these would be the less-suspicious types) and another set that declines to provide a sample (these would be the more-suspicious types, plus the guilty party). Declining to provide a sample results in an inference that the agent is, at best, a suspicious type and, at worst, the guilty party. Although each agent has a right to privacy (absent probable cause), it is waivable, and the coercive effect of waivable privacy is clear: there will be some (perhaps very many) types who would prefer not to provide their DNA, but end up doing so in order to avoid the adverse inference.

5.3 Examples Wherein $\alpha < 0$ and $\beta > 0$

Here the setting is an action that creates individual esteem but contributes to a public bad. When α is negative but $\beta > 0$, then again a policy of privacy is always *ex ante* optimal.

Student Rankings

An interesting policy is the University of California at Berkeley Boalt Hall Law School's policy of not reporting class rank to potential employers.[29] The law school does not report class rank; indeed, it provides the student with this information in non-verifiable form and makes it an honor code violation to reveal it, so this privacy is not even waivable. In terms of our model, g_i is student i's competitive effort and θ_i is agent i's talent; more talented students experience higher marginal utility of effort and optimally work harder. Moreover, the student's final class rank is likely to affect positively the esteem in which the student is held by peers and his or her future employment options, so $\beta > 0$. The use of such a policy suggests that the Law School views the resulting expansion in competitive effort (if class rank were public) as being dissipative or even counter-productive to educational objectives; thus α is small or even negative.

Inadmissibility of Settlement Offers at Trial

Another example in which non-waivable privacy is the prevailing policy is Federal Rule of Evidence

---

[29] This policy is available at www.law.berkeley.edu/students/registrar/academicrules/ ; see Section 3.06. Exceptions are made for students applying for clerkships or academic positions that require the information. We thank Eric Talley for drawing this policy to our attention.

408, which makes the details of failed settlement negotiations inadmissible in court. As shown in Daughety and Reinganum (1995),[30] a plaintiff with private information about her damages has a greater incentive to inflate her demand when it will also be observed by the judge should the case come to trial, as a higher demand is inferred to reflect higher damages implying a higher reward at trial (and thus, $\beta > 0$). This results in more failed negotiations, more cases coming to trial, and thus greater court congestion and increased litigation costs, which is a public bad (and thus, $\alpha < 0$). Privacy is the *ex ante* optimal policy in this case, and privacy is ensured by Rule 408.

5.4 Examples Wherein $\alpha < 0$ and $\beta < 0$

When both α and β are negative, then a policy of publicity may be *ex ante* preferred; by making their actions public, agents are induced to choose lower levels, thus contributing less to a public bad. However, a policy of publicity will be preferred to a policy of privacy only if α is sufficiently negative.[31]

Electricity and Water Usage During Periods of Voluntary Rationing

Electricity or water usage are examples wherein greater use contributes to a public bad (air pollution and depletion of the water table, respectively), and may be viewed adversely by other members of society. In this application, g_i is agent i's consumption of electricity or water, while θ_i represents an attribute such as selfishness or wastefulness. In this case, publicity (or "shaming") can induce reductions in use; this may or may not be socially-optimal, since individuals' intrinsic utility will also be reduced as they consume less than their intrinsically-optimal level of electricity or water. Although, currently, agents' usage of electricity or water tends to be private, technologies are being developed that would allow individuals to elect publicity,

---

[30] The model in Daughety and Reinganum (1995) differs in the payoff structure, but contains the same elements: intrinsic utility, a public-bad aspect, and a continuation value based on inferred type.

[31] The following examples are about actions that generate a relatively mild public bad, as actions that generate a severe public bad are generally the subject of criminal fines and penalties. Moreover, we have restricted the analysis to a symmetric agent model, so that would require that all agents are willing to engage in such activities and create a public bad, which is more likely to be true of a mild public bad than a severe one. Finally, focusing on a severe public bad raises the separate question as to whether a perpetrator's utility belongs in the social welfare function.

and compulsory publicity is not outside the realm of possibility. For instance, according to Thompson (2008), a recent invention (called the Wattson) "not only shows your energy usage but can also transmit the data to a Web site, letting you compare yourself with other Wattson users worldwide." The idea is that "You'd work harder to conserve so you don't look like a jackass in front of your peers." Goodman (2007) reports that during a recent extended drought in Georgia, a Marietta man's water usage was disclosed by the local ABC affiliate (approximately 14,700 gallons per day, compared to an average in the Atlanta area of about 183 gallons per day). A public relations specialist for the man indicated that he "had only recently become aware of the severity of the water crisis and was now taking steps to conserve." Cobb County subsequently released the names of ten more major water users.

Shaming Speeders

Driving faster than the posted speed limit is another potential application of our model. Here, the amount of the public bad is best viewed as the average amount of speeding by drivers, rather than the aggregate amount of speeding, but this is a trivial modification of the model. Everyone engages in speeding, to some extent; let g_i denote agent i's speed in excess of the limit, and let θ_i represent an attribute such as selfishness or carelessness toward others. Depending on its extent, speeding has substantial negative externalities in terms of the risk of accident and injury to others (so $\alpha < 0$) and selfish or careless types are likely to receive social disapproval (so $\beta < 0$). Thus, a policy of publicity, which is predicted to reduce speeding, may be optimal. Such a policy has been instituted recently in the UK. As observed in a recent press release concerning this policy: "The London Safety Camera Partnership has installed England's first fixed speed indicator device with automatic number plate recognition. Drivers who break the 30 mph limit as they approach Richmond Circus will see their speed and number plate flash up on the roadside screens. It is hoped the embarrassment of seeing their illegal driving illuminated in this way will encourage motorists to stick to the limit. If the trial proves successful, the device could be rolled out London-wide" (see London Borough of Richmond upon Thames, 2008). A similar policy and device has been tested on the M42 in the Midlands, UK, and "almost half of drivers breaking the limit slowed" (see Auto Express News, 2006).

Earmark Publicity

A recent political issue that has arisen is the propriety of allowing legislators to pass bills with "earmarks" which fund pet projects without identifying the sponsor of the earmark; the earmark directs funds and typically avoids the standard process wherein funds go through a federal agency and are subject to executive control. Earmarks are frequently added to legislation after the basic bill has been passed by both houses, during the "conference" phase; as such they occur essentially in secret. Let g_i represent Senator i's proposed spending on earmarks and let θ_i reflect Senator i's willingness to re-direct public funds to pet projects; this suggests that $\beta < 0$, reflecting social disapproval. From the members' point of view, α is negative (earmarks result in a bloated budget or come at the cost of other, more worthy projects). The perception that earmarking was out of control led to the Legislative Transparency and Accountability Act of 2006 which passed the Senate on a vote of 90 to 8; the Act required a formal listing of earmarks and of each earmark's sponsor's name (this legislation, however, eventually failed to pass both houses of Congress).[32]

6. Summary and Conclusions

In this paper we develop an economic model of privacy, concentrating on privacy of action. Privacy of action means that relevant actions are not publically observable but rather are protected from the public glare. Under privacy of action, agents choose their full-information optimal actions. Our model incorporates three primary elements: 1) an intrinsic value for the activity involved; 2) esteem (or in some examples, social disapproval); and 3) consumption of any public-good (or public-bad) aspects that arise from the aggregate activity of all individuals. We show that privacy can be welfare-enhancing in both *ex ante* expected social welfare terms and in *interim* (that is, type-specific) terms, though a conflict can readily arise between the policy that is *ex ante* best and the policy that the median type *interim* prefers. If a pure policy of privacy (or

[32] See http://www.sourcewatch.org/index.php?title=Earmarks (accessed April 10, 2008) for a more detailed discussion of the earmark process and recent legislation (which has failed) to publicize earmarks.

a pure policy of publicity) is sufficiently costly to enforce, then a policy of waivable privacy can be *ex ante* socially preferred; otherwise, waivable privacy is never the socially-preferred policy. A policy of waivable privacy gives rise to two types of pure-strategy equilibria: 1) one wherein all types choose to publically reveal their actions (with the concomitant distortion of action) and 2) one wherein a subset of types (possibly empty) chooses to make their actions public while the rest choose to keep their actions private. Thus, for example, we should expect that agents will differ in their use of opt-out clauses in contracts wherein an agent's private information used by a company is proposed to be shared with other commercial interests.

We applied our model in a number of settings, but the bottom line is that there is an *ex ante* expected social preference for privacy when the effects of esteem and the marginal utility of the public good enter the agent's utility function with different signs. On the other hand, there will be ranges of social preference for privacy or publicity when these forces work in the same direction, with the primary dividing point depending upon the magnitude of the public-good (or public-bad) effect: only when this effect sufficiently outweighs the disutility of the action distortion due to publicity will it be optimal to choose publicity. Finally, we found that privacy of action provides a form of protection for privacy of type that may be more effective and (since under privacy of action, an agent chooses his full-information optimal action) more efficient than direct privacy of type itself.

References

Acquisti, Alessandro and Hal R. Varian. "Conditioning Prices on Purchase History," Marketing Science 24 (2005), 367-381.

Andreoni, James. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," Journal of Political Economy 97 (1989), 1447-1458.

Andreoni, James. "Toward a Theory of Charitable Fund-Raising," Journal of Political Economy 106 (1998), 1186-1213.

Becker, Gary S. "A Theory of Social Interactions," Journal of Political Economy 82 (1974), 1063-1093.

Benabou, Roland and Jean Tirole. "Incentives and Prosocial Behavior," American Economic Review 96 (2006), 1652-1678.

Bernheim, B. Douglas. "A Theory of Conformity," Journal of Political Economy 102 (1994), 841-877.

Calzolari, Giacomo and Alessandro Pavan. "On the Optimality of Privacy in Sequential Contracting," Journal of Economic Theory 130 (2006), 168-204.

Cho, In-Koo and David M. Kreps. "Signaling Games and Stable Equilibria," Quarterly Journal of Economics 102, 179-221.

Daughety, Andrew F. and Jennifer F. Reinganum. "Keeping Society in the Dark: On the Admissibility of Pretrial Negotiations as Evidence in Court," The RAND Journal of Economics 26 (1995), 203-221.

Daughety, Andrew F. and Jennifer F. Reinganum. "Communicating Quality: A Unified Model of Disclosure and Signaling," Working Paper No. 07-W03, Department of Economics, Vanderbilt University, January 2007 (revised January 2008).

Drobner, Fred W. "DNA Dragnets: Constitutional Aspects of Mass DNA Identification and Testing," Capital University Law Review 28 (2000), 479-511.

Fudenberg, D. and Tirole, J. *Game Theory*. Cambridge, MA: MIT Press, 1991.

Glazer, Amihai and Kai Konrad. "A Signaling Explanation for Charity," American Economic Review 86 (1996), 1019-1028.

Goodman, Brenda (2007), "Amid Drought, a Georgian Consumes a Niagara," *New York Times*, November 15, 2007; available at www.nytimes.com/2007/11/15/us/15water.html?fta=y; accessed April 4, 2008.

Grossman, Sanford. "The Informational Role of Warranties and Private Disclosure about Product Quality," Journal of Law and Economics 24 (1981), 461-483.

Harbaugh, William T. "What Do Donations Buy? A Model of Philanthropy Based on Prestige and Warm Glow," Journal of Public Economics 67 (1998), 269-284.

Hermalin, Benjamin E. and Michael L. Katz. "Privacy, Property Rights & Efficiency: The Economics of Privacy as Secrecy," Quantitative Marketing and Economics 4 (2006), 209-239.

Jong, Lee Tee, "Koreans Trash the Waste Problem," *Straits Times*, September 30, 2007, available at www.wildsingapore.com/news/20070910/070930-1.htm; accessed 4/9/08.

Lee, Sang Don (2003), "Waste Reduction and Recyling Law in Korea," American Bar Association available at www.abanet.org/environ/committees/intenviron/newsletter/feb03/korea/; accessed 4/9/08.

Linardi, Sera and Margaret A. McConnell. "Volunteering and Image Concerns," working paper, California Institute of Technology, January 2008.

Marx, Leslie M. and Steven A. Matthews. "Dynamic Voluntary Contribution to a Public Project," Review of Economic Studies 67 (2000), 327-358.

Milgrom, Paul. "Good News and Bad News: Representation Theorems and Applications," Bell Journal of Economics 12 (1981), 380-391.

Posner, Richard A. "The Right of Privacy," Georgia Law Review 12 (1978), 393-422.

Posner, Richard A. "The Economics of Privacy," The American Economics Review 71 (1981), 405-409.

Romano, Richard and Huseyin Yildirim, "Why Charities Announce Donations: A Positive Perspective," Journal of Public Economics 81 (2001), 423-447.

"Shaming Speed Deterrent a National First,"February 15, 2008, London Borough of Richmond upon Thames, available at http://www.richmond.gov.uk/shaming_speed_deterrent_a_national_first; accessed April 4, 2008.

"Speeders' Plates in Lights,"Auto Express News, August 31, 2006, available at http://www.autoexpress.co.uk/news/autoexpressnews/202483/speeders_plates_in_lights.html; accessed April 4, 2008.

Spence, Michael. "Job Market Signaling," Quarterly Journal of Economics 87 (1973), 355-374.

Stigler, George J. "An Introduction to Privacy in Economics and Politics," Journal of Legal Studies 9 (1980), 623-644.

Taylor, Curtis R. "Consumer Privacy and the Market for Customer Information," RAND Journal of Economics 35 (2004), 631-650.

Thompson, Clive (2008), "Desktop Orb Could Reform Energy Hogs;" *Wired*, July 24, 2007 available at www.wired.com/print/techbiz/people/magazine/15-08/st_thompson; accessed April 4, 2008.

Vesterlund, Lise. "The Informational Value of Sequential Fundraising," Journal of Public Economics 87 (2003), 627-657.

Vesterlund, Lise. "After You – Endogenous Sequencing in Voluntary Contribution Games," Journal of Public Economics 89 (2005), 1399-1419.

Appendix

This appendix provides a more detailed derivation of the separating equilibrium (both for $\beta > 0$ as described in Proposition 1 and for $\beta < 0$); details of the computations discussed in the text. Proofs of comparative statics results and an analysis of *interim* preferences over privacy, publicity, and waivable privacy can be found in the Technical Appendix.

Derivation of the Separating Equilibrium for $\beta > 0$

In a separating equilibrium it must be that $g^O(\theta_i)$ is either everywhere increasing or everywhere decreasing on $[0, \bar{\theta}]$ (if this function were not monotonic, then two types would be choosing the same action and the equilibrium would not be separating). Therefore, there is a smallest and a largest value of $g^O(\theta_i)$; denote these by a and b, respectively. Furthermore, the equilibrium action function for agent i, $g^O(\theta_i)$, must be continuous on $[0, \bar{\theta}]$. This follows because, in a separating equilibrium, the utility for type θ_i is $V_i(g^O(\theta_i), \theta_i, \theta_i, G_{-i})$, which varies continuously in each of its arguments, so that if $g^O(\theta_i)$ were discontinuous then two arbitrarily-close types would make distinctly different utility payoffs, causing one of them to defect and pool with the other. The fact that $g^O(\theta_i)$ is monotonic and continuous on $[0, \bar{\theta}]$ means that it is invertible to obtain $B^O(g)$ for g in [a, b]; moreover, the function thus derived is monotonic and continuous on [a, b]. This further implies that $B^O(g)$ is differentiable almost everywhere on [a, b]. Since $V_i(g_i, \theta_i, B^O(g_i), G_{-i})$ is differentiable in all its arguments, the total derivative with respect to g_i exists almost everywhere on [a, b]. Differentiating $V_i(g_i, \theta_i, B^O(g_i), G_{-i})$ totally with respect to g_i yields the first-order condition:

$$g_{min} - g_i + \theta_i + \beta B^{O\prime}(g_i) = 0.$$

If the optimal g_i for some θ_i occurs at a point of non-differentiability of $B^O(g)$, then this equation will not hold there, but this can only happen for isolated values of θ_i. Employing consistency (that is, requiring that $B^O(g^O(\theta)) = \theta$) allows us to write the above first-order-condition as a differential equation that must hold almost everywhere:

$$dB^O(g)/dg + B^O(g)/\beta + (g_{min} - g)/\beta = 0.$$

Finally, the relevant boundary condition is $B^O(g_{min}) = 0$; the lowest type, which will be revealed as such in the separating equilibrium, need not distort its action choice. This differential equation has a one-parameter family of solutions; imposing the boundary condition (and, thereafter, the continuity of $B^O(g)$) selects a unique solution, which turns out to be a well-behaved (e.g., twice differentiable) function. This solution is:

$$B^O(g) = g - g_{min} - \beta(1 - \exp[- (g - g_{min})/\beta]).$$

Thus, the function $V_i(g_i, \theta_i, B^O(g_i), G_{-i})$ is actually twice differentiable in g_i, so the first-order condition must hold at $g^O(\theta_i)$. The second-order necessary condition for a maximum is that $-1+ \beta B^{O\prime\prime}(g^O(\theta_i)) \leq 0$. Differentiating $B^O(g)$ twice and evaluating at $g_i = g^O(\theta_i)$ yields $B^{O\prime\prime}(g^O(\theta_i)) = (1/\beta)\exp[-(g_i^O(\theta) - g_{min})/\beta]$. Thus, the second-order necessary condition $-1+ \beta B^{O\prime\prime}(g^O(\theta_i)) \leq 0$ requires that $-1 + \exp[- (g^O(\theta_i) - g_{min})/\beta] \leq 0$; that is, $g^O(\theta_i) \geq g_{min}$.

Differentiating the expression $B^O(g)$ and plugging $B^{O\prime}(g)$ into the first-order condition (or, equivalently, inverting the function $B^O(g)$) results in an equation that defines $g^O(\theta_i)$ implicitly:

$$g^O(\theta_i) = g_{min} + \theta_i + \beta(1 - \exp[- (g^O(\theta_i) - g_{min})/\beta]). \quad \text{(A1)}$$

We have already shown that the maximizing value $g^O(\theta_i)$ is at least g_{min}; therefore the third term in this equation is non-negative. But this implies that $g^O(\theta_i) > g_{min}$ for $\theta_i > 0$, and hence the third term in this equation is strictly positive for $\theta_i > 0$. This finding can be used to prove that (for $\theta_i > 0$) $g^O(\theta_i)$ is strictly increasing in θ_i and $g^O(\theta_i) > g^P(\theta_i)$, and to sign various comparative statics effects.

Computational Results on the Effect of β on α^{PO}

The Table below displays computational results for four density functions: 1) the Uniform density, with $h(\theta) = 1$; 2) the Left Triangle density, with $h(\theta) = 2 - 2\theta$; 3) the Middle Triangle density, with $h(\theta) = 4\theta$ when $\theta \leq ½$, and $h(\theta) = 4 - 4\theta$ when $\theta > ½$; and 4) the Right Triangle density, with $h(\theta) = 2\theta$. Notice that the Uniform density is a mean-preserving spread of the Middle Triangle density.

Table: Computed Effect of β on $E(\delta^2)/E(\delta)$ for Alternative Densities of θ

density↓ β→	0.5	1.0	2.0
Uniform	0.40859	0.69264	1.14159
Left Triangle	0.36996	0.61131	0.96546
Middle Triangle	0.41363	0.69296	1.10361
Right Triangle	0.43900	0.75341	1.22101

The Table suggests that, for a given density, increasing β increases α^{PO}, so that $\Phi^{PO}(\alpha^{PO})$ shifts up, associating more values of α with privacy than were associated with the lower value of β. Also, note that, holding β constant, the computed values of $E(\delta^2)/E(\delta)$ increase as we move from the Left to the Middle to the Right Triangle distributions. Thus, the Table is consistent with the conjecture that a shift in H to a new distribution H′, where H′ first-order stochastic dominates H, results in higher values of α^{PO} as well (i.e., upward shifts of Φ^{PO}, too).

Material on Deriving Waiver Equilibria

Recall the definition of $\Delta(\theta_i; \hat{\theta})$ from the text: $\Delta(\theta_i; \hat{\theta}) \equiv V_i(g^P(\theta_i), \theta_i, \mu(\hat{\theta}), G_{-i}(\hat{\theta})) - V_i(g^O(\theta_i), \theta_i, B^O(g^O(\theta_i)), G_{-i}(\hat{\theta}))$ is the net value for agent i of type θ_i of privacy over publicity, given that all other agents use the strategy and beliefs specified in the text. It is straightforward to show that this net value is decreasing in type. Let θ^W denote an equilibrium value of $\hat{\theta}$. There are three possible types of equilibria with waiver. First, $\theta^W = 0$ is an equilibrium if and only if $\Delta(0; 0) \leq 0$; to see this, note that if all other agents' types choose O and $\Delta(0; 0) \leq 0$, then $\Delta(\theta_i; 0) < 0$ for all $\theta_i > 0$, so all of agent i's types will also choose O. This is a full-waiver equilibrium, in which every type discloses his action and chooses his action according to $g^O(\theta)$.

Second, $\theta^W = \bar{\theta}$ is an equilibrium if and only if $\Delta(\bar{\theta}; \bar{\theta}) \geq 0$; to see this, note that if all other agents' types choose P and $\Delta(\bar{\theta}; \bar{\theta}) \geq 0$, then $\Delta(\theta_i; \bar{\theta}) > 0$ for all $\theta_i < \bar{\theta}$, so all of agent i's types will also choose P. This is a no-waiver equilibrium, in which no type discloses his action and every type chooses his action according to $g^P(\theta)$. Finally, $\theta^W \in (0, \bar{\theta})$ is an equilibrium if and only if $\Delta(\theta^W; \theta^W) = 0$; to see this, note that if all other agents' types choose P when $\theta < \theta^W$ and O when $\theta \geq \theta^W$ and if $\Delta(\theta^W; \theta^W) = 0$, then $\theta_i = \theta^W$ is indifferent between P and O (and hence willing to choose O). Moreover, $\Delta(\theta_i; \theta^W) > 0$ for $\theta_i < \theta^W$ and $\Delta(\theta_i; \theta^W) < 0$ for $\theta_i > \theta^W$; that is, agent i will choose P for $\theta_i < \theta^W$ and O for $\theta_i > \theta^W$. Thus, we have confirmed that if all other agents choose P (and the action $g^P(\theta)$) when $\theta < \theta^W$ and choose O (and the action $g^O(\theta)$) when $\theta \geq \theta^W$, then it will be a best response for agent i to do so as well. When $\theta^W \in (0, \bar{\theta})$, we will refer to this as a partial-waiver equilibrium in which some agent types disclose their actions, while others keep their actions private.

Calculation yields: $\Delta(\theta_i; \theta^W) = (\delta(\theta_i; \beta))^2/2 + \beta[\mu(\theta^W) - \theta_i]$. Thus, there is always an equilibrium at $\theta^W = 0$, since $\Delta(0; 0) = 0$. In this equilibrium, all types choose to waive privacy and we obtain the full-publicity outcome discussed in Section 3.3 as an equilibrium. Moreover, using the fact that $g^{O\prime}(\theta) = 1/(1 - \exp[-(g^O(\theta) - g_{min})/\beta]) > 0$, it is straightforward to show that $d\Delta(\theta; \theta)/d\theta = \beta[\mu'(\theta) - (1 - \exp[-(g^O(\theta) - g_{min})/\beta])] > 0$ when evaluated at $\theta = 0$, so that there is at least one more equilibrium. If $\Delta(\bar{\theta}; \bar{\theta}) \geq 0$ then there is a no-waiver equilibrium wherein all types choose not to waive privacy and we obtain the full-privacy outcome discussed in Section 3.2. Finally, if $\Delta(\bar{\theta}; \bar{\theta}) < 0$ then a no-waiver equilibrium does not exist, but there is at least one partial-waiver equilibrium as described earlier. A necessary and sufficient condition for such an interior equilibrium to exist is that distorting so as to signal type is not too costly in the sense that $\mu < \bar{\theta} - (\delta(\bar{\theta}; \beta))^2/2\beta$. Moreover, such an interior equilibrium will be unique if $\Delta(\theta^W; \theta^W)$ is concave; a sufficient condition for this to hold is that $\mu''(\theta) \leq 0$ for all θ.

Derivation of the Separating Equilibrium for $\beta < 0$

All of the arguments given above for the case of $\beta > 0$ still apply to the case of $\beta < 0$, with one exception (the boundary condition). Thus, the ordinary differential equation $dB^O(g)/dg + B^O(g)/\beta + (g_{min} - g)/\beta = 0$ still characterizes the equilibrium relationship between beliefs and actions. However, now the "weakest" type is type $\bar{\theta}$, so this is the type which need not distort its action to be identified. Consequently, the relevant boundary condition becomes $g^O(\bar{\theta}) = g^P(\bar{\theta}) = g_{min} + \bar{\theta}$ or, in terms of the beliefs, $B^O(g_{min} + \bar{\theta}) = \bar{\theta}$.

This differential equation has a one-parameter family of solutions; imposing the boundary condition (and, thereafter, the continuity of $B^O(g)$) selects a unique solution, which turns out to be a well-behaved (e.g., twice differentiable) function. This solution is:

$$B^O(g) = g - g_{min} - \beta(1 - \exp[(g_{min} + \bar{\theta} - g)/\beta]).$$

Thus, the function $V_i(g_i, \theta_i, B^O(g_i), G_{-i})$ is actually twice differentiable in g_i, so the first-order condition must hold at $g^O(\theta_i)$. The second-order necessary condition for a maximum is that $-1 + \beta + B^{O\prime\prime}(g^O(\theta_i)) \leq 0$. Differentiating $B^O(g)$ twice and evaluating at $g = g^O(\theta_i)$ yields $B^{O\prime\prime}(g^O(\theta_i)) = (1/\beta)\exp[(g_{min} + \bar{\theta} - g^O(\theta_i))/\beta]$. Thus, the second-order necessary condition $-1 + \beta B^{O\prime\prime}(g^O(\theta_i)) \leq 0$ requires that $-1 + \exp[(g_{min} + \bar{\theta} - g^O(\theta_i))/\beta] \leq 0$; that is (since $\beta < 0$), $g^O(\theta_i)) \leq g_{min} + \bar{\theta}$.

Differentiating the expression $B^O(g)$ and plugging $B^{O\prime}(g)$ into the first-order condition (or, equivalently, inverting the function $B^O(g)$) results in an equation that defines $g^O(\theta_i)$ implicitly:

$$g^O(\theta_i) = g_{min} + \theta_i + \beta(1 - \exp[(g_{min} + \bar{\theta} - g^O(\theta_i))/\beta]). \tag{A2}$$

We have already shown that the maximizing value $g^O(\theta_i)$ does not exceed $g_{min} + \bar{\theta}$; therefore the term in the exponential function is non-positive, and overall the third term in (A2) is non-positive. But this implies that $g^O(\theta_i) < g_{min} + \bar{\theta}$ for $\theta_i < \bar{\theta}$, and hence (since β is negative) the third term in this equation is strictly negative for $\theta_i < \bar{\theta}$. Thus, it follows that (for $\theta_i < \bar{\theta}$) $g^O(\theta_i)$ is strictly increasing in θ_i and that $g^O(\theta_i) < g^P(\theta_i)$. Thus, the agent's action is now downward-distorted under publicity, with the greatest distortion occurring for the lowest type, $\theta_i = 0$. In order to ensure that $g^O(\theta_i) \geq 0$, it is sufficient to assume that $g_{min} \geq -\beta$.

Technical Appendix for
"Privacy, Publicity, and Choice"
by Andrew F. Daughety and Jennifer F. Reinganum

This technical appendix includes proofs of comparative statics results; the proof of the claim made in the text that if $\beta' > \beta$, then $H^\delta(\delta; \beta')$ first-order stochastic dominates $H^\delta(\delta; \beta)$; and detailed analysis of the *interim* preferences over policies.

Comparative Statics

The functions $g^P(\theta_i)$ and $g^O(\theta_i)$ depend on θ_i and the parameters α, β, γ, and p.

Comparative statics of $g^P(\theta_i)$

Since $g^P(\theta_i) = g_{min} + \theta_i$ and $g_{min} = \gamma + \alpha - p$, it is obvious that $g^P(\theta_i)$ is an increasing function of θ_i, and that the function $g^P(\theta_i)$ shifts upward with an increase in α and γ, and shifts downward with an increase in p. Finally, the function $g^P(\theta_i)$ is always independent of β. Since the utility function is quasilinear, $g^P(\theta_i)$ is independent of income, I_i.

Comparative statics of $g^O(\theta_i)$

Since $g^O(0) = g^P(0)$, $g^O(0)$ behaves as described above with respect to the parameters. Thus, in what follows, we will consider only $\theta_i > 0$. Let RHS $\equiv g_{min} + \theta_i + \beta(1 - \exp[- (g^O(\theta_i) - g_{min})/\beta])$. For any parameter m, the implicit function in (A1) can be differentiated to obtain $\partial g^O/\partial m = (\partial RHS/\partial m) + (\partial RHS/\partial g^O)(\partial g^O/\partial m)$. Collecting terms implies that $\partial g^O/\partial m = (\partial RHS/\partial m)/(1 - \exp[-(g^O(\theta_i) - g_{min})/\beta])$. Since the denominator is positive, the sign of $\partial g^O/\partial m$ is the same as the sign of $(\partial RHS/\partial m)$. To save on notation, it will be useful to define the function $z^O(\theta_i) \equiv (g^O(\theta_i) - g_{min})/\beta$, and to use z to denote an arbitrary (positive value).

Since $\partial RHS/\partial\theta_i = 1$, it follows that $g^{O\prime}(\theta_i) = 1/(1 - \exp[-z^O(\theta_i)]) > 0$; that is, the equilibrium action under a policy of publicity (openness) is increasing in type.

Since the parameters α, γ, and p appear only in g_{min}, and $(\partial RHS/\partial g_{min}) = (1 - \exp[-z^O(\theta_i)])$, it is straightforward to show that $\partial g^O(\theta_i)/\partial g_{min} = 1$. Therefore $\partial g^O(\theta_i)/\partial\alpha = \partial g^O(\theta_i)/\partial\gamma = 1$ and $\partial g^O(\theta_i)/\partial p = -1$.

Differentiating and collecting terms yields $\partial g^O(\theta_i)/\partial\beta = (1 - \exp[-z^O(\theta_i)] - z^O(\theta_i)\exp[-z^O(\theta_i)])/(1 - \exp[-z^O(\theta_i)])$. The function $1 - \exp[-z] - z\exp[-z]$ is easily shown to be positive for $z > 0$; thus, $\partial g^O(\theta_i)/\partial\beta > 0$.

Comparative Statics of the Extent of Distortion $g^O(\theta_i) - g^P(\theta_i)$

Let $\delta(\theta_i; \beta) \equiv g^O(\theta_i) - g^P(\theta_i) = \beta(1 - \exp[-(z^O(\theta_i)])$ denote the extent of (upward) distortion as a function of θ_i. The distortion is increasing in type; that is, $\delta'(\theta_i; \beta) = \exp[-(z^O(\theta_i)]g^{O\prime}(\theta_i) > 0$. Thus, the highest type distorts his action the most.

We have already seen that $\partial g^O(\theta_i)/\partial g_{min} = 1$; this yields the immediate (and very useful) result that $\partial z^O(\theta_i)/\partial g_{min} = (\partial(g^O(\theta_i) - g_{min})/\partial g_{min})/\beta = 0$. This implies that the distortion $\delta(\theta_i; \beta)$ is independent of the parameters α, γ, and p.

Since $g^P(\theta_i)$ is independent of β, then $\partial\delta(\theta_i; \beta)/\partial\beta = \partial g^O(\theta_i)/\partial\beta > 0$.

Proof of claim that if $\beta' > \beta$, then $H^\delta(\delta; \beta')$ first-order stochastic dominates $H^\delta(\delta; \beta)$

Recall that $\delta(\theta; \beta) = \beta(1 - \exp[- (g^O(\theta) - g_{min})/\beta])$, and let $\bar{t}(\beta) \equiv \delta(\bar{\theta}; \beta)$ for any given β; since $\delta(\bar{\theta}; \beta)$ is increasing in β, so is $\bar{t}(\beta)$. Therefore the support of $H^\delta(t; \beta)$ induced by $H(\theta)$ and $\delta(\theta; \beta)$ is $[0, \bar{t}(\beta)]$. Then, fixing β:

$H^{\delta}(t;\beta) \equiv \Pr\{\delta(\theta_i;\beta) \le t\} = \Pr\{\theta \le (g^O)^{-1}(\beta\ln(\beta/(\beta - t) + g_{min})\} = H((g^O)^{-1}(\beta\ln(\beta/(\beta - t) + g_{min}))$.

Thus, $\partial H^{\delta}(t;\beta)/\partial\beta = h(t)[((g^O)^{-1}(t))'(\ln(\beta/(\beta - t) + g_{min})][\ln\beta + 1 - \ln(\beta - t) - \beta/(\beta - t)]$, so that $\partial H^{\delta}(t;\beta)/\partial\beta < 0$ if and only if $\ln\beta + 1 - \ln(\beta - t) - \beta/(\beta - t) < 0$. Note that $H^{\delta}(0;\beta) = 0$ and $H^{\delta}(\bar{t}(\beta);\beta) = \Pr\{\delta(\bar{\theta};\beta) \le \bar{t}(\beta)\} = 1$ for any given value of β, so we are interested in $\partial H^{\delta}(t;\beta)/\partial\beta$ for $t \in (0, \bar{t}(\beta))$.[1] Note that $\ln\beta + 1 - \ln(\beta - t) - \beta/(\beta - t) < 0$ if and only if $\ln(\beta/(\beta - t)) < 1 - \beta/(\beta - t)$ for t in this open interval. Note that $t < \beta$ since $(1 - \exp[-(g^O(\theta) - g_{min})/\beta]) < 1$. Thus, we may restate the problem as: is $\ln x < x - 1$ for $x \ge 1$? In fact, the line $x - 1$ is the tangent to $\ln x$ at $x = 1$, so $\ln x < x - 1$ for $x > 1$ and the two functions are equal at $x = 1$. Therefore, $\partial H^{\delta}(t;\beta)/\partial\beta < 0$ for $t \in (0, \bar{t}(\beta))$, so that if $\beta' > \beta$, then $H^{\delta}(t;\beta') < H^{\delta}(t;\beta)$ for $t \in (0, \bar{t}(\beta))$; that is, $H^{\delta}(t;\beta')$ first-order stochastic dominates $H^{\delta}(t;\beta)$.

Material on Interim Preferences over Policies P and O

This material pertains to Proposition 4. Two results follow from equation (6). First, comparing with equation (4), we see that $E(\Gamma^{PO}(\theta,\alpha)) = \Phi^{PO}(\alpha)/n$, so that when evaluated at $\alpha = \alpha^{PO}$, $E(\Gamma^{PO}(\theta,\alpha^{PO}),\alpha^{PO}) = 0$. Since differentiating $\Gamma^{PO}(\theta_i,\alpha)$ shows that it is a monotonically decreasing function of θ_i for each value of α, this implies that $\Gamma^{PO}(0,\alpha^{PO}) > 0$ and $\Gamma^{PO}(\bar{\theta},\alpha^{PO}) < 0$, so that on an *interim* basis, if $\alpha = \alpha^{PO}$, then some types will (*interim*) prefer P to O and some will (*interim*) prefer O to P. Define two other values of α, namely $\underline{\alpha}^{PO}$ such that $\Gamma^{PO}(\bar{\theta},\underline{\alpha}^{PO}) = 0$ (that is, the value of α such that all types will *interim* prefer P to O for any $\alpha \le \underline{\alpha}^{PO}$), and $\bar{\alpha}^{PO}$ such that $\Gamma^{PO}(0,\bar{\alpha}^{PO}) = 0$ (that is, the value of α such that all types will *interim* prefer O to P for any $\alpha \ge \bar{\alpha}^{PO}$). By construction, $\underline{\alpha}^{PO} < \alpha^{PO} < \bar{\alpha}^{PO}$. Furthermore, when $\alpha \le \underline{\alpha}^{PO}$, the *ex ante* social preference for P over O is therefore reinforced by *interim* unanimity for P over O, while when $\alpha \ge \bar{\alpha}^{PO}$, the *ex ante* social preference for O over P is reinforced by *interim* unanimity for O over P. However, when α lies between $\underline{\alpha}^{PO}$ and $\bar{\alpha}^{PO}$, some types prefer P to O while the rest of the types prefer O to P, so that for all α in the interval $(\underline{\alpha}^{PO}, \bar{\alpha}^{PO})$ there is disagreement about the preferred policy at the *interim* stage, and there will not be unanimous reinforcement of any *ex ante* policy choice.

Proof of Proposition 6(a)

Proposition 6(a) provides the following ordering of the α-values at which there is *ex ante* indifference between any two policies: $0 < \alpha^{WO} < \alpha^{PO} < \alpha^{PW}$. To see that $0 < \alpha^{WO} < \alpha^{PO}$, let

$$\eta(t) \equiv \int_0^t (\delta(\theta;\beta))^2 h(\theta)d\theta / \int_0^t \delta(\theta;\beta)h(\theta)d\theta.$$

Then $\alpha^{WO} = \eta(\theta^W)$, which is clearly positive, while $\alpha^{PO} = \eta(\bar{\theta})$. It is straightforward to show that $\text{sgn}\{\eta'(t)\} = \text{sgn}\{\delta(t;\beta)\int_0^t \delta(\theta;\beta)h(\theta)d\theta - \int_0^t (\delta(\theta;\beta))^2 h(\theta)d\theta\} > 0$ for all $t > 0$. Therefore, it follows that $\alpha^{PO} = \eta(\bar{\theta}) > \eta(\theta^W) = \alpha^{WO}$.

To see that $\alpha^{PO} < \alpha^{PW}$, let

$$\nu(s) \equiv \int_s^{\bar{\theta}} (\delta(\theta;\beta))^2 h(\theta)d\theta / \int_s^{\bar{\theta}} \delta(\theta;\beta)h(\theta)d\theta.$$

Then $\alpha^{PO} = \nu(0)$, while $\alpha^{PW} = \nu(\theta^W)$. It is straightforward to show that $\text{sgn}\{\nu'(s)\} = \text{sgn}\{\int_s^{\bar{\theta}} (\delta(\theta;\beta))^2 h(\theta)d\theta -$

---

[1] Note that increasing β increases the right end-point, so this means we must extend $H^{\delta}(t;\beta)$ to be 1 on the interval $[\bar{t}(\beta), \bar{t}(\beta')]$ when we compare it to the distribution $H^{\delta}(t;\beta')$, so that they are on the same support.

$\delta(s; \beta)\int_s^{\bar{\theta}} \delta(\theta; \beta)h(\theta)d\theta\} > 0$ for all $s < \bar{\theta}$. Therefore, it follows that $\alpha^{PO} = \nu(0) < \nu(\theta^W) = \alpha^{PW}$. QED.

Material on Interim Preferences over Policies P, O and W

Throughout this discussion we assume that $\theta^W \in (0, \bar{\theta})$; if not, then the policy W coincides with either O or P and there are not three distinct policies to be compared.

Recall that the conditional mean is $\mu(\theta^W) = \int_T th(t)dt/H(\theta^W)$, where $T = [0, \theta^W]$. Furthermore, let $E(g^O - g^P)$ denote the expected distortion under a policy of O versus a policy of P, and similarly for $E(g^W - g^P)$ and $E(g^O - g^W)$. Then:

(a) $E(g^O - g^P) = \int \delta(t; \beta)h(t)dt$, where the integral is taken over $[0, \bar{\theta}]$;
(b) $E(g^W - g^P) = \int_{TC} \delta(t; \beta)h(t)dt$, where the integral is taken over $TC = [\theta^W, \bar{\theta}]$;
(c) $E(g^O - g^W) = \int_T \delta(t; \beta)h(t)dt$, where the integral is taken over $T = [0, \theta^W]$.

The integral in part (a) reflects the fact that every type (except the lowest) distorts her action under a policy of O while no type distorts her action under a policy of P. The integral in part (b) reflects the fact that only those types in $TC = [\theta^W, \bar{\theta}]$ distort their actions. Finally, the integral in part (c) reflects the fact that only those types in $T = [0, \theta^W]$ do not distort their actions.

These definitions allow us to summarize the type-specific value of one policy over another. Let $\Gamma^{PO}(\theta_i, \alpha) \equiv V_i(g^P(\theta_i), \theta_i, \mu, G^P_{-i}) - V_i(g^O(\theta_i), \theta_i, \theta_i, G^O_{-i})$ denote the type-specific value of a policy of privacy over a policy of publicity. Then:

$$\Gamma^{PO}(\theta_i, \alpha) = \beta(\mu - \theta_i) + (\delta(\theta_i; \beta))^2/2 - \alpha(n - 1)E(g^O - g^P).$$

Similarly, let $\Gamma^{PW}(\theta_i, \alpha) \equiv V_i(g^P(\theta_i), \theta_i, \mu, G^P_{-i}) - V_i(g^W(\theta_i), \theta_i, \theta_i, G^W_{-i})$ denote the type-specific value of a policy of privacy over a policy of waiver. Then:

$$\Gamma^{PW}(\theta_i, \alpha) = \beta(\mu - \mu(\theta^W)) - \alpha(n - 1))E(g^W - g^P) \text{ for } \theta_i < \theta^W.; \text{ and}$$
$$= \beta(\mu - \theta_i) + (\delta(\theta_i; \beta))^2/2 - \alpha(n - 1))E(g^W - g^P), \text{ for } \theta_i \geq \theta^W.$$

Finally, let $\Gamma^{WO}(\theta_i, \alpha) \equiv V_i(g^W(\theta_i), \theta_i, \theta_i, G^W_{-i}) - V_i(g^O(\theta_i), \theta_i, \mu, G^O_{-i})$ denote the type-specific value of a policy of waiver over a policy of publicity. Then:

$$\Gamma^{WO}(\theta_i, \alpha) = \beta(\mu(\theta^W) - \theta_i) + (\delta(\theta_i; \beta))^2/2 - \alpha(n - 1)E(g^O - g^W), \text{ for } \theta_i < \theta^W; \text{ and}$$
$$= - \alpha(n - 1)E(g^O - g^W), \text{ for } \theta_i \geq \theta^W, \text{ for } \theta_i \geq \theta^W.$$

The functions $\Gamma^{PO}(\theta_i, \alpha)$, $\Gamma^{PW}(\theta_i, \alpha)$, and $\Gamma^{WO}(\theta_i, \alpha)$ are continuous in both arguments and strictly decreasing in α; the latter two functions have portions that are constant with respect to θ_i, but they are strictly decreasing in θ_i over the non-constant regions.

We first determine conditions under which there will be non-trivial sets of types who prefer each policy in a binary comparison. In particular, let $\bar{\alpha}^{IJ}$, for IJ = PO, PW, WO, be the value of α for which $\theta_i = 0$ is indifferent between policy I and policy J (for this and any higher value of α, policy J will be preferred to policy I for all types). Then $\bar{\alpha}^{IJ}$ is defined uniquely by $\Gamma^{IJ}(0, \bar{\alpha}^{IJ}) = 0$, yielding:

$\bar{\alpha}^{PO} = \beta\mu/(n - 1)E(g^O - g^P)$;
$\bar{\alpha}^{PW} = \beta(\mu - \mu(\theta^W))/(n - 1)E(g^W - g^P)$;

$\bar{\alpha}^{WO} = \beta\mu(\theta^W)/(n - 1)E(g^O - g^w)$.

Provided that $\alpha < \min \{\bar{\alpha}^{IJ}\}$, there will be at least some (low) types who prefer policy I to policy J in a binary comparison. In order to have at least some (high) types who prefer policy J to policy I in a binary comparison, it must be that $\Gamma^{IJ}(\bar{\theta}, \alpha) < 0$; our hypothesis that $\theta^W < \bar{\theta}$ is enough to guarantee that this holds for all $\alpha > 0$.

Claim 1: If $0 < \alpha < \min \{\bar{\alpha}^{IJ}\}$, then:
(i) there exists a unique $\theta^{IJ}(\alpha) \in (0, \bar{\theta})$ such that $\Gamma^{IJ}(\theta^{IJ}(\alpha), \alpha) = 0$;
(ii) moreover, $\theta^{WO}(\alpha) < \theta^W < \theta^{PW}(\alpha)$ and $\theta^{WO}(\alpha) < \theta^{PO}(\alpha) < \theta^{PW}(\alpha)$.

Proof. By construction, if $0 < \alpha < \min \{\bar{\alpha}^{IJ}\}$, then $\Gamma^{IJ}(0, \alpha) > 0$ and $\Gamma^{IJ}(\bar{\theta}, \alpha) < 0$, for all IJ. First consider IJ = PO. The function $\Gamma^{PO}(\theta, \alpha)$ is continuous and strictly decreasing in θ; therefore there exists a unique value $\theta^{IJ}(\alpha) \in (0, \bar{\theta})$ such that $\Gamma^{IJ}(\theta^{IJ}(\alpha), \alpha) = 0$. Next consider IJ = PW. The function $\Gamma^{PW}(\theta, \alpha)$ is constant at a positive level for $\theta_i < \theta^W$, and $\Gamma^{PW}(\theta, \alpha) = \Gamma^{PO}(\theta, \alpha) + E(g^O - g^W)$ for $\theta_i \geq \theta^W$. Since this is a continuous and strictly decreasing function, there is a unique value $\theta^{PW}(\alpha) \in (\theta^W, \bar{\theta})$ such that $\Gamma^{PW}(\theta^{PW}(\alpha), \alpha) = 0$. Moreover, this implies that $\Gamma^{PO}(\theta^{PW}(\alpha), \alpha) = - E(g^O - g^W) < 0$, so $\theta^{PO}(\alpha) < \theta^{PW}(\alpha)$. Finally, consider IJ = WO. The function $\Gamma^{WO}(\theta, \alpha)$ is constant at a negative level for $\theta_i \geq \theta^W$; it is a continuous and strictly decreasing function for $\theta_i < \theta^W$. Therefore, there is a unique value $\theta^{WO}(\alpha) \in (0, \theta^W)$ such that $\Gamma^{WO}(\theta^{WO}(\alpha), \alpha) = 0$. Moreover, evaluating Γ^{PO} at this level yields $\Gamma^{PO}(\theta^{WO}(\alpha), \alpha) = \Gamma^{PW}(0, \alpha) > 0$, so $\theta^{WO}(\alpha) < \theta^{PO}(\alpha)$. QED

Note that for the special case of $\alpha = 0$ the claim above still holds with the following minor modifications. Now the function $\Gamma^{WO}(\theta, \alpha)$ starts out positive and declines to zero at θ^W; moreover, it remains constant at zero for $\theta_i \geq \theta^W$. Thus, the equation $\Gamma^{WO}(\theta^{WO}(\alpha), \alpha) = 0$ is satisfied by all members of the set $[\theta^W, \bar{\theta}]$; we take the left-most element as $\theta^{WO}(\alpha)$, and thus $\theta^{WO}(\alpha) = \theta^W$. The rest of the claim continues to hold as stated.

Given the ordering $\theta^{WO}(\alpha) < \theta^{PO}(\alpha) < \theta^{PW}(\alpha)$ derived above, it is straightforward to show that no type finds W to be the best policy. The preference orderings are as follows:

For $\theta \in [0, \theta^{WO}(\alpha))$	$P \succ W \succ O$	(with $W \sim O$ at $\theta^{WO}(\alpha)$)
For $\theta \in (\theta^{WO}(\alpha), \theta^{PO}(\alpha))$	$P \succ O \succ W$	(with $P \sim O \succ W$ at $\theta^{PO}(\alpha)$)
For $\theta \in (\theta^{PO}(\alpha), \theta^{PW}(\alpha))$	$O \succ P \succ W$	($O \succ P \sim W$ at $\theta^{PW}(\alpha)$)
For $\theta \in (\theta^{PW}(\alpha), \bar{\theta}]$	$O \succ W \succ P$	

Now we relax the assumption that $\alpha < \min \{\bar{\alpha}^{IJ}\}$, IJ = PO, PW, WO. It is straightforward to show that $\bar{\alpha}^{PO}$ must lie between $\bar{\alpha}^{PW}$ and $\bar{\alpha}^{WO}$, but we are unable to determine in general whether $\bar{\alpha}^{PW} < \bar{\alpha}^{WO}$ or $\bar{\alpha}^{WO} < \bar{\alpha}^{PW}$ (however, if $\bar{\alpha}^{WO} < \bar{\alpha}^{PW}$, then W can never be *interim*-optimal for any type because $\Gamma^{WO}(0, \alpha) < 0$, implying that O is preferred to W for all types).

As claimed in the text, there are conditions under which some types will most-prefer a policy of W; these conditions are now described. First, it can be shown that $\bar{\alpha}^{PW} < \bar{\alpha}^{WO}$ for the case in which θ is distributed uniformly on $[0, \bar{\theta}]$. For $\bar{\alpha}^{PW} < \alpha < \bar{\alpha}^{WO}$, all types strictly prefer P to W, while those in $[0, \theta^{WO}(\alpha))$ also strictly prefer W to O. So it is possible for some types to *interim*-prefer W to both P and O (however, this set is limited by the fact that $\theta^{WO}(\alpha) < \theta^W$ still holds). Notice that the types who *interim*-prefer W to both P and O will exercise privacy under a policy of W (since they are $< \theta^W$), but hope to gain both from higher types who also choose privacy and from the disclosures and distortions of even higher types.