PRIVACY, PROPERTY RIGHTS & EFFICIENCY: The Economics of Privacy as Secrecy

9 July 2004

Benjamin E. Hermalin
and
Michael L. Katz*

Abstract

There is a long history of governmental efforts to protect personal privacy and strong debates about the merits of such policies. A central element of privacy is the ability to control the dissemination of personally identifiable data to private parties. Posner, Stigler, and others have argued that privacy comes at the expense of allocative efficiency. Others have argued that privacy issues are readily resolved by proper allocation of property rights to control information. Our central findings challenge both views. We find: (a) privacy can be efficient even when there is no "taste" for privacy *per se*, and (b) to be effective, a privacy policy may need to ban information transmission rather than simply assign individuals control rights to their personally identifiable data.

* The authors would like to thank Thomas Davidoff, Hal Varian, Michael Waldman, and each other for helpful discussions of these issues. An earlier version of this paper was titled "Is Privacy Efficient? The Economics of Privacy as Secrecy."

I. INTRODUCTION

There is a long history of contentious policy debates and governmental efforts to protect personal privacy. A central element of privacy is the ability to maintain control over the dissemination of personally identifiable data—privacy as secrecy.[1] Recent technological developments in information collection and processing have heightened privacy concerns.[2] Today, Amazon.com knows what you like to read, your TiVo reports your viewing habits back to the company's central database,[3] and airlines keep a record of where you travel. Each year sees a number of privacy bills introduced in state legislatures and the U.S. Congress in response to these concerns, but there is little consensus on the appropriate approach.[4] There are many calls for strong governmental intervention to restrict the use of personally identifiable data. However, there are also calls simply to establish appropriate property rights to information on the grounds that market forces will then lead to efficient privacy levels that government fiat cannot be expected to achieve.

Broadly speaking, there are two reasons that an individual might wish to withhold personal information from others. First, an individual may have an intrinsic taste for privacy (*e.g.,* someone simply does not like the thought of strangers looking at her medical records or her neighbors reviewing her online pornography purchases). Second, an individual may wish to conceal information from a potential trading partner because revelation of the information would

---

[1] A distinct conception of privacy is autonomy, both from the state (*e.g.,* the right to choose to have an abortion) and from annoyance by other private parties (*e.g.,* the ability to be free of telemarketing calls). For an early discussion by an economist of privacy as autonomy, see Hirshleifer (1980).

[2] Even this trend is not new. Concerns about increasing surveillance and data processing led to the amendment of the California State Constitution in the early 1970s to include an explicit right to privacy. There is no explicit right to privacy in the United States Constitution.

[3] See, *e.g.*, "Greeting Big Brother With Open Arms," *New York Times*, January 17, 2004, B9.

lead the partner to take actions that would have adverse consequences for the person about whom the information was revealed.[5] For example, revelation that a policyholder smokes might lead to less favorable insurance rates, or someone who revealed he was HIV positive would, to many, be a less attractive sexual partner.

We are concerned with two questions. First, is privacy efficient? Second, is there an assignment of property rights to personally identifiable information that leads to an optimal level of privacy or disclosure? Clearly, the provision of privacy can be efficient when individuals have a taste for privacy. In what follows, we assume that there are no such tastes and, instead, focus on the second motive for privacy. We assume that both parties to the trading relationship are private agents and neither party has the power to compel the other to take action involuntarily. We also assume that both private parties are economically rational, expected-utility maximizers.

The central tenet of the rational-actor theory of information is that an agent will collect, disseminate, or conceal information to maximize his or her *private* benefits. Private and social incentives can diverge whenever an action changes an agent's *share* of the total net benefits. Consequently, distributional effects can lead to distortions in agents' incentives to manage information flows.

At least since the work of Hirshleifer (1971), it has been recognized that parties may invest inefficient amounts in collecting (or concealing) information in order to affect the

---

4 For a recent summary of federal legislation, see Smith (2003).

5 Another situation in which privacy concerns arise is one in which the individual wishes to prevent a trading partner from intentionally or unintentionally sharing information with a third party (*e.g.,* the sale of mailing lists or the failure to take adequate measures to secure a database of credit card numbers). Of course, the two cases are linked when the third party obtaining the information is also one of the individual's trading partners. For a recent analysis of third-party sharing, see Kahn *et al.* (2000).

distribution of rents.[6] For example, in order to increase its revenues from a given level of sales, a seller may expend resources to collect information about potential customers' willingness to pay for its product that is privately valuable but socially worthless. This observation suggests that there can be an efficiency-enhancing role for privacy regulations.[7] However, exactly the same logic suggests that there can be an efficiency-enhancing role for making illegal any effort to keep information private! The reason is that the parties who possess information that is unfavorable to them (*e.g.,* workers with low expected marginal revenue products) have incentives to expend resources to conceal that fact even though there may be no social value from doing so.

In many situations, the administrative or transactions costs associated with disseminating information to—or concealing it from—a trading partner are low. A more fundamental question is how revelation of information to a potential trading partner affects the efficiency of the equilibrium actions taken by the parties. The Chicago School, most notably Posner (1981) and Stigler (1980), asserts that privacy is harmful to efficiency because it stops information flows that would otherwise lead to improved levels of economic exchange. There are at least three mechanisms though which these adverse effects may arise. First, a lack of information can prevent the realization of matching benefits. For example, it might be efficient for an employer to provide the most extensive training to those employees with the best long-term health prospects. A privacy policy that limited the disclosure of health information would be an obstacle to such matching. Second, privacy can lead to informational asymmetries that destroy

---

6 For a recent analysis along these lines, see Taylor (2004).

7 This does not, however, imply that protecting privacy will promote efficiency. Depending on the elasticity of demand for information, implementing a privacy policy that raises the cost of collecting information might actually worsen the inefficiency by leading to higher levels of socially unproductive expenditures.

markets and prevent efficient exchange. This would be true, for example, when individuals have significant private knowledge about their likelihood of suffering a particular harm and insurance companies consequently face severe adverse selection problems.[8] Lastly, privacy can discourage productive investments. If one cannot reveal one's productivity, there can be less incentive to invest in increased productivity. An example would be policies that prohibit business school students from revealing their grades to potential employers.

Under the Chicago School view, more information is better, at least if costlessly obtained, and thus privacy is generally inefficient unless some parties have a demand for privacy for its own sake. One author summarized the situation as follows:

> In grossly oversimplified terms, the consensus of the law and economics literature is this: more information is better, and restrictions on the flow of information in the name of privacy are generally not social wealth maximizing, because they inhibit decisionmaking, increase transactions costs, and encourage fraud.[9]

A superficial analysis appears to support the Chicago position: Rational-actor models predict that, in the absence of transactions costs, perfectly informed parties will undertake all efficient trades, while it is well-established that imperfectly informed agents may fail to trade efficiently. This argument is, however, incomplete in three important respects.

First, welfare depends on more than *ex post* trade efficiency. Specifically, there can be *ex ante* efficiency effects on the provision of insurance and on investment incentives. Consider the effects on insurance. Privacy protection can create insurance that would otherwise be destroyed. For instance, if a potential policyholder can be tested for the likelihood of developing a fatal health condition, life insurance companies might demand to test potential policy holders and

---

[8] We observe that, in order to understand the full effects of privacy policies, one must also examine other potential market responses to privacy, such as insurance suppliers' relying on employer-purchased plans to reduce self-selection.

adjust prices according to the test results. The competitive equilibrium would be *ex post* efficient: each risk-averse person would purchase full insurance at an actuarially fair rate based on his or her test results. However, from an *ex ante* perspective, individuals would bear the risks associated with the outcomes of their test results. If testing—by either individuals or insurance companies—were banned, then the competitive equilibrium would entail all risk-averse individuals' buying full insurance at a common rate. Welfare would be greater than under the testing equilibrium both because the (socially wasteful) costs of testing would be avoided and because risk-averse individuals would bear less risk.

With respect to investment, privacy policies affect not only the dissemination of information, but also its acquisition or investments in complementary assets (*e.g.,* developing a statistical model of consumer behavior to target marketing efforts based on household characteristics). Specifically, if a party has to disclose any information that it has collected, then it has reduced incentives to collect the information.[10] For example, absent the ability to keep information confidential, people may not collect information about themselves (*e.g.,* individuals might forgo AIDS testing if disclosure were mandatory), resulting in unintended adverse consequences. When the information that would otherwise be collected has social value, this is a bad thing.[11] Similarly, policies that influence the cost of obtaining information will also influence the incentives to make complementary investments.[12]

---

9 Richard S. Murphy (1996) at 2382.

10 The structure of the argument is isomorphic to the logic of granting patents and other intellectual property rights.

11 Curiously, despite reaching his overall conclusion that privacy is harmful, Stigler (1980) also observes that disclosure can discourage efficient investment in obtaining information. He apparently failed to notice that this fact can be construed as an argument that privacy protection can be efficiency enhancing.

12 For a recent analysis in a related context, see Kahn *et al.* (2000).

A second consideration omitted from the Chicago School argument is that there may be other market imperfections that interact with privacy. Specifically, in the presence of price rigidities (say due to regulations or social norms), markets can fail to adjust efficiently to additional information. For example, rather than reducing an employee's wage offer in response to unfavorable information, a potential employer may simply refuse to make any offer at all. Consequently, increased information dissemination can reduce the efficiency of the resulting trading equilibrium.

In a sense, each of the two issues just summarized concerns an extension of the model covered by the Chicago School argument. The next point, however, goes to the heart of the argument itself. The supporting logic given above considers only the limiting case of full information and does *not* establish that intermediate increases in information are efficiency enhancing. One of the main contributions of the present paper is to demonstrate that, in fact, additional intermediate levels of information can reduce *ex post* trade efficiency.

There are two types of situation to examine: (a) pure price discrimination (*e.g.,* a seller learns something about the buyer's valuation, but the information tells the seller nothing about its costs), and (b) benefit-relevant information (*e.g.,* an insurance company obtains information about a potential buyer's health status or a firm learns about a potential employee's work record). The key distinction between these two cases is that, in the first but not the second, the transactions price is a sufficient statistic for calculating the benefits of trading with the specific partner.

In Section III, we examine price discrimination by a monopolist or monopsonist that seeks personally identifiable data concerning potential trading partners. This example is of interest, in part, because many privacy advocates have asserted that e-commerce and other technologies (*e.g.*, supermarket frequent-buyer cards) are going to lead to pervasive, inefficient price discrimination and that state intervention is necessary to prevent this outcome.[13] We show through a simple example that additional, intermediate levels information can raise or lower total surplus. We also show that the equilibrium outcome is independent of the assignment of rights to the personally identifiable data. While this latter result is reminiscent of the Coase (1960) Theorem, in that the assignment of property rights is irrelevant to the determination of total welfare, there are three important differences. First, in most applications, the Coase Theorem implies that the assignment of property rights affects the distribution of surplus but not the total. Here, the assignment of property rights has no effects on either the distribution or the total level of surplus. Second, the Coase Theorem often fails when, as here, the parties bargain under asymmetric information.[14] Third, the resulting outcomes in our applications are inefficient.

In Section IV, we examine a competitive market, which we describe in terms of an employment example. Privacy advocates strongly argue that workers should be protected from invasive questioning. But proponents of the Chicago School argue that efficiency is harmed by governmental limitations on employers' abilities to seek and act on information from or about potential employees.[15] Our analysis demonstrates that there are complicated tradeoffs missed by both sides of the debate. We also show that, here too, privacy rights in the form of intellectual property rights can be worthless.

The paper closes with a brief conclusion in Section V.

---

13 For discussions of the Internet and price discrimination, see Acquisti and Varian (2001) and Odlyzko (2003).

14 See, *e.g.*, Hermalin and Katz (2005) for a discussion.

15 See, *e.g.*, Posner (1981) at 405.

II. A MODEL

In this section, we describe the general model. We are interested in situations in which one side of the market, "firms," would like to learn the value of certain individually identifiable data concerning the other, "households." The labels, firms and households, are purely for expositional convenience. The analysis would apply equally well to any situation in which one party seeks information about another.

A type-θ household earns utility $u(\theta, t)$ from a transaction with a firm, where t is the monetary transfer *from the household to the firm*. It is common knowledge that each household knows its type. A household's outside opportunity (its payoff in the absence of trade with a firm) is $\underline{u}(\theta)$. A firm earns $\pi(\theta) + t$ from a transaction with a type-θ household. In the case of a firms selling a unit of some commodity at price p with marginal cost c that is independent of the buyer's type, we have $t = p$ and $\pi(\theta) = -c$. In the case of insurance, the cost could be a function of the buyer's type: $\pi(\theta) < 0$ would be the expected payable claim of a type-θ household. For an employer paying wage w to a risk-neutral worker with marginal revenue product θ, we have $t = -w$, $\pi(\theta) = \theta$, and $u(\theta, t) = w$.

There is an indicator variable, σ, that is informative with respect to a household's type, θ. We assume that σ takes on a finite number of values. We also assume each household knows the value of its indicator variable but firms cannot observe a household's indicator variable unless the information is released to them. Throughout, we assume that the indicator variable is hard evidence (*i.e.,* it can be concealed, but its value cannot be distorted or lied about if it is revealed). For example, a potential employee might be asked to release his or her medical records, or an e-tailer might seek to track a consumer's purchase history. In contrast, the value of θ is soft evidence, in that it is neither observable nor verifiable, so that a household can choose to misreport its value. For example, θ may be a measure of a household's willingness to pay for some product.

Many people have argued that the assignment of intellectual property rights to personally identifiable information is an important element of privacy policy.[16] Thus, in what follows, we are interested in comparing the equilibrium when firms have the right to compel revelation of the indicator variable with the equilibrium when households have the right to conceal this information if they choose. We will also examine what happens when firms are legally prevented from making use of σ, whether or not households would otherwise consent to its release and use.

The structure of the game is as follows. Firms simultaneously make offers. Firm *i* makes offer M_σ^i, which is a menu of options, or a mechanism, whose structure is conditional on whether a household's indicator variable is concealed (which, in a slight abuse of notation, we denote by $\sigma = 0$) and, if the variable is revealed, its value. Households then simultaneously decide whether to reveal their indicator variables and which, if any, offers to accept. Accepting an offer means that the household agrees to play according to the mechanism proposed by the corresponding firm. Our equilibrium concept is perfect Bayesian equilibrium.

III. MARKET POWER AND PRICE DISCRIMINATION

We begin by considering a market in which there is a single firm. Although the analysis applies equally well to a monopsonist, we describe the analysis in terms of a monopoly seller seeking information to serve as the basis of price discrimination.

---

[16] See, for instance, Varian (1997).

Our first result is that additional information can raise or lower total surplus. It is already known that additional information can harm total surplus if one does not allow the monopolist to make full use of the information (*i.e.,* if one restricts the set of allowable mechanisms). In particular, it is well known that moving from simple monopoly pricing to third-degree price discrimination has ambiguous welfare effects (see, *e.g.,* Varian, 1989). One can interpret allowing the monopolist to segment the market into two or more groups as an increase in information and thus the ambiguous welfare effects of third-degree price discrimination can be interpreted as ambiguous effects of increased information. However, because uniform pricing within each segment typically does not make full use of the information available to the monopolist, one cannot be sure that the welfare effects are due to changes in information or some ad hoc restriction imposed on the seller.[17] For this reason, we examine a model that permits the monopolist to employ a pricing mechanism that is optimal conditional on the information available to it.

Because the finding that additional information can raise or lower welfare clearly would continue to hold in more general models, we make this point in a very simple example. A monopoly seller faces two types of consumers, each of which has a linear demand curve with slope -1. As shown in Figure 1, the vertical intercept of an individual's demand curve is θ_L for a low-demand buyer and $\theta_H = \theta_L + \delta$ for a high-demand consumers, δ a positive constant. There are m low-demand buyers and n high-demand buyers. We have normalized demand so that the constant marginal costs of production are equal to 0.

[17] The restriction to third-degree price discrimination is as an example of the price rigidities discussed in the introduction.

By the revelation principle, the seller can do no better than to offer a set of options and let the buyers self-select among them. It is well known that the optimal mechanism for the seller to use presents the consumers with two options, purchase quantity x_L for a total payment of r_L or purchase quantity x_H for a total payment of r_H.[18] It is readily shown that a high-demand buyer must purchase at least as much as a low-demand buyer; that is, $x_H \geq x_L$.

As is also well known, the seller will induce the first-best level of consumption for high-value consumers: $x_H^* = \theta_H$. The reason is that, at any lower level of consumption, the monopolist could offer to sell additional output to its customers for an incremental amount equal to the relevant area under the demand curve of a high type. Clearly, doing so would not attract any low types to mimic a high type, and it would increase the seller's profits as long as the new quantity did not exceed θ_H.

As shown in Figure 1, the most that the monopolist could charge a low-demand buyer for x_L units is given by the area under the lower demand curve up to the purchase quantity, area *E* in the figure. The most that the seller can charge a high-demand buyer for $x_H^* = \theta_H$ is equal to the sum of areas *E*, *G*, and *J*. In particular, the seller must allow each high-demand customer to earn an information rent equal to area *F,* which is the surplus a high-demand consumer would enjoy from purchasing x_L for a total outlay equal to area *E*. Raising x_L increases *F* and thus imposes a cost on the seller. This cost induces the seller to reduce the consumption of each low-value customer below the first-best level, which is θ_L.

Assuming an interior solution, the first-order condition for the profit-maximizing

---

18 See Katz (1983) for details of the analysis.

quantity is $m(\theta_L - x_L) - n\delta = 0$, or $x_L^* = \theta_L - \frac{n}{m}\delta$. The resulting deadweight loss is

$\frac{m}{2}(\frac{n}{m}\delta)^2 = \frac{n^2}{m}\frac{\delta^2}{2}$ (which, when $x_L = x_L^*$, is equal to area *J* in Figure 1 multiplied by *m*).

For high enough values of $\frac{n}{m}$, there is no interior solution. Specifically, when

$\frac{\theta_L}{\delta} - \frac{n}{m} \leq 0$, it is privately optimal for the monopolist to exclude low-demand consumers from

the market, and the monopolist sets $x_L^* = 0$. The resulting deadweight loss is $m\frac{\theta_L^{\ 2}}{2}$. Observe

for future reference that this is the maximal equilibrium level of deadweight loss.

Now, suppose that, based on some indicator, the monopolist is able to divide the population into two sub-populations, $i = 1,2$, where group i has m_i low types and n_i high types. For example, the seller might be an e-commerce web site that is able to match its customers to certain demographic data (*e.g.*, home ownership) or identify whether customers have made particular purchases in the past.

As long as $\frac{n_1}{m_1} \neq \frac{n_2}{m_2}$, this division corresponds to an increase in the seller's information

and a reduction in the degree of informational asymmetry between the seller and the buyers. The improvement in the seller's information can be seen by considering the two sub-populations of buyers as being two urns from which realizations are drawn. The indicator variable can be seen as knowledge of the urn from which a buyer has been chosen. In the initial case—in which the seller does not have access to data on the values of consumers' indicator variables—it is as if

both urns have ratios of $\frac{n}{m}$. After the seller has access to the indicator variable, the two urns have different ratios.[19]

Proposition 1: *Consider an improvement in the seller's information as described above.*

(a) If $\frac{n}{m} \geq \frac{\theta_L}{\delta}$*, then total surplus weakly rises.*

(b) If $\max_i\{\frac{n_i}{m_i}\} \leq \frac{\theta_L}{\delta}$*, then total surplus strictly falls .*

(c) If $\frac{n}{m} < \frac{\theta_L}{\delta} < \max_i\{\frac{n_i}{m_i}\}$*, then total surplus may rise or fall depending on the parameter values.*

Proof: Consider each case in turn:

(a) Absent the improvement in information, the monopolist sets $x_L^* = 0$, and deadweight loss is at its maximal equilibrium level.

(b) The change in deadweight loss due to the additional information is equal to $\delta^2/2$ times

$$\frac{n_1^2}{m_1} + \frac{n_2^2}{m_2} - \frac{(n_1 + n_2)^2}{m_1 + m_2} = \frac{(n_1 m_2 - n_2 m_1)^2}{m_1 m_2 (m_1 + m_2)} \geq 0 \ ,$$

with strict inequality unless $\frac{n_1}{m_1} = \frac{n_2}{m_2}$.

(c) If the seller is able to sort the two types perfectly (*e.g.*, $m_1 = m$, $n_2 = n$, and $n_1 = 0 = m_2$), then there is no deadweight loss following the information improvement. Instances in which the information improvement lowers total surplus can be constructed by making use of (b) and considering values of $\max_i\{\frac{n_i}{m_i}\}$ that are just above $\frac{\theta_L}{\delta}$. **Q.E.D.**

---

19 It is readily shown that this is an improvement in information in the Blackwell (statistical) sense. It can be shown that there exists a garbling matrix that maps the assignment probabilities of the second case into those of the first case, which means the second is more informative than the first.

Although we analyze splitting households into two subgroups, the analysis could be iterated to treat any number of subgroups to capture additional information improvements.

The intuition underlying the finding that information can be harmful is that the increased information gives the monopolist greater potential market power, which it then exercises in an inefficient manner. Although total surplus falls, the seller's *share* rises by an amount more than sufficient to offset the fall in the total.

This analysis demonstrates that both sides of the e-commerce privacy debate have overstated their cases. Proposition 1 shows that those who assert that additional information gathering necessarily will improve efficiency are incorrect. Thus, there might appear to be scope for efficient governmental intervention. However, claims by privacy advocates that consumers necessarily are harmed by the loss of privacy are also incorrect. A simple corollary shows that increased information collection may benefit or harm consumers, depending on the specifics of the market under consideration.

Corollary: *In some cases an improvement in the seller's information raises equilibrium consumer surplus, and in other cases the improvement lowers equilibrium consumer surplus.*

Proof: Regardless of the seller's information, low-demand consumers derive no surplus in equilibrium, so our interest is in high-demand consumers. Consider two scenarios.

$\frac{n}{m} > \frac{\theta_L}{\delta}$: Absent access to the indicator variable, the seller sets $x_L^* = 0$, and high-demand consumers enjoy no surplus (area *F* in Figure 1 collapses). Suppose that an improvement in information leads to $0 < \frac{n_i}{m_i} < \frac{\theta_L}{\delta}$ for at least one sub-population. Now, high-demand consumers in that sub-population earn positive surplus. Because the surplus of the high-demand consumers in the other sub-population can't be less than zero, total consumer surplus has increased.

$\frac{n}{m} < \frac{\theta_L}{\delta}$: Absent access to the indicator variable, the seller chooses $x_L^* > 0$, and high-demand consumers enjoy positive surplus equal to area *F*. Suppose that $n_1 = 0$ and $\frac{n_2}{m_2} > \frac{\theta_L}{\delta}$. In this case, neither consumer type captures any surplus, and the information improvement lowers equilibrium consumer surplus.[20] ***Q.E.D.***

This analysis tells us that a simple ban on the seller's collecting and using personally identifiable information may raise or lower equilibrium total surplus. Similarly, the ban may raise or lower equilibrium consumer surplus.

Some people have argued that consumer welfare could be maximized by giving households property rights to their personally identifiable information. In many cases, such rights would, however, be worthless because a seller with sufficient market power can *costlessly* compel revelation of the information. Specifically, the monopolist can: (1) refuse to sell output to any consumer who does not reveal the sub-population to which he or she belongs; and (2) offer the menus described above, conditional on sub-population membership, to each consumer who does reveal the sub-population to which he or she belongs. For example, an e-merchant could design its web site so that consumers must enroll and provide personal information before being allowed to shop. The resulting outcome will be identical to the case above in which disclosure of a consumer's demographic characteristic is mandatory.

This result holds more broadly than our simple example. Formally,

---

20 One can also find examples in which increased information lowers consumer surplus and m_1, m_2, n_1, and n_2 all are strictly positive. The formula for equilibrium consumer surplus is

$$\delta \sum_i \left[n_i \max\{0, \theta_L - \frac{n_i}{m_i}\delta\} \right].$$

Proposition 2: *Suppose that there is single firm, which can commit to offering mechanisms on a take-it-or-leave-it basis. The equilibrium outcome is the same whether the property rights to personally identifiable data are given to the firm or to households.*

Proof: First, suppose that the firm can directly compel revelation of the indicator variable and offers mechanism M_σ, which is contingent on the observed value of σ. Define $V(M,\theta)$ as the value that a type-θ household derives under mechanism M. The firm must choose a mechanism that satisfies the household's participation constraint: $V(M_\sigma,\theta) \geq \underline{u}(\theta)$.[21] Clearly, offering a family of mechanisms contingent on σ is weakly more profitable than offering a single, non-contingent mechanism.

Now, suppose that households have the property (privacy) rights to their indicator variables. That is, the firm must induce revelation of the indicator variable. Suppose the firm makes the following take-it-or-leave-it offer to households: (a) reveal σ and be offered mechanism M_σ as above, or (b) refuse to reveal σ and be refused a trade. A household decides whether to reveal its information by comparing $V(M_\sigma,\theta)$ with $\underline{u}(\theta)$. Thus, the original participation constraint implies that the consumer will choose to reveal the value of the indicator variable. In other words, the firm can costlessly induce revelation and thus will choose to do so. ***Q.E.D.***

The surprising part of this result is not that revelation is induced, but that there is no cost to the monopolist to do so. In other words, the assignment of property rights generates no information rents for consumers. This result reflects a critical difference between hard and soft

---

[21] Because the space of mechanisms can be expanded without loss of generality to include the "no-trade" mechanism, there is no loss of generality in assuming that participation constraint is met in equilibrium.

information—with hard information there is no way for one type to mimic another's information. This result demonstrates that assigning privacy rights is not enough to protect consumers or promote efficiency.

IV. COMPETITIVE MARKETS AND ADVERSE SELECTION

In the previous example, the exercise of market power was the source of distortion. We next examine the adverse effects of incremental information in a competitive market where the personally identifiable information is directly payoff relevant to a trading partner. As above, because the finding that additional information can raise or lower welfare clearly would continue to hold in more general models, we make this point in a simple example. Specifically, we consider a competitive employment market in which there are m low-quality workers, each with ability θ_L, and n high-quality workers, each with ability θ_H, where $\theta_L < \theta_H$. We assume both m and n are strictly positive and that there are more than $m + n$ potential employers, each of whom hires at most one employee.[22]

A worker's utility is $\theta(1\text{-}x) + y$, where $x \in \{0,1\}$ denotes whether she is employed in this sector ($x = 1$) or uses her time in some other way ($x = 0$), and y is income. Observe that she is willing to accept a job if the wage offered is at least θ. Letting the worker's reservation wage be correlated with her ability reflects that the value of her outside option (*e.g.,* becoming self-employed or working in some other sector) is likely an increasing function of her ability. Let the proportion of high-ability (*i.e.*, $\theta = \theta_H$) workers in a given set be denoted by λ. For the entire

[22] One can equivalently imagine this as a used car market in which workers are the sellers of used cars and employers are potential buyers of used cars along the lines of Akerlof's (1970) seminal article.

population in our example, this proportion is $\lambda^p \equiv \frac{n}{m+n}$. Denote the average worker ability when the proportion of high-ability workers is λ as $\theta_A(\lambda) \equiv \lambda\theta_H + (1-\lambda)\theta_L$.

An employer's profit is $v\theta x - y$, where θ is the realized ability of the worker hired and $v > 1$ is the marginal revenue product of ability. Note that, because $v > 1$, the first-best outcome entails all $m + n$ workers' being employed in this sector. An employer cannot observe θ directly and must, instead, form a prediction of its value. A rational employer will pay a worker up to $v\theta_A(\lambda^e)$, where λ^e is the employer's belief about the proportion of high-ability job applicants conditional on the information available to him. To make the problem nontrivial, we assume that the amount rational employers would bid for workers known to be of low productivity is less than the reservation wage of a high-quality employee: $v\theta_L < \theta_H$.

Employers simultaneously bid for workers by announcing wage offers, where the wage offer to a worker can be made conditional on the worker's personally identifiable data if they are available to the employer. For example, if the data are available, wages can be made contingent on past employment or the potential employee's health status. Once the wage offers have been made, workers decide which, if any, offers to accept.

We now characterize the competitive equilibrium for a population where employers know λ and have no additional information. We consider two cases, which depend on whether λ is greater or less than $\lambda^* \equiv \frac{\theta_H - v\theta_L}{v(\theta_H - \theta_L)}$. In the *high-productivity case,* $\lambda > \lambda^*$ and, hence, $v\theta_A(\lambda) \geq \theta_H$. In words, the average productivity of workers in this group is greater than the reservation wage of a high-ability worker. Given this relationship, there exists an equilibrium in which all workers accept a wage of $v\theta_A(\lambda)$.[23] Because all workers are employed, the outcome is efficient. Thus, the asymmetric information about worker ability does not adversely affect the market outcome.

In the *low-productivity case*, $\lambda < \lambda^*$ and, thus, $\theta_H > v\theta_A(\lambda)$. In the low-productivity case, the asymmetry of information leads to an adverse-selection problem. If employers thought all workers would be in the market, the most that firms would bid is $v\theta_A(\lambda)$, which is less than θ_H. Consequently, high-ability workers would exit the labor force. In equilibrium, employers anticipate this exit, so the equilibrium wage offer is $v\theta_L$, and high-quality workers are unemployed. The resulting deadweight loss is $n(v\text{-}1)\theta_H$.

Our interest is in the role of information. Suppose that the indicator variable takes one of two possible values. Label the two resulting sub-populations identified on the basis of σ as 1 and 2, and let λ_i denote the proportion of high-ability workers in sub-population *i*.

Suppose that, initially, $\lambda_1 = \lambda^p = \lambda_2$. Then the indicator variable is uninformative, and revelation of the indicator variable has no effect on efficiency. Now consider the following thought experiment. Exchange a low-ability worker from sub-population 2 with a high-ability worker from sub-population 1. λ_2 rises and λ_1 falls. It is readily shown that this exchange corresponds to an increase in the informativeness of the indicator variable. Additional exchanges would similarly increase information. As we will now show, the welfare effects of these information increases depend on the ranges in which various parameters fall.

---

23 Because $v\theta_L < \theta_H$, there is also a perverse Nash equilibrium in which employers expect workers of ability θ_H not to seek employment and, thus, employers never bid above $v\theta_L$. However, this is not Bayesian perfect under the market structure assumed here. If an employer deviated and offered a wage of $\theta_H + \varepsilon$, ε an arbitrarily small positive number, then all workers would be willing to be employed by that firm and the deviating employer would earn positive expected profit for sufficiently small ε because $v\theta_A(\lambda) - \theta_H > 0$.

First, suppose that $\lambda^* < \lambda^p$. Then the equilibrium with the indicator variable concealed is efficient and additional information is weakly harmful. As long as $\lambda^* < \lambda_1$, the increased information has no effect. But suppose $0 < \lambda_1 < \lambda^* < \lambda^p < \lambda_2$. When firms can make offers conditional on the indicator variable, workers in sub-population 2 all are hired at wage of $v\theta_A(\lambda_2)$. But the equilibrium wage for sub-population 1 is $v\theta_L$, and high-ability workers in that group are inefficiently unemployed. The additional information destroys what would otherwise have been efficient pooling. Observe, however, each additional exchange of this type would both increase the informativeness of the indicator variable and also increase total surplus at the margin because it would reduce the number of unemployed high-ability workers in sub-population 1. However, until one got to the point where $\lambda_1 = 0$, the level of total surplus would still be lower than when the indicator variable was concealed.

Next, suppose that $\lambda^p < \lambda^*$. In this case, the equilibrium with the indicator variable concealed is inefficient: no high-ability worker is employed. Hence, improved information weakly increases total surplus. Consider a sequence of improvements in information as described above. As long as $\lambda_2 < \lambda^*$, the outcome is unchanged. However, if $0 < \lambda_1 < \lambda^p < \lambda^* < \lambda_2$, high-ability workers in sub-population 2 are efficiently employed. Moreover, additional shifts of workers that increase the homogeneity of each sub-population lead to increased total surplus as the number of (employed) high-ability workers in sub-population 2 rises and the number of (unemployed) high-ability workers in sub-population 1 falls.

In summary, although complete employer information about worker types induces the efficient outcome, total surplus is a non-monotonic function of intermediate levels of information.

Proposition 3: *As long as* $0 < \lambda_1 < \lambda^* < \lambda^p < \lambda_2$*, releasing personally identifiable data strictly lowers total surplus. In other cases, releasing these data has no effect or raises total surplus.*

In addition to characterizing total surplus effects, we can examine the distributional consequences of information flows from workers to employers. Employers always make zero expected profits conditional on their information. In the high-productivity case, low-ability workers are made increasingly worse off by more information when $\lambda_1 < \lambda^* < \lambda^p$. The reason is that members of the sub-population are no longer paid a pooled wage, and, as λ_1 falls, a higher percentage of low-ability workers are in sub-population 1. In other circumstances, low-ability workers can gain from increased information. In particular, they gain when release of the indicator variable leads some of them to be pooled with high-ability workers and such pooling would not arise absent revelation of the indicator variable (*i.e.*, when $\lambda^p < \lambda^* < \lambda_2 < 1$). Moreover, continued increases in λ_2 reduce the number of low-ability workers benefiting from pooling, but increases the wage, $v\theta_A(\lambda_2)$, of those who are pooled. It can be shown that the net effect is ambiguous.[24]

High-ability workers gain from increased information when $\lambda^* < \lambda_1 < \lambda^p < \lambda_2$. In this case, everyone is employed with or without release of the indicator data, but high-ability workers are less co-mingled with low-ability workers and thus high-ability workers appropriate a higher

---

[24] Tedious calculation demonstrate that the sign of the change is equal to the sign of $m_2 - n_2 - 2$.

percentage of their revenue product for themselves. In this case, the interests of high- and low-ability workers are opposite one another. In other cases, however, both types may gain. Specifically, when $\lambda_1 < \lambda^p < \lambda^* < \lambda_2 < 1$, both high- and low-ability households gain from being pooled in sub-population 2 in comparison with the original situation in which the equilibrium wage was $v\theta_L$ and only low-ability workers were employed. Households in sub-population 1 are unaffected. In yet other cases, both types may be harmed by increased information. This happens, for example, when $\lambda_1 < \lambda^* < \lambda^p < \lambda_2$ but λ_1 is almost equal to λ_2. In this situation, the dominant effect is the collapse of pooling in sub-population 1 and both types of household see their expected wages fall.

Summarizing the analysis of expected wages, we see again that there is no clear policy even if the goal is simply raising workers' welfare.

Corollary: *Depending on the parameter values, an improvement in employers' information can: (a) benefit both types of worker; (b) harm both types of worker; or (c) benefit high-ability workers at the expense of low-ability workers.*

One might wonder whether employers would have incentives to compete with one another by respecting privacy. If workers had an explicit taste for privacy, the answer could be yes. But, in the present setting, any such employer would be unable simultaneously to attract a worker and earn non-negative profits unless it offered a wage equal to $v\theta_A(\lambda_1)$, where $\lambda_1 < \lambda_2$. This can be seen as follows. If the firm offered a wage above $v\theta_A(\lambda_1)$ but below $v\theta_A(\lambda_2)$, it would attract only workers in sub-population 1 and would lose money on average. If the firm offered a wage greater than or equal to $v\theta_A(\lambda_2)$, it would attract all workers and lose money on average. In other words, any firm that unilaterally respected privacy would have to make wage offers based on the assumption that its workers had the less favorable value of the indicator variable.

Now, consider a public policy of assigning to workers the property rights to their personally identifiable information, as many writers have advocated.[25] In the previous section, we saw that such rights are worth nothing when the uninformed side of the market has sufficient market power. We will now demonstrate that a parallel result holds even when firms are competitive, as in our employment example. The reason is that each household chooses to reveal information based on his or her self-interest. Consequently, households may have incentives to disseminate their information in ways that harm other households, as well overall efficiency.[26]

Suppose that workers have the right to keep their health status secret from potential employers, and consider the high-productivity case in our example. As shown above, if no worker revealed her health status, the equilibrium wage would be $v\theta_A(\lambda^p)$. But a worker who chose to reveal her good health status (*i.e.*, that she is in sub-population 2) would receive an offer of $v\theta_A(\lambda_2)$. In the high-productivity case, $v\theta_A(\lambda_2) > v\theta_A(\lambda^p) > \theta_H$, so all workers with good health status would reveal their health status. An employer would believe that any worker who declined to reveal her health status must have poor health. Hence, there would be complete information (indicator) revelation even if the workers had the right to keep their health status from potential employers.

---

25 For example, Shapiro and Varian (1997, pages 29 and 30) argue that:

> The right way to think about privacy, in our opinion, is that it is an externality problem. I may be adversely affected by the way people use information about me and there may be no way that I can easily convey my preferences to these parties. The solution to this externality problem is to assign property rights in information about individuals to those individuals. They can then contract with other parties, such as direct mail distributors, about how they might use the information.

The next result makes this argument more generally.[27]

Proposition 4: *Suppose that firms are competitive. Then the equilibrium outcome is the same whether the property rights to personally identifiable data are given to the firm or to households.*

Proof: Let $\pi^e(\sigma)$ denote a firm's expected value of $\pi(\theta)$ conditional on σ. Order the I values of σ so that $\pi^e(\sigma_i) < \pi^e(\sigma_j)$ if and only if $i < j$.[28]

If firms can compel revelation of σ, then they will do so and competition among firms will lead them to offer $t = -\pi^e(\sigma)$ to a household with indicator value σ (*i.e.*, a *payment* of $\pi^e(\sigma)$ to a household with indicator value σ).

Now, suppose firms don't know a household's value of σ unless the household chooses to reveal it. Suppose, counterfactually, that there exists an equilibrium in which households with value σ_I do not reveal their indicator variables. They cannot be the only group that conceals their indicator variables; if they were, the payment to households who conceal their indicator variables would be $\pi^e(\sigma_I)$, and all households would choose to conceal. If at least one other group conceals its indicator variable in equilibrium, then σ_I-households must receive a payment that is less than $\pi^e(\sigma_I)$. But then a firm could profitably deviate by offering to pay $\pi^e(\sigma_I) - \varepsilon$, where ε is a small positive number, to any household that revealed it was type σ_I: for sufficiently small ε, the deviating firm would induce σ_I-households to reveal themselves and the

---

26 Thus, the argument presented by Shapiro and Varian (1997) is incomplete. Their proposed assignment of rights does not solve the cross-consumer externality.

27 For an early application of this type of unraveling argument see, *e.g.,* Grossman (1981).

28 Recall that the set of possible values of σ is finite. If two different values of σ give rise to the same value of $\pi^e(\sigma)$, there is no loss of generality in treating them as being the same.

firm would earn positive profits. A similar argument can be applied inductively to households with other values of σ_i. ***Q.E.D.***

Observe that the statement of Proposition 4 is similar to that of Proposition 2, but the underlying logic is very different. In each case, there are no information rents associated with hard information. But in the first case, this is due to market power, and the second it is due to externalities across household types.

V. CONCLUSION

With so many people making extreme claims in discussions of privacy and related public policy, and with so little understanding of the underlying economics, it is important to identify the fundamental forces clearly. Our central finding is that, contrary to the Chicago School argument, the flow of information from one trading partner to the other can reduce *ex post* trade efficiency when the increase in information does not lead to symmetrically or fully informed parties.

Turning to public policy, there are many issues to consider before reaching conclusions beyond the general one that one should proceed with caution. However, our analysis shows that, whether the uninformed side of the market is monopolized or is perfectly competitive, the assignment of privacy rights to personally identifiable information may have no effect on agents' equilibrium welfare levels and need not lead to an efficient equilibrium privacy level. In some situations, the only effective policy would be an outright ban on the dissemination and use of such information. However, our analysis also demonstrates that it is extremely difficult to determine when such a ban would increase efficiency.

REFERENCES

Acquisti, Alessandro and Hal R. Varian (2002) "Conditioning Prices on Purchase History," unpublished manuscript, University of California, Berkeley.

Akerlof, George A. (1970) "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics,* **84(3):** 488-500.

Coase, Ronald (1960) "The Problem of Social Cost," *Journal of Law and Economics*, **3(1):** 1-44.

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

Hermalin, Benjamin E. and Avery W. Katz (2005) "The Law & Economics of Contracts," *The Handbook of Law and Economics* (A.M. Polinsky & S. Shavell, eds.), Amsterdam: North-Holland.

Hirshleifer, Jack (1971) "The Private and Social Value of Information and the Reward to Inventive Activity," *The American Economic Review,* **61(4):** 561-574.

Hirshleifer, Jack (1980) "Privacy: Its Origin, Function, and Future"*The Journal of Legal Studies,* **9(4):**649-664.

Kahn, Charles M., James McAndrews, and William Roberds (2000) "A Theory of Transactions Privacy," Working Paper 2000-22, Federal Reserve Bank of Atlanta.

Katz, Michael L. (1983) "Non-uniform Pricing, Output and Welfare under Monopoly," *Review of Economic Studies,* **L:** 37-56.

Murphy, Richard S. (1996) "Property Rights in Personal Information: An Economic Defense of Privacy," *Georgetown Law Journal* **84:**2381-2417.

Odlyzko, Andrew (2003) "Privacy, economics, and price discrimination on the internet," ACM, *Fifth International Conference on Electronic Commerce*.

Posner, Richard A. (1981) “The Economics of Privacy,” *American Economic Review* **71(2):** 405-409.

Shapiro, Carl, and Hal R. Varian (1997) “US Government Information Policy,” unpublished manuscript, University of California, Berkeley.

Stigler, George J. (1980) “An Introduction to Privacy in Economics and Politics,” *The Journal of Legal Studies,* **9(4):**623-644.

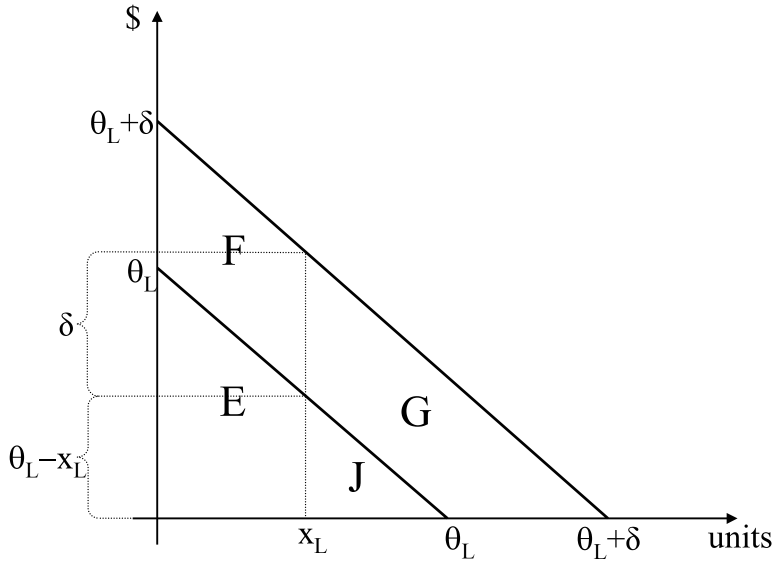
Smith, Marcia S. (2003) “Internet Privacy: Overview and Pending Legislation,” Report for Congress, Congressional Research Service, The Library of Congress, updated February 6, 2003.

Taylor, Curtis R. (2004) “Privacy and Information Acquisition in Competitive Markets,” unpublished manuscript, Duke University.

Varian, Hal R. (1989) “Price Discrimination” in *Handbook of Industrial Organization*, Vol. 1, R. Schmalensee and R. Willig, eds., Amsterdam: North-Holland.

___________ (1997) “Economic Aspects of Personal Privacy,” in *Privacy and Self-regulation in the Information Age,* Washington, D.C.: U.S. Department of Commerce.

Figure 1


$
$\theta_L+\delta$
F
θ_L
δ
E
G
θ_L-x_L
J
x_L
θ_L
$\theta_L+\delta$
units