Privacy and Endogenous Monitoring Choice When Private Information is a Public Good[1]

Stefan Dodds

Department of Economics
Queen's University
99 University Ave.
Kingston, Ontario, Canada
K7L 3N6
email: doddss@qed.econ.queensu.ca
JEL Classification: D82, H41

September 1, 2002

Abstract:
This paper examines why economies endow agents with a degree of personal privacy, even when (a) "no privacy" is ex-post (Pareto) efficient, and (b) a costless monitoring technology exists. A government can provide more of a public good only by identifying "valuable" agents from a population of n. All agents report their type to the government — truthfully or not — unsure if they, or others, are being observed. When n is small, it is shown that increasing monitoring effectiveness can actually lead to ex-post inefficiency. Political equilibria are also characterized, where agents vote to constrain the government's monitoring effectivenes but not its ability to levy penalties or rewards. When n is large, all such equilibria are efficient; however, a utilitarian government may not implement taxes to reward honest reporting, nor impose penalties to punish it, even when these options ensure full revelation. Legislating a "right to privacy", by contrast, is always inefficient.

[1]Comments are very welcome. Thanks to Robin Boadway, Jonathan Pincus, Alasdair Roberts, Ruqu Wang, and seminar participants at Queen's University, 2001 PhD Conference on Business and Economics at the University of Western Australia, and the 2002 Association for Public Economic Theory conference for useful suggestions. All errors are my own.

1 Introduction

The concept of "privacy" — studied extensively by researchers in many disciplines — has received relatively little consideration by economic theorists. In its most widely accepted meaning, privacy refers to the ability of individuals to be alone from the larger society. This "aloneness" may take a number of forms. To be physically left alone, to have a sphere of influence over which one is free from coercion, to have one's personal thoughts, characteristics and behaviour remain unknown to the rest of society — all these are examples of being endowed with privacy. It is widely accepted that individuals desire a degree of privacy, although many explanations have been given as to why this should be the case. In most developed nations, privacy laws govern many issues associated with information gathering, retention and dissemination, for firms and governments as well as for individuals. Most of the popular concern with privacy, however, focuses on the ability of individuals to retain a "private sphere" in the face of intrusion by government agencies and other citizens. Frequently, this means establishing institutions to keep personal information *secret*, in the sense that it cannot be accessed by other groups or individuals in society.

Political, sociological and legal scholars have noted an increasing popular concern with privacy, which has coincided with technological advances in electronic record-keeping, computing power and monitoring technologies. Consider, for example, the advent of DNA testing, video surveillance, electronic databases and so-called Smart-Cards. According to Lyon (2001), these new forms of surveillance are advocated "for efficiency or public order" although their secondary effect is to "destabilize the public/private boundary" between individuals and society. Privacy issues have received attention both in the workplace (Connerley et. al, 1999, Gilliom, 1997) and in the home (Crabb, 1999). However, most emphasis in these fields has dealt with privacy as a "right" or an intrinsic "freedom". For example, Brin (1998) and Garfinkel (2000) argue that the surveillance of behaviour and characteristics should be controlled to preserve democratic objectives. Implicit in their arguments is the view that privacy is an inherently valuable *primary good*, as opposed to a feature of the economic environment which affects welfare via outcomes.

Why might individuals and governments be concerned about observing the private behaviour and personal characteristics of other members in society? For individuals, others' information may be valuable in itself (as in the case of "nosy preferences", where individuals obtain utility from knowing the behaviour or consumption of others), or it may provide strategic benefits (for example, knowing the characteristics of competitors for employment may allow individuals to better tailor their own applications). Governments seek individuals' information for a variety of reasons, usually to improve the effectiveness of public transfer policies (e.g. welfare, UI) and to efficiently allocate government expenditures (e.g. via information gathered by a census).

The existence of some degree of privacy is tacitly acknowledged in the massive literature on adverse selection and moral hazard problems. In an early paper on this issue Posner (1981) conjectured that the institution of worker-applicant privacy hurts the efficiency of labour markets by allowing inferior employees to hide bad characteristics from potential employers. This paper departs from previous work by focusing on the origins of asymmetric information rather than optimal policies in this environment. For instance, a fundamental problem in public finance is to determine optimal policies to implement *when certain information is not observable by the government.* It is rarely if ever asked *why* such information is unobservable, especially when *full information* almost always leads to greater efficiency in the Pareto sense. That is, full information on individual types expands the set of efficient feasible allocations that can be implemented by policy-makers. Consequently, there appears to be a tension between the degree to which individuals may hide their personal information and economic efficiency. Is this always the case? If not, it may be beneficial for a benevolent government *not* to have full information about individuals' types even when it is feasible to do so.

This paper has two goals. The first purpose is normative: to investigate the nature of the trade-off between privacy and welfare, using a simple model where two individuals must simultaneously choose whether to report their true characteristics to a government that monitors all agents with some probability. These characteristics are used directly in the production of a public good. One group in the economy ("high types") have valuable information, while the other group ("low types") do not. When revealing valuable information — whether truthfully or after being observed by the government — is costly, high type agents have an incentive to free-ride on the revealed information of other high types. However, this incentive is tempered by their desire for the public good and inability to directly observe the types of others. It is shown that although full information (zero privacy) and no monitoring (full privacy) can both lead to efficient outcomes, a unique equilibrium can exist for intermediate levels of monitoring which guarantees *inefficiency.* Therefore it is not the case that increased monitoring (reduced privacy) leads to greater welfare.

The second purpose of the paper is positive: to endogenously determine what amount of leeway agents will give the government to observe their types, given the government's endogenous desire to observe. The model outlined above is extended by letting the number of agents become large and adding two prior stages to the reporting game. First, after learning their types, agents are given the power to vote on a binding level of monitoring to which they will be subjected; second, a welfare-maximizing government is given the option to levy fines and set transfers between individuals after they commit to a certain level of monitoring. The objective here is not to replicate an actual political process, but rather to model a situation where governments are constrained (for example, legally) from observing citizens to whatever extent they wish. In this setting, many political equilibria are possible, but two are striking. Even though high type agents have "something to hide", they may vote to eliminate privacy in order to capture all other high types' information and compel the government to provide compensation. And even though low type agents have "nothing to hide", they may vote for less than full privacy and allow some high types to escape being monitored, so as not to subsidize compensation for all high types. Returning to the normative issue, it is shown that if high types ever prefer some monitoring to none, then a "right to privacy" is always inefficient even though voluntary contribution of information is allowed at all times.

The model in this paper borrows elements from several existing literatures in public economics and political economy. The structure of individuals' behaviour is similar to that of the literature on tax evasion (eg. Allingham and Sandmo (1972), Beck, Davis and Jung (2000), Boadway and Sato, (2000), Chander and Wilde (1998)), in which agents submit true or false reports subject to a probability of audit, and receive a reward or penalty based on their honesty or deception. The nature of the monitoring technology is similar to the "tagging technologies" studied by Akerlof (1978) and Parsons (1996), in which direct identification of individual's types improve the efficiency of government transfer schemes to approach full information outcomes. This paper departs from these studies by adding strategic behaviour with other individuals via public good provision, and by assuming costless monitoring. Elsewhere, Konrad (2001) has written specifically on the issue of privacy. He shows that when the government cannot commit ex ante to limit its observation of individuals' characteristics ex post, restricting the government's ability to collect information about individuals may be Pareto-improving. This paper, by contrast, allows individuals to restrict the monitoring available to the government as part of a political equilibrium.

There is also a literature on the desirability or feasibility of targeting individuals as a means to improve economic allocations. Targeting may be too costly relative to its benefits (Besley (1990), Jacquet and van der Linden (2001)), or politically unsupportable (DeDonder and Hindriks (1998)). The relationship between targeting and privacy is summarized by Lindbeck (1988):

> A basic reason why *privacy* is threatened by far-reaching welfare state policy ... is that such policies, if they include high marginal tax and transfer rates, may make honesty towards the state quite "expensive" for the individual because of the gains from giving incorrect information in connection to taxes

> and public benefits ... The government will therefore be induced to build up an elaborate system of information of the lives of individuals and private organizations. (311-312)
>
> Of course there is an alternative to a control state, namely to accept slack in the system, in the sense that a considerable amount of cheating with taxes, breaking of regulations and misuse of benefits are accepted by the authorities ... Obviously, democratic societies tend to opt for some *combination* of control and slack. (314, original emphasis) [1]

The model outlined in the following sections — that of treating private information as an input to a public good — is used to consider the question of what degree of "slack" emerges from the model, and why.

The outline of the paper is as follows. Section 2 describes the model. Section 3 presents some solutions to in a two-agent setting, while Section 4 extends the model to n-agents. The choice of monitoring in a political equilibrium is addressed in Section 5. Section 6 concludes.

2 The basic model

The model used is a variant of a voluntary contribution mechanism. Individuals of two distinct types simultaneously contribute to a government-collected pool of information used to produce a public good. By reporting their types, truthfully or otherwise, they make a claim about their usefulness in producing the public good. However, agents are exposed to a given monitoring technology when reporting their type. The following is an outline of the basic economic setting.

Environment: There are n agents (indexed by $j = 1 \dots n$) and a central agency (government). The agents are divided into two types, "high" and "low", so agents may be of type $i \in \{H, L\}$. Assume that each agent has an ex ante probability of being type H with probability α, and type L with probability $(1 - \alpha)$. When n is large, this implies that the ex post distribution of types is $(H, L; \alpha, 1 - \alpha)$ and $n = n^H + n^L$. Assume that α is common knowledge to all actors in the economy.[2] Each agent is endowed with equal economic resources, Y.

[1] A recent example pertains to the department of human resources in the Canadian government (HRDC), which under public pressure was forced to destroy a longitudinal database containing a large set of personal information on Canadian citizens. Almost simultaneously, the department was criticized for its inefficient allocation of federal grants and poor record-keeping, problems which could ostensibly be remedied through improved targeting of benefits based on personal and firm characteristics. Thus, while individuals were concerned with their own privacy, they were also concerned with inefficiencies that their personal information could help alleviate. (This example was suggested by Alasdair Roberts).

[2] This assumption is maintained for simplicity, but could be relaxed to allow for a richer set of predictions from the model.

Preferences: Each type of agent i, $i \in \{H, L\}$ has preferences over a public good, G, and a composite private good, x: $u^i(G, x_i)$. Assume that u is twice differentiable in each argument, that $u' > 0$ and $u'' < 0$ in each of G and x, and that $u_x^H / u_G^H < u_x^L / u_G^L$ for all (x, G). That is, high types like the public good relatively more than low types.

Public Good: The public good is produced according to the following technology:

$$G = F(\tilde{n}^H) \tag{1}$$

where

$$\tilde{n}^H = (\# \text{ of high types observed by the government}) \text{ , } \tilde{n}^H \leq n^H$$

Let $F'(\cdot) > 0$ and $F''(\cdot) \leq 0$. In words, more of the public good can be produced when more high type agents are identified, but at a declining rate. **Only by identifying high types can the production of G be increased.**[3] Upon being identified, each high type pays a compliance cost, $c \geq 0$, of which a portion β accrues to the government for production of G.

Mechanism with Monitoring: Assume that each agent's type is his or her *private information*, unobservable to other agents or to the government. The mechanism put in place by the government takes the following form:

1. There exists a costless[4] monitoring technology which can identify an agent's true type with known probability $m \in [0, 1]$. (At this stage, assume that the particular technology has been predetermined.) m is known to the government and all agents.
2. Agents simultaneously submit a type report $\hat{H}$ or $\hat{L}$ to the government. This report can be different from the agent's true type, H or L.
3. The government observes true types according to m and checks its observations with the reports. If they agree, the government does nothing. If they disagree, the government levies a penalty $p \geq 0$ on each agent j who misrepresents himself. A high type agent caught cheating pays the compliance cost c as well as the penalty p.
4. Allocations are made and utilities are determined.

[3] This specification allows the amount of public good to be non-zero when no high types are identified: e.g. $G = \bar{G} + f(\tilde{n}^H)$.

[4] The assumption of costlessness is used to simplify the analysis. Costly monitoring is also omitted to eliminate outcomes where privacy is chosen simply because monitoring technologies are prohibitively expensive. However, improvements in technology, especially in the realm of electronic surveillance, suggest that many techniques can gather alot of information about individuals for little cost. e.g. consider hidden video-surveillance, or computer algorithms able to scan email, etc. (see Lyon, 2001).

It is immediate that low types (L) will always report $\hat{L}$ to the government. To see this, note that low types are incapable of adding to the public good, and would only incur a penalty if caught reporting $\hat{H}$.

Monitoring ability is used as a proxy for the level of privacy that agents enjoy: lower monitoring implies greater privacy. Readers may object that when $m > 0$ there is an *a priori* case for the elimination of privacy.[5] Alternatively, m may be considered one measure of "slack" observed in public policy. Monitoring has two opposite incentive effects for high types in this model. First, increased monitoring encourages these individuals to report their true information despite compliance costs, given the threat of punishment if they are found to be misrepresenting themselves. Second, increased monitoring raises the chance that *other* individuals will be forced to submit their true information, raising the incentive to a given individual to withhold *her* true information and save compliance costs.

Note that an observed high type must pay c either voluntarily or if caught cheating. This cost may be seen as a time cost of participating in a survey or study, the inherent cost of relinquishing a piece of personal information, or simply as a monetary fee levied against high type individuals. This definition follows that of Besley (1990). βc is the amount of useful resources — expressed in monetary (or private good) terms — that the government can obtain from each high type agent it can identify.[6] High types reporting untruthfully but caught by the monitoring system are indistinguishable from low types, and so do not incur these costs.

Several examples can provide intuition for this model. These include:

- The public good is to find a cure for a genetic disease. High types carry a predisposition for the disease, which can be identified with a simple DNA test. If discovered, high types must submit to medical study or testing regimen which is costly from their perspective but increases the probability of finding a cure.

- The public good is a reduction in violent crime. High types are law abiding citizens who own guns, but the government proposes a registry to track guns which are resold, stolen, and used in crimes. With access to purchasing records, the government could identify gun owners. Gun owners — if discovered — must fill out

[5]Any *threat* that one's actions are being observed could be construed as an invasion of privacy. In the present context, the government announces the level of monitoring m to those potentially being observed. Coupled with a public scheme for punishing those who misrepresent themselves, this mechanism allows agents to compute their expected payoffs from different reporting strategies.

[6]Implicitly, the amount $(1-\beta)c$ is a deadweight loss to society for a "found" high type. Such a loss is analogous to studies of tax evasion in which costly auditing is needed to identify cheaters (e.g. Beck, Davis & Jung, 2000), although here the costs are borne by monitored agents, not the government. The exact level of β does not affect any qualitative results of the paper.

lengthy forms to register their guns, be subject to inspection, and participate in safety classes.

- The public good is the success of a sports team. High type players like to consume alcohol and stay up late, while low types do not. If high types stay in and don't drink, the team will do better but they will personally suffer a utility cost equal to c from abstaining. The coach monitors players' behaviour after hours to identify these high types.

This model is applicable to other scenarios as well. Governments evaluate projects based on the characteristics of those who stand to principally benefit from them. Academic studies typically require identification of research subjects to increase their efficacy. And government programs can become more efficient as certain population subgroups are identified (as in the tagging and targeting literatures), freeing up resources for other uses.

3 The Model with Two Agents

The purpose of this section is to examine the tradeoff between the degree of privacy enjoyed by agents, and the efficiency of the outcomes generated in the mechanism described above. Because $n = 2$ is small, the ex post distribution of types need not be $(H, L; \alpha, 1-\alpha)$; in general it will be $[(H,H),(H,L),(L,H),(L,L);\alpha^2,\alpha(1-\alpha),(1-\alpha)\alpha,(1-\alpha)^2]$. A relevant question is then to ask, "how much monitoring is needed to ensure an efficient outcome?" In a setting where agents must decide to reveal or withhold their true information, unsure of (a) the information held by other agents and (b) whether they are being monitored or not, the answer to this question is unclear. This section will provide basic intuition for the n agent case.

Given the assumptions of the previous section, and the straightforward behaviour of low types, the set of possible allocations for an agent can be characterized. For every low type agent, this set is $(Y, F(\tilde{n}^H))$, where Y is the agent's endowment. For every high type agent j, the set of possible allocations is given by $(Y - c_j - p_j, F(\tilde{n}^H))$, where $c_j = c$ if agent j truthfully reports $\hat{H}$ — or if j reports $\hat{L}$ and is caught lying — and 0 otherwise. Also, $p_j = p$ if agent j is caught reporting $\hat{L}$ and 0 otherwise.

Define $\overline{x} = Y$ and $\underline{x} = Y - c$. With two players, there are two possible allocations for low types: $(\overline{x}, G_m)$, $(\overline{x}, G_l)$ and six for high types: $(\overline{x}, G_m)$, $(\overline{x}, G_l)$, $(\underline{x}, G_h)$, $(\underline{x}, G_m)$ and $(\underline{x} - p, G_h)$, $(\underline{x} - p, G_m)$, where h, m, l denote "high", "medium" and "low". To clarify, $G_h = F(2)$, $G_m = F(1)$ and $G_l = F(0)$.

Utilities from these allocations are ranked according to:

$$u^L(\overline{x}, G_m) > u^L(\overline{x}, G_l)$$

$u^H(\overline{x}, G_m) > u^H(\underline{x}, G_h) > u^H(\underline{x}, G_m) > u^H(\overline{x}, G_l)$

$u^H(\underline{x} - p, G_h) > u^H(\underline{x} - p, G_m)$

Both high and low types prefer having a high level of the private good and a medium level of the public good to all other alternatives. Nothing is assumed about the relative rankings of outcomes for high types where a penalty is assessed, except that they converge to the "no penalty" payoffs as p tends to 0.

Efficiency is determined in this setting by the Pareto criterion: an outcome is efficient if it is not possible for one agent to be made better off without reducing the other agent's welfare. An equilibrium of the above game is *ex post* efficient if it is Pareto efficient following the (possibly hypothetical) revelation of each agent's type.

The (Bayesian) Nash equilibrium concept is used to find equilibria. This criterion requires agents to play their best strategies for each possible type they may have, contingent on the Nash strategies played by other agents and the distribution over types. To simplify the analysis, suppose that $\alpha = 1/2$. When the probability of being each type is equal and independently drawn for each agent, the joint distribution of types is $[(H,H),(H,L),(L,H),(L,L);\ 1/4,1/4,1/4,1/4]$. This distribution is maintained for the extent of this section. The strategy set for each agent is the set of their possible type-reports, high or low, given their type, high or low. It has been established that low types will always report $\hat{L}$. Thus, for $i = 1, 2$, the strategies are (Report $\hat{H}$ or $\hat{L}$ given type is H, Report $\hat{L}$ given type is L), or in shorthand: $(\hat{H}\hat{L}), (\hat{L}\hat{L})$. The equilibrium concept involves calculating the best responses of each player to his opponent's possible strategies.

3.1 Full Privacy ($m = 0$)

With no monitoring, agents are free to report their types without fear of punishment should they misrepresent themselves to the government. Although low types report $\hat{L}$, a high-type may also report $\hat{L}$. If he does so, and his opponent reports $\hat{H}$, he will receive $u^H(\overline{x}, G_m)$, where $G_m = F(1)$.

To illustrate, consider the expected payoff to a high type agent of playing $\hat{H}$, contingent on his opponent playing $(\hat{H}\hat{L})$. This payoff is given by:

$$\begin{aligned} EU^H(\hat{H}|\hat{H}\hat{L}) &= \frac{1}{2}u^H(Y-c, F(2)) + \frac{1}{2}u^H(Y-c, F(1)) \\ &= \frac{1}{2}u^H(\underline{x}, G_h) + \frac{1}{2}u^H(\underline{x}, G_m) \end{aligned}$$

Equivalently, one can find $EU^H(\hat{L}|\hat{H}\hat{L})$, $EU^L(\hat{H}|\hat{L}\hat{L})$ and $EU^L(\hat{L}|\hat{L}\hat{L})$, and use these values to determine the best responses of a high type to his opponent's possible type-contingent strategies (recall that low types always have the best response $\hat{L}$). Proceeding in this manner to find Nash equilibria, we arrive at the following conclusions:

Remark 1: **(1)** If $u^H(\underline{x}, G_h) + u^H(\underline{x}, G_m) > u^H(\overline{x}, G_m) + u^H(\overline{x}, G_l)$, reporting $\hat{H}$ for type H is a dominant strategy. The unique equilibrium is $(\hat{H}\hat{L}, \hat{H}\hat{L})$ — the "truth telling" case — and all outcomes are ex post efficient. **(2)** If $u^H(\underline{x}, G_h) + u^H(\underline{x}, G_m) < u^H(\overline{x}, G_m) + u^H(\overline{x}, G_l)$, there are two pure strategy equilibria — $(\hat{H}\hat{L}, \hat{L}\hat{L})$ and $(\hat{L}\hat{L}, \hat{H}\hat{L})$ — and one mixed strategy equilibrium where each agent randomizes over $(\hat{H}\hat{L})$ and $(\hat{L}\hat{L})$. Not all equilibria are ex post efficient.

When both agents report truthfully as in **(1)** of the remark, the ex post set of outcomes are

1. $u^H(\underline{x}, G_h)$ for each if the types are (H, H).
2. $u^L(\overline{x}, G_l)$ for each if the types are (L, L).
3. $u^H(\underline{x}, G_m)$ and $u^L(\overline{x}, G_m)$ if the types are (H, L).

Examining the rankings of payoffs indicates that no other *feasible* allocations are Pareto superior than those obtained in this simple mechanism. Note also that truth-telling gives the high type a relatively low level of utility when his opponent turns out to be a low type. Case **(2)** of the Remark 1 is the interesting case for the purposes of analysis, since the ex post outcomes can produce inefficiency. Here the outcomes from pure strategies are

1. $u^H(\underline{x}, G_m)$ for one agent and $u^H(\overline{x}, G_m)$ for the other if the types are (H, H).
2. $u^L(\overline{x}, G_l)$ for each if the types are (L, L).
3. $u^H(\underline{x}, G_m)$ for the high-type agent and $u^L(\overline{x}, G_m)$ for the low-type **or** $u^H(\overline{x}, G_l)$ for the high-type and $u^L(\overline{x}, G_l)$ for the low if the types are (H, L).

Situations 1. and 2. here generate efficient outcomes. Situation 3. (when there is one agent of each type) generates an inefficient outcome when the high type agent plays $(\hat{L}\hat{L})$ and the low type plays $(\hat{H}\hat{L})$. In this case, welfare could be improved by compelling the high type to report truthfully, thereby raising G. Note also the distributional consequences of situation 1. Although each agent turns out to be identically a high type, one contributes his true information while the other does not. In this equilibrium, one agent always ends up misrepresenting.

The mixed strategy equilibrium is characterized by agents playing the strategy $(\hat{H}\hat{L})$ with probability q and $(\hat{L}\hat{L})$ with probability $(1-q)$, where q is defined by,

$$q = \frac{2[u^H(\underline{x}, G_m) - u^H(\overline{x}, G_l)]}{u^H(\overline{x}, G_m) + u^H(\underline{x}, G_m) - u^H(\underline{x}, G_h) - u^H(\overline{x}, G_l)}. \tag{2}$$

$q \in (0,1)$ when situation **(2)** holds. The mixed strategy concept does not proscribe a certain outcome, but rather allows many efficient and inefficient outcomes to occur with probabilities dependent on q. For instance, consider the strategy-profile $(\hat{L}\hat{L}, \hat{L}\hat{L})$ which would be expected to occur with probability $(1-q)^2$. If the ex-post distribution of types is (H, H), each player receives payoff $u^H(\overline{x}, G_l)$, the lowest ranked welfare level ex post.

The inability of pure privacy to guarantee fully efficient outcomes encourages one to ask: can reductions in privacy, construed as increased monitoring, increase the likelihood of efficient outcomes? The answer to this question is given in the following section. As it happens, if the penalty is not set too high, there exists an interesting pattern of behaviour as the level of monitoring increases. Here, increased monitoring can actually guarantee ex post *inefficiency* when one or both of the players are high types.

3.2 Monitoring $(0 < m \leq 1)$

As described above, the government now monitors agents and informs them that they are subject to a monitoring technology which can detect their true type with probability m.

Again, attention can be restricted to the strategies $(\hat{H}\hat{L})$ and $(\hat{L}\hat{L})$, since no low-type agent ever has an incentive to receive the allocation attainable by reporting himself as a high type. Now, however, the expected payoffs to the agents also depend upon m and p. Consider the expected payoff to an agent of playing $(\hat{L}\hat{L})$ given that his opponent also plays $(\hat{L}\hat{L})$. With probability $\frac{1}{4} = \frac{1}{2} \times \frac{1}{2}$, both agents are high-types and their strategies specify that they should each lie about their type. In this case, the payoff to an agent following such a strategy is:

$$\begin{aligned} EU^H(\hat{L}|\hat{L};(H,H)) \quad &= \quad m^2 u^H(\underline{x} - p, G_h) + m(1-m)u^H(\underline{x} - p, G_m) + \\ & \quad\quad m(1-m)u^H(\overline{x}, G_M) + (1-m)^2 u^H(\overline{x}, G_l) \end{aligned}$$

The first term in the above expression represents the allocation obtained when both agents are monitored and caught. Here, each agent pays the penalty[7], p, and is forced to pay compliance cost c. However, the highest value of G is produced, $G_h = F(2)$. The second term represents the allocation when the agent in question is monitored but the other is

[7] I assume that the penalty is discarded after its collection; i.e. it is not returned to the agents. This assumption is reconsidered in section 6.

not, and the third term the allocation when the agent in question is not monitored but his competitor is. With probability $(1-m)^2$, neither agent is monitored, and the outcome is that where each agent retains his private information but little of the public good is provided. By calculating the payoffs obtainable under every possible joint distribution of types, given the strategies played, the equilibria of this game are obtained.

An indepth description of the calculation of equilibria is omitted here. However, high type agents' behaviour can be characterized as follows. The report $\hat{H}$ is a best response to $(\hat{H}\hat{L})$ if:

$$\begin{aligned} R(m) \quad \equiv \quad & [u^H(\overline{x}, G_m) + u^H(\overline{x}, G_l) - u^H(\underline{x} - p, G_h) - u^H(\underline{x} - p, G_m)]m - \\ & [u^H(\overline{x}, G_m) + u^H(\overline{x}, G_l) - u^H(\underline{x}, G_h) - u^H(\underline{x}, G_m)] \geq 0 \end{aligned} \tag{3}$$

Given p, $R(m)$ is an increasing linear function in m, with $R(0) < 0$ and $R(1) \geq 0$, with equality if $p = 0$. Define $\dot{m}$ such that $R(\dot{m}) = 0$. Then $(\hat{L}\hat{L})$ is a best response to $(\hat{H}\hat{L})$ for all $m < \dot{m}$. The report $\hat{H}$ is a best response to $(\hat{L}\hat{L})$ if:

$$\begin{aligned} S(m) \quad \equiv \quad & [u^H(\overline{x}, G_m) + u^H(\underline{x} - p, G_m) - u^H(\underline{x} - p, G_h) - u^H(\overline{x}, G_l)]m^2 \\ & [u^H(\underline{x}, G_h) + 3u^H(\overline{x}, G_l) - u^H(\overline{x}, G_m) - u^H(\underline{x}, G_m) - 2u^H(\underline{x} - p, G_m)]m \\ & +2[u^H(\underline{x}, G_m) - u^H(\overline{x}, G_l)] \geq 0 \end{aligned} \tag{4}$$

Notice that $S(m)$ attains a unique minimum if $u^H(\overline{x}, G_m) + u^H(\underline{x} - p, G_m) > u^H(\underline{x} - p, G_h) + u^H(\overline{x}, G_l)$. If so, call this minimum $\breve{m}$. Since $S(0) > 0$ and $S(1) > 0$, if $S(\breve{m}) \leq 0$ for some $m \in (0, 1)$, and because $S(m)$ is quadratic in m, then there exist values of m, $m' \leq m''$, such that $(\hat{L}\hat{L})$ is a best response to $(\hat{H}\hat{L})$ for all $m \in (m', m'')$. Figure 1 illustrates the shapes of $R(m)$ and $S(m)$.

The equilibria to the game with monitoring are characterized in the following set of remarks:

Remark 2: When $m = 1$ (full monitoring or "no privacy"), each agent has a dominant strategy to report his true type for any $p \geq 0$.[8] Thus, the unique equilibrium is $(\hat{H}\hat{L}, \hat{H}\hat{L})$, and this outcome is ex post efficient.

Remark 3: Let $S(\breve{m}) \leq 0$ with $m' \leq m''$ and $m' < \dot{m}$. Then **(1)** For $m \in [0, m']$, the equilibria are $(\hat{H}\hat{L}, \hat{L}\hat{L})$, $(\hat{L}\hat{L}, \hat{H}\hat{L})$ and one mixed N.E. **(2)** For $m \in (m', \min\{\dot{m}, m''\})$, the unique equilibrium is $(\hat{L}\hat{L}, \hat{L}\hat{L})$. **(3)** For $m \in [\min\{\dot{m}, m''\}, \max\{\dot{m}, m''\})$, the equilibria

[8] In fact, there may exist some parameter values which sustain non-truthtelling equilibria for $p = 0$. These equilibria are eliminated by assuming that an agent plays $(\hat{H}\hat{L})$ no matter what his opponent plays when he is indifferent between $(\hat{H}\hat{L})$ and $(\hat{L}\hat{L})$.

are $(\hat{H}\hat{L}, \hat{L}\hat{L})$, $(\hat{L}\hat{L}, \hat{H}\hat{L})$ and one mixed N.E. **(4)** For $m \in [\max\{\dot{m},\ m''\}, 1]$, the unique equilibrium is $(\hat{H}\hat{L}, \hat{H}\hat{L})$. Several of these outcomes are ex post inefficient, and in situation **(2)**, all outcomes where there is at least one high type in the game are ex post inefficient.[9]

Remark 2 simply states that when agents are sure of being monitored they will never lie about their types. Thus, consider the polar cases of $m = 0$ and $m = 1$, maintaining the assumption that $u^H(\underline{x}, G_h) + u^H(\underline{x}, G_m) < u^H(\overline{x}, G_m) + u^H(\overline{x}, G_l)$. When $m = 0$, full privacy holds, and inefficiency occurs in some equilibria. When $m = 1$, there is no privacy, but efficiency is guaranteed.

Remark 3 points out that efficient outcomes are not necessarily more plausible as the level of monitoring rises. The intuition behind this result is as follows. Suppose agents are named A and B, and now consider the problem from agent A's perspective, where A is a high type. A knows that B will also be a high type with probability $\frac{1}{2}$. When B uses the strategy $(\hat{L}\hat{L})$ and $m = 0$, A's best response is to play $\hat{H}$ (see Remark 1). Now suppose m were to steadily increase. m' — if it exists — is the level of m at which it becomes a best response for A to play $\hat{L}$ given that B plays $(\hat{L}\hat{L})$. Why should such a threshold level of m exist? Essentially, A's selfish motive is to retain his private information while maximizing the level of G. But as m rises, the probability that a "high-type" B will be caught while playing $(\hat{L}\hat{L})$ increases, and when B is caught he is forced to contribute to G. At some point, it is possible that this incentive causes A to submit to the same possibility of getting caught, meaning that after m' is reached he will play $\hat{L}$ given that B is playing $(\hat{L}\hat{L})$. But suppose m rises further still. We know that at $m = 1$, $\hat{H}$ becomes a dominant strategy for A. If m' exists, then $m'' \geq m'$ is the level of monitoring at which A again chooses to play $\hat{H}$ given that B plays $(\hat{L}\hat{L})$. Intuitively, a level of m exists where the threat of being caught when playing $\hat{L}$ for A is too great; i.e. the possible loss from being caught outweighs the possible gain from a "high type" B being caught and paying towards G.

The threshold level $\dot{m}$ refers to the level of m such that agent A will begin playing $\hat{H}$ as a best response to B using the strategy $(\hat{H}\hat{L})$. When $m = 0$, A will use $\hat{L}$ as a best response in pure strategies: since a "high type" A stands no chance of being caught, he can attain a high level of utility by lying in the case where his opponent is also a high type *and* tells the truth. As m rises, the chance of being caught rises and A's gain from misrepresentation falls. Thus, $\dot{m}$ is the level of monitoring at which truth-telling $(\hat{H})$ becomes a best response to B's truth-telling $(\hat{H}\hat{L})$.

These two effects combine to produce different equilibria at different levels of m, espe-

[9]The mixed-strategy equilibrium is defined by the level of $q(m)$ which equates $q(m)EU(HL|HL) + (1 - q(m))EU(\hat{H}\hat{L}|\hat{L}\hat{L}) = q(m)EU(\hat{L}\hat{L}|\hat{H}\hat{L}) + (1 - q(m))EU(\hat{L}\hat{L}|\hat{L}\hat{L})$.

cially when there is a significant difference in high types' utility between the allocations $(\overline{x}, G_m)$ and $(\underline{x}, G_h)$. This is a case where compliance costs are high, and $G_l = F(0)$ is very low. Consider the following example (which lists only the utility to high-types since equilibrium payoffs to low-types' utilities are irrelevant in the calculation):

Example 1 Let $u^H(\overline{x}, G_m) = 13$, $u^H(\underline{x}, G_h) = 5$, $u^H(\underline{x}, G_m) = 4$, $u^H(\overline{x}, G_l) = 3$, $u^H(\underline{x} - p, G_h) = 4$ and $u^H(\underline{x} - p, G_m) = 3$. Then the following threshold values obtain: $m' = 1/3$, $m'' = 2/3$, $\dot{m} = 7/9$. The equilibrium outcomes are:

$$\begin{aligned}
(\hat{H}\hat{L}, \hat{L}\hat{L}) \text{ and } (\hat{L}\hat{L}, \hat{H}\hat{L}) \text{ and 1 mixed} \quad &\Longleftrightarrow \quad m \in [0, 1/3] \\
(\hat{L}\hat{L}, \hat{L}\hat{L}) \quad &\Longleftrightarrow \quad m \in (1/3, 2/3) \\
(\hat{H}\hat{L}, \hat{L}\hat{L}) \text{ and } (\hat{L}\hat{L}, \hat{H}\hat{L}) \text{ and 1 mixed} \quad &\Longleftrightarrow \quad m \in [2/3, 7/9) \\
(\hat{H}\hat{L}, \hat{H}\hat{L}) \quad &\Longleftrightarrow \quad m \in [7/9, 1]
\end{aligned}$$

It is clear to see that this mechanism *guarantees* efficiency only for levels of monitoring above 7/9. Indeed, for "moderate" levels of m such as 1/2, the only equilibrium is that where both agents misrepresent themselves when they are high-types. In this case, the only guaranteed ex-post efficient situation is that where both agents are low types; when they are both high types, each agent receives an allocation based on whether they or their opponent is monitored. Figure 1 is drawn using the numbers from this example.

Being monitored and caught is *never* efficient, even though monitoring is costless. Because agents who are caught must pay a fine *and* report their true type, they could have done better by telling the truth to begin with. Decreasing p decreases the slope of $R(m)$ and "flattens" $S(m)$. To illustrate, consider example 1, but with $p = 0$; that is, $u^H(\underline{x} - p, G_m) = u^H(\underline{x}, G_m) = 4$ and $u^H(\underline{x} - p, G_h) = u^H(\underline{x}, G_h) = 5$. In this case, $\dot{m} = 1$, while $m' = 0.22$ and $m'' = 1$. These parameter values indicate a wider range of monitoring values which sustain the "worst" outcome at $(\hat{L}\hat{L}, \hat{L}\hat{L})$, indicating that the penalty makes misrepresentation more unattractive for high types. The issue of what penalty might be levied in this sort of game is examined in the next section. One result there is that a benevolent government may not wish to ensure full compliance with strict penalties.

Example 1 required high type agents to place a very high value on the public good. If the G is not as valuable relative to x, or compliance costs are small, high type agents are less likely to risk misrepresenting themselves:

Example 2 Let $u^H(\overline{x}, G_m) = 12$, $u^H(\underline{x}, G_h) = 5$, $u^H(\underline{x}, G_m) = 4$, $u^H(\overline{x}, G_m) = 3$, $u^H(\underline{x} - p, G_h) = 4$ and $u^H(\underline{x} - p, G_m) = 3$. Then the following threshold values obtain:

$m' = m'' = 1/2$, $\dot{m} = 3/4$. The equilibrium outcomes are:

$$
\begin{aligned}
(\hat{H}\hat{L},\hat{L}\hat{L}) \text{ and } (\hat{L}\hat{L},\hat{H}\hat{L}) \text{ and 1 mixed} \quad &\Longleftrightarrow \quad m \in [0,1/2) \\
(\hat{L}\hat{L},\hat{L}\hat{L}) \quad &\Longleftrightarrow \quad m = 1/2 \\
(\hat{H}\hat{L},\hat{L}\hat{L}) \text{ and } (\hat{L}\hat{L},\hat{H}\hat{L}) \text{ and 1 mixed} \quad &\Longleftrightarrow \quad m \in (1/2,3/4) \\
(\hat{H}\hat{L},\hat{H}\hat{L}) \quad &\Longleftrightarrow \quad m \in [3/4,1]
\end{aligned}
$$

If all other parameter values are maintained, but $u^H(\overline{x}, G_m)$ is dropped to 11, $\dot{m}$ becomes 5/7, but m' and m'' cease to exist since $S(m)$ never dips below 0 for any m. That is, there is no level of monitoring for which a high type agent may be induced to report $\hat{L}$, given that his opponent is using the strategy $(\hat{L}\hat{L})$.

3.3 Other Results

3.3.1 Cheap Talk Between Agents

An important result to note is that pre-play communication in this game does not affect the equilibria. That is, if agents have access to a costless signal with which they can inform their opponent of their type (ie. "cheap talk"), then the equilibrium from this signalling process is "babbling" in that it conveys no relevant information (as in Morris (2001)). Suppose that agents can simply announce their types to each other before making their type-announcement to the government. Then it can be shown that all agents have a weakly dominant strategy to announce that they are low types. In this case, the announcement stage does not change agents' beliefs about their opponents' types and so does not change the reports made to the government.

3.3.2 Three Agents

If $n = 3$, the nature of agents' play is changed because there is now an additional potential contributor to the public good/aggregate information set. A high-type third agent can potentially add to the stock of G and increase the incumbent agents' utilities and possibly his own as well. If the incumbents were both low types, this would be an unambiguous welfare improvement. However, it is uncertain whether such welfare improving cases will occur. For example, suppose that there is a two-agent game, with both agents being high-types ex post and m being such that $(\hat{H}\hat{L},\hat{L}\hat{L})$ is the unique equilibrium of the mechanism. If a third player were added, $(\hat{H}\hat{L})$ might be the best response to her opponents' strategies $(\hat{H}\hat{L},\hat{L}\hat{L})$. To maintain the equilibrium $(\hat{H}\hat{L},\hat{L}\hat{L},\hat{H}\hat{L})$ — which would give agents 1 and 2 the same x_i regardless 3's of type, and greater G if 3 turns out to be a high type —

it must also be the case that $(\hat{L}\hat{L})$ is a best response to $(\hat{H}\hat{L}, \hat{H}\hat{L})$ for agent 2 (agent 1's problem is symmetric to that of agent 3). Whether this is the case depends upon the level of m and the way in which G enters agents' utilities. In general, it can be shown that truth-telling equilibria more unlikely in the three-player case — that is, for a given p, m must be higher to induce all agents to tell the truth. Intuitively, high type agents prefer to misrepresent themselves as the number of total agents rises, to take advantage of the fact that others may be caught and compelled to contribute to G.

4 The Model with n Agents

Individual behaviour in the type-reporting game depends on the total number of other individuals subject to the monitoring mechanism. This result follows because every agent cares about others' actions through the amount of G eventually provided. As monitoring increases, one blunt effect is simply to force more high type individuals to contribute the public good by directly observing their types. However, depending on the penalty, increasing m may reduce the tendency of high types to report truthfully and simply "take their chances" under m. Since the level of m is a rough substitute to the number of high types who voluntarily report $\hat{H}$, high type agents may be encouraged to "free-ride" on others' information if p is not too great.

The n-player model is constructed as follows. First, assume that n is large, so that a measure m of individuals are monitored and $(1-m)$ are not. For the same reasons as given above, low type agents always report $\hat{L}$ in equilibrium. High type agents, may report either $\hat{H}$ or $\hat{L}$ in equilibrium. Let $q \in [0, 1]$ be the probability that a high type agent reports $\hat{H}$ as part of a mixed strategy, or equivalently, since n is large, the proportion of high types who play $\hat{H}$ as a pure strategy. The derivation of q in equilibrium will be clarified below. Also, since n is large, the ex ante and ex post distributions of types are identical: $\{(H, L); \alpha, (1-\alpha)\}$.

The total amount of the public good provided ex post, as a function of (q, m) is:

$$\begin{aligned} G(q,m) &= F(\tilde{n}^H) \\ &= F(\alpha n q + \alpha n(1-q)m) \\ &= F(\alpha n[m + q(1-m)]) \end{aligned} \quad (5)$$

Each (q, m) pair defines a (potentially non-unique) level of G. For example,

$$G(0,m) = F(\alpha n m) \quad (6)$$

and

$$G(1,m) = G_{max} = F(\alpha n) \tag{7}$$

The latter value of G is independent of m since all high type agents contribute βc whether they are monitored or not. However, m is taken as given, whereas q is determined as a consequence of equilibrium play.

Consider the problem of any high-type agent facing $n-1$ other agents, where n is large. Define the equilibrium proportion of high-type agents playing $\hat{H}$ (or equivalently, the mixed-strategy probability of all high type agents playing $\hat{H}$) by $q^*(m,p) \in [0,1]$. Then, the choice problem of a high type agent is to report either $\hat{H}$ or $\hat{L}$ according to the rule $EU^H(\hat{L}, q, m) \lesseqqgtr EU^H(\hat{H}, q, m)$, or,

$$\begin{aligned}
(1-m)u^H(F(\alpha(n-1)[m+q(1-m)]), Y) +\\
mu^H(F((\alpha(n-1)[m+q(1-m)]) + 1), Y - c - p)\\
\lesseqqgtr u^H(F((\alpha(n-1)[m+q(1-m)]) + 1), Y - c)
\end{aligned}$$

In shorthand notation, this becomes

$$(1-m)u^H(G_l(q,m), \overline{x}) + mu^H(G_h(q,m), \underline{x} - p) \lesseqqgtr u^H(G_h(q,m), \underline{x}) \tag{8}$$

where "l" and "h" indicate "low" and "high" amounts of G, for any given (q,m) pair. Expression (8) is the best-response rule for a high-type agent when facing a proportion q of other high types reporting truthfully. If $m=1$, it is clearly a best-response to report $\hat{H}$ regardless of others' actions. Otherwise, the equilibrium proportion of truth-telling high type agents is $q^*(m,p)$, since, when q is viewed as a mixed strategy, this level ensures $EU^H(\hat{L}, q^*(m,p), m) = EU^H(\hat{H}, q^*(m,p), m)$.

Under certain forms of preferences, one can solve explicitly for $q^*(m,p)$. Here, preferences of the form $u^i(G,x) = b_i ln(G) + v(x)$ are used, where $v(\cdot)$ is an increasing concave function and $b_H > b_L$.

When preferences take this form, and $m \in [0,1)$,

$$\frac{G_l(q^*,m)}{G_h(q^*,m)} = h(m,p) \equiv exp\left\{ \left(\frac{1}{(1-m)b_H} \right) [v(\underline{x}) - mv(\underline{x}-p) - (1-m)v(\overline{x})] \right\} \tag{9}$$

When $h(m,p) \geq 1$, it must be that all high types are "honest" ($q^* = 1$). Given p, there will always be such a value of $m \in [0,1)$. To see this, note that $h(m,p) > 0$ for all $m \in [0,1)$, but $h(m,p) \geq 1$ for some values of m in this range as well. When this case occurs, there is no value of q for which $\frac{G_l}{G_h} = h(m,p)$. Specifically, $h(\tilde{m},p) = 1$ where

$$\tilde{m} = \frac{v(\overline{x}) - v(\underline{x})}{v(\overline{x}) - v(\underline{x} - p)} \tag{10}$$

Clearly, $0 < \tilde{m} \leq 1$. Since it can be shown that $h'(m,p) > 0$ for all m, there exists a $\tilde{m}$ such that $h(m,p) \geq 1$ for all $m \geq \tilde{m}$. In this range, reporting $\hat{H}$ is always a best response, regardless of other agents' actions. Thus, truthful revelation is guaranteed as an equilibrium given any level of $m > 0$ by imposing a sufficiently high fine, p.[10] Lastly, $h(m,p)$ reaches a minimum either if $m = 0$ or $p = 0$; call this h_{min}. That is,

$$h_{min} = h(0,p) = h(m,0) = exp\left\{ \left(\frac{1}{b_H}\right) [v(\underline{x}) - v(\overline{x})] \right\}. \tag{11}$$

4.1 Equilibrium

Equilibrium play requires optimal behaviour from high type agents. There are three possibilities in equilibrium:

1. All high types report $\hat{H}$ so $q^*(m,p) = 1$.
2. All high types report $\hat{L}$ so $q^*(m,p) = 0$.
3. A proportion $q^*(m,p) \in (0,1)$ report $\hat{H}$ and $(1-q^*(m,p))$ report $\hat{L}$, or, equivalently, all high types report $\hat{H}$ with probability $q^*(m,p)$.[11]

Analogous to the two-agent model, the interesting case for analysis is that where some high-type agents do not report truthfully. Assume therefore that $h(m,p) < 1$. This assumption rules out the case where p is set sufficiently high to ensure truth-telling by high-type agents. For simplicity, assume also that

$$F(\tilde{n}^H) = (\beta c[\alpha n(m + q(1-m))])^{\delta} \tag{12}$$

where $\delta \in (0,1]$ describes the public-good productivity of identifying additional high type agents. Expression (9) can be then be re-arranged to yield:

$$q^*(m,p) = \frac{m}{m-1} + \frac{1}{\alpha(n-1)(1-m)} \left(\frac{F^{-1}(h(m,p))}{[1 - F^{-1}(h(m,p))]} \right) \tag{13}$$

[10]This condition is not necessary to induce a truth-telling equilibrium for all agents, but it is sufficient. i.e. For a given p there can exist a $m' < \tilde{m}$ for which $q^*(m',p) = 1$.

[11]Again, this equivalence depends on n being large, a difference from the two player case.

For it to be an equilibrium proportion of high-type agents reporting $\hat{H}$, $q^*(m,p) \in [0,1]$. Letting the RHS of (13) vary with m, however, the calculated value of $q(m,p)$ can also be negative or greater than one. In the former case, set $q^*(m,p) = 0$ (no high types report $\hat{H}$) and in the latter, set $q^*(m,p) = 1$ (all high types report $\hat{H}$).

This setting ensures that some proportion of high types will reveal voluntarily even when $m = 0$, as in the standard case where G is provided privately. To see this, note that $h(0,p) = h_{min} \in (0,1)$. Then $F^{-1}(h_{min}) = (h_{min})^{1/\delta} \in (0,1)$ also, and

$$q^*(0,p) = \frac{1}{\alpha(n-1)} \left(\frac{F^{-1}(h_{min})}{[1 - F^{-1}(h_{min})]} \right) \in (0,1) \tag{14}$$

However, the reader should note that although $h(0,p) = h(m,0)$, it is not the case that $q^*(0,p) = q^*(m,0)$. As will be seen in the next section, it is quite possible that even though $q^*(0,p) > 0$, $q^*(m,0) = 0$.

Remark 4 For future reference, it is useful to note the comparative statics of $q^*(m,p)$. Specifically, $\partial q^*/\partial n < 0$, $\partial q^*/\partial \alpha < 0$ and $\partial q^*/\partial h(m) > 0$ always. As the number of agents in the economy rises, *ceteris paribus*, the tendency in equilibrium for high types to report truthfully falls. Similarly, the tendency for high types to report truthfully rises as the function $h(m)$ increases in value. Since $\partial h(m,p)/\partial p > 0$, $\partial q^*(m,p)/\partial p > 0$: higher penalties induce more truth-telling in equilibrium. However, the sign of $\partial q^*(m,p)/\partial m$ is ambiguous in general. Depending on other parameters, and the shapes of $v(\cdot)$ and $F(\cdot)$, higher levels of monitoring can either increase or decrease the equilibrium proportion of high types who tell the truth. Essentially, this reflects the dual incentive effects of monitoring: it increases the private-good cost of cheating due to a greater chance of being caught, but increases the level of G by increasing the number of cheaters who *are* caught.

To this point, nothing has been said about how m and p are determined; rather, only the agents' behaviour *given* these parameters has been analyzed. The next section analyzes what level of m will be chosen, as well as what levels of penalties and transfers will obtain under this model.

5 Monitoring and Policy Choice

Consider the following game in an economy with n agents of equal endowments Y, α of whom are high types (H) and $(1-\alpha)$ of whom are low types (L). As in the previous section, preferences take the form $u^i(x_i, G) = b_i ln(G) + v(x_i)$, with $b_H > b_L$. The government provides (ex post) an amount of the public good $G = (\beta c \tilde{n}^H)^\delta$ where $\delta \in (0,1]$ describes the productivity of high types' information. Given the potentially low level of public good provided with no monitoring ($m = 0$), described by equation (14), the government proposes the introduction of a monitoring scheme to identify high types directly:

Stage 1: Agents vote in a referendum in which they indicate their preferred level of monitoring, $m \in [0, 1]$. All monitoring technologies are feasible and costless, and the government implements the level of m which has the plurality. The results of the vote become common knowledge.

Stage 2: Given m, the government announces a level of penalty, $p \in [0, Y]$ for agents it finds to be misrepresenting their type. It also sets a *transfer* of T_H to high type agents reporting $\hat{H}$ whether they are monitored or not. Call this a "reward". To finance the reward, a tax payment, T_L, is levied on all non-monitored agents who report $\hat{L}$ and on monitored low type agents, who always report $\hat{L}$.[12] Transfers can be either positive or negative. Denote the policy vector as $\tau = (T_H, T_L, p)$.

Stage 3: With the monitoring system in place, agents make their reports to the government. The behaviour of agents is characterized by the analysis of the preceding section, except now the government's set of instruments is expanded to τ from simply p.

The (subgame perfect) equilibria of this game can be found by solving the game backwards in the conventional manner:

Stage 3: Given the values of m and τ, $h(m; \tau)$ is given by:

$$h(m;\tau) = exp\left\{ \left(\frac{1}{(1-m)b_H} \right) [v(\underline{x} - T_H) - mv(\underline{x} - p) - (1-m)v(\overline{x} - T_L)] \right\} \tag{15}$$

and the equilibrium proportion of high-type agents reporting truthfully, $q^*(m; \tau) \in [0, 1]$, is given by (13). The threshold value of m, $\tilde{m}$, inducing truthful revelation (for which $h(m, \tau) \geq 1$) is,

$$\tilde{m} = \frac{v(\overline{x} - T_L) - v(\underline{x} - T_H)}{v(\overline{x} - T_L) - v(\underline{x} - p)} \tag{16}$$

[12] This is analogous to Akerlof's (1978) choice of different consumption levels dependent on acquiring a "tag". Here, the tags are "high type, cheater", "high type, not a cheater", "high or low type, reporting $\hat{L}$".

Levels of $m \geq \tilde{m}$ ensure that $q^* = 1$. Conversely, levels of m such that

$$h(m;\tau) < F\left(\frac{\alpha(n-1)m}{1+\alpha(n-1)m}\right) \tag{17}$$

must hold to ensure $q^* = 0$. Since the right hand side of (17) approaches 1 as n grows large, the range of m for which $q(m,\tau) \in (0,1)$ becomes very small away from $m = 0$. For this reason, the focus below is on government policies which set either $q^* = 0$ or $q^* = 1$.

Stage 2: Following Parsons (1996), the government is benevolent, and sets τ to maximize the sum of ex post utilities weighted by the proportion of each group in the population[13] :

$$\begin{aligned} U^{gov} \quad = \quad & \alpha n[q(m;\tau)u^H(G(q(m;\tau),m),\underline{x}-T_H)+ \\ & (1-q(m;\tau))mu^H(G(q(m;\tau),m),\underline{x}-p)+ \\ & (1-q(m;\tau))(1-m)u^H(G(q(m;\tau),m),\overline{x}-T_L)]+ \\ & (1-\alpha)nu^L(G(q(m;\tau),m),\overline{x}-T_L) \end{aligned} \tag{18}$$

subject to budget balance,

$$\alpha nq(m;\tau)T_H + [\alpha n(1-m)(1-q(m;\tau)) + (1-\alpha)n]T_L + m(1-q(m;\tau))p = 0 \tag{19}$$

In this formulation, the government cares only about the final welfare of each agent, independent of their previous behaviour. Since the government anticipates the play of agents in the next stage (as given by $q(m,\tau)$), it can usually induce $q^* = 0$ or $q^* = 1$ with a certain set of policy, τ. The play of the agents then determines the level of G. The set of policies generating $q^*(m;\tau) \in (0,1)$ are less easily observed due to the recursive nature of the solution: in stage 3, $q(m,\tau(q))$ via (15). A solution (q^*,τ^*) with $0 < q^* < 1$ — if it exists— is the fixed point of such a calculation.

Three relevant policy options exist for the government given the instruments at its

[13]The assumption of benevolence means that the government maximizes a function of agents' utilities. The inclusion of *all* agents utilities in the objective function may be seen as problematic: ie. why would the government care about the welfare of cheaters? There is a literature in which the government may exclude certain characteristics from compensation since they are seen as the agents' responsibility (see e.g. Fleurbaey and Maniquet (1999) and Roemer (1998)). In one sense, this is a benchmark case; the government could maximize (18) by excluding the utilities of the $(1-q)$ high types who cheat. Since the government can influence the degree of cheating (i.e. $q(m;\tau)$), it could set policies to increase q to one, thus increasing its own objective. However, doing so would alter the objective of the government significantly — it would be concerned not just with welfare broadly defined, but also with eliminating cheating for its own sake. Alternatively, one could interpret (18) as that objective when the government cannot commit to a more stringent welfare measure. Confronted with a choice of m from stage 1, the government cannot help but maximize all agents' utilities in stage 2.

disposal:

1. Set $T_H = T_L = p = 0$. Call this policy M.
2. Set $p = 0$, and T_H, T_L as determined by (19). Call this policy T.
3. Set $p > 0$, and $T_H = T_L = 0$. Call this policy P.

Each of these policies can be examined in turn. The government will then choose, given m, the policy that maximizes (18).

Policy M If the government does nothing, $G(0, m) = F(\alpha n[m+q(1-m)])$, with $q(m; 0)$. "Cheating" high types are not penalized beyond having to pay their compliance costs, c.

Policy T (19) can be rearranged as:

$$T_H = -\frac{[1 - \alpha(m + q(1 - m))]}{\alpha q} T_L = -\omega T_L \tag{20}$$

Suppose the government were to maximize (18) for an arbitrary q. Then the first order condition is:

$$\alpha q \omega u_x^H(\underline{x} + \omega T_L) = \alpha(1 - q)(1 - m)u_x^H(\overline{x} - T_L) + (1 - \alpha)u_x^L(\overline{x} - T_L) \tag{21}$$

When $u_x^H = u_x^L = v'(x)$, (21) rearranges to the condition:

$$v'(\underline{x} + \omega T_L) = v'(\overline{x} - T_L). \tag{22}$$

Solving for T_L yields,

$$T_L(q) = \frac{\alpha q c}{1 - \alpha m(1 - q)} \tag{23}$$

The government can naively set a tax based on what it expects q to be in Stage 3. However, given any q it desires, it sets $v(Y - c + \omega T_L) = v(Y - T_L) > v(Y - c)$, no matter what q is eventually played. But by (8), even with $p = 0$, a high type agent will never report $\hat{L}$. Thus, a solution (q^*, T_L^*) of the government's choice is to set $T_L(q = 1) = \alpha c$. Indeed, this is what the government would choose if types were *perfectly observable*. Note that p is redundant in any case when a tax/transfer scheme is in place, hence $p = 0$ is a valid conjecture. This policy results in the highest possible value of $G_{max} = F(\alpha n)$, and spreads the burden of compliance costs across the society. It is also invariant to the level of m selected in stage 1.

Policy P Instead of a tax/transfer scheme, the government might choose to apply a penalty, p, on monitored high types reporting $\hat{L}$. From the definition of $h(m;\tau)$, the government can always set p sufficiently high to implement $q^* = 1$ in stage 3, given m, but then it will never be collected. Alternatively, depending on the choice of m, it may set a penalty sufficiently low to set $q^* = 0$ in stage 3. But by inspection of (18), if $m > 0$ and $q^* = 0$, the government would never choose such a penalty, since it would only reduce the utility of monitored high types reporting $\hat{L}$. Similarly, suppose the government set p to induce $0 < q^* < 1$ with the proceeds transferred equally to all other agents. With $m > 0$, this policy would be identical to the tax transfer scheme in which case p is again redundant. Thus, if $p > 0$, it must be high enough to force $q^* = 1$. Then $G_{max} = F(\alpha n)$, but the compliance costs are borne entirely by high types.

Stage 1: Suppose that all agents are asked to vote in a secret ballot as to the level of monitoring which should be implemented in this setting. This ballot is essentially blank, with each agent writing in his preferred level of monitoring.[14] Suppose further that the m implemented is that level which obtains a plurality of votes. Since it is the case that all high types and all low types vote in the same way, respectively, then the winning level of m is simply the preferred level of the type that is in the majority (ie. the median voter).

Agents of each type anticipate the best response of the government to various levels of m and vote for that which brings their type the highest expected utility. Which policy is chosen depends on several parameters in the economy, including α, $F(\cdot)$, $(Y - c)$ and n (via G). The government's choice between a tax/transfer scheme and a penalty on cheaters is invariant to m, however. Thus, it is always feasible for the government to make the agents' choice of m redundant.

5.1 Equilibria in Monitoring and Policy Choice

It has already been shown that the introduction of a tax/transfer scheme — for any level of m chosen in the first stage — produces $q^* = 1$. Also, a benevolent government will never impose a penalty which leads to $q^* < 1$. An ambiguity remains in the case where no policy tools beyond m are used ($\tau = (0, 0, 0)$). However, by inspection of (11), (15) and (17), this means $h(m, \tau) = h_{min} < 1$ and if n is sufficiently large, $q^*(m; 0) \in (0, 1)$

[14] Of course, one can think of many alternate mechanisms for choosing the level of monitoring. A simple alternative, for example, might include the government *naming* a level of m, and asking agents to vote "yes" or "no" on that level, with the alternative being the absence of a monitoring technology (full privacy). In this case, the government would be more able to implement its preferred level of monitoring by offering a majority of the agents better welfare levels than they would obtain in the full privacy case. If agents do not know their types when casting their ballots, it is reasonable to assume that their preferences would be exactly those of the government, in which case stage 1 would be redundant.

only for levels of m equal to and approaching zero. Assuming that $q^* = 0$ for all $m > 0$ permits a more straightforward analysis of expected payoffs to agents and the government for levels of $m > 0$. Define an equilibrium of this game as the triplet (m^*, Policy$\{M, P$ or $T\}, q^*$).

With preferences specified as above, and the condition that $q^*(m; 0) = 0$ for $m > 0$, the payoffs from the three policy regimes (M, P, T) can be defined.

High Types: Under each policy, the Stage 1 payoffs for a high type agent are:

$$\begin{aligned}
U_M^H &= mv(Y - c) + (1 - m)v(Y) + b_H \ln(G(0, m)) \\
U_T^H &= v(Y - \alpha c) + b_H \ln(G_{max}) \\
U_P^H &= v(Y - c) + b_H \ln(G_{max})
\end{aligned} \tag{24}$$

Low Types: Under each policy, the Stage 1 payoffs for a low type agent are:

$$\begin{aligned}
U_M^L &= v(Y) + b_L \ln(G(0, m)) \\
U_T^L &= v(Y - \alpha c) + b_L \ln(G_{max}) \\
U_P^L &= v(Y) + b_L \ln(G_{max})
\end{aligned} \tag{25}$$

Government: Under each policy, the Stage 2 payoffs for the government are (from (18)):

$$\begin{aligned}
U_M^{gov} &= [(1 - \alpha m)v(Y) + \alpha mv(Y - c) + [\alpha b_H + (1 - \alpha)b_L] \ln(G(0, m))]n \\
U_T^{gov} &= [v(Y - \alpha c) + [\alpha b_H + (1 - \alpha)b_L] \ln(G_{max})]n \\
U_P^{gov} &= [\alpha v(Y - c) + (1 - \alpha)v(Y) + [\alpha b_H + (1 - \alpha)b_L] \ln(G_{max})]n
\end{aligned} \tag{26}$$

Lemma 1: (a) The penalty scheme, P, is *never* preferred by high type agents, (b) P is *always* preferred by low type agents, (c) P is *never* preferred by the government.

These results are intuitive. Low types would always prefer a penalty to a tax since in the former case they receive G_{max} without bearing any private costs. They also bear no private costs under a monitoring scheme, but receive more G under the penalty scheme. High types always prefer a tax scheme to a penalty since they will report truthfully under each scheme, but will be compensated for some of their compliance costs under the former. It may be the case that $U_P^H \geq U_M^H$ over some ranges of m. However, it follows from Lemma 1 that if $U_M^H \geq U_T^H$ for some range of m, then $U_M^H > U_P^H$ over this range by the transitive property.

5.1.1 The Government's Response to m

Lemma 1 ensures that there will be no equilibrium in which a penalty scheme is used. Thus, further analysis can abstract from comparisons of U_M^{gov} and U_P^{gov}, since this pairwise ranking will not affect how the government chooses between M and T. The general rule for comparing the tax scheme with the monitoring scheme for the government is the following: A tax scheme (T) is preferred to a monitoring scheme (M) if $U_T^{gov} > U_M^{gov}$, or:

$$\ln\left(\frac{G_{max}}{G(m,0)}\right)[\alpha b_H + (1-\alpha)b_L)] > (1-\alpha m)v(Y) + \alpha m v(Y-c) - v(Y-\alpha c) \qquad (27)$$

Given m, the left hand side of (27) represents the additional amount of G obtainable from moving to a tax scheme (under which $q^* = 1$), whereas the right hand side is the extra private benefit from retaining a monitoring scheme only. This leads to a second result:

Lemma 2: With $q^* = 0$, there always exists a level of $m < 1$, $\hat{m}$, such that the government prefers a tax scheme to a monitoring scheme for all $m > \hat{m}$.

Lemma 2 states that the government always has a best response to choose scheme T if agents choose a sufficiently high m in stage 1. However, note that (27) does not rule out the case that $U_T^{gov} > U_M^{gov}(m)$ for *all* m. In this scenario, the government always chooses to implement scheme T in stage 2 regardless of what m is chosen in stage 1. If such is the case, then there is no voting equilibrium in m since the final allocations will be invariant to the choice of m. This result holds no matter which type is in the majority.

Remark 5: If $U_T^{gov} > U_M^{gov}(m)$ for all m, the set of equilibria of the monitoring/policy choice game is $(m^* \in [0,1], T, q^* = 1)$.

The complementary case is that where the government prefers M to T over some range of m. Note that the function $U_M^{gov}(m)$ is hump-shaped with respect to m.[15] Denote as $\tilde{m}^{gov}$ the value of m which maximizes $U_M^{gov}(m)$. Then if $U_M^{gov}(\tilde{m}^{gov}) > U_T^{gov}$, then $U_M^{gov}(m) > U_T^{gov}$ for some range of m around $\tilde{m}^{gov}$. Due to Lemma 2, there will be some upper bound to this range, given by $\hat{m}$.

[15] This may be observed by taking first and second derivatives — if a certain level of m solves the first order condition, then it is a maximum since the second order condition is non-positive.

5.1.2 Agents' Voting Behaviour

In the first stage, the agents must anticipate the government's best policy action in the second stage when voting. All high type agents vote the same, as do all low type agents. Thus, the winning level of m is that of the majority group, determined by α, the proportion of high types.

High Types in the Majority: **($\alpha > 1/2$)** $U_M^H(m)$ is also a hump-shaped curve, attaining a maximum at $\tilde{m}^H$. High types will vote for the specific level of monitoring that maximizes $U_M^H(m)$, conditional on (a) $U_M^H(\tilde{m}^H) > U_T^H$ *and* (b) $U_M^{gov}(\tilde{m}^H) > U_T^{gov}$. Fortunately, the following lemmas establish that condition (a) implies condition (b).

Lemma 3: $\tilde{m}^{gov} > \tilde{m}^H$.

Lemma 4: If $U_M^H(\tilde{m}^H) > U_T^H$ for some $m \in [0, 1)$, then $U_M^{gov}(\tilde{m}^H) > U_T^{gov}$. However, $U_M^H(\tilde{m}^H) < U_T^H$ does not necessarily imply $U_M^{gov}(\tilde{m}^{gov}) < U_T^{gov}$.

Lemma 4 states that if high types ever prefer M to T, the government will also prefer M to T at the level preferred by high types. Thus, high types will vote for $\tilde{m}^H$ when $U_M^H(\tilde{m}^H) > U_T^H$. However, the additional result is that even if high types *never* prefer M to T, it is possible that the government *does* over some range of m. Thus, in equilibrium, high types will never vote for m in such a range if they themselves always prefer T to M.

Remark 6: Given $U_M^{gov}(\tilde{m}^{gov}) > U_T^{gov}$ and $\alpha > 1/2$, (a) If $U_M^H(\tilde{m}^H) > U_T^H$, the unique equilibrium of the monitoring/policy choice game is $(m^* = \tilde{m}^H, M, q^* = 0)$. (b) If $U_M^H(\tilde{m}^H) < U_T^H$, the set of equilibria of the monitoring/policy choice game is $(m^* = \{m \in [0, 1]$ such that $U_M^{gov}(m) < U_T^{gov}\}, T, q^* = 1)$.

Low Types in the Majority: **($\alpha < 1/2$)** From (25), observe that $U_M^L(m)$ is a strictly increasing function in m. Lemma 1 established that low types would always prefer a penalty scheme as government policy, but that the government would never want such a scheme when a tax/transfer scheme is available. Indeed, from Lemma 2, the government will always implement scheme T if m is sufficiently high. Thus, the voting behaviour of low types is governed by choosing a level of m, $\tilde{m}^L$, which maximizes $U_M^L(m)$ conditional on (a) $U_M^L(\tilde{m}^L) > U_T^L$ and (b) $U_M^{gov}(\tilde{m}^L) > U_T^{gov}$. The following lemma establishes that any m satisfying condition (b) also satisfies condition (a).

Lemma 5: If $U_M^{gov}(m) > U_T^{gov}$ for some $m \in [0,1)$, then $U_M^L(m) > U_T^L$. However, $U_M^L(m) > U_T^L$ for some m does not necessarily imply $U_M^{gov}(m) > U_T^{gov}$.
Low types' voting behaviour is governed by the following remark:

Remark 7: Given $U_M^{gov}(\tilde{m}^{gov}) > U_T^{gov}$ and $\alpha < 1/2$, the equilibrium of the monitoring/policy choice game is $(m^* = \tilde{m}^L = \hat{m}, M, q^* = 0)$.

Recall from Lemma 2 that $\hat{m}$ is the highest level of m at which the government still prefers M to T. Given that $U_M^{gov}(\tilde{m}^{gov}) > U_T^{gov}$ for $\tilde{m}^{gov} \in (0,1)$, $0 < \tilde{m}^{gov} \leq \hat{m} < 1$. Thus, even when low types are in the majority, the equilibrium outcome will generate $m^* < 1$. This result is surprising, since low types themselves have nothing valuable to hide under a monitoring scheme. It is informative to examine some cases by example.

Example 3: Let $Y = 5$, $c = 3$, $\alpha = 0.4$, $\beta = 0.2$, $n = 100$, $b_H = 0.3$, $b_L = 0.2$. Also let $v(x) = \sqrt{x}$ and $F(\cdot) = \sqrt{\cdot}$. Under scheme M, $\tilde{m}^{gov} = 0.36$ ($U_M^{gov}(0.36) = 237$). Under scheme T, $U_T^{gov} = 233$ while $U_T^L = 2.26$. Low types vote for $\tilde{m}^L = \hat{m} = 0.78$, above which the government would switch to a tax scheme. Ironically, $U_M^H(\tilde{m}^H) < U_T^H$ for this example, implying that high types here would actually prefer a tax scheme to any monitoring level. See Figure 2.

Example 4: Assume the same parameters as above, except $\alpha = 0.6$ so that high types are in the majority. Under scheme M, $\tilde{m}^{gov} = 0.25$ ($U_M^{gov}(0.25) = 239$), while $\tilde{m}^H = 0.18$ ($U_M^H(0.18) = 2.36$). Under scheme T, $U_T^{gov} = 225$ while $U_T^H = 2.32$. But because the government prefers M to T for $0.05 < m < 0.85$, the high types freely select their preferred level of m and so $m^* = 0.18$. See Figure 3.

Example 5: Again, assume the same parameters as in example 4, except with $\alpha = 0.6$, $Y = 7$ and $b_H = 0.5$. High types are again in the majority. Under scheme M, $\tilde{m}^{gov} = 0.42$ ($U_M^{gov}(0.42) = 302$), while $\tilde{m}^H = 0.37$ ($U_M^H(0.37) = 3.07$). However, under scheme T, $U_T^{gov} = 297$ and $U_T^H = 316$. Thus, high types always prefer T to M, although for $m \in [0.25, 0.8]$, the government does not. Therefore, $m^* \in [0, 0.25) \cup (0.8, 1]$, and scheme T is implemented. The low types in this example would have preferred $\hat{m} = 0.8$. See Figure 4.

Examples can also be constructed for which full privacy is chosen endogenously in the model. Typically, this result requires a public good function such as $G = \bar{G} + f(\cdot)$.

5.2 Welfare

The level of monitoring chosen by agents and the policy chosen by the government affect the relative benefits of high and low type agents. For example, suppose low type agents are in the majority and vote for a high level of monitoring (but less than unity). It may be the case that high type agents would prefer scheme T at this level of m, although a switch to a lower level of m or to scheme T would make low types worse off. In this sense, all outcomes of the monitoring/policy choice game are *interim* Pareto efficient, i.e., from the perspective where agents know their types but not the exact outcome of the mechanism. Under scheme M, with $q^* = 0$, a proportion m of high types will be caught and receive $v(Y - c) + b_H \ln(G(0, m))$ ex post. Clearly these agents would have preferred to report honestly in stage 3 of the game, but rationally chose to take their chances under the monitoring mechanism, given their type.

Notably, the interim inefficient outcomes which exist in the model are for levels of monitoring below $\tilde{m}^H$ under scheme M, when $U_M^H(\tilde{m}^H) > U_T^H$. These levels of m will never be implemented in equilibrium, although one can imagine cases where legal restrictions prevent the introduction of a monitoring scheme and set $m = 0$ (with $q^* > 0$ but small). In other words, *if high types ever prefer some monitoring to none, a "right to privacy" is never efficient even though it allows for voluntary reporting.* Combining Lemmas 4 and 5, $U_M^H(\tilde{m}^H) > U_T^H$ implies $U_M^L(\tilde{m}^H) > U_T^L$. Thus, reducing m below $\tilde{m}^H$ creates a situation where all agents prefer more monitoring to less at the interim stage when they know their types. This observation shows that the introduction of an imperfect monitoring scheme may be Pareto superior to the absence of any monitoring scheme with only voluntary contributions. Levels of $m > \tilde{m}^H$ are preferred by both the government (see Lemma 3) and low types, but implementing this level of monitoring can only harm high types under policy M. These differences in preferences over m reflect the observed tension between citizens and the government with respect to private information and the government's policies for gathering it.

Finally, note the welfare consequences which would result if agents were asked to vote over m at the ex ante stage, i.e. when agents did not know their types, but only α. In this case their preferences would be identical to those of the government, and the outcomes would reflect the government's desires from stage 2 of the game above. We may then view $\tilde{m}^{gov}$ as the optimal level of monitoring ex ante, as well as the optimal level from the government's perspective. As shown above, this level is different from that of both high and low types, meaning that when stage 1 is included the monitoring level chosen is always suboptimal for the government.

5.3 Extension: Different Valuation of G

To this point, it has been assumed that high types prefer the public good relatively more than low types. In the present model, this condition means that $b_H > b_L$. One can think of situations where such an assumption is unreasonable: consider the case of "deviant" behaviour by high types where G represents a measure of public safety or ambiance. For instance, suppose that high type agents have the tendency to dump used paint down their drains, polluting local waterways. These agents may care to an extent about environmental quality, but not to the same extent as low type agents, who do not dump paint. The government may consider implementing a monitoring technology to identify the "dumpers" and thus increase the level of G. Under a policy such as M, these agents may be forced to take their used paint to a recycling center; under a policy such as T, the government may set up a program to collect paints for free from people who come forward, by taxing those who do not use the program.

Suppose then that $b_L > b_H \geq 0$. It can be verified that Lemmas 1-3 are unaffected by this change, but that Lemmas 4-5 cease to hold with certainty. Thus, for example, it is not the case that if high types ever prefer policy M to T that the government will as well. Because their preference for the public good holds less weight in the government's objective, the government will be more likely to prefer the same policies as do low-types — namely, greater compliance and contribution to G. Thus, voting equilibria over m will change when high types are in the majority.

As an extreme case, suppose that high types have no preference for the public good, $b_H = 0$. It is easy to see in this case that high types will always prefer $m = 0$, with no tax/transfer scheme, since this structure guarantees such agents the payoff $v(Y)$. The government's utility from a monitoring scheme is:

$$U_M^{gov} = [(1-\alpha m)v(Y) + \alpha m v(Y-c) + (1-\alpha)b_L \ln(G(0,m))]n \tag{28}$$

and from a tax/transfer scheme is:

$$U_T^{gov} = [v(Y-\alpha c) + (1-\alpha)b_L \ln(G_{max})]n \tag{29}$$

Note that if $b_H = 0$, high types will never contribute voluntarily even if $m = 0$, so $q^* = 0$ for this level of m. If $G(q,m) = G(0,0) = 0$, then both the government and low types receive arbitrarily low utility from the "full privacy" case that would result if high types were in the majority. In this case, it may be necessary for high types to be extra-compensated (above Y) for truthful reporting.

In cases where b_H is significantly less than b_L, but still positive, the government's best policy response is unclear. The shape of the public good technology, $F(\cdot)$, and the value of α may matter a great deal. Two polar examples illustrate the intuition. Suppose high types litter, and $\alpha > 1/2$. Then the government (and high types) may prefer a very small amount of monitoring of this behaviour, because capturing everyone who litters imposes many costs without major benefits. Conversely, suppose high types are sex offenders and α is very small. Then the government (and low types) will prefer a very high amount of monitoring, since leaving even a few sexual offenders uncaught may impose small compliance costs relative to the larger public benefit.

6 Concluding Remarks

The model outlined in this paper provides some intuition about the relationship between privacy, monitoring and welfare. When the number of agents is small, this simple model shows how perfect levels of monitoring are not necessary to guarantee ex post efficiency, but that privacy can lead to inefficient outcomes. These inefficiencies occur because, ex post, the distribution of agents may be such that each agent would have been better off playing a different strategy. In some cases, only very high levels of m may *guarantee* efficiency even though some efficient equilibria exist at lower levels. As m rises, so does the probability that agents' types (information) are revealed and contributions are made to G. But such cases may be "worse" than simple truth-telling if cheating high-types must pay a penalty of p.

When n is large, the ex post inefficiencies due to skewed ex post distributions of agents' types are eliminated, since *on average* the distribution of types $(\alpha, 1-\alpha)$ is stable. Changes in the level of monitoring often help one group and hurt another, exemplifying the tradeoff implicit in privacy issues. From the government's perspective, agents choose either too much and too little privacy. However, it is not necessarily the case that high type individuals want more privacy and low type individuals want less. When the government can control redistribution and penalties, a range of equilibria are possible.

As in the earlier quote, the model can generate a rationale behind "slack" frequently observed in government policy. When monitoring is less than perfect, a portion of the population can cheat and not be caught. This behaviour is accepted by the government, since full compliance would bring too few benefits for the costs it generates. Notably, full compliance may not be optimal for the government even when its costs can be spread socially. However, such an outcome can be generated when agents vote for a high level of monitoring before the government decision is made.

This paper also answers the question: why might not all agents agree to agree to either zero or full privacy? The answer given here is that an incentive exists for individuals with valuable information to submit false reports when monitoring is imperfect. In the model, individuals are unconcerned with privacy for its own sake. However, more complex behaviour likely generates the demand for privacy and the way it enters into welfare. The present analysis has abstracted away from a number of features, including dynamic concerns (confidentiality), non-benevolent government, and errors in monitoring ability. These issues are likely also important to the privacy issue in economics.

Appendix: Proofs

Lemma 1: (a) It suffices to find a policy instrument that is always preferred to P. Clearly $U_T^H > U_P^H$ since $v(Y - \alpha c) > v(Y - c)$. (b) $U_P^L > U_T^L$ since $v(Y) > v(Y - \alpha c)$. $U_P^L > U_M^L$ since $G_{max} \geq G(0, m)$ for all $m \leq 1$, by definition. Thus P is always preferred. (c) Since $v(\cdot)$ is concave, $v(Y - \alpha c) > \alpha v(Y - c) + (1 - \alpha)v(Y)$. Thus, $U_T^{gov} > U_P^{gov}$ always.

Lemma 2: From the result of Lemma 1, if $m = 1$ then the RHS of (27) is negative. Thus, $U_T^{gov} > U_M^{gov}(1)$ always. Observe that $(1 - \alpha m)v(Y) + \alpha m v(Y - c)$ is decreasing monotonically in m, and that $v(Y) > v(Y - \alpha c)$. Thus, there exists a monitoring level, $\check{m} \in (0, 1)$, such that $v(Y - \alpha c) = (1 - \alpha\check{m})v(Y) + \alpha\check{m}v(Y - c)$. Levels of $m \geq \check{m}$ are therefore *sufficient* to make $U_T^{gov} > U_M^{gov}(m)$, since the LHS of (27) is positive by definition.

Lemma 3: From the first order condition of maximizing $U_M^H(m)$, $\tilde{m}^H$ must satisfy:

$$v(Y) - v(Y - c) = \frac{b_H G'(0, \tilde{m}^H)}{(G(0, \tilde{m}^H)}$$

while $\tilde{m}^{gov}$ must satisfy:

$$v(Y) - v(Y - c) = \frac{b_H G'(0, \tilde{m}^{gov})}{G(0, \tilde{m}^{gov})} + \frac{(1 - \alpha)}{\alpha} \frac{b_L G'(0, \tilde{m}^{gov})}{G(0, \tilde{m}^{gov})}$$

For both conditions to hold, it must be that $\frac{G'(0,\tilde{m}^H)}{G(0,\tilde{m}^H)} > \frac{G'(0,\tilde{m}^{gov})}{G(0,\tilde{m}^{gov})}$, or equivalently, $\frac{G'(0,\tilde{m}^H)}{G'(0,\tilde{m}^{gov})} > \frac{G(0,\tilde{m}^H)}{G(0,\tilde{m}^{gov})}$. If $\tilde{m}^H > \tilde{m}^{gov}$, the RHS of this expression must be greater than one, and the LHS must be less than one, due to the properties of $F(\cdot)$. But then the inequality cannot hold. Therefore, $\tilde{m}^{gov} > \tilde{m}^H$.

Lemma 4: From Lemma 3, we know that $\tilde{m}^{gov}$ and $\tilde{m}^H$ take different values. From (24), we know that if $U_M^H(\tilde{m}^H) > U_T^H$, then

$$\tilde{m}^H v(Y - c) + (1 - \tilde{m}^H)v(Y) + b_H \ln(G(0, \tilde{m}^H)) = v(Y - \alpha c) + b_H \ln(G_{max}) + \Omega$$

where $\Omega > 0$. Suppose that, towards establishing a contradiction, $U_M^{gov}(\tilde{m}^H) \leq U_T^{gov}$. Then from (26)

$$(1 - \alpha\tilde{m}^H)v(Y) + \alpha\tilde{m}^H v(Y - c) + [\alpha b_H + (1 - \alpha)b_L] \ln(G(0, \tilde{m}^H)) \leq \\ v(Y - \alpha c) + [\alpha b_H + (1 - \alpha)b_L] \ln(G_{max})$$

Adding $(1-\alpha)b_H \ln G_{max}$ to each side, and substituting the above expression, we have:

$$\begin{aligned}(1-\alpha\tilde{m}^H)v(Y) + \alpha\tilde{m}^H v(Y-c) + [\alpha b_H + (1-\alpha)b_L]\ln(G(0,\tilde{m}^H)) \\ + (1-\alpha)b_H \ln G_{max} \leq \\ (1-\tilde{m}^H)v(Y) + \tilde{m}^H v(Y-c) + b_H \ln(G(0,\tilde{m}^H)) - \Omega + (1-\alpha)b_L \ln(G_{max})\end{aligned}$$

Simplifying and rearranging,

$$\tilde{m}^H[v(Y) - v(Y-c)] + (b_H - b_L)[\ln G_{max} - \ln(G(0,\tilde{m}^H))] + \frac{\Omega}{(1-\alpha)} \leq 0$$

But the left hand side of this expression must be strictly positive for any $m \geq 0$, a contradiction. Thus, $U_M^{gov}(\tilde{m}^H) > U_T^{gov}$.

To show that $U_M^H(\tilde{m}^H) < U_T^H$ does not necessarily imply $U_M^{gov}(\tilde{m}^H) < U_T^{gov}$, the same exercise can be repeated with $U_M^H(\tilde{m}^H) + \Phi = U_T^H$, $\Phi > 0$. This shows that $U_M^H(\tilde{m}^H) < U_T^H$ does not imply $U_M^{gov}(\tilde{m}^H) < U_T^{gov}$, for Φ sufficiently small. Since $\tilde{m}^{gov} > \tilde{m}^H$ from Lemma 3, this result also shows that it is possible to have $U_M^{gov}(\tilde{m}^{gov}) > U_T^{gov}$ along with $U_M^H(\tilde{m}^H) < U_T^H$.

Lemma 5: The proof is analogous to that of Lemma 4. Setting $U_M^{gov}(m) = U_T^{gov} + \Psi$, $\Psi > 0$, shows that there is no $\Psi > 0$ which can permit $U_M^L(m) \leq U_T^L$.

REFERENCES:

1. Akerlof, George A. “The Economics of ‘Tagging’ as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning” *American Economic Review*, 68 (1), 1978, 8-19.

2. Allingham, Michael and Agnar Sandmo. “Income Tax Evasion: A Theoretical Analysis”, *Journal of Public Economics*, 29, 1972, 323-338.

3. Beck, Paul J., John S. Davis and Woon-oh Jung. “Taxpayer Disclosure and Penalty Laws”, *Journal of Public Economic Theory*, 2 (2), 2000.

4. Besley, Timothy. “Means Testing Versus Universal Provision in Poverty Alleviation Programmes,” *Economics*, 57, 1990, 119-129.

5. Boadway, Robin and Motohiro Sato (2000), “The Optimality of Punishing Only the Innocent: The Case of Tax Evasion”, *International Tax and Public Finance*, 7, 641-664.

6. Brin, David. *The transparent society: Will technology force us to choose between privacy and freedom?*. Addison Wesley, 1998.

7. Chander, Parkash and Louis L. Wilde (1998), “A General Characterization of Optimal Income Tax Enforcement”, *Review of Economic Studies*, 65, 165-183.

8. Connerley, Mary L. et. al. (1999), “ ‘Don’t Ask — Please Tell’: Selection Privacy from Two Perspectives”, *Journal of Occupational and Organizational Psychology*, 72, 405-422.

9. Crabb, Peter B. (1999), "The Use of Answering Machines and Caller ID to Regulate Home Privacy", *Environment and Behaviour*, 31 (5), 657-670.

10. De Donder, Phillipe and Jean Hindriks (1998), "The Political Economy of Targeting", *Public Choice*, 95, 177-200.

11. Etzioni, Amitai. *The Limits of Privacy.* Basic Books, 1999.

12. Fleurbaey, Marc and Francois Maniquet, "Compensation and Responsibility", Mimeo, 1999.

13. Garfinkel, Simson. *Database Nation: The death of privacy in the 21st Century.* O'Reilly, 2000.

14. Gilliom, John (1994), *Surveillance, Privacy and the Law: Employee Drug Testing and the Politics of Social Control*, Ann Arbor: University of Michigan Press.

15. Immonen, Ritva et. al. "Tagging and Taxing: The Optimal Use of Categorical and Income Information in Designing Tax/Transfer Schemes", *Economica*, 65, 1998, 179-92.

16. Jacquet, Laurence and Bruno van der Linden (2001), "The normative analysis of 'Tagging' revisited: Dealing with Stigmatization", Unpublished Manuscript.

17. Konrad, Kai A. "Privacy and time-consistent optimal labour income taxation", *Journal of Public Economics*, 79, 2001, 503-519.

18. Lindbeck, Assar. "Individual Freedom and Welfare State Policy", *European Economic Review*, 32, 1988, 295-318.

19. Lindbeck, Assar, Nyberg, Sten and Jorgen W. Weibull (1999), "Social Norms and Economic Incentives in the Welfare State", *Quarterly Journal of Economics*, 114 (1), 1-35.

20. Lyon, David. *Surveillance Society: Monitoring Everyday Life.* Open University, 2001.

21. Morris, Stephen (2001), "Political Correctness", *Journal of Political Economy*, 109 (2), 231-265.

22. Parsons, Donald O. "Imperfect 'tagging' in social insurance programs", *Journal of Public Economics*, 62, 1996, 183-207.

23. Posner, Richard. "The Economics of Privacy", *American Economic Review*, 71 (2), 1981, 405-409.

24. Roemer, John E. *Equality of Opportunity.* Harvard University Press, 1998.

25. Young, John B. (ed.) *Privacy.* John Wiley, 1978.

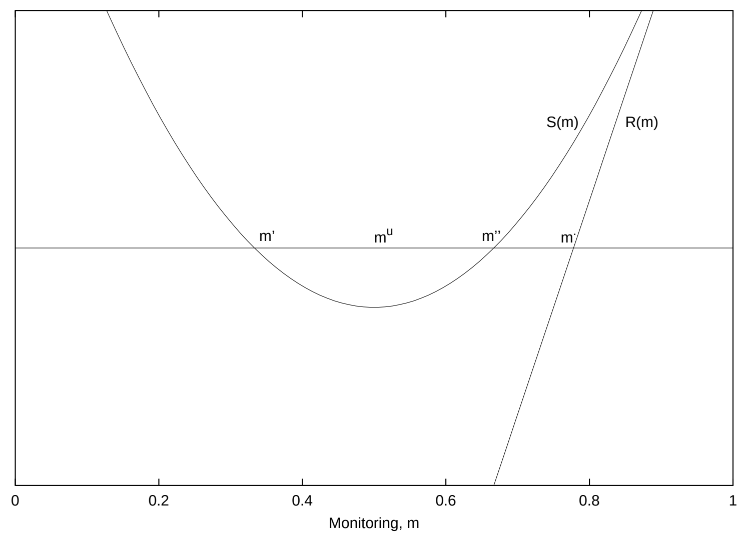

Figure 1: Best Responses: $R(m) > 0$ means $\hat{H}$ is the best response of type H to $\hat{H}\hat{L}$; $S(m) > 0$ means $\hat{H}$ is the best response of type H to $\hat{L}\hat{L}$. This figure is drawn using the numerical values from Example 1.

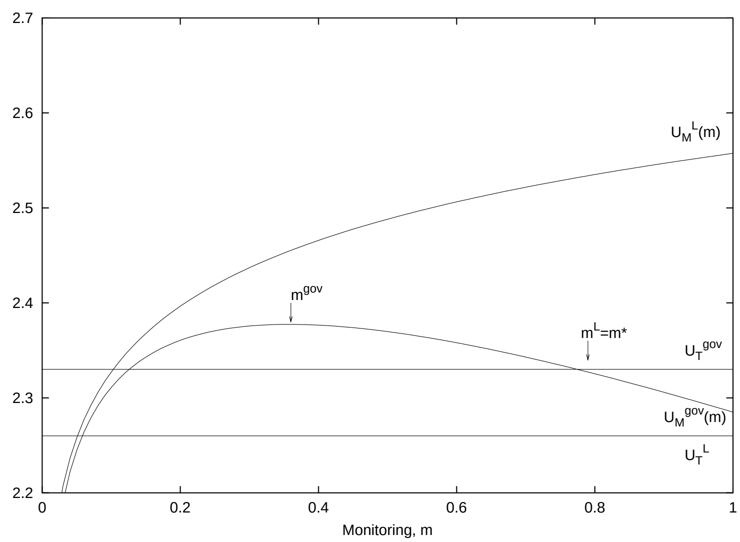

Figure 2: Example 3, Low types in the majority.

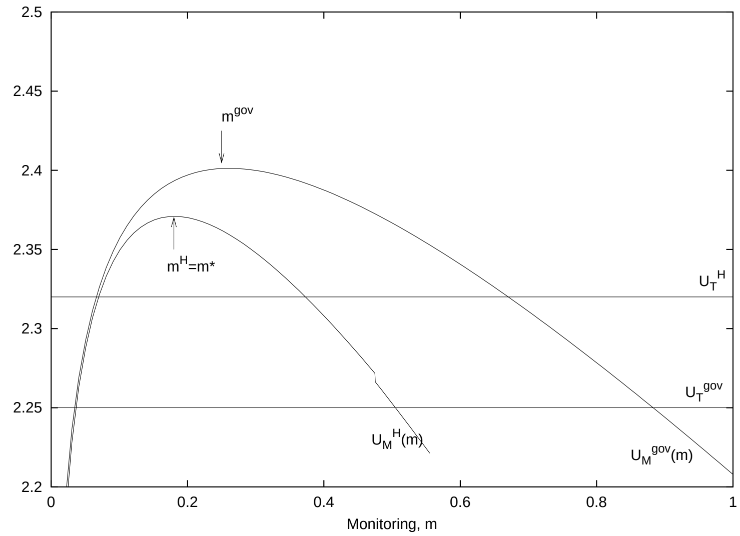

Figure 3: Example 4, High types in the majority.

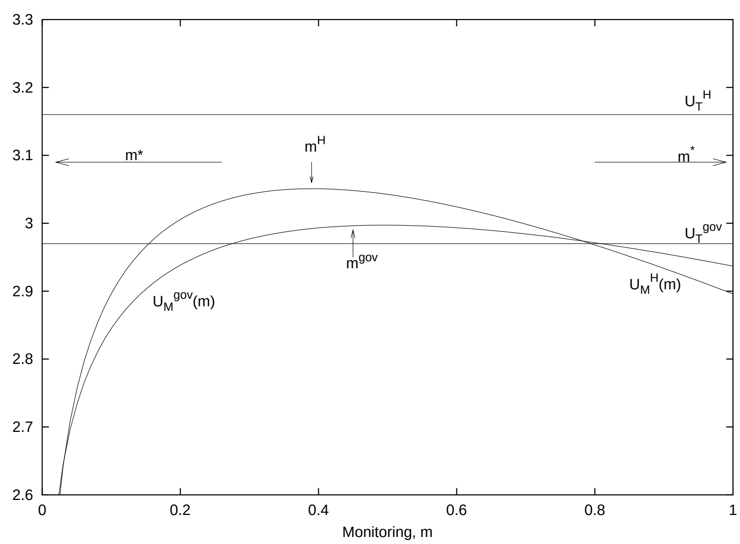

Figure 4: Example 5, High types in the majority.