日本銀行ワーキングペーパーシリーズ

プライバシーの経済学入門

宇野洋輔*
yousuke.uno@boj.or.jp

園田章*

別所昌樹*
masaki.bessho@boj.or.jp

No.21-J-10
2021 年 6 月

日本銀行
〒103-8660　日本郵便（株）日本橋郵便局私書箱 30 号

* 決済機構局

プライバシーの経済学入門*

宇野洋輔† 園田章‡ 別所昌樹§

2021-5-31（2021-07-07 改訂）

要旨

本稿では、「プライバシーの経済学」と呼ばれる分野のサーベイを行う。インターネット空間における個人情報の取扱いに対する関心がグローバルに高まるなか、プライバシーの経済学は、近年のプライバシー保護規制当局による規制強化の動きと軌を一にしつつ急速に発展している。プライバシーの経済学が教えるところでは、社会的に望ましいプライバシー保護水準をどう決めるか、個人情報データが有する「負の外部性」に起因するプライバシーの侵害にどう対処するか、といった問題を市場メカニズムによって解決することは難しい。こうした認識は、デジタル決済システムを利用する人々に安心感を与えつつデータの利活用をどう進めていくかを考える際に、重要な示唆を与えうるものである（*JEL* D62、D82、D83、K20、M31、M37）。

1 はじめに

インターネット空間におけるプライバシーの保護は、近年、グローバルに重大な関心を集めている。ひとつの契機になったのは、Cambridge Analytica による個人情報データの不正利用スキャンダルである。Cambridge Analytica は、選挙コンサルティング目的で Facebook ユーザー 5,000 万人分の個人情報データを不正に入手し、2016 年の米国大統領選挙や英国の EU 離脱国民投票の際に利用したといわれている（Guardian 2018; New York Times 2018）[1]。

*本稿の作成にあたっては、渡辺安虎氏、神山一成氏、副島豊氏、奥野聡雄氏、山田健氏、鳩貝淳一郎氏、北條真史氏、宇根正志氏、菅和聖氏、白木紀行氏、高野裕幸氏から有益なコメントを頂いた。記して感謝したい。もちろん、あり得べき誤りは筆者らに属する。また、本稿に示される内容や意見は、筆者ら個人に属するものであり、日本銀行の公式見解を示すものではない。

†日本銀行（yousuke.uno@boj.or.jp）

‡日本銀行

§日本銀行（masaki.bessho@boj.or.jp）

[1]2016 年の米国大統領選挙について、Allcott and Gentzkow（2017）は、ソーシャルメディア上の「フェイク・ニュース」が選挙結果に影響を与えた実証的証拠はないと主張している。

このスキャンダルは、プラットフォーマーが如何に膨大な個人情報を有しているか、それが不正に利用されると如何に大きなインパクトがあるかを人々に認識させた。同時に、膨大な個人情報を収集して収益化するプラットフォーマーのビジネスモデルにも批判的な目が向けられるようになってきた[2]。たとえば、SNS ユーザーは、友人とのコミュニケーションのために写真に文章やタグを付けて投稿するが、SNS プラットフォーマーは、これを「ラベル付きの画像データ」として AI の訓練に利用する。この仕組みは、人々に一定程度の利便性を提供する一方で、得られるアップサイドの収益すべてをプラットフォーマーが得る構造ともみなせるため、「テクノロジー封建制（technofeudalism）」と呼ばれることがある（Posner and Weyl 2018）[3]。Lanier（2013）は、これが人々に適切なインセンティブを与えていないとして強い危機感を示している。

こうした背景もあって、主要な法域の規制当局は、近年、プライバシー保護に関する規制を相次いで導入している。欧州連合では、2018 年 5 月、消費者等のデータ主体による同意の条件などを厳格に定めた、一般データ保護規則（General Data Protection Regulation, GDPR）が施行されている。欧州におけるプライバシー保護規制の歴史は古く、GDPR は、1995 年に成立した、EU データ保護指令（Data Protection Directive）の後継の規制として生まれたものである。また、米国カリフォルニア州では、2020 年 1 月、消費者の個人情報の取扱いを厳格に定めた、カリフォルニア州消費者プライバシー法（California Consumer Privacy Act of 2018, CCPA）が施行されている。

規制当局の動きと軌を一にしつつ、アカデミアでも、プラットフォーマーの行動様式をどう理解すればいいのか、人々のプライバシーをどうすれば保護できるのか、GDPR や CCPA などのプライバシー保護規制はどう機能するのか、といった点に関する議論が活発になっている。プラットフォーマーの行動様式を理解するためには、機械学習やコンピューター・サイエンスといった領域の知識が不可欠である。また、GDPR や CCPA などの法規制の基本的な考え方を理解するためには、法学の知識も必要になる。さらに、望ましいプライバシー保護のあり方を考えるうえでは、どこにどのようなトレードオフが存在するのかを明らかにしようとする経済学的思考も有益である。「プライバシーの経済学」と呼ばれる分野では、これら複数の領域の知識やツールを用いながら、望ましいプライバシー保護のあり方が議論されている[4]。

[2] プラットフォーマーは、インターネット上の多くの場所で個人情報データを収集して人々の行動を追跡している。そうした行為を最も積極的に行っているのは、Google、Facebook、Twitter、Amazon、AdNexus、Oracle の 6 社である（Englehardt and Narayanan 2016）。

[3] この「封建的な」状態を脱するために、Posner and Weyl（2018）や Arrieta-Ibarra et al.（2018）は、インターネット空間で創出された個人情報データを労働の産物とみなす、「労働としてのデータ（data as labor）」という考え方を提唱している。

[4] Acquisti et al.（2016）の整理によれば、最近の「プライバシーの経済学」のブームは、三度目の波にあたるものである。第一の波は、1970 年代から 1980 年代にかけて、Posner（1978, 1981）や Stigler（1980）など「シカゴ・スクール」によって牽引されたもので、そこでは、プライバシ

本稿は、人々が安心して利用できるデジタル決済システムのあり方を考えるため、プライバシーの経済学が蓄積してきた共通認識を整理するものである。もとより、決済処理とは一定程度の個人情報（送金人、受取人、金額、日時）の受け渡しを伴うものである。加えて、近年では、決済のデジタル化やインターフェイスの多様化などにより、決済を取り巻くエコシステム内での個人情報データの受け渡しが様々なかたちで行われるようになってきている。また、プラットフォーマーは、新たな個人情報を求めてデジタル決済サービスの分野にも参入してきており（FSB 2019）、グローバルなプライバシー保護に対する関心の高まりが決済システムにおける個人情報データの取扱いに及ぶ可能性も高い。実際、2020 年に欧州中央銀行が行った一般利用型中央銀行デジタル通貨に関する市中協議では、「デジタル・ユーロが備えるべき特性」に対する回答として最も多かったのは、「プライバシー保護」（全体の 43%）であった（ECB 2021）。こうした状況を踏まえると、人々が安心して利用できるデジタル決済システムがどうあるべきか、プライバシーの経済学の共通認識を一度整理しておく意義は大きいと思われる。これは、決済システムやその周囲のエコシステム内におけるデータの利活用をどう進めるかという議論の前提となるものである。

本稿の構成は次のとおりである。2節では、プライバシーに関する基本的な事柄を整理する。3節では、企業が支払うプライバシー保護のコストを扱う。プライバシー保護は、消費者にベネフィットをもたらす一方、個人情報データの利用の制約を通じてコストを生む。このコストは、直接的には、個人情報データを利用して収益化する企業が支払うことになる。ただし、プライバシーの経済学では、消費者のプライバシー保護が企業側にもベネフィットをもたらすという興味深い報告もなされている。4節では、個人情報データが有する「負の外部性」について整理する。この負の外部性は、ある人が秘匿した情報が予測によって類推されてしまう状況において生じる。これがプライバシー保護に及ぼす影響は深刻であり、プライバシーの経済学における最も重要な論点となっている。5節は、本稿のまとめである。

2　プライバシーの基礎

本節では、プライバシーに関する基本的な事柄を整理する。具体的には、2.1節において、経済学におけるプライバシーの考え方を確認したあと、2.2節では、プライバシー保護の度合いを表現する、「差分プライバシー（differential privacy）」と呼ばれる概念を紹介する。

ー保護は有用な情報を隠すために非効率性を生むといった主張がみられた。第二の波は、1990 年代中頃で、暗号技術の役割や個人情報データの二次利用の含意が議論された。たとえば、Varian（1996）は、第三者に提供される個人情報が「少な過ぎる」場合に消費者が負担するコストが大きくなりうるとした。

差分プライバシーをツールとして用いると様々な政策的議論が可能になる。もっとも、2.3節でみるように、現時点では、差分プライバシーのパラメータを観察データから明らかにすることは難しい面がある。また、2.4節でみるように、合理的な差分プライベート・システムを市場メカニズムによって設計することも困難であると考えられている。

2.1 プライバシーとは何か

プライバシーの経済学では、自分自身と他者との間に境界を設けるとき、自分自身の側に属するものをプライバシーと考える。このため、プライバシーは、人それぞれに異なるものであると考えられている（Posner 1978, 1981; Acquisti et al. 2016）。たとえば、ある人にとっては雇用状態が、別の人にとっては健康状態がプライバシーになるかもしれない。さらにいえば、同じ人であっても、文脈によってプライバシーが異なることもありうるものと考えられている。たとえば、ある人にとって、レストランに食事に行ったという情報は、一緒に食事に行った相手が誰かによってプライバシーになりうるものである。

また、プライバシーの経済学では、プライバシーとは、人々の効用に内在的（intrinsic）に存在するものというより、道具的（instrumental）な価値を有するものと考えることが多い（Posner 1978, 1981; Acquisti et al. 2016）。たとえば、プライバシーとは、個人情報データのかたちでマーケティング目的に利用されることなどによって価値を生むものと整理されている。もっとも、最近では、2.3.1節でみるように、プライバシーが有するこのふたつの側面を実験によって明らかにしようする取組みもみられている（Lin 2021）。

なお、4節で詳しく議論するように、プライバシーの経済学の重大な関心は個人情報データが有する負の外部性にあり、これは、人々のプライバシーがそれぞれ異なっていることが暗黙の前提となっている。ただし、代表的な消費者を想定したマクロの議論を行う場合など、人々のプライバシーの差異が想定されないこともある（Jones and Tonetti 2020）。

2.2 差分プライバシーと呼ばれるツール

プライバシーの経済学では、プライバシー保護の度合いを表現する場合、差分プライバシーと呼ばれるツールが用いられることが多い（Ghosh and Roth 2011; Pai and Roth 2013; Hsu et al. 2014; Abowd and Schmutte 2019）。ここでは、差分プライバシーについて概観したあと（2.2.1節）、差分プライバシーの概念をプライバシーに関する効用関数に埋め込む（2.2.2節）。

2.2.1 差分プライバシー

差分プライバシーは、コンピューター・サイエンスの分野では比較的古くから知られている概念である（Dwork et al. 2006; Dwork 2006）。今、個人情報データを含むデータセットを D とし、このデータセット D から何らかの情報を取り出すための問合せ行為を「クエリ」と呼び、Q と書く。容易に想像できるように、問合せの結果 $Q(D)$ は、データセット D に含まれる個人情報データを暴露してしまう可能性がある。

たとえば、寺田（2019）の例では、男女 20 人のクラスにおける試験の結果を記録したデータセットから、「受験者の性別」と「合否結果」を抽出するクエリを実行する。問合せの結果、合格者がすべて女性だった場合、クラスの男性で自らの試験結果を秘匿したいと思っていた者がいたとしても、その男性のプライバシーは保護されない。なぜなら、先の問合せの結果、すべての男性が落第したことが暴露されているためである。

そこで、$Q(D)$ をそのまま出力するのではなく、プライバシーを保護するための何らかの加工を行ったうえで出力することを考える。このプライバシー保護のための加工を $\mathcal{M}$ とし、プライバシー保護のための加工が済んだ問合せ結果を $\mathcal{M}(D)$ と書く。このとき、プライバシー保護方式 $\mathcal{M}$ の安全性をどう評価できるだろうか。差分プライバシーの基本的なアイデアは、ある個人 1 名分のデータが含まれているかどうかだけが異なるふたつのデータセット $D_1 \in D$ と $D_2 \in D$ を考え、それらに対する問合せ結果 $\mathcal{M}(D_1)$ と $\mathcal{M}(D_2)$ の「差分」から何らかの情報を読み取れるか否かにより、$\mathcal{M}$ の安全性を評価しようとするものである。

具体例を用いて説明しよう。D_1 をある部署のある月の職員の給与が記録されたデータセットとする。翌月に新たに職員 1 名が雇用され、1 名分のデータが追加されたデータセットを D_2 とする。Q として、給与の平均値を問い合わせるクエリを考える。$\mathcal{M}$ は、データセット D_1 と D_2 に対して、確率的にランダムなノイズを付加したうえで、$Q(D_1)$ と $Q(D_2)$ を出力する手続きとする。

仮に、D_1 と D_2 に対して、クエリ Q を 100 回実行できるとすると、ノイズを含んだ平均給与額 $\mathcal{M}(D_1)$ と $\mathcal{M}(D_2)$ をそれぞれ 100 回観察することができるため、ふたつの分布の形状の違いから新たに雇用された職員の給与額を類推することができる。もし、ふたつの分布が完全に一致すれば、新たに雇用された職員の給与額についての手がかりが何もないため、その職員のプライバシーは完全に保護されていると考えることができる。もっとも、その状態は、データセット D_2 には、追加的に有用な情報が何もないことも同時に意味している。

ここで、新たに雇用された職員のプライバシー保護とデータセット D_2 の有用性の確保との間のトレードオフについて考える。差分プライバシーの画期的

なアイデアは、このトレードオフをひとつのパラメータ $\varepsilon > 0$ だけで表現しようとするものである。フォーマルには、ひとつの要素だけが異なるデータセットのペア (D_1, D_2) とすべての $R \subseteq Range(\mathcal{M})$ に対して、以下が満たされるとき、あるプライバシー保護方式 $\mathcal{M}$ は、差分プライバシーの意味で安全であるとされる。

$$\frac{\Pr(\mathcal{M}(D_1) \in R)}{\Pr(\mathcal{M}(D_2) \in R)} \leq \exp(\varepsilon)$$

直観的には、ε は、ふたつの確率分布 $\Pr(\mathcal{M}(D_1))$ と $\Pr(\mathcal{M}(D_2))$ の「ズレの大きさ」を表現している。ε が大きければ、ふたつの確率分布は大きくズレているため、新たに雇用された職員の給与額について何らかの情報を得ることができる。これは、データセットの有用性の確保を優先したものと解釈できる。逆に、ε を小さくするほどふたつの確率分布は区別できなくなるため、プライバシー保護を優先しようと思えば、ゼロに近い ε を設定すればよいということになる[5]。なお、ε は、プライバシー・ロスやプライバシー・バジェットと呼ばれている。

残念ながら、ε の大きさが 0.01 ならどの程度安全なのかを直観的に理解することはできない。重要なことは、差分プライバシーの考え方を使えば、プライバシーが保護されているか否かの二者択一でなく、プライバシー保護に関する定量的な指標が得られることである[6]。

2.2.2 効用関数の中の差分プライバシー

Ghosh and Roth（2011）は、差分プライバシーのパラメータ ε を用いて、プライバシーに関する消費者 i の効用を u_i を

$$u_i = p_i - v_i\varepsilon \tag{1}$$

と定式化している。ここで、p_i はプライバシーが侵害される場合の対価、v_i はプライバシーが侵害される場合のコスト（不効用）、ε はプライバシー・バジェットである。なお、p_i は、必ずしも金銭的なものとは限らない。利便性の高いアプリケーションやオンライン・サービスもこれに含まれると考えることが多い。

[5] この議論では、データセットを管理する主体は、ノイズが付加される前の個人情報データにアクセスできることが暗黙的に想定されている。もっとも、データセットの管理主体に対しても、自らの個人情報データを開示したくないと考える人々もおそらく存在する。こうした状況に対処するために、データセットの管理主体にさえ個人情報データを開示しないという意味でプライバシーをより厳格に保護することができる「局所差分プライバシー（local differential privacy）」と呼ばれる技術が提案されている（Kasiviswanathan et al. 2011; Duchi et al. 2013）。局所差分プライバシーについての詳細は、補論A節を参照。

[6] 差分プライバシーなどの定量指標を利用せず、たとえば、あるデータセットの中から住所と電話番号を除けば個人を特定できないだろう、と安易に考えることは危険である。コンピューター・サイエンスの分野では、どの情報を秘匿するかを直感で決めてしまったために、個人情報を再識別（re-identification）されてしまった有名な事案がいくつか知られている（Narayanan and Shmatikov 2008; Heffetz and Ligett 2014）。

2.3 ε を観察することの難しさ

データから ε を観察することは、ふたつの意味で難しい。ひとつは、ε を識別するのに十分な情報（(1) 式の u_i と p_i と v_i）を入手することが難しいという意味である。たとえば、地図上でナビゲーションを行うオンライン・サービスは、通常、利用者の現在地データと引き換えに提供される。サービスの提供が個人情報データの提供と同時に生じるため、そのサービスの利用の有無を把握できるデータが入手可能だとしても、そのサービスを利用しているのが、利用者の現在地データの提供に対する懸念が薄い（(1) 式の $v_i\varepsilon$ が小さい）からなのか、そのサービスの利便性を高く評価している（(1) 式の p_i が大きい）からなのかを識別することができない。このような困難にもかかわらず、2.3.1節でみるように、多くの研究がプライバシーに関する効用の特徴を明らかにしようとしてきた[7]。

もうひとつは、人々がアンケート調査などで回答する望ましいプライバシー保護の度合いと実際の行動の間に乖離があるという意味である。これは、コンピューター・サイエンスの分野では「プライバシー・パラドックス」と呼ばれている（Acquisti 2004; Barnes 2006）。2.3.2節では、このパラドックスを概観する。

2.3.1 プライバシーに関する効用の計測

Huberman et al.（2005）は、「年齢」と「体重」という個人情報データをいくらで提供するかというリバース・オークションを行い、(1) 式の p_i を観察しようとする。オークションの結果として、p_i に大きなばらつきがあることを報告している。Goldfarb and Tucker（2012a）は、(1) 式の $v_i\varepsilon$ を計測しようとする。具体的には、2001 年から 2008 年までの期間で、人々のプライバシー保護に対する懸念 $v_i\varepsilon$ が年々高まってきたこと、高齢層は若年層に比べて情報を開示しない傾向が強く（$v_i\varepsilon$ が大きく）、そのギャップが年々拡大していることを指摘している。

Kummer and Schulte（2019）は、2012 年から 2014 年にかけて Google Play ストア上で観察された、約 30 万件のスマートフォン・アプリに関するデータから、人々のプライバシーに関する効用を計測しようとする。まず、アプリ開発者がプライバシーに関するパーミッション情報を事前に選択するという Google Play ストア上の仕組みを利用して、プライバシー・センシティブなパーミッション情報がアプリに含まれるかどうかを判別する。そのうえで、プライバシー・アクセスに関する認可の有無が「プライバシー市場」の需要と供給に影響を与えるかを調べている。推計の結果として、プライバシー・センシティブな情報

[7] わが国の人々のプライバシー保護意識の特徴については、補論B節を参照。

へのアクセスを求めることで、需要側ではインストール数が 25% 減少し、供給側ではアプリ開発者が価格を有意に引下げることを示した。

Lin（2021）は、ε の観察を企図したものではないが、巧妙に設計された実験を行うことにより、プライバシーに関する効用を精緻に計測しようとしている。具体的には、Becker（1980）の効用モデルにもとづいて、プライバシーに関する効用を内在的（intrinsic）な部分と道具的（instrumental）な部分に分解したうえで、プライバシーの内在的な部分に関する価値評価は、人々の間でのばらつきが大きく、同時に、一部に極端に価値評価が高い人が存在する（分布の右側の裾が長い）ことを明らかにしている。

2.3.2 プライバシー・パラドックス

Athey et al.（2017）は、2014 年にマサチューセッツ工科大学で行われた社会実験のデータを使い、デジタル・プライバシー・パラドックスが観察されたことを報告している[8]。その実験では、被験者は、対照群と処置群にランダムに分けられ、被験者の友人のメールアドレスを実験者に伝えるように指示される。ただし、処置群の被験者にだけは、友人と無料でピザを食べるクーポンが与えられている。実験の結果、処置群では、友人と無料でピザを食べるクーポンという非常に僅かな対価にも関わらず、正しくないメールアドレスを伝える確率が 54% 低くなった。この結果は、事前に表明していたプライバシー保護に対する意識の違いを勘案したとしても変わらなかった。

Acquisti et al.（2013）は、ピッツバーグのショッピングモールでの実験結果から、プライバシー保護を獲得するためにいくら支払うか（Willingness To Pay, WTP）と既に手に入れているプライバシー保護を手放す場合にいくらを要求するか（Willingness To Accept, WTA）が異なることを明らかにした。これは、保有効果（endowment effect）と呼ばれる、行動経済学においてよく知られた現象で、人々が既に手に入れている物を高く評価するというバイアスである。Acquisti et al.（2013）は、プライバシーに関しては、WTA と WTP の比率が 5.47 と、通常の財の 2.92 に比べて、はるかに大きいことを指摘している[9]。

最近では、プライバシー・パラドックスを解明しようとする実証研究もみられている。Chen et al.（2021）は、Alipay ユーザーに対して、プライバシー意識に関するサーベイを実施し、その回答結果とユーザーの管理データ（administrative data）をマッチングしてプライバシー意識と個人情報データの提供行動の関係を分析している。分析の結果、ユーザーの属性を制御したとしても、プライバ

[8] 実験の詳細は、Catalini and Tucker（2016, 2017）を参照。
[9] プライバシーに関する WTA と WTP のギャップについては、Hui and Png（2006）によるサーベイも参照。

シー保護に関する懸念と個人情報データの提供行動との間には統計的に有意な関係がなく、むしろプライバシー保護に関する懸念が強いユーザーほどデジタル・サービスを積極的に利用していることを明らかにした。この逆説的な結果は、デジタル・サービスの積極的な利用経験から多くを学ぶこと（(1) 式の p_i が大きいこと）により、プライバシー保護に関する懸念が大きくなった（(1) 式の $v_i\varepsilon$ が大きくなった）ものと解釈されている。p_i と $v_i\varepsilon$ が相関しているのであれば、個人情報データの提供行動が $v_i\varepsilon$ と整合しないという「プライバシー・パラドックス」を説明することができる。なお、この解釈は、人々のプライバシー保護に対する懸念が先天的（innate）なものではなく、デジタル・サービスの利用を通じて徐々に形成されたもの（privacy as a developed preference）であることを示唆している。

2.4 ε を巡る政策的議論

差分プライバシーを用いると、様々な政策的議論が可能になる。たとえば、Abowd and Schmutte（2019）は、政府統計の正確性とプライバシー保護のトレードオフのなかで、社会的に望ましい ε の水準が定まるとしている。また、Ghosh and Roth（2011）や Hsu et al.（2014）は、より一般的な状況のもとで政策当局が直面するトレードオフについて、次のような議論をしている。

あるシステムを運営する政策当局が ε をゼロ近くの値に設定すると、人々のプライバシーが厳格に保護されるため、プライバシー保護を理由にそのシステムへの参加を躊躇する人はいない。ただし、システム内で収集されるデータにはプライバシー保護のために多くのノイズが含まれるため、そのデータの利用価値は低くなる。他方、データの利用価値を高めるために、当局が ε を大きくする（ノイズを少なくする）と、プライバシー保護意識の高い（(1) 式の v_i が相対的に大きい）人々がそのシステムから退出する可能性が高くなる。その場合、そのシステムには、プライバシー保護意識の低い人々が相対的に多くなる。よく知られているように、そうしたシステムにおいて収集されるデータには、サンプル選択バイアスが含まれる（Heckman 1979）。このように当局は、ε の決定に際して、(1) システムを利用するユーザー数、(2) ノイズの意味でのデータの利用価値、(3) バイアスの意味でのデータの利用価値という三つのうち、少なくともひとつを諦めなければならなくなる。

こうしたトレードオフのもとで、政策当局は、ε の水準を合理的に決定する市場メカニズムを設計できるだろうか。Ghosh and Roth（2011）は、それぞれの人々が望むプライバシー保護度合いを正直に表明させるような差分プライベート・メカニズムが「存在しない」ことを指摘している[10]。具体的には、プライバ

[10] ここでは、差分プライバシーは、ひとつの均衡概念（解の概念）として扱われている。すなわち、差分プライバシーは、ある個人のデータを含むデータセット D_1 とそれを含まないデータ

シーが保護されない場合の不効用（(1) 式の $v_i\varepsilon$）とその場合の対価（(1) 式の p_i）が相関する場合に、直接顕示原理（direct revelation mechanism）が働かないことを明らかにした。たとえば、自分が感染症にかかっていることをプライバシー情報だと考える人は、感染症にかかっているかどうかの情報をいくらで提供するかというリバース・オークションにおいて、高い入札価格を提示することを躊躇する。これは、高い入札価格自体が自らが感染症にかかっていることを暴露してしまうためである。このように、プライバシーに関しては、直接顕示原理によって ε の水準を合理的に決定することが難しくなる[11]。

また、Ichihashi（2020c）は、消費者とプラットフォーマーの間での動学ゲームの枠組みを用いて、プライバシー保護規制について議論している。消費者の限界的なプライバシーコストが逓減するという仮定のもとで、政策当局がより厳格なプライバシー保護規制を導入する（ε をゼロ近くに設定する）と、短期的には、消費者の厚生が改善するとともに、消費者の限界的なプライバシーコストが低下することでプラットフォーム上での活動水準が上昇する。消費者のプラットフォーム上での活動水準が上昇すると、より多くの個人情報データが創出され、消費者の限界的なプライバシーコストがさらに低下するため、長期的には、消費者のプラットフォーム上での活動水準がきわめて高くなり、消費者は自らのプライバシーを完全に失ってしまう。このことは、長期的にみると、厳格なプライバシー保護規制が所期の目的を達成できない可能性があることを示唆する[12]。

現時点で、筆者らの知る限り、ε の水準を低コストで合理的に決定する手段は存在しない[13]。Heffetz and Ligett（2014）が「the time seems ripe for more economists to join the conversation」としたように、今後もこの分野でより多くの知見が蓄積されていくことが期待される。

3　企業が支払うプライバシー保護のコスト

プライバシー保護は人々にベネフィットをもたらす一方で、個人情報データの利用が制約されることを通じてコストも生む。このコストは、直接的には、個

セット D_2 がほとんど同じ結果を返すことを保証するものと捉えられている。この均衡概念としての差分プライバシーという点は、McSherry and Talwar（2007）によって、比較的早い時期から指摘されている。

[11] Ghosh and Roth（2011）の帰結は、いくつかの条件を追加・変更することにより、改善できることも指摘されている（Ligett and Roth 2012）。

[12] Ichihashi（2020c）の議論のポイントは、プライバシー保護に関する事前と事後の規制では、その効果が異なりうるという点である。すなわち、事前のプライバシー保護規制が消費者の厚生を低下させる可能性がある一方、3.2節でみるように、消費者の「忘れられる権利」を保護するような事後的な規制は、消費者の厚生を高める可能性がある。

[13] コンピューター・サイエンスの分野では、ε は、0.01 から 10 までの範囲で設定されることが多いが、その範囲に実証的な根拠はない（Hsu et al. 2014）。

人情報データを保有・利用することによって収益化している企業が支払うことになる。差分プライベートなシステムでは、ε の水準を低く設定するほど（人々のプライバシー保護を厳格に行うほど）、企業が支払うコストは高くなる。現実のプライバシー保護規制において ε が設定されているわけではないが、企業が支払うコストは観察可能である。

企業は、3.1節でみるように、プライバシー保護のコストを基本的には支払うことになるが、3.2節と3.3節でみるように、常にコストを支払うわけではなく、ベネフィット（負のコスト）を得ることもある。

3.1 プライバシー保護規制のコスト

プライバシー保護の規制としては、EU における規制がよく知られている。Goldfarb and Tucker（2011）は、2001 年から 2008 年の間のオンライン・キャンペーンに関するデータを利用して、EU が 2002 年に定めた「プライバシーと電子コミュニケーション指令（Privacy and Electronic Communications Directive）」により、オンライン広告の効果が 65% 減少したことを実証的に示している。

最近では、2018 年 5 月に施行された、EU の GDPR が経済に及ぼした影響を調べる実証研究が蓄積されてきている。Goldberg et al.（2019）は、Adobe によって提供されたデータを利用して、GDPR の導入によって、EU に所在する企業のウェブサイトの閲覧が 9.7% 減少し、電子商取引のウェブサイトに限れば、閲覧が 4.2%、収入が 8.3%、それぞれ減少したことを明らかにしている。また、Jia et al.（forthcoming）は、2014 年 1 月から 2019 年 4 月までのベンチャー投資のデータを利用して、GDPR の導入によって EU のベンチャー企業への投資のディール数が 26.1% 減少したことを報告している[14]。

医療の世界では、プライバシー保護の制度設計が人の生死を左右するような影響をもたらす場合がある。Miller and Tucker（2009）は、米国の州ごとのプライバシー保護規制の違いを利用して、プライバシー保護規制が強いと電子医療記録（Electronic Medical Records, EMR）の導入が拡大しない傾向があることを示した。さらに、Miller and Tucker（2011）は、EMR の導入が拡大すれば、新生児の死亡率が有意に低下することを報告している。こうしたことから、Goldfarb and Tucker（2012b）は、プライバシー保護政策とイノベーション政策は不可分であることを強調している。

[14]ただし、論文のタイトルにもあるように、この結果が一時的なものである可能性には留意が必要である。

3.2　忘れられる権利を保護するコスト

2.3.2節でみたように、プライバシー保護に関する人々の行動はしばしば合理的でない。これを踏まえると、プライバシー保護において、過去の自らの意思決定を撤回する機会が存在することは重要である。この点、人々には「忘れられる権利（right to be forgotten）」があるという指摘がある（Rosen 2012）。実際、Ichihashi（2020c）は、消費者とプラットフォーマーの間での動学ゲームの枠組みを用いて、「忘れられる権利」が保護される場合に、消費者がこれを行使することで厚生が高まることを明らかにしている。

EU の GDPR では、データ主体が過去にした同意を撤回する権利（同意撤回権）が明示的に認められている。わが国では、同権利を明示的に保護する法令は存在していないものの、最高裁は、サーチエンジンによるプライバシー情報を含むウェブサイトの URL 情報の提供が法的な保護の対象となるか否かについて、一定の判断基準を示している（最決平成 29 年 1 月 31 日民集 71 巻 1 号 63 頁）。

個人情報データを利用する企業からみると、個人情報データは、長期間にわたって保有・蓄積されれば、より大きな付加価値を生み出す可能性がある。このことは、人々の「忘れられる権利」が何らかの仕組みによって保護される場合、個人情報データの保持期間が制約され、その結果として、個人情報データの利活用による便益の縮小というコストが生じる可能性があることを示唆する。

他方、Chiou and Tucker（2017）は、検索者の個人情報データの保持期間の変更がサーチエンジンの検索クオリティに与える影響を検証し[15]、統計的に有意な影響が確認できないことを報告している。このことは、サーチエンジンによる検索サービスにおいて過去の個人情報データはさほど重要でないこと、言い換えると、サーチエンジンによる検索サービスを提供する企業にとっては、人々の「忘れられる権利」を保護するコストが小さいことを示唆している。

Chiou and Tucker（2017）の結果は、日々新しい言葉が検索されるサーチエンジンに固有のものである可能性には留意が必要である。それでも、データの品質がアルゴリズムの精度に大きく影響することや時間の経過とともにデータのドメインが変化する可能性を踏まえると、古い時点のデータを捨てたとしても悪い結果につながるわけはないという指摘は、ある程度もっともらしいように思われる。

[15] ここでの検索クオリティは、サーチエンジンのユーザーが表示された検索結果に従ってウェブサイトを訪れるか、検索をやり直すかを計測した指標である。

3.3　プライバシー保護による「負の」コスト

プライバシー保護は、企業側に「負の」コスト、すなわちベネフィットをもたらすという興味深い実証研究がある。これらは、同じ精度の結果を得るためには、プライバシーを保護しないアンケート調査より保護する方が低コストで済むとした、Hsu et al.（2014）のエクササイズとも整合的である。

3.3.1　プライバシー・ポリシー変更による広告効果の向上

Tucker（2014）は、2010 年 5 月 28 日に実施された、Facebook のプライバシー・ポリシーの変更が広告効果に与えた影響を実証的に分析している。データは、米国の教育関連 NPO が Facebook 上で行った広告キャンペーンの結果として得られた、Facebook ユーザーのクリック・スルー・レートである。データ期間は、5 月 28 日を挟む 2.5 週間分で、この NPO の広告キャンペーン中に Facebook のプライバシー・ポリシーが変更されたことは偶然である。

Facebook のプライバシー・ポリシーは、2010 年 5 月に変更される以前は非常に複雑だとされており[16]、170 にも及ぶオプションを選択しないとプライバシーをコントロールすることができない仕様であったが、この時の変更により、すべてのプライバシー・コントロールがひとつに集約されたほか、第三者の個人情報へのアクセスをワン・クリックで拒否できるようになった。これは、Facebook ユーザーの個人情報データのコントロール権が大幅に強くなったものと解釈できる。

こうした変化は、広告効果を減少させることが事前に予想されたが、得られた結果は逆であった。プライバシー・ポリシー変更後、クリック・スルー・レートは、変更前の約 2 倍になった。このことは、消費者の交渉力を強めることが広告効果の向上というかたちで企業側にベネフィットをもたらすことを意味している。

3.3.2　GDPR の導入による企業側のベネフィット

Aridor et al.（2020）は、旅行関連プラットフォームのデータを利用して、GDPR 導入の影響を調べている。まず、GDPR の導入により、クッキーが 12.5% 減少した[17]。これは、GDPR の導入によってクッキーの共有を明示的に拒否す

[16] こうした傾向は、Facebook に限らない。Ramadorai et al.（2019）は、米国企業 4,078 社のプライバシー・ポリシーについて読みやすさの指標（Gunning Fog Index）を作成し、中央値レベルのプライバシー・ポリシーを理解するには、少なくとも大学卒業程度の教育レベルが必要であることを指摘している。

[17] クッキーとは、消費者がウェブサイトを閲覧する際に、ウェブ・サーバから消費者のウェブ・ブラウザに対して送信される、その消費者の情報を保存しておくためのテキスト・ファイルであ

る消費者が増加したためと考えられる。

しかし同時に、GDPR の導入後にクッキーの共有に明示的に同意した消費者については、驚くべきことに追跡可能性（trackability）が 8% 増加していた[18]。これについて、Aridor et al.（2020）は、これまでブラウザベースのクッキー・ブロック・ツールを使っていた消費者がクッキーの共有を明示的に拒否することでウェブ・サーバ側のデータから欠落したためにノイズが減少した結果ではないかとしている。

敷衍すると、ブラウザベースのクッキー・ブロック・ツールは、ウェブサイトを訪問するたびに新しいクッキーを再生成させるため、ウェブ・サーバ側では同一人物が複数のクッキーを有するデータが観察されることになり、消費者を特定した分析を行う際にはノイジーなデータとなってしまう。他方、GDPR のもとで消費者がクッキーの共有を明示的に拒否すれば、その消費者のクッキー情報はウェブ・サーバに送信されないため、ウェブ・サーバ側のデータにおいて同一人物が複数のクッキーを有するという意味でのノイズは減少することになる。Aridor et al.（2020）は、この興味深い帰結を GDPR による正の外部性と評価している。

4　個人情報データの負の外部性

2節と3節は、大きく括れば、いずれも差分プライバシーのパラメータ ε に関する議論である。他方、本節は、プライバシー保護方式 $\mathcal{M}$ に関する議論を扱う。具体的には、あるデータが十分安全に秘匿されていたとしても、他のデータからその秘匿されたデータを類推することができてしまう状況について考える。これは、個人情報データの負の外部性として、よく知られた問題である（4.1節）。プライバシーの経済学では、この負の外部性が社会に及ぼす深刻な影響についてコンセンサスが得られており（4.2節）、こうした状況下で有効なプライバシー保護方式も提案されている（4.3節）。

なお、ここでいう「負の」とは、人々の効用水準にマイナスの影響（不効用）を与えるという意味である。したがって、負の外部性とは逆に、人々の効用にプラスの影響を与える正の外部性も存在する。本稿の射程はプライバシー保護であるため、ここでは負の外部性のみを扱うが、プライバシーを含めた人々の効用全体への影響を考える場合には、Ichihashi（2020b）や Fainmesser et al.（2021）のように、正の外部性についても包括的に議論する必要がある。

る。ウェブ・ブラウザからのリクエストにクッキー情報が含まれることにより、ウェブ・サーバ側でのセッション管理が可能になる。

[18] ここでの追跡可能性は、あるウェブサイトおいて、一定期間に同じクッキーが何度観察されるかを指標化したものである。

4.1　個人情報データの負の外部性とは何か

個人情報データの負の外部性は、ある人が秘匿したデータが別の人が開示したデータから類推されてしまう状況において生じる。こうした状況が生じるのは、秘匿されたデータと開示されたデータとの間に相関関係があるためである。なお、この相関関係を簡単に扱うために、プライバシーの経済学では、人々の個人情報データを適当な確率変数として表現することが多い。

山本（2016）は、負の外部性が顕在化する事例として、次の仮想的な事例を挙げている。仮に、女性がうつ状態にあるときに化粧品を購入しやすいという因果関係が知られていたとする。そこで、化粧品メーカーは、複数の一般的な個人情報から、それぞれの個人がうつ状態にあるかどうかを予測し、今まさにうつ状態にあると予測された女性を特定して化粧品のターゲット広告を配信する[19]。

この事例のポイントは、人々が自らの健康状態を開示していないとしても、そうしたセンシティブな情報が一般的な個人情報から類推されうることである。プライバシーの経済学では、この個人情報データが有する負の外部性は、最も重要な論点であると考えられている。

4.2　負の外部性がプライバシー保護に及ぼす影響

プライバシーの経済学では、負の外部性があるもとで人々が合理的に行動するとき、プライバシーの侵害が必ず生じることが明らかにされている[20]。具体的には、負の外部性が存在する場合に、プラットフォーマーに提供される個人情報データが「過剰」になり、その対価が非常に安くなることが知られている（Choi et al. 2019; Acemoglu et al. forthcoming; Bergemann et al. 2020; Ichihashi 2020a, 2020b; Fainmesser et al. 2021）。こうした帰結は、一部プラットフォーマーに膨大な個人情報データが安価で提供されている現状を的確に説明している[21]。

メカニズムはシンプルである。人々は、自らが負の外部性の影響を受けていることを知っており、自らのプライバシー情報を秘匿したいと思ってもできない可能性があることを理解しているため、僅かな対価で自らの個人情報データを提供することが合理的になる。他方、もし、人々の効用水準を最大化する主

[19] Wei et al.（2020）は、実際に Twitter のターゲット広告にどのような属性が利用されているのかを調べている。それによると、Twitter のターゲット広告に最も利用される属性は、言語、年齢、場所であり、性別という属性が利用されることは少ない。

[20] プライバシーの侵害とは別の観点になるが、Ichihashi（2020a）は、負の外部性があるもとで人々が合理的に行動すると、消費者が直面する財の価格が引き上げられることで消費者のペイオフが悪化することを指摘している。

[21] 外部性とは異なるメカニズムながら、Ichihashi（2020c）は、消費者の限界プライバシーコストが逓減するという仮定のもとで、長期的には、消費者が個人情報データを過剰に提供して自らのプライバシーを完全に失ってしまう可能性があることを明らかにしている（2.4節の議論も参照）。

体（ソーシャル・プランナー）が存在するのであれば、それぞれの人々のプライバシーが侵害されない程度に個人情報データの提供が抑制されることになる。このように、人々が別々に合理的な判断をすると、ソーシャル・プランナーが人々の個人情報データの提供量を決める場合と比べて、提供される個人情報データが「過剰」になる。結果として、社会全体で「過剰」にプライバシーの侵害が生じる[22]。

この負の外部性がもたらす帰結は、人々が自らの求めるプライバシー保護度合いを正直に表明できなくなることに他ならない。人々は、本来であれば、より強いプライバシー保護を望んでいるにも関わらず、それを選択しないことが合理的になるような状況に置かれてしまう。

Choi et al.（2019）が強調するように、こうした状況は、人々への教育や啓蒙などでは解決できない。人々は、プライバシーの侵害を予想できていないわけではなく、それを予想したうえで個人情報データを提供することが合理的になっているのである。

また、プラットフォーマー間の競争促進も、こうした状況を必ずしも改善しない。Choi et al.（2019）は、競争があったとしても同じ状況が生じると主張し、Acemoglu et al.（forthcoming）は、プラットフォーマー間での競争によって状況がかえって悪化する可能性があることを明らかにしている。Acemoglu et al.（forthcoming）は、場合によっては、個人情報データ市場をシャットダウンする方が経済全体にとって望ましいことさえあるとしている。

さらに、やや逆説的だが、こうした状況は、わが国の個人情報保護法によっても対処できない可能性がある。なぜなら、推論（プロファイリング）による要配慮個人情報の生成が要配慮個人情報の「取得」に該当するかは解釈問題とされているためである（宇賀 2018）。該当しないとの解釈によった場合には、プラットフォーマーが要配慮個人情報に該当しない個人情報をオプトアウト方式により第三者に提供し、その第三者が推論を行ったとしても、個人情報保護法には違反していないと考えられる。

4.3　負の外部性に対処するプライバシー保護方式

4.2節でみたように、負の外部性が存在するもとでは、個人情報保護法が守られていたとしても、プライバシーが過剰に侵害されるという意味での「非効率性」が発生する可能性がある。この非効率性を改善できるのであれば、公的な

[22] この負の外部性がもたらす帰結は、2.3.2節で紹介した「プライバシー・パラドックス」を説明できる可能性がある（Bergemann et al. 2020）。すなわち、負の外部性があるもとでは、それぞれの人々の個人情報データの価値が非常に低くなるため、友人と無料でピザを食べるクーポンといった、非常に僅かな対価でも個人情報データを手放すことが合理的になると考えられる。

介入は正当化されうる。ここでは、負の外部性に対処することで非効率性を改善する、いくつかのプライバシー保護方式を紹介する。

第一は、外部性を「内部化」するような「パーソナライズされたピグー税」である（Acemoglu et al. forthcoming）。負の外部性があるもとでデータが過剰に提供されるという非効率性が発生するのは、それぞれの人々が外部性のコストを全く負担しないことが原因である。したがって、人々の個人情報データ間の相関構造に応じて、税金を負担させればよいということになる。別の人との相関が強い人は、相対的に多くの税金を負担することで、データ提供のインセンティブをそがれることになり、経済全体では、ソーシャル・プランナーが存在する場合と同じ効率的な状況が実現できる[23]。もっとも、「パーソナライズされたピグー税」は、さすがに非現実的である。たとえば、1,000 万人の利用者を有するプラットフォームにおいて、個人情報データの相関行列にもとづいて最適な税負担額を随時計算することは、およそ現実的なスキームとは思われない。

第二は、「価格差別なしのオプト・イン同意規制」である（Choi et al. 2019）。オプト・イン同意規制とは、データ提供の際、消費者が明示的な同意を事前に行うことを求めるものである。EU の GDPR では、オプト・イン同意規制が課されており、消費者に提示される同意のチェックボックスに予めチェックが入っている状態は認められていない。Choi et al.（2019）は、社会的に望ましい水準を上回ってデータを収集する際にオプト・インを求めるような規制によって状態が改善しうるとする。これは、望ましい水準を上回らないように、オプト・インによるコストを設けるというものであり、本質的には、第一の「パーソナライズされたピグー税」と同じ発想といえる。

第三は、「相関除去（de-correlation）メカニズム」である（Acemoglu et al. forthcoming; Ichihashi 2020b）。これは、非効率性の原因となっている、個人情報データ間の相関構造を消してしまうという発想である。具体的には、信頼できる第三者が一旦すべての個人情報データを収集し、個人情報データ間の相関をすべてゼロになるようにプラットフォーマーに開示するデータと開示しないデータを選択するというものである。このスキームは、経済全体の余剰を必ず改善する。

相関除去メカニズムのひとつの実装事例として、「Randomized Aggregatable Privacy-Preserving Ordinal Response（RAPPOR）」と呼ばれるアルゴリズムを挙げることができる（Erlingsson et al. 2014）。RAPPOR は、「Chromium」と呼ばれるオープンソースのウェブ・ブラウザ開発プロジェクトにおいて開発された技術で[24]、局所差分プライバシーを満足する、ふたつのプライバシー保護メカ

[23] Fainmesser et al.（2021）は、プラットフォーマーに対して収集したデータ量に応じた税金を課すことで、「パーソナライズされたピグー税」と同じ状況を達成できるとしている。

[24] Google が開発しているウェブ・ブラウザ「Chrome」は、このプロジェクトで開発されたソースコードを利用している。

ニズムからなる。ふたつの匿名化ステップを設けるのは、直観的にいえば、データが有する相関構造を利用した攻撃に対処するためである。これは、現時点では特定の相関構造に限られてはいるが、データの有する外部性に対処しようするものである。

5 まとめ

本稿では、プライバシーの経済学と呼ばれる分野のサーベイを行った。プライバシーの経済学は、インターネット空間における個人情報の取扱いに対する関心がグローバルに高まるなか、近年急速に発展してきている。そこで蓄積されている共通認識は、デジタル決済システムを利用する人々に安心感を与えつつデータの利活用をどう進めていくかを考える際に、重要な示唆を与えうるものである。それらをまとめると、以下のとおりである。

- プライバシーとは、人それぞれに異なるものである
- プライバシー保護の度合いは、差分プライバシーによって表現できる
- 人々が求めるプライバシー保護度合いを推定・観察することは困難（「プライバシー・パラドックス」）
- プライバシー保護は、企業側のコストを伴う
- ただし、消費者のプライバシー保護に取組むことで、企業側にベネフィットが生じることもある
- 社会的に望ましいプライバシー保護の水準の決定と、個人情報データが有する負の外部性への対処を市場メカニズムによって実現することはできない
- 社会的に望ましいプライバシー保護の水準を決めることは難題
- 個人情報データが有する負の外部性に完全に対処することもまた難題

参考文献

宇賀克也 (2018)「個人情報保護法の逐条解説（第 6 版）」有斐閣.

総務省 (2017)「安心・安全なデータ流通・利活用に関する調査研究」https://www.soumu.go.jp/johotsusintokei/linkdata/h29_02_houkoku.pdf

総務省 (2020)「データ流通環境等に関する消費者の意識に関する調査研究」https://www.soumu.go.jp/johotsusintokei/linkdata/r02_04_houkoku.pdf

寺田雅之 (2019)「差分プライバシーとは何か」システム／制御／情報 63 巻 2 号, 58-63.

山本龍彦 (2016)「ビッグデータ社会とプロファイリング」論究ジュリスト 18 号.

Abowd, John M. and Ian M. Schmutte (2019) "An Economic Analysis of Privacy Protection and Statistical Accuracy as Social Choices," *American Economic Review* 109(1), 171–202.

Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asuman Ozdaglar (forthcoming) "Too Much Data: Prices and Inefficiencies in Data Markets," *American Economic Journal: Microeconomics*.

Acquisti, Alessandro (2004) "Privacy in Electronic Commerce and the Economics of Immediate Gratification," *Proceedings of the 5th ACM Conference on Electronic Commerce*, 21–29.

Acquisti, Alessandro, Leslie K. John, and George Loewenstein (2013) "What is Privacy Worth?," *Journal of Legal Studies* 42(2), 249-274.

Acquisti, Alessandro, Curtis Taylor, and Lian Wagman (2016) "The Economics of Privacy," *Journal of Economic Literature* 54(2), 442-492.

Allcott, Hunt and Matthew Gentzkow (2017) "Social Media and Fake News in the 2016 Election," *Journal of Economic Perspectives* 31(2), 211-36.

Aridor, Guy, Yeon-Koo Che, and Tobias Salz (2020) "The Effect of Privacy Regulation on the Data Industry: Empirical Evidence from GDPR," Available at https://www.ftc.gov/system/files/documents/public_events/1548288/privacycon-2020-guy_aridor.pdf

Arrieta-Ibarra, Imanol, Leonard Goff, Diego Jiménez-Hernández, Jaron Lanier, and E. Glen Weyl (2018) "Should We Treat Data as Labor? Moving beyond 'Free'," *AEA Papers and Proceedings* 108, 38-42.

Athey, Susan, Christian Catalini, and Catherine E. Tucker (2017) "The Digital Privacy Paradox: Small Money, Small Costs, Small Talk," NBER Working Paper.

Barnes, Susan B. (2006) "A Privacy Paradox: Social Networking in the United States," *First Monday* 11(9).

Becker, Gary. S. (1980) "Privacy and Malfeasance: A Comment," *Journal of Legal Studies* 9(4), 823–826.

Bergemann, Dirk, Alessandro Bonatti, and Tan Gan (2020) "The Economics of Social Data," *arXiv*.

Catalini, Christian and Catherine E. Tucker (2016) "Seeding the S-Curve? The Role of Early Adopters in Diffusion," NBER Working Paper.

Catalini, Christian and Catherine E. Tucker (2017) "When Early Adopters Don't Adopt," *Science* 357(6347), 135-136.

Chen, Long, Yadong Huang, Shumiao Ouyang, and Wei Xiong (2021) "The Data Privacy Paradox and Digital Demand," NBER Working Paper.

Chiou, Lesley and Catherine E. Tucker (2017) "Search Engines and Data Retention: Implications for Privacy and Antitrust," NBER Working Paper.

Choi, Jay Pil, Doh-Shin Jeon, and Byung-Cheol Kim (2019) "Privacy and Personal Data Collection with Information Externalities," *Journal of Public Economics*, 173, 113-124.

Duchi, John C., Michael I. Jordan, and Martin J. Wainwright (2013) "Local Privacy and Statistical Minimax Rates," *Proceedings of the 2013 IEEE 54th Annual Symposium on Foundations of Computer Science (FOCS)*, 429-438.

Dwork, Cynthia (2006) "Differential Privacy," *Proceedings of the International Colloquium on Automata, Languages and Programming (ICALP)*, 1-12.

Dwork, Cynthia, Frank McSherry, Kobbi Nissim, and Adam Smith (2006) "Calibrating Noise to Sensitivity in Private Data Analysis," *Proceedings of the Third Theory of Cryptography Conference TCC, volume 3876 of Lecture Notes in Computer Science*, 265-284.

Englehardt, Steven and Arvind Narayanan (2016) "Online Tracking: A 1-million-site Measurement and Analysis," *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 1388–1401.

Erlingsson, Ulfar, Vasyl Pihur, and Aleksandra Korolova (2014) "RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response," *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security - CCS'14*, 1054–1067.

European Central Bank (2021) "Eurosystem Report on the Public Consultation on a Digital Euro," Available at https://www.ecb.europa.eu/paym/digital_euro/html/pubcon.en.html

Fainmesser, Itay Perah and Andrea Galeotti, and Ruslan Momot (2021) "Digital Privacy," HEC Paris Research Paper. Available at https://ssrn.com/abstract=3459274

Financial Stability Board (2019) "BigTech in Finance," Available at https://www.fsb.org/2019/12/bigtech-in-finance-market-developments-and-potential-financial-stability-implications/

Ghosh, Arpita and Aaron Roth (2011) "Selling Privacy at Auction," *Proceedings of the 12th ACM Conference on Electronic Commerce*, 199–208.

Goldberg, Samuel, Garrett Johnson, and Scott Shriver (2019) "Regulating Privacy Online: The Early Impact of the GDPR on European Web Traffic & E-Commerce Outcomes," NBER Working Paper.

Goldfarb, Avi and Catherine E. Tucker (2011) "Privacy Regulation and Online Advertising," *Marketing Science* 57(1), 57-71.

Goldfarb, Avi and Catherine E. Tucker (2012a) "Shifts in Privacy Concerns," *American Economic Review* 102(3), 349–353.

Goldfarb, Avi and Catherine E. Tucker (2012b) "Privacy and Innovation," *Innovation Policy and the Economy* 12, 65-90.

Guardian (2018) "How Cambridge Analytica Turned Facebook 'Likes' into a Lucrative Political Tool," Available at https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm

Heckman, James, J. (1979) "Sample Selection Bias as a Specification Error," *Econometrica* 47(1), 153–161.

Heffetz, Ori and Katrina Ligett (2014) "Privacy and Data-Based Research," *Journal of Economic Perspectives* 28(2), 75-98.

Hsu, Justin, Marco Gaboardi, Andreas Haeberlen, Sanjeev Khanna, Arjun Narayan, Benjamin C. Pierce, and Aaron Roth (2014) "Differential

Privacy: An Economic Method for Choosing Epsilon," *Proceedings of 27th IEEE Computer Security Foundations Symposium*, 398-410.

Huberman, Bernardo A., Eytan Adar, and Leslie Fine (2005) "Valuating Privacy," *IEEE Security and Privacy* 3(5), 22–25.

Hui, Kai-Lung and Ivan Paak Liang Png (2006) "The Economics of Privacy," In *Handbooks in Information Systems: Volume 1: Economics and Information Systems*, edited by Terrence Hendershott, 471-498. Bingley, UK: Emerald.

Ichihashi, Shota (2020a) "Online Privacy and Information Disclosure by Consumers," *American Economic Review* 110(2), 569-595.

Ichihashi, Shota (2020b) "The Economics of Data Externalities," Available at https://shota2.github.io/research/externality.pdf

Ichihashi, Shota (2020c) "Dynamic Privacy Choices," Available at https://shota2.github.io/research/dynamicPrivacy.pdf

Jia, Jian, Ginger Zhe Jin, and Liad Wagman (forthcoming) "The Short-Run Effects of GDPR on Technology Venture Investment," *Marketing Science*.

Jones, Charles I. and Christopher Tonetti (2020) "Nonrivalry and the Economics of Data," *American Economic Review* 110(9), 2819-2858.

Kasiviswanathan, Shiva Prasad, Homin K. Lee, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith (2011) "What Can We Learn Privately?," *SIAM Journal on Computing* 40(3), 793-826.

Konečný, Jakub, H. Brendan McMahan, Daniel Ramage, and Peter Richtárik (2016) "Federated Optimazation: Distributed Machine Learning for On-Device Intelligence," *arXiv*.

Konečný, Jakub, H. Brendan McMahan, Felix X. Yu, Ananda Theertha Suresh, Dave Bacon, and Peter Richtárik (2017) "Federated Learning: Strategies for Improving Communication Efficiency," *arXiv*.

Kummer, Michael and Patrick Schulte (2019) "When Private Information Settles the Bill: Money and Privacy in Google's Market for Smartphone Applications," *Management Science* 65(8), 3470-3494.

Lanier, Jaron (2013) *Who Owns the Future?*, New York: Simon & Schuster.

Ligett, Katrina and Aaron Roth (2012) "Take It or Leave It: Running a Survey When Privacy Comes at a Cost," *Proceedings of the 8th International Conference on Internet and Network Economics*, 378–391.

Lin, Tesary (2021) "Valuing Intrinsic and Instrumental Preferences for Privacy," Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3406412

McMahan, H. Brendan, Eider Moore, Daniel Ramage, and Blaise Agüera y Arcas (2016) "Federated Learning of Deep Networks using Model Averaging," *arXiv*.

McSherry, Frank and Kunal Talwar (2007) "Mechanism Design via Differential Privacy," *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, 94-103.

Miller, Amalia R. and Catherine E. Tucker (2009) "Privacy Protection and Technology Diffusion: The Case of Electronic Medical Records," *Management Science* 55(7), 1077-1093.

Miller, Amalia R. and Catherine E. Tucker (2011) "Can Health Care Information Technology Save Babies?," *Journal of Political Economy* 119(2), 289-324.

Narayanan, Arvind and Vitaly Shmatikov (2008) "Robust De-anonymization of Large Sparse Datasets," *Proceedings of the 2008 IEEE Symposium on Security and Privacy*, 111–125.

New York Times (2018) "How Trump Consultants Exploited the Facebook Data of Millions," Available at https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html

Pai, Mallesh M. and Aaron Roth (2013) "Privacy and Mechanism Design," *ACM SIGecom Exchanges* 12(1), 8-29.

Posner, Eric A. and E. Glen Weyl (2018) *Radical Markets: Uprooting Capitalism and Democracy for a Just Society*, Princeton, NJ: Princeton University Press.

Posner, Richard A. (1978) "The Right of Privacy," *Georgia Law Review* 12(3), 393-422.

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

Ramadorai, Tarun, Antoine Uettwiller, and Ansgar Walther (2019) "The Market for Data Privacy," CEPR Discussion Paper.

Rosen, Jeffrey (2012) "The Right to be Forgotten," *Stanford Law Review Online* 64, 88.

Stigler, George J. (1980) "An Introduction to Privacy in Economics and Politics," *Journal of Legal Studies* 9(4), 623–44.

Tucker, Catherine E. (2014) "Social Networks, Personalized Advertising, and Privacy Controls," *Journal of Marketing Research* 51(5), 546-562.

Varian, Hal R. (1996) "Economic Aspects of Personal Privacy," In *Privacy and Self-Regulation in the Information Age*, Washington, DC: US Department of Commerce, National Telecommunications and Information Administration.

Wei, Miranda, Madison Stamos, Sophie Veys, Nathan Reitinger, Justin Goodman, Margot Herman, Dorota Filipczuk, Ben Weinshel, and Michelle L. Mazurek (2020) "What Twitter Knows: Characterizing Ad Targeting Practices, User Perceptions, and Ad Explanations Through Users' Own Twitter Data," *Proceedings of the 29th USENIX Security Symposium* 145-162.

Xiong, Xingxing, Shubo Liu, Dan Li, Zhaohui Cai, and Xiaoguang Niu (2020) "A Comprehensive Survey on Local Differential Privacy," *Security and Communication Networks*.

A　局所差分プライバシー

局所差分プライバシー（local differential privacy）は、人々がそれぞれの端末において、プライバシー保護のためのノイズを付加する加工を行ったうえで、その加工済みのデータをサーバに送ってデータセットを構築する技術である。局所差分プライベートなシステムのもとでは、人々の個人情報データは、それぞれの「ローカルな」端末内にしか存在しない[25]。サーバ内のデータセットには、プライバシー保護のためのノイズが付加された後のデータしか格納されていないため、データセットの管理主体は、「真の」個人情報データを観察することができない。言うまでもなく、そのデータセットから個人情報データが漏洩することもあり得ない。

局所差分プライバシーが実装されている事例として、4.3節でも紹介した、RAPPOR を挙げることができる。RAPPOR は、それぞれの端末のブラウザ側でノイズを付加する加工を行ったうえで、その加工済みデータをサーバに送信するアルゴリズムである。RAPPOR 以外にも、近年、Apple、Microsoft、Samsung といった企業が開発するソフトウェアにおいて、局所差分プライバシーを満足するアルゴリズムが実装されている（Xiong et al. 2020）。

デジタル決済システムでは、決済処理を行うサーバにおいてノイズが付加される前のデータが必要になるため、局所差分プライバシーの技術を全面的に適用することは難しいと思われる。それでも、決済処理に必要な個人情報データとその他の個人情報データを区別できるのであれば、後者に対して、局所差分プライバシーを満たすアルゴリズムを適用することは技術的には可能である。

B　わが国の人々のプライバシー保護意識

総務省（2020）は、日本、米国、ドイツ、中国の 4 か国において、それぞれ 1,000 名を対象としたウェブ上でのアンケート調査を行っている。そのアンケート調査によれば、わが国は、企業等に個人情報データを提供することに不安を感じる人の割合が 78.2% と、4 か国中で最も高い。過去の同様の調査でも、同計数は 84.1% と、その 4 か国の中では最も高かった（総務省 2017）。

また、わが国は、個人情報データの提供を判断するうえで、そのサービスや

[25] 機密性の高いデータをローカルな端末に保持しつつ、サーバ上で何らかの集計や学習を行うという発想は、McMahan et al.（2016）や Konečný et al.（2016, 2017）らによって提案された、「協調学習（federated learning）」と呼ばれる技術にも通底する。協調学習は、学習するモデルを共有したうえで、それぞれの端末でデータを保持したまま別々に学習を行い、学習によって得られたパラメータをサーバに送信することでパラメータの更新を行っていく手法である。機械学習の分野では、協調学習は、プライバシー保護に関する重要な貢献をする可能性が高いとみなされている。

アプリによるメリットを重視する人の割合が 57.3% と最も低い（最も高いのは中国の 92.5%）。これと整合的に、個人情報データ・ストアや情報銀行を利用したいとする人の割合も 34.7% と、わが国が最低である（最も高いのは中国の 79.0%）。ちなみに、信用スコアリング・サービスについて「抵抗なし」とした人の割合は、わが国と米国・ドイツが 3 割程度となっている一方、中国は 72.3% となっている。

さらに、国によって定められるプライバシーやデータ保護に関する規制やルールについて、わが国は、「安心・安全性を重視」する人の割合が 78.5%（他方、「便利・快適性を重視」する人の割合が 21.5%）と、4 か国中で最も高い（最も低いのは中国の 50.9%）。

ただし、このアンケートは、それぞれの国に居住する者を対象とするため、実際に経験している利便性・快適性の程度や享受しているサービスやアプリのメリットが異なった人々を比較していることには留意が必要である。たとえば、中国以外の国の人々は、実際に信用スコアリング・サービスを利用したことがない一方、中国の人々は実際に同サービスを利用した結果として「抵抗なし」と回答している。

2.3.2節で紹介した、Acquisti et al.（2013）の議論を踏まえると、わが国の人々のプライバシー保護意識が相対的に保守的にみえるのは、既に手に入れているプライバシー保護を高く評価していることの裏返しであるとも考えられる。逆にいえば、何らかの利便性と引き換えに既に手に入れているプライバシー保護を手放す経験をすると、わが国の人々のプライバシー保護意識が異なるかたちで観察される可能性もある。また、プライバシー・パラドックスを考慮すると、こうしたアンケート結果が実際の行動と整合しないことも考えられる。いずれにしても、わが国の人々のプライバシー保護意識について、このアンケート調査だけでは評価できないと思われる。