Sara Krehbiel*

Choosing Epsilon for Privacy as a Service

Abstract: In many real world scenarios, terms of service allow a producer of a service to collect data from its users. Producers value data but often only compensate users for their data indirectly with reduced prices for the service. This work considers how a producer (data analyst) may offer differential privacy as a premium service for its users (data subjects), where the degree of privacy offered may itself depend on the user data. Along the way, it strengthens prior negative results for privacy markets to the pay-for-privacy setting and develops a new notion of endogenous differential privacy. A positive result for endogenous privacy is given in the form of a class of mechanisms for privacy-as-a-service markets that 1) determine ϵ using the privacy and accuracy preferences of a heterogeneous body of data subjects and a single analyst, 2) collect and distribute payments for the chosen level of privacy, and 3) privately analyze the database. These mechanisms are endogenously differentially private with respect to data subjects' privacy preferences as well as their private data, they directly elicit data subjects' true preferences, and they determine a level of privacy that is efficient given all parties' preferences.

Keywords: differential privacy, data-dependent privacy, mechanism design, markets for privacy

DOI 10.2478/popets-2019-0011
Received 2018-05-31; revised 2018-09-15; accepted 2018-09-16.

1 Introduction

As consumers have become more privacy conscious, producers of goods and services who collect masses of data as a matter of course have begun to advertise data privacy as a feature. Like many services, privacy has a cost to the provider because it requires restricting access to otherwise profitable data. This cost is visible in several instances of creative pricing schemes. For example, Progressive's Snapshot program allows a customer to save on car insurance premiums by volunteering to plug a device into their car that sends Progressive information about their driving behavior [19]. When AT&T first offered gigabit internet service in Austin, the base rate was \$70/month, but for \$99/month customers could opt out of targeted advertising based on collected browsing history [2]. In both cases, data subjects paid the data collector a base rate for a primary service and a premium to keep their data private.

The choice between data access for producers and privacy for users need not be binary. The field of differential privacy offers methods to ensure against leaking individuals' private data while permitting meaningful statistical analysis. A data access or sanitization mechanism is said to be ϵ-*differentially private* if no row of the database has too much effect probabilistically on the output of the mechanism, where privacy parameter $\epsilon > 0$ quantifies the magnitude of this effect. Varying ϵ corresponds to smoothly transitioning from perfect privacy (no data access) to no privacy (full data access). There is by now a rich body of research that establishes how to conduct a wide variety of statistical analysis goals while maintaining differential privacy, but much of this literature is agnostic to the choice of ϵ.

Different applications intuitively require different levels of privacy; medical records, for example, may require a higher standard of privacy than Netflix preferences. But what are the analytical attributes of an optimal compromise between privacy and data access, and how can it be implemented in the presence of conflicting interests between data subjects and analysts? A database curator might set ϵ according to some expert discretion about the relative needs for privacy versus meaningful data analysis, but she may not have good information about the privacy preferences of the data subjects or the value of accurate data for analysts. Indeed, we would expect data subjects and analysts to inflate their stated respective needs for privacy and accurate data if there is no downside to doing so. A non-optimal value of ϵ may be economically inefficient in that an analyst may be willing to pay data subjects to accept a weaker privacy guarantee in order to improve the accuracy of a statistical analysis, or data subjects may be willing to pay more for stronger privacy.

Existing frameworks monetizing privacy fall short of solving the market problem of finding an optimal level of privacy in a manner that itself preserves privacy. Mechanism design and computational market equilib-

***Corresponding Author: Sara Krehbiel:** University of Richmond, E-mail: email@email.edu

rium techniques have been applied to similar settings in which multiple parties' competing preferences jointly determine socially efficient pricing and allocation of goods. This work helps bridge the gap between database privacy and market economics by 1) formalizing a new privacy definition for mechanisms that provide input-dependent privacy guarantees, including those that simulate privacy markets by making monetary transfers reflecting the value of the data and/or privacy guarantees, and 2) adapting an existing mechanism for efficient allocation of public goods to this privacy setting.

1.1 Related Works

The problem of quantifying the optimal tradeoff between privacy and accuracy is considered in [1, 13]. The setting in [1] is motivated by a statistical agency such as the U.S. Census Bureau charged with the responsibility to publicly release statistical information about the population. Data collectors and data subjects are not modeled as separate entities with opposing preferences, but rather data subjects' utility functions reflect preference for accuracy as well as privacy. They model privacy as a public good and propose that the best choice of ϵ is that which maximizes society's aggregate utility for privacy and accuracy subject to the constraints imposed by the differential privacy technology in use. In [13], the authors consider the choice of ϵ in terms of the probability of specific events in concrete applications.

A number of works explore the ramifications of modeling the monetary cost of privacy loss to individual data subjects. A primary focus of many of these works is *incentive compatibility*, also referred to as *truthfulness*, which is the property that each data subject maximizes his utility by revealing his true private data. In [22], authors provides a generic transformation of any truthful mechanism into one that is differentially private, but truthfulness breaks if privacy itself influences utility. Several other papers provide mechanisms that are both truthful and private [3, 15, 17]. In their mechanisms, data subject utility depends on the output of the mechanism as well as privacy, so the mechanism itself may provide value to the data subjects.

Other works assume that private data has already been collected, so mechanisms in this setting needn't elicit truthful data, but they must elicit truthful privacy preferences. Because data subjects' private data and privacy preferences may be correlated, the privacy of both must be protected. To capture the setting of data collectors soliciting information by paying would-be data subjects, [14] and [8] design survey mechanisms that incentivize enough voluntarily participation to guarantee accurate differentially private analysis. The survey mechanism of [14] causes some data subjects to experience a net negative utility, i.e., the mechanism is not *individually rational*, while [8] rely on knowledge of a prior distribution of privacy preferences for each private data type. Dropping the assumption of a known prior distribution, [18] develop an individually rational mechanism that truthfully elicits privacy preferences directly, and they guarantee a relaxed notion of privacy.

In all of the above mechanisms that consider the preferences of individual data subjects, ϵ must be chosen exogenously, before looking at these preferences or the private data itself. In contrast, [10] propose mechanisms that solicit these privacy preferences, determine an appropriate value of ϵ, charge the analyst some payment in exchange for a noisy statistic on the data, and distribute this payment among data subjects to compensate for their ϵ loss of privacy. Their mechanisms privately estimate a counting query while also guaranteeing incentive compatibility and individual rationality. This result is extended in [5], who provide a mechanism estimating a more general linear predictor. An alternate model is proposed in [9], in which an analyst proposes some differentially private computation designed so the privacy parameter ϵ decreases with the number of data subjects that voluntarily opt in.

A weakness of the insensitive value model of [10] and the opt-in model of [9] is that while differential privacy is guaranteed at any level output by the mechanism, privacy is with respect to the private data only. This means that if individuals' data are correlated with their preferences or opt-out decisions, the mechanism may indirectly leak information about private data. To address this concern, [10] also propose a stronger *sensitive value model* that requires differential privacy with respect to privacy preferences as well as private data. However, they prove that no mechanism can simultaneously offer this privacy guarantee, accuracy of the published statistic, and several desirable economic properties. This result and new extensions are discussed more fully in Section 3.

1.2 Overview

The first major contribution of this work is a new definition of privacy for mechanisms that select privacy parameters as a function of data, whose necessity is justified in part by new negative results strengthening those

of [10]. We argue that because the standard definition of differential privacy is parameterized by a fixed ϵ, it cannot enforce a meaningful relationship between the privacy guarantee of a mechanism whose internal privacy-preserving behavior is a function of the data and the inputs themselves. For example, [10] require the outputs of their mechanisms on any two neighboring databases to be ϵ close for the smallest ϵ the mechanism may output on any input database; we call this ϵ the minimum privacy selection of the mechanism. This requirement compels such mechanisms to effectively always choose the minimal value of ϵ, ignoring the inputs:

Definition 1.1 (Privacy in [10], informal)**.** A mechanism $\mathcal{M}$ that internally selects $\epsilon \in \mathbb{R}^+$ as a function of its inputs $\mathbf{d} \in \mathcal{D}^n$ and publicly outputs some $R \in \mathcal{R}$ is *differentially private* if for all neighboring $\mathbf{d}, \mathbf{d}' \in \mathcal{D}^n$ and event $E \subseteq \mathcal{R}$, we have

$$\Pr[\mathcal{M}(d) \in E] \leq \exp(\epsilon) \cdot \Pr[\mathcal{M}(\mathbf{d}') \in E],$$

where ϵ is the minimum privacy selection of $\mathcal{M}$ on any input.

The negative result of [10] states that no mechanism guaranteeing privacy as above can simultaneously offer non-trivial accuracy and individual rationality while charging the analyst any finite quantity to compensate data subjects for their privacy loss. This result is presented formally in Section 3. The intuition is that if a data subject can have arbitrarily high cost associated with privacy loss, then an individually rational mechanism must be prepared to charge the analyst an arbitrarily high payment for any fixed ϵ. Maintaining differential privacy with respect to these privacy preferences means this high lower bound on the analyst payment must hold even when privacy preferences are moderate.

In Section 3 we provide new negative results in the pay-for-privacy setting in which an analyst is a producer that already has access to user data and wishes to charge a premium for privacy. These new negative results suggest that while forcing users to pay for privacy rather than forcing producers to pay for data makes it easier to simultaneously realize individual rationality and incentive compatibility, we need a new data-dependent endogenous privacy definition to more fundamentally circumvents the impossibility result of [10].

The proposed definition of *endogenous differential privacy* is a strict relaxation, still requiring that the mechanism's output on a particular database is close to the output on a neighboring database, but only by an amount determined by the ϵ chosen for that database:

Definition 1.2 (Endogenous privacy, informal)**.** A mechanism $\mathcal{M}$ that internally selects $\epsilon, \delta \in \mathbb{R}^+$ as a function of its inputs $\mathbf{d} \in \mathcal{D}^n$ and publicly outputs some $R \in \mathcal{R}$ is *endogenously differentially private* if for all neighboring $\mathbf{d}, \mathbf{d}' \in \mathcal{D}^n$, event $E \subseteq \mathcal{R}$, and ϵ, δ the privacy selection of $\mathcal{M}(\mathbf{d})$, we have

$$\Pr[\mathcal{M}(\mathbf{d}) \in E] \geq \exp(-\epsilon) \cdot \Pr[\mathcal{M}(\mathbf{d}') \in E] + \delta,$$
$$\Pr[\mathcal{M}(\mathbf{d}) \in E] \leq \exp(\epsilon) \cdot \Pr[\mathcal{M}(\mathbf{d}') \in E] + \delta.$$

Both inequalities are necessary since ϵ is specific to $\mathbf{d}$, breaking the symmetry of $\mathbf{d}$ and $\mathbf{d}'$. The formalization of this definition (Definition 3.6) includes syntactic differences for added generality and applicability to the market setting of interest (Model 1), but here we first offer a concrete example of a simple endogenously private mechanism without any market considerations. Suppose a database consists of one bit b_i per person $i \in [n]$, and we wish approximate the count of 1-bits with an increasingly strong privacy guarantee as this count increases. For example, b_i could indicate whether i cares about privacy, and we want to estimate how many people care about privacy while respecting the aggregate preference. We can do this by adding $\text{Lap}(\sqrt{\sum b_i + 1})$ noise, producing an approximation of $\sum b_i$ that has error at most $\sqrt{\sum b_i + 1} \ln 1/\beta$ with all but probability β and guaranteeing endogenous differential privacy for $\epsilon = 1/\sqrt{\sum b_i + 1}$ and $\delta = \exp(-2\sqrt{\sum b_i})$. In the extreme case that no one cares about privacy, these bounds correspond to $\ln 1/\beta$ error and no guarantee of privacy ($\epsilon = 1, \delta = 1$). This result follows from the structure of the proof of Lemma 4.3, which pertains to the more general Model 1 and so involves more parameters that are not made explicit here for simplicity. We note that the $\delta > 0$ is unavoidable because two neighboring inputs receive noise from distributions with differing variance.

By adding data-dependent noise, this example appears similar to mechanisms that add noise proportional to local sensitivity, which are susceptible to attack when the local sensitivity of neighboring databases may differ dramatically [16]. In the above example and the class of mechanisms presented in Section 3, the data-dependent noise is varying with the choice of ϵ, which depends on the input data smoothly in that it has finite global sensitivity. In this way, mechanisms can add different amounts of noise as ϵ varies (smoothly).

After formally introducing the new notion of endogenous privacy, we present a class of endogenously private privacy-as-a-service market mechanisms. This class of mechanisms that receive the private data, so-

licit privacy and accuracy preferences from the data subjects and analyst, choose a level of privacy consistent with these preferences, make monetary transfers, and get a noisy statistic by running a standard differentially private mechanism on the data at the previously determined level of privacy. These mechanisms captures the idea of privacy as a premium service, assuming that users have already divulged their private data to an entity that they can expect will try to profit from it, and they are willing to pay more for privacy versus a baseline of unrestricted data access. Viewing privacy as a good, it is easy to design mechanisms that do not overcharge data subjects for privacy, ensuring individual rationality. The new challenge is to construct a payment scheme that discourages data subjects from understating their individual preferences for privacy, letting others pay for the privacy enjoyed by all. The cumulative works of [4, 11, 12, 21] provide an elegant solution to this "free-rider problem" as it exists more generally in neoclassical economies, achieving the market goal of a *Pareto efficient* (see Definition A.4) level of production while incentivizing consumers to report their true preferences.

The class of mechanisms presented in Section 4 adapts this solution to the free-rider problem to our privacy market framework. A mechanism $\mathcal{M}$ in our class computes the level of privacy that maximizes the value of privacy to data subjects less the cost of accuracy loss to the analyst, it charges data subjects individual payments that align individual utility with social utility, and it pays the analyst noisy compensation for her associated accuracy loss. It then runs some standard differentially private computation to approximate the desired query on the database at the market-determined level of privacy. We prove that our mechanisms are endogenously differentially private, incentive compatible, individually rational when the sum of data subjects' values for privacy is not too small relative to analyst cost, Pareto efficient for appropriate choices of parameters, and collect non-negative revenue in expectation.

2 Preliminaries

Let $\mathbb{R}$ denote the set of real numbers and $\mathbb{R}^+$ denote the nonnegative reals. For families of functions $\mathcal{F}$ that are isomorphic to the reals, we write $\mathcal{F} \equiv \mathbb{R}$. In this case, we often let a single letter denote both the function and the associated real, e.g., for the family of constant functions, we may write $c(x) = c$. Let $x^+ = \max(0, x)$ for $x \in \mathbb{R}$. Let $\sum$ or $\sum_i$ denote $\sum_{i\in[n]}$, and $\sum_{j\neq i}$ denotes $\sum_{j\in[n]\setminus\{i\}}$ when n is clear. For any n-dimensional vector $\mathbf{v}$, let v_i denote the ith entry, and denote the vector without v_i as $\mathbf{v}_{-i}$. Let $\mathbf{v}_{-i}\|v_i'$ denote $\mathbf{v}$ with v_i replaced with v_i'. For n-dimension vectors $\mathbf{v}, \mathbf{v}'$, let $\mathbf{v} \sim \mathbf{v}'$ denote that they differ on only one row, and we call such vectors neighboring.

The following is the standard definition of (approximate) differential privacy [7] with respect to neighboring databases differing in at most one row:

Definition 2.1. A mechanism $\mathcal{M} : \mathcal{D}^n \to \mathcal{R}$ is (ϵ, δ)-differentially private if for all neighboring databases $\mathbf{d}, \mathbf{d}' \in \mathcal{D}^n$ differing on one row and $S \subseteq \mathcal{R}$,

$$\Pr[\mathcal{M}(\mathbf{d}) \in S] \leq \exp(\epsilon) \cdot \Pr[\mathcal{M}(\mathbf{d}') \in S] + \delta.$$

For any $b \in \mathbb{R}^+$, let $\mathrm{Lap}(b)$ denote the real random variable with pdf $p(x) = \frac{1}{2b}\exp(-|x|/b)$. For any query $f : \mathcal{D}^n \to \mathbb{R}$ with sensitivity $\max_{\mathbf{d}\sim\mathbf{d}'}|f(\mathbf{d}) - f(\mathbf{d}')| \leq \Delta$, the Laplace mechanism [7] is an $(\epsilon, 0)$-differentially approximation of f computed as $f(\mathbf{d}) + \mathrm{Lap}(\Delta/\epsilon)$.

3 Lower Bounds for Positive Value Privacy Markets

The question of whether a pay-for-privacy market can simultaneously satisfy privacy, accuracy, and desired economic properties is in large part motivated by the negative results for the pay-for-data sensitive value model of [10]. In this section, we provide additional negative results in the pay-for-privacy setting when their worst-case (non-endogenous) privacy is required. These results motivate a new notion of endogenous privacy and suggest that this definition and not pay-for-privacy is the core change that allows us to circumvent the impossibility result of [10] in Section 4.

3.1 A General Model for Privacy Markets

We start by describing a generalization of the setting considered in [10]; this generalized setting is formalized in Model 1. A mechanism $\mathcal{M}$ in the generalized model internally chooses some privacy policy $q \in \mathcal{Q}$, possibly as a function of its inputs, and it publishes a statistic $R \in \mathcal{R}$ summarizing the data. By allowing $\mathcal{Q}$ to be general, q may determine the amount of noise reflected in R, as in the example in the overview and the mechanisms in Section 4, or it may be a vector of ϵ privacy

requirements for each data subject, as in the results in this section. Each data subject $i \in [n]$ has a true privacy preference $v_i^* \in \mathcal{V} \subseteq \{\mathcal{Q} \to \mathbb{R}\}$, where $v_i^*(q)$ indicates the utility (positive or negative) realized by i when the mechanism enacts privacy policy q. The mechanism receives reported privacy preferences $\mathbf{v} \in \mathcal{V}^n$ and verifiable data $\mathbf{d} \in \mathcal{D}^n$ from the data subjects as well as a function $c \in \mathcal{C} \subseteq \{\mathcal{Q} \to \mathbb{R}\}$ from the analyst, which restricts the allowed monetary transfers. In [10], we can think of $\mathcal{C}$ as being the set of constant functions imposing a budget constraint; in Section 4, we will let $\mathcal{C}$ represent inaccuracy cost functions, so that $c(q)$ represents the analyst's loss in profit due to the expected inaccuracy from privacy policy q. After selecting privacy policy q, the mechanism collects payments $\mathbf{p} \in \mathbb{R}^n$ from the data subjects and pays the analyst $P \in \mathbb{R}$. In [10], the p_i and P are negative because the analyst must pay for the data; in Section 4, they are positive.

Model 1 Privacy Market

1: Upon initialization, there exists a mechanism $\mathcal{M} : \mathcal{D}^n \times \mathcal{V}^n \times \mathcal{C} \to \mathcal{Q} \times \mathbb{R}^n \times \mathbb{R} \times \mathcal{R}$ for general types $\mathcal{D}, \mathcal{R}, \mathcal{Q}$ and $\mathcal{V}, \mathcal{C} \subseteq \{\mathcal{Q} \to \mathbb{R}\}$. Each data subject $i \in [n]$ has verifiable data $d_i \in \mathcal{D}$ and some true privacy preference $v_i^* \in \mathcal{V}$, and an analyst has a cost function $c \in \mathcal{C}$.
2: $\mathcal{M}$ receives the verifiable data, and data subjects and the analyst report their preferences to $\mathcal{M}$.
3: $\mathcal{M}$ internally selects endogenous privacy parameter $q \in \mathcal{Q}$.
4: $\mathcal{M}$ makes transfers $\mathbf{p} \in \mathbb{R}^n$ and $P \in \mathbb{R}$ from and to the data subjects and analyst, resp.
5: $\mathcal{M}$ publishes statistic $R \in \mathcal{R}$.
6: Each data subject $i \in [n]$ realizes utility $v_i^*(q) - p_i$.

Mechanisms in this model are expected to simultaneously satisfy privacy with respect to reported preferences $\mathbf{v}$ and verifiable data $\mathbf{d}$ and accuracy of R with respect to some desired statistical goal. It is important that privacy must be with respect to both these inputs in the event that they are related, or even identical, as in the example in the introduction. In addition, mechanisms should satisfy the market properties of incentive compatibility (a data subject cannot benefit from misreporting v_i^*), individual rationality (a data subject gains non-negative utility from the mechanism), balanced budget (the mechanism collects as much money as it pays out), and Pareto efficiency (no other choice of q will benefit one data subject without harming another, subject to the payment constraints imposed by c). These properties are formally defined in the context of this model in Appendix A.

For the rest of this section, we consider $\mathcal{D} = \{0, 1\}$, $\mathcal{R} = \mathbb{R}$ and the task of private bit approximation as in [10]. We also consider $\mathcal{Q} = (\mathbb{R}^+)^n$ and interpret a choice of q as a vector of privacy parameters ϵ_i for $i \in [n]$. Privacy and accuracy as in [10] using the notation of Model 1 are formalized as follows.

Definition 3.1 (Accuracy of bit approximation [10])**.** A mechanism $\mathcal{M} : \{0,1\}^n \times \mathcal{V}^n \times \mathcal{C} \to (\mathbb{R}^+)^n \times \mathbb{R}^n \times \mathbb{R} \times \mathcal{R}$ is called α-accurate if for any $\mathbf{v} \in \mathcal{V}^n, \mathbf{d} \in \{0,1\}^n, c \in \mathcal{C}$, we have $\Pr[|\mathcal{M}_R(\mathbf{d}, \mathbf{v}, c) - \sum d_i| > \alpha n] \leq 1/3$, where $\mathcal{M}_R(\mathbf{d}, \mathbf{v}, c)$ denotes the random variable describing the statistic published by $\mathcal{M}$ on the specified inputs.

Definition 3.2 (Privacy in [10])**.** For any $\boldsymbol{\epsilon} \in (\mathbb{R}^+)^n$, a mechanism $\mathcal{M}$ is $\boldsymbol{\epsilon}$-differentially private if for any $i \in [n]$, $c \in \mathcal{C}$, neighboring $(\mathbf{d}, \mathbf{v}) \sim (\mathbf{d}', \mathbf{v}')$ differing on row i, and $E \subseteq \mathbb{R} \times \mathcal{R}$,

$$\Pr[\mathcal{M}_{R,P}(\mathbf{d}, \mathbf{v}, c) \in E] \leq \exp(\epsilon_i) \cdot \Pr[\mathcal{M}_{R,P}(\mathbf{d}', \mathbf{v}', c) \in E],$$

where $\mathcal{M}_{R,P}(\mathbf{d}, \mathbf{v}, c)$ denotes the joint random variable of the published outputs (R, P) of $\mathcal{M}$ on specified inputs.

For their mechanisms choosing $\boldsymbol{\epsilon} \in \mathcal{Q}$, [10] require $\boldsymbol{\epsilon}$-differential privacy for $\boldsymbol{\epsilon}$ consisting of the entrywise smallest values of ϵ_i the mechanism can choose on any inputs (possibly with $c \in \mathcal{C}$ fixed).

3.2 New Lower Bounds

The main negative result of [10] (Theorem 3.3) is that a mechanism cannot be non-trivially accurate, individually rational, and budget-balanced if privacy of data subjects' privacy valuations as well as private data is required in the above sense for any ϵ in the privacy support of the mechanism.[1] The crux of this result is that individual rationality means analyst payment *must* be arbitrarily large for some databases. Since the payment is small with probability zero on such a database, privacy then precludes a small payment for any database.

1 Note that all the lower bounds in this section hold even for non-truthful mechanisms, i.e., mechanisms that do not incentivize data subjects to report their true privacy valuation as required by Definition A.1.

We note that for ϵ chosen independent of the input data, any standard moneyless differentially private mechanism is trivially budget balanced and individually rational if its privacy provides data subjects with nonnegative utility, and such a mechanism is accurate and ϵ-differentially private for this single exogenous value of ϵ. Although this technically circumvents the impossibility result, assumption of nonnegative privacy utility alone does not make any real progress towards selecting a sensible value of ϵ endogenously without also modifying the definition of privacy. On the contrary, we show that such trivial mechanisms are essentially the *only* mechanisms possible in this positive-value scenario (Theorem 3.4). If we allow mechanisms to run a deficit at most half the time, we get a slightly weaker but still unsatisfying negative result in this model (Theorem 3.5). Ultimately, our endogenous privacy definition (along with valuing privacy positively and permitting the mechanism to sometimes run a deficit) allows our framework to admit meaningful privacy markets such as those in Section 4, circumventing the key negative result of [10].

In proving their impossibility result, [10] assume for simplicity that $\mathcal{V} \equiv \mathbb{R}^+$ with $v_i^*(\boldsymbol{\epsilon}) = -v_i^* \cdot \epsilon_i$ and show that nontrivial accuracy and privacy together require $\sum \epsilon_i \geq \ln(4/3)$. They could more generally assume that the privacy support of the mechanism is nontrivial, i.e., it contains some nonzero $\boldsymbol{\epsilon} \in (\mathbb{R}^+)^n$, and that the privacy cost function family is unbounded, i.e. for any payment $-P > 0$ from the analyst and nonzero $\boldsymbol{\epsilon} \in (\mathbb{R}^+)^n$, there exists some $\mathbf{v} \in (\mathbb{R}^+)^n$ with $\sum v_i \cdot \epsilon_i > -P$. We present the proof for their result in this more general case for comparison to the new lower bounds in our framework. Our new results for positive value markets analogously concern mechanisms in Model 1 with arbitrarily small value to the data subjects, i.e., $\mathcal{V}$ such that for any payment $P > 0$ to the analyst, there exists a $\mathbf{v} \in \mathcal{V}^n$ with $\sum v_i(\epsilon_i) < P$ for any $q = \boldsymbol{\epsilon}$ in the privacy support of $\mathcal{M}$ on $\mathbf{v}$.

Theorem 3.3 (Theorem 5.1 [10])**.** Any mechanism $\mathcal{M}$ in Model 1 with nontrivial privacy support and unbounded cost function family that is ϵ_i-differentially private for each $i \in [n]$ for every $\boldsymbol{\epsilon}$ it outputs, individually rational, and budget-balanced can never charge the analyst any finite payment.

Proof summary. Assume for contradiction that there exists some private, individually rational, and budget-balanced mechanism $\mathcal{M}$ that charges the analyst some finite $-P'$ on some fixed $\mathbf{v}'$ and other inputs with positive probability. Consider $\mathbf{v}$ with $\sum v_i \cdot \epsilon_i > -P'$ for some nontrivial $\boldsymbol{\epsilon}$ in the privacy support of $\mathcal{M}$. By individual rationality and the balanced budget assumption, $\mathcal{M}$ running on $\mathbf{v}$ must charge the analyst more than $-P'$, i.e., $\mathcal{M}$ charges at most $-P'$ with probability zero. Then by privacy, the probability that $\mathcal{M}$ on $\mathbf{v}'$ outputs $-P'$ is also zero, contradicting the initial assumption. □

When privacy has positive value to data subjects, individual rationality is easy to achieve with these other properties in mechanisms that make no monetary transfers. Clearly any moneyless (trivially budget-balanced) standard differentially private mechanism with some fixed privacy parameter (that it privately outputs) satisfies differential privacy and individual rationality when the mechanism has non-negative utility for the data subjects. The following theorem shows that these are the *only* positive-value approximation mechanisms that satisfy all these properties.

Theorem 3.4. Any mechanism $\mathcal{M}$ in Model 1 with $\mathcal{Q} = (\mathbb{R}^+)^n$ and arbitrarily small value to the data subjects that is ϵ_i-differentially private for each $i \in [n]$ for every $q = \boldsymbol{\epsilon}$ it outputs, individually rational, and budget-balanced can never make any positive payment to the analyst.

Proof. Fix any $P > 0$ and let $\mathbf{v}$ be a set of valuations such that $\sum v_i(\epsilon_i) < P$ for $\boldsymbol{\epsilon}$ in the privacy support of $\mathcal{M}$ on $\mathbf{v}$. Then by individual rationality and the balanced budget assumption, $\mathcal{M}$ must pay the analyst less than P when running on $\mathbf{v}$, and by privacy, the probability that $\mathcal{M}$ pays the analyst at least P when running on any inputs is also zero. □

Underlying the result of Theorem 3.4 is the zero-probability event that a strictly budget-balanced mechanism pays the analyst more than it can charge data subjects for some fixed inputs and corresponding privacy guarantees. By permitting the mechanism to *sometimes* lose money by paying the analyst more than the individual rationality-mandated maximum that the data subjects can be charged, arbitrarily small $\sum v_i(\epsilon_i)$ no longer forces an arbitrarily small upper bound on the noisy P. However, if we make only the weak assumption that for any fixed inputs, the mechanism loses money at most half the time, the standard definition of differential privacy still imposes a strict upper bound on the probability of the analyst receiving any positive payment.

Theorem 3.5. Let $\mathcal{M}$ be any mechanism in Model 1 with $\mathcal{Q} = (\mathbb{R}^+)^n$ and arbitrarily small value to the data subjects that is ϵ_i-differentially private for each $i \in [n]$

for every $q = \boldsymbol{\epsilon}$ it outputs, individually rational, and satisfies $\Pr[P > \sum p_i] \leq 1/2$ for any inputs. Let $\epsilon_{i,\inf}$ denote the infimum of $q_i = \epsilon_i$ in the support of the mechanism. Then for any fixed $\bar{P} > 0$ and any inputs, $\mathcal{M}$ pays the analyst more than $\bar{P}$ with probability at most $\exp(\sum \epsilon_{i,\inf})/2$.

Proof. Fix any $\bar{P} > 0$ and let $\mathbf{v}$ be a set of valuations such that $\sum v_i(\epsilon_i) < \bar{P}$ for $\boldsymbol{\epsilon}$ in the privacy support of $\mathcal{M}$ on $\mathbf{v}$. By individual rationality, we have $\sum p_i \leq \sum v_i(\epsilon_i)$, so we must have $\Pr[P > \bar{P}] \leq 1/2$. Then by privacy, we have that for any inputs, $\Pr[P > \bar{P}] \leq \exp(\sum \epsilon_{i,\inf})/2$. □

In other words, any mechanism capable of making strong privacy guarantees must pay the analyst nothing almost half the time *if the standard notion of input-independent differential privacy is used.*

3.3 Endogenous Privacy

In addition to yielding the negative results of the previous subsection, the strength of the privacy requirement of [10] counterintuitively breaks the relationship between the privacy guarantee output by a mechanism and a data subject's utility for that run of the mechanism. Since outputs of *any* neighboring inputs are guaranteed to be ϵ-close for the smallest possible value of ϵ, it does not make sense that a data subject would value privacy as a function of the ϵ chosen on a particular run of a mechanism. The usual definition of differential privacy captures the idea that an individual cares about the difference between two output distributions: that of the mechanism run on the true database, and that of the mechanism run on the same database with his row changed. We argue that privacy for all ϵ in the support of the mechanism goes far beyond the true concerns of a data subject. We propose the following new privacy definition, which endogenizes the privacy guarantee by requiring that the output distribution of the mechanism on any set of reference inputs is close to the output distribution of the mechanism on any neighboring set of inputs, where closeness is determined *only by the ϵ supported by mechanism running on the reference inputs.* Using the syntax of Model 1, we define privacy with respect to fixed functions $\epsilon, \delta : \mathcal{Q} \to \mathbb{R}^+$ that map the mechanism's internally chosen privacy level q to its provable privacy guarantee as follows:

Definition 3.6 (Endogenous differential privacy)**.** For fixed $\epsilon, \delta : \mathcal{Q} \to \mathbb{R}^+$, a mechanism $\mathcal{M}$ is (ϵ, δ)-endogenously differentially private if for all neighboring $(\mathbf{d}, \mathbf{v}) \sim (\mathbf{d}', \mathbf{v}')$, $c \in \mathcal{C}$, q in the privacy support of $\mathcal{M}(\mathbf{d}, \mathbf{v}, c)$, and $E \subseteq \mathcal{R} \times \mathbb{R}$,

$$\Pr[\mathcal{M}(\mathbf{d}, \mathbf{v}, c) \in E] \leq e^{\epsilon(q)} \cdot \Pr[\mathcal{M}(\mathbf{d}', \mathbf{v}', c) \in E] + \delta(q),$$
$$\Pr[\mathcal{M}(\mathbf{d}', \mathbf{v}', c) \in E] \leq e^{\epsilon(q)} \cdot \Pr[\mathcal{M}(\mathbf{d}, \mathbf{v}, c) \in E] + \delta(q).$$

To see how this definition circumvents the negative results of the previous section, we consider the probability bound in Theorem 3.5 showing that any mechanism capable of making strong privacy guarantees has limited ability to collect significant rents for the analyst. This bound arises from collapsing the bounds in probabilities of the event that the mechanism pays the analyst $P > \bar{P}$ on neighboring pairs of databases in the chain of databases $\mathbf{v}^{(0)} = \mathbf{v}', \ldots, \mathbf{v}^{(j)} = (v_1, \ldots, v_j, v'_{j+1}, \ldots, v'_n), \ldots, \mathbf{v}^{(n)} = \mathbf{v}$ where $\mathbf{v}$ is such that $\sum v_i(\epsilon_i) < \bar{P}$ and $\mathbf{v}'$ is arbitrary.

With (ϵ, δ)-endogenous privacy for ϵ the identity function and δ the zero function, the reference databases and not $\{\epsilon_{i,\inf}\}$ determine the probability differences across neighboring databases. Let $\boldsymbol{\epsilon}^{(j)}$ denote the entry-wise minimum privacy parameters in the support of $\mathcal{M}(\mathbf{v}^{(j)})$. Then endogenous privacy yields the bound:

$$\begin{aligned}\Pr[P > \bar{P} \mid \mathbf{v}'] &\leq \exp(\epsilon_1^{(1)}) \Pr[P > \bar{P} \mid \mathbf{v}^{(1)}] \\ &\leq \cdots \\ &\leq \exp(\sum \epsilon_i^{(i)}) \Pr[P > \bar{P} \mid \mathbf{v}] \\ &\leq \exp(\sum \epsilon_i^{(i)})/2.\end{aligned}$$

For $\mathbf{v}$ small enough for $\sum v_i(\epsilon_i) < \bar{P}$, we expect the $\epsilon_i^{(i)}$ for large i to be large, making this bound[2] loose if not trivial, and suggesting that the definition of endogenous privacy permits mechanisms exhibiting reasonable behavior.

The mechanisms providing positive results in Section 4 use $\mathcal{Q} = \mathbb{R}^+$ and offer the same privacy guarantee to each data subject. The privacy policy q selected internally corresponds to the quality of the privacy guarantee, because noise proportional to q is added to the publicly released quantities R and P. In the Laplace mechanism, noise magnitude directly implies some level of ϵ-differential privacy, but we present the definition of endogenous privacy for general $\mathcal{Q}$ because in some settings it may be useful to choose among a more general set of privacy policies which data subjects may value in more

2 We also get a similar bound for every permutation π on $[n]$ identifying a different chain of hybrid databases between $\mathbf{v}'$ and $\mathbf{v}$, but in all of these cases, the $\epsilon_{\pi(i)}^{(i)}$ for large i should be large.

subtle ways. For example, [3, 15, 18] argue that ϵ alone can only provide an *upper bound* on the information leaked by a mechanism. This is because output distributions of a mechanism on neighboring databases may only be ϵ apart for an extremely unlikely set of events and closer otherwise, or the analyzed upper bound on ϵ may itself be loose. By allowing q to be of a general form, future mechanisms in this framework can potentially release more specific information about their privacy policies that may allow tighter analysis of privacy loss. Furthermore, data subjects' utilities for a mechanism selecting a general privacy parameter need not be limited to the mechanism's privacy properties. For example, q may include the publicly released analysis R, allowing the new framework to model outcome-dependent utility [3, 15, 17, 18], which depends on differential privacy guarantees and public outputs.

We also note that the choice of q in Model 1 does not itself mandate a privacy requirement. In the same spirit of the parametrized notion of differential privacy, the strength of an endogenous privacy guarantee is parametrized through carefully chosen functions ϵ, δ connecting the privacy policy to the privacy guarantee, permitting a meaningful relationship between the utility realized by the data subjects and the data-dependent private behavior of the mechanism.

Finally we remark that although the syntax of Definition 3.6 is consistent with that or Model 1, the definition does not rely on this specific framework for privacy markets or even access to privacy preferences.

4 Endogenous Privacy Markets

In this section, we present a class of mechanisms in our framework satisfying the previously discussed properties. Our mechanisms will discourage users from overstating their privacy preferences by charging premiums that increase with the amount of privacy demanded. The challenge is that users may try to avoid higher individual premiums by understating their preference for privacy, letting others pay for the privacy enjoyed by all. This "free-rider problem" is solved in a much more general public goods setting in the cumulative works of [4, 11, 12, 21].

In the setting studied by [4, 11, 12], consumers communicate to some central body, called the *government*, their valuation $v_i(\cdot)$ of a certain public good. The government chooses the level of public good that optimizes social utility, and it levies taxes designed to align individual consumers' utilities with social utility in order to avoid free-riding [4]. Specifically, the government pays a producer $c(q)$ to produce $q \geq 0$ units of the good for the level q maximizing consumer surplus, $\sum v_i(q) - c(q)$. Each consumer i receives utility $v_i(q)$ for the public good and is charged the amount he diminishes others' surplus: $c(q) - \sum_{j \neq i} v_j(q) + \max_{q_{-i}}(\sum_{j \neq i} v_j(q_{-i}) - \frac{n-1}{n} c(q_{-i}))$. With these allocation and tax rules, consumers are incentivized to communicate their true preferences, and sufficient funds are raised to produce a Pareto efficient level of the public good. (See Appendix B.)

4.1 A Class of Pareto Efficient Privacy Markets

Our class of mechanisms assumes the statistical analysis goal is characterized by some query $f : \mathcal{D}^n \to \mathcal{R}$, and that there exists some differentially private $\mathcal{M}_f : \mathcal{D}^n \to \mathcal{R}$ that is (ϵ_f, δ_f)-differentially private in (in the standard sense) when instantiated on any ϵ_f, δ_f in some legal set, including arbitrarily small ϵ_f. We fix $\mathcal{Q}, \mathcal{C}, \mathcal{V}$ as follows:

- $\mathcal{Q} = \mathbb{R}^+$, with larger values implying more noise and (because the ϵ, δ parametrizing the privacy guarantee are decreasing functions of q) stronger privacy.
- $\mathcal{C}$ is the set of functions $c(q) = cq$ for $c \in \mathbb{R}^+$, consistent with [4, 12].
- $\mathcal{V}$ is the set of functions $v_i(q) = v_i \ln(q + 1)$ for $v_i \in \mathbb{R}^+$, chosen for arithmetic convenience and capturing the notion of diminishing marginal returns; Appendix D considers other $\mathcal{V}$.

The below mechanism also fixes a truncation parameter $\Delta \in \mathbb{R}^+$ and function $h : \mathbb{R}^+ \to \mathbb{R}^+$ specifying how much noise to add for privacy. The following results hold for arbitrary Δ, h, but privacy is meaningful for choices specified by Lemma 4.4. For concreteness, consider

$$\Delta = \ln n \quad h(q) = \sqrt{q + \Delta}.$$

Mechanism 2 first truncates the reported privacy valuations depending on Δ and then computes the Pareto efficient privacy level q and incentive-compatible charges to the data subjects, using the allocation and tax rules from [12]. The analyst payment P includes noise that is a function of q, and the statistic R is computed by $\mathcal{M}_f$ running with privacy parameters ϵ_f, δ_f, where ϵ_f is some other function of q and δ_f is a privacy parameter required by $\mathcal{M}_f$, possibly 0 depending on the mechanism.

Mechanism 2 Privacy Market

Inputs: Database $\mathbf{d} \in \mathcal{D}^n$, privacy valuations $\mathbf{v} \in (\mathbb{R}^+)^n$ and cost $c \in \mathbb{R}^+$.

1: $\bar{v}_i \leftarrow \min(v_i, c \cdot \Delta)$ for all $i \in [n]$.
2: $q \leftarrow (\frac{\sum_i \bar{v}_i}{c} - 1)^+$.
3: $p_i \leftarrow cq - \sum_{j \neq i} \bar{v}_j \ln(q+1)$
$+ \max_{q_{-i} \geq 0} \left(\sum_{j \neq i} \bar{v}_j \ln(q_{-i}+1) - \frac{n-1}{n} c q_{-i} \right)$
for all $i \in [n]$.
4: $P \leftarrow c(q + \gamma)$ with γ drawn from $\text{Lap}(h(q))$.
5: $R \leftarrow \mathcal{M}_f(\mathbf{d})$ with privacy parameters $\epsilon_f = \frac{\Delta}{h(q-\Delta)}$ and δ_f.
6: Publish R, collect p_i from each $i \in [n]$, and pay analyst P.

Theorem 4.1. Fix any $\mathcal{M}_f$ differentially private in the standard sense, $\Delta \in \mathbb{R}^+$, and increasing, differentiable, concave $h : \mathbb{R}^+ \to \mathbb{R}^+$ with $h'(0) \leq 1$. Then Mechanism 2 is (ϵ, δ)-endogenously differentially private for $\epsilon(q) = \frac{3\Delta}{h(q-\Delta)}$ and $\delta(q) = \delta_f + 1/\exp(1/h'(q-\Delta))$. Mechanism 2 is incentive compatible with respect to privacy valuations, individually rational when $c \leq \sum \bar{v}_i / e$, budget-balanced in expectation, and Pareto efficient when $v_i \leq c\Delta$.

This theorem follows from a series of lemmas we outline below. We show that truncation allows us to bound the sensitivity of the privacy policy q (Lemma 4.2) without harming incentive compatibility (Lemma C.1). Then we can establish endogenous privacy (Lemma 4.3) based on the noise added to the analyst payment and the privacy of $\mathcal{M}_f$. Individual rationality is guaranteed in all but extreme cases with very low privacy valuations relative to accuracy cost (Lemma C.2). Symmetric noise along with the fact that payments are structured so $\sum p_i \geq cq$ guarantees a balanced budget in expectation.

After providing proofs for these lemmas, we provide guidance for reasonable choices of parameters (Lemma 4.4). Good concrete choices to keep in mind are $\Delta = \ln n$ and $h(q) = \sqrt{q + \ln n}$. With these choices and values of $v_i, c \in \Theta(1)$, Theorem 4.1 gives privacy for ϵ, δ decreasing functions of n, incentive compatibility, individual rationality, balanced budget in expectation with a negligible chance of a large deficit, and Pareto efficiency.

4.2 Sensitivity of Privacy Policy q

After fixing $c(q) = cq$ for $c, q \in \mathbb{R}^+$ following [4, 12], the main goal is to release $P \approx cq$ in an (ϵ, δ)-endogenously differentially private manner for some appropriate (positive and decreasing) functions $\epsilon(q)$ and $\delta(q)$. Classical differential privacy techniques would suggest first bounding the sensitivity of $\arg\max_{q \geq 0} \sum v_i(q) - cq$ by some Δ, and then $P = c(q + \text{Lap}(\Delta/\epsilon))$ is ϵ-differentially private for some fixed target ϵ. While our endogenous privacy parameter ϵ will be a function of q, we nonetheless first attempt to bound the sensitivity of q.

For consistency with the perspective of q as a public good, willingness-to-pay functions $v_i(q)$ should be nonnegative, increasing, and concave. With $c(q) = cq$, the unique consumer surplus-maximizing level of privacy will have $\sum \frac{d}{dq} v_i(q) = c$ unless $q = 0$. When $v_i(q) = v_i \ln(q+1)$, this optimal level is given by $q = (\sum v_i / c - 1)^+$ as in Step 2. By first truncating the v_i to $\bar{v}_i = \min(v_i, c \cdot \Delta)$, sensitivity of q is immediate from $|(\sum \bar{v}_i/c - 1)^+ - (\sum \bar{v}'_i/c - 1)^+| = |(v_i - v'_i)/c| \leq \Delta$:

Lemma 4.2 (Sensitivity of q)**.** For any $c \in \mathbb{R}^+$ and neighboring $\mathbf{v} \sim \mathbf{v}'$, we have:

$$|(\sum \bar{v}_i/c - 1)^+ - (\sum \bar{v}'_i/c - 1)^+| \leq \Delta$$

We may worry that bounding v_i will generate sample bias (see [10]). Indeed, if data subjects have negative utilities for privacy loss, a mechanism operating on truncated costs will not be able to adequately compensate data subjects, and if these privacy-sensitive data subjects are able to opt out of the mechanism, this may bias the data. Alternatively, we may worry that truncation might break truthfulness. However, Lemma C.1 shows that truncation preserves the incentive compatibility argument of [12], and Lemma C.2 shows that our positive-value mechanism is individually rational for all data subjects as long as analyst cost is not too high relative to data subject valuations.

4.3 Heteroskedastic Noise for Endogenous Privacy

With non-negative and increasing v_i, truncation in Step 1 preserves the incentive compatibility argument of [12]. The proof that the mechanism is individually rational for all data subjects as long as analyst cost is not too high relative to data subject valuations is presented in the appendix. Noting that truncation ensures that q has sensitivity Δ, it remains to argue that releasing noisy

$P \approx cq$ preserves endogenous privacy for some appropriate (positive and decreasing) functions $\epsilon(q), \delta(q)$. We prove privacy with respect to an arbitrary noise function h that is a parameter of a mechanism; Lemma 4.4 describes choices for which privacy is meaningful.

Note that in our mechanism, noise is *heteroskedastic* in that the variance of the Laplace noise added to q is not uniform across all values of q. This deviates significantly from the usual privacy scenario and creates a new challenge in proving endogenous privacy.

Lemma 4.3 (Privately publishing q)**.** Define $\epsilon(q) = 2\Delta/h(q-\Delta)$ and $\delta(q) = \exp(-1/h'(q-\Delta))$ for increasing, differentiable, concave $h : \mathbb{R}^+ \to \mathbb{R}^+$ with $h'(0) \leq 1$, fixed $\Delta \in \mathbb{R}^+$, and any $q \in \mathbb{R}^+$. For any $c \in \mathbb{R}^+$ and neighboring $\mathbf{v} \sim \mathbf{v}'$, let $q = (\sum \bar{v}_i/c - 1)^+, q' = (\sum \bar{v}'_i/c - 1)^+$, and let γ and γ' denote random variables with distributions $\text{Lap}(h(q))$ and $\text{Lap}(h(q'))$, respectively. Then for any $T \subseteq \mathbb{R}$, we have

$$\Pr[q+\gamma \in T] \leq \exp(\epsilon(q)) \cdot \Pr[q'+\gamma' \in T] + \delta(q), \quad (1)$$

$$\Pr[q'+\gamma' \in T] \leq \exp(\epsilon(q)) \cdot \Pr[q+\gamma \in T] + \delta(q). \quad (2)$$

The proof sketch is as follows. If we show 1 for $q < q'$ and then for $q' < q$, then 2 follows by symmetry. First note that if $q < q'$, then $\Pr[q+\gamma = t] \leq \Pr[q'+\gamma' = t]$ for t sufficiently far from q', and otherwise their ratios differ maximally at $t = q$. Then it is enough to show $\Pr[q+\gamma = q]/\Pr[q'+\gamma' = q] \leq \epsilon(q)$, which can be shown using the sensitivity bound on q with the concavity and other assumptions about h.

Now consider $q' < q$. Since $\Pr[q+\gamma = t]/\Pr[q'+\gamma' = t]$ grows with t, we will not be able to achieve $(\epsilon, 0)$-differential privacy for any ϵ. Instead note that $\int_{t^*}^{\infty} \Pr[q+\gamma = t]dt = \exp(-1/h'(q-\Delta))/2 = \delta(q)/2$ for $t^* = q + h(q)/h'(q-\Delta)$. Then we can bound $\Pr[q+\gamma = t]/\Pr[q'+\gamma' = t]$ only for $t \in [q - h(q)/h'(q-\Delta), t^*]$. Since t^* is the point in this range where the pdfs of $q+\gamma$ and $q'+\gamma'$ differ maximally, it is enough to show $\Pr[q+\gamma = t^*]/\Pr[q'+\gamma' = t^*] \leq \epsilon(q)$. As in the first case, this bound is achieved using the sensitivity of q and the assumptions about h.

Note that $q = (\sum \bar{v}_i/c - 1)^+$ is the *unique* q in the privacy support of $\mathcal{M}(\mathbf{d}, \mathbf{v}, c)$ for any inputs $\mathbf{d}, \mathbf{v}, c$. Therefore, Lemma 4.3 establishes endogenous differential privacy (Definition 3.6) of $P = c(q + \text{Lap}(h(q)))$ for the ϵ, δ in the lemma statement. Endogenous differential privacy of the overall mechanism (Theorem 4.1) is an immediate corollary assuming the differential privacy of $\mathcal{M}_f$ in the standard sense and using basic composition.

4.4 Choosing Δ and h for Pareto Efficiency and Accuracy

The internally chosen consumer surplus maximizing privacy level q is noiseless, so its Pareto efficiency follows immediately by the arguments of [12] whenever truncation is avoided, i.e., when each $v_i \leq \Delta \cdot c$. If we expect constant v_i and c, we should set $\Delta = \omega(1)$ to avoid truncation. An immediate consequence of the taxation scheme of [12] is that the budget balances in expectation since $\sum p_i \geq cq$. Accuracy of R is inherited directly from $\mathcal{M}_f$, so the parameters should be set so that $\epsilon(q), \delta(q)$ decrease with n. Good choices are $\Delta = \ln n$ and $h(q) = \sqrt{q+\Delta}$. The following lemma gives more general conditions on Δ, h for which we simultaneously achieve all desired properties:

Lemma 4.4. Let $\Delta = \omega(1)$, and let h be increasing, differentiable, and concave with $h(-\Delta) = 0, h(V) = o(n), h'(0) \leq 1, h'(V) = o(1/\ln n)$ for any $V = \Theta(n)$. If $v_i, c = \Theta(1)$ for $i \in [n]$, then Mechanism 2 on inputs $\mathbf{v}, c$ and any database $\mathbf{d}$ is Pareto efficient and accurate as determined by $\mathcal{M}_f$ with $\epsilon_f \leq O(\Delta/h(V))$ and δ_f for some $V \in \Theta(n)$, endogenously private for $\epsilon(q) = o(1)$ and $\delta(q) = \delta_f + 1/\exp(\Theta(h(n))) \leq \delta_f + 1/\text{poly}(n)$, incentive compatible, individually rational, and has a balanced budget in expectation and a deficit greater than t with probability at most $\exp(-t/h(\Theta(n)))/2$.

5 Open Questions

Our class of endogenously private mechanisms is a special case of the schema for Pareto efficient allocation of goods described in [12]. Their framework allows for multiple public goods with different production prices. This generality could be readily exploited to create markets for privacy with multiple analysts, possibly with different levels of ϵ for different databases or different queries.

Although the negative results extending those of [10] do not hold up under endogenous privacy, the techniques used in Section 4 rely heavily on the view of privacy as a public good. It remains an interesting open question whether the endogenous differential privacy relaxation alone is enough to give a positive result that circumvents the negative result of [10] when data subjects have disutility for imperfect privacy.

Acknowledgements

The author would like to thank Chris Peikert for many helpful discussions. This research was supported by the National Science Foundation under CAREER Award CCF-1054495.

References

[1] J. M. Abowd and I. Schmutte. Revisiting the economics of privacy: Population statistics and confidentiality protection as public goods. Document 22, Labor Dynamics Institute, Jan. 2015.

[2] J. Brodkin. At&t offers gigabit internet discount in exchange for your web history. arstechnica.com/information-technology/2013/12/att-offers-gigabit-internet-discount-in-exchange-for-your-web-history/, Posted: 12/11/2013.

[3] Y. Chen, S. Chong, I. A. Kash, T. Moran, and S. Vadhan. Truthful mechanisms for agents that value privacy. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, EC '13, pages 215–232, New York, NY, USA, 2013. ACM.

[4] E. H. Clarke. Multipart pricing of public goods. *Public Choice*, 11(1):17–33, 1971.

[5] P. Dandekar, N. Fawaz, and S. Ioannidis. Privacy auctions for inner product disclosures. *CoRR*, abs/1111.2885, 2011.

[6] C. Dwork and J. Lei. Differential privacy and robust statistics. In *Proceedings of the 41st annual ACM symposium on Theory of computing*, pages 371–380. ACM, 2009.

[7] C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In S. Halevi and T. Rabin, editors, *Theory of Cryptography*, volume 3876 of *Lecture Notes in Computer Science*, pages 265–284. Springer Berlin Heidelberg, 2006.

[8] L. Fleischer and Y.-H. Lyu. Approximately optimal auctions for selling privacy when costs are correlated with data. In *ACM Conference on Electronic Commerce*, pages 568–585, 2012.

[9] A. Ghosh and K. Ligett. Privacy and coordination: computing on databases with endogenous participation. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, EC '13, pages 543–560, New York, NY, USA, 2013. ACM.

[10] A. Ghosh and A. Roth. Selling privacy at auction. In *Proceedings of the 12th ACM Conference on Electronic Commerce*, EC '11, pages 199–208, New York, NY, USA, 2011. ACM.

[11] T. Groves. *The Allocation of Resources Under Uncertainty: The Informational and Incentive Roles of Prices and Demands in a Team*. Technical report (University of California, Berkeley. Center for Research in Management Science). University of California, 1970.

[12] T. Groves and J. O. Ledyard. Optimal allocation of public goods: a solution to the "free rider" problem. *Econometrica*, 45(4):783–809, May 1977.

[13] J. Hsu, M. Gaboardi, A. Haeberlen, S. Khanna, A. Narayan, B. C. Pierce, and A. Roth. Differential privacy: An economic method for choosing epsilon. *CoRR*, abs/1402.3329, 2014.

[14] K. Ligett and A. Roth. Take it or leave it: running a survey when privacy comes at a cost. In *Proceedings of the 8th International Conference on Internet and Network Economics*, WINE'12, pages 378–391, Berlin, Heidelberg, 2012. Springer-Verlag.

[15] K. Nissim, C. Orlandi, and R. Smorodinsky. Privacy-aware mechanism design. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, EC '12, pages 774–789, New York, NY, USA, 2012. ACM.

[16] K. Nissim, S. Raskhodnikova, and A. Smith. Smooth sensitivity and sampling in private data analysis. In *Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computing*, STOC '07, pages 75–84, New York, NY, USA, 2007. ACM.

[17] K. Nissim, R. Smorodinsky, and M. Tennenholtz. Approximately optimal mechanism design via differential privacy. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, ITCS '12, pages 203–213, New York, NY, USA, 2012. ACM.

[18] K. Nissim, S. Vadhan, and D. Xiao. Redrawing the boundaries on purchasing data from privacy-sensitive individuals. In *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science*, ITCS '14, pages 411–422, New York, NY, USA, 2014. ACM.

[19] Progressive. Snapshot plug-in device terms and conditions. www.progressive.com/auto/discounts/snapshot/snapshot-terms-conditions/, Last updated: 5/11/2017.

[20] P. A. Samuelson. The pure theory of public expenditure. *The Review of Economics and Statistics*, 36(4):387–389, Nov. 1954.

[21] W. Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, 16(1):8–37, 03 1961.

[22] D. Xiao. Is privacy compatible with truthfulness? In *In Proc. ITCS 2013*, pages 67–86, 2013.

A Market Properties

The principal mechanism design goal of both [10] and the current work is to elicit data subjects' true privacy valuations v_i^*. In order to reasonably assume that data subjects report these private types truthfully, we would like our mechanisms to be *incentive compatible*, meaning that each data subject maximizes his expected utility by reporting his true type:

Definition A.1 (Incentive compatibility)**.** A mechanism $\mathcal{M}$ is incentive compatible if for any $i \in [n]$, privacy valuation $v_i^* \in \mathcal{V}$, and any inputs $\mathbf{d} \in \mathcal{D}^n, \mathbf{v}_{-i} \in \mathcal{V}^{n-1}, c \in \mathcal{C}$,

$$v_i^* \in \arg\max_{v_i} \mathrm{E}[v_i^*(q) - p_i],$$

where (q, p_i) is draw from $\mathcal{M}(\mathbf{d}, \mathbf{v}_{-i} \| v_i, c)$ according to the randomness of $\mathcal{M}$.

Note that this definition is only consistent with utility models in which the value of privacy is tied to the privacy output of the particular run of the mechanism rather than a potentially input-independent overall privacy property of the mechanism. The same is true for the definition of *individual rationality*, the property that no one is worse off having participated in the mechanism:

Definition A.2 (Individual rationality)**.** A mechanism $\mathcal{M}$ is individually rational if for any $i \in [n]$ and inputs $\mathbf{d} \in \mathcal{D}^n, \mathbf{v} \in \mathcal{V}^n, c \in \mathcal{C}$,

$$\mathrm{E}[v_i(q) - p_i] \geq 0,$$

where (q, p_i) is draw from $\mathcal{M}(\mathbf{d}, \mathbf{v}_{-i} \| v_i, c)$ according to the randomness of $\mathcal{M}$.

As in [10], we require incentive compatibility of privacy valuations but not of private data, assuming instead that the true private data is already held somewhere. In this case it may not be possible for data subjects to opt out of the privacy market even if they do not expect it to be individually rational for them. For this reason, we consider individual rationality to be a secondary fairness goal and permit mechanisms with qualified individual rationality.

Mechanisms in [10] are required to be *budget-balanced*, i.e., $-P \geq \sum -p_i$ for any run of the mechanism so the mechanism always raises enough funds to pay the data subjects. We will add noise to the analyst's payment to protect privacy and must therefore relax this requirement:

Definition A.3 (Expected balanced budget)**.** A mechanism $\mathcal{M}$ is budget-balanced in expectation if for any inputs $\mathbf{d} \in \mathcal{D}^n, \mathbf{v} \in \mathcal{V}^n, c \in \mathcal{C}$,

$$\mathrm{E}[\textstyle\sum p_i - P] \geq 0,$$

where $(P, \mathbf{p})$ is drawn from $\mathcal{M}(\mathbf{d}, \mathbf{v}, c)$ according to the randomness of $\mathcal{M}$.

A mechanism with a balanced budget in expectation can be thought of as having a cyclically balanced budget, in that its surpluses will offset its deficits over time across many runs. The left tail of $\sum p_i - P$ should be tightly bounded so the mechanism is unlikely to ever run a large deficit. The particular class of mechanisms we study will further ensure that the analyst achieves a fair target revenue, $\mathrm{E}[P] = c(q)$, where the expectation is over the randomness of $\mathcal{M}$ conditioned on the event that q was selected.

Rather than seeking to minimize analyst payment subject to a minimum accuracy requirement or maximize accuracy subject to a maximum budget as in [10], our mechanisms take into account an analyst's desired tradeoff between money and accuracy by soliciting $c \in \mathcal{C} \subseteq \{\mathcal{Q} \to \mathbb{R}\}$, which indirectly describes this tradeoff by assigning a monetary value to each possible privacy level q, representing the cost to the analyst of R generated by the mechanism running on q compared to a noiseless statistic. Given the preferences of data subjects and the analyst, our mechanisms seek to find a *Pareto efficient* (or *Pareto optimal*) level of privacy, meaning one where no data subject can be made strictly better off without making another strictly worse off, subject to collecting enough total funds for the analyst. $\mathcal{V}$ and $\mathcal{C}$ should be chosen so that for any $\mathbf{v} \in \mathcal{V}^n, c \in \mathcal{C}$, there exists some Pareto efficient $q \in \mathcal{Q}$.

Definition A.4 (Pareto efficiency)**.** Privacy level $q \in \mathcal{Q}$ is Pareto efficient for $\mathbf{v} \in \mathcal{V}^n, c \in \mathcal{C}$ if there exist payments $\mathbf{p} \in \mathbb{R}^n$ such that $\sum p_i \geq c(q)$, and for all q' and $\mathbf{p}'$ such that $\sum p'_i \geq c(q')$ and $v_i(q') - p'_i > v_i(q) - p_i$ for some $i \in [n]$, there exists some $j \in [n]$ with $v_j(q') - p'_j < v_j(q) - p_j$.

B Properties of Non-Private Public Goods Allocation

To verify incentive compatibility of the public goods mechanisms discussed in Section 4 [4, 11, 12, 21], note that $\max_{q_{-i}}(\sum_{j \neq i} v_j(q_{-i}) - \frac{n-1}{n} c(q_{-i}))$ is independent of v_i, so to maximize his utility, i should report $\arg\max_{v_i} v_i^*(q(\mathbf{v}_{-i} \| v_i, c)) - (c(q(\mathbf{v}_{-i} \| v_i, c)) - \sum_{j \neq i} v_j(q(\mathbf{v}_{-i} \| v_i, c))$. Because $q(\mathbf{v}_{-i} \| v_i^*, c) = \arg\max_q v_i^*(q) + \sum_{j \neq i} v_j(q) - c(q)$, this quantity is indeed maximized when $v_i = v_i^*$.

A sufficient condition for Pareto efficiency is the Samuelson condition [20], that the sum of the marginal benefit of a public good over all consumers equals its marginal cost. With this allocation rule, the quantity maximizing consumer surplus is q such that $\sum \frac{d}{dq} v_i(q) = \frac{d}{dq} c(q)$. Assuming incentive compatibility, this is equivalent to the condition that the sum of the marginal benefit of q is equal to the marginal cost, so the allocation rule is Pareto efficient.

It can be easily verified that the sum of payments is at least $c(q)$. Since it may be strictly greater, the payments collected are not guaranteed to be Pareto efficient. This is a problem addressed in [12] through dif-

ferent tax and allocation rules, but these modifications complicate the privacy utility model in our setting.

C Privacy Market Proofs

For notational simplicity, we define the following functions for the optimal privacy level and individual taxes computed by the mechanism for inputs $\mathbf{v} \in (\mathbb{R}^+)^n, c \in \mathbb{R}^+$, recalling that $\bar{v}_i = \min(v_i, c \cdot \Delta)$:

$$q(\mathbf{v}, c) = \left(\frac{\sum \bar{v}_i}{c} - 1\right)^+$$

$$p_i(\mathbf{v}, c) = cq(\mathbf{v}, c) - \sum_{j \neq i} \bar{v}_j \ln(q(\mathbf{v}, c) + 1) + \max_{q_{-i}} \left(\sum_{j \neq i} \bar{v}_j \ln(q_{-i} + 1) - \tfrac{n-1}{n} c q_{-i}\right).$$

Lemma C.1. Mechanism 2 is incentive compatible.

Proof. Fix any $i \in [n]$ and $v_i^* \in \mathbb{R}^+$, and denote

$$\begin{aligned} U_i(\mathbf{v}, c) &= v_i^*(q(\mathbf{v}, c)) - p_i(\mathbf{v}, c) \\ &= (v_i^* + \sum_{j \neq i} \bar{v}_j) \ln(q(\mathbf{v}, c) + 1) - cq(\mathbf{v}, c) \\ &\quad - \max_{q_{-i} \geq 0} \left(\sum_{j \neq i} \bar{v}_j \ln(q_{-i} + 1) - \tfrac{n-1}{n} c q_{-i}\right). \end{aligned}$$

We need to show that for any $\mathbf{v}_{-i} \in (\mathbb{R}^+)^{n-1}$ and $c \in \mathbb{R}^+$, we have $v_i^* \in \arg\max_{v_i} U_i(\mathbf{v}_{-i} \| v_i, c)$.

Observing that $\max_{q_{-i} \geq 0} (\sum_{j \neq i} \bar{v}_j \ln(q_{-i} + 1) - \frac{n-1}{n} c q_{-i})$ has no dependence on v_i, we see that $U_i(\mathbf{v}, c)$ increases with q until $q = (v_i^* + \sum_{j \neq i} \bar{v}_j)/c - 1$. Therefore, by declaring $v_i = v_i^*$, $q(\mathbf{v}, c)$ coincides with i's optimal value of q if $v_i^* \leq c \cdot \Delta$, and it it maximizes i's utility subject to truncation otherwise. □

Lemma C.2. Mechanism 2 is individually rational on inputs $\mathbf{v} \in (\mathbb{R}^+)^n, c \leq \sum \bar{v}_i / e$.

Proof. First note that $v_i \ln(q(\mathbf{v}, c) + 1) = v_i \ln \frac{\sum \bar{v}_i}{c} \geq \bar{v}_i$, so it is enough to show that $p_i(\mathbf{v}, c) \leq \bar{v}_i$. Bound p_i as follows:

$$\begin{aligned} p_i(\mathbf{v}, c) &= cq(\mathbf{v}, c) - \sum_{j \neq i} \bar{v}_j \ln(q(\mathbf{v}, c) + 1) \\ &\quad + \max_{q_{-i}} \left(\sum_{j \neq i} \bar{v}_j \ln(q_{-i} + 1) - \frac{n-1}{n} c q_{-i}\right) \\ &= \left(\sum \bar{v}_i - c\right) - \sum_{j \neq i} \bar{v}_j \ln \frac{\sum \bar{v}_i}{c} \\ &\quad + \sum_{j \neq i} \bar{v}_j \left(\left(\ln \frac{\sum_{j \neq i} \bar{v}_j}{\frac{n-1}{n} c}\right) - 1\right) + \frac{n-1}{n} c \\ &= \bar{v}_i - \frac{c}{n} - \sum_{j \neq i} \bar{v}_j \left(1 + \ln \frac{\sum \bar{v}_i}{c} - \ln \frac{\sum_{j \neq i} \bar{v}_j}{\frac{n-1}{n} c}\right) \\ &= \bar{v}_i - \frac{c}{n} - \sum_{j \neq i} \bar{v}_j \ln \frac{e^{\frac{n-1}{n}} \sum \bar{v}_i}{\sum_{j \neq i} \bar{v}_j}. \end{aligned}$$

Since the $\bar{v}_i$ are nonnegative, it is enough to show that $e^{\frac{n-1}{n}} \geq 1$, which clearly holds for any $n \geq 2$. □

Note that the conditions for Lemma C.2 hold whenever $c \leq n$ and $\sum \min(v_i, c \cdot \Delta)/n \geq e$. The mechanism could easily be modified to enforce $c \leq n$ and in many scenarios it may be a reasonable to assume a distribution on v_i satisfying the latter requirement. Note that these qualifications do not affect the impossibility result in [10], which relies on the existence of v_i implying arbitrarily high costs for any fixed ϵ_i. They *do* affect Theorem 3.4 since a mechanism running on this restricted set of inputs cannot output a privacy level with arbitrarily small value to the data subjects. However, note that $\mathbf{v} = (0, \ldots, 0, ce)$ satisfies $c \leq \sum \bar{v}_i / e$ for $\Delta \geq 1$. Mechanism 2 outputs $q = \sum \bar{v}_i / c - 1 = e - 1$ on $\mathbf{v}, c$, and $\sum v_i (e - 1) = ce \ln(e - 1 + 1) = ce$. Then with standard instead of endogenous differential privacy, individually rational mechanisms running on the restricted set of inputs can pay the analyst at most ce with almost $1/2$ probability.

If qualified individual rationality is undesired, one might consider applying the propose-test-release strategy of [6] and aborting as a first step if the conditions of Lemma C.2 are not met. However, note that $\sum \bar{v}_i / c$ has sensitivity Δ. Adding noise $\mathrm{Lap}(\Delta/\epsilon)$ for differential privacy (although there would be some modifications to make this endogenously private) would overwhelm the threshold e when $\Delta = \omega(1)$ as in the usual case, so this strategy seems unlikely to work directly. We leave the issue of unqualified individual rationality as a question for future work.

D Generalized Privacy Valuations

Mechanism 2 relies on the assumption each data subject's utility for the level of q provided by the mechanism is represented by some $v_i(q) = v_i \ln(q+1)$. This choice of logarithmic utility functions was the convenient one, since it allows us to easily bound the sensitivity of q using a simple truncation rule. However, many other non-negative, increasing, concave functions of q may be appropriate models of the utility to data subjects of q.

Consider the case that each data subject has valuation function $v_i(q) = v_i q^{1/a}$ for $a > 1$. As before, we first truncate the v_i so that $\bar{v}_i = \max(v_i, v_{\max})$ for some $v_{\max}$ to be determined later to adequately control the sensitivity of q, which is the level of privacy that maximizes consumer surplus, i.e., $q(\mathbf{v}, c) := \arg\max_{q \geq 0} \sum \bar{v}_i(q) - cq$. Then we have:

$$q(\mathbf{v}, c) = \left(\frac{\sum \bar{v}_i'}{ac}\right)^{\frac{a}{a-1}}$$

$$\begin{aligned}
&|q(\mathbf{v}_{-i} \| v_i, c) - q(\mathbf{v}, c)| \\
&\quad = (ac)^{\frac{a-1}{a}} \cdot |(v_i' + \sum_{j \neq i} v_j)^{\frac{a}{a-1}} - (v_i + \sum_{j \neq i} v_j)^{\frac{a}{a-1}}| \\
&\quad \leq (ac)^{\frac{a-1}{a}} \cdot \left((v_{\max} + \sum_{j \neq i} v_j)^{\frac{a}{a-1}} - (\sum_{j \neq i} v_j)^{\frac{a}{a-1}} \right) \\
&\quad \leq (ac)^{\frac{a-1}{a}} \cdot \left((n v_{\max})^{\frac{a}{a-1}} - ((n-1) v_{\max})^{\frac{a}{a-1}} \right) \\
&\quad \leq \left(\frac{(n-1) v_{\max}}{ac} \right)^{\frac{a}{a-1}} \cdot \left((1 + \frac{1}{n-1})^{\frac{a}{a-1}} - 1 \right) \\
&\quad \leq \left(\frac{(n-1) v_{\max}}{ac} \right)^{\frac{a}{a-1}} \cdot \left(\frac{\frac{a}{a-1}(1+1/n)^{\frac{1}{a-1}}}{n} \right)
\end{aligned}$$

In the case that $a \geq 2$, we have $|q(\mathbf{v}_{-i} \| v_i, c) - q(\mathbf{v}, c)| \leq (n v_{\max}/c)^{\frac{a}{a-1}}/n$. Then if we set $v_{\max} = c\Delta^{\frac{a-1}{a}}/n^{1/a}$, the sensitivity of q is Δ for some fixed Δ as before, and privacy follows as in Lemma 4.3. Incentive compatibility also follows as in Lemma C.1. Individual rationality, however, does not appear to hold for agents with low privacy sensitivity. In particular, when $v_i = 0$, i will always be charged $p_i(\mathbf{v}, c) > 0$ whenever $\sum_{j \neq i} v_j > 0$. With the exception of individual rationality, the other properties of Lemma 4.4 hold with $\Theta(n)$ replaced with $\Theta(n^{a/(a-1)})$. It remains an open problem to identify further classes of valuation functions for which our mechanism or variants of it satisfy all desired properties for endogenous privacy markets.