The Law and Economics of Consumer Privacy Versus Data Mining

Peter H. Huang*

Law School
University of Pennsylvania
3400 Chestnut Street
Philadelphia, PA 19104-6204
E-mail: phuang@oyez.law.upenn.edu
(215) 573-6018

First Draft: February 8, 1997

This Version: May 27, 1998

Abstract: The collection, storage, processing, recombination,[1] and (re)sale of consumers' data has grown tremendously. Consumer data is now usually maintained in on-line databases and collected without remuneration towards or even the informed consent of consumers. The mining of such consumer data raises a host of legal and economic issues. This paper considers the tradeoff between allocative efficiency from matching direct mail advertising with demographic niches versus consumer privacy; potential misuse or abuse of data mining, ranging from outright criminal theft to inaccurate or outdated records; piggy-backing of law enforcement agencies on privately created data mines; whether default rules about data mining effectively become mandatory; and finally, who does and should own such data as consumers' names, addresses, phone numbers and purchase histories.

Traditional consumer protection law deals with such problems of fraudulent business practices as deceptive advertising and misrepresentation in product warranties. Such laws have the economics of information as their common intellectual basis.[2] This paper applies psychological games and behavioral economics to investigate distinct justifications for and types of regulation of data mining. This paper also considers personhood and commodification perspectives critical of applying economic analysis to study legal rules and institutions, in particular, treating consumers' data as private property.

Introduction

In the movie, The Net,[3] Sandra Bullock portrays a systems analyst, whose life is threatened by a covert organization, whose members learn about her vacation travel plans, favorite movie, and other personal information from her computerized credit card purchase history. They alter her computerized records at the California DMV, create a criminal history for her in the LAPD database, alter her friend's medical records in a hospital's computer, and replace her on-line identity with that of an entirely fictitious one they create. While the movie is hopefully fictional, there are many ways in which the Net has the potential to threaten consumers' informational privacy.[4] Advances in the computerized recording of consumer information, as well as decreases in the costs of its collection, storage, processing, recombination or manipulation in conjunction with other databases have altered such existing businesses as direct mail marketing and created a whole new industry engaged in the sale of consumer profiles based upon data mining.

Some irate consumers have called for the regulation of these data mining intrusions into their privacy. But, consumers' privacy interests are not absolute; they must be weighed against such other interests as the convenience to consumers of using credit cards on the Net and receiving e-mail “catalogues”; efficiencies from better matching of demographic and preference niches with businesses which cater to such groups; and reduced criminal enforcement costs from having more knowledge about

consumers. A balancing analysis is prudent before undertaking any regulation, so as to determine whether and what regulation is called for.

The purpose of this paper is to apply behavioral economics to consider distinct justifications for and types of data mining regulations. The first justification for and type of data mining regulation involves a psychological game-theoretic model of self-regulation through changing industry norms of consumer privacy. Are there competitive pressures to drive firms towards or away from respecting consumer privacy? The second justification for and type of data mining regulation is voluntary regulation through default privacy rules which can be opted out of or contracted into by individual consumers' and/or businesses' choices. Should privacy defaults be chosen to be what a majority of the parties involved would hypothetically choose themselves or chosen in order to force better informed parties to reveal their private knowledge? The third justification for and type of data mining regulation is mandatory state regulation to improve market performance. Does the choice of a privacy default result in that practice becoming mandatory due to network externalities? Does making consumers aware of certain privacy issues raise consumers' consciousness and induce data mining businesses to alter their behavior? A final justification for and type of data mining regulation is also mandatory state regulation, but from a property-based perspective. Do the personhood concerns and commodification issues data mining raise lead to regulation in order to achieve incomplete commodification?

I. Economics of Data Mining Versus Consumer Privacy

Any consumer purchase on the Net will also incidentally create as a natural by-product of that sale consumer data. An example of such data is the quantity and price of the product, used respectively for inventory control and to make refunds if necessary. Such data existed with even off-the-Net cash transactions and before the spread of computerized record keeping. Data from credit card transactions include consumers' names and credit card numbers. The transaction of getting a credit card results in the collection and verification of data regarding addresses, spouse names, number of dependents, social security numbers, employment histories, financial records, and credit histories. Thus, while there has always been consumer data, the amount, type, and uses of such data have all increased with the advent of automatic recording of data by computers at the point of sale (POS). In addition, recent advances in computer technology have greatly lowered the costs of disseminating, sharing, and correlating consumer data once it has been collected. These cost decreases in manipulating consumer data have led to "data mining" for new potential customers. The packaging and statistical analysis of consumer data once would have been prohibitively costly to do, but is now cheap and commonplace. In response to technological progress, an entire industry has arisen to sell consumer data to marketers looking for a competitive advantage. Like any type of information,

consumer data has both costs and benefits. The costs of consumer data include the loss of privacy and all of its attendant problems, such as the potential for either criminal misuse or state abuse. The benefits of consumer data include the matching of consumers with businesses which can satisfy their preferences and better safety or security for consumers in verifying identification or detecting criminals.

A. Allocational Efficiency Versus Consumer Privacy

While using data mining to better match up consumers with firms which can cater to their tastes increases allocative efficiency on the Net; it raises at least four consumer privacy loss issues. First, there is the question of who has access to data mines. Are there adequate safeguards to protect data mines from unintended eyes? Second, there is a standard problem with any type of information, namely its ease of resale. Will data from different sources be complied together into data mines of consumer profiles and sold to vendors without knowledge or consent of consumers? Third, will data mines be kept accurately in spite of the inevitable potential for human error? Finally, how often are data mines to be updated to avoid inaccuracy from obsolescence? All of these consumer privacy concerns are heightened by the fact that data mines are primarily used to better target direct mail marketing or home phone solicitation, both of which can possibly be intrusions into an individual's home or "castle". Similar uses of data mining to market at the

POS on or off-the-Net, for example, would probably not be viewed as intrusive of consumer privacy.[5]

Against such consumer privacy concerns must be weighed the allocative efficiencies of businesses only contacting those consumers who have some interest in their product. Neither firms nor consumers want e-mail to be deleted or mail order catalogues or brochures to be thrown into the trash or phone solicitation to become bothersome burdens. People differ in their preferences towards receiving unsolicited e-mail or mail catalogues or cold calls. Some may feel honored to be among the selected many to receive such attention, while others might feel annoyed by any uninvited form of contact. While knowledge concerning these sorts of preferences would be useful in allocative efficiency, this is not usually the sort of information contained in data mines and certainly not a cause for consumer privacy concerns.

Although firms have no desire to waste resources on e-mail, snail mail, or phone calls on those who would not be interested to become consumers, let alone possibly antagonize potential customers who might otherwise have become actual buyers; businesses know from experience there will always be a positive yield from e-mail lists, mass mailing, or cold calling. Some consumers' preferences will be swayed by well-designed e-mails, colorful mailers or persuasive callers. Not only do people experience weakness of human will and impulse buying, but also there are people who might genuinely learn about products or services they did not know about but come to realize they can not live without. The net dollar benefit of unwanted contact is

ultimately an empirical question, the answer to which varies from case to case depending on the particulars of the product, social norms, and the nature and manner of the unsolicited contact.

B. Consumer Privacy Versus Consumer Security

Another trade-off exists between consumer privacy and consumer security. This trade-off exists because certain types of detailed personal consumer data can be used to verify the identity of the consumer. While more specific consumer data protects consumers from unauthorized use, it also means that consumers have less privacy. In particular, two groups besides private firms may obtain access to data mines, namely criminals and law enforcement officials. While the latter group's desire for consumer data is motivated by catching the first group, both groups desire to get access to data mines not collected by them. An invasion of consumer privacy occurs whether there is criminal theft from or law enforcement access to data mines. The view that privacy is not an issue when an individual has nothing to hide equates privacy with merely the hiding secrets.[6] While privacy has many possible interpretations,[7] at its heart it simply involves the idea of "the right to be left alone."[8]

Another way in which consumer privacy and consumer security conflict is illustrated by the recent proposal that airlines collect more detailed information about passengers when they book flights in order to detect and hopefully deter potential terrorists. The baggage x-ray process in U.S. airports is an

inconvenience which most of us feel is a necessary price to pay for heightened security and airline safety. Obviously, greater deterrence can be obtained by mandatory hand searches of all bags, but such an intrusion of individual privacy probably would not be tolerated by most people nor would it even be necessarily cost-effective in terms of the required labor and time delays. Any criminal detection system will experience type 1 errors (false acquittals) and type 2 errors (false convictions). What levels of these legal enforcement errors we as a society will tolerate is technology dependent and preference driven. Both the state of technology and individual as well as social preferences over the costs and benefits of enforcement will undoubtedly change over time, often in unexpected or even unforeseeable ways.

II. Self-Regulation on the Net: Private Norms of Consumer Privacy

The first justification and type of regulation about data mining on the Net involves firms on the Net regulating themselves. Private industry often adopts its own standards of behavior in part to forestall impeding government regulation. A recent example is the new rating system for television programs. Other examples of self-regulation are the professional codes of ethics accountants, doctors, and lawyers have respectively adopted. The Net may adopt norms of privacy regarding data mining if and when mandatory regulation by the FTC appears likely. But, even before then, are there profit-maximizing forces towards self-regulation or will market competition result in a race to the bottom?

Businesses engaged in data mining could find it in their own monetary self-interest to respect consumer privacy if consumers are willing and able to pay more for privacy than it costs to provide. Viewing consumer privacy merely as an obstacle to growth of the Net is myopic and wrong. Purely as a matter of public relations, informing consumers of the safeguards businesses take in protecting the consumer privacy of data mines will result in consumer goodwill. Empirically, the questions are how much will consumers pay for their privacy and what is the market value of consumer goodwill from respecting privacy? Obtaining data from consumers who have provided informed consent is easier than obtaining data from consumers who are suspicious or antagonistic towards the data mining industry because of a

lack of well-established consumer privacy business norms. But, industry norms by their very definition involve the potential for a free-rider or collective action problem. Counteracting this incentive is the competitive loss to firms which do not respect consumer privacy.

A. A Psychological Game-Theoretic Model of Industry Norms

Psychological game theory has been applied to a number of legal issues.[9] For example, psychological game-theoretic models of monopoly pricing and personnel decisions explain why firms do not always charge monopoly prices nor fire workers in the manner predicted by neoclassical labor economics.[10] Managers are often motivated by fairness concerns and behave less "selfishly" than myopic thin self-interest predicts. Psychological game-theoretic models explain how endogenously created emotions can increase the frequency of suits going to trial instead of settling.[11] Especially in divorce or child custody proceedings, parties may be motivated by anger over what they believe the other party should not have done.[12] Psychological versions of a one-sided prisoner's dilemma explain how to control bureaucratic corruption and tax evasion.[13] In principal-agent relationships, agents may be motivated by guilt or shame from breaching trust if they believe others are honoring trust. Taxpayers may be similarly motivated to pay their taxes if they feel shame at not doing so when they believe others are paying their fair share. Psychological games have even been applied to legal education to

suggest that first year law students might become less cooperative in playing a two-person prisoner's dilemma than before they entered law school because they believe others will play less cooperatively than they believed before they entered law school.[14] Finally, psychological games have been applied to analytically model the concerns over commodification and the possible domino effects of monetary commensurability.[15]

A psychological game is defined as a game where at least one player's utility depends not only on the strategy choices of players, and hence indirectly on beliefs about strategic choices through their influence on such choices, but also *directly* on the beliefs of a player about another player's choices and, possibly, beliefs about such beliefs about choices, and so forth. Thus, psychological game theory overcomes Ellickson's[16] criticisms of too narrow a conception of human behavior in law and economics. Consider this psychological game-theoretic model of privacy norms in the data mining industry. Suppose there is a private cost of $K for any single data mining firm to use technology ensuring consumer privacy. The whole data mining industry is modeled as the unit interval: [0, 1].[17] Assume that all data mining industry members face a binary decision: to use a particular consumer privacy technology or not. For simplicity, assume that all data mining firms have identical preferences regarding consumer privacy technology with the status quo of not using the consumer privacy technology in question having a normalized profit of zero. Assume that any single data mining firm's profits from adopting the consumer privacy technology depends on

the number of other firms who will adopt the privacy technology, and not on the identity of those firms who adopt the privacy technology.

Let P be the proportion of the data mining industry choosing to adopt the consumer privacy technology. Let R be the proportion of the data mining industry that any single firm expects will adopt the consumer privacy technology. Suppose that the net profits, Π, to any firm of adopting the consumer privacy technology are described by the equation: $\Pi(P) = B - K$, where the benefits, B, can be decomposed into an absolute component, A, and a relative component, C(P). Assume that $C(P) = CP$, for some constant with $C > 1$. Because P is unknown at the time that any single firm has to decide whether to adopt the privacy technology, assume that all data mining firms are risk-neutral and maximize expected profits, $E\Pi(R) = A + CR - K$. This functional form means that the more other data mining firms are expected to adopt the consumer privacy technology, the more an single data mining firm benefits from adopting it. This captures the idea that not adopting the consumer privacy technology in question is more costly the more other firms adopt it. Assume that $A + C > K$; this means that if all other data mining firms adopt the consumer privacy technology, any particular data mining firm also has to adopt it just to "keep up" with the other firms. Furthermore, assume that $A < K$, for otherwise the consumer privacy technology has such large absolute benefits that even when no other data mining firm adopts it, any single firm would choose to adopt it; that is, $\Pi(0) < 0$. A data mining firm

compares its profits $\Pi = A + CR - K$ from adopting the consumer privacy technology when it expects the proportion R other firms to adopt the consumer privacy technology with 0, the normalized value of net profits from not adopting the consumer privacy technology.

B. Endogenously Determined Equilibrium Norms on the Net

Part of the definition of a psychological game-theoretic equilibrium is that players' beliefs are correct in the sense of being self-fulfilling. This means that in a psychological equilibrium the expected proportion of data mining firms choosing to use the consumer privacy technology must equal the proportion that actually chooses to use it, that is $R = P$. The second part of the definition of a psychological equilibrium is that players' strategies are best responses to each other given their beliefs.[18] There are several psychological equilibria. In the first equilibrium, no data mining firm uses the consumer privacy technology, because with $R = 0$, $A - K < 0$ and $P = 0$. In the second equilibrium, all data mining firms adopt the consumer privacy technology, because with $R = 1$, $A + C - K > 0$ and $P = 1$.

If $(K - A) < C$; there is a third equilibrium, in which an intermediate proportion P^* of the data mining industry adopts the consumer privacy technology in question, while the complementary proportion $(1 - P^*)$ does not adopt this consumer privacy technology. The critical value of P^* is solved for by setting $\Pi = 0$,[19] or $A + CR - K = 0$. Thus, $P^* = (K - A)/C$. Only at that

value of P will a data mining firm be indifferent between adopting the consumer privacy technology and not. This last equilibrium might be unstable in the sense that if P is perturbed away from P*, it may either go to zero or one, instead of returning to P* (this depends on the industry dynamics of adjustment). Of the two remaining equilibria, the one where no data mining firm adopts the consumer privacy technology is the status quo. The one where all data mining firms adopt the consumer privacy technology was assumed to be individually profitable for all data mining firms (by making the assumption that $A + C > K$).

The above model suggests how self-regulation can allow the data mining industry to overcome the potential free-rider or collective action problem by selecting a set of beliefs about what other data mining firms will do as focal. For example, the data mining industry could announce that it recommends adoption of the consumer privacy technology in question. If the recommendation is believed to be genuine, then it will become self-enforcing. This phenomenon is analogous to how the F.D.I.C.'s insurance of individual bank deposits (for amounts less than $100,000) can itself prevent any panic-induced bank run. Self-regulation thus offers the data mining industry a way to select as focal a particular equilibrium set of beliefs among several possible ones.

III. Voluntary Regulation on the Net: Consumer Privacy Defaults

The second justification and type of data mining regulation is voluntary regulation by consumers or data mining industry members opting into default rules about consumer privacy. The language of default rules itself suggests that such privacy defaults about how to deal with data mining can be opted into or out of by data mining market participants, be they consumers or industry members. The use of the word "default" is intended to stress the voluntary nature of such rules as opposed to mandatory or immutable rules which can not be opted out of by individual choice or contractual agreement. There is a vast literature in contract law about default rules concerning such issues as excusable reasons for contractual nonperformance (for example, fraud or duress) or appropriate remedies for contractual breach (for example, monetary damages or specific performance).[20] A focus of this literature is on how legal defaults are selected. The two most commonly proposed type of default rules are the so-called majoritarian defaults[21] and penalty defaults.[22] Each of these will be considered in the context of consumer privacy defaults about data mining.

A. Majoritarian or Hypothetical Consumer Privacy Defaults

Majoritarian contractual default rules mimic what a majority of contracting parties would have chosen to agree upon had they bargained explicitly over the relevant issues. The rationale for majoritarian defaults can be a utilitarian one because they minimize transactions costs for most parties which do not have to

negotiate the issues provided for by those defaults. But, another rationale for majoritarian defaults is a nonutilitarian one, based on the hypothetical consent of contracting parties because the defaults reflect what most parties would have consented to had they actually bargained over the relevant issues.[23] This rationale for majoritarian defaults is often called a contractarian one because moral philosophical arguments which are based on what people would have chosen to agree upon in some idealized bargain are often labeled contractarian. Of course, the distinction between these alternative rationales for majoritarian defaults is moot if contracting parties always choose to maximize the joint expected value of their contracts. Whatever the rationale for majoritarian defaults, they clearly have intuitive appeal. Does this intuition carry over from general contractual settings to the specific area of consumer privacy over data mining?

Privacy defaults over data mining are not likely to be negotiated over between individual consumers and businesses. Instead such defaults are likely to be buried in the boilerplate of so-called contracts of adhesion which consumers must accept on a take it or leave it basis. Thus, the hypothetical bargain rationale probably does not apply because a bargaining environment in which one side makes take it or leave it offers to the other side is a highly specialized one and one which most consumers probably would argue is procedurally, if not substantively biased in favor of businesses. Of course, the typical justification for contracts of adhesion is a transactions

cost based one: it is prohibitively costly for a firm to bargain with individual customers. In addition, the argument goes on, firms will not exploit consumers because of competitive pressures. But, that argument presumes the market in question is competitive and that contracts of adhesion have not already become industry custom or a de facto standard. Both of these presumptions may fail to hold for data mining and even if they do hold, they may not continue to hold. The other basis for majoritarian defaults also does not apply to a data mining firm's unilaterally chosen consumer privacy defaults because they would suffer from the same problem that what a majority of consumers would have bargained for is almost surely not what a data mining firm would set as the default. If instead of having defaults chosen by a data mining firm, they are set by law as is the case with contracts; the rationale of a majoritarian or hypothetical default is more plausible. Nonetheless, there are reasons to worry about such defaults.[24]

B. Informational Forcing Penalty Consumer Privacy Defaults

Penalty defaults are designed to avoid inefficiencies which result when a contracting party with hidden information conceals that knowledge for private strategic benefit in bargaining. The rationale for penalty defaults is to select rules which result in outcomes particularly unfavorable to a party with private information. Such defaults, in turn, force the better informed party to reveal its private knowledge to both the other

contracting party and third parties, such as courts to avoid the penalty outcomes provided by the default. Penalty defaults make sense if there are informational asymmetries between contracting parties themselves and with third parties in the legal system. Do penalty defaults make sense as privacy defaults about data mining?

While consumers obviously know more about themselves than data mining businesses (hopefully), choosing defaults to force consumers to reveal such information to data mining firms or the courts does not seem to be really necessary. After all, the reason for consumers to not voluntarily reveal this information under a penalty default scenario is to gain a strategic advantage in bargaining. But, as is clear from the discussion in the previous section, consumers are unlikely to be negotiating individually with data mining firms. While consumers will clearly differ in and know more about consumers' attitudes towards privacy than data mining firms or the courts; consumers also have no incentive to hide this sort of data. Indeed, those consumers who desire privacy have ample reasons to make their preferences over their privacy known to data mining firms to avoid unwanted solicitation. How consumers would effectively communicate their privacy preferences is also far from clear given the absence of any real individual negotiation over consumer privacy.

On the other hand, data mining firms usually do so without consumer knowledge or informed consent. These two informational asymmetries (general unawareness of data mining practices and

specific unawareness of what is being done by a particular data mining firm) form the rationale for Peter P. Swire's proposal that cyberbanking privacy laws be designed as penalty default rules.[25] Similar defaults in the data mining context would be skewed in favor of consumers who, even if they know of the depth and breadth of data mining, face high transactions costs in trying to keep track of how their specific profiles are being (ab)used by individual data mining firms. Penalty defaults would be designed to force data mining firms to reveal their private knowledge about data mining in general and their specific data mining practices in particular by way of corporate consumer privacy policy statements or warnings to consumers as a class. Penalty defaults may also force data mining firms to reveal their private knowledge about data mining practices specific to an individual consumer in the form of providing notice or access to informational trails or logs. Of course, such penalty defaults might also effectively become mandatory for a variety of reasons.[26] But, as opposed to hypothetical or majoritarian defaults, which may unintentionally penalize "discrete and insular" minorities; penalty defaults are intentionally designed to penalize the better informed party (which also is the party with more bargaining power, despite most law and economics approaches to contract law being silent about bargaining strength inequities between contracting parties). The overall consumer ignorance regarding the data mining industry also suggests a different justification for mandatory data mining regulation.[27]

IV. Mandatory State Regulation on the Net: Improving Performance

The third justification and type of data mining regulation is mandatory state regulation in order to improve the performance of markets. The mandatory part of mandatory state regulation implies that voluntary regulation is problematic. The state part of mandatory state regulation is traditionally justified in the framework of neoclassical economics as the way to correct market failures[28]. A prototypical market failure is the phenomenon of externalities. An externality occurs if the market system is not complete in the sense that some activity is generating costs (or benefits) which are not being priced. The textbook example of a (negative) externality is pollution of some kind where the polluter does not bear or pay the full social cost of polluting. An externality receiving much recent attention in economics and to a lesser extent in the law is a network externality.[29] Network externalities occur when an activity's benefits increase with the number of expected others who will also be engaged in the same activity. An example of a network externality is the adoption of a particular computer operating system or protocol. The next section explains how default rules can become effectively mandatory because of network externalities.

A. The Unavoidable Transformative Nature of Default Rules

The literature on contract default rules considers the possibility that defaults can have a transformative effect on the

preferences of contracting parties.[30] One possible way in which this might unavoidably happen is via network externalities in the preferences of individuals. In the context of the model of data mining consumer privacy norms presented in section II.A. of this paper, a legal default rule chosen by the state might have the same effect as a focal point selected by the data mining industry. Additionally, it may overcome the credibility issue of whether data mining firms believe that other data mining firms will really adopt consumer privacy technology. Collective movement between multiple equilibria can be accomplished by mandatory state regulation, which serves to change social norms.[31]

Besides the likely fact data mining firms' preferences will exhibit network externalities, there are other reasons to believe so-called voluntary default rules will become in reality mandatory rules. These are based on cognitive psychology experiments which find that people exhibit status quo biases. In the context of data mining, this means that whatever is the default consumer privacy rule might be hard to move away from due to a human and even greater corporate tendency for staying put.[32] Risk aversion by managers of data mining firms provides yet another incentive to "not rock the boat" and follow defaults.

B. Raising Consumer Consciousness and Altering Firm Behavior

An entirely different justification for mandatory state regulation to improve market performance is based on raising the

consciousness of consumers and in so doing also changing the behavior of firms.[33] The central idea of this alternative justification for regulation is that the mere act of regulation itself will bring to consumers' attention and make salient data mining practices which they would not otherwise have even been aware of. The regulation itself may also induce some data mining firms to alter their behavior so as to avoid being covered by the regulation in question. This alternative justification for regulation explicitly acknowledges that most consumers today are not aware of data mining, let alone its rapid growth. At least initially, regulation of the data mining industry would have such a consciousness raising effect on previously unaware consumers. Under a consciousness-raising rationale, regulations about data mining would still be evaluated by whether they led to the data mining industry making cost-justified increases in consumer privacy.

V. Mandatory State Regulation on the Net: Property Rationales

The final justification for and type of regulation is also mandatory state regulation, but based on property rationales. The underlying idea of this perspective on regulation is that data mines form a type of personal or intellectual property. Being information, however, data mines exhibit all of the usually problematic features as property that any type of information does, namely difficulty in valuation and ease of resale due to a public good quality. The value of data mines are not so much

unknowable as it is low on a per individual and per fact basis. Much of the value of data mines occurs through aggregation into massive databases and is technology driven. The public good nature of data mines simply means that once collected or packaged, data mines are very cheap to sell again and again. Thus, because data mines have a low marginal cost of redistribution, there is a natural temptation to resell data mines once they have been acquired. Thus, an active resale market in data mines is almost inevitable. The next two sections argue for mandatory state regulation to counter such tendencies.

A. Personhood Concerns

The nature of data mines are such that if they are viewed as property, it raises all of the personhood concerns Professor Radin is troubled by in solely economic and individualistic understandings of property.[34] The information in data mines can be said to be almost a part of who consumers are in a constitutive sense, much like a home or wedding ring. In fact, the data mine analogues of those two types of personal property are home address and marital status. Not surprisingly, personal property is related to notions of privacy and liberty. Consumer profiles can be viewed as personal property because not only would they not exist without consumers, but also because consumer profiles often effectively “capture” consumers’ tastes as revealed by their market choices in the form of “digital snapshots”. As a form of personal property, consumer profiles

reduce the privacy and liberty of consumers by making their purchase histories and likely future tendencies commodities for sale to the highest bidder. But, while most consumers would regard certain (particularly “major”) purchases as private, some consumers may not (for example, the snob appeal of designer labels, expensive car models, or exclusive residential neighborhoods). In addition, some consumers may regard certain (particularly “minor”) purchases as not particularly worthy of privacy protection, such as grocery shopping records.

B. Commodification Issues

Data mining also raises all of the commodification issues Professor Radin is troubled by in the use of market rhetoric.[35] Radin suggests that the dominance of market rhetoric generally in legal, policy, and social discourse can have undesirable side-effects. In the particular context of data mines,[36] both literal and metaphorical commodification occurs in laws and rhetoric treating data mines as intellectual property. Mandatory state regulation of data mines may be necessary to be a countervailing force of friction against the so-called "slippery slope" of commodification. Mixing or switching metaphors, the "domino effect" Radin is worried about might be interrupted by the "circuit breaker" of incomplete commodification via mandatory state regulation. The words “may” and “might” in the preceding two sentences reflect what is ultimately an empirical question as to whether data mining must lead inevitably to yet further

commodification on and off-the-Net. After all, public schools do co-exist with private schools,[37]

The underlying idea of psychological game theory that people often feel emotions which depend on their beliefs about how they or others will choose to behave also resonates with Radin's personhood perspective to property.[38] Personhood theory involves a thicker rather than thinner theory of the self.[39] Those objects such as one's home or wedding ring, which are "personal" and noncommodified are so because of emotions based on beliefs about what one has and others do not have the right to do regarding those objects. Those same objects become "fungible" property in relation to others and commodified because they do not share the same beliefs and emotions about what they can do about those objects. Thus, a landlord or government official feels no emotional attachment to a tenant's residence and views that "home" as a commodity, being interchangeable with other "places to live" or even money.

The fact that universal noncommodification and universal commodification are the endpoints of a continuum along which the degree of commodification resides can be formally captured by the idea of probability beliefs that lie in the interval with endpoints 0 and 1 over what can or should be metaphorically conceived of as commodities. The view that the culture of commodification is socially constructed is captured by the idea that beliefs have to be in equilibrium and thus consistent with actual social behavior. The shared beliefs which human emotions depend on will vary with the particular context. The difficulty

and possibility of changing the prevailing legal/moral culture is reflected by the fact that individuals unilaterally are unable to move among but collectively can move among multiple equilibria. This paradox may explain the ambivalence many feel over contested commodities, such as body parts[40] or "human capital assets"[41] formed during marriage.

Universal commodification of an object translates into no emotional attachments to that object due to beliefs placing a probability of unity on trading money for that object. Universal noncommodification of an object translates into strong emotional attachments to that object due to beliefs placing a probability of zero on trading money for that object. Incomplete commodification of an object translates into conflicted emotional attachments to that object due to beliefs placing a probability strictly between zero and unity on trading money for that object. It is straightforward to construct a formal psychological game-theoretic model in which there will be these three equilibrium degrees of commodification.[42] These degrees of commodification are associated with a set of probability beliefs, which can be interpreted as endogenously determined corresponding social norms.

The above model suggests a natural role that regulation can have in selecting an intermediate focal point among the multiple possible equilibrium outcomes by selecting an intermediate set of probability beliefs about what others will do as being salient or focal. In so doing, regulation alters the degree of market discourse or culture and helps to construct preferences. An

alternative to regulation as expressing incomplete commodification is a public policy of an "educational" campaign aimed at changing people's beliefs about what other people will do and thus people's belief-dependent preferences.

Conclusions

In summary, data mining is a technological and economic reality which is here to stay. Data mining raises legitimate concerns regarding consumer informational privacy. If and what sort of regulation can alleviate these concerns are important social policy questions. This paper offered a taxonomy of possible types of and justifications for regulations over data mining on the Net. It did so by utilizing both traditional and behavioral law and economics to analyze such regulation. Ultimately, the answers to the questions of whether and if so, how to regulate data mining will not be uniform across different categories of data. Instead of such a “unified field theory” of data mining, the appropriate justifications for and types of regulations are more nuanced and varies according to the type of information contained in particular data mines. A single transaction on the Net can lead to many different categories of data mines. Thus, any particular data mine should be "unpacked" into its component pieces, each of which may require a separate justification for and type of regulation. This view of data mines is akin to the modern financial perspective that (under certain technical conditions[43]) any security is really a

portfolio formed over an underlying set of basic securities.[44] Just as that economic perspective forms the intellectual basis of financial engineering,[45] this viewpoint of data mines can be said to underlie "informational engineering".

A first cut for unbundling data mining is the dichotomy between sensitive and non-sensitive data mines. Examples of the former are such financial information as credit card numbers or bank account records or patients' mental health information (which some insurance companies insisted on placing into computer networks[46]). An example of the latter might[47] be compact disc purchase histories.

As with any type of classification scheme, this one suffers from a possible lack of consensus among consumers of how to treat particular data mines. For example, consumers differ in whether they view their names, addresses, or phone numbers as sensitive or non-sensitive data mines as evidenced by different consumers' views about anonymous remailers,[48] digital cash,[49] or Caller ID.[50] One solution is to just let individual consumers decide whether they want a particular data mine to receive sensitive or non-sensitive treatment. But, the problems with such a solution include the lack of truthful preference revelation (if sensitive data mines receive heightened legal or technical protection and possible monetary compensation); endogenous preferences (if consumers have not even thought about how they feel about certain kinds of data mines or are unaware of the data mining industry); and nonuniformity of legal and possibly economic treatment across otherwise identical consumers (which may raise issues of whether

such treatment is constitutional and whether price discrimination is socially desirable).

A second cut for unbundling data mining is the dichotomy between data obtained from adults versus that obtained from kids. Children who answer questions regarding their parents, living situations, and own preferences to be able to gain access to web pages, play new games on the Net, or join on-line clubs do so more often than not without realizing this information will become part of data mines. Such data mining practices could be seen as being unconscionable because minors lack the understanding which a contractual paradigm presupposes.

In conclusion, consumers' data mines have now become and will continue to be a valuable commodity. The questions are whether and if so, how to regulate the growing practice of data mining. This paper applied behavioral law and economics to introduce and consider distinct justifications for and types of data mining regulation. The next step is to get down to the "nitty gritty" of determining which distinct justifications for and types of regulations fit which types of data mines. An important part of that determination will involve the laws of privacy, contract, (intellectual) property, and antitrust as well as the behavioral economics of network externalities and information processing.

* Thanks to Peter P. Swire for helpful discussion and the Institute for Law and Economics at the University of Pennsylvania for financial support during July and August 1997.

[1] "Recombinant information" describes a composite profile of a consumer made up of individual "pieces" of information about that consumer. This clever term was introduced by Erik Larson, The Naked Consumer: How Our Private Lives Become Public Commodities, 12 (1992).

[2] The author served as a staff economist in the Division of Consumer Protection at the Federal Trade Commission, 1984-85.

[3] The Net, Columbia Pictures, 1995.

[4] Ann Cavoukian & Don Tapscott, Who Knows: Safeguarding Your Privacy In A Networked World (1997).

[5] For an excellent history of privacy as both a philosophical and legal concept, see Priscilla M. Regan, Legislating Privacy: Technology, Social Values, And Public Policy, 24-41, 69-108, 214-43 (1995).

[6] See Richard A. Posner, *The Right of Privacy*, 12 Ga L Rev 393 (1978).

[7] See Kim Lane Scheppele, Legal Secrets, 181-265 (1988).

[8] Samuel D. Warren & Louis D. Brandeis, *The Right To Privacy*, 4 Harv L Rev 193 (1890).

[9] For the precise technical definition of a psychological game, see John D. Geanakoplos et al., *Psychological Games and Sequential Rationality*, 1 Games & Econ Behav 60 (1989) and Van Kolpin, *Equilibrium Refinement in Psychological Games*, Games &

Econ Behav 218 (1992).

[10] Matthew Rabin, *Incorporating Fairness into Game Theory and Economics*, 83 Am Econ Rev 1281 (1993).

[11] Peter H. Huang & Ho-Mou Wu, *Emotional Responses in Litigation*, 12 Int'l Rev L & Econ 31 (1992).

[12] William Morrison & Glenn Feltham, “Getting Angry and Getting Even: Emotional Behavior and Civil Disputes,” July 1997 (unpublished manuscript).

[13] Peter H. Huang & Ho-Mou Wu, *More Order without More Law: A Theory of Social Norms and Organizational Cultures*, 10 J L Econ & Org 390 (1994).

[14] Peter H. Huang, Does Being a 1L Foster Distrust and Preemptive Dishonesty? June 1995 (unpublished manuscript).

[15] Peter H. Huang, Dangers of Monetary Commensurability: A Psychological Game Model of Contagion, 146 U Pa L Rev (forthcoming, 1988).

[16] Robert C. Ellickson, *A Critique of Economic and Sociological Theories of Social Control*, 16 J Legal Stud 67 (1987).

[17] This technical assumption just ensures that no single business is "large" relative to the data mining industry as a whole, so that no individual firm's choice by itself affects the proportion of the data mining industry making that choice.

[18] This is the standard Nash equilibrium requirement. *See, e.g.,* John Nash, *Equilibrium Points in n-Person Games*, 36 Proc Nat'l Acad Sci 48 (1950); *Non-Cooperative Games,* 54 Annals Math

286 (1951). In other words, a psychological game-theoretic equilibrium is a set of beliefs and strategies for players such that: (1) the beliefs are rational, that is, consistent with the strategies; and (2) the strategies themselves form a Nash equilibrium given the beliefs.

[19] Setting $\Pi = 0$ makes a firm indifferent between using the privacy technology, thus getting a payoff of Π and not using the privacy technology, thus getting a payoff of 0 (by assumption, status quo payoffs were normalized to be 0).

[20] See generally, Richard Craswell & Alan Schwartz, 16-30 Foundations Of Contract Law (1994).

[21] Charles E. Goetz & Robert E. Scott, *The Mitigation Principle: Toward a General Theory of Contractual Obligation*, 69 Va L Rev 967 (1983).

[22] Ian Ayres & Robert Gertner, *Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules*, 94 Yale L J 97 (1989); *Strategic Contractual Inefficiency and the Optimal Choice of Legal Rules*, 101 Yale L J 729 (1992).

[23] See, for example, Randy E. Barnett, *The Sound of Silence: Default Rules and Contractual Consent*, 78 Va L Rev 821 (1992).

[24] See discussion *infra* part IV.A.

[25] Peter P. Swire, Cyberbanking and Privacy: The Contracts Model, presented at the Seventh Conference on Computers, Freedom, and Privacy (March 12, 1997), *in* CFP'97: Commerce And Community, March 1997, at 148-52.

[26] See discussion *infra* part IV.A.

[27] See discussion *infra* part IV.B.

[28] W. Kip Viscusi, et al. 654-86 Economics Of Regulation And Antitrust (1992).

[29] *See, e.g.*, Michael Klausner, *Corporations, Corporate Law, and Networks of Contracts*, 81 Va L Rev 757 (1995).

[30] *See, e.g.*, Alan Schwartz, *The Default Rule Paradigm and the Limits of Contract Law*, 3 So Cal Interdiscip L J 389 (1993).

[31] Cass R. Sunstein, Social Norms and Social Roles, 32-69, in Cass R. Sunstein, Free Markets And Social Justice (1997).

[32] *See, e.g.*, Donald C. Langevoort, Beliefs, Biases and Organizational Behavior: The Epistemology of Corporate-Securities Lawyering (March 1997, unpublished manuscript).

[33] Michael Barsa, California's Proposition 65 and the Limits of Information Economics, 49 Stan L Rev 1223 (1997).

[34] Margaret Jane Radin, Reinterpreting Property (1993).

[35] Margaret Jane Radin, Contested Commodities: The Trouble With Trading In Sex, Babies, Body Parts And Other Things (1996).

[36] See also Margaret Jane Radin, *Property Evolving in Cyberspace*, 15 J L & Com 509 (1996).

[37] Mark Kelman, *Health Care Rights: Distinct Claims, Distinct Justifications*, 3 Stan L & Policy Rev 90, 99 (1990).

[38] Radin, *supra* note 34

[39] *Id.* at 25-27.

[40] *Id.* at 15.

[41] *Id.* at 33.

[42] As in the model in section II.A, individuals' utility functions could involve emotions which depend, for simplicity. linearly on probability beliefs.

[43] See Kenneth J. Arrow, *The Role of Securities in the Optimal Allocation of Risk-Bearing*, 31 Rev Econ Stud 91 (1964).

[44] See William F. Sharpe, *Nuclear Financial Economics, in* Risk Management: Problems & Solutions 17 (William H. Beaver & George Parker, eds., 1995)

[45] *Id*. at 18.

[46] Maggie Scarf, *Keeping Secrets*, NY Times, June 16, 1996, § 6 (Magazine), at 38.

[47] I say might because some individuals believe privacy is a basic human right (the European model) and view the collection and use of consumers' purchase histories as invasions of their privacy unless prior notice and informed consent occur.

[48] See A. Michael Froomkin, *Flood Control on the Information Ocean: Living with Anonymity, Digital Cash, and Distributed Databases*, 15 J L & Com 395, 402-49.

[49] *Id*. at 450-79.

[50] See Steven P. Oates, *Caller ID: Privacy Protector or Privacy Invader?* 1992 U Ill L Rev 219.