Privacy Economics: From information theory to privacy as an asset

Carly Beckerman[*] and Julian Williams

Synonyms

Information economics; value-of-privacy

Definitions

Privacy economics is a branch of economics that studies the value, protection, and market for privacy and private information. This subfield has several defining features: it considers the construction and nature of information in economic activities, deals typically with personal information, considers the disclosure and dissemination of that information by one or more third parties, and questions the utility and impact of these activities if disclosure is outside of the originating owner's control.

Background

Privacy economics is a subfield of information economics. It focuses on the disclosure and economic use of a type of information that is typically generated and owned by individuals. Privacy economics provides an umbrella term for understanding when this information may be disclosed, subsequently aggregated and used by a third party in a variety economic activities. Debates surrounding privacy economics largely fall within three frames of reference: agency, asymmetric information, and public good.

Importantly, the notion of agency gives meaning to concepts of private information and valuation. Privacy economics seeks to understand the broad collection of economic activities associated with the collection, aggregation, transmission, and utilization of private information, as well as how

[*] corresponding author

these activities affect individual and collective welfare. In this context, agency refers to an individual, or collection of individuals, that is treated as a primitive homogenous entity possessing specific attributes and preferences. The agent may prefer to conceal information or to disclose it. Although the broad literature on privacy economics does not always refer explicitly to the importance of agency, most of the relevant scholarship reflects this core tenet. This is evident in the long history of economic approaches to modelling privacy, with Posner (1981) being an early explicit example and Acquisti et al (2016) representing a more recent survey.

Early work on privacy and economics focused on the payoff structure resulting from asymmetric information. This ’first wave’ of scholarship has made many formative contributions. However, while it recognises that privacy is a multi-faceted concept, the analysis has been focused on how asymmetric information impacts various aspects of economic decision-making. Posner (1981) for example, refers to privacy narrowly as ”the concealment of information”, ”peace and quiet”, or as ”a synonym for freedom and autonomy” (Posner 1981, p405:1).

Subsequent theoretical constructions then modelled economic activities involving the asymmetric possession of private information. For instance, Grossman and Stiglitz (1976), Grossman and Stiglitz (1980) and Hellwig et al (1980) instigated basic models of private information in securities trading. Hart (1983) outlined the importance of private information in wage-bargaining and the setting of optimal employment contracts. Using similar modelling strategies, Rothschild and Stiglitz (1976) and Hellwig (1988) have addressed the value of information in determining relative value for insurance contracts. There are also models of asymmetric information in credit markets. Bester and Hellwig (1987) and later Pagano and Jappelli (1993), amongst others, study the welfare effects of hidden action and private information in loan contracts.

This classic literature focuses on how asymmetric information affects standard economic contracts, which are traditionally modelled under full information. In this sense, the mechanism of privacy is a deliverable tangible economic benefit, allowing individuals to engage in hidden actions and exploit information asymmetries. Within this ’first wave’-inspired research, the value of private information is via a second order effect. There is no explicit welfare gain or loss derived from the exposure of information.

This contrasts to the ’second wave’ of privacy economics scholarship, which investigates privacy as an asset coupled to the notion of individual and group psychological wellbeing. See, for instance, Brandimarte and Acquisti (2012) and Loewenstein et al (2014).

Further works, such as Tucker (2012) and Acquisti and Fong (2020) have revisited the classical paradigm while appreciating that privacy, as well as the perception, realization, and associated risks of disclosure, carry explicit costs and benefits. This means that the intrinsic value of an individual piece of information changes as it moves from a private to public sphere of consumption. This is also true in the reverse. The digital economy of sharing via social media generates broad welfare effects, but third parties will harvest, collect, codify, and analyze vast tracts of this personal information. The cost of large scale information collection is offset by the sizeable socio-economic and political gains expected from the careful modelling and forecasting of individual behavior across a range of contexts. Anticipated gains will vary widely. Within liberal democracies, for example, perceived breaches of citizen privacy may trigger audience costs Fearon (1994). Within authoritarian regimes, benefit may be derived from the suppression of online dissent and the tracking of activists and journalists. These more varied types of effects are documented in Fjell et al (2010), Tsai et al (2011), and Egelman and Peer (2015), amongst others across the social sciences.

Such activities have also sparked new debates regarding regulation. Tene (2013), for example, provides an overview of global privacy laws, while Kerber (2016) and Romanosky et al (2011) provide theoretical and empirical perspectives on the effectiveness of legal mechanisms. Public support for global privacy protection varies widely, however. In the United States, for example, support for international laws and rights is strongly correlated with political ideology and level of individualized morality. See Stolerman and Lagnado (2019).

Theory

Models that embed the notion of privacy within the welfare function of individuals must formalize the mechanisms that define the information object and degree of privacy.

A simple construction is as follows. Consider an information universe with three spheres: private, collection, and utilized. Let $\mathscr{P}$ be the abstract set of information production, private valuation and retention; $\mathscr{C}$ be the abstract set of collected information and collection mechanisms; and $\mathscr{U}$ be the abstract set of processed public information, analytical tools, payoffs, and actions. The analogy to social media, would be that $\mathscr{P}$ is the set of private activities, by one or more individuals, that generates a collection of private information. This might include images, purchasing preferences, and personal messages. $\mathscr{C}$ is the compehensive or partial codified data that is collected and stored by some technology from $\mathscr{P}$, which can be combined with other information collected across a range of sources. Finally, $\mathscr{U}$ is a series of tools, including search engines, public-facing social media outlets, and private analytical frameworks such as machine learning algorithms. The payoff component of $\mathscr{U}$ contains the ex-post welfare for agents whose information sits within the non-private sphere.

Within the three information spheres, strategic actions are determined by individual agents. The technology of sharing determines the rules under which these strategic actions are determined. Most rules are imposed by one or a combination of infrastructure providers, legal mechanisms or social norms, see Campbell et al (2015).

Agents then make strategic decisions based on local frames of reference. Let $\mathscr{P}$, $\mathscr{C}$ and $\mathscr{U}$ define classes of agents within the spheres $\mathscr{P}$, $\mathscr{C}$ and $\mathscr{U}$ respectively. Agents can belong to one or more of the classes, and the transmission and transformation of information between spheres is a function of each individual agent's actions. Compositionally, this is via a set of general functions $p(\cdot)$, $c\overset{\cdot}{()}$ and $u(\cdot)$, where $\cdot$ is a placeholder for a variety of inputs. The action to disclose is defined by a set $\mathscr{A}$. The technology underpinning the information system then determines how $\mathscr{A}$ operates. The simplest form is a binary decision for each agent $A_i(X_j) \in \{0,1\}$ to disclose an individual piece of information X_j from the private to the public sphere, $\forall i, j \in \{1,\ldots,I\}$. In systems where disclosure is the individual agent's responsibility, then $j = i$ and the collection of all other actions $\boldsymbol{a}_{-i}$ is irrelevant. In other cases, subsets $\boldsymbol{d} \subset \{1,\ldots,I\}$ will have different voting rules, such as full consent for agents with enforceable property rights on X_j.

The information production mechanism for an individual is derived from the primitive object Y_i, which is an unstructured set of histories for a given agent i, $\forall i \in \{1,\ldots,I\}$. Let $\boldsymbol{y}_{-i}$ be the set of all histories for all agents excluding i. The production of personal information can be described by $\{X_i, M_i\} \leftarrow p(Y_i, \boldsymbol{y}_{-i})$, where X_i is a packet of specific personal information and M_i is the associated meta-data. The difference between Y_i and X_i is not strictly relevant. However, structurally, Y_i is a psychological representation, whereas X_i is a tangible object. Indeed, X_i may not be consciously understood by agent i and might be inferred through observation of X_{-i} and the identification of some prior information set X_i'.

Let $\mathscr{M}(M_i)$ be an operator that assigns ex-ante and ex-post property rights of X_i across the collection of agents. This notion is complex and can be specific, such as an enforceable copyright on the social network-use of a picture. When X_i is within the private sphere $\mathscr{P}$, then the value of X_i to all agents is given by $\boldsymbol{p}(X_i)$. This could be zero to all agents except i, which is not necessarily true for shared private information.

Transition of X_i from $\mathscr{P}$ is via a technology represented by an abstract function $c\overset{\cdot}{()}$. In the most general construction all agents have some actions $\mathscr{A}$ that determine the outcome of the collection function as it operates on X_i and M_i. This is denoted by $Z_i \leftarrow c(X_i, M_i, \mathscr{X}, \mathscr{M}, \mathscr{A})$, where $\mathscr{X}$ and $\mathscr{M}$ are the set of other possible information packets (these can of course be empty sets). This component represents the transformation of personal information into stored data. Denoting $\boldsymbol{z} \supset \{Z_i, \boldsymbol{z}_{-i}\}$ as the set of all information collected and processed from the private sphere.

The third sphere represents the public frame of reference. $\boldsymbol{u} = u(\boldsymbol{Z}, \mathscr{A})$ denotes welfare derived from the modelling and analysis of $\boldsymbol{z}$ within the public sphere, meaning that $\boldsymbol{u} := \{u_1, \ldots, u_I\}$. For a given X_i, the difference in welfare for the set of agents is given by $\Delta\boldsymbol{u} = \boldsymbol{u} - \boldsymbol{p}(X_i)$. This is list of changes in welfare for all agents following the transfer of X_i from the private to the public sphere.

When the gain from disclosure $u_i - p_i(X_j)$, $\forall j \in \{1,\ldots,I\}$ is positive, there is a positive welfare effect on agent i to allow their information asset X_j to be transformed and incorporated into the public information set $\boldsymbol{z}$. This is denoted by $\mathscr{U}_i[X_j] \succ \mathscr{P}_i[X_j]$. When $u_i - p_i(X_j)$ is negative, then i has a preference for this asset to not be included in the public information set $\boldsymbol{z}$. This is denoted by $\mathscr{U}_i[X_j] \prec \mathscr{P}_i[X_j]$.

The degree of reversibility in the original information indicates the degree of privacy preservation during its transition from the private to the public sphere. Consider the collection of inverse problems, whereby underlying information is reconstructed from the public presence:

$$p^{-1^*}(X_i^*, M_i^*) \rightarrow \{Y_i^*, y_{-i}\}, \quad (1)$$
$$c^{-1^*}(\boldsymbol{Z_i}) \rightarrow \{X_i^*, M_i^*\}, \quad (2)$$
$$u^{-1^*}(\boldsymbol{u}) \rightarrow \{\boldsymbol{Z_i}*\}, \quad (3)$$

where $f^{-1^*} \in \{p^{-1^*}, c^{-1^*}, u^{-1^*}\}$ is an approximate inverse of $f \in \{p, c, u\}$. We can then define a norm $\Delta(V) = ||V - V^*||^{-1}$ in Euclidean space where "$-$" is an abstract comparator operator element by element of the attributes of sets V and V^*. When $\Delta(V) \rightarrow \emptyset$, then the data is reconstructed perfectly. Starting in reverse, the first inverse maps the payoffs from the analysis of the public data set and the set of individual payoffs. The second inverse reconstructs approximations of the individual private information from the public dataset. The first inverse reconstructs an approximation of underlying personal histories from the reconstruction of the personal information and meta data.

Consider some examples of collective policies in regard to the sharing of private information:

Individual consent. The simplest case when the action $\mathscr{A}$ is solely at the discretion of the individual, hence disclosure only occurs when $u_i - p_i(X_i) > 0$.

Fully informed consent. In this case, data is transferred to the public sphere only with the consent of the individuals with claim $\mathscr{M}(M_i)$ above some threshold $\bar{M}$. If those individuals have positive welfare gains from disclosure, then the transfer actions are positive $\mathscr{A}$. A specific example of this type of action could be the disclosure of patent information from collaborative research for public review.

The greatest good policy. Consider a third party information aggregator, such as a network provider that monitors the production of X_i. A utilitarian approach aggregates the welfare payoffs for disclosure $\sum_i^I u_i - p_i(X_j)$ and operates on the principle of the greater good in terms of determining the action $\mathscr{A}$ to disclose. In a non-utilitarian case, a policy function $g(\mathscr{X})$ is used to determine disclosure. Examples of this type of disclosure can be found in command and control economies where information sharing is imposed. Other examples include enforced information sharing during a time of crisis, such as a natural disaster or pandemic. A further example is in the monitoring of physical and mental health for agents involved in critical activities, such as airline pilots and medical professionals.

Incentive schemes. This indicates a payment of measured adjusted welfare $\boldsymbol{\varpi} := \{\varpi_1, \ldots, \varpi_I\}$ made to agents to ensure that $u_i - p_i(X_j) + \varpi_i > 0$ is sufficient to warrant the transfer X_j to the public sphere, $\mathscr{A}$. The simplest example is providing sufficient welfare (either from payments or provision of in-kind services) to incentivize the disclosure of private information. Examples include requiring users to enter contact details to play online games, allowing the tracking of habits on social media, and paying directly for access to personal records.

It is also worth noting that part of the key value of social networks is in status signalling gained from the deliberate exposure of private information to a broad public group. The gains from exploiting privacy as a property right appear asymmetric and develop with the evolution of social norms. Hence, a privacy based Kuznet curve, equivalent to the environmental and social Kuznet curves surveyed and discussed in Dinda (2004), Bazillier and Sirven (2008) and Dobson and Ramlogan (2009) could be constructed for differentiated degrees of development in privacy, privacy technology and privacy regulation.

The desire for differentiated privacy across collections of personal objects can be viewed as a straight forward set of preferences following the conventional choice axioms. For example, we can envision a consistent choice experiment with participants considering bundles of privacy options. Standard choice axioms would be applied to the privacy bundles using tools such as the general form of axiomatic revealed preference (GARP). This is helpful in understanding the consistency of choices revealed on social media versus other platforms such as mobile phone networks. It is then possible to evaluate the convexity of preferences trading-off between the control of information-disclosure, the desire for dissemination and the value of choice. Techniques such as Afriat's theorem are useful in this regard, see Afriat (1967) and Fostel et al (2004).

Open problems and Future directions

Moving forward, privacy economics is concerned with four domains of interest. First is the interaction between privacy economics and network economics. This has been explored qualitatively in Bramoullé et al (2014), but there is scope to embed notions of privacy and disclosure within a fully realized economic model. Second, the interaction between Artificial Intelligence and Machine Learning has implications for privacy economics. AI and ML algorithms are currently used to analyze harvested data, with and without consent. See, for instance, the use of Facebook data by Cambridge Analytica documented in Isaak and Hanna (2018), among others. Third, there is considerable work on pro-social behavior, see Bénabou and Tirole (2006). Observed social networks have communities that formulate codes of practice without the need for a central coordinating entity. Privacy and privacy protection in online communities offers a deeper insight into collective human behavior. Finally, new technologies, such as zero knowledge proofs, permit far more complex privacy preserving arrangements while still allowing some insight into behaviors that may have marketable value. Kumaresan et al (2015), for example, uses a combination of cryptocurrencies and multi-party computation to generate a poker game that has tangible financial stakes but requires no centralized authority, even for the game phase. The card dealing and winner validation is all accomplished under mutual distrust and fully preserves privacy. This type of entity has interesting implications for the future of privacy economics. Indeed, this fully mutually distrustful approach has been implemented in several cases for information architectures that have privacy embedded in their design in such a way that, even when compelled under legal notice, the infrastructure provider is physically unable to circumvent the privacy preserving features.

Cross-References

Privacy in social networks, Privacy laws and directives, Privacy-preserving data mining; economics of surveillance

References

Acquisti A, Fong C (2020) An experiment in hiring discrimination via online social networks. Management Science 66(3):1005–1024

Acquisti A, Taylor C, Wagman L (2016) The economics of privacy. Journal of economic Literature 54(2):442–92

Afriat SN (1967) The construction of utility functions from expenditure data. International economic review 8(1):67–77

Bazillier R, Sirven N (2008) Core labour standards and inequalities: Is there a social kuznets curve? Journal of Development Studies 44(7):913–934

Bénabou R, Tirole J (2006) Incentives and prosocial behavior. American economic review 96(5):1652–1678

Bester H, Hellwig M (1987) Moral hazard and equilibrium credit rationing: An overview of the issues. In: Agency theory, information, and incentives, Springer, pp 135–166

Bramoullé Y, Kranton R, D'amours M (2014) Strategic interaction and networks. American Economic Review 104(3):898–930

Brandimarte L, Acquisti A (2012) The economics of privacy

Campbell J, Goldfarb A, Tucker C (2015) Privacy regulation and market structure. Journal of Economics & Management Strategy 24(1):47–73

Dinda S (2004) Environmental kuznets curve hypothesis: a survey. Ecological economics 49(4):431–455

Dobson S, Ramlogan C (2009) Is there an openness kuznets curve? Kyklos 62(2):226–238

Egelman S, Peer E (2015) The myth of the average user: Improving privacy and security systems through individualization. In: Proceedings of the 2015 New Security Paradigms Workshop, pp 16–28

Fearon JD (1994) Domestic political audiences and the escalation of international disputes. American political science review 88(3):577–592

Fjell K, Foros Ø, Steen F (2010) The economics of social networks: The winner takes it all?

Fostel A, Scarf HE, Todd MJ (2004) Two new proofs of afriat's theorem. Economic Theory 24(1):211–219

Grossman SJ, Stiglitz JE (1976) Information and competitive price systems. The American Economic Review 66(2):246–253

Grossman SJ, Stiglitz JE (1980) On the impossibility of informationally efficient markets. The American economic review 70(3):393–408

Hart OD (1983) Optimal labour contracts under asymmetric information: An introduction. The Review of Economic Studies 50(1):3–35

Hellwig MF (1988) A note on the specification of interfirm communication in insurance markets with adverse selection. Journal of Economic Theory 46(1):154–163

Hellwig MF, et al (1980) On the aggregation of information in competitive markets. Journal of Economic Theory 22(3):477–498

Isaak J, Hanna MJ (2018) User data privacy: Facebook, cambridge analytica, and privacy protection. Computer 51(8):56–59

Kerber W (2016) Digital markets, data, and privacy: competition law, consumer law and data protection. Journal of Intellectual Property Law & Practice 11(11):856–866

Kumaresan R, Moran T, Bentov I (2015) How to use bitcoin to play decentralized poker. In: Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, pp 195–206

Loewenstein G, Sunstein CR, Golman R (2014) Disclosure: Psychology changes everything. Annu Rev Econ 6(1):391–419

Pagano M, Jappelli T (1993) Information sharing in credit markets. The Journal of Finance 48(5):1693–1718

Posner RA (1981) The economics of privacy. The American economic review 71(2):405–409

Romanosky S, Telang R, Acquisti A (2011) Do data breach disclosure laws reduce identity theft? Journal of Policy Analysis and Management 30(2):256–286

Rothschild M, Stiglitz J (1976) Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. The Quarterly Journal of Economics 90(4):629–649

Stolerman D, Lagnado D (2019) The moral foundations of human rights attitudes. Political Psychology

Tene O (2013) Privacy law's midlife crisis: A critical assessment of the second wave of global privacy laws. Ohio St LJ 74:1217

Tsai JY, Egelman S, Cranor L, Acquisti A (2011) The effect of online privacy information on purchasing behavior: An experimental study. Information systems research 22(2):254–268

Tucker CE (2012) The economics of advertising and privacy. International journal of Industrial organization 30(3):326–329