American Economic Journal: Microeconomics 2 (May 2010): 191–221
http://www.aeaweb.org/articles.php?doi=10.1257/mic.2.2.191

Public Goods, Social Pressure, and the Choice Between Privacy and Publicity†

By Andrew F. Daughety and Jennifer F. Reinganum*

We model privacy as an agent's choice of action being unobservable to others. An agent derives utility from his action, the aggregate of agents' actions, and other agents' perceptions of his type. If his action is unobservable, he takes his full-information optimal action and is pooled with other types, while if observable, then he distorts it to enhance others' perceptions of him. This increases the public good, but the disutility from distortion is a social cost. When the disutility of distortion is high (low) relative to the marginal utility of the public good, a policy of privacy (publicity) is optimal. (JEL D82, H41)

We develop and explore a new model of the economics of privacy. Individuals with private information about an immutable personal attribute (the agent's "type") engage in actions that have a private benefit, contribute to a public good (bad) and, if observable, may reveal the agent's type resulting in social approval (disapproval) or increased (decreased) future trading opportunities. We consider "privacy of action" as a limitation on the public observability of an individual agent's choice of activity level, 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 how much they contribute to a charity, or over the level of health services they 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 (wherein all types take the same, observable, action), privacy of action means that type-pooling occurs even though

* Daughety: Department of Economics and Law School, Vanderbilt University, VU Station B #351819, 2301 Vanderbilt Place, Nashville, TN 37235–1819 (e-mail: andrew.f.daughety@vanderbilt.edu); Reinganum: Department of Economics and Law School, Vanderbilt University, VU Station B #351819, 2301 Vanderbilt Place, Nashville, TN 37235–1819 (e-mail: jennifer.f.reinganum@vanderbilt.edu). We gratefully acknowledge research support under National Science Foundation grant SES08-14312. 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 audiences at the University of Alberta; the ACLE Workshop, University of Amsterdam; the Summer Workshop on Industrial Organization, 2009, Auckland University; the Georgetown Law Center; the Institut d'Anàlisi Econòmica, Barcelona; the Max Planck Institute for Research on Collective Goods, Bonn; the National Bureau of Economic Research Law and Economics Summer Institute 2008; Northwestern University Law School; PET 09–Galway; Universitat Pompeu Fabra, Barcelona; the TILEC Law and Economics Workshop, Tilburg University; and Vanderbilt University for comments. We thank John Hillas, Rebecca Morton, Chris Snyder, Joel Sobel, Eric Talley, Abraham Wickelgren, Ben Zissimos, and two anonymous referees for helpful comments and suggestions.

† To comment on this article in the online discussion forum, or to view additional materials, visit the articles page at http://www.aeaweb.org/articles.php?doi=10.1257/mic.2.2.191.

different types choose their (different) individual full-information levels of the action. This notion of privacy leads to conditions wherein it can be individually and socially preferred to the alterative, which we refer to as "publicity," where individual choice is influenced by social pressure because actions are publically observable. 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 find that when the activity contributes to the provision of a public good and generates social approval, then a trade-off arises. Publicity induces agents to engage in more activity than they would choose under privacy, which constitutes a social loss. On the other hand, this increased activity increases the amount of the public good provided, which constitutes a social gain. If agents' marginal utility for the public good is sufficiently high, then a policy of publicity is preferred. Otherwise, privacy is optimal. A similar tradeoff arises when the activity contributes to a public bad and generates social disapproval. Publicity induces agents to engage in less activity than they would choose under private information, which constitutes a social loss, but this reduces the extent of the public bad, which constitutes a social gain. When the activity contributes to a public good, but generates social disapproval (or contributes to a public bad but generates social approval), then a policy of privacy of action is preferred to a policy of publicity. As an example, suppose that treatment for substance abuse, if observed, results in an inference that the agent may be an unreliable partner or worker, generating social disapproval or reduced future trading opportunities. Then agents will consume less treatment than they would under a policy of privacy, wherein their actions are rendered unobservable. On the other hand, consuming appropriate treatment contributes to a public good (in addition to its private benefits). In this case, it is socially optimal to provide privacy protection to agents' consumption decisions regarding treatment for substance abuse.[1]

We further consider an intermediate policy of waivable privacy, wherein an agent may choose to make his actions observable. Despite the absence of an exogenous 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 compliance with a pure policy either of privacy or of publicity. This somewhat surprising result occurs because the private and social preferences over which types "ought to" waive privacy are in conflict. The private benefit to waiver is highest for "better" types (because they garner the most social approval), but the social benefit to waiver is lowest for better types (because they distort their actions the most per unit of increment to the public good). However, waivable privacy turns out to be socially preferred to policies of either pure privacy or pure publicity when compliance with those pure policies is sufficiently costly to enforce, and many of the privacy policies in our society are of the waivable form (from the Fifth Amendment's right to silence to opt-outs from privacy restrictions in insurance policies).

In summary, we find three main results. First, privacy of action contributes to welfare, since agents can pursue their full-information optimal actions (instead of either

[1] There may be a social gain from the release of such private information, so that, for example, better labor market matching with firms might occur. We abstract from this consideration in most of our analysis below, but return to this issue in Section V.

conforming to a common action in order to pool or distorting their actions in order to signal). Second, the tradeoff between policies of privacy and publicity involves the social gain from this source versus the social benefit of harnessing the distortionary effects of publicity so as to increase the contribution to the public good. Third, a policy which grants discretion to the agents (waivable privacy) is only socially preferred to the best pure policy if enforcing the pure policy is sufficiently costly.

Finally, we discuss applications of such policies to open-source software development; charitable giving; consumption of health services; conspicuous consumption in a recession; 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.

Section I provides a brief review of some papers on privacy and also some related work on conformity and on pro-social activity. Section II provides the basic analysis of pure policies of privacy of action and publicity of action. We characterize how a policy of privacy or of publicity affects ex ante social welfare, and we indicate why 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. Section III adds the policy of waivable privacy to the welfare comparisons. Section IV provides a number of examples that illustrate our results, while Section V provides a summary and suggestions for future research. An Appendix provides primary derivations. A Web Technical Appendix provides supplementary results.

I. Related Literature

The early literature on the economics of privacy centered on what might be called "privacy of type," wherein privacy rules are used to protect private information against direct observation.[2] The basic idea is that individual agents have immutable 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. This literature also assumed (though typically implicitly) that this private information could be disclosed costlessly and credibly. 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 protection is welfare-reducing in the economy. For example, Posner likens privacy protection to the protection of fraudulent claims.[3]

Subsequent research in information economics, especially in the form of signaling and screening models, assumed that private information about type was prohibitively

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

[3] "The basic point … 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, 406).

costly to disclose. 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 costly screening of the privately-informed agents, or because conditions obtain wherein distortionary signaling by the privately-informed agents ensues.

This paper is related to both 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).[4] 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 IV, 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 Roland Benabou and Jean 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 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.[5]

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 antisocial activities, and those that have

[4] Early papers on disclosure include Sanford J. Grossman (1981) and Paul R. Milgrom (1981). They show that if disclosure is costless and credible, then all types will disclose. An early paper on signaling is A. Michael Spence (1973). Daughety and Reinganum (2008) provide a discussion and unification of the disclosure and signaling models in the context of private information about product quality.

[5] Sera Linardi and Margaret A. 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.

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).[6]

Several recent papers consider providing privacy protection in the context of market transactions between a buyer with private information and a sequence of sellers (for a related survey, see Drew Fudenberg and J. Miguel Villas-Boas 2006). Although the details of the models differ, they obtain similar results. Curtis R. Taylor (2004) considers a market in which a buyer with private information buys sequentially from two sellers. 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 via screening. Alessandro Acquisti and Hal R. Varian (2005) analyze a related model in which the first and second seller are the same firm. Giacomo Calzolari and Alessandro Pavan (2006) use a principal-agent model in which a buyer contracts sequentially with two sellers. 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) with respect to the second seller. These papers address privacy of action in the sense that the buyer's first-period action is unobservable to the second seller when the first seller promises privacy. Buyers, however, are still engaged in a game of asymmetric information with the first (and second) seller, so their decisions are still distorted away from their full-information optimal levels. By contrast, in our model, 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, Benjamin E. Hermalin and Michael L. Katz (2006) consider costlessly waivable privacy in a model in which knowledge of a buyer's type may facilitate price discrimination or improve matching between buyers and sellers. They find that privacy must be mandated rather than waivable in order to have any effect, due to the classic "unraveling" result (see footnote 4). By contrast, we consider a costlessly waivable privacy of *action*. By waiving privacy, the agent can reveal his action, but not his type, costlessly. Although waiving privacy of action is costless, it 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.

[6] Gary S. Becker's (1974) complete information model of social interactions includes intrinsic utility and a status motive for engaging in an activity. B. Douglas Bernheim's (1994) incomplete information model of conformity involves intrinsic utility and esteem 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 adjust their public actions in the same direction (either up or down, 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.

II. 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. Later in this paper, we will modify the model slightly to consider actions that might generate public bads and/or social disapproval. We address three policies: complete privacy (that is, no individual's action is observable by others); complete publicity (that is, each individual's action is observable by all, though each individual's type is not observable directly); and waivable privacy, wherein each agent can choose to make his 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.

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

A. *Model Setup*

There is a continuum of agents of mass M, each with income denoted I.[7] Agent i's utility over the consumption of a composite good c_i and an action g_i is 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 $\theta_i \in [0, \overline{\theta}]$ for all i where $\overline{\theta}$ is finite. Furthermore, assume that each agent observes his type privately, but it is common knowledge that each agent's type is an independently and identically distributed draw from a commonly known distribution, H, with support $[0, \overline{\theta}]$, 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 examples in Section IV, 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 nonnegative real line.[8] Since the agents are identical (except for

[7] Given the assumed quasilinearity of utility, we could allow each agent to have a possibly different income without affecting the results in most of the paper. Only in the case of charitable giving, wherein contributions are related to income (to be discussed in Section IV), would we further require that income be observable. We simplify the discussion and presentation of results by assuming all agents have the same income.

[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, 332–33).

their privately-observed types), from any agent i's point of view, the aggregate of all agents' actions is given by $G \equiv \int E_{\theta_j}(g(\theta_j))\,dj = ME_\theta(g(\theta))$, where $g(\theta)$ is agent i's conjecture of the action of every other agent as a function of that agent's realized type. Since agent i is of measure zero, his contribution to the aggregate is negligible, and his optimal action does not depend upon others' choices. The value of the public good to agent i is given by αG. The nonnegative parameter α represents the marginal utility of the public good.

We further assume that overall utility also depends upon the esteem (social approval) in which agent i is held by others. Note that esteem might be a personal consumption value, or it might reflect continuation values from future trading opportunities.[9] 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.[10] 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.

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 $c_i + pg_i = I$. 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, we assume that 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.

Our model of agent i's choice problem, after substituting for the numeraire 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,[11] denoted as $V_i(g_i, \theta_i, \tilde{\theta}_i, G)$, under the specified rule of privacy (that is, complete privacy, complete publicity, or waivable privacy), where

$$(1) \qquad V_i(g_i, \theta_i, \tilde{\theta}_i, G) \equiv \gamma g_i - (g_i - \theta_i)^2/2 + \alpha G + \beta\tilde{\theta}_i + I - pg_i.$$

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). Note that, with a continuum of agents, no

[9] The related concept of norms is a popular topic in the law literature. Richard H. McAdams (1997) notes that norms induce conformity by denying esteem to those who violate the norm. This is similar in notion to Bernheim's (1994) discussion of conformity. McAdams (1997) further argues that privacy can make norms unenforceable by making violations of the norm undetectable. Thus, when the norm is itself inefficient, privacy is beneficial.

[10] In Section IIE, we allow α to be negative, so as to consider public bads. We also allow β to be negative to reflect social disapproval. The details of the analysis for $\beta < 0$ are in the Appendix. In the conclusion, we briefly discuss the implications of making this term nonlinear in $\tilde{\theta}_i$.

[11] In this formulation, every type receives positive esteem. Alternatively, one might specify 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.

inference can be drawn about any other agent's action based solely on observing the total amount of the public good.

B. *Privacy of Action*

Assume that the society of agents operates under a policy of privacy of action, so that each agent's perceived type is the average of the group, $\mu = \int_{\mathcal{A}} \theta h(\theta) d\theta$, where $\mathcal{A} = [0, \overline{\theta}]$. We use a superscript "$P$" to denote the actions under this regime. Since V_i is quasilinear (and thus the optimal action is independent of agent i's income) and the term G 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 $V_i(g_i, \theta_i), \mu, G)$, which again reflects the symmetry of the agents. Since V_i is strictly concave in g_i, the optimal solution for agent i is $g^P(\theta_i) \equiv \gamma - p + \theta_i$. In order to simplify the notation, let $g_{min} \equiv \gamma - p$, and assume that $\gamma > p$. Thus, agent i's equilibrium level of action under a policy of privacy is

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

That is, the agent will choose a positive amount of the action, and that amount is increasing in type. Finally, note that the agent's optimal value function, $V_i(g_i^P(\theta_i), \mu, G)$, is 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 alone (wherein an agent's type is not observable, but could be inferred from observable behavior).

C. *Publicity of Action*

Consider the same society, but now operating under a policy of publicity (or, for notational convenience, "O" for "observable"). In what follows, we will characterize a separating equilibrium. Publicity assures that actions are observable while separation assures that each action allows inference of the type that would take that action in equilibrium. Although other equilibria exist, we focus on the separating equilibrium because it provides the greatest contrast (in terms of information revelation) with a policy of privacy.[12] Since agents are identical in all observable

[12] This problem satisfies the conditions of Garey Ramey (1996), implying that the separating equilibrium is the unique perfect Bayesian equilibrium satisfying condition D1 (In-Koo Cho and David M. Kreps 1987). However, refinement is controversial, and not the focus of our inquiry. An analysis based on pooling equilibria is

aspects, beliefs about which type will have chosen any particular observed action are assumed to be 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 $B(g_i) \in [0, \overline{\theta}]$. Agent i choosing observable action level g_i, with true type θ_i and perceived type $\tilde{\theta}_i = B(g_i)$ obtains utility $V_i(g_i, \theta_i, B(g_i), G)$. 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' actions. While G contains conjectures about the (expected) actions of the other agents, this term is merely a constant in agent i's objective function. Thus, the symmetric 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 correctly infer 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 symmetric separating perfect Bayesian equilibrium consists of the action function,* $g^O(\cdot)$, *and beliefs,* $B^O(\cdot)$, *such that for all agents and for all* $\theta_i \in [0, \overline{\theta}]$:

(a) $$V_i(g^O(\theta_i), \theta_i, \theta_i, G) \geq \max_g V_i(g, \theta_i, B^O(g), G)$$

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

The structure of the payoff function allows us to use George J. Mailath's (1987) conditions that guarantee and characterize a unique separating equilibrium outcome. Since $\beta > 0$, type $\theta = 0$ is the "weakest" type. This type need not alter its action to be identified, yielding the boundary condition $g^O(0) = g^P(0)$.

PROPOSITION 1: *The Unique Symmetric Separating Equilibrium Outcome under Publicity.*

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

- *The increasing solution to the following equation is* $g^O(\theta_i)$:
  $g^O(\theta_i) = g_{min} + \theta_i + \beta(1 - \exp[-(g^O(\theta_i) - g_{min})/\beta])$ *for all* $\theta_i \in [0, \overline{\theta}]$.

- *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 $g \in [g_{min}, g^O(\overline{\theta})]$, *with out-of-equilibrium beliefs:*
$B^O(g) = 0$ *for* $g < g_{min}$, *and* $B^O(g) = \overline{\theta}$ *for* $g > g^O(\overline{\theta})$.

---

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.

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

(iii) *G in the publicity equilibrium strictly exceeds G in the privacy equilibrium:* $G^O > G^P$.

Part (i) follows from Mailath (1987); part (ii) follows from the boundary condition and the fact that $g^O(\theta_i)$ is increasing; and, therefore, part (iii) holds by the definitions of G^O and G^P.

An informative illustration of the relationship between $g^O(\theta_i)$ and $g^P(\theta_i)$ is shown in Figure 1, 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 function $g^O(\theta_i)$ is increasing and concave. It starts at the same intercept and lies everywhere above the $g^P(\theta_i)$-line. Moreover, the slope of $g^O(\theta_i)$ is greater than one (see the Web Technical Appendix). Therefore, the vertical distance between $g^O(\theta_i)$ and $g^P(\theta_i)$ is increasing in θ_i, indicating that (under a policy of publicity) higher types inflate more to separate from lower types who would otherwise mimic the higher types so as to garner an increase in esteem. That is, the "action differential due to signaling," $\delta(\theta_i; \beta) \equiv g^O(\theta_i) - g^P(\theta_i) = \beta\big(1 - \exp[-(g^O(\theta_i) - g_{min})/\beta]\big)$, is increasing in θ_i.

Again, due to the envelope theorem, agent i's optimal value function, $V_i\big(g^O(\theta_i), \theta_i, B^O\big(g^O(\theta_i)\big), G^O\big)$ is readily shown to be increasing in θ_i. It is clear that neither g^O nor δ depend upon α or M. With a little more work, one can show that $\partial g^O(\theta_i)/\partial \gamma = 1$, $\partial g^O(\theta_i)/\partial p = -1$, and $\partial g^O(\theta_i)/\partial \beta > 0$. Finally, for future use, we note that the action differential $\delta(\theta_i; \beta)$ is independent of γ and p, and increasing in β. That is, an increase in the importance to the agent of esteem results in an increased action differential in order to signal type.

D. *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 choose depends upon an interesting trade-off that involves the importance of the public good, α, and the disutility from the action differential $\delta(\theta_i; \beta)$. To be more precise, we define the ex ante expected social payoff from a policy of publicity for a representative agent i, denoted as SW^O, as $E_\theta[V_i\big(g^O(\theta_i), \theta_i, B^O\big(g^O(\theta_i)\big), G^O\big)]$, where the expectation is taken over the possible types of the representative agent, and G^O is the aggregate level of the public good provided by all agents when operating under a policy of publicity. Thus,

$$\begin{aligned} SW^O &= \gamma E_\theta(g^O(\theta)) - (\tfrac{1}{2})E_\theta\big((g^O(\theta) - \theta)^2\big) + \alpha M E_\theta\big(g^O(\theta)\big) + \beta\mu \\ &\quad + I - pE_\theta\big(g^O(\theta)\big) \\ &= (g_{min} + \alpha M)E_\theta\big(g^O(\theta)\big) - (\tfrac{1}{2})E_\theta\big((g_{min} + \delta(\theta; \beta))^2\big) + \beta\mu + I. \end{aligned}$$

Similarly, we define the ex ante expected social payoff from a policy of privacy for a representative agent i, denoted as SW^P, as $E_\theta[V_i\big(g^P(\theta_i), \theta_i, \mu, G^P\big)]$, where G^P is

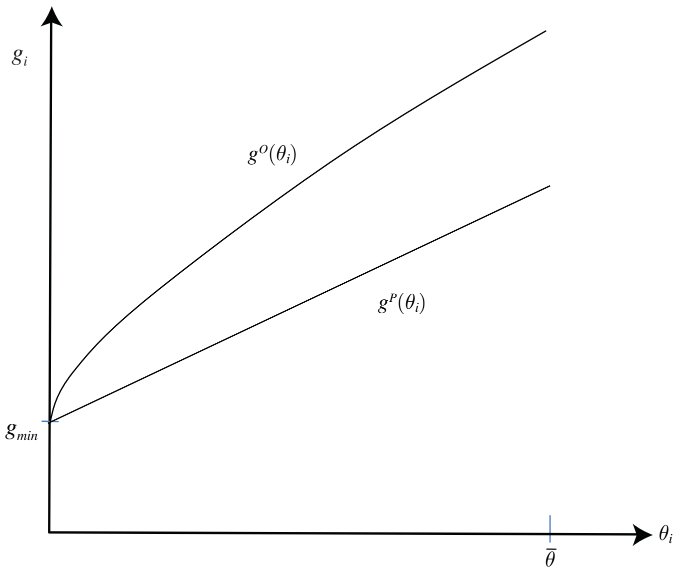

FIGURE 1. COMPARISONS OF ACTION FUNCTIONS

the aggregate level of the public good provided by all agents when operating under a policy of privacy. Thus,

$$SW^P = (g_{min} + \alpha M)E_\theta\big(g^P(\theta)\big) - (\tfrac{1}{2})(g_{min})^2 + \beta\mu + I.$$

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

$$\Phi^{PO}(\alpha) = (\tfrac{1}{2})E(\delta^2) - \alpha M E(\delta), \tag{3}$$

where, for readability, we let $E(\delta^2)$ denote $E_\theta\big((\delta(\theta;\,\beta))^2\big)$, the second moment of $\delta(\theta;\,\beta)$; and let $E(\delta)$ denote $E_\theta\big(\delta(\theta;\,\beta)\big)$, 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 $\Phi^{PO}(\alpha)$.[13]

The sign of the right-hand-side of equation (3) is influenced by the marginal utility of the public good, α; the size of the society in question, M; and the distribution of $\delta(\theta;\,\beta)$, which is influenced by the distribution of θ, H, and the esteem parameter, β. First, since $\Phi^{PO}(\alpha)$ is linear and decreasing in α, it is greater than, equal to, or less than zero as the marginal utility of the public good, α, is less than, equal to, or

[13] If an agent's continuation value, $\beta\tilde{\theta}_i$, were a more general function of his own or all agents' types, then there might be an additional social value (loss) associated with revelation of agents' types and, hence, with publicity. This would increase (decrease) the range of parameters for which publicity is an ex ante optimal policy, but the essential trade-off identified in this paper would remain. We briefly discuss the implications of such nonlinearity in Section V.

greater than $E(\delta^2)/[2ME(\delta)]$. Let $\alpha^{PO} \equiv E(\delta^2)/[2ME(\delta)]$ denote the value of α that yields social indifference between P and O.[14] That is, for values of $\alpha < \alpha^{PO}$, privacy is socially preferred to publicity because the 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 value of the public good is sufficiently high that each agent would prefer to bear the expected disutility from 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. Thus, unless the public good is of sufficient importance to the agents, the ex ante preferred policy is privacy despite the public-goods aspect to the actions of the agents. This reflects the costly effects on the agents of the action differential induced by the esteem term.[15] This 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 $\alpha > \alpha^{PO}$.

Second, holding all else constant, increasing the size of the society, M, decreases α^{PO}, thereby increasing the set of α-values consistent with publicity being ex ante socially preferred to privacy. Third, 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 rearrangement of equation (3) reveals that a mean-preserving spread of H^{δ} results in an increased social preference for privacy. To see this, observe that equation (3) can be written as

$$\Phi^{PO}(\alpha) = \left(\frac{1}{2}\right)\left[E(\delta^2) - \big(E(\delta)\big)^2\right] + E(\delta)[E(\delta)/2 - \alpha M].$$

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 and M). 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}.

[14] 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 $\beta(0) > 0$, privacy is still preferred for sufficiently low α. If $d\beta/d\alpha$ is sufficiently small, there is still a unique value of α beyond which a policy of publicity is preferred. If $d\beta/d\alpha$ 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] 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 be preferred (see Section IIE).

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 Web 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 $E(\delta)$. Thus, since α^{PO} is proportional to $E(\delta^2)/E(\delta)$, the parameter β will potentially affect α^{PO}, but in a nonobvious manner, since there is no inherited property for ratios. 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.[16] First, assuming that H is the uniform distribution, we have computed the necessary and sufficient condition for $\partial(E(\delta^2)/E(\delta))/\partial\beta$ to be positive (and therefore, for $\partial\alpha^{PO}/\partial\beta > 0$) for a broad range of positive values of β and found it 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 Web Technical Appendix for details. The computations suggest that, for a given density, increasing β shifts $\Phi^{PO}(\alpha)$ up, thus increasing α^{PO}. That is, a larger set of α-values are associated 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.[17] That is, 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 inflate action) and common knowledge that realized types are more likely to be high-θ substitutes for the need for individuals to inflate their actions 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.*

E. *Results for Other Sign Patterns for α and β*

In the foregoing, we assumed that $\alpha > 0$ 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 action differential is downward. Every type (except the highest, which is now viewed as the weakest type) chooses a lower level of the action when it is observable than when it remains private. If $\alpha < 0$ as well, so that the activity generates a public bad and (if observable) social disapproval, then the

[16] Calculations were produced using Mathematica 6, the notebook is available upon request.

[17] The table in the Web Technical Appendix suggests that no such generalization to mean-preserving spreads of H is possible. See the results for the Uniform and Middle Triangle Densities Section C.

earlier tradeoff arises in a similar fashion. If the marginal disutility of the public bad is sufficiently great, then publicity is socially preferred to privacy. This is sometimes referred to as "shaming."

If $\alpha > 0$, but $\beta < 0$, then the downward action differential reduces both the agent's intrinsic utility and his contribution to the public good. Here, a policy of privacy is always socially preferred to a policy of publicity. Similarly, when $\alpha < 0$, but $\beta > 0$, the same result obtains since actions are distorted upward, leading to a greater public bad. Again, privacy is always socially preferred to publicity. We consider examples of all four possible sign patterns in Section IV.

F. *Interim Comparisons of Privacy and Publicity*

We now compare *interim* (that is, type-specific) utility payoffs for each individual under a policy of privacy versus a policy of publicity. Our primary result will be a proposition that summarizes *interim* preferences over privacy and publicity. Moreover, we find that whatever the ex ante policy adopted, reasonable conditions exist such that the median type will *interim*-prefer the opposite policy. This leads to the next section on 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, denoted $\Gamma^{PO}(\theta_i, \alpha)$, for $\theta_i \in [0, \overline{\theta}]$, where

$$\Gamma^{PO}(\theta_i, \alpha) \equiv V_i\big(g^P(\theta_i), \theta_i, \mu, G^P\big) - V_i\big(g^O(\theta_i), \theta_i, \theta_i, G^O\big). \tag{4}$$

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 $g^P(\theta)$ and $g^O(\theta)$ into equation (4) and simplifying yields

$$\Gamma^{PO}(\theta_i, \alpha) = \beta(\mu - \theta_i) + \big(\delta(\theta_i; \beta)\big)^2/2 - \alpha ME(\delta). \tag{5}$$

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 action differential squared) is strictly positive when θ_i is positive. Finally, as before, the term $\alpha ME(\delta)$ is the individual's gain in utility from the public good if all other agents' actions are public rather than private. It is straightforward to show that $\Gamma^{PO}(\theta_i, \alpha)$ is monotonically decreasing in θ_i.

Taking expectations of both sides of equation (5), we get that $E\big(\Gamma^{PO}(\theta, \alpha)\big) = \Phi^{PO}(\alpha)$. 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. Let $\overline{\alpha}^{PO}$ be defined by $\Gamma^{PO}(0, \overline{\alpha}^{PO}) = 0$; $\overline{\alpha}^{PO}$ is clearly positive. Furthermore, if $\mu \geq \overline{\theta} - \big(\delta(\overline{\theta}; \beta)\big)^2/2\beta$, then let $\underline{\alpha}^{PO}$ be defined by $\Gamma^{PO}(\overline{\theta}, \underline{\alpha}^{PO}) = 0$; $\underline{\alpha}^{PO}$ is nonnegative. If $\mu < \overline{\theta} - \big(\delta(\overline{\theta}; \beta)\big)^2/2\beta$, then such a value does not exist, so for convenience in the proposition, define $\underline{\alpha}^{PO}$ as zero. This *interim* preference for privacy is summarized in the following proposition (details of the derivation are provided in the Web Technical Appendix).

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

(i) *If* $\mu \geq \overline{\theta} - (\delta(\overline{\theta};\beta))^2/2\beta$, then $0 \leq \underline{\alpha}^{PO} < \overline{\alpha}^{PO}$, *and:*
- *all types interim-prefer P to O for any* $\alpha \leq \underline{\alpha}^{PO}$;
- *all types interim-prefer O to P for any* $\alpha \geq \overline{\alpha}^{PO}$;
- *lower types prefer P to O and higher types prefer O to P when* $\alpha \in (\underline{\alpha}^{PO}, \overline{\alpha}^{PO})$.

(ii) *If* $\mu < \overline{\theta} - (\delta(\overline{\theta};\beta))^2/2\beta$, *then:*
- *not all types interim-prefer P to O for any* $\alpha \geq 0$;
- *all types interim-prefer O to P for any* $\alpha \geq \overline{\alpha}^{PO}$;
- *lower types prefer P to O and higher types prefer O to P when* $\alpha \in [0, \overline{\alpha}^{PO})$.

(iii) *Furthermore, the* α*-value such that society is ex ante indifferent between P and O,* α^{PO}*, is always in this interval:* $\alpha^{PO} \in (\underline{\alpha}^{PO}, \overline{\alpha}^{PO})$. *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. For a detailed discussion, see the Web Technical Appendix. Further, notice that $\mu \geq \overline{\theta} - (\delta(\overline{\theta};\beta))^2/2\beta$ if and only if $(\delta(\overline{\theta};\beta))^2 \geq 2\beta(\overline{\theta} - \mu)$. Since in a separating equilibrium, the action differential is independent of the shape of the distribution (and only depends upon the support), then for distributions H with sufficiently high mean μ, the inequality above will hold. Thus, for a sufficiently right-weighted density h, part (i) of Proposition 4 will hold. Can the reverse hold? That is, will part (ii) sometimes be the relevant case? Since $\delta(\overline{\theta};\beta) < \beta$, then a sufficient condition for part (ii) to hold is that $2\beta(\overline{\theta} - \mu) > \beta^2$. Alternatively put, if $\overline{\theta} > \mu + \beta/2$, then part (ii) will hold. Thus, a sufficiently left-weighted density h will produce this result when $\overline{\theta} > \beta/2$.

III. Waivable Privacy

As noted above, some types will *interim*-prefer to change from the ex ante optimal policy, which raises two issues. First, even if enforcing the ex ante optimal policy of either privacy or publicity is costless, it is conceivable that allowing individual choice at the *interim* stage might be socially preferred to a pure policy. Thus, we consider a third policy wherein privacy is waivable. That is, we will 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 of his action. Second, if it is costly to enforce a pure policy of either privacy or publicity, say at a per-person cost of $k > 0$, then waiver's social value would be further increased.

Without a cost of disclosure, one might expect that all types would disclose, in keeping with the unraveling result mentioned in Section I. That is, one might readily expect that all types would 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 in which 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 is higher 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 in which only some types waive privacy can exist. Here, we characterize such an equilibrium and then compare it with the pure policies discussed earlier.

A. *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 an arbitrary strategy for all other agents that results in an aggregate level of the public good of $\hat{G}$ and arbitrary beliefs following nondisclosure of $\hat{\mu}$. Moreover, suppose that all agents maintain the separating beliefs $B^O(g)$ if the agent chooses O and takes observable action g. Then it is clear that agent i's payoff under P, $V_i(g_i, \theta_i, \hat{\mu}, \hat{G})$, is maximized at $g^P(\theta_i)$, while agent i's payoff under O, $V_i(g_i, \theta_i, B^O(g_i), \hat{G})$, is maximized at $g^O(\theta_i)$, since $\hat{\mu}$ and $\hat{G}$ do not affect the optimal solutions. The net value of privacy over publicity for agent i of type θ_i is given by:

$$V_i(g^P(\theta_i), \theta_i, \hat{\mu}, \hat{G}) - V_i\big(g^O(\theta_i), \theta_i, B^O(g^O(\theta_i)), \hat{G}\big).$$

It is straightforward to show that the derivative of this net value with respect to θ_i is $g^P(\theta_i) - g^O(\theta_i) < 0$ for $\theta_i > 0$. Thus, lower values of θ_i are associated with higher net values of privacy, regardless of the form of other agents' strategies or the beliefs following nondisclosure.

Consequently, in what follows, we will consider strategies wherein low θ-types choose P, while high θ-types choose O. Denote, for any $\hat{\theta} \in [0, \overline{\theta}]$, the *waivable-privacy strategy* for agent j as $(W, g^W(\theta; \hat{\theta}))$, where $(W, g^W(\theta; \hat{\theta})) = (P, g^P(\theta_j))$ if $\theta_j \in [0, \hat{\theta})$ and $(W, g^W(\theta; \hat{\theta})) = (O, g^O(\theta_j))$ if $\theta_j \in [\hat{\theta}, \overline{\theta}]$. The marginal type $\hat{\theta}$ thereby characterizes a cutoff rule for every agent $j \neq i$. Then the perceived type for any agent who chooses P is given by the conditional mean, $\mu(\hat{\theta}) \equiv \int_{\mathcal{J}} th(t)dt/H(\hat{\theta})$, with $\mathcal{J} = [0, \hat{\theta})$. Let $G^W(\hat{\theta}) \equiv ME_\theta(g^W(\theta; \hat{\theta}))$ denote the aggregate amount of the public good. The net value of privacy over publicity for agent i of type θ_i, given that all other agents use the strategy and beliefs specified above, is

$$\text{(6)} \qquad \Delta(\theta_i; \hat{\theta}) \equiv V_i(g^P(\theta_i), \theta_i, \mu(\hat{\theta}), G^W(\hat{\theta})) - V_i\big(g^O(\theta_i), \theta_i, B^O(g^O(\theta_i)), G^W(\hat{\theta})\big).$$

As above, this net value is decreasing in θ_i since $\partial\Delta(\theta_i; \hat{\theta})/\partial\theta_i = g^P(\theta_i) - g^O(\theta_i) < 0$ for $\theta_i > 0$.

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 given the separating beliefs B^O.

PROPOSITION 5: *Symmetric Equilibria under Waivable Privacy.*

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

(ii) *If $\mu \geq \overline{\theta} - (\delta(\overline{\theta};\ \beta))^2/2\beta$, 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 $g^P(\theta)$. This equilibrium is supported by the following beliefs: if the agent chooses P, then $\tilde{\theta} = \mu$; if the agent chooses (O, g), then $\tilde{\theta} = B^O(g)$. A sufficient condition for this to be the only equilibrium involving privacy is that the conditional mean $\mu(\theta)$ is concave in θ.*

(iii) *If $\mu < \overline{\theta} - (\delta(\overline{\theta};\ \beta))^2/2\beta$, then a <u>no-waiver equilibrium does not exist</u>, but at least one partial-waiver (that is, interior) equilibrium does exist, where $\theta^W \in (0, \overline{\theta})$ solves $\Delta(\theta^W;\ \theta^W) = 0$. This equilibrium is supported by the following beliefs: if the agent chooses P, then $\tilde{\theta} = \mu(\theta^W)$; if the agent chooses (O, g), then $\tilde{\theta} = B^O(g)$. A sufficient condition for this interior equilibrium to be the only equilibrium involving privacy is that the conditional mean $\mu(\theta)$ is 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 concave[18]), the function $\Delta(\theta;\ \theta)$ is concave, so that this second type of equilibrium is unique. Therefore, in what follows, we assume that $\mu(\theta)$ is concave. Figure 2 illustrates $g^W(\theta;\ \theta^W)$ for the partial waiver equilibrium. The base curves are those shown in Figure 1 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 in private, but being believed to be the mean type in $[0, \theta^W)$, and distorting his public action to obtain the esteem associated with his true type. Thus, equilibrium action choice is always made along the solid portion of the curves.

We have used the full-publicity equilibrium beliefs to support the partial-waiver equilibrium. That is, actions that are disclosed, and that are in the interval $[g^O(\theta^W), g^O(\overline{\theta})]$, are taken to be from the corresponding type in the interval $[\theta^W, \overline{\theta}]$ given by $B^O(g)$. What if, however, one observes an action g that is less than $g^O(\theta^W)$? This is now an out-of-equilibrium action choice, which has been made public. This could be due to one of two possible sources of error: perhaps the agent

[18] 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, then $\mu''(\theta) \leq 0$ for all θ. Thus, this sufficient condition appears to hold for a highly relevant class of distributions.

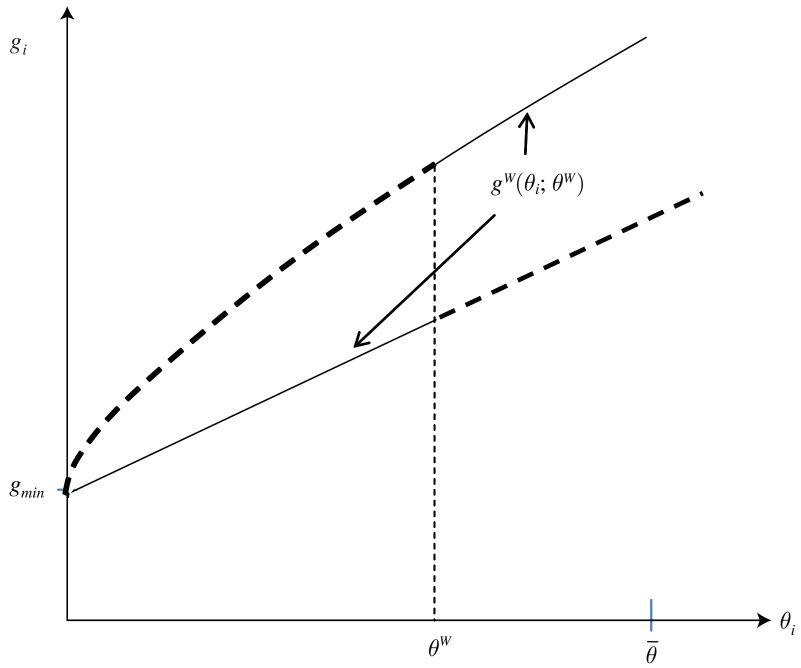

Figure 2. Action Function in Partial Waiver Equilibrium

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 since there is always a second equilibrium at $\theta^W = 0$, perhaps there was a coordination failure in which 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 2 to the inferred value of θ provides the needed disincentive to keep types from strategically engaging in such behavior.

Although disclosure of action is costless, the jump in the action function reflects the endogenous cost of disclosing type that accompanies choosing a public, rather than a private, action and deriving any esteem based on this choice (thereby causing the positive action differential).

B. *Ex Ante Comparisons of Waivable Privacy with Policies of Privacy and of Publicity*

Recall from Section IID the computed versions of the ex ante expected social payoff associated with P (SW^P) and O (SW^O). In a similar fashion, the ex ante expected social payoff associated with waivable privacy, denoted as SW^W, is[19]

$$SW^W = (g_{min} + \alpha M)E_\theta\big(g^W(\theta; \theta^W)\big) - \left(\frac{1}{2}\right)E_\theta\Big(\big(g^W(\theta; \theta^W) - \theta\big)^2\Big) + \beta\mu + I.$$

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

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

$$\Phi^{PW}(\alpha) = SW^{P} - SW^{W} = \left(\frac{1}{2}\right)E^{W}(\delta^2) - \alpha M E^{W}(\delta).\text{[20]} \tag{7}$$

Analogous to before, if $\alpha = \alpha^{PW} \equiv E^{W}(\delta^2)/(2ME^{W}(\delta))$, 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 waivable privacy, while if $\alpha > \alpha^{PW}$, then 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

$$\Phi^{WO}(\alpha) = SW^{W} - SW^{O} = \left(\frac{1}{2}\right)[E(\delta^2) - E^{W}(\delta^2)] - \alpha M[E(\delta) - E^{W}(\delta)]. \tag{8}$$

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

Note that the equations providing $\Phi^{PO}(\alpha)$, $\Phi^{PW}(\alpha)$, and $\Phi^{WO}(\alpha)$ (that is, equations (3), (7), and (8), respectively) are all of the same form, except for the terms having to do with the first and second moments of the action differential $\delta(\theta;\beta)$. 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 $[0, \overline{\theta}]$), 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 $[\theta^{W}, \overline{\theta}]$, 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.*

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

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

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

[20] $E^{W}(\delta^2) \equiv \int(\delta(\theta;\beta))^2 h(\theta)d\theta$ and $E^{W}(\delta) \equiv \int\delta(\theta;\beta)h(\theta)d\theta$, with both integrals on $[\theta^{W}, \overline{\theta}]$.

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

(v) *If the social cost of enforcing publicity is* $k > |\Phi^{PO}(\alpha^{PW})|$, *and waivable privacy is costless to enforce, 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. To see what is going on, consider item (iii) of Proposition 6. Here, waiver (W) is second best when $\alpha < \alpha^{WO}$ because O performs so poorly vis-à-vis P, and W has some of the advantages of the P-policy, since some types choose privacy. This is similarly true when $\alpha > \alpha^{PW}$. In this case, the P-policy performs very poorly in comparison with the O-policy, and the W-policy has some of the virtues of the O-policy, since some types do choose to make their actions public. The reason that the W-policy is third best in item (ii) of Proposition 6 is that now the P- and O-policies perform more closely to one another, but the negative attributes of the W-policy become more important: the wrong types are choosing to reveal their actions (the higher types, who incur a higher action differential per unit of public good generated); and the marginal type is being "pressured" in equilibrium to reveal his action when he prefers privacy (but not at the cost of being lumped in with the lower types). The rankings in Proposition 6(ii) and (iii) are shown in Figure 3, 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 $H(\theta)$ is the Uniform distribution on $[0, 1]$ and $\beta = 1$.

Figure 3 also helps us illustrate the effect of enforcement costs, as discussed in items (iv) and (v) 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 down. If that cost exceeds $\Phi^{PW}(\alpha^{WO})$, then the crossing point of the *PO* and *PW* Φ-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.

The foregoing may explain why, for example, a pure policy of privacy is rarely observed (though we will discuss two examples in Section IV). 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 (such as esteem, or future trading opportunities). 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 Fifth 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

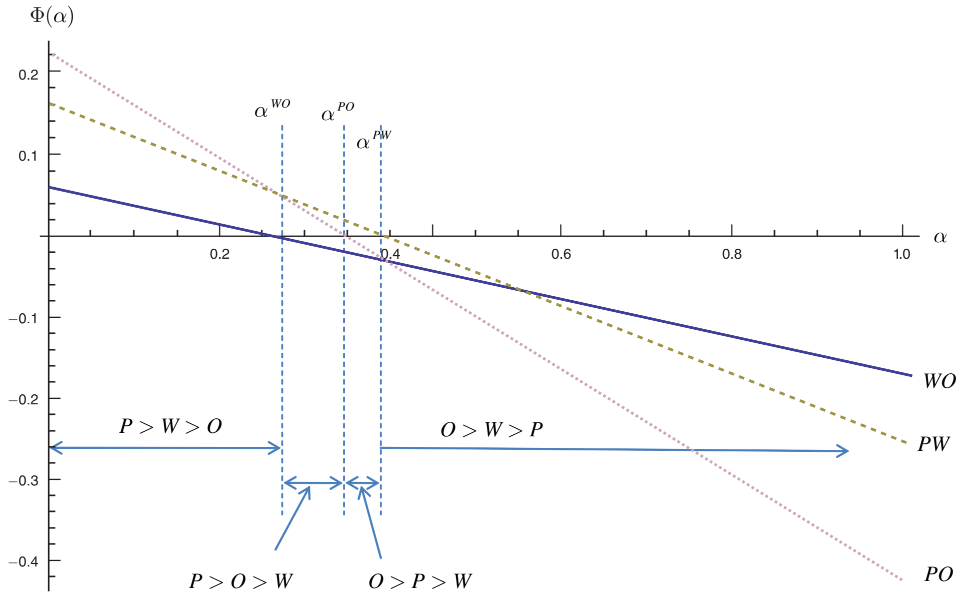

FIGURE 3. EX ANTE SOCIAL ORDERING OVER *P*, *O*, AND *W* POLICIES

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).

IV. 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. Recall that, from Section IIE, 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.

A. *Examples Wherein* $\alpha > 0$ *and* $\beta > 0$

Open-Source Software Development.—In open-source software development, independent programmers contribute toward the improvement of a program that is freely available 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). Josh Lerner and Tirole (2002, 213) argue that delayed benefits (such as possible job opportunities, future access to venture capital, or ego gratification due to peer recognition, presumably all based on others' inferences about talent) accrue to programmers. They refer to these delayed benefits collectively as the "signaling incentive," and provide empirical evidence regarding benefits accrued subsequently by 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 free, 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 (James Andreoni 1989), and the desire to signal some attribute to acquire status (William T. Harbaugh 1998; Amihai Glazer and Kai Konrad 1996).[21] 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 increase associated with public giving, and this can result in too much of the public good. However, they do not include intrinsic utility and do not evaluate welfare under alternative policies.

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. Assume that agent i derives intrinsic utility ("warm glow") from giving according to the utility function $\gamma g_i - (g_i - \theta_i I)^2/2$, where g_i denotes agent i's gift, and θ_i reflects agent i's generosity.[22] 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, $\alpha > 0$ and $\beta > 0$. It is straightforward to show that agent i's gift under a policy of privacy is $g^P(\theta_i; I) = g_{min} + \theta_i I$ for all $\theta_i \in [0, \overline{\theta}]$, while his gift under a policy of publicity is given implicitly by $g^O(\theta_i; I) = g_{min} + \theta_i I + (\beta/I)(1 - \exp[-I(g^O(\theta_i; I) - g_{min})/\beta])$. 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

[21] There is a literature on fundraising 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 (Lise Vesterlund 2003; Jan Potters, Martin Sefton, and Vesterlund 2005); there may be complementarities in giving (Richard Romano and Huseyin Yildirim 2001); and this may solve coordination problems when a fixed total contribution is required (James Andreoni 1998; Leslie M. Marx and Steven A. Matthews 2000). Our model does not involve any of these effects, so the timing of contributions is irrelevant.

[22] If income varies over agents but is observable, then the extension to incorporate this is immediate (replacing I by I_i everywhere).

discrete classifications, rather than actual dollar amounts[23]), and an agent is likely to benefit substantially from others' contributions to his 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.[24]

B. *Examples Wherein* $\alpha > 0$ *and* $\beta < 0$

Consumption of Health Care.—A prime example with this pattern is the consumption of health care (perhaps especially the consumption of mental health care or treatment for addiction). 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 needs has a higher marginal intrinsic utility for health care.[25] 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, privacy results in an increased consumption of health care (relative to a policy of publicity), which provides both a private and a public good.

Conspicuous Consumption in Times of Recession.—In the current recession, high levels of consumption are drawing disapproval. As observed in a recent newspaper article:

> "Even some of the very affluent said they were reluctant to be conspicuous in their spending. 'It's disrespectful to the people who don't have much to flaunt your wealth.' " (Shalia Dewan 2009).[26]

In terms of the model, θ_i is a measure of agent i's insensitivity to plight of the less fortunate. A higher θ_i leads to higher consumption g_i, but if such consumption is public (i.e., conspicuous), the agent accrues social disapproval. This will act to reduce consumption below its full-information level. Properly accounting for the

[23] Both Harbaugh (1998) and Luis Rayo (forthcoming) show that schemes in which coarser information is revealed (for example, contributions are grouped within ranges, so that separate pools are created) can further enhance giving.

[24] Giving to either a local symphony or a global relief fund may generate a high intrinsic value for the giver. In both cases, the policy is waivable. One can contribute anonymously to the symphony, and one can post a receipt for a contribution to a global relief fund on the Web, if desired.

[25] 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.

[26] Dewan, Shaila. "One U.S. Recession Casualty: Conspicuous Consumption," *International Herald Tribune*, March 10, 2009; available at http://www.iht.com/articles/2009/03/10/america/10reset.php, accessed March 21, 2009.

presence of the "less fortunate" requires income to possibly vary over the agents, but quasilinearity implies that g_i is independent of income. Thus, such a modification of the model is easily accommodated. As the private virtue of saving can be a public vice in a recession, α is positive while β is negative. In this case, the visibility of consumption may limit the speed of recovery from the recession.

C. *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 Law School's policy of not reporting class rank to potential employers.[27] The law school does not report class rank (which is arguably more informative than grades). Indeed, it provides the student with this information in nonverifiable 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 student *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 positively affect the esteem in which the student is held by peers and his or her future employment options, so $\beta > 0$. This 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.[28]

Inadmissibility of Settlement Offers at Trial.—Another example in which nonwaivable privacy is the prevailing policy is Federal Rule of Evidence 408, which makes the details of failed settlement negotiations inadmissible in court. As shown in Daughety and Reinganum (1995),[29] a plaintiff with private information about his damages has a greater incentive to inflate his 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.

[27] 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.

[28] Michael Ostrovsky and Michael Schwarz (2010) provide other examples in which schools engage in partial suppression of information through (for example) issuing "noisy" transcripts. Their model does not involve endogenous effort on the part of students (and therefore does not consider effort-based externalities among students), and does not examine the social desirability of privacy.

[29] Daughety and Reinganum (1995) use a different payoff structure, but include the same elements: intrinsic utility, a public-bad aspect, and a continuation value based on inferred type.

D. *Examples Wherein $\alpha < 0$ and $\beta < 0$*

When $\alpha < 0$ and $\beta < 0$, then 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.[30]

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 or the risk of a brown-out 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, and compulsory publicity is not outside the realm of possibility. For instance, according to Clive 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."[31] Brenda Goodman (2007) reports that during a recent extended drought in Georgia, a Marietta man's water usage was disclosed by the local American Broadcasting Company affiliate (approximately 14,700 gallons per day, compared to an average in the Atlanta area of about 183 gallons per day).[32] 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

[30] 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.

[31] Thompson, Clive. "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.

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

the United Kingdom. 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).[33]

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 redirect 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 a number of ongoing efforts to disclose and control earmarks and their sponsors (see Megan Suzanne Lynch 2008, for the rules adopted by the 110th Congress).[34]

V. 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: an intrinsic value for the activity involved; esteem (or in some examples, social disapproval); and 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 compliance with a pure policy of privacy (or publicity) is sufficiently costly to enforce, then waivable privacy can be ex ante socially preferred. Otherwise, waivable privacy is never socially preferred. A policy of waivable privacy gives rise to two types of pure-strategy equilibria: one wherein all types choose to reveal their actions (with

[33] "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.

[34] Lynch, Megan Suzanne. "Earmark Disclosure Rules in the Senate: Member and Committee Requirements," CRS Report for Congress, Order Code RS22867, April 29th 2008. Avail. at: http://8vsb.files.wordpress.com/2008/04/earmark-disclosure-rules-in-the-senate.pdf, accessed March 21, 2009.

the concomitant change in action level) and one in which a subset of types (possibly empty) chooses to make their actions public while the rest choose to keep their actions private.

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 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 differential due to publicity will it be optimal to choose publicity.

We have formulated the model at a rather abstract level in order to identify the fundamental trade-off between the expected disutility due to signaling and the increased contribution to the public good that generates the social choice of privacy policy. We believe that this basic model can serve as a useful foundation for models that are more tailored to specific applications, possibly involving finitely many agents with greater strategic interaction, and more detailed modeling of the continuation value. We briefly discuss each of these extensions below.

We have considered a continuum of agents. Adapting the model to a society with a small number of agents will involve two complications that arise from the fact that each agent is now nonnegligible in the aggregate amount of the public good provided. First, having a small number of agents may create information flows even under a policy of privacy of action (knowledge of the aggregate amount of the public good and one's own contribution can be used to update beliefs about others' contributions and hence their types). Second, having a small number of agents may create strategic interactions via substitution effects or complementarities among their actions. This would necessitate the use of a simultaneous signaling model in which each agent's action as a function of his type satisfies not only the incentive compatibility conditions required for separation, but also the best-response condition necessary for a Nash equilibrium.

We have assumed that the continuation value is a linear function of an agent's inferred type. While this is a common specification, it omits another influence that may affect the social preference over privacy versus publicity. Consider the case in which the agents' actions contribute to a public good (that is, $\alpha > 0$) and higher types are accorded greater esteem (that is, $\beta > 0$), and suppose that the continuation value as a function of inferred type is given by $CV(\tilde{\theta})$. Then the ex ante expected continuation value would be $E\big(CV(\theta)\big)$ under a policy of publicity resulting in a separating equilibrium, and it would be $CV\big(E(\theta)\big)$ under a policy of privacy. If $E\big(CV(\theta)\big) > CV\big(E(\theta)\big)$ (as would occur, for instance, if the function $CV(\cdot)$ were convex), then this additional effect would favor publicity. That is, publicity would be the socially optimal policy for somewhat lower values of the marginal utility of the public good (i.e., α^{PO} would be lower); publicity might be optimal even for somewhat negative values of α. On the other hand, if $E\big(CV(\theta)\big) < CV\big(E(\theta)\big)$ (as would occur, for instance, if $CV(\cdot)$ were concave), then this additional effect would favor privacy. That is, privacy would be the socially optimal policy for somewhat higher values of the marginal utility of the public good (i.e., α^{PO} would be higher). Similar results would follow if the continuation values were functions of other agents' types as well (as would occur, for instance, if the continuation game involved matching of agents). Again, inclusion of these effects could move the social indifference point in one direction or the other, depending on the fine details of the continuation value function.

Two further extensions would be to address more precisely the continuation value issue raised above and to allow for a richer range of agent behavior. In the first case, one could marry the analysis in the current paper to a matching model so as to address more carefully the issue of efficiency losses or gains generated by privacy protection. This would endogenize the continuation value to the agents. Second, the current model assumes all agents are fundamentally the same except for a one-dimensional type space. One might want to allow for agents to have multi-dimensional types, so that some agents in the population might entertain engaging in activities that others would eschew: some agents might engage in criminal or reprehensible behavior (for example, driving under the influence or patronizing prostitutes) that others would never consider. The extension to multi-dimensional types would also allow for an examination of situations wherein two agents of very different types take the same observable actions (would reading about the life and times of a famous terrorist be associated with someone interested in understanding the world, or is it a signal of a budding terrorist?). These extensions would provide for an even more nuanced analysis of the social value of policies of privacy, publicity, and waivable privacy.

Appendix

This appendix provides the derivation of the waiver equilibria and a discussion of the separating equilibrium when $\beta < 0$. Proofs of comparative statics results, computational results, and an analysis of *interim* preferences over privacy, publicity, and waivable privacy can be found in the Web Technical Appendix.

A. *Material on Deriving Waiver Equilibria*

Recall the definition of $\Delta(\theta_i; \hat{\theta})$ from the text: $\Delta(\theta_i; \hat{\theta}) \equiv V_i\big(g^P(\theta_i), \theta_i, \mu(\hat{\theta}), G(\hat{\theta})\big) - V_i\big(g^O(\theta_i), \theta_i, B^O(g^O(\theta_i)), G(\hat{\theta})\big)$ 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, $\theta^W = \overline{\theta}$ is an equilibrium if and only if $\Delta(\overline{\theta}; \overline{\theta}) \geq 0$. To see this, note that if all other agents' types choose *P* and $\Delta(\overline{\theta}; \overline{\theta}) \geq 0$, then $\Delta(\theta_i; \overline{\theta}) > 0$ for all $\theta_i < \overline{\theta}$, 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, $\theta^W \in (0, \overline{\theta})$ 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 $\theta^W \in (0, \overline{\theta})$, 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 (incorporating the equilibrium beliefs) 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 IIC 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 $d\Delta(\theta; \theta)/d\theta = \beta\big[\mu'(\theta) - (1 - \exp[-(g^O(\theta) - g_{min})/\beta])\big] > 0$ when evaluated at $\theta = 0$, so that there is at least one more equilibrium. If $\Delta(\overline{\theta}; \overline{\theta}) \geq 0$ 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 IIB. If $\Delta(\theta; \theta)$ is concave, then this is the unique equilibrium involving privacy. Finally, if $\Delta(\overline{\theta}; \overline{\theta}) < 0$ 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 $\mu < \overline{\theta} - (\delta(\overline{\theta}; \beta))^2/2\beta$. Moreover, such an interior equilibrium will be unique if $\Delta(\theta; \theta)$ is concave; a sufficient condition for this to hold is that $\mu''(\theta) \leq 0$ for all θ.

B. *Discussion of the Separating Equilibrium for $\beta < 0$*

All of the arguments given in the text 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 $\overline{\theta}$, so this is the type which need not distort its action to be identified. Consequently, the relevant boundary condition becomes $g^O(\overline{\theta}) = g^P(\overline{\theta}) = g_{min} + \overline{\theta}$ or, in terms of the beliefs, $B^O(g_{min} + \overline{\theta}) = \overline{\theta}$.

As before, this differential equation has a one-parameter family of solutions. Imposing the boundary condition selects a unique solution:

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

Inverting the function $B^O(g)$ results in an equation that defines $g^O(\theta_i)$ implicitly:

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

The separating equilibrium $g^O(\theta_i)$ is still increasing in θ_i; therefore the term in the exponential function is negative, and hence (since β is negative) the third term in this equation is strictly negative for $\theta_i < \overline{\theta}$. Thus, the agent's action is now downward-distorted under publicity, with the greatest action differential 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 $g_{min} \geq -\beta$.

REFERENCES

Acquisti, Alessandro, and Hal R. Varian. 2005. "Conditioning Prices on Purchase History." *Marketing Science*, 24(3): 367–81.

Andreoni, James. 1989. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy*, 97(6): 1447–58.

Andreoni, James. 1998. "Toward a Theory of Charitable Fund-Raising." *Journal of Political Economy*, 106(6): 1186–213.

Becker, Gary S. 1974. "A Theory of Social Interactions." *Journal of Political Economy*, 82(6): 1063–93.

Benabou, Roland, and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review*, 96(5): 1652–78.

Bernheim, B. Douglas. 1994. "A Theory of Conformity." *Journal of Political Economy*, 102(5): 841–77.

Calzolari, Giacomo, and Alessandro Pavan. 2006. "On the Optimality of Privacy in Sequential Contracting." *Journal of Economic Theory*, 130(1): 168–204.

Cho, In-Koo, and David M. Kreps. 1987. "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics*, 102(2): 179–221.

Daughety, Andrew F., and Jennifer F. Reinganum. 1995. "Keeping Society in the Dark: On the Admissibility of Pretrial Negotiations as Evidence in Court." *RAND Journal of Economics*, 26(2): 203–21.

Daughety, Andrew F., and Jennifer F. Reinganum. 2008. "Communicating Quality: A Unified Model of Disclosure and Signaling." *RAND Journal of Economics*, 39(4): 973–89.

Edwards, Donald, Elizabeth Monk-Turner, Steve Poorman, Maria Rushing, Stephen Warren, and Jarita Willie. 2002. "Predictors of Hand-Washing Behavior." *Social Behavior and Personality*, 30(8): 751–56.

Fudenberg, Drew, and Jean Tirole. 1991. *Game Theory*. Cambridge, MA: MIT Press.

Fudenberg, Drew, and J. Miguel Villas-Boas. 2006. "Behavior-Based Price Discrimination and Customer Recognition." In *Economics and Information Systems*, ed. Terrence Hendershott, 377–436. Amsterdam: Elsevier.

Glazer, Amihai, and Kai A. Konrad. 1996. "A Signaling Explanation for Charity." *American Economic Review*, 86(4): 1019–28.

Grossman, Sanford J. 1981. "The Informational Role of Warranties and Private Disclosure About Product Quality." *Journal of Law and Economics*, 24(3): 461–83.

Harbaugh, William T. 1998. "What Do Donations Buy? A Model of Philanthropy Based on Prestige and Warm Glow." *Journal of Public Economics*, 67(2): 269–84.

Hermalin, Benjamin E., and Michael L. Katz. 2006. "Privacy, Property Rights and Efficiency: The Economics of Privacy as Secrecy." *Quantitative Marketing and Economics*, 4(3): 209–39.

Lerner, Josh, and Jean Tirole. 2002. "Some Simple Economics of Open Source." *Journal of Industrial Economics*, 50(2): 197–234.

Linardi, Sera, and Margaret A. McConnell. 2008. "Volunteering and Image Concerns." California Institute of Technology Division of the Humanities and Social Sciences Working Paper 1282.

Mailath, George J. 1987. "Incentive Compatibility in Signaling Games with a Continuum of Types." *Econometrica*, 55(6): 1349–65.

Marx, Leslie M., and Steven A. Matthews. 2000. "Dynamic Voluntary Contribution to a Public Project." *Review of Economic Studies*, 67(2): 327–58.

McAdams, Richard H. 1997. "The Origin, Development, and Regulation of Norms." *Michigan Law Review*, 96(2): 338–433.

Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics*, 12(2): 380–91.

Ostrovsky, Michael, and Michael Schwarz. 2010. "Information Disclosure and Unraveling in Matching Markets." *American Economic Journal: Microeconomics*, 2(2): 34–63.

Posner, Richard A. 1981. "The Economics of Privacy." *American Economic Review*, 71(2): 405–9.

Potters, Jan, Martin Sefton, and Lise Vesterlund. 2005. "After You—Endogenous Sequencing in Voluntary Contribution Games." *Journal of Public Economics*, 89(8): 1399–419.

Ramey, Garey. 1996. "D1 Signaling Equilibria with Multiple Signals and a Continuum of Types." *Journal of Economic Theory*, 69(2): 508–31.

Rayo, Luis. Forthcoming. "Monopolistic Signal Provision." *The B.E. Journal of Theoretical Economics*.

Romano, Richard, and Huseyin Yildirim. 2001. "Why Charities Announce Donations: A Positive Perspective." *Journal of Public Economics*, 81(3): 423–47.

Spence, A. Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics*, 87(3): 355–74.

Stigler, George J. 1980. "An Introduction to Privacy in Economics and Politics." *The Journal of Legal Studies*, 9(4): 623–44.

Taylor, Curtis R. 2004. "Consumer Privacy and the Market for Customer Information." *RAND Journal of Economics*, 35(4): 631–50.

Vesterlund, Lise. 2003. "The Informational Value of Sequential Fundraising." *Journal of Public Economics*, 87(3–4): 627–57.