A Network-Economic Policy Study of Identity Management Systems and Implications for Security and Privacy Policy

Repkine, Alexandre and Hwang, Junseog

Seoul National University

August 2004

A Network-Economic Policy Study of Identity Management Systems and Implications for Security and Privacy Policy

Alexandre Repkine
Seoul National University, Techno-Economics and Policy Program, Seoul, Republic of Korea
repkine@snu.ac.kr

Junseok Hwang
Seoul National University, Techno-Economics and Policy Program, Seoul, Republic of Korea
Syracuse University, School of Information Studies, NY, USA

Abstract

Solving the problems associated with identity management in the "virtual" world is proving to be one of the keys to full realization of the economic and social benefits of networked information systems. By definition, the virtual world lacks the rich combination of sensory and contextual cues that permit organizations and individual humans interacting in the physical world to reliably identify people and authorize them to engage in certain transactions or access specific resources. Being able to determine who an online user is and what they are authorized to do thus requires an identity management infrastructure. Some of the most vexing problems associated with the Internet (the deluge of spam, the need to regulate access to certain kinds of content, securing networks from intrusion and disruption, problems of inter-jurisdictional law enforcement related to online activities, impediments to the sharing of distributed computing resources) are fundamentally the problems of identity management. And yet, efforts by organizations and governments to solve those problems by producing and consuming identity systems may create serious risks to freedom and privacy. Thus the implementation and maintenance of identity management systems raises important public policy issues.

The identity management systems (the IMS-s) often tend to require more information from the consumers than would otherwise be necessary for the authentication purposes.

The typical choice being analyzed in IMS is the one between a completely centralized or integrated system (one ID - one password, and a single sign-on) and the one comprising a plethora of (highly) specialized IMS-s (multiple ID-s and passwords). While the centralized system is the most convenient one, it is also likely to require too much personal information about the users, which may infringe on their rights to privacy and which definitely will result in serious damage should this personal information be stolen and/or abused. When more than two IMS-s interconnect (more of a practical side with various types of commercial values), they share the private information with each other, thus increasing consumers' exposure to possible information misuse. It is thus rather obvious that the public policy plays an important role to maintain the structure of identity management systems ensuring the existence of a sound balance between the authentication requirements and consumers' rights to privacy. The focus of this paper is on investigating this type of tradeoff by employing a theoretical framework with agents whose utility depends on the amount of private information revealed, and on making policy recommendations related to the issue of interconnection between alternative IMS-s. Our model derives optimal process of interconnection between IMS-s in the simple case of three IMS-s, then generalizing it to the case of more than three firms. The socially optimal outcome of the interconnection process in our model implies encouraging the interconnection between smaller rather than larger IMS-s.

JEL-Codes: L14, D85, D78, L25, L43, L51
Keywords: Networks, Interconnection, Identity Management, Regulation Policy

I. Introduction

Diverse as they are in terms of both scale and scope, the provision of almost all online services involves soliciting personal information from the customers. In fact, as the information technology progresses, more and more personal information is being collected from the consumers (Bennett, 1992.) Thus, the insurance companies want to know one's age, marital status and smoking habits, while credit card companies are concerned whether one has ever defaulted on his or her debt. The mortgage company is interested in the size of one's annual income and the online music store is anxious to know about your music tastes. Becoming a client of this or that service provider thus typically involves disclosing some sort of personal (private) information, which naturally raises concerns for privacy. Consumers are often unaware of the reuse and disclosure of personal information they provide to others during daily transactions.

Naturally, when an organization such as a business or the Government gets hold of the citizens' personal information, it gives them some sort of power of control over these individuals' lives. In particular, by learning more information about an individual may result in the information-soliciting party to extract more of the individual's consumer surplus. Indeed, whereas in the more classical treatment the price-discriminating firms or agencies had to design self-selecting contracts in order to charge each type their most appropriate price, the same sort of discrimination can be carried out by simply making an individual fill out a questionnaire, either in paper or online. Obviously, the increasing presence of online services makes the process of collecting such personal information much easier.

This paper presents an attempt at economic analysis of the privacy issues. The concept of privacy was introduced in the literature almost four decades ago in Westin (1967): "Privacy is the ability of individuals to exercise control over the disclosure and subsequent uses of their personal information." Gavison's (1980) definition is more straightforward and consists of three elements: secrecy ('the extent to which we are known to others'), solitude ('the extent to which others have physical access to us') and anonymity ('the extent to which we are the subject of others' attention').The importance of privacy protection has received

special recognition with the issue of the OECD's guidelines on privacy published in the year of 1980 (OECD, 1980). In particular, these guidelines identified the unlawful or inaccurate storage of personal data, abuse or unauthorized disclosure of personal information as violation of fundamental human rights. The report concluded that there should be some constraints as to the personal information to be collected by (government) organizations and firms alike, such as the solicited personal data should be relevant to the purposes which they are being collected for. At the same time the report acknowledged the existence of a tradeoff between the extent of privacy and the minimal amount of control over the information-providing citizens. That such control is necessary becomes clear once one thinks about the issues of identity theft, secure access to online banking and the like.

It is becoming a common business practice to provide services together with a note on privacy policy that typically guarantees that the information solicited from the customer either explicitly (through e.g. the questionnaires) or implicitly (by keeping record of customers' purchases in individual profiles) will not be sold to the third parties or otherwise abused. Moreover, as an alternative to the centralized depositories of private information such as Microsoft Passport online industry-backed organizations like Liberty Alliance (backed by the Sun Microsystems and Intel, among others) have developed a federated authentication procedure such that the businesses affiliated with the Alliance only get access to the information on an individual necessary for each particular transaction rather than the whole file of information on a person.

Recent surveys have found that four out of five Net users are concerned about threats to their privacy when they are online (Hansen and Berlich, 2003). Yet only 6% of them have actually experienced privacy abuses. If electronic commerce is going to thrive, this fear is going to have to be dealt with by laws and by industry practices. This paper argues that such laws are indeed necessary since our results suggest counting on the industry to regulate itself is not necessarily conducive to the socially optimal outcome

In this study we present a model that formalizes the concepts that are being widely used in a fairly loose way, such as 'personal information', 'identity management' and 'privacy'. We postulate that the amount of services

consumers obtain is directly proportional to the amount of personal information they reveal to the third parties. Consumers' privacy concerns are modeled as that amount of personal information they disclose beyond which their anxiety related to possible information misuse outweighs the benefits from enjoying more services whose provision is made possible due to more disclosure of personal information. The existence of such privacy threshold imposes certain constraints on the behavior of both policy makers and service providers.

This paper is organized as follows. Section II presents a discourse on the concept of identity and identity management systems. Section III provides an overview of the literature on the subject. Section IV proceeds with setting up a formal framework for analyzing the issues of identity management. Section V derives and analyzes the outcomes predicted by the model developed in the previous section. Section VI summarizes the paper and offers several policy implications.

II. Identity and Identity Management Systems[1]

To our knowledge, there has been no commonly accepted definition of identity so far. The way the concept is being viewed and perceived differs widely depending on the context and the academic area. In very general terms, the concept of identity can be said to emphasize the difference between an individual and a person (Mead, 1934). The way any person is interacting with the society is in many instances the way his or her identity interacts with the social system. In other words, the concept of a person may consist of a plethora of individuums each one carrying a particular identity, or equivalently, each person may possess multiple identities (Lehnhardt, 1995).

In this study we are interested in developing a framework for analyzing the process of interaction of a person (and hence, his identities) with the social system(s) in a well-defined, formal way. For that reason, out of all the multitude of the definitions of identity we limit ourselves to the usage of *digital identity* for online identity. The term *digital identity* refers to the process of the attribution of properties to a person, which can be technically formalized, listed and put into a readily accessible digital form, hence the name of the concept. One's digital identity can be a single E-mail address or a list of answers to the questions like "What is your age?" or "Are you married?". Clarke (1999) elaborates on the concept of digital identity in the following way: *"Digital identity is the means whereby data is associated with a digital persona. Organizations which pursue relationships with individuals can generally establish an identifier for use on its master file and on transactions with or relating to the individual. […] There are three approaches whereby a digital identity can be constructed from multiple sources: a common identifier, multiple identifiers, correlated; and multi-attributive matching."*

Up until recent developments in the area of automatic management of personal data such as electronic banking, on-line filling out of tax forms, e-commerce and loan applications, to name just a few, the most broadly accepted definition of legal person is "a human being to which the legal system refers rights, privileges and obligations" (Kelsen, 1966). However, as it is becoming

[1] Discussion in this section largely relies on the study by the Independent Centre for Privacy Protection, 2003.

increasingly easier to get access to and manipulate substantial amounts of personal data, and hence abuse them, the individual identity, including the digital one, started to receive protection from the main legal sources, such as constitutions, international treaties (e.g. treaties of the European Union and its directives), national laws and other international regulations.

Among the many aspects of regulating the use of one's personal data, the issue of giving the personal data owner (i.e. a person) the most control possible on its own identity and personal data has been receiving increasingly more attention Thus, the European Directive 95/46/CE about data protection postulates the following principles for providing each person with these rights:

- Personal data must always be processed fairly and lawfully
- Personal data must be collected for explicit and legitimate purposes and used accordingly
- Personal data must be relevant and not excessive in relation to the purpose for which they are processed
- Data that identify individuals must not be kept longer than necessary
- Appropriate technical and organizational measures should be taken against unauthorized or unlawful processing of personal data

In our everyday lives we face various environments that require us to present our identities, or in other words, reveal our personal data at least to some extent. These environments can be civic administration, supermarkets, schools, offices and shopping malls. Even if the personal information we reveal in those environments does not necessarily uniquely identifies or reveals everything about us, frequently this information is capable of giving an indication of who we really are, giving potential scope to discovering more information about us compared to what we actually have revealed. For example, it is prohibited by law in the U.S. to ask for one's marital status in a credit card application. However, this same application requires other information such as name and physical address, if combined with the easily accessible public records such as Lexis-Nexis, will provide information on one's marital status just as easily. This single particular example illustrates the more general principle (ICPP 2003): "*In general it is not possible to successfully manage one's partial identities without knowing when and where they may be involuntarily disclosed. This is not only the case with data trails in digital networks, but also capturing biometrics, e.g.*

by video surveillance, is often possible without knowledge and consent of the individual."

As mentioned before, our personal data can be used in various context and fashions by various business and public entities such as shops and tax offices. In order for these entities to manage the data on our personal identities, especially the digital ones, they need to employ *identity management systems*. As is the case with the concept of identity, there is a variety of the definitions of identity management systems (the IMS). In this study we understand the IMS to be an infrastructure within one or between several organizations, which have agreed upon a mutual model of trust in managing and using identities. This definition also includes an implementation of identity management encompassing a whole society. (ICPP 2003).

We concentrate on the relationship between a person (possibly with multiple identities) and an organization that employs identity management systems (such as a shopping mall or a tax office). In general, one can represent such a relationship as a digital transaction between a user and an organization, e.g. an e-commerce or an e-government service provider, offering its digital services. In this type of digital transactions, the issue of privacy protection and anonymity emerge to be very important. Whenever sensitive data such as credit card numbers or medical records are to be transmitted through the Internet, users often balk away from submitting their data electronically for fear of these data being stolen or misused, the latter including using these data in the way not intended by the users. Most people would therefore like to individually control what data will be transmitted to whom and for what purpose. However, since providing the users with such freedom often lies outside the scope of incentives of the services providers, we believe it is the scope of the government policy to constrain the freedom of the information-collecting agencies so as to comply with the basic guidelines on individual rights and freedoms.

III. A Brief Overview of the Literature on Digital Identity Management

Designing policies aimed at regulating the type and/or the amount of personal information solicited from individuals is commonly referred to as identity management (Clarke, 2004). Although identity management has many aspects to it (e.g. prevention of identity theft or development of authentication requirements), we concentrate on the issue of multiple identities, or alternatively, the issue of the optimal choice of the structure of the identity management systems. Our choice is motivated by two factors. First, most existing literature on identity management concentrates on the supply-side of the phenomenon, such as security of access, authentication algorithms etc. Second, despite of the postulated necessity (OECD, 1980) for the consumers to be able to control the process of collecting and using their personal data, consumers are as yet not able to exercise sufficient control over the information they disclose. One reason behind the latter might be that the only lobbying party with respect to identity management appears to be the representatives of identity management systems themselves who are naturally interested in soliciting as much information from the consumers as possible. In particular, the issue of multiple identities has been receiving a considerable amount of attention in recent years. The focal point of discussion is the choice between an all-encompassing single identity management system that knows everything about everyone on one side of the IMS spectrum and a multitude of highly specialized small IMS-s that perform narrowly defined operations and solicit minimum information from each individual at a time. Such keen attention to the issue appears to be primarily caused by the fact that individuals as well as businesses get increasingly more concerned about the way their personal information is being solicited and used. Thus, the Microsoft Passport IMS fell far short of the expectations of its creators because consumers did not like the idea of a lot of personal information about themselves collected from different places be stored in one place on the one hand, and on the other hand, businesses who foresaw such unwillingness to take place were unwilling to pay thousands of dollars a year for access to the Passport services. Another salient example is the recent debates on the introduction of a single identifier in the countries of European Union and Australia. Especially in the latter case, recognizing all of the perils associated with multiple identities abuse, the Australian government has forsaken the idea of a uniform omnipotent identity card for its citizens.

An example of how unregulated interconnection may infringe on a individual's privacy is given by Davis (2000). While the U.S. law forbids soliciting information on race and marital status by the credit issuing agencies, such information can be fairly easily obtained by the credit agency since this is contained public records (e.g. Lexis-Nexis) that require the very basic information such as name and address for access. It is rather obvious that national governments have an important role to play in maintaining the structure of identity management systems ensuring the existence of a sound balance between the authentication requirements and consumers' rights to privacy (JHU, 2003). While it appears natural that more safety in transactions involving personal identification requires more information on consumers' identity (see e.g. Ogata et al., 2004) consumers will be also likely to be more reluctant to reveal their personal data as they are required for more such data to be revealed (Olivero and Lund, 2004). The focus of this paper is on investigating this type of tradeoff by employing a theoretical framework with agents whose utility depends on the amount of private information revealed, and on making policy recommendations related to the issue of interconnection between alternative IMS-s.

Clarke (2004) notes that "Many scheme designers fail to demonstrate any appreciation of the need that individuals have to sustain many identitites, and to avoid linkage among them." The existence of multiple identities is often frowned upon as an inconvenience to individuals (hence the introduction of Microsoft Passport for example), with the role of multiple identities as a means of protecting people's privacy often overlooked. PINGID: "The issue is how to manage the linkage or sharing multiple identities." Liberty Alliance: "to gain access to portions of the user's identity information that may be distributed across multiple providers." Proponents of sharing: "consumers are concerned about need to have to remember multiple username/password pairs; consumers are concerned about re-authentication requirements; want dealings with multiple organizations to be seamless." Clarke 2004: "Consumers would like to avoid being subjected to large amounts of personal data disclosure, and that are able to continually add to that data in order to locate and track them."

We judge on the social desirability of alternative IMS structures by comparing

the values of total consumer utility accruing to each one of the alternatives. We model consumers' utility as a function of the extent of revealed partial identity (Clauss and Kohntopp, 2001) and build on the network interconnection model by Heal and Kunreuther (2002) to model incentives of the identity management firms. The resulting theoretical model yields several inferences and policy implications stemming from the predicted relationship between private and social optima with respect to the interconnection issue.

Since there clearly is a tradeoff between the extent to which privacy is protected and the effectiveness with which the Government is able to control individuals' actions, privacy must be compromised to a certain degree. Or, as the Open Group put it, "… the desire for privacy and individual dignity must be reconciled with the desire for effective government and with legal needs and national security needs." (Open Group) Limited acceptance of Microsoft Passport due to "the reluctance of the public to trust any single organization to provide a universal identity management solution, reinforced by the fact that security question marks have been raised relating to the specific Passport implementation." (Open Group) Hansen et al. (2003) "On the one hand, in particular legal contexts reliable identification of a person is necessary; and, on the other, the structuring and representation of identity is based in human rights law."

In this study we are looking for the scope for balanced solutions to the problem of identity management focusing on the issue of interconnection between alternative IMS firms. We are especially interested in identifying the type of environment in which the individual incentives of IMS firms push them to interconnect in a way that is socially suboptimal. Designing policies for this type of environment is a challenging goal for the policy makers that can be better achieved when backed by a better understanding of the economic processes behind interconnection.

IV. Incentive Structure of Consumers and IMS Firms

Clarke (1994) suggests there exists an important difference between the real world of physical existence and an abstract world of information. Physical entities (such as people or organizations) possess attributes (such as name and age). In the same way as people perform many roles in different contexts, a physical entity poses as a different identity in each type of these contexts. For example, the same man can be a client of his bank, an employee of his boss, a goalkeeper in the football team etc. Each one of the roles this man plays is associated with different sets of attributes such as credit history, number of cars sold or the amount of goals kept. In this way, an entity may have more than one identity associated with it.

Following Heal and Kunreuther (2002), we postulate that consumers are completely identified by a set of informational atoms a_i belonging to set A called one's complete (or full) identity. Any subset of *A* is called partial identity. Partial identity can be thought of as a piece of personal information an individual is willing to reveal in order to obtain a specific type of service provided by an information management system. We assume that the amount of services provided for the consumer by an IMS firm is equal to the extent to which a consumer has revealed her partial identity to the IMS firm. To be more explicit, we impose that the maximum amount of the provided services is equal to the norm of the information set A which for simplicity we take to be the interval of [0,1].

Denote S to be the set of all information services s_i that cannot be split any further. An IMS is defined as an entity that exchanges any set of elementary information services s_i for consumers' partial identities, that is, subsets of A. Each service s_i is assigned a numerical value from [0,1] such that (without loss of generality) the sum of all s_i-s is normalized to 1.

Connecting to an IMS produces two effects. Getting access to a subset of S associated with any given IMS delivers certain utility. On the other hand, disutility results as well since consumers are aware of the risks associated with possible misuse of their private information. We thus represent our utility

function as composition of utility and disutility of revealing information with the general property that

$$\begin{cases} U(s_j) > U(s_i), s^* > s_j > s_i \\ U(s_j) < U(s_i), s^* < s_i < s_j \end{cases} \quad (1)$$

where s^* is the threshold of privacy disclosure beyond which the disutility effect caused by privacy concerns starts dominating the positive utility effect caused by the increase in provided informational services. Utility function (1) can be thought of as reflecting both utility and disutility effects produced by disclosing a certain amount of private information. A construct similar to (1) is the decreasing deadline utility function used in health economics (Murthy and Sarkar, 1997). Alternatively, revealing partial identity can be thought of as the combined consumption of goods (the services provided by IMS-s) and bads (risk of misuse of information, identity theft and the like).

The population as a whole consumes all of the available information services. The amount of private information (partial identity) that each consumer reveals to any single IMS is directly proportional to the amount of information services provided to her in exchange, so that consuming more services entails revealing more partial identity. We assume that the number of consumers connected to each IMS is directly proportional to the amount of services this particular IMS provides. There are two ways in which we rationalize this assumption. First, it is unlikely that smaller IMS-s will provide a wide array of services due to e.g. lack of economies of scale and scope. Second, a wider array of services will likely to be designed in such a way as to serve various types of consumers in order to cater to a wider segment of the population, naturally increasing the market share of the IMS with a wide array of services. For that latter reason, latter firms will need to solicit more private information from the consumers so as to be able to discern between the types.

In line with the main conclusions of the literature on price discrimination, the firms that know more about their customers' types will be able to extract a wider

fraction of their consumer surplus. Since using the IMS services in our model is costless, consumers' utility from connecting to an IMS is equal to consumer surplus. Without loss of generality we assume that the IMS firms extract all consumer surplus.

In order to impose more structure on the consumers' utility function described above, we postulate a quadratic form as follows:

$$U(s) = -as^2 + bs \qquad (2)$$

where $a > 0, b > 0$[2]. Recall that apart from measuring the amount of consumed information services, s is also proportional to the market share of the IMS that provides these services. Since this function reaches its maximum at $s^* = \frac{b}{2a}$, more privacy-conscious consumers will be characterized by a lower ratio of $\frac{b}{a}$.

Interconnection between any two IMS-s in our model produces two effects. First, as mentioned already, the firms share private identities of their existing customers with each other. Second, interconnection results in a greater market share commanded by the two interconnected IMS together. Thus, if IMS_1 commanding market share s_1 interconnects with IMS_2 that enjoys market share s_2 the new IMS resulting from the interconnection of the previous two will command market share $s_1 + s_2$.

In our framework, if an IMS is commanding a market share of s for a group of consumers characterized by parameters a and b in (2), its payoff $P(s)$ is defined as:

$$P(s) = s\left(-as^2 + bs\right) \qquad (3)$$

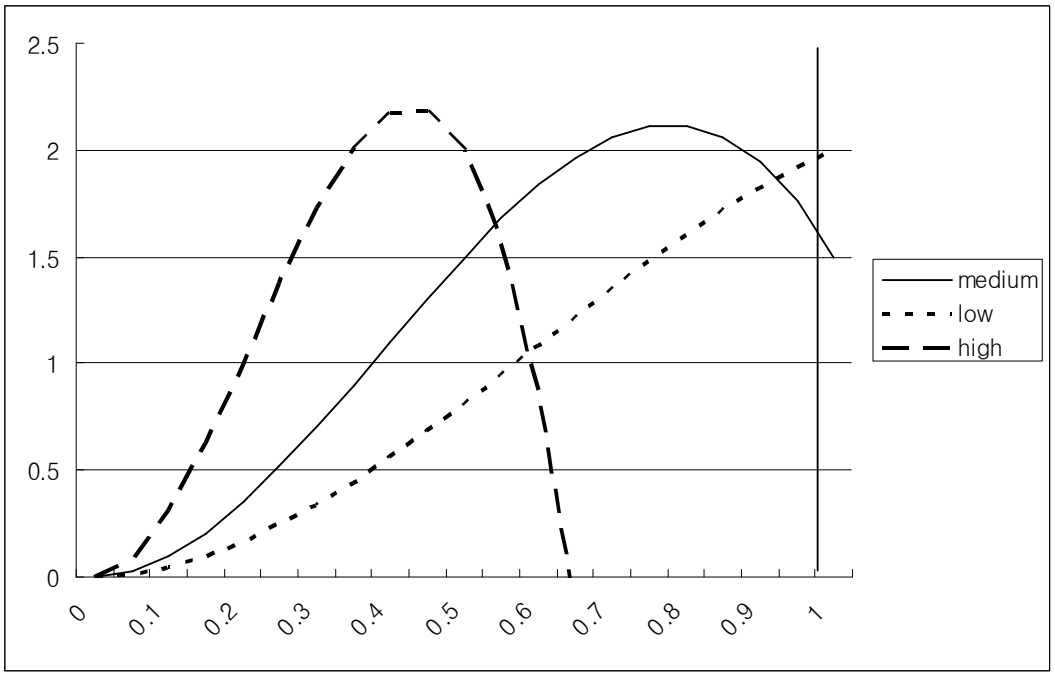
[2] Specification (2) can be thought of as the second-order Taylor approximation of the composition of utility and disutility effects. In this way (2) accommodates any utility function where the former effect dominates first, while the latter dominates later.

Payoff function (3) reaches its maximum at $s^{**} = \frac{2b}{3a}$, which is the market share beyond which IMS-s' incentives to interconnect disappear. We consequently call s^{**} the privacy threshold. Note that in our framework the IMS payoff is equivalent to combined consumers' surplus since it is the product of the fraction of consumers who use this particular IMS and the level of their utility and since using the IMS services incurs no charges. (We do not need to assume, as we do here implicitly, that the IMS firms are able to extract all of the consumer surplus since nothing changes if we constrain the IMS-s to be only able to extract a *fixed* fraction thereof.)

The schedule of payoff function (3) for different levels of privacy thresholds is given in Graph 1 below:

Graph 1

Plots of Payoff Functions for Medium, Low and High Levels of Privacy Thresholds



Note: the fraction of partial identity revealed is marked on the horizontal axis, the vertical axis marks the levels of the payoff functions. The privacy threshold is defined as that level of partial identity revealed at which the IMS's payoff function reaches its maximum.

Privacy threshold s^{**} may or may not exceed the maximum possible amount of the provided information services. In the former case consumers are fine with the dominant monopolistic IMS (the dotted line in Graph 1 representing consumers with low privacy concerns). In contrast, in the latter case the disutility effect of revealing private identity starts dominating the utility effect before the complete identity is revealed so that the IMS-s find it optimal to stop growing at the level of partial identity that is short of complete identity (the dashed and normal lines in Graph 1). The greater are consumers' privacy concerns, the smaller the maximum amount of information the IMS-s will find it

optimal to stop at soliciting. In Graph 1 above the dashed line corresponds to consumers with high privacy concerns, representing the privacy threshold that is lower than that of consumers with the medium privacy concerns represented by the normal line.

Those IMS firms that are relatively small market participants (commanding smaller market shares) would have incentives to interconnect with their larger peers, but these incentives will be constrained by the working of the privacy concerns effect that gets increasingly stronger once the privacy threshold is exceeded.

V. Deriving the Optimal Structure of Interconnection

Denote $\tilde{s}$ the smallest solution to

$$P(s)=1 \tag{4}$$

It is socially optimal for the smaller IMS (of size $s<\tilde{s}$) to interconnect to the IMS of any size, while only certain size classes will be socially attractive (in terms of interconnection) for the IMS firms that are sized in the interval of $\tilde{s}<s<s^{**}$. Namely, IMS-s in the range $[0, s'-s]$, where s' is the size of IMS that results in the same value of payoff function $P(s)$ as the one currently enjoyed by the IMS in question. It is easy to see that the closer the IMS firm is to size s^{**} the narrower the size class it would like to choose firms for interconnection with. The IMS firms of size greater than s^{**} will not be willing to interconnect with anyone.

In this section we investigate which IMS should interconnect with each other in order to result in the highest consumer surplus rather than whether the interconnection should occur at all. We consider a hypothetical case of three IMS firms of sizes s_i, s_j and s_k with the restriction that their sum does not exceed unity. The interconnection process may result in an omnipotent IMS structure when one single firm provides all of the services and has size $s_i + s_j + s_k = 1$, the structure with two IMS-s one of which is an interconnection of (s_i, s_j), (s_i, s_k) or (s_j, s_k), and a completely fragmented structure which is the original structure. The way we judge about the preferability of any one of these five outcomes is by looking at the associated payoff and comparing them to each other. We first consider the case when both omnipotent and fragmented systems are suboptimal relative to the structure where two IMS-s interconnect.

In order to see which type of interconnection delivers the highest value of the IMS-s' joint payoff function, it is sufficient to find the conditions under which connecting any given IMS firm (say, size s_k) to any other one of the remaining two (say, size s_i) is optimal relative to connecting s_k to s_j. We thus denote the corresponding values of the IMS structures' joint payoff functions as $P(ik, j)$ and $P(jk, i)$ and consider their difference:

$$\Delta P(i, j) = P(ik, j) - P(jk, i) \tag{5}$$

Substituting (3) into (5) yields the following condition for the interconnection structure (ik, j) to be preferable:

$$\Delta P(i, j) = s_k (s_j - s_i)[2b - 3a] > 0 \tag{6}$$

According to the optimality condition (6), the total payoff function will be

maximized if IMS_k interconnects with the smallest IMS in case privacy constraints are binding (that is, if the maximum of payoff function $P(s)$, $s^{**} = \frac{2b}{3a}$, does not exceed the maximum possible amount of information services, which in our framework we normalized to 1. We can thus summarize our main finding in the following proposition:

Proposition 1.

a) Fragmented structure is preferred to an interconnected structure if the privacy concerns are large enough (namely, if $\frac{b}{a} < \frac{3}{2}(s_i + s_j)$)

b) Monopolistic structure is preferred to an interconnected structure if the privacy concerns are small enough (namely, if $\frac{b}{a} > \frac{1 - 3s_i s_j (s_i + s_j) - (s_i^3 + s_j^3 + s_k^3)}{s_i + s_j - (s_i - s_j)^2}$).

(*The proof of the above proposition can be found in the Appendix*.)

Proposition 2.

In case of the binding privacy constraints, the following holds:

a) If an interconnected structure in an original three IMS system is preferable to both monopolistic and fragmented structures, for any two alternative interconnection structures the one that connects any given IMS to the smaller alternative maximizes the total payoff function.
b) Given the conditions in a) hold, the structure of IMS that interconnects two smallest firms will maximize total payoff function.

(*The proof of the above proposition can be found in the Appendix*.)

The message of Proposition 2 can be extended to the case of more than three IMS firms. Indeed, consider a system of *N* IMS firms such that, without loss of generality, $s_1 < s_2 < ... < s_N$. Consider a subsystem of three IMS firms that includes the first smallest two and any third one, say IMS_3 sized s_3. Since the total payoff function is additive in the payoff functions of individual IMS firms, we can consider the subsystem of three IMS-s as a separate entity so that the optimal interconnection within this system will imply the optimal interconnection

structure for the whole system of *N* IMS-s. In case of the three systems, however, Proposition 2 implies the interconnection between the two smallest systems is optimal. We can thus extend Proposition 2 to the following:

Corollary

In case the IMS structure in which some two of the original IMS firms interconnect is preferable to both monopolistic and fragmented structures, interconnecting the two smallest IMS firms is optimal.

VI. Summary and Policy Implications

In this study we offered a framework for designing a policy for interconnection between alternative identity management systems (the IMS-s). We believe the scope for such policy stems from the fact that consumers have concerns about privacy, that is, the utility gains from acquiring the many services in exchange for supply of private information (such as e.g. the credit card or house loan services) can be mitigated or even offset by increases in the disutility of providing this private information. Our key assumption is that the process of interconnection results in the IMS-s sharing the private information on consumers they obtain prior to such interconnection, thus increasing their ability to extract consumer surplus. Our key finding is that interconnecting the two smallest firms with each other is preferable to any other type of interconnection, given the interconnected structure is itself preferable to both monopolistic and fragmented structures.

Our major policy implication is that the process of soliciting private information from individuals by businesses should be regulated since the socially optimal outcome may differ from the privately optimal outcome. For example, even if interconnecting the two smallest firms in a system of three or more IMS-s results in a larger overall consumer surplus, such interconnection is not necessarily what either or each one of these firms will find it to do optimal for itself.

Second, given the fact that privacy as a measurable concept has not yet been defined, we suggest to design a uniform way in which one's identity can be measured and recorded. While realizing that it is impossible to digitize each and every element that constitutes the very complex human personality, we believe a substantial part of it can and should be, helping to make the design of identity-related policies based more on rigorous analysis rather than on someone's arbitrary judgment.

We plan to extend this research in several ways. First, we find important to derive the conditions under which the divergence of private and social interests mentioned above might occur. In other words, we want to know when the unregulated process of interconnection will result in the IMS structure that is

socially suboptimal. Second, in our framework we postulated the quadratic "downward-looking" utility function that depends on the amount of provided informational services. Naturally, more general forms of such function should be considered in order to render our findings more robust to the functional specification. Along the same lines, it appears to be worthwhile to consider the case of more than three IMS-s. Third, we assumed that there are no externalities to merging two pieces of information, that is, combining these two pieces does not reveal anything about any third element of one's partial identity. Finally, an essential assumption in our work is that no two IMS firms provide the same kind of informational services to consumers, or equivalently, the intersection of the partial identities used by any two IMS-s is an empty set. This is clearly not the real-world case since parts of identity like one's name or age are likely to be asked for in most questionnaires.

Keeping all the caveats outlined in the paragraph above in mind, we believe this study is a useful step on the way of formalizing the identity-related policy design which may help make policy decisions in the area of identity management more educated.

Appendix

Proof of Proposition 1.

Part a: fragmented structure

Denote $P(ij,k)$ the value of total payoff function for the optimal interconnection structure, in this case (ij,k) without loss of generality. Similarly, $P(i,j,k)$ will be the total payoff function's value in case of the completely fragmented structure of IMS. The conditions under which the difference of the former with the latter is negative are also the conditions for the fragmented structure to be preferable compared to the interconnected one.

$$\Delta P^F = P(ij,k) - P(i,j,k) = -as_i^3 - as_i s_j^2 - 2as_i^2 s_j + bs_i^2 + bs_i s_j - as_i^2 s_j - as_j^2 - 2as_i s_j^2 +$$
$$+bs_i s_j + bs_j^2 + as_i^3 + as_j^3 + as_k^3 - bs_i^2 - bs_j^2 - bs_k^2 = \{s_i + s_j + s_k = 1\} = s_i s_j (-3as_i - 3as_j + 2b)$$

wherefrom $\Delta P^F < 0 \Leftrightarrow \frac{b}{a} < \frac{3}{2}(s_i + s_j)$.

Part b: monopolistic structure

Similarly to the argument in part a) above, consider the difference

$$\Delta P^M = P(ij,k) - P(ijk) = -a(s_i^3 + s_j^3 + s_k^3) - 3as_i s_j (s_i + s_j) + 2bs_i s_j + b(s_i^2 + s_j^2 + s_k^2) + a - b$$

where $P(ij,k)$ is the most preferable interconnection structure. Exploiting the fact that the sum of all original IMS systems's sizes is equal to 1 yields the following conditions on the utility function parameters (or on the extent of privacy concerns) that ensure the monopolistic structure is optimal:

$\Delta P^M < 0 \Leftrightarrow \frac{b}{a} > \frac{1 - 3s_i s_j (s_i + s_j) - (s_i^3 + s_j^3 + s_k^3)}{s_i + s_j - (s_i - s_j)^2}$, which completes the proof.

Proof of Proposition 2.

Part a)

Let IMS_k consider interconnecting with either IMS_i or IMS_j. The sizes of these IMS-s are s_k, s_i and s_j, respectively. Denote the total payoff function for the interconnection structure whereby IMS_k connects to IMS_i as $P(ik, j)$. It follows from (3) that

$$P(ik, j) = (s_i + s_k)\left[-a(s_i + s_k)^2 + b(s_i + s_k)\right] + s_j\left[-as_j^2 + bs_j\right] \quad \text{(A1)}$$

Similarly,

$$P(jk, i) = (s_j + s_k)\left[-a(s_j + s_k)^2 + b(s_j + s_k)\right] + s_i\left[-as_i^2 + bs_i\right] \quad \text{(A2)}$$

Expanding the squares and products in (A1) and (A2) yields the following expression for the difference between values of total payoff function for the two alternative types of the interconnection structure:

$$\begin{aligned}
&\Delta P = P(ik, j) - P(jk, i) = -as_i s_k^2 - 2as_i^2 s_k + 2bs_i s_k - as_i^2 s_k - 2as_i s_k^2 + \\
&+as_j s_k^2 + 2as_j^2 s_k - 2bs_j s_k + as_j^2 s_k + 2as_j s_k^2 = \\
&= s_k\left[-as_i s_k - 2as_i^2 + 2bs_i - as_i^2 - 2as_i s_k + as_j s_k + 2as_j^2 - 2bs_j + as_j^2 + 2as_j s_k\right] = \\
&= s_k\left[-3as_i s_k - 3as_i^2 + 2bs_i + 3as_j s_k + 3as_j^2 - 2bs_j\right] = \\
&= s_k(s_i - s_j)\left[2b - 3as_k - 3a(s_i + s_j)\right]
\end{aligned} \quad \text{(A3)}$$

Since by definition $s_k = 1 - s_i - s_j$, we can rewrite the last expression as:

$$\Delta P = s_k(s_i - s_j)(2b - 3a) \quad \text{(A4)}$$

In case of the binding privacy constraints (which is the case of interest corresponding to $2b < 3a$) ΔP will be positive if and only if $s_i < s_j$. In other words, the IMS_k will prefer to interconnect with IMS_i rather than to IMS_j if the former commands a smaller market share. Shortly, in case of the binding privacy constraints interconnecting to a smaller IMS maximizes the total payoff

function of the providers of IMS services.

Part b)

Suppose without loss of generality that $s_i < s_j < s_k$ and (ik, j) is an optimal structure. That implies $(ik, j) \succ (ij, k)$ in the sense of $\Delta P > 0$ in (A3). The latter implies, by the proof of part a) above, that $s_k < s_j$, which is a contradiction to our initial conditions. Similarly, $(jk, i) \succ (ji, k) \Rightarrow s_k < s_i \lhd\rhd s_i < s_j < s_k$.

End of Proof.

References

Bennett C (1992) Regulating Privacy: Data Protection and Public Policy in Europe and the United States (Ithaca: Cornell University Press).

Clauss, S., and Kohntopp, M., “Identity management and its support of multilateral security”, Computer Networks, 2001.

Davis, J.C., “Protecting privacy in the cyber era”, 2000.

Economides, N., “The economics of networks”, International Journal of Industrial Organization, 1996.

Gavison R (1980) 'Privacy and the Limits of the Law', Yale Law Journal 421.

Hansen, M., and Berlich, P., “Identity Management Systems: Gateway and Guardian for Virtual Residences”, Emtel Conference, 2003.

Heal, G., and Kunreuther, H., “Interdependent security: the case of identical agents”, mimeo, 2002.

ICPP and SNG, “Identity management systems (IMS): identification and comparison study”, Independent Centre for Privacy Protection (ICPP) and Studio Notarile Genghini (SNG), Contract No. 19960-2002-10 F1ED SEV DE, September 2003.

Kelsen, 1966. Quoted from Pizzorusso, Scialoja, Branca—page 3. Galgano, Struttura logica e contenuto normative del concetto di persona giuridica, Riv. Dir. Civ., I, 553-633. Kelsen, La dottrina pura del diritto, Einaudi, 1966, 200; Teoria generale del diritto e dello stato, Etas, 1978.

Lehnhardt, M., “Identitat im Netz: das Reden von der “Multiplen Personlichkeit””, in: Martin Rost (Ed.): Die Netzrevolution—Auf dem Weg in die Weltgesellschaft, Eichborn, Frankfurt am Main, 1995.

Mead, G.H., “Mind, Self and Society”, Chicago Press, 1934.

Olivero, N., and Lunt, P., “Privacy versus willingness to disclose in e-commerce exchanges: the effect of risk awareness on the relative role of trust and control”, Journal of Economic Psychology, Vol. 25, pp. 243-262, 2004.

Wakaha, O., et al., “New combinatorial designs and their applications to authentication and secret sharing schemes”