Confidentiality Rules and Welfare: A Dynamic Contracting Approach with Two Principals[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, D23, K12

January 1, 2003

Abstract:

Individual characteristics are indirectly revealed in mechanisms where agents self-select from a menu of options set by a principal. These (revealed) characteristics are potentially valuable to other principals if they are correlated with economic variables of interest. This paper uses a dynamic contracting model — in which agents contract sequentially with two principals involved in distinct activities — to examine the incentives for the "upstream" principal to keep agents' contract-choices confidential. Perfect Bayesian Equilibria are characterized when the principals cannot commit to confidentiality or long-term contracts. If agents are farsighted and sufficiently patient, (i) *principals* seek confidentiality legislation as a commitment device, and (ii) *agents* as a whole do strictly worse with a confidentiality law. This result is contrary to the popular view that confidentiality legislation is enacted to protect the interests of those who *provide* information. The paper also derives conditions when the sharing of information maximizes total surplus: these results have policy implications for the implementation of confidentiality rules in the public and private sectors.

[1]Comments are very welcome. Thanks to Robin Boadway, Sumon Majumdar and seminar participants at Queen's University for useful comments.

> "Much of the modern concern with the right to privacy arises in these consensual situations. Information is easily divisible and reproducible; unlike land it can be given away with one hand and retained by the other." (Richard Epstein, 2000, p.21)

1 Introduction

Much debate on personal privacy has in fact focussed on what might be more appropriately called confidentiality. The two terms can be differentiated in the following way: privacy refers to the ability of individuals to prevent certain organizations — firms, governments, and other individuals — from directly observing their personal information without their consent; confidentiality refers to their ability to restrict the dissemination of this information once it has been revealed to some organization (consensually or otherwise). For example, individuals' income-information is not "private" per se: it must be reported to a central tax authority which also has the power to audit the report for its accuracy. However, individuals often expect that information they report to the tax authority will remain confidential; i.e. that it will not be gathered and then distributed to firms, society at large or even other branches of government.[1]

Indeed, this expectation is often codified legally. The past 25 years has seen the advent of "privacy" legislation in most developed nations.[2] Such laws frequently encode aspects of what Bennett (1992, 2001) and others have termed the "Fair Information Principles Doctrine" (FIPD), a set of normative information-policy guidelines for organizations that gather personal information. Among others, the doctrine includes the following tenets:

- An organization should not use or disclose personal information for purposes other than those identified, except with the consent of the individual.
- An organization should be open about its policies and practices and maintain no secret information system.

The implicit rationale behind the FIPD is that it protects information-providers (individuals) from the self-interested actions of information-gatherers (organizations). Le-

[1] The U.S. Tax Reform Act of 1976 "protects the confidentiality of tax returns and return-related information and limits the dissemination of individual tax return data to other federal agencies" (Valentine, 2000).

[2] See Bennett (2001), Appendix 1, for a summary of dates and nations where such legislation has been enacted. Sweden instituted a "privacy law" as early as 1973, the United States in 1974, and Canada in 1982. More recently have been South Korea (1994) and Latvia (2000).

gal sanctions therefore act as devices which *commit* organizations to a certain set of rules about how individual information may be shared. There is a wide variety of information-policy laws between jurisdictions. For example, legislation may cover the information-sharing activities of government and/or those of the private-sector as well. North American privacy legislation has mostly focused on the public sector, while European legislation typically covers both the public and private sectors (Valentine, 2000). However, the increasing ease with which electronic databases of information can be established, and the simplicity of transferring this information between organizations, have led North American legislators to consider laws governing private-sector information practices generally, and confidentiality in particular.[3]

This paper asks two related questions: first, when are confidentiality laws welfare-enhancing? and second, who benefits from such laws? Observation of the privacy debate suggests that individuals stand primarily to benefit from confidentiality (see Bennett, 1992). One (trivially) obvious explanation for the rise of confidentiality legislation is that benevolent governments act on behalf of individuals, who care inherently about their information "getting out". They simply prefer that others know as little as possible about them as possible. An opposite explanation lends itself more to economic interpretation, and is intuitively understandable. In this view, individuals recognize that their information is valuable because its knowledge allows other economic agents to exploit it strategically in the future. Thus, in asymmetric-information settings where the truthful

[3]In Canada, the Personal Information Protection and Electronic Documents Act (Bill C-6) was enacted in 2001 to broaden the original Privacy Act of 1982. As of January 2002, it has prohibited the sharing of health information in the private sector without the consent of those from whom the information has been gathered; by 2004, the Act will also prevent unconsented dissemination of information gathered in the commercial sector. In the United States, laws dealing with confidentiality are varied across states and areas of coverage (e.g. tax information, credit information, retail data), while federal laws have focused on information usage in the public sector (Bennett, 1992; Valentine, 2000). One exception is Title V of the Gramm-Leach-Bliley Act of 1999, which stipulates that private firms which collect *financial information* (e.g. banking, credit, leasing), must allow customers to "opt out" of any scheme in which their information is shared with other firms. However, the law is sufficiently incomplete to permit a good deal of information-dissemination *without* consent. A U.S. Federal Trade Commission document, "Privacy Choices for your Personal Financial Information" (http://www.ftc.gov/bcp/conline/pubs/credit/privchoices.htm) reports:

> ...you cannot opt out and completely stop the flow of all your personal financial information ...Among other things, your financial company can provide to non-affiliates ...records of your transactions — such as your loan payments, credit card or debit card purchases, and checking and savings account statements — to firms that provide data processing and mailing services for your company.

Contrary to the European (and recently, Canadian) approaches to confidentiality rules, which have attempted to broaden the use of "fair information practices" with widely-applicable laws, the American response has been to legislate rules for specific sectors on a piecemeal basis (Valentine, 2000).

revelation of an agent's information (for example, "type") can only be achieved by the surrender of information rents by a principal, individuals may seek to preserve the rents they are allocated by keeping information asymmetric. Confidentiality rules may then be a barrier to the "appropriation" of rents when an agent's true information is known to others with greater probability.[4] [5]

Yet in many economic settings, sharing information can *expand* total rents. This is true, for example, when firms can extract more output from workers they discover to be of high ability. If the gains to sharing information accrue mostly to the organizations that collect and disseminate it, then a clear incentive exists for those who provide information to lobby for confidentiality regulations. On the other hand, if those who provide information can be compensated for giving up this information, they may then be unconcerned about its release if doing so increases their total welfare.

One purpose of this paper is to characterize the behaviour of information-gathering organizations *independent of legislation.* A criterion for the imposition of confidentiality laws is that they should increase aggregate welfare from the case where there are no such laws. If organizations can improve aggregate welfare without legislation, then the legislation is redundant from a purely utilitarian standpoint. A crucial assumption in the model below is that an organization which indirectly collects information from individuals cannot commit to confidentiality without legislation. In addition, consent by agents is modeled as an "up front" process: dealing with an organization implies that an individual consents (trivially) to the sharing of her information.[6] The equilibrium results in this environment are somewhat counterintuitive: confidentiality laws (a) favour information gatherers over information providers, and (b) are actually welfare-decreasing

[4]For example, suppose a buyer is willing to pay a maximum of \$90 for a certain good both today and tomorrow, and this information is private. Seller X can sell the good today, seller Y can sell it tomorrow, and each knows that the distribution of the individuals' willingness to pay, v, is distributed uniformly between \$0 and \$100. If the good is has zero cost, X will offer to sell it for \$50 today, and the buyer will receive a surplus of \$40. But tomorrow, if Y can observe the price of the sale today, he will optimally offer to sell for \$75 since he now knows that the buyer's valuation is at least \$50. The buyer will then accept and receive a surplus of \$15. If the history of sales were confidential,the buyer would have received the \$40 surplus again, since seller Y's problem would have been identical to that of seller X. In this example, there is a fixed "pie" that buyer and seller divide. Therefore, the outcome of the sale is efficient whether information is revealed after the first sale or not, so long as the good is actually sold. Taylor (2002) constructs a dynamic pricing model where information can be resold and evaluates the welfare losses of such a model in terms of foregone sales.

[5]Epstein (2002) argues that personal information is akin to private property; hence, uncompensated confiscation of this information can be seen as a "taking".

[6]i.e. The right to an individual's information is transferred to an organization as soon as an individual becomes involved with that organization. This assumption is relaxed in Section 7.

when providers tend to discount the future heavily.

The model outlined below can be applied to the dissemination of information between firms, between government agencies, or between the government and the private sector. For example, it can offer a rationale for why information between government departments may not be shared, reinforcing arguments for dividing government functions into separate departments.[7]

For simplicity, the model is based purely on adverse-selection. Agents of different bi-dimensional types contract first with a principal concerned solely about element x of the agent's type. That is, x is important to the payoff of this principal; this element may be, for example, a productivity level or a taste parameter of the agent. As in the screening literature, the choice of a specific contract reveals information about the individual; when agents are fully separated, their choice of contract is an indirect (but perfect) indicator of the agent's x-type. Agents can then contract separately with another principal concerned about the element y of an agent's information set. When x and y are correlated, information from the first stage of contracting is useful to the second principal since it permits the identification of an agent's y-type with greater certainty. The possibility of greater rents generates a "price" for the information about first-period contracts, but also distorts the incentive of agents to reveal information in the first period by their choice of contract.[8] In a regime of "no confidentiality" (sharing permitted by law), information may be sold between principals to exploit the greater efficiencies associated with sharing information. However, if agents are sophisticated enough to anticipate information-sharing, they must be left *extra* rents to reveal their type truthfully in the first period. Some agents have an incentive to choose the "wrong" contract from the first principal, since they may anticipate a higher payoff in the future once their choice of contract is revealed. In a regime of confidentiality where information sharing is prohibited, agents need not be extra-compensated for truthful revelation in the first stage of contracting, although the efficiencies associated with information-sharing will then not be exploited either. A utilitarian law-maker will not want to impose confidentiality in the case where the first principal nevertheless finds it optimal (in the

[7] A related argument for "separation" is outlined in Laffont and Matrimort (1997).

[8] This is an example of the "ratchet effect"; see Freixas, Guesnerie and Tirole (1985), Laffont and Tirole (1988), Dewatripont and Maskin, (1990). These papers consider principal agent settings with a single principal and repeated tasks, whereas the current study models two distinct principals conducting different tasks. The division of this model into two principals is somewhat arbitrary (as Taylor, 2002, notes); however, a single principal could set bidimensional screening contracts as described by Armstrong and Rochet (1999). This possibility is not modeled here because of the focus on information-sharing.

no-commitment case) to compensate agents to choose the "right" contract.

A novel aspect of this model is that it explores the incentives for the first principal to protect information. When agents discount the future to the same extent as the first principal, that principal would always like to have a commitment device not to share information. By assumption, the only such device is a confidentiality law. However, if agents discount the future relatively heavily, the first principal may not want such a device. In the absence of a confidentiality law, such a principal may wish to exaggerate *when* it may share information to persuade agents to accept lower levels of compensation for revealing their types ex ante.

This paper is close in spirit to two recent studies. Taylor (2002) builds a dynamic selling model in which sellers may exchange information regarding buyers' purchasing histories, also in a no-commitment framework. Buyers have uncertain (but correlated) valuations for different goods sold in two periods. He finds similar welfare effects from allowing the sale of information as in this paper (e.g. the desire of the first principal to commit to confidentiality), although welfare losses are characterized in terms of deadweight-losses from foregone sales. By contrast, this paper looks at a screening framework where principals can offer *menus* of contracts which can potentially augment the size of available economic rents. Calzolari and Pavan (2002) characterize the optimal information-sharing policies — adopted by principals — in a similar setting where distinct principals sequentially set contract menus for agents. The main difference of that study is the assumption that both principals can fully commit to contract schemes and disclosure (confidentiality) policies, which introduces more strategic interaction between the principals. They also assume that agents' types are perfectly correlated between the payoff-relevant variables of each principal.

The issue of time consistency and ratcheting in repeated principal-agent settings has long been studied (e.g. Freixas, Guesnerie and Tirole (1985), Laffont and Tirole (1988)). More recent work in public economics has reexamined this issue in terms of optimal government policy. For example, Dillen and Lundholm (1996) and Konrad (2001) both find that governments can optimally increase welfare by "throwing away" past information about agents, or committing not to observe agents' past decisions. Thum and Thum (2001) examine the optimality of using different transfer instruments, where agents anticipate the use of their revealed choices of in-kind transfers in future government policy. Each of these studies examines a single government's behaviour over time. By contrast, this model may be extended to look at the optimality of sharing information between different branches of government.

Section 2 sets up the basic model of the paper and describes the nature of the equilibrium. The equilibrium with "no commitment" is solved in Section 3. Section 4 compares the welfare of principals and agents in the no-commitment equilibrium with that under a confidentiality law. Section 5 extends the model to include the case where agents discount the future, and a numerical example is given in Section 6. Section 7 discusses the role of consent in the model. Section 8 concludes.

2 The Basic Model (No Confidentiality)

To study the incentives for information-sharing between organizations, consider a model in which many agents contract separately and sequentially with two principals called X and Y. Generally speaking, agents can be described by their personal characteristics ("types", including physical attributes, tastes, abilities, income level, etc.), and histories of behaviour (past actions and choices). Here, every agent is described by two elements, x and y, which jointly represent an agent's type. Agents know their own types, but this information is inherently *unobservable* by others, including the principals. In this sense agents' information is private. Agent i's type is described by the couplet (x^i, y^i), where $x^i \in \{\overline{x}, \underline{x}\}$ and $y^i \in \{\overline{y}, \underline{y}\}$. Thus, there are four possible types of agent: $(\overline{x}, \overline{y}), (\overline{x}, \underline{y}), (\underline{x}, \overline{y}), (\underline{x}, \underline{y})$.

2.1 The structure of the problem

The joint distribution of (x, y) across agents is commonly known. Let the joint distribution of $\{x, y\}$ be symmetric, so that:

$$\Pr(x = \overline{x}, y = \overline{y}) = \alpha$$

$$\Pr(x = \overline{x}, y = \underline{y}) = \Pr(x = \underline{x}, y = \overline{y}) = \beta$$

$$\Pr(x = \underline{x}, y = \underline{y}) = \gamma$$

The symmetry of the distribution implies that the marginal probabilities are also symmetric for x and y:

$$\Pr(x = \overline{x}) = \Pr(y = \overline{y}) = f = \alpha + \beta$$

$$\Pr(x = \underline{x}) = \Pr(y = \underline{y}) = 1 - f = \gamma + \beta$$

Conditional probabilities are then given by:

$\Pr(y = \overline{y} | x = \overline{x}) = \frac{\alpha}{f}$

$\Pr(y = \underline{y} | x = \overline{x}) = \frac{\beta}{f}$

$\Pr(y = \overline{y} | x = \underline{x}) = \frac{\beta}{1-f}$

$\Pr(y = \underline{y} | x = \underline{x}) = \frac{\gamma}{1-f}$

and similarly for $\Pr(x|y)$. After observing information about x, one can update with more precision beliefs about the value of y. Here, it is assumed that x and y are positively correlated[9]. If correlation, ρ, is positive:

$$\rho = \alpha\gamma - \beta^2 > 0 \tag{1}$$

Note that positive correlation might still allow either α *or* γ to be less than β.[10] Given the assumption of symmetry, and the requirement that $\alpha + 2\beta + \gamma = 1$, the distribution can be characterized in terms of α and f alone. Specifically, $\beta = f - \alpha$ and $\gamma = (1 - f) - \beta = 1 - 2f + \alpha$. Then

$$\rho > 0 \Leftrightarrow \alpha\gamma - \beta^2 > 0 \Leftrightarrow \alpha - f^2 > 0 \Leftrightarrow f < \sqrt{\alpha} \tag{2}$$

is required for positive correlation. Suppose one fixes $\alpha = \Pr(x = \overline{x}, y = \overline{y}) = 0.25$; then, $f < 0.5$. It is obvious that, given α, a higher f implies lower correlation since β grows. Similarly, given f, a higher α implies greater correlation since β shrinks.

The problem of each principal is to offer the *set* of contracts to agents which maximizes that principal's expected payoffs. As stated earlier, each principal is only concerned with one aspect of the agents' type[11]. A natural model to examine is one in which each principal hires agents to perform a discrete task on his behalf. For ease of exposition, suppose that principal X hires agents to produce some amount of a good z_x. The return to the production of z_x for principal X is $r_x(z_x)$, with $r' > 0$ and $r'' \leq 0$.

[9] Positive correlation is arbitrary in the model: the crucial assumption is that x is a useful but imperfect indicator of y, and vice versa.

[10] For example, $\rho > 0$ can still hold with $\Pr(y = \overline{y}, x = \overline{x}) < \Pr(y = \underline{y}, x = \overline{x})$ — and hence $\Pr(y = \overline{y} | x = \overline{x}) < \Pr(y = \underline{y}, x = \overline{x})$ — so long as $\Pr(x = \underline{x}, y = \underline{y}) > \Pr(x = \overline{x}, y = \underline{y})$.

[11] Under this assumption, principal X cannot contract over both x and y dimensions. This situation might be thought of as two firms producing very distinct goods, or two levels of government concerned with different public programs.

The contract payment made to an agent in return for z_x is s_x. Along the same lines, let principal Y hire agents to produce some amount of a different good, z_y. Y pays agents s_y in return for z_y, but now let $r_y(z_y) = \omega r_x(z_y)$. Here, ω is a parameter describing the difference between X and Y's returns to the production of a unit of the respective goods, z_x and z_y. This is a rough measure of the relative value of production between X and Y. When $\omega > 1$, we can say that Y earns a higher return per unit of z_y than does X per unit of z_x.

The variables x and y refer to the abilities of the agents at producing goods z_x and z_y, respectively. An agent with $x = \overline{x}$, (resp. $y = \overline{y}$) has the cost function $c_{\overline{x}}(z)$ (resp. $c_{\overline{y}}(z)$), whereas an agent with $x = \underline{x}$, (resp. $y = \underline{y}$) has the cost function $c_{\underline{x}}(z)$ (resp. $c_{\underline{y}}(z)$). For simplicity, suppose that the cost functions $c_x(\cdot) = c_y(\cdot)$, so that a $\overline{x}$ (resp. $\underline{x}$) type is equally adept at producing good z_x as is a $\overline{y}$ (resp. $\underline{y}$) at producing good z_y. Calling an $\bar{\cdot}$ type "high ability" leads to the assumption that $c_{\overline{x}}(z) < c_{\underline{x}}(z)$ and $c_{\overline{y}}(z) < c_{\underline{y}}(z)$. Furthermore, let $c' > 0$ and $c'' \geq 0$ for all x and y.

It is assumed that the return functions $r_x(\cdot)$, $r_y(\cdot)$ and cost functions $c_{\bar{\cdot}}(\cdot)$, $c_{\underline{\cdot}}(\cdot)$ are commonly known. To act as a benchmark case, let there initially be no discounting on the part of principals or agents. This assumption will be relaxed in section 5.

2.2 Timing of Decisions

Because x and y are correlated, the results of principal X's contracts with the agents are potentially useful to Y when the latter is designing his own contracts. Confidentiality, however, implies that X prevents Y from discovering at least the identities of agents signing particular contracts. As suggested earlier, confidentiality might be enforced through legislation, or, if credible, by the commitment of principle X to voluntarily enact a policy of no information-sharing.

If principal X has an incentive to voluntarily commit to confidentiality, then it must be the case that commitment can only improve his payoff[12]. The initial assumption is that X *cannot* commit to a confidentiality policy, and there are no laws in place which prohibit information-sharing. For this reason, the model is constructed with sequentially rational players, so that every actor (principal or agent) at each node of the decision tree makes the best possible decision based on his current information set and expected

[12] Laffont and Tirole (1990) point out that, in repeated principal-agent settings, commitment is generally never harmful for a principal. In the present setting, committing to confidentiality might be sub-optimal depending on the ex-ante behavioural response of agents to the possibility that their information will be shared.

future payoffs. That is, a player cannot commit to use actions further down the tree which may be suboptimal when that node is actually reached.

The timing of the model is as follows:

Stage 1. Principal X announces a menu of contracts, $(\mathbf{z_x}, \mathbf{s_x})$ which he offers to all agents. Assume that the there are K such contracts, $k = 1, 2 \ldots K$, so that $(\mathbf{z_x}, \mathbf{s_x}) \equiv \{(z_{x1}, s_{x1}), (z_{x2}, s_{x2}), \ldots, (z_{xK}, s_{xK})\} \equiv \{(z_x, s_x)^1, (z_x, s_x)^2, \ldots, (z_x, s_x)^K\}$. To save notation, assume that one of the contracts offered is a "trivial" one: $(0, 0)$. This option may be interpreted as "no contract", since there is no production and no payment.

Stage 2. Each agent i chooses his or her preferred contract, $(z_x, s_x)^i$, from the menu $(\mathbf{z_x}, \mathbf{s_x})$. Each may also choose not to sign a contract, in which case we write that $(z_x, s_x)^i = (0, 0)$.

Stage 3. Principal X decides whether or not to sell his history of contract information to Y. If he decides not to sell (NS), the game moves directly to stage 5. If he decides to sell (S), X compiles a list of contracts he has offered, $(\mathbf{z_x}, \mathbf{s_x})$ as well as aggregate data on the proportions of agents choosing each contract offered. Denote the proportion of agents choosing contract $(z_x, s_x)^k$ by $\eta^X(k)$, where $\eta^X(\mathbf{z_x}, \mathbf{s_x}) \equiv \{\eta^X(1), \eta^X(2), \ldots, \eta^X(K)\}$ lists the proportion of agents choosing each contract offered. Assume that $(\mathbf{z_x}, \mathbf{s_x})$ and $\eta^X(\mathbf{z_x}, \mathbf{s_x})$ are verifiable to Principal Y.[13] X then makes a "take it or leave it" offer of p to Y for the rights to match particular contracts, $(z_x, s_x)^k$, with particular agents, i, who have chosen them.

Stage 4. Given X's offer of p, the list of contracts offered, $(\mathbf{z_x}, \mathbf{s_x})$, and the proportions who have chosen them, $\eta^X(\mathbf{z_x}, \mathbf{s_x})$, Y decides whether or not to buy the information. If she buys (B), she is able to match agents' identities with particular contracts. If she does not buy (NB), Y remains ignorant of who chose which contract in stage 2.

Stage 5. Principal Y sets an overall menu of $l = 1, 2, \ldots, L$ contracts: $(\mathbf{z_y}, \mathbf{s_y})$. Since Y can use any information she gains in stage 4 when designing the contracts she offers, she can restrict the contracts offered to agents who have been "tagged" according to their choice of X-contract. Assume in general that an agent who has chosen the k^{th} contract

[13]This assumption rules out cases where X lies about the results of contracting to improve his selling price.

at the X-stage is offered a sub-menu containing M_k contracts at the Y stage. Thus, an agent identified as having chosen contract $(z_x, s_x)^k$ is offered the *menu* $(\mathbf{z_y}, \mathbf{s_y})^k = \{(z_{yk}, s_{yk})^1, \ldots, (z_{yk}, s_{yk})^{M_k}\}$. This requires that the union of all contracts offered to each of the k types is the overall set of L contracts:

$$\bigcup_{k=1}^{K} (\mathbf{z_y}, \mathbf{s_y})^k = (\mathbf{z_y}, \mathbf{s_y}) \equiv \{(z_y, s_y)^1, (z_y, s_y)^2, \ldots, (z_y, s_y)^L\}$$

It can happen that each of the k types is offered the entire set of contracts — the most obvious case where this happens is where Y has not purchased information and so does not know which X contracts specific agents have chosen.[14] Therefore, it is useful to write down some notation for what Y comes to know at stage 5. Call Y's information set in general Ω, where:

$$\Omega = \Omega_b \equiv \{\alpha, f, (\mathbf{z_x}, \mathbf{s_x}), \eta^X(\mathbf{z_x}, \mathbf{s_x}), \text{ and } (z_x, s_x)^i \ \forall \ i\}$$

if X sells and Y buys information,

$$\Omega = \Omega_e \equiv \{\alpha, f, (\mathbf{z_x}, \mathbf{s_x}), \eta^X(\mathbf{z_x}, \mathbf{s_x})\}$$

if X sells (exhibits) information but Y does not buy, and

$$\Omega = \Omega_o \equiv \{\alpha, f\}$$

if X does not sell information. Ω_o is therefore the "prior" information set. The basic advantage to purchasing information is the ability to "tag" individuals according to the X-contract they have accepted. This includes the case where agents have not accepted any contract from the menu offered; i.e. $(z_x, s_x)^i = (0, 0)$.

Stage 6. Each agent i, having chosen contract k at the X-stage, chooses his or her preferred contract $(z_y, s_y)^i$ from the menu he or she is offered, $(\mathbf{z_y}, \mathbf{s_y})^k$. They may also choose not to sign a contract, in which case $(z_y, s_y)^i = (0, 0)$.

[14] Y is also free to set $M_k = 0$ for some k; i.e. to not offer any contract to agents choosing contract $(z_x, s_x)^k$ from X.

2.3 Payoffs

To provide a benchmark case, assume that the payoffs to each player in the sequential game outlined above are undiscounted. Call agents' final utilities U_{xy}, which depend on their type (x, y) and the contracts they *actually* sign in stages 2 and 6. An agent's single-period payoff from the X-contract is u_{xy}^X, and that from the Y-contract is u_{xy}^Y. If an agent eventually chooses a contract (z_x, s_x) from X and a contract (z_y, s_y) from Y, then

$$U_{xy}(z, s) = u_{xy}^X(z_x, s_x) + u_{xy}^Y(z_y, s_y) \tag{3}$$

where

$$\begin{aligned} u_{xy}^X(z_x, s_x) &= s_x - c_x(z_x) \\ u_{xy}^Y(z_y, s_y) &= s_y - c_y(z_y). \end{aligned}$$

Let $u_{xy}^X(0,0) = u_{xy}^Y(0,0) = 0$, which implies $c_x(0) = c_y(0) = 0$. At stage 2, however, agents are uncertain of the set of contracts they will be offered, or the particular one (if any) they will sign at stage 6. Because Y can buy information from X, the X-contract an agent signs can affect the Y-contracts that an agent is offered at stage 5. Depending on Y's information set, Ω, the menu of contracts offered may change. Therefore, the expected payoff of an agent who chooses contract $(z_x, s_x)^k$ from X is:

$$EU_{xy}(z, s) = u_{xy}^X((z_x, s_x)^k) + Eu_{xy}^Y((z_{yk}, s_{yk})^{m_k} | (z_x, s_x)^k, \Omega_b) \tag{4}$$

where m_k denotes the contract an agent purchases from the submenu with M_k options.

Let the payoff to principal X be Π_X. This payoff depends upon the (direct) expected return he receives from the contract-menu he designs, $E\pi$, as well as the funds he can raise by selling information about this contract-menu to Y. The price X sets — if he sells — depends in turn on the value of the information that Y gets from X. Letting $p(\Omega_b)$ be the price at which X can sell information to Y, X's expected payoff when designing the contract menu $(\mathbf{z_x}, \mathbf{s_x})$ is:

$$\begin{aligned} E\Pi_X(\mathbf{z_x}, \mathbf{s_x}) &= E\pi_X(\mathbf{z_x}, \mathbf{s_x}) + Ep(\Omega_b) \\ &= \sum_{k=1}^{K} [r_x(z_{xk}) - s_{xk}] \eta^X(k) + Ep(\Omega_b). \end{aligned} \tag{5}$$

Similarly, denote the payoff to Y by Π_Y. If X sells and Y buys information, she does so at the price X offers, $p(\Omega_b)$. Given Ω_n — which Y knows whether or not she purchases information, given that X sells — Y can infer the distribution of agents' y-types among X-contracts that have been signed, and hence the potential value of purchased information. Write Y's *posterior beliefs* about the distribution of y-types as $\Phi(\Omega_e)$[15]. Crucially, Y *can only act upon her beliefs if she buys information.* Therefore, Y's "useful" beliefs when she does not buy information are the same as her prior beliefs, which (abusing notation slightly) are denoted here by Φ_o. Then noting that Y treats $p(\Omega_b)$ as a parameter of her problem, one can write:

$$\begin{aligned} E\Pi_Y(\mathbf{z_y}(\Phi), \mathbf{s_y}(\Phi)) &= E\pi_Y(\mathbf{z_y}(\Phi), \mathbf{s_y}(\Phi)) - p \\ &= \sum_{l=1}^{L} [r_y(z_{yl}(\Phi)) - s_{xl}(\Phi)]\eta^Y(l) - p \end{aligned} \tag{6}$$

when X sells and Y buys information, and

$$\begin{aligned} E\Pi_Y(\mathbf{z_y}(\Phi_o), \mathbf{s_y}(\Phi_o)) &= E\pi_Y(\mathbf{z_y}(\Phi_o), \mathbf{s_y}(\Phi_o)) \\ &= \sum_{l'=1}^{L'} [r_y(z_{yl'}(\Phi_o)) - s_{xl'}(\Phi_o)]\eta^Y(l') \end{aligned} \tag{7}$$

when Y does not buy information, regardless of whether X was willing to sell or not. (Y may offer different overall numbers of contracts depending on her information set.) $\eta^Y(l)$ is the overall proportion of all agents who choose contract $l \in (\mathbf{z_y}, \mathbf{s_y})$.

2.4 Strategies and Definition of Equilibrium

The possible strategies for an agent i of type (xy) are summarized by $\sigma^i_{xy} \equiv \{(z_x, s_x)^i, (z_y, s_y)^i\}$. That is, given that an agent is offered two menus of contracts (sequentially), that agent selects exactly one from each menu. Principal X's strategy profile is $\sigma_X = \{(\mathbf{z_x}, \mathbf{s_x}), (S, NS), p\}$ and principal Y's is $\sigma_Y = \{(\mathbf{z_y}(\mathbf{\Phi}), \mathbf{s_y}(\mathbf{\Phi})), (B, NB)\}$. Each principal can offer a large and varied set of contracts over (z, s), and principal X is free to set any p he wants, including negative offers where X compensates Y for taking his information. Let $\sigma = (\sigma^i_{xy}, \sigma_X, \sigma_Y)$ be the profile of all strategies.

The equilibrium concept is that of Perfect Bayesian Equilibrium: (a) the strategy

[15]Since Ω_e is all that is required for Y to update her beliefs, and Y always comes to know Ω_n when X chooses to sell, the argument Ω_e is dropped to avoid clutter.

profile of each player is sequentionally rational, given the players' beliefs of being at a specific node in their information sets, and (b) the players' beliefs are derived from Bayes' rule where possible. In the present framework, principal Y is the crucial player. When X sells information, Y always comes to know at least Ω_e and can update her beliefs to $\Phi(\Omega_e)$ using Bayes' Rule. These beliefs describe the probability Y assigns to the *actual* y-type of agents, given the X-contracts they have been offered and those which they have selected. (Again, Y must purchase information to actually "tag" agents according to their X-contract.) Sequential rationality then requires that Y offer $(\mathbf{z_y}(\mathbf{\Phi}), \mathbf{s_y}(\mathbf{\Phi}))$ or $(\mathbf{z_y}(\mathbf{\Phi_o}), \mathbf{s_y}(\mathbf{\Phi_o}))$ optimally. In any equilibrium it must be the case that agents would not choose different contracts in Stage 2, and that principal X would not offer a different menu in Stage 1. What Y must surmise for every individual who contracts at this stage is the probability that that individual is a y type. Recall that each agent chooses one of K contracts offered by X; specifically, $(z_x, s_x)^k$. When X sells and Y buys information, Y's posterior belief about the y-type of a particular agent who has chosen contract k is $\Phi_k = Pr(\overline{y}|(z_x, s_x)^k, \Omega_b)$. When Y doesn't buy information, she maintains the prior belief about any particular agent: $\Phi_o = Pr(\overline{y}|\Omega_o) = f$.

Formally, for strategy profile $\sigma^* = \{\sigma_{xy}^{i*}, \sigma_X^*, \sigma_Y^*\}$ to be a Perfect Bayesian equilibrium,

1. σ^* must be sequentionally rational, given Φ, for each agent i, and principals X and Y.
2. $\Phi_k = Pr(\overline{y}|(z_x, s_x)^k, \Omega)$, from Bayes' Rule, when Y buys information.
3. Φ is the believed distribution of $\overline{y}$, given Φ_k for all k.

3 The Equilibrium of the No-Confidentiality Model

In this section, the solution to the model at hand is characterized. As a benchmark, the section 3.1 briefly describes the solution to the standard principal-agent model with a single period. Section 3.2 describes the solution to the dynamic problem outlined above.

3.1 Solution to the basic single-period problem with adverse selection

The results of a single-period adverse selection problem in a principal-agent setting are straightforward.[16] Let ε be the variable of interest with agents having types $\overline{\varepsilon}$ with

[16] See Laffont and Matrimort (2002) for details.

probability f and $\underline{\varepsilon}$ with probability $(1-f)$. Let the returns to the principal be $r(z_\varepsilon)$ and the costs to the agents be $c_{\underline{\varepsilon}}(z_\varepsilon) > c_{\overline{\varepsilon}}(z_\varepsilon)$. The principal chooses a menu $(z_\varepsilon, s_\varepsilon)$ to maximize:

$$\Pi = f[r(z_{\overline{\varepsilon}}) - s_{\overline{\varepsilon}}(z_{\overline{\varepsilon}})] - (1-f)[r(z_{\underline{\varepsilon}}) - s_{\underline{\varepsilon}}(z_{\underline{\varepsilon}})]. \quad (8)$$

Let each agent's reservation utility be 0. If types were directly observable, the principal would offer contracts which induced every agent to participate in the contract scheme and which left each of the agents at their reservation utilities. That is, $s_{\overline{\varepsilon}}(z_{\overline{\varepsilon}}) = c_{\overline{\varepsilon}}(z_{\overline{\varepsilon}})$ and $s_{\underline{\varepsilon}}(z_{\underline{\varepsilon}}) = c_{\underline{\varepsilon}}(z_{\underline{\varepsilon}})$. Then by choosing $z_{\overline{\varepsilon}}$ and $z_{\underline{\varepsilon}}$ optimally,

$$r'(z_{\overline{\varepsilon}}) = c'_{\overline{\varepsilon}}(z_{\overline{\varepsilon}}) \quad (9)$$

and

$$r'(z_{\underline{\varepsilon}}) = c'_{\underline{\varepsilon}}(z_{\underline{\varepsilon}}). \quad (10)$$

Together, (9) and (10) represent the *first best*: since each output level is set where marginal cost equals marginal benefit, the total surplus from production of z_ε is maximized. Denote any $z_{\overline{\varepsilon}}$ satisfying (9) as $z^F_{\overline{\varepsilon}}$ and any $z_{\underline{\varepsilon}}$ satisfying (10) as $z^F_{\underline{\varepsilon}}$. Full observability maximizes the principal's return because the total surplus is the largest of any mechanism and every agent is left with zero utility.

When types are *not* directly observable, the principal must optimally redesign the contract since $\overline{\varepsilon}$ types would receive positive utility from signing contracts designed for $\underline{\varepsilon}$ types. The revelation principle states that any direct revelation mechanism in which agents honestly report their types to the principal yields the highest expected payoffs. Consequently, the principal maximizes (8) subject to incentive and participation constraints (not shown for the sake of brevity). The solution to this *second-best* problem is characterized by:

$$r'(z_{\overline{\varepsilon}}) = c'_{\overline{\varepsilon}}(z_{\overline{\varepsilon}}) \quad (11)$$

and $s_{\overline{\varepsilon}}(z_{\overline{\varepsilon}}) = c_{\overline{\varepsilon}}(z_{\overline{\varepsilon}}) + R(z_{\underline{\varepsilon}})$, where

$$R(z_{\underline{\varepsilon}}) = c_{\overline{\varepsilon}}(z_{\underline{\varepsilon}}) - c_{\underline{\varepsilon}}(z_{\underline{\varepsilon}}) > 0 \quad (12)$$

indicates the information rents which must be given to a $\overline{\varepsilon}$ agent to truthfully reveal his type. Because (11) is identical to (9), $z_{\overline{\varepsilon}} = z^F_{\overline{\varepsilon}}$ even in the second-best case. Note that

$R'(z) > 0$ and $R''(z) > 0$. The optimal choice of $z_{\underline{\varepsilon}}$ is given by

$$r'(z_{\underline{\varepsilon}}) = c'_{\underline{\varepsilon}}(z_{\underline{\varepsilon}}) + \frac{f}{1-f} R'(z_{\underline{\varepsilon}}) \tag{13}$$

and $s_{\underline{\varepsilon}}(z_{\underline{\varepsilon}}) = c_{\underline{\varepsilon}}(z_{\underline{\varepsilon}})$. This is the standard "no distortion at the top" result. Since the utility of an agent is given by $U_{\varepsilon} = s_{\varepsilon}(z_{\varepsilon}) - c_{\varepsilon}(z_{\varepsilon})$, the ex post utilities of agents are $U_{\overline{\varepsilon}} = R(z_{\underline{\varepsilon}})$ and $U_{\underline{\varepsilon}} = 0$. Thus, $\overline{\varepsilon}$ agents always do better in this scheme than in the full information case, although the size of the overall rents is diminished; for this reason also, the principal does strictly worse: $\Pi^F > \Pi^{SB}$, where "SB" stands for "second-best".

Let any $z_{\underline{\varepsilon}}$ solving (13) be denoted $z_{\underline{\varepsilon}}^{SB}$. It has already been established that $z_{\overline{\varepsilon}}^{SB} = z_{\overline{\varepsilon}}^{F}$.

3.2 Solution to the no-commitment, information-sharing model

Since any Perfect Bayesian Equilibrium requires sequential rationality on the part of each player at each stage of the game given beliefs, backwards induction is used:

Stage 6. Given that principal Y offers an agent choosing contract k at the X-contract stage the menu $(\mathbf{z_y}, \mathbf{s_y})^k$, the agent chooses a specific contract, $(z_{yk}, s_{yk})^i$ to maximize:

$$u_{xy}^{Y}(z_{yk}, s_{yk}) = s_{yk} - c_y(z_{yk}) \tag{14}$$

Because this stage represents the end of the game, and every agent's reservation utility is zero, the agent will choose no contract if $u_{xy}^{Y}(z_{yk}, s_{yk}) < 0$ for all $(\mathbf{z_y}, \mathbf{s_y})^k$.

Stage 5. Suppose Y has purchased information from X. Then Y's structure of beliefs, Φ, is the conditional distribution of $\overline{y}$ contingent on which X-contracts, $(z_x, s_x)^k$, agents have chosen from the menu they were offered. Recall that $\Phi = \{\Phi_1, \Phi_2, \ldots, \Phi_K\}$, with:

$$\Phi_k = \Pr(\overline{y} | (z_x, s_x)^k, \Omega_b). \tag{15}$$

This expression says that if principal Y, having purchased information from X, sees that agent i has chosen contract k, then she believes that this agent is of type $\overline{y}$ with probability Φ_k. Similarly,

$$(1 - \Phi_k) = \Pr(\underline{y} | (z_x, s_x)^k, \Omega_b). \tag{16}$$

In any Bayesian equilibrium, the calculation of Φ_k must be consistent with Bayes' rule:

$$\begin{aligned}\Phi_k &= \Pr(\overline{y}|(z_x,s_x)^k,\Omega_b)\\ &= \frac{\Pr(\overline{y})\ \Pr((z_x,s_x)^k,\Omega_b|\overline{y})}{\Pr(\overline{y})\ \Pr((z_x,s_x)^k,\Omega_b|\overline{y})+\Pr(\underline{y})\ \Pr((z_x,s_x)^k,\Omega_b|\underline{y})}.\end{aligned} \tag{17}$$

Notice that the denominator of (17) is equivalent to $\Pr((z_x,s_x)^k,\Omega_b)$, which is itself just $\eta^X(k)$, the fraction of all agents who have chosen contract $(z_x,s_x)^k$. This information is available to Y since it is in Ω_b. We can also write $\Pr((z_x,s_x)^k,\Omega_b|y)$ as (for $y=\{\overline{y},\underline{y}\}$):

$$\Pr((z_x,s_x)^k|y)=\frac{\Pr((z_x,s_x)^k,\Omega_b,y)}{\Pr(y)}=\frac{\phi_k}{\Pr(y)}. \tag{18}$$

Write $\Pr((z_x,s_x)^k,\Omega_b,\overline{y})$ as $\overline{\phi}_k$ and $\Pr((z_x,s_x)^k,\Omega_b,\underline{y})$ as $\underline{\phi}_k$. Posterior beliefs about an agent's y-type, conditional on seeing that this agent has chosen contract $(z_x,s_x)^k$ are:

$$\Phi_k=\frac{\overline{\phi}_k}{\overline{\phi}_k+\underline{\phi}_k}=\frac{\overline{\phi}_k}{\eta^X((z_x,s_x)^k)}. \tag{19}$$

Example: Suppose that Principal X offers two contracts at Stage 1: $(z_x,s_x)^1$ and $(z_x,s_x)^2$. Suppose further that all $\overline{x}$ types choose contract 1, all $\underline{x}$ types choose contract 2, *and* that Y believes that this behaviour on the part of agents is consistent with the menu offered by principal X. If Y is to believe that all $\overline{x}$ have chosen contract 1, and all $\underline{x}$ have chosen contract 2, a minimal requirement is $\eta^X(1)=f$ and $\eta^X(2)=1-f$, since any other values of η^X would violate Y's belief that all $\overline{x}$ types choose contract 1 and all $\underline{x}$ types choose contract 2. Then if Y's posterior beliefs about $\overline{y}$ are:

$$\Phi_1=\frac{\overline{\phi}_1}{\overline{\phi}_1+\underline{\phi}_1}=\frac{\alpha}{\alpha+\beta}=\frac{\alpha}{f}$$

$$\Phi_2=\frac{\overline{\phi}_2}{\overline{\phi}_2+\underline{\phi}_2}=\frac{\beta}{\beta+\gamma}=\frac{\beta}{1-f}$$

which are exactly the conditional distributions of the prior joint distribution of x and y. In this case, one can say that "Y believes that X has fully separated agents along their x dimension of type (xy)."If Y purchases information, she can apply these beliefs to particular agents she identifies. If Y does not purchase this information, her beliefs about any specific agent must be represented by Φ_o. With no identity information from

X, her best estimate of a particular agent's type is the prior unconditional distribution of y: $\Pr(y = \overline{y}) = f$ and $\Pr(y = \underline{y}) = 1 - f$.

The following Lemma establishes Principal Y's behaviour under the assumption of sequential rationality:

Lemma 1: If Y has beliefs Φ_o at Stage 5, she will offer one menu of $l = 2$ contracts to all agents:

$$(\mathbf{z_y}(\mathbf{\Phi_o}), \mathbf{s_y}(\mathbf{\Phi_o})) = \{(z_{\overline{y}}^F, c_{\overline{y}}(z_{\overline{y}}^F) + R(z_{\underline{y}}^{SB})), (z_{\underline{y}}^{SB}, c_{\underline{y}}(z_{\underline{y}}^{SB}))\}$$

as in section 3.1. If Y has beliefs $\Phi \neq \Phi_o$ at Stage 5, she will separate agents according to the contract k each agent accepted from X, and offer each of the $k = 1, 2, \ldots, K$ subsets $M_k = 2$ contracts:

$$\begin{aligned}(\mathbf{z_y}(\mathbf{\Phi}), \mathbf{s_y}(\mathbf{\Phi})) &= \left\{ \bigcup_{k=1}^{K} (\mathbf{z_y}(\mathbf{\Phi}), \mathbf{s_y}(\mathbf{\Phi}))^k \right\} \\ &= \left\{ \bigcup_{k=1}^{K} \Big((z_{\underline{y}k}, c_{\underline{y}}(z_{\underline{y}k})), (z_{\overline{y}}^F, c_{\overline{y}}(z_{\overline{y}}^F) + R(z_{\underline{y}k}) \Big) \right\} \end{aligned} \tag{20}$$

Moreover, when $\Phi \neq \Phi_o$ and contract $k \neq k'$, $z_{\underline{y}k} > z_{\underline{y}k'}$ if $\Phi_k < \Phi_{k'}$.

Proof: From section 3.1, we know that the optimal contract menu that Y can offer when her beliefs are Φ_o is to set the "second-best" contract menu. Since this menu is incentive-compatible for all agents, Y could split up agents into arbitrary groups, and offer each group this exact menu. The choices of agents would then be identical to those if agents were not split up. Thus, the single menu $\{(z_{\overline{y}}^F, c_{\overline{y}}(z_{\overline{y}}^F) + R(z_{\underline{y}}^{SB})), (z_{\underline{y}}^{SB}, c_{\underline{y}}(z_{\underline{y}}^{SB}))\}$ is incentive feasible for *any* division of agents ex post. However, if Y has beliefs Φ, then she has K separate static problems which are identical in form to that of 3.1, except that her posterior beliefs are potentially different for each k group. Her problem then is to set up k different sets of contracts to replicate the problem of section 3.1:

$$\begin{aligned}\max_{(\mathbf{z_y},\mathbf{s_y})^k} \Pi_Y &= \sum_{k=1}^{K} \eta^X((z_x, s_x)^k)\{\Phi_k[r_y(z_{\overline{y}k}) - s_{\overline{y}}(z_{\overline{y}k})] + (1-\Phi_k)[r_y(z_{\underline{y}k}) - s_{\underline{y}}(z_{\underline{y}k})]\} \\ &= \sum_{k=1}^{K} \{\overline{\phi}_k[r_y(z_{\overline{y}k}) - s_{\overline{y}}(z_{\overline{y}k})] + \underline{\phi}_k[r_y(z_{\underline{y}k}) - s_{\underline{y}}(z_{\underline{y}k})]\} \end{aligned} \tag{21}$$

subject to participation constraints ($u_{xy} \geq 0$ for all (xy)) as well as the relevant incentive constraints *within* each of the k groups.

Whenever $\Phi_k \neq \Phi_o$, Y will set $z_{\overline{y}k}$ and $z_{\underline{y}k}$ such that:

$$r'(z_{\overline{y}k}) = c'_{\overline{y}}(z_{\overline{y}k}) \tag{22}$$

$$r'(z_{\underline{y}k}) = c'_{\underline{y}}(z_{\overline{y}k}) + \frac{\overline{\phi}_k}{\underline{\phi}_k} R'(z_{\underline{y}k}). \tag{23}$$

Thus, $z_{\overline{y}k} = z^F_{\overline{y}}$ as in the section 3.1 problem. However, $z_{\underline{y}k} \neq z^{SB}_{\underline{y}}$ unless $\Phi_k = f$ and $(1-\Phi_k) = 1-f$. But since $z^{SB}_{\underline{y}}$ is feasible even when agents are split into k subgroups, principal Y must be doing better by following her belief system characterized by Φ. Moreover, note that:

$$\frac{dz_{\underline{y}k}}{d(\frac{\overline{\phi}_k}{\underline{\phi}_k})} = \frac{R'(\cdot)}{r''_y(\cdot) - c''_{\underline{y}}(\cdot) - R''(\cdot)} < 0.$$

By inspection of (19),

$$\frac{\overline{\phi}_k}{\underline{\phi}_k} = \frac{\Phi_k}{1-\Phi_k} \tag{24}$$

So if for the X contracts k and k', $\Phi_k < \Phi_{k'}$, then $z_{\underline{y}k} > z_{\underline{y}k'}$. ∎

Call principal Y's payoff from this policy of splitting agents into k distinct subgroups as $\Pi^{split}_Y(\Phi)$. Lemma 1 implies that $\Pi^{split}_Y(\Phi) > \Pi^{SB}_Y(\Phi_o)$. For future reference, define $W(\Phi)$ as Y's "maximum willingness to pay" for X's information:

$$W(\Phi) = \Pi^{split}_Y(\Phi) - \Pi^{SB}_Y(\Phi_o) \geq 0 \tag{25}$$

Clearly, the more closely that Φ can be used to identify $\overline{y}$ types given X-contract information, the more valuable is this information. For example, if Y believed that virtually every agent who chose contract j was of type $\overline{y}$, she would set $z_{\underline{y}j}$ very low. This in turn would mean that the rents paid to $\overline{y}$ agents within the j subset would be very low, showing that the payoff from the j subset for the principal would approach the "first best" payoff, at least for this subset.

Stage 4. The problem for Y in stage 4 is simply to buy (B) or not buy (NB) information from X. When making her decision, Y can examine the menu of contracts that X has proposed, $(\mathbf{z_x}, \mathbf{s_x})$, as well as the proportions of agents who have chosen each contract, $\eta^X(\mathbf{z_x}, \mathbf{s_x})$. Thus, she is able to update her beliefs about the proportions of $\overline{y}$ types in each group, and can therefore calculate $W(\Phi)$. To gain the privilege of matching contracts with identities, she is willing to pay any amount up to $W(\Phi)$, since her next best option is to implement the second-best contract menu, corresponding to her prior beliefs Φ_o.

Two special cases are of interest. The first is where principal X sets a menu of only one contract, and this contract is chosen by all agents.[17] In this case, $\Phi = \Phi_1$ trivially and $\eta^X(1) = 1$. Then by (19), $\Phi = \phi(\overline{y}) = f$ and $(1 - \Phi) = \phi(\underline{y}) = 1 - f$. Since these are the same beliefs as Φ_o, Y would pay nothing for this information. In the second case, it may be that X's information forces Y to adjust her beliefs such that, by (24), the ratio $\frac{\overline{\phi}_k}{\underline{\phi}_k}$ is the same for all subsets, k. This implies $\Phi_k = f$ for all k. Here, Y would set $z_{\underline{y}k} = z_{\underline{y}}^{SB}$: the second-best contract menu. Again Y would not be willing to pay for this information. The next result follows:

Lemma 2: In any Bayesian equilibrium, Y will pay for information gathered by X only if this information changes Y's beliefs about the relative proportions of $\overline{y}$ types in the subsets of agents defined by the contracts they have chosen at the X-contract stage. If Y does not pay for information, then she will offer the second-best-payoff menu of contracts in stage 5 to all agents.

For $p \geq 0$ Y's strategies given the price offered by X are:

$$\begin{aligned} \text{Buy, } (B) \quad &\Longleftrightarrow \quad p \leq W(\Phi) \\ \text{No Buy, } (NB) \quad &\Longleftrightarrow \quad \text{otherwise} \end{aligned}$$

Clearly, Y will "buy" any information when $p < 0$, even under beliefs Φ_o.

[17] If X set one contract that was *not* accepted by all agents, Y could still draw inferences about the proportions of y types amongst those who chose not to sign the contract.

Stage 3. Principal X has a strong incentive to sell information once it has been gathered from agents, since contracts have already been signed. Sequential rationality requires that X act in his best interest once stage 3 has been reached. Thus, given the contracts X has offered in stage 1, and the proportions of agents who have accepted them, X will sell his information for the $p(\Omega_b)$ which maximizes his payoff at stage 3.

Note that X can calculate the value of his own information to principal Y.[18] Then he will charge $p(\Omega_b) = W(\Phi)$, extracting all possible surplus from Y, given that he chooses to sell. The only situation in which X potentially will *not* sell is if the information sold would not change Y's beliefs, leaving Y at Φ_o. In this case, Y would be willing to pay exactly zero for the information. Therefore, whether or not X chooses to sell when his information generates Φ_o does not affect the outcome of the game from Stage 3 onwards, so assume that X would sell for zero if $W(\Phi) = W(\Phi_o) = 0$.

Lemma 3. In any Bayesian equilibrium, X will always choose to sell his acquired information.

Stage 2. Agents anticipate the dissemination of their information in the next stage (by Lemma 3) and the use of this information by principal Y in the design of contracts they will be offered in Stage 5 (Lemma 1). Thus, given that they are presented with a K-element menu $(\mathbf{z_x}, \mathbf{s_x})$ by principal X, an agent chooses contract $(z_x, s_x)^k$ from the menu to maximize (4):

$$EU_{xy}((z_x, s_x)^k) = u^X_{xy}((z_x, s_x)^k) + Eu^Y_{xy}((z_{yk}, s_{yk})^{m_k}|(z_x, s_x)^k, \Omega_b)$$

Every agent can easily calculate his current payoff u^X_{xy}, $s_x - c_x(z_x)$. However, the choice of a particular contract can affect the *menu* of contracts he will be offered next period, directly through $(z_x, s_x)^k$ and indirectly through Y's posterior beliefs, Φ. Agents must therefore also consider the contract choices of other agents and how these other choices will affect principal Y's beliefs, since the characteristics of the menu offered by Y depend essentially on *what proportion of* $\overline{y}$*-types* Y believes have chosen each contract offered by X.

Solving recursively, agents know what contracts will be offered by Y, given Φ; specif-

[18] Since Φ is derived from Bayes' rule, and X is aware of his own menu of offered contracts, he and Y share the same information set once information has been sold: Ω_b. X also knows that Y, being sequentially rational, must choose to use any information which changes her beliefs.

ically, they have a choice between two contracts, one designed for $\overline{y}$-types choosing k and one for $\underline{y}$-types choosing k. Since they know their type in the y-dimension, they know which of these contracts they will accept since stage 6 is the end of the game. If they are of type $\underline{y}$, they will receive $u^Y_{x\underline{y}} = 0$ for certain, no matter which X-contract they choose:

$$EU_{x\underline{y}}((z_x, s_x)^k) = u^X_{x\underline{y}}((z_x, s_x)^k) \tag{26}$$

Similarly, if agents are of type $\overline{y}$, they know they will receive $R(z_{\underline{y}k})$ since $u^Y_{x\overline{y}} = s_{\overline{y}k} - c_{\overline{y}}(z^F_{\overline{y}})$ and $s_{\overline{y}k} = c_{\overline{y}}(z^F_{\overline{y}}) + R(z_{\underline{y}k})$ by Lemma 1. Thus, $u^Y_{x\overline{y}} = R(z_{\underline{y}k})$ and $Eu^Y_{x\overline{y}} = R(Ez_{\underline{y}k})$:

$$EU_{x\overline{y}}((z_x, s_x)^k) = u^X_{x\overline{y}}((z_x, s_x)^k) + R(Ez_{\underline{y}k}) \tag{27}$$

The expectation of $z_{\underline{y}k}$ depends on the expectation of other agents' choices of X-contract. Because $z_{\underline{y}k}$ depends on Φ, Φ depends on Ω_b, and Ω_b includes $\eta^X(\mathbf{z_x}, \mathbf{s_x})$, every agent's choice of which X-contract to choose ultimately affects $z_{\underline{y}k}$.

Recall that agent i's strategy is $\sigma^i_{xy} = ((z_x, s_x)^i, (z_y, s_y)^i)$. Let $\sigma^{-i}_{xy} = (\sigma^{j \neq i}_{xy}, \sigma_X, \sigma_Y)$ be the profile of every strategy for every player except i. Given any σ^{-i}_{xy}, the *best response* of agent i is to choose the x contract which maximizes (26) for agents of type $(x\underline{y})$, and (27) for agents of type $(x\overline{y})$. Thus, the best response for $(x\underline{y})$ agents is to choose contract k over any other contract, k' if:

$$u^X_{x\underline{y}}((z_x, s_x)^k | \sigma^{-i}_{xy}) \geq u^X_{x\underline{y}}((z_x, s_x)^{k\prime} | \sigma^{-i}_{xy}) \tag{28}$$

and for $(x\overline{y})$ agents:

$$u^X_{x\overline{y}}((z_x, s_x)^k | \sigma^{-i}_{xy}) + R(Ez_{\underline{y}k} | \Omega_b(\sigma^{-i}_{xy})) \geq u^X_{x\overline{y}}((z_x, s_x)^{k\prime} | \sigma^{-i}_{xy}) + R(Ez_{\underline{y}k'} | \Omega_b(\sigma^{-i}_{xy})) \tag{29}$$

In the case of $(x\underline{y})$ agents, the best response is simply to choose their one-period preferred contract offered by X, since by Lemma 1, Y will always give these agents a payoff of zero next period. However, $(x\overline{y})$ agents' decisions are governed by their expectation of z_{yk} for all k, conditioned by the choices of other agents, $\sigma^{j \neq i}_{xy}$. If fewer *other* $(x\overline{y})$ agents choose contract k than k', and if Y's beliefs correctly reflect these agents' choices, then by (24), $E(z_{\underline{y}k} | \Omega_b(\sigma^{-i}_{xy})) > E(z_{\underline{y}k'} | \Omega_b(\sigma^{-i}_{xy}))$ and it is relatively more lucrative to choose contract k. This effect must be balanced against the payoff $u^X_{x\overline{y}}((z_x, s_x)^k)$, which is indirectly determined by X's menu-offer.

Stage 1. It has been established that the static second-best optimal contract menu for X to offer is characterized by (11) and (13). However, the dynamic nature of the current problem can destroy the validity of the revelation principle and alter X's optimal contract menu away from a direct-revelation situation (see e.g. Laffont and Tirole, 1988). Fortunately, Principal X's problem is immediately simplified by the following result:

Lemma 4. Principal X offers at most a menu of two contracts.

Proof: See appendix.

The optimal contract menu offered by X must take into account the agents' dynamic problem in stage 2. From section 3.1, we know that X could always feasibly set the "second-best" menu of contracts. However, by doing so, agents will anticipate the release of their information, and adjust their behaviour accordingly. Suppose that X (naively) sets the menu: $(\mathbf{z_x}, \mathbf{s_x}) = \{(z_x, s_x)^1, (z_x, s_x)^2\} = \{(z_{\overline{x}}^F, c_{\overline{x}}(z_{\overline{x}}^F) + R(z_{\underline{x}}^{SB})), (z_{\underline{x}}^{SB}, c_{\underline{x}}(z_{\underline{x}}^{SB}))\}$. Then we can state:

Proposition 1. If principal X sets the one-period second-best optimal menu of contracts with no confidentiality, and agents are fully rational, then the value of information to principal Y is zero and Y also sets the one-period second-best optimal menu of contracts.

Proof: See appendix.

Proposition 1 establishes that the inability of Principal X to commit to keep information confidential leads to the destruction of any value of information that he might want to sell. Any situation where $R(Ez_{y2}|\Omega_b(\sigma_{xy}^{-i})) > R(Ez_{y1}|\Omega_b(\sigma_{xy}^{-i}))$ leads a proportion $(1-h)$ of $(\overline{xy})$ types to choose the "wrong" contract from X's perspective in stage 2. When X offers the one-period second-best optimal menu:

$$\underline{h} = \left(\frac{f-\alpha}{\alpha}\right)\left(\frac{f}{1-f}\right) < 1.$$

Therefore, if X naively sets his contract menu at what it would be in the single-period case he obtains a strictly smaller payoff than in the one period "second-best" case, Π_X^{SB}.

Call X's payoff under this strategy $\Pi_X^{SB'}$. Then this payoff is:

$$\begin{aligned}\Pi_X^{SB'} &= (\beta + \underline{h}\alpha)[r_x(z_{\overline{x}}^F) - c_{\overline{x}}(z_{\overline{x}}^F) - R(z_{\underline{x}}^{SB})] + \\ &\quad (1 - f + (1 - \underline{h})\alpha)[r_x(z_{\underline{x}}^{SB}) - c_{\underline{x}}(z_{\underline{x}}^{SB})]. \end{aligned} \tag{30}$$

With some rearranging:

$$\Pi_X^{SB} - \Pi_X^{SB'} = \alpha(1 - \underline{h})\left\{\left(r_x(z_x^F) - c_{\overline{x}}(z_x^F)\right) - \left(r_x(z_{\underline{x}}^{SB}) - c_{\overline{x}}(z_{\underline{x}}^{SB})\right)\right\} \tag{31}$$

This expression is always greater than zero, since by definition z_x^F maximizes $r_x(\cdot) - c_{\overline{x}}(\cdot)$. Essentially, (31) represents the deadweight loss of no-confidentiality: the lack of protection of information causes the "most productive" ($\overline{xy}$) types to choose a less-productive contract. Whenever X offers the second-best contract menu, however, it must be that Y can also do no better that to set the second-best menu herself. Therefore, *strictly less* is produced when this situation holds than would be the case under confidentiality.[19]

A reasonable question to ask is: when would principal X offer such a (naive) second-best menu? The answer is: only if he could be sure it would generate $h = 1$! Recall from Lemma 4 that X will offer either 1 or 2 contracts in his menu. If he offers 2 contracts, then he can always do better than to offer $z_{\underline{x}}^{SB}$. To see why, note that $\underline{h}$ is invariant to the terms of $(\mathbf{z_X}, \mathbf{s_X})$. However, because X anticipates the subsequent behaviour of agents (characterized by $\underline{h}$), he should set $z_{\underline{x}}$ to maximize the general form of (30) as opposed to (8). The optimal contract menu for X should reflect the ratio $\frac{\beta + \underline{h}\alpha}{1 - f + (1 - \underline{h})\alpha}$, not $\frac{f}{1-f}$. Since

$$\frac{\beta + \underline{h}\alpha}{1 - f + (1 - \underline{h})\alpha} = \frac{\beta}{\gamma} < \frac{f}{1 - f} \iff \rho > 0$$

X should optimally reduce $z_{\underline{x}}$ from $z_{\underline{x}}^{SB}$ according to:

$$r_x'(z_{\underline{x}}) = c_{\underline{x}}'(z_{\underline{x}}) + \frac{\beta}{\gamma} R'(z_{\underline{x}}). \tag{32}$$

Principal X may, however, wish to set a contract which generates $h \in (\underline{h}, 1]$, in which

[19] This phenomenon is a consequence of the ratchet effect in a no-commitment model (Freixas, Guesnerie and Tirole, 1985).

case $z_{\underline{x}}$ should be set at $z_{\underline{x}}(h)$, where:

$$r_x'(z_{\underline{x}}(h)) = c_{\underline{x}}'(z_{\underline{x}}(h)) + \frac{\beta + h\alpha}{1 - f + (1-h)\alpha} R'(z_{\underline{x}}(h)). \tag{33}$$

Call the $z_{\underline{x}}(h)$ which satisfies (33) $z_{\underline{x}}^{NC}$ where "NC" stands for "no commitment" or "no confidentiality". Replacing $z_{\underline{x}}^{SB}$ with $z_{\underline{x}}^{NC}$, and $\underline{h}$ with the more general h in (31) demonstrates that $0 < \Pi_X^{SB} - \Pi_X^{NC} < \Pi_X^{SB} - \Pi_X^{SB'}$, or $\Pi_X^{NC} > \Pi_X^{SB'}$. Thus, X will always set these levels of z if he sets 2 contracts.

Given that the behaviour of all players in Stages 2-5 has been characterized, X's choice is essentially between setting either a single contract (z_x, s_x) or else the contract menu:

$$(\mathbf{z_X}, \mathbf{s_X})^{TB} \equiv \{(z_x, s_x)^1, (z_x, s_x)^2\} \equiv \left\{ \Big(z_{\overline{x}}^F, c_{\overline{x}}(z_{\overline{x}}^F) + R(z_{\underline{x}}(h)) + S(h) \Big), \Big(z_{\underline{x}}(h), c_{\underline{x}}(z_{\underline{x}}(h)) \Big) \right\}$$

where h is the proportion of $(\overline{xy})$ types choosing contract 1, and $S(h) \geq 0$ is a *supplementary payment* given to those choosing contract 1 for inducing a certain level of h. Recall that an agent of type $(\overline{xy})$ will choose contract 1 if:

$$\begin{aligned} R(z_{\underline{x}}(h)) + R(Ez_{y1}(h)|\Omega_b(\sigma_{xy}^{-i})) + S(h) &> R(z_{\underline{x}}(h)) + R(Ez_{y2}(h)|\Omega_b(\sigma_{xy}^{-i})) \\ R(Ez_{y1}(h)|\Omega_b(\sigma_{xy}^{-i})) + S(h) &> R(Ez_{y2}(h)|\Omega_b(\sigma_{xy}^{-i})). \end{aligned} \tag{34}$$

Therefore, if X wants to set a level of $h \in (\underline{h}, 1]$, S should be set so that:

$$\begin{aligned} S(h) &= R(Ez_{y2}(h)) - R(Ez_{y1}(h)) \\ &= U_{\overline{xy}}((z_x, s_x)^2, h) - U_{\overline{xy}}((z_x, s_x)^1, h) \end{aligned} \tag{35}$$

If X wants to set $h = \underline{h}$, he sets $S(\underline{h}) = 0$. Because setting $\underline{h}$ is always feasible, X can always choose his menu treating $\underline{h}$ as a parameter of his problem. But by the logic of section 3.1, X will optimally choose to set a menu which separates different types, even imperfectly. Therefore, setting a single contract is never optimal and will never be chosen by X in equilibrium.

Following equation (5), the total payoff to principal X in any no-commitment equilibrium is:

$$\Pi_X^{NC}\Big((\mathbf{z_x},\mathbf{s_x})^{NC}\Big) = (\beta+h\alpha)[r_x(z_{\overline{x}}^F)-c_{\overline{x}}(z_{\overline{x}}^F)-R(z_{\underline{x}}^{NC}(h))-S(h)] \\ +(1-f+(1-h)\alpha)[r_x(z_{\underline{x}}^{NC}(h))-c_{\underline{x}}(z_{\underline{x}}^{NC}(h))]+ \\ +W(\Phi(h)) \tag{36}$$

where $W(\Phi(h)) = p(\Omega_b)$ is the maximum price X can charge Y for information about particular individuals' contract-choices when h obtains in equilibrium. i.e. This information must induce beliefs in Y, $\Phi(h)$, that are consistent with the proportion h of $(\overline{xy})$ types choosing contract 1 from X. Y has access to the menu of contracts X has offered, as well as the proportion of agents who have chosen each. Therefore, given $(\mathbf{z_x},\mathbf{s_x})^{NC}$, $\eta^X(1) = (\beta+h\alpha)$ and $\eta^X(2) = (1-f+(1-h)\alpha)$, Y must surmise:

$$\begin{aligned} \overline{\phi}_1 &= \alpha h \\ \underline{\phi}_1 &= \beta \\ \overline{\phi}_2 &= \beta+(1-h)\alpha = f-\alpha h \\ \underline{\phi}_2 &= \gamma. \end{aligned} \tag{37}$$

So long as $h \neq \underline{h}$, Y's optimal strategy in Stage 5 requires her to set two contract menus: one for those who have chosen contract 1 from X, and one for those who have chosen contract 2. These contracts are indexed by $k = 1, 2$, from (20). Accordingly, she sets the menus 1 and 2 such that:

$$\begin{aligned} (\mathbf{z_y}(\mathbf{\Phi(h)}),\mathbf{s_y}(\mathbf{\Phi(h)})) &= \Big\{(\mathbf{z_y}(\mathbf{\Phi(h)}),\mathbf{s_y}(\mathbf{\Phi(h)}))^1,(\mathbf{z_y}(\mathbf{\Phi(h)}),\mathbf{s_y}(\mathbf{\Phi(h)}))^2\Big\} \\ &= \Big\{\Big(\Big(z_{\overline{y}}^F, c_{\overline{y}}(z_{\overline{y}}^F)+R(z_{\underline{y}1}(h))\Big),\Big(z_{\underline{y}1}(h), c_{\underline{y}}(z_{\underline{y}1}(h))\Big)\Big) \\ &\quad \Big(\Big(z_{\overline{y}}^F, c_{\overline{y}}(z_{\overline{y}}^F)+R(z_{\underline{y}2}(h))\Big),\Big(z_{\underline{y}2}(h), c_{\underline{y}}(z_{\underline{y}2}(h))\Big)\Big)\Big\} \end{aligned} \tag{38}$$

where $z_{\underline{y}1}(h))$ solves:

$$r_y'(z_{\underline{y}1}) = c_{\underline{y}}'(z_{\underline{y}1}) + \frac{\alpha h}{\beta}R'(z_{\underline{y}1})$$

and $z_{\underline{y}2}(h))$ solves:

$$r'_y(z_{\underline{y}2}) = c'_{\underline{y}}(z_{\underline{y}2}) + \frac{f - \alpha h}{\gamma} R'(z_{\underline{y}2}).$$

What the expressions above say is that Y must incorporate her knowledge that some agents do not choose their "targeted" contract from X, when she sets her own contracts. If she sees $h = \underline{h}$, her beliefs remain at Φ_o, and she sets the "second best" menu. Otherwise, greater values of h mean better information: she can optimally set $z_{\underline{y}1}(h)$ and $z_{\underline{y}2}(h)$ to extract a higher payoff. Indeed, if $h = 1$, then she has the best information possible, and can identify $(x\overline{y})$ types with the greatest certainty.

The problem for X is therefore what level of h to induce, knowing that Y's value of information rises when h is set above $\underline{h}$. We can immediately state the following:

Lemma 5: Principal Y cannot compensate X for inducing *any* level of h above $\underline{h}$.

Proof: See appendix.

X's objective in stage 1 is to set S as low as possible while still inducing an increase in h. But X's inherent problem is that every marginal increase in S given to $(\overline{xy})$ types must also be given to $(\overline{x}\underline{y})$ types as well, since X can never set $u_{\overline{xy}}(z_x, s_x)^k > u_{\overline{x}\underline{y}}(z_x, s_x)^{k\prime}$ without $(\overline{x}\underline{y})$ types choosing the contract k. However, Y's *marginal benefit* of an increased h is the first-order reduction in extra rents, $R(z_{\underline{y}2}(h)) - R(z_{\underline{y}1}(h)) = S(h)$ which she saves by inducing incremental $\overline{xy}$ types to choose contract 1 instead of 2. Although X can capture Y's marginal benefit through p, this benefit cannot cover the payments $S(h)$ which are given to $\overline{x}\underline{y}$ types. This situation is illustrated by Figure 1 at the end of paper (using numerical values from the example in section 6).

Nevertheless, increasing h is beneficial for X in the sense that it encourages more $(\overline{xy})$ types to select the "first-best" level of z_x. Overall, the change in X's payoff as h increases from $\underline{h}$ is (making use of the envelope theorem since Π_X^{NC} is a maximum-value function):

$$\begin{aligned}\frac{\partial \Pi_X^{NC}}{\partial h} &= \alpha\left[\left(r_x(z_{\overline{x}}^F) - c_{\overline{x}}(z_{\overline{x}}^F) - R(z_{\underline{x}}^{NC}(h))\right) - \left(r_x(z_{\underline{x}}^{NC}(h)) - c_{\underline{x}}(z_{\underline{x}}^{NC}(h))\right)\right] \\ &\quad -\left[\frac{\partial[(\beta+\alpha h)S]}{\partial h} - \frac{\partial W(\Phi(h))}{\partial h}\right] \\ &= \alpha\left[\left(r_x(z_{\overline{x}}^F) - c_{\overline{x}}(z_{\overline{x}}^F)\right) - \left(r_x(z_{\underline{x}}^{NC}(h)) - c_{\overline{x}}(z_{\underline{x}}^{NC}(h))\right)\right] - \left[(\beta+\alpha h)\frac{\partial S}{\partial h}\right]\end{aligned} \tag{39}$$

The first bracketed term above is the marginal benefit to X of raising h; the second is his marginal (net) cost of doing so. Because by lemma 5, Y cannot compensate X for raising the level of h, X must bear some of this cost.

Proposition 2: In the *Perfect Bayesian Equilibrium* of the game with no confidentiality, if

$$\frac{\partial^2 z_{y2}}{\partial h \partial \omega} \geq 0 \text{ and } \frac{\partial^2 z_{y1}}{\partial h \partial \omega} \leq 0$$

then

i. For ω sufficiently small, the equilibrium exhibits $h = 1$ and $(\overline{x}y)$ agents fully separate according to their x-type with extra compensation $S(1) = R(z_{y2}(1)) - R(z_{y1}(1))$.

ii. For ω sufficiently large, the equilibrium exhibits $h = \underline{h}$, these agents receive no extra compensation, and separation is imperfect.

Proof: See appendix.

The results of Proposition 2 can hold even if the above condition on the cross-derivatives above does not hold, but not with certainty. Intuitively, when ω rises Y optimally sets $z_{\underline{y}k}$ *higher* for both $k = 1, 2$.[20] X must also pay S to all $\overline{x}y$ agents to increase h. If $z_{\underline{y}2} - z_{\underline{y}1}$ *grows* as ω rises, then $S(h)$ becomes costlier for all $h > \underline{h}$. Since by lemma 5, Y cannot fully compensate X for his extra-payments, S, X is less likely to increase h the larger is ω. This result is somewhat counterintuitive. Recall that ω is a rough measure of the productivity of principal Y relative to X, so a higher ω

[20]Mathematically, $\frac{dz_{yk}}{d\omega} = \frac{-r_x'(z)}{\omega r_x''(z) - c_{\underline{y}}''(z) - g_k(h)R''(z)} > 0$, for $k = 1, 2$.

implies that individuals' information is relatively more productive to Y than to X.[21] Proposition 2 states that even though "better" information is more valuable to Y, the larger is ω the less likely she will obtain this information in any equilibrium. The payoff to Principal X for implementing different values of h given "high" and "low" levels of ω is illustrated in Figure 2.

We can now state the overall equilibrium $\sigma^* = \{\sigma^*_X, \sigma^*_Y, \sigma^{i*}_{xy}\}$ of the game:

$$\begin{aligned}
\sigma^*_X &= \{\text{set a menu of two contracts } (\mathbf{z_x}, \mathbf{s_x})^{TB}, S(\text{ell}), p = W(\Phi(h))\} \\
\sigma^*_Y &= \{B(\text{uy}), \text{set two menus of two contracts } (\mathbf{z_y}(\boldsymbol{\Phi}(\mathbf{h})), \mathbf{s_x}(\boldsymbol{\Phi}(\mathbf{h})))^{k=1,2}\} \\
\sigma^{i*}_{\underline{xy}} &= \{\text{choose } (z_x, s_x)^2, \text{choose } (z_{y2}, s_{y2})^2\} \\
\sigma^{i*}_{\underline{x}\overline{y}} &= \{\text{choose } (z_x, s_x)^2, \text{choose } (z_{y2}, s_{y2})^1\} \\
\sigma^{i*}_{\overline{x}\underline{y}} &= \{\text{choose } (z_x, s_x)^1, \text{choose } (z_{y1}, s_{y1})^2\} \\
\sigma^{i*}_{\overline{xy}} &= \{\text{choose } (z_x, s_x)^1 \text{ w.p. } h, \text{choose } (z_x, s_x)^2 \text{ w.p. } (1-h), \\
&\quad \text{choose } (z_{y1}, s_{y1})^1 \text{ if 1 chosen from } X, \text{choose } (z_{y2}, s_{y2})^1 \text{ if 2 chosen from } X\}
\end{aligned}$$

4 Confidentiality and Welfare

To this point we have assumed that principal X could offer no assurances that information gathered from agents would remain secret from principal Y. As a consequence, agents rationally anticipate X's decision to sell his contract-information, leading to an equilibrium in which the value of this information is destroyed. A law regarding confidentiality may be seen as a strong commitment device for X not to share information.

A confidentiality law's effect is to eliminate the incentive for agents to "choose the wrong contract" from X. In particular, if $(\overline{xy})$ types believe that Principal X will not share information, they will all choose the contract 1 that X offers. If confidentiality holds, he simply sets the one-period optimal contract menu and receives Π^{SB}_X. In this case, Y's beliefs remain at Φ_o, and Y too sets the one-period optimal contract menu to receive Π^{SB}_Y. Therefore:

Proposition 3: Under confidentiality, X always receives a higher payoff than in the no-commitment equilibrium. Y's payoff is the same in either case.

[21]To see this result mathematically, differentiate $\frac{\partial \Pi^{split}_Y}{\partial h}$ with respect to ω. This cross-derivative is positive for all h.

Proof: Y fares equally well in either case, since in the no-commitment equilibrium, X is able to set a price for information that extracts $W(\Phi(h))$ from Y, regardless of $h \in [\underline{h}, 1]$. Examining (36), $\Pi_X^{SB} = \Pi_X^{NC}$ only if $h = 1$ and $fS = W(\Phi(1))$. But from Lemma 5, $fS > W(\Phi(1))$ always. therefore $\Pi_X^{NC} < \Pi_X^{SB}$. ■

Indeed, for X *not* to want confidentiality when such a law is available, it must be that his payoff from no confidentiality exceeds Π_X^{SB}. Proposition 3 showed that this outcome is impossible in the no-commitment equilibrium with no discounting.

The comparison of agents' welfare under the confidentiality and no-confidentiality cases is summarized in the following result:

Proposition 4: Agents of every type do at least as well as under confidentiality as under confidentiality.

Proof: Agents receive ex-post utility $U_{xy}(z, s) = u_{xy}^X(z_x, s_x) + u_{xy}^Y(z_y, s_y)$. Denote the no-confidentiality (or no-commitment) payoff by "NC" and the confidentiality payoff by "C". Then, noting that $S(h) = [R(z_{\underline{y}2}(h)) - R(z_{\underline{y}1}(h))]$:

$$\begin{aligned}
U_{\underline{xy}}^{NC} &= 0 + 0 = 0 = U_{\underline{xy}}^{C} \\
U_{\overline{x}\underline{y}}^{NC} &= R(z_{\underline{x}}(h)) + [R(z_{\underline{y}2}(h)) - R(z_{\underline{y}1}(h))] + 0 > R(z_{\underline{x}}^{SB}) = U_{\overline{x}\underline{y}}^{C} \\
U_{\underline{x}\overline{y}}^{NC} &= 0 + R(z_{\underline{y}2}(h)) \geq R(z_{\underline{y}}^{SB}) = U_{\underline{x}\overline{y}}^{C} \\
U_{\overline{xy}}^{NC} &= R(z_{\underline{x}}(h)) + [R(z_{\underline{y}2}(h)) - R(z_{\underline{y}1}(h))] + R(z_{\underline{y}1}(h)) \\
&= R(z_{\underline{x}}(h)) + R(z_{\underline{y}2}(h)) > R(z_{\underline{x}}^{SB}) + R(z_{\underline{y}}^{SB}) = U_{\overline{xy}}^{C}
\end{aligned} \tag{40}$$

Although $z_{\underline{y}1}(h) = z_{\underline{y}2}(h) = z_{\underline{y}}^{SB}$ when $h = \underline{h}$, $(\overline{x}\underline{y})$ and $(\overline{xy})$ types still do strictly better because $R(z_{\underline{x}}(h)) > R(z_{\underline{x}}^{SB})$ unless $h = 1$. When $h > \underline{h}$, $(\underline{x}\overline{y})$ do strictly better as well. ■

Proposition 4 presents an interesting result, namely that although agents destroy the value of information in the no-commitment equilibrium, they end up extracting at least as high payoffs from the two principals. Indeed, if $h > \underline{h}$ — that is, some extra-compensation, $S(h) > 0$ is given to $(\overline{x}y)$ agents — all types of agent except for $(\underline{xy})$ types do strictly better. From Proposition 2, this outcome is more likely the less productive is principal Y, as characterized by ω.

Corollary: For ω sufficiently small, and $\frac{\partial^2 z_{y2}}{\partial h \partial \omega} \geq 0$ and $\frac{\partial^2 z_{y1}}{\partial h \partial \omega} \leq 0$, all agents do at least as well in the no-commitment equilibrium as under confidentiality, and a proportion $(1-\gamma)$ do strictly better.

When X cannot withhold information with Y, and Y is less productive than X, it turns out that agents would prefer that their information be shared than kept secret. Still, if X can, he will set a confidentiality policy in which information will not be shared. In all these cases, confidentiality is strictly in principal X's best interest, forming an interesting counterpoint to claims that firms must be forced to provide confidentiality in order to preserve consumers' interests. Notably, this result is robust to changes in a number of features of the model, including the distribution of types, the level of (positive) correlation between x and y, and the form of the rent-function, $R(\cdot)$.

5 Discounting

Thus far, confidentiality has emerged as a dominant policy for principal X. If so, how can information-sharing be explained in equilibrium? One explanation is that agents discount the future to an extent that X finds it profitable to compensate agents for *eventually* selling their information to principal Y.

Suppose all agents discount the future at the rate $\delta < 1$, or, more specifically that the discount factor of agents exceeds that of principal X by a factor of δ.[22] Then agents' ex-ante payoffs are given by:

$$EU_{xy}(z,s) = u^X_{xy}((z_x, s_x)^k) + \delta R(E(z_{yk}, s_{yk})^{m_k} | \Omega_b(\sigma^{-i}_{xy})) \ , \ k = 1, 2. \tag{41}$$

For additional $(\overline{xy})$ agents to choose contract 1 from X — i.e. to increase h — they must be offered:

$$S = [R(z_{\underline{y}2}(h)) - R(z_{\underline{y}1}(h))] \tag{42}$$

X can set the contract menu $\left\{ \left(z^F_{\overline{x}}\right), c_{\overline{x}}(z^F_{\overline{x}}) + R(z^{SB}_{\underline{x}}) + \delta S \right), \left(z^{SB}_{\underline{x}}, c_{\underline{x}}(z^{SB}_{\underline{x}})\right) \right\}$, setting S high enough to ensure $h = 1$. This policy will cost the amount $f\delta S$, since in this equilibrium, all f $(\overline{x}y)$ agents must be paid S but value S discounted at the rate δ. When stage 4 is reached, Y values this information at $W(\Phi(h = 1))$.

[22] This is actually a critical distinction: an alternative definition would be that Principal X does not discount the future, although agents do.

The marginal (gross) cost for X to provide a marginal increase in h is now:

$$\frac{\partial[-\delta(\beta+\alpha h)S]}{\partial h} = -\delta\Big\{\alpha S + (\beta+\alpha h)\frac{\partial S}{\partial h}\Big\}. \tag{43}$$

X can anticipate receiving the amount $\frac{\partial W(\Phi(h))}{\partial h} = \alpha S$ for every incremental rise in h he induces. For X to receive a *higher payoff than* Π_X^{SB} at $h = 1$, it must be the case that

$$\frac{\partial[\delta(\beta+\alpha h)S]}{\partial h} = \delta\big(\alpha S + f\frac{\partial S}{\partial h}\big) < \alpha S = \frac{\partial W(\Phi(1))}{\partial h}$$

This in turn implies:

$$\delta < \frac{\alpha S}{\alpha S + f\frac{\partial S}{\partial h}} \tag{44}$$

is *necessary* (but not sufficient!) for X to forego confidentiality in favour of sharing information. Note that if $\delta = 1$, as we have assumed previously, then this condition can never hold. A sufficient condition to forego confidentiality is $W(\Phi(1)) > \delta f S$. This implies $\delta < \frac{W(\Phi(1)}{fS}$. Lemma 5 implies that $\frac{W(\Phi(1))}{fS} < 1$, since otherwise information selling would always dominate confidentiality when $\delta = 1$, which it does not. Thus

$$\delta < \frac{W(\Phi(1))}{fS} \iff \frac{W(\Phi(1))}{fS} < 1 \text{ but } \frac{W(\Phi(1))}{fS} > 0 \text{ by definition.}$$

Therefore, there must be some $\delta \in (0,1)$ such that $W > \delta f S$. Call the cutoff level of the discount factor $\hat{\delta}$ so $W = \hat{\delta} f S$.

Proposition 5: If agents discount the future sufficiently more than does principal X (i.e. $\hat{\delta}$ is sufficiently low), the payoff to Principal X from the no-commitment equilibrium is strictly greater than that under a confidentiality law.

Lastly, consider possible commitment by Y. Ex ante, Y can always announce that she will pay an extra δS to those choosing contract 1 from X, of whom there will be α in an equilibrium inducing $h = 1$. Y's benefit from doing so at Stage 5 is αS; her cost is $\delta\alpha S$. Therefore, the net benefit to Y from this policy at stage 5 is $\alpha S(1-\delta)$. Moreover, the discount rate of agents, δ need not be exceptionally low in this case — simply less than 1. For any discount rate less than 1, we have the conclusion that when Y is able to commit to a policy which extra-compensates $(\overline{xy})$ types at the Y-contracting stage, X is able to fully separate agents along the x dimension *and* sell this information to Y to receive a payoff greater than Π_X^{SB}. For Y to commit to such a policy, however, she must be guaranteed a payoff greater than Π_Y^{SB}, which would require an ex-ante side payment from X.[23] Such a payment must be small enough for X to prefer commitment on the part of Y to his no-commitment payoff. Because X does not have to pay compensation in the Y-commitment scenario, and can still sell information for $p = W(\Phi(1))$, his payoff must be strictly greater than in the no-confidentiality, X-compensation case. He would therefore have a surplus he could split with Y in return for Y's credible commitment.[24]

As a final note on this model, consider the position of an $(\overline{xy})$ agent at the start of Stage 5, when X has extra-compensated these agents the amount $\delta S(1)$ when making his offer at Stage 1. Because this policy induces $h = 1$, it must be that all such agents choose contract 1 from X, and so receive the payoff $R(z_{\underline{y}1}(1)) < R(z_{\underline{y}}^{SB})$ at Stage 5. This situation gives rise to a posterior "complaint" from such agents: namely, that they would have done better had they not chosen contract 1 from X initially. That is, agents may be trapped by their own present-bias when making choices at Stage 2.

6 Example

To illustrate the mechanics of the model, this section presents a numerical example. A simple model is one in which $c(\cdot)$ is linear, and $r_x(\cdot)$ is concave. Let $r_x(z_x) = z_x^{\frac{1}{2}}$, $r_y(z_y) = \omega z_y^{\frac{1}{2}}$, $c_{\overline{\cdot}}(z) = \overline{\theta} z$ for x and y, and $c_{\underline{\cdot}}(z) = \underline{\theta} z$ for x and y. Note that $\overline{\theta} < \underline{\theta}$, and $R(z) = \overline{\theta} z - \underline{\theta} z = \Delta\theta z$. By applying the formulas described earlier:

$$\begin{aligned} z_{\overline{x}}^{FB} &= \left(\frac{1}{2\overline{\theta}}\right)^2 \qquad z_{\underline{x}}^{FB} = \left(\frac{1}{2\underline{\theta}}\right)^2 \\ z_{\overline{y}}^{FB} &= \left(\frac{\omega}{2\overline{\theta}}\right)^2 \qquad z_{\underline{y}}^{FB} = \left(\frac{\omega}{2\underline{\theta}}\right)^2 \end{aligned} \tag{45}$$

are the "first-best" levels of z,

$$z_{\underline{x}}^{SB} = \left(\frac{1}{2(\underline{\theta} + \Delta\theta(\frac{f}{1-f})}\right)^2 \qquad z_{\underline{y}}^{SB} = \left(\frac{\omega}{2(\underline{\theta} + \Delta\theta(\frac{f}{1-f})}\right)^2 \tag{46}$$

[23]Once stage 3 has been reached, Y would not be able to refuse a price of at least $p = W(\Phi(1))$. Accepting at this price would set Y's payoff at Π_Y^{SB}.

[24]This story may explain the popularity of "reward" programs, in which indirect information gathered from (here, consumer information), is accumulated to redeem for future rewards.

are the "second-best" levels of $z_{\underline{x}}$ and $z_{\underline{y}}$, and

$$\begin{aligned} z_{\underline{x}}(h) &= \left(\frac{1}{2(\underline{\theta}+\Delta\theta(\frac{f-\alpha(1-h)}{1-f+(1-h)\alpha}))}\right)^2 \\ z_{\underline{y}1}(h) &= \left(\frac{\omega}{2(\underline{\theta}+\Delta\theta(\frac{\alpha h}{f-\alpha}))}\right)^2 \\ z_{\underline{y}2}(h) &= \left(\frac{\omega}{2(\underline{\theta}+\Delta\theta(\frac{f-\alpha h}{1-2f+\alpha}))}\right)^2 \end{aligned} \tag{47}$$

are the optimal levels of $z_{\underline{x}}$, $z_{\underline{y}1}(h)$ and $z_{\underline{y}2}(h)$ chosen when X induces $h \in [\underline{h}, 1]$. The reader can verify that this set-up satisfies the condition:

$$\frac{\partial^2 z_{y2}}{\partial h \partial \omega} \geq 0 \text{ and } \frac{\partial^2 z_{y1}}{\partial h \partial \omega} \leq 0.$$

Calculations are made using two separate values of ρ and two separate values of ω, without $\delta = 1$. Results are given in Table 1. The last rows of the table refer to the *total* returns — for $r_x(\cdot)$ and $r_y(\cdot)$ attainable under different informational settings, and the aggregate payoffs to all principals and agents, respectively. Both of these are measures of aggregate welfare. There are two points in particular to notice. First, in cases A and B, if X cannot commit to confidentiality, he chooses to set $h = \underline{h}$ and selects the overall worst aggregate outcome (although the best for himself). However, X's payoff in this case is still lower than it would have been with confidentiality: U_X^{SB}. Thus, the role of commitment (or legislation) here is effectively to increase X's payoff at the expense of agents (see proposition 4). Second, in cases C and D, if X cannot commit to confidentiality, he chooses to set $h = 1$ and selects a *better* aggregate outcome than would obtain under confidentiality. The role of confidentiality is again to increase X's payoff, but here at the expense of agents and principal Y. However, confidentiality can produce a worse aggregate outcome than would have obtained under no-confidentiality, even though X's payoff is higher. The magnitude of this loss increases as the level of ρ rises; i.e. as Y's possible gains from learning information increase. The conclusion is that as ρ grows, other things equal, *confidentiality becomes less desirable from the perspective of increasing aggregate welfare.*

For the case of discounting, consider for the moment column B. It can be shown that if $\delta < 0.47$, X will find it in his interest to pay δS to $\overline{xy}$ agents, inducing $h = 1$ with no confidentiality. Then the ex-post (i.e. undiscounted) aggregate welfare level dominates that under confidentiality.

7 A Note about Consent

As discussed in the introduction, many modern confidentiality laws contain provisions that personal information must be kept confidential unless consent is given from the individuals concerned for it to be released. This paper has abstracted from this issue. If a confidentiality law contains a provision that individuals must give consent for their information to be released to principal Y from X, will this affect the solution of the model?

The critical issue appears to the timing of the "consent" question. At present, consent is automatically built into the contracting choice. Suppose δ is sufficiently small that X would find it in his interest not to commit to confidentiality. Then in this case, he essentially asks agents, "will you sign contract k, knowing that by doing so you give consent to release your contract information to Principal Y?" The next best option for an individual is to sign no contract, leaving a utility for that agent of zero at the X-contract stage. Because X will pay S in this case, to induce $h = 1$, agents will always be willing to consent ex ante, since staying out of the X contract cannot improve their overall payoff given Y's subsequent beliefs.

Suppose that in the same environment, X must legally obtain consent to sell information to Y *after* X-contracts have been signed. To model this feature, one can add two additional stages: stage 3(a), in which X decides whether or not to sell information, and stage 3(b), in which individual agents decide whether or not to consent to share their information. If all agents would give consent at this stage, then the game is identical to the one described above. Now, the crucial feature appears to be the behaviour of $(x\underline{y})$ types. Suppose all these types choose to consent — since they will receive a payoff of zero in either case — and all $(x\overline{y})$ types do not. Y will then observe β agents consenting who have chosen contract 1 and γ agents consenting who have chosen contract 2. She will then surmise that all agents not consenting are of type $(x\overline{y})$, and offer these agents a particularly unrewarding contract. This equilibrium would be unaffected by the timing of consent, because X, knowing all agents will consent to release their information, will always ask in stage 3(a).

A similar line of argument can be used in cases where δ is sufficiently high that X

has an incentive to commit to confidentiality. If consent must be given after contracts are signed, again all agents would consent. X would then want to sell information prior to this stage, causing agents to earlier behave according to h. But if this were the case X has an incentive to commit *never even to ask the question about consent*. In this sense, he must go even further than confidentiality-law would stipulate, telling agents at Stage 1: "When you sign this contract, I will never ask your permission to pass this information along to Y." X may then need a stronger confidentiality policy than the law stipulates, possibly imposing a penalty on himself for ever disclosing information.

The issue becomes more clouded when all $(x\underline{y})$ agents would not choose to consent if asked, or indeed if there are more than two types in each dimension. Such would be the case, if, for example, individuals had an *inherent* desire for confidentiality. Examination of this issue is left for future research.

8 Conclusion:

This study has used a dynamic contracting model to explain the incentives for sharing information by principals, when agents indirectly reveal their characteristics through their choices. When agents anticipate the future behaviour of principals and do not discount the future at any greater rate, they may choose "the wrong contract" from the first principal if not extra-compensated for revealing their information. The no-commitment equilibrium, in which the first principal cannot commit to keep contract information secret, displays the following results:

- Without extra-compensation, a proportion $(1 - h)$ of $(\overline{xy})$ types always choose their "non-targeted" contract from X, and this collective behaviour destroys the value of information for Y.

- X can increase the level of h by extra-compensating all $(\overline{x}y)$ types, but will only be induced to do so if Y is sufficiently less productive than X. Y can never fully compensate X for inducing $h = 1$, which would maximize the value of information to Y.

- Agents always fare at least as well under no-confidentiality as under confidentiality whether or not they are extra-compensated by X, while X always fares better under confidentiality.

- Confidentiality legislation provides a commitment device for principal X not to share information.
- Confidentiality legislation can increase aggregate welfare from the no-commitment equilibrium, especially if Y is relatively more productive than X. However, it can decrease welfare from the no-commitment equilibrium when X is relatively more productive than Y. This decrease appears to be more acute, the greater is the correlation between x and y.
- Only if δ — the agents' discount rate — is sufficiently small relative to principals X and Y, would X not want a confidentiality law. In this case, prohibiting information sharing can harm aggregate welfare.

The arguments of this paper are counterintuitive in the sense that they seem to run against the "popular" desire to keep information confidential. In this paper, contrary to one tenet of the Fair Information Principles doctrine, agents should be keen to disclose their information for purposes other than those for which it was gathered. Assuming rational actors does not lead agents to fear the revelation of their information, when they are sufficiently farsighted to understand the uses to which it would be put. Some explanations of this puzzle include: (a) agents are not rational with respect to the disclosure of their personal information, (b) principals do not set optimal policies in the face of agents' behaviour, (c) agents are naive about the use of their information, or uncertain about with whom it will be shared and for what purpose. The last explanation seems intuitively appealing, given the casual observation that personal information gathered by firms and governments seems to float freely in electronic databanks, and is shared covertly.

Appendix

Proof of Lemma 4: Suppose X offers two menus: Menu $A \equiv \{(z_x, s_x)^1, (z_x, s_x)^2, \ldots, (z_x, s_x)^K\}$ and Menu $B \equiv \{(z_x, s_x)^1, (z_x, s_x)^2, \ldots, (z_x, s_x)^K, (z_x, s_x)^{K+1}\}$. Menus A and B are identical except that B includes one option which is not in A. Suppose further that:

$$\sum_{k=1}^{K} \eta_A^X(k) = 1 = \sum_{k=1}^{K} \eta_B^X(k).$$

That is, $\eta_B^X(K+1) = 0$: when menu B is offered, no agent chooses contract $(K+1)$. Because the rest of the menus are identical, it must be that $\eta_A^X(k) = \sum_{k=1}^{K} \eta_B^X(k)$ for all $k = 1, \ldots, K$. Then Y must have the same beliefs about the proportions of $\overline{y}$ types who have chosen each contract, k; or, $\Phi(\Omega_b^A) = \Phi(\Omega_b^B)$. In words, when all agents are choosing one of K contracts, adding an additional contract which is preferred by no agent does not change the beliefs of principal Y.

Let $k = 1, \ldots, K$ denote the contracts offered by X, and consider agents of the type $(\overline{x}y)$. It must be that all such agents choose the *same* contract. Suppose they do not. By definition, $u_{\overline{xy}}^X(\cdot) = u_{\overline{x}\underline{y}}^X(\cdot)$, so the contract (z_x, s_x) gives each type the same "X" utility. If k and k' are designed so that $u_{\overline{xy}}^X((z_x, s_x)^{k\prime}|\sigma_{xy}^{-i}) > u_{\overline{x}\underline{y}}^X((z_x, s_x)^k|\sigma_{xy}^{-i})$, then all $(\overline{x}\underline{y})$ types would choose the contract k'. For $(\overline{x}\underline{y})$ types to continue choosing k, while $(\overline{xy})$ types choose k', it must be that $u_{\overline{x}y}^X((z_x, s_x)^k|\sigma_{xy}^{-i}) > u_{\overline{x}y}^X((z_x, s_x)^{k\prime}|\sigma_{xy}^{-i})$, and

$$u_{\overline{xy}}^X((z_x, s_x)^{k\prime}|\sigma_{xy}^{-i}) + R(Ez_{\underline{y}k'}|\sigma_{xy}^{-i}) > u_{\overline{xy}}^X((z_x, s_x)^k|\sigma_{xy}^{-i}) + R(Ez_{\underline{y}k}|\sigma_{xy}^{-i})$$

This condition implies that $R(Ez_{\underline{y}k'}|\sigma_{xy}^{-i}) > R(Ez_{\underline{y}k}|\sigma_{xy}^{-i})$. For this condition to be an equilibrium, Y must believe that $\Phi_k > \Phi_k'$ (by lemma 1). However, at Stage 5, Y is endowed with information Ω_e, and would only believe $\Phi_k > \Phi_k'$ if k and k' were offered such that some or all $(\underline{x}\overline{y})$ types chose k *and* some or all $(\underline{xy})$ types chose k'. To see why this is the case, note that if only $(\underline{x}\overline{y})$ types chose k, then after receiving Ω, Y would set $\Phi_k = 0$ and $\Phi_{k'} = 1$. However, for $(\underline{xy})$ types to choose contract k' — and not contract k — requires that $u_{\underline{xy}}^X((z_x, s_x)^{k\prime}|\sigma_{xy}^{-i}) > u_{\underline{xy}}^X((z_x, s_x)^k|\sigma_{xy}^{-i})$. But since $u^X(\cdot) = s_x - c_x(z_x)$, this condition implies $u_{\overline{x}y}^X((z_x, s_x)^{k\prime}|\sigma_{xy}^{-i}) > u_{\overline{x}y}^X((z_x, s_x)^k|\sigma_{xy}^{-i})$, a contradiction. Therefore it must be that $u_{\overline{x}y}^X((z_x, s_x)^k|\sigma_{xy}^{-i}) = u_{\overline{x}y}^X((z_x, s_x)^{k\prime}|\sigma_{xy}^{-i})$, meaning that k and k' are the same contract.

The fact that the contracts chosen by $(\underline{xy})$ and $(\underline{x}\overline{y})$ must be the same is established in a similar manner.

Therefore, from a menu of K contracts offered, agents will never choose more than two different contracts. But as established above, setting $K \geq 3$ contracts when all agents choose one of only two contracts implies that $K-2$ contracts are redundant. Since their presence does not affect the selling price of information, p, nor the current payoff of X (see eq. 5), X's problem is equivalent to only offering at most two contracts.
■

Proof of Proposition 1: To begin with, note that under this menu of contracts all agents will participate by choosing some contract. Let the two contracts be denoted 1 and 2, where 1 is targeted at $(\overline{x}y)$ types and 2 is targeted at $(\underline{x}y)$ types. It must be the case, from (28), that under this menu, all $(\overline{x}\underline{y})$ types choose 1 and all $(\underline{x}\underline{y})$ types choose 2. At a *minimum*, therefore, $\eta^X(1) = \beta$ and $\eta^X(2) = \gamma$. What remains is to see what contract the other types will choose. If an $(\underline{x}\overline{y})$ type chooses 1, he receives:

$$c_{\overline{x}}(z_{\overline{x}}^F) + R(z_{\underline{x}}^{SB}) - c_{\underline{x}}(z_{\overline{x}}^F) + R(Ez_{y1}|\Omega_b(\sigma_{xy}^{-i})) = R(z_{\underline{x}}^{SB}) - R(z_{\overline{x}}^F) + R(Ez_{y1}|\Omega_b(\sigma_{xy}^{-i}))$$

if he chooses 2, he receives:

$$0 + R(Ez_{y2}|\Omega_b(\sigma_{xy}^{-i})) = R(Ez_{y2}|\Omega_b(\sigma_{xy}^{-i}))$$

So a particular agent of this type will choose 1 if $[R(z_{\underline{x}}^{SB}) - R(z_{\overline{x}}^F)] + R(Ez_{y1}|\Omega_b(\sigma_{xy}^{-i})) > R(Ez_{y2}|\Omega_b(\sigma_{xy}^{-i}))$. Since the term in square brackets is negative, it must be certainly be the case that $Ez_{y1} > Ez_{y2}$. Note that $Ez_{y1} = Ez_{y2}$ is insufficient for any $(\underline{x}\overline{y})$ type to choose contract 1.

If an $(\overline{x}\overline{y})$ type chooses 2, he receives:

$$c_{\underline{x}}(z_{\underline{x}}^{SB}) - c_{\overline{x}}(z_{\underline{x}}^{SB}) + R(Ez_{y2}|\Omega_b(\sigma_{xy}^{-i})) = R(z_{\underline{x}}^{SB}) + R(Ez_{y2}|\Omega_b(\sigma_{xy}^{-i}))$$

if he chooses 1, he receives:

$$R(z_{\underline{x}}^{SB}) + R(Ez_{y1}|\Omega_b)(\sigma_{xy}^{-i}))$$

So they choose 2 if $R(Ez_{y2}|\Omega(\sigma_{xy}^{-i})) > R(Ez_{y1}|\Omega(\sigma_{xy}^{-i}))$, which must mean that for any $(\overline{x}\overline{y})$ type to choose contract 2, $Ez_{y2} > Ez_{y1}$. This is the opposite condition to induce $(\underline{x}\overline{y})$ types to choose contract 1, meaning that either some or all $(\overline{x}\overline{y})$ types choose 2 along with all $(\underline{x}\overline{y})$ types, or some or all $(\underline{x}\overline{y})$ types choose 1 along with all $(\overline{x}\overline{y})$ types.

To show that the latter case can never occur, recall that $z_{y1} > z_{y2}$ iff $\Phi_1 < \Phi_2$. This may also be written from (24) as:

$$\frac{\Pr\ ((z_x, s_x)^1, \Omega_b, \overline{y})}{\Pr\ ((z_x, s_x)^1, \Omega_b, \underline{y}))} = \frac{\overline{\phi}_1}{\underline{\phi}_1} < \frac{\overline{\phi}_2}{\underline{\phi}_2} = \frac{\Pr\ ((z_x, s_x)^2, \Omega_b, \overline{y})}{\Pr\ ((z_x, s_x)^2, \Omega_b, \underline{y}))}$$

If all $(\overline{x}\overline{y})$ types choose contract 1, and only a proportion $0 < g < 1$ of $(\underline{x}\overline{y})$ types choose contract 2 — and this behaviour is believed by Y upon seeing Ω — then $\overline{\phi}_1 = \alpha+\beta(1-g)$ and $\underline{\phi}_1 = \beta$. The condition $z_{y1} > z_{y2}$ requires:

$$\frac{\alpha+\beta(1-g)}{\beta} < \frac{\beta g}{\gamma} \iff g > \frac{\gamma(\alpha+\beta)}{\beta(\gamma+\beta)} = \frac{\gamma f}{\beta(1-f)}$$

The requirement $0 < g < 1$ means $\frac{\gamma f}{\beta(1-f)} < 1$. But $\frac{\gamma f}{\beta(1-f)} < 1 \iff \alpha - f^2 < 0$, or $\rho < 0$. Since $\rho > 0$, therefore $\Phi_1 \geq \Phi_2$ and $z_{y1} \leq z_{y2}$.

Suppose that all $(\overline{x}\overline{y})$ types choose contract 1 and all $(\underline{x}\overline{y})$ choose contract 2. From an earlier example, it was shown that Y would see $\eta^X(1) = f$ and $\eta^X(2) = 1-f$ in Ω, and would also know — since she can also see the contract menu offered by X — that β of the f agents choosing 1 are type $(\overline{x}\underline{y})$, and γ of the $(1-f)$ agents choosing 2 are type $(\underline{x}\underline{y})$. Since no proportion g of $(\underline{x}\overline{y})$ types would choose 1 under these beliefs, Y must reason that all types are choosing their "targeted" contract and set $z_{y2} > z_{y1}$. However, this case cannot be an equilibrium: if all other $(\overline{x}\overline{y})$ types are choosing 1, then an incentive exists for a particular $(\overline{x}\overline{y})$-type agent, i, to deviate and choose contract 2! This incentive only disappears when $z_{y2} = z_{y1}$, or:

$$\frac{\overline{\phi}_1}{\underline{\phi}_1} = \frac{\overline{\phi}_2}{\underline{\phi}_2}$$

Thus, let h be the proportion of $(\overline{x}\overline{y})$ agents choosing contract 1. Then $\overline{\phi}_1 = \alpha h$, $\eta^X(1) = \alpha h + \beta = f - \alpha(1-h)$ and

$$\Phi_1 = \frac{\alpha h}{f - \alpha(1-h)}$$

Similarly, $\overline{\phi}_2 = \beta + (1-h)\alpha = f - \alpha h$, $\eta^X(2) = 1 - f + \alpha(1-h)$ and

$$\Phi_2 = \frac{f - \alpha h}{1 - f + \alpha(1-h)}$$

Call $\underline{h}$ the level of h such that Y believes the proportion of $(x\overline{y})$ types choosing each contract is the same. Then

$$\underline{h} = \frac{\beta f}{\alpha(\gamma+\beta)} = \left(\frac{f-\alpha}{\alpha}\right)\left(\frac{f}{1-f}\right) < 1 \iff \rho > 0$$

Since Y updates her beliefs such that $z_{y2} = z_{y1}$, therefore $\Pi_Y^{split} = \Pi_Y^{SB}$, so $W(\Phi) =$

$W(\Phi_o) = 0$. The reader can verify that Φ_1 evaluated at $\underline{h}$ equals f, which is the same as the prior beliefs. Thus, the value of information to principal Y is zero. ■

Proof of Lemma 5: For any incremental increase in h, X must increase the payment S to any agent choosing contract 1. As established in the proof of Lemma 4, $u^X_{\overline{x}\overline{y}} = u^X_{\overline{x}\underline{y}}$; that is, X must also pay S to agents of type $(\overline{x}\underline{y})$ since they all choose contract 1. Thus, the incremental cost to X of increasing h is:

$$\begin{aligned}\frac{\partial[-(\beta+\alpha h)S]}{\partial h} &= -\left\{\alpha S + (\beta+\alpha h)\frac{\partial S}{\partial h}\right\} \\ &= -\left\{\alpha\left[R(z_{\underline{y}2}(h)) - R(z_{\underline{y}1}(h))\right]\right. \\ &\quad \left. +(\beta+\alpha h)\left[R'(z_{\underline{y}2}(h))\left(\frac{dz_{\underline{y}2}}{dh}\right) - R'(z_{\underline{y}1}(h))\left(\frac{dz_{\underline{y}1}}{dh}\right)\right]\right\} \qquad (48)\end{aligned}$$

The incremental compensation from Y is given by the change in the price Y pays for information. Using the envelope theorem since Π^{split}_Y is a maximum-value function:

$$\begin{aligned}\frac{\partial W(\Phi(h))}{\partial h} &= \frac{\partial \Pi^{split}_Y}{\partial h} \\ &= \left[r_y(z^F_{\overline{y}}) - c_{\overline{y}}(z^F_{\overline{y}})\right]\left(\frac{d\overline{\phi}_1}{dh} + \frac{d\overline{\phi}_2}{dh}\right) \\ &\quad -R(z_{\underline{y}1}(h))\left(\frac{d\overline{\phi}_1}{dh}\right) - R(z_{\underline{y}2}(h))\left(\frac{d\overline{\phi}_2}{dh}\right) \\ &= \left[r_y(z^F_{\overline{y}}) - c_{\overline{y}}(z^F_{\overline{y}})\right](\alpha - \alpha) - \alpha R(z_{\underline{y}1}(h)) + \alpha R(z_{\underline{y}2}(h)) \\ &= \alpha\left[R(z_{\underline{y}2}(h)) - R(z_{\underline{y}1}(h))\right] \\ &= \alpha S. \qquad (49)\end{aligned}$$

Since $\frac{dz_{\underline{y}1}}{dh} < 0$ and $\frac{dz_{\underline{y}2}}{dh} > 0$, $\frac{\partial S}{\partial h} > 0$; i.e. for h to rise, S must also rise. Therefore,

$$-\left[\frac{\partial[(\beta+\alpha h)S]}{\partial h}\right] + \frac{\partial W(\Phi(h))}{\partial h} = -(\beta+\alpha h)\frac{\partial S}{\partial h} < 0,$$

meaning that Y can never fully compensate X for increasing h for *any* $h \in (\underline{h}, 1]$. ■

Proof of Proposition 2: $\frac{\partial \Pi_X^{NC}}{\partial h}$ represents the change in X's equilibrium payoff as h changes. Thus, the slope of Π_X^{NC} is non-negative if:

$$\alpha\left[\left(r_x(z_{\overline{x}}^F) - c_{\overline{x}}(z_{\overline{x}}^F)\right) - \left(r_x(z_{\underline{x}}^{NC}(h)) - c_{\overline{x}}(z_{\underline{x}}^{NC}(h))\right)\right] \geq \left[(\beta + \alpha h)\frac{\partial S}{\partial h}\right] \tag{50}$$

and negative otherwise. Meanwhile:

$$\begin{aligned}
\frac{\partial^2 \Pi_X^{NC}}{\partial h \partial \omega} &= \frac{\partial^2 S}{\partial h \partial \omega} \\
&= R''(z_{\underline{y}2}(h))\left(\frac{\partial z_{\underline{y}2}}{\partial h}\right)\left(\frac{\partial z_{\underline{y}2}}{\partial \omega}\right) + R'(z_{\underline{y}2}(h))\left(\frac{\partial^2 z_{\underline{y}2}}{\partial h \partial \omega}\right) \\
&- R''(z_{\underline{y}1}(h))\left(\frac{\partial z_{\underline{y}1}}{\partial h}\right)\left(\frac{\partial z_{\underline{y}1}}{\partial \omega}\right) - R'(z_{\underline{y}1}(h))\left(\frac{\partial^2 z_{\underline{y}1}}{\partial h \partial \omega}\right)
\end{aligned} \tag{51}$$

It can be shown easily that $\frac{\partial z_{\underline{y}k}}{\partial \omega} > 0$ for $k = 1, 2$. Therefore, a *sufficient* condition for the rate of extra-compensation, S, to increase as ω increases is that claimed in the proposition. If so, then a sufficiently small ω will ensure that $\frac{\partial \Pi_X^{NC}}{\partial h}$ is everywhere upward-sloping, and a sufficiently large ω will ensure that it is everywhere downward-sloping. ■

References:

1. Armstrong, Mark and Jean-Charles Rochet (1999): "Multi-dimensional screening: A User's Guide", *European Economic Review*, 43, 959-979.

2. Bennett, Colin J. (1992): *Regulating Privacy: Data Protection and Public Policy in Europe and the United States*, Ithaca, N.Y.: Cornell University Press.

3. Bennett, Colin J. (2001): "What Government Should Know about Privacy: A Foundation Paper", mimeo.
(http://www.gov.on.ca/MBS/english/fip/pub/wgskap.doc)

4. Calzolari, Giacomo and Alessandro Pavan (2002): "Optimal Design of Privacy Policies", Mimeo, Department of Economics, Northwestern University.

5. Dillen, Mats and Michael Lundholm (1996): "Dynamic Income Taxation, Redistribution, and the Ratchet Effect", *Journal of Public Economics*, 59, 69-93.

6. Dranove, David et. al. (2002): "Is More Information Better? The Effects of 'Report Cards' on Health Care Providers", NBER Working Paper Series 8697.

7. Epstein, Richard A. (2000): "Deconstructing Privacy: And Putting it Back Together Again", in Ellen Krankel Paul, Fred D. Miller Jr., and Jeffrey Paul (eds.), *The Right to Privacy*, Cambridge: Cambridge University Press.

8. Epstein, Richard A. (2002): "HIPAA on Privacy: Its Unintended and Intended Consequences", *Cato Journal*, 22 (1), 13-31.

9. Freixas, X., R. Guesnerie and J. Tirole (1985): "Planning Under Incomplete Information and the Ratchet Effect" *Review of Economic Studies*, 52, 173-192.

10. Konrad, Kai (2001): "Privacy and Time-Consistent Optimal Labour Taxation", *Journal of Public Economics*, 79, 503-519.

11. Laffont, Jean-Jacques and Jean Tirole (1988): "The Dynamics of Incentive Contracts", *Econometrica*, 56 (5), 1153-1175.

12. Laffont, Jean-Jacques and Jean Tirole (1992): "Should Governments Commit?", *European Economic Review*, 36, 345-353.

13. Laffont, Jean-Jacques and David Matrimort (1997): "The Firm as A Multi-Contract Organization", *Journal of Economics and Management Strategy*, 6(2), 201-234.

14. Laffont, Jean-Jacques and David Matrimort (2002): *The Theory of Incentives: The Principal-Agent Model*, Princeton: Princeton University Press.

15. Linowes, David F. and Ray C. Spencer (1997): "How Employers Handle Employees' Personal Information: Report of A Recent Survey", *Employee Rights and Employment Policy Journal*, 1 (1), 153-172.

16. Taylor, Curtis R. (2002): "Private Demands and Demands for Privacy: Dynamic Pricing and the Market for Customer Information", *Duke Economics Working Paper 02-02*.

17. Thum, Claudio and Marcel Thum (2001): "Repeated Interaction and the Public Provision of Private Goods", *Scandinavian Journal of Economics*, 103 (4), 625-643.

18. Valentine, Deborah A. (2000): "Privacy on the Internet: The Evolving Legal Landscape", (http://www.ftc.gov/speeches/other/dvsantaclaraspeech.htm)

Table 1: Numerical Example

Case:	A	B	C	D
	$\omega = 2$ $f = 0.5$ $\alpha = 0.3$ $\underline{h} = 0.66$	$\omega = 2$ $f = 0.5$ $\alpha = 0.45$ $\underline{h} = 0.11$	$\omega = 0.5$ $f = 0.5$ $\alpha = 0.3$ $\underline{h} = 0.66$	$\omega = 0.5$ $f = 0.5$ $\alpha = 0.45$ $\underline{h} = 0.11$
Π_X^F	0.1875	0.1875	0.1875	0.1875
Π_X^{SB}	0.1666	0.1666	0.1666	0.1666
$\Pi_X^{NC}(h=1)$	0.1401	0.109	0.165	0.163
$\Pi_X^{NC}(h=\underline{h})$	0.1562	0.132	0.156	0.131
Π_Y^F	0.75	0.75	0.0468	0.0468
Π_Y^{SB}	0.666	0.666	0.0416	0.0416
$S(h=1)$	0.0590	0.216	0.0037	0.0135
$(\sum r(\cdot))^{FB}$	1.875	1.875	0.469	0.469
$(\sum r(\cdot))^{SB}$	1.666	1.666	0.416	0.416
$(\sum r(\cdot))^{NC}$, $h=1$	1.673	1.769	0.417	0.423
$(\sum r(\cdot))^{NC}$, $h=\underline{h}$	1.596	1.645	0.395	0.346
$(\sum \Pi + \sum U)^{FB}$	0.9375	0.9375	0.234	0.234
$(\sum \Pi + \sum U)^{SB}$	0.9027	0.9027	0.2255	0.2255
$(\sum \Pi + \sum U)^{NC}$, $h=1$	0.9028	0.9130	0.2256	0.2260
$(\sum \Pi + \sum U)^{NC}$, $h=\underline{h}$	0.8925	0.8594	0.2150	0.1820