Differential Privacy: An Economic Method for Choosing Epsilon

Justin Hsu  Marco Gaboardi  Andreas Haeberlen  Sanjeev Khanna  
Arjun Narayan  Benjamin C. Pierce  Aaron Roth  
University of Pennsylvania

ABSTRACT

Differential privacy is becoming a gold standard for privacy research; it offers a guaranteed bound on loss of privacy due to release of query results, even under worst-case assumptions. The theory of differential privacy is an active research area, and there are now differentially private algorithms for a wide range of interesting problems.

However, the question of when differential privacy works in practice has received relatively little attention. In particular, there is still no rigorous method for choosing the two key parameters, \( \epsilon \) and \( \delta \), which control the crucial tradeoff between the strength of the privacy guarantee and the accuracy of the published results.

In this paper, we examine the role that these parameters play in concrete applications of differential privacy, identifying the key challenges that must be addressed when choosing specific values. This choice requires balancing the interests of two different parties: the data analyst and the prospective participant, who must decide whether to allow their data to be included in the analysis. We propose a simple model that expresses this balance as formulas over a handful of parameters. We illustrate the model with two worked examples: a hypothetical medical study and an educational statistics application. We also explore a surprising insight: in some circumstances, a differentially private study can be more accurate than a non-private study for the same cost, under our model. Finally, we discuss a number of simplifying assumptions in our model and outline a research agenda for further refinements.

1 Introduction

Protecting privacy is hard. Experience has repeatedly shown that when owners of sensitive datasets release derived data, the released data can easily reveal more information than intended. Even careful efforts to protect privacy, e.g., by anonymizing or aggregating, often prove inadequate. A notable example is the Netflix prize competition, which released more than 100 million movie ratings by more than 480,000 users. Although the data was carefully anonymized, Narayanan and Shmatikov were later able to “de-anonymize” some of the private records [21].

A common reason for privacy breaches is that the owner of the dataset uses an incorrect threat model—e.g., they make wrong assumptions about the knowledge available to attackers. In the Netflix example, Narayanan and Shmatikov had access to auxiliary data in the form of another public data set (from IMDB) that contained similar ratings but was not anonymized. Such errors are difficult to prevent: doing so seems to require reasoning about all the information that could potentially be available to an attacker.

One way through this dilemma is to make sure that every computation on sensitive data satisfies differential privacy [8]. This gives a very strong guarantee: if an individual’s data is used in a differentially private computation, the probability of any given result changes by at most a factor of \( e^\epsilon \), where \( \epsilon \) is a parameter controlling the tradeoff between privacy and accuracy. In comparison to other privacy frameworks, differential privacy impresses by the long list of assumptions it does not require: it is not necessary to predict what auxiliary information an attacker might have, whether attackers might collude with one another, or what the attackers might be looking for.

And yet, there is one question that users of differential privacy cannot avoid: how to choose the privacy parameter \( \epsilon \). This is the central parameter in differential privacy: \( \epsilon \) controls the number of queries that can be answered privately as well as the achievable accuracy, so the choice of \( \epsilon \) determines whether the analysis will be feasible, how much it will cost, and how useful the results will be. But \( \epsilon \) is also a rather abstract quantity, and it is not immediately clear how to choose an appropriate value in a given situation. This is evident in the literature on differential privacy, where algorithms have been evaluated with \( \epsilon \) ranging from as little as 0.01 to as much as 7, often without explanation. A similar concern applies to a second parameter \( \delta \), which appears in a widely used generalization of differential privacy [7].

In this paper, we take a step towards a more principled approach by examining the impact of \( \epsilon \) and \( \delta \) on the different actors in a differentially private study: the data analyst, and the potential participants who contribute private data. We propose a simple model that can be used to calculate a range of acceptable values of \( \epsilon \) and \( \delta \), based on a small number of parameters of the study. Our model is based on the assumption that the participants are rational and will choose to contribute their data if their expected benefits from the study are greater than their associated risks.

To explore the usefulness of our model, we develop two case studies in detail: a clinical study of the connection between smoking and tuberculosis, and a study of educational data. In both cases, the parameters of our model can easily be estimated from public sources. We also find that—somewhat counterintuitively—a study with strong differential privacy guarantees can sometimes be cheaper or (given a fixed budget) more accurate than an equivalent study without any privacy protections: while a differentially

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright 20XX ACM X-XXXXX-XX-X/XX/XX ...$15.00.
private study requires considerably more participants to account for
the additional noise, it substantially reduces the risks of each par-
ticipant and thus lowers the compensation that rational participants
should demand in return for their participation.

Our model provides a principled way to choose a reasonable
value for \( \varepsilon \) based on parameters that, at least in some cases, are
simple to estimate. For many applications of differential privacy,
this may already be good enough. However, like any model, ours
relies on a number of simplifying assumptions; for instance, we
assume that participants fear some specific bad events when partici-
pating in the study, and that they can estimate their expected cost
from these events even when they do not participate in the study.
Some applications may require a more detailed model.

Our main contributions are: first, a principled approach to choos-
ing the privacy parameter \( \varepsilon \) for differentially private data analysis
(Section 4); second, two case studies: a clinical study (Section 5)
and an educational statistics application (Section 7); and third, an
extension of our model to \( \delta \)-differential privacy (Section 6).
Along the way, we consider when a differentially private study can
be cheaper than a non-private study (Section 6). We discuss possi-
ble extensions of our model in Section 9, review related work in
Section 10, and conclude in Section 11.

2 Background: Differential privacy

Let us begin by reviewing the core definitions of \( \varepsilon \)-differential
privacy. We defer the generalization of \( (\varepsilon, \delta) \)-differential privacy to
Section 8.

Differential privacy [8] is a quantitative notion of privacy that
bounds how much a single individual’s data contributes to a pub-
lished output. The standard setting involves a database of private
information and a mechanism that calculates an output given the
database. More formally, a database \( D \) is a multiset of records be-
longing to some data universe \( \mathcal{X} \), where a record corresponds to
one individual’s private data. We say that two databases are neigh-
bors if they are the same size, and identical except for a single
record. A mechanism \( M \) is a randomized function that takes the
database as input and outputs an element of the range \( \mathcal{X} \).

**Definition 1** ([8]). Given \( \varepsilon \geq 0 \), a mechanism \( M \) is \( \varepsilon \)-
differentially private if, for any two neighboring databases \( D \) and
\( D' \) and for any subset \( S \subseteq \mathcal{X} \) of outputs,
\[
Pr[M(D) \in S] \leq e^{\varepsilon} \cdot Pr[M(D') \in S].
\]
(1)

Note that \( S \) in this definition is any subset of the mechanism’s range.
In particular, when \( S \) is a singleton set \( \{x\} \), the definition states that the
probability of releasing \( x \) on a database \( D \) is at most \( e^{\varepsilon} \) times the
probability of releasing \( x \) on any neighboring database \( D' \).

The definition also implies a lower bound: swapping \( D \) and \( D' \)
yields \( e^{-\varepsilon} \cdot Pr[M(D) \in S] \geq Pr[M(D') \in S], \) or
\[
Pr[M(D) \in S] \geq e^{-\varepsilon} \cdot Pr[M(D') \in S].
\]
(2)

That is, the probability of an output in \( S \) on a database \( D \) is at
least \( e^{-\varepsilon} \) times the probability of an output in \( S \) on a neighboring
database \( D' \). If we combine this with the original guarantee, we can
see that, if \( Pr[M(D) \in S] = 0 \) for some \( D \) and \( S \), it must also be
that \( Pr[M(D') \in S] = 0 \) for all databases \( D' \) — if some outputs are
impossible on one input database, they must be impossible on all
inputs. When \( |e| \ll 1 \), we can approximate \( e^\varepsilon \) by \( 1 + \varepsilon \), thus, for
neighboring \( D, D' \), the differential privacy guarantee is
\[
(1 - \varepsilon) Pr[M(D') \in S] \leq Pr[M(D') \in S] \leq (1 + \varepsilon) Pr[M(D') \in S],
\]
1For example, \( e^{0.5} = 1.65 \), while \( e^{0.1} = 1.11 \).

**The Laplace mechanism** The standard example of a differenti-
ally private mechanism is the Laplace mechanism.

**Theorem 2** ([8]). Suppose \( \varepsilon, c > 0 \). A function \( g \) that maps
databases to real numbers is \( c \)-sensitive if \( |g(D) - g(D')| \leq c \) for all
neighboring \( D, D' \). For such a function, the Laplace mechanism is
defined by \( L_{\varepsilon, c}(D) = g(D) + v \), where \( v \) is drawn from the Laplace
distribution \( \text{Lap}(c/\varepsilon) \). This mechanism is \( \varepsilon \)-differentially private.

Figure 7 shows what the distribution for the Laplace mech-
anism looks like. Intuitively, the scale \( c/\varepsilon \) of the Laplace distri-
bution controls its spread: the distribution is wider for more sensitive
functions (larger \( c \)) or stronger privacy guarantees (smaller \( \varepsilon \)), giv-
ing a higher probability of adding more noise.

For example, suppose that we have a database \( D \) of medical in-
formation and we wish to compute the proportion of smokers in a
differentially private way. If the database has \( N \) records, de-
define \( g(D) = \frac{1}{N} \#(\text{smokers in } D) \). Notice that on any two neigh-
boring databases, this proportion changes by at most \( 1/N \), since the
numerator changes by at most 1 when we modify a single
record. Thus, we can define a differentially private mechanism
\( L(D) = g(D) + v \), where \( v \sim \text{Lap}(1/N) \).

**Benefits** A key benefit of differential privacy lies in its straight-
forward, worst-case assumptions. When formalizing privacy, it is
difficult to reason about what auxiliary information might be avail-
able to the adversary with which to reconstruct private data. For ex-
ample, public rating information from the Internet Movie Database
was used in a surprising way to re-identify users in anonymized
Netflix data [21].

Differential privacy avoids this problem by making no assump-
tions about the adversary’s knowledge. Even knowing all but one
record of the database does not help the adversary learn the last
record: the output of a differentially private mechanism has approx-
imately the same distribution no matter what that record contains.
(Of course, in typical scenarios, the adversary has far less informa-
tion than this, but making this worst-case assumption avoids losing
sleep wondering exactly what the adversary knows or can deduce.)

A second convenient feature of differential privacy is its flexible
framework—the statistical guarantees provided by differential pri-

cacy hold regardless of the particular form of the records, the space
of possible outputs, and the way in which the mechanism operates
internally. Furthermore, these guarantees are preserved under arbi-
trary post-processing: given a differentially private mechanism \( M \)
and a function \( f \) on the outputs, the composition \( f \circ M \) is differenti-
ally private. Hence, outputs of a differentially private mechanism
can be further transformed with no additional risk to privacy.

A third useful property of differential privacy is composition-
ality—the privacy guarantee degrades gracefully when compos-
ing private mechanisms together. For example, running \( k \) \( \varepsilon \-
differentially private mechanisms (in series or in parallel) will yield

\[
\text{Pr}(T \in S) \leq e^{\varepsilon k} \cdot \text{Pr}(M(T) \in S)
\]

for any \( T \) and any \( k \) differentially private mechanisms \( M \) and
mechanisms \( T \) and \( M' \), where \( M' \) is a mechanism that outputs
\( M(T) \).

**Figure 1:** Probability distributions of the Laplace mechanism
for a \( c \)-sensitive function on two neighboring databases.
a $k\epsilon$-private mechanism [10]. This allows straightforward construction of more complex algorithms out of simpler primitives, while preserving privacy.

3 Interpreted of $\epsilon$

Before presenting our model for setting this parameter, let us take a closer look at what $\epsilon$ means. For an intuitive reading of Definition 1, let $x$ be an individual, and let $D$ contain the same data as $D'$, except with one record replaced by the private data of $x$. Here, the differential privacy guarantee states that the probability of any output of mechanism $M$, whether or not $x$ decides to participate, is within an $\epsilon'$ multiplicative factor. Hence, the parameter $\epsilon$ controls how much the distribution of outputs can depend on data from the individual $x$. We first review generic upper and lower bounds on $\epsilon$, which follow from the definition of differential privacy. Then, we consider interpretations of $\epsilon$ that depend on the nature of the data.

Upper bounds on $\epsilon$ While the definition of differential privacy is formally valid for any value of $\epsilon$ [9], values that are too large or too small give weak guarantees. For large values of $\epsilon$, the upper bound on the probability $\Pr[M(D) \in S]$ can arise above one and thus become meaningless: for instance, if $\epsilon = 20$, Equation (1) imposes no constraint on the mechanism’s output distribution unless $\Pr[M(D') \in S] \leq e^{-20}$. To demonstrate this, we describe an $\epsilon$-private mechanism for large $\epsilon$ which is not intuitively private.

Consider a mechanism $M$ with range $\mathcal{X}$ equal to data universe $\mathcal{X}$, and consider a targeted individual $i$. When she is in the database, it publishes her private record with probability $p^* > 1/|\mathcal{X}|$, otherwise it releases a record at random.

We first show that this mechanism is $\epsilon$-differentially private, for large $\epsilon$. Let $j$ be $i$’s record, and let

$$p = \frac{1 - p^*}{|\mathcal{X}| - 1} < \frac{1}{|\mathcal{X}|}$$

be the probability of releasing a record $s \neq j$, when she is in the database. Consider two databases, $D \cup i$ and $D \cup j$, where $i$ is any record. For $M$ to be $\epsilon$-differentially private, it must satisfy that

$$e^{-\epsilon} \Pr[M(D) = j] \leq \Pr[M(D \cup i) = j] \leq e^{\epsilon} \Pr[M(D) = j]$$

and

$$e^{-\epsilon} \Pr[M(D \cup i) = s] \leq \Pr[M(D \cup j) = s] \leq e^{\epsilon} \Pr[M(D \cup i) = s],$$

for all $s \neq j$. Rewriting this means

$$e^{-\epsilon} |\mathcal{X}| \leq p^* \leq e^{\epsilon} \frac{1}{|\mathcal{X}|}$$

and

$$e^{-\epsilon} \frac{1}{|\mathcal{X}|} \leq p \leq e^{\epsilon} \frac{1}{|\mathcal{X}|}.$$ 

By assumption, the left inequality in the first constraint and the right inequality in the second constraint are satisfied. Thus, if

$$\epsilon \geq \ln(p^* |\mathcal{X}|),$$

the first constraint is satisfied. Since the probabilities over all outputs sums to one, we also know $p^* + \left(1 - p^* - 1\right) = 1$. So,

$$\epsilon \geq \ln \left(1 - \frac{p^*}{p(1 - p^*)}\right) \geq \ln \left(\frac{|\mathcal{X}| - 1}{|\mathcal{X}| (1 - p^*)}\right)$$

(4)

suffices to satisfy the second constraint.

Therefore, $M$ is $\epsilon$-differentially private if $\epsilon$ satisfies these equations. For instance, suppose $|\mathcal{X}| = 10^6$, and $p^* = 0.99$. Mechanism $M$ almost always publishes $i$’s record (probability 0.99) if she is in the database, but it is still $\epsilon$-differentially private if $\epsilon \geq 14$.

Clearly, a process that always publish a targeted individual’s data if she is in the database, and never publish her data if she is not in the database, is blatantly non-private. This $\epsilon$-private mechanism does nearly the same thing: with probability $p^* = 0.99$, it publishes $i$’s record with probability at least $p^* = 0.99$ if she is in the database, and with probability $1/|\mathcal{X}| = 10^{-6}$ if she is not. Evidently, this large value of $\epsilon$ does not give a very useful privacy guarantee in this situation.

Lower bounds on $\epsilon$ While choosing $\epsilon$ too large will compromise the privacy guarantee, choosing $\epsilon$ too small will ruin accuracy—the mechanism must behave similarly for databases that are very different. For example, let $D, D'$ be arbitrary databases of size $N$, and let $0 < \epsilon \leq \frac{\ln N}{4}$. Since the two databases have the same size, we can change $D$ to $D'$ by changing at most $N$ rows. By the sequence of intermediate neighboring databases $D_1, \ldots, D_{N-1}$.

By differential privacy,

$$\Pr[M(D) \in S] \leq e^{\epsilon} \Pr[M(D_1) \in S]$$

$$\Pr[M(D_1) \in S] \leq e^{\epsilon} \Pr[M(D_2) \in S]$$

$$\cdots$$

$$\Pr[M(D_{N-1}) \in S] \leq e^{\epsilon} \Pr[M(D') \in S].$$

Combining, $\Pr[M(D) \in S] \leq e^{2\epsilon} \Pr[M(D') \in S]$. Similarly, we can use Equation (2) to show $\Pr[M(D) \in S] \geq e^{-2\epsilon} \Pr[M(D') \in S]$. But we have taken $\epsilon \leq \frac{\ln N}{4}$; thus, the exponents are at most 1 and at least $-1$. So, the probability of every event is fixed up to a multiplicative factor of at most $e$, whether the input is $D$ or $D'$. (Differential privacy with $\epsilon = 1$ guarantees this for neighboring databases, but here $D$ and $D'$ may differ in many—or all!—rows.) Such an algorithm is probably useless: its distribution over outputs depends only weakly on its input.

These simple calculations show that $\epsilon$ should neither be too large nor too small. However, the numbers involved can be many orders of magnitude apart! While it is possible to further tighten these bounds, let us turn instead to the question of how to interpret $\epsilon$ and differential privacy more concretely.

Reasoning about “bad events” One natural interpretation of differential privacy is in terms of bad events. For concreteness, let the mechanism correspond to a scientific study, and suppose the individual has a choice to contribute data. Let $\mathcal{U}$ be the space of all real-world events, and suppose $\mathcal{E} \subseteq \mathcal{U}$ is the set of events that cannot observe an individual’s participation in the mechanism—if

$^3$Even though the upper bound may not guarantee anything, differential privacy still gives some guarantee. For instance, suppose $\epsilon = 20$ and $\Pr[M(D') \in S] = 1/2 > e^{-20}$. The upper bound gives

$$\Pr[M(D) \in S] \leq e^{-20} \cdot 1/2 \approx 10^9,$$

which is useless. However, consider the outputs $S = \mathcal{X} \setminus S$: we know that $\Pr[M(D') \in S] = 1/2$, so by Equation (2),

$$\Pr[M(D) \in S] \geq e^{-20} \cdot 1/2 \approx 10^{-9},$$

which is a nontrivial bound. In particular, it implies that

$$\Pr[M(D) \in S] = 1 - \Pr[M(D) \in \bar{S}] \leq 1 - 10^{-9}.$$
the output of the mechanism is fixed, an individual’s participation should have no effect on the probabilities of events in \( \mathcal{E} \). Formally, conditioning on the output of \( M \), the events in \( \mathcal{E} \) are independent of \( x \)’s participation. Note that the individual should consider this property when thinking about privacy; as we shall see, events not in \( \mathcal{E} \) may not be “protected” by differential privacy.

Let \( \Theta_{bad} \subseteq \mathcal{E} \) be a set of bad events that may become more likely if an individual participates in a private study, such as a rise in their health insurance premium. To connect the output of the mechanism to the real events in \( \Theta_{bad} \), we imagine two runs of the mechanism: one with \( x \)’s data and one without, where the other records are the same in both runs. Let \( x_p \) be the event “\( x \) participates,” \( x_{np} \) be the event “\( x \) does not participate,” and \( R \) be the output of the mechanism (a random variable). For any event \( e \in \Theta_{bad} \), the probability of bad event \( e \) if \( x \) participates in the study is

\[
\Pr[e | x_p] = \sum_{r \in R} \Pr[e | x_p, R = r] \cdot \Pr[R = r | x_p].
\]

Now, the requirement that events cannot “see” \( x \)’s participation means that all the differences between the two trials are due to differences in the output: if the output is the same in both trials, the adversary knows that John Doe does not participate, and \( R \) is the output of the mechanism (a random variable). For any event \( e \in \Theta_{bad} \), the probability of bad event \( e \) if \( x \) participates in the study is

\[
\Pr[e | x_p] = \sum_{r \in R} \Pr[e | x_p, R = r] \cdot \Pr[R = r | x_p].
\]

Now, the requirement that events cannot “see” \( x \)’s participation means that all the differences between the two trials are due to differences in the output: if the output is the same in both trials, the probability of events in \( \Theta_{bad} \) must be the same. That is, the first probability under the summation is the same assuming event \( x_p \) or \( x_{np} \). Also, by differential privacy (Equation (1)), the second probability is bounded by \( e^\varepsilon \) times the probability of output \( r \) if \( x \) does not participate. Hence:

\[
\Pr[e | x_p] \leq \sum_{r \in R} \Pr[e | x_{np}, R = r] \cdot \Pr[R = r | x_p] + e^\varepsilon \cdot \sum_{r \in R} \Pr[e | x_{np}, R = r] \cdot \Pr[R = x_{np} | x_{np}] = e^\varepsilon \cdot \Pr[e | x_{np}].
\]

Therefore, the probability of any bad event in \( \Theta_{bad} \) increases by at most a factor \( e^\varepsilon \) if the individual participates, compared to when the individual does not participate.

While in principle \( \Theta_{bad} \) could contain every possible event that an individual considers bad, this could cause the individual to fear events that will not be affected by her participation. On the other hand, if the individual selects too few events for \( \Theta_{bad} \), she could face unexpected dangers by participating.

By a similar calculation applying Equation (2), we have \( \Pr[e | x_{np}] \leq e^{-\varepsilon} \Pr[e | x_p] \). Of course, these calculations are valid for all \( e \in \mathcal{E} \), not just \( \Theta_{bad} \). In particular, if \( e \) is a “good” event in \( \mathcal{E} \), this bound means that \( e \) will not become much less likely if the individual decides to participate.

Hence, the interpretation of privacy in terms of events: under differential privacy, events in \( \mathcal{E} \) (both good and bad) will happen with nearly the same probability, regardless of whether an individual participates or not.

It is important to note that if the bad event is already quite likely, this guarantee may allow a large absolute increase in the probability of bad events. Thus, the differential privacy guarantee is stronger for bad events that are very unlikely to happen if the individual does not participate. This is true of many unpleasant events concerning private data: for instance, the probability that an individual’s genome is released if they do not participate in a study is low.

A cautionary example: protected and unprotected events

One might wonder why we have defined \( \mathcal{E} \) in such a technical way—why not simply include all possible events? As alluded to before, the reason is that not all events are protected by differential privacy—the probability of some events may become a lot more likely if an individual participates. For a trivial example, the event “John Doe contributes data to the study” is very likely if John Doe participates, and very unlikely if John Doe does not participate.

However, not all cases are so obvious. Consider the following scenario: an adversary believes that a differentially private study is conducted on either a population of cancer patients, or a control group of healthy patients. The adversary does not know which is the case, but the adversary knows that John Doe is part of the study.

Now, suppose the study provider releases a noisy, private count of the number of cancer patients in the study. For this answer to be remotely useful, it must distinguish between the case where all the participants have cancer and the case where none of the participants do. However, this means that the adversary will be reasonably certain whether the participants in the database have cancer, and hence whether John Doe has cancer. This seems to violate differential privacy: by participating in the study, John Doe has revealed private information about himself. Where did we go wrong?

The key subtlety is whether the adversary can observe John Doe’s participation. More precisely, suppose that the released (noisy) count of cancer patients is \( n \), and suppose the bad event John Doe is worried about is “the adversary thinks that John Doe has cancer.” In order for this event to be protected by differential privacy, it must happen with the same probability whether John Doe participates or not, whenever the noisy count is \( n \). If the adversary can truly observe John Doe’s participation, i.e. he can tell if John Doe participates or not, then clearly this is not the case—if John Doe participates, the adversary will believe that he probably has cancer, and if John Doe does not participate, the adversary will not believe this.

On the other hand, if the adversary merely believes that John Doe participated, even if he did not actually participate, the bad event is protected by differential privacy—if John Doe participates and the adversary discovers that he has cancer, the adversary would probably still think he has cancer even if he did not participate.

This example illustrates a fine point about the notion of privacy implicit in differential privacy: while an informal notion of privacy concerns an adversary correctly learning an individual’s secret data, differential privacy deals instead with the end result of privacy breaches. If John Doe does not participate but the adversary thinks he has cancer, John Doe should not be happy just because the adversary did not learn his private data—his insurance premiums might still increase. In this sense, differential privacy guarantees that he is harmed nearly the same, whether he elects to participate or not.

Introducing cost

In reality, some bad events are merely annoying, while some bad events are truly terrible. To capture this distinction, a natural interpretation is to assign a cost to each event. Specifically, suppose the potential participant has a non-negative event cost function \( f_e \) on the space of events \( \mathcal{E} \). Let \( R \) again be the output of mechanism \( M \), and define the associated output cost function \( f \) on the space of outputs \( \mathcal{R} \) by

\[
f(r) = \mathbb{E}_{e \in \mathcal{E}}[f_e(e) | R = r].
\]

Note that

\[
\mathbb{E}_{e \in \mathcal{E}}[f_e(e) | x_p] = \mathbb{E}_{e \in \mathcal{E}}[f_e(e) | x_{np}],
\]

and similarly with \( x_{np} \) instead of \( x_p \), so bounds on the expected value of \( f \) carry over to bounds on the expected value of \( f_e \). Thus, the individual does not need to reason about the set of outputs \( \mathcal{R} \) and the output cost function \( f \) directly; she can simply reason about costs of actual events, represented by \( f_e \). This is convenient, since reasoning about which outputs cause which events is impractical.
Using the differential privacy guarantee, we can bound the expected cost of participating in the study: \(^5\)

\[
e^{-\epsilon} \mathbb{E}_{r \in \mathbb{R}} [f(r) \mid x_{np}] \leq \mathbb{E}_{r \in \mathbb{R}} [f(r) \mid x_p] \leq e^\epsilon \mathbb{E}_{r \in \mathbb{R}} [f(r) \mid x_{np}],
\]

(5)

In other words, the expected cost of \(x\) participating in a study is within an \(e^\epsilon\) factor of the expected cost of declining.

Again, the participant needs to reason carefully about the set \(\mathcal{E}\) to make sure the events she is considering do not depend directly on her participation in the study. Also, if she includes bad events that will not actually be affected by the mechanism, such as the event that an asteroid impact destroys civilization, the participant’s perceived increase in expected cost might be prohibitively (and unjustly) large.

In general, the question of what events a differentially private study may be responsible for is a social question, which might be handled by the legal system—just as laws describe who is liable for bad events, perhaps laws could also describe which events a private mechanism is liable for.

The challenge of setting \(\epsilon\) The calculations we’ve sketched (and others found in the literature) show how to think about \(\epsilon\) abstractly, as a “knob” that can be used to trade off between privacy and accuracy. However, most prior work (with a few exceptions, which we consider in Section 10) focuses on how the knob works rather than how it should be set.

High-level discussions of the privacy implications of possible settings of \(\epsilon\) tend to offer fairly generic guidance, for example reasoning that a 10% increase in the probability of a bad event that is already low probability does not seem very significant, so \(0.1\) is a sensible value for \(\epsilon\). On the other hand, experimental evaluations of differential privacy, where a concrete choice of \(\epsilon\) is required, often just pick a value (with choices from 0.01 [24] to 7 [19]) without detailed justification.

In a sense, the difficulty of choosing \(\epsilon\) is a hidden consequence of a main strength of differential privacy: its extreme simplicity. The \(\epsilon\) parameter is difficult to think about because it rolls up into a single parameter a fairly complex real-world situation involving at least two parties with opposing interests (the analyst and the participants), as well as considerations such as how individuals should be compensated for their risk.

Our goal in this paper is to unpack some of this complexity and offer a more ramified model with more intuitive parameters. We now turn to this task.

4 A two-party model

This section proposes a simple model for choosing \(\epsilon\), involving two rational parties: the analyst and the individuals.

The analyst’s view The analyst’s goal is to conduct a study by running a private mechanism, in order to learn (and publish) some useful facts about a population. The analyst’s main concern is the accuracy of the mechanism’s result, with respect to a benchmark.

One natural benchmark is the sample statistic: the answer for the non-differentially private version of the study. Compared to this standard, the error in the differentially private study is due entirely to noise added to preserve privacy. This error is determined partly by \(\epsilon\), but also (for typical statistical queries where the analyst is trying to learn information about an entire population) by \(N\), the number of records in the analyst’s database: for a larger number of records, there is less privacy loss to any individual, so less noise needs to be added to protect privacy.\(^6\) So, once the analyst chooses a mechanism, he can trade off between \(\epsilon\) and \(N\) to achieve a target level of accuracy.

Another possible benchmark is the population statistic: the true answer on the entire population. This is the goal of statistical inference: infer properties of the population, given only a random sample of individual data (here, the database). If this is the benchmark, an additional source of error is sampling error: the sample is not perfectly representative of the population. Typically, this error decreases as \(N\) increases: as the sample grows, it becomes more representative. This error is independent of differential privacy, and so is independent of \(\epsilon\).

Since these errors both decrease as \(N\) increases, an analyst would conduct huge studies without the second constraint: budget. Each individual needs to be compensated for their participation, so the analyst can only afford studies of limited size. We can now define a model for the analyst:

**Definition 3 (Analyst Model).** The analyst runs a private mechanism \(M\) parameterized by \(\epsilon\) and by \(N\), the number of records. \(M\) has a real-valued accuracy function \(A_M(\epsilon, N)\), where smaller values of \(A_M(\epsilon, N)\) correspond to more accurate results. (We will omit the subscript when the mechanism is clear.) The analyst wants a target accuracy \(\alpha\), and so requires that \(A_M(\epsilon, N) \leq \alpha\). Finally, the analyst has a budget \(B\) to spend on the study.

Depending on what the analyst is trying to learn, he may be able to tolerate a lower or higher total error. For example, if \(A_M\) tracks the additive error of calculating a value in the interval \((0, 1)\), having \(A_M \geq 1\) would be useless. In general, the analyst may even have a utility function that quantifies how bad a specific amount of error is. For simplicity, we assume that the analyst cannot tolerate any inaccuracy beyond the target level \(\alpha\), and accepts any inaccuracy below this level.

The individual’s view We next consider the individuals who might want to contribute their information to a database in exchange for payment. Participants may face personal harm if their data is revealed, so they are willing to join the study only if they are adequately compensated for the risk they take. A simple way to model the individual’s risk is via a cost function \(f\) as described in the Section 2.

In our setting, the individual is offered a choice between participating in a study and declining, but the study is assumed to happen either way. Our model does not decide whether or not to run the study—are the study’s potential discoveries worth the potential harm to individuals? Instead, we assume that someone (perhaps a scientist or an ethics board) has decided that the study will take place, and the individual only gets to decide whether to participate or not. Thus, the individual participates only if she is compensated for her marginal increase in expected cost.

There is a hidden complication in modeling this increase: the individuals do not know exactly what the other records in the database will be. Instead, they only have a belief about the other

\(^5\)For one direction, \n\[
\mathbb{E}_{r \in \mathbb{R}} [f(r) \mid x_p] = \sum_{r \in \mathbb{R}} \Pr[R = r \mid x_p] \cdot f(r)
\]
\[
\leq \sum_{r \in \mathbb{R}} e^\epsilon \Pr[R = r \mid x_{np}] \cdot f(s) = e^\epsilon \cdot \mathbb{E}_{r \in \mathbb{R}} [f(r) \mid x_{np}]
\]

The other direction is similar, appealing to Equation (2)

\(^6\)For example, if the Laplace mechanism is used to release an average value, the sensitivity of the underlying function depends on \(N\); as \(N\) increases, the sensitivity decreases, and less noise is needed to achieve a set level of privacy.
records. This belief also encodes what the individual believes the study will discover.

To calculate the marginal increase in expected cost for a specific individual, let \( D \) be the set of all possible databases of the other individual's data, and let \( E \) be the individual's expected cost if she decides not to participate. Unpacking,

\[
E = \mathbb{E}[f(M(D))] = \sum_{D' \in \mathcal{D}} \Pr[D = D', s = M(D)] \cdot f(s)
\]

Similarly, if \( C \) is the individual's cost for when participating and \( y \) is some record in \( D \),

\[
C = \mathbb{E}[f(M(D \cup x \setminus y))]
\]

But the inner summation is the individual’s expected cost when the rest of the database is known to be \( D' \). By Equation (5), we bound the increase of this cost if \( x \) participates:

\[
\sum_{x \in \mathcal{D}} \Pr[s = M(D \cup x \setminus y) | D = D'] \cdot f(s) \leq \epsilon^2 \sum_{x \in \mathcal{D}} \Pr[s = M(D) | D = D'] \cdot f(s).
\]

Repacking the expressions for \( E \) and \( C \), we get \( C \leq \epsilon^2 E \), hence the individual’s marginal cost of participating is at most \( C - E \leq \epsilon^2 E - E = (\epsilon^2 - 1)E \).

As before, the costs considered must be independent of the individual’s participation, conditioning on the output of the mechanism—our model does not apply if the events depend directly on an individual’s participation in the study.

Now, we are ready to define our model for the individual.

**Definition 4 (Individual Model).** The individuals are offered a chance to participate in a study with a set level of \( \epsilon \) for some payment. Each individual considers a space of real-world events that, conditioned on the output of the study and the database size, are independent of their participation.

Each individual also has a non-negative cost function on this space, which implies a non-negative cost function \( f \) on the space of outputs of the mechanism (as discussed above), and base cost \( \mathbb{E}[f(R)] \), where \( R \) is the random output of the mechanism without the individual’s data.\(^7\) Let \( E \) be an upper bound on the individual’s base costs. The individual participates only if she is compensated for the worst-case increase in her expected cost by participating, or \((\epsilon^2 - 1)E\).

\( E \) represents an upper bound on the individuals’ beliefs about how much the study will cost them if they do not participate in the study, since the study may discover something that causes her costs to increase even if she declines to participate. For instance, a study might discover that people in a certain town are likely to have cancer—this knowledge could harm all the residents of the town, not just the participants.

Also, note the requirement on the space of bad events: we condition on the output of the mechanism, as well as the size of the database. Intuitively, this is because the size of the database is also published. While such information is certainly private,\(^8\) it is hard to imagine conducting a study without anyone knowing how many people are in it—for one thing, it controls the budget for a study. By the conditioning, we require that an adversary cannot infer an individual’s participation even if he knows the database size and the output of the mechanism.

**Combining the two views** To couple the two views, we assume that the analyst compensates the participants. Since \( N \) individuals need to be paid \((\epsilon^2 - 1)E\) each and the analyst has total budget \( B \), we have the following budget constraint:

\[
(\epsilon^2 - 1)EN \leq B
\]

This constraint, combined with the analyst’s accuracy constraint \( A_M(\epsilon, N) \leq \alpha \), determines the feasible values of \( N \) and \( \epsilon \). In general, there may be no be any feasible values: in this case, using mechanism \( M \) will not meet the requirements of the study. On the other hand, there may be multiple feasible values. These trade off between the analyst’s priorities and the individual’s priorities: larger values of \( \epsilon \) and smaller values of \( N \) make the study smaller and more accurate, while smaller values of \( \epsilon \) and larger values of \( N \) give a stronger guarantee to the individuals. In any case, feasible values of \( N \) and \( \epsilon \) will give a study that is under budget, achieves the target accuracy, and compensates each individual adequately for their risk.

Note that the payments depend on the particular study only through the \( E \) parameter—different studies require different data, which may lead to different base costs—and the \( \epsilon \) parameter, which controls the privacy guarantee. Internal details about the study do not play a role in this model—using other details to justify that a particular \( \epsilon \)-differentially private mechanism is somehow safer than a generic \( \epsilon \)-differentially private mechanism circumvents the privacy guarantee.

By using differential privacy as an abstraction, the model automatically covers differentially private mechanisms in many settings: offline, interactive, distributed, centralized, and more. Further, the model can be applied whether the analyst has benevolent intentions (conducting a study) or malicious ones (violating someone’s privacy). Since differential privacy does not make these distinctions, neither does the model.

**Deriving the cost \( E \)** While the expected cost of not participating in a study is an intuitively simple idea, it depends on a combination of factors. For instance, the cost depends on what the individuals believe about the outcome of the study, as well as what bad events individuals are worried about. The cost even depends on prior private studies an individual has participated in—the more studies, the higher the base cost.

Since individuals have potentially different belief about this cost, the analyst must be sure to offer enough payment to cover each individual’s expected cost. Otherwise, there may be sampling bias: individuals with high cost could decline to participate in the study. While the analyst would like to offer each individual just enough compensation to incentivize them to participate, the costs can depend on private data. Thus, we model the analyst as assuming all individuals have some maximum cost, and offering everyone the same compensation.

---

\(^7\)Note that there are two sources of randomness here: from the mechanism itself, and from the individual’s uncertainty about the rest of the data that goes into the mechanism.

\(^8\)In the worst case, an adversary may know the exact count of individuals with some disease to within 1—publishing the number of individuals with the disease could violate an individual’s privacy.

\(^9\)In reality, there is often a cost for the analyst that scales according to the size of the study. It is not difficult to incorporate this into our model, but for simplicity we leave it out.
Since this maximum expected cost is hard to perfectly calculate in reality, it should be estimated. This can be done in many ways: reasoning about specific bad events and their costs, conducting surveys about what people think their expected cost is, etc. While these estimates may be imprecise, we can build in a margin of error by overestimating the costs. This has the effect of compensating individuals more for their privacy at the cost of reduced accuracy. This cost could also be found via an auction [12, 18, 2, 5, 23].

5 Case Study: Insurance for smokers

In this section, we will show how to apply our model in a study of smokers. The participants are afraid that their insurance company will discover that they smoke, so the analysts must protect this data.

A basic example: answering one query Suppose we are the analyst, and we want to answer a single query: say, estimate the proportion of smokers in the general population with tuberculosis. For accuracy, we will fix the desired error for estimating the proportion of smokers in the sample (the sample mean) from the population mean by at least \( T/N \) for large enough \( T \). Let \( \mu \) be the population proportion of smokers in the general population with tuberculosis. The analyst, and we want to answer a single query: say, estimate the proportion of smokers in the sample (the sample mean) from the population mean by at least \( T/N \) for large enough \( T \). Let \( \mu \) be the population proportion of smokers in the general population with tuberculosis.

Now that we have specified a mechanism, we need to reason about the accuracy and budget of this study. For accuracy, we first need to derive the tradeoff between \( \epsilon, N \), and accuracy for this mechanism, represented by \( A(\epsilon, N) \), in order to use the analyst model.

In general, there are several choices of what \( A \) can mean. In this example, we will fix the desired error \( T/N \) for large enough \( T \) and deviate from the true population mean that we are interested in. Note that the sample mean is free of sampling bias for this to hold—inferring population statistics from a non-representative sample will skew the estimates. This is why participants should be compensated so that they are incentivized to participate, regardless of their private data.

We use the following result to bound the probability of the Laplace noise being large.

**Theorem 6 (Tail bound on Laplace distribution).** Let \( \nu \) be drawn from \( \text{Lap}(\rho) \). Then,

\[
\Pr[|\nu| \geq T] \leq \exp \left(-\frac{T}{\rho}\right).
\]

Now, since the total error of the mechanism is the difference between the sample mean and the population mean plus the Laplace noise, if the output of the Laplace mechanism deviates from the population mean by at least \( T/N \), then either the sample mean deviates by at least \( T/2N \), or the Laplace noise added is of magnitude at least \( T/N \). Therefore, we can bound the failure probability \( A(\epsilon, N) \) by

\[
\Pr[|g(D_N) - \mu| \geq T/2] + \Pr[|\nu| \geq T/2] \geq \frac{1}{2}. \tag{7}
\]

Likewise, we can use the tail bound on the Laplace distribution to bound the second term by

\[
\Pr[|\nu(N)| \geq T/2] \leq \exp \left(-\frac{T\epsilon}{2}\right),
\]

since we added noise with scale \( \rho = 1/\epsilon N \). Therefore, we define the accuracy constraint to be

\[
A(\epsilon, N) := 2\exp \left(-\frac{N\epsilon^2}{12}\right) + \exp \left(-\frac{T\epsilon}{2}\right) \leq \alpha. \tag{7}
\]

For the budget side of the problem, if our budget is \( B \), we have the budget constraint Equation (6): \( \epsilon(\epsilon - 1)EN \leq B \).

Our goal is to find \( \epsilon \) and \( N \) that satisfy this budget constraint, as well as the accuracy constraint Equation (7). While it is possible to use a numerical solver to find a solution, here we derive a more illuminating closed-form solution. Eliminating \( \epsilon \) and \( N \) from these constraints is difficult, so we find a sufficient condition on feasibility instead. First, for large enough \( \epsilon \), the sampling error dominates the error introduced by the Laplace noise. Since for \( \epsilon \geq T/6 \),

\[
\exp \left(-\frac{TN\epsilon}{2}\right) \leq \exp \left(-\frac{N\epsilon^2}{12}\right),
\]

it suffices to satisfy this system instead:

\[
3\exp \left(-\frac{N\epsilon^2}{12}\right) \leq \alpha \quad \text{and} \quad (\epsilon^2 - 1)EN \leq B.
\]

Solving, we need

\[
N \geq \frac{12}{T^2} \ln \frac{3}{\alpha \epsilon^2}.
\]

Taking equality gives the loosest condition on \( \epsilon \), when the second constraint becomes

\[
\epsilon \leq \ln \left(1 + \frac{BT^2}{12E \ln \frac{\alpha}{\epsilon}}\right).
\]

Thus, combining with the lower bound on \( \epsilon \), we have

\[
\frac{T}{6} \leq \epsilon \leq \ln \left(1 + \frac{BT^2}{12E \ln \frac{\alpha}{\epsilon}}\right), \tag{8}
\]

then the study can be done at accuracy \( \alpha \), budget \( B \).

For a concrete instance, suppose we want the true proportion to be \( \pm 0.005 \) accuracy (0.5% additive error), so we take \( T = 0.005 \). We want this accuracy except with at most \( \alpha = 0.05 \) probability, so that we are 95% confident of our result.
Rather, participating in a study may lead to a payment. This is not impossible—perhaps an insurance agency employees the individual smoking outside. However, it also is not very likely—insurance agencies generally do not spy on individuals, trying to catch smokers.

So, suppose the participants think there is a moderate, 20% chance that the insurance company finds out that an individual smokers, even if she does not participate in the study. This is not impossible—perhaps an insurance agency employee observes the individual smoking outside. However, it also is not very likely—insurance agencies generally do not spy on individuals, trying to catch smokers.

For the precision constraint, let the precision be $\varepsilon = 4.5 \times 10^{-5}$. To estimate each individual’s base cost, we need to reason about the individual’s costs that might be affected by this study.

For the sake of example, suppose that the health insurance company does not know that the individual smokes, and the individual is afraid that the insurance company will discover this and raise her premiums. Taking some average figures, the average health insurance premium for smokers is $1274$ more, compared to nonsmokers [20]. Thus, some participants fear a price increase of $1274$.

To calculate how much individuals should be compensated, we need to reason about the probability that the insurance company finds out that an individual smokes, even if she does not participate in the study. This is not impossible—perhaps an insurance agency employee observes the individual smoking outside. However, it also is not very likely—insurance agencies generally do not spy on individuals, trying to catch smokers.

Figure 2: Feasible $\varepsilon, N$, for accuracy $\alpha$ and budget $B$.

Figure 3: Constant accuracy curves for $\alpha_1 < \alpha_2$, constant budget curves for $B_1 < B_2$.

A more realistic example: answering many queries. The previous example has a significant drawback in that the mechanism only answers a single query. Any realistic study, medical or otherwise, will need many more queries. In this section, we will see how to carry out calculations for a more sophisticated algorithm from the privacy literature: the multiplicative weights exponential mechanism (MWEM) [13, 14].

MWEM is a mechanism that can answer a large number (exponential in $N$) of counting queries: queries of the form “What fraction of the records in the database satisfy property $P$?” For a concrete example, suppose that the space of records is bit strings of length 20, i.e., $\mathcal{X} = \{0, 1\}^{20}$. Each individual’s bit string can be thought of as a list of attributes: the first bit might encode the gender, the second bit might encode the smoking status, the third bit might encode whether the age is above 50 or not, etc. Then, queries like “What fraction of subjects are male, smokers and above 50?”, or “What proportion of subjects are female nonsmokers?” are counting queries.

We will use an accuracy bound for MWEM [14]. For a data universe $\mathcal{X}$, set of queries $\mathcal{X}$ and number of records $N$, the $\varepsilon$-private MWEM answers all queries in $\mathcal{X}$ to within additive error $T$ with probability at least $1 - \beta$, where

$$T = \left( \frac{128 \ln |\mathcal{X}| \ln \left( \frac{32 |\mathcal{X}| \ln |\mathcal{X}|}{\varepsilon T^2} \right)}{\varepsilon N} \right)^{1/3}.$$  

To fit this into our model, we define the accuracy function $A(\varepsilon, N)$ to be the probability of exceeding error $T$ on any query, i.e. $\beta$ above. Solving, we can set

$$A(\varepsilon, N) := \beta = \frac{32 |\mathcal{X}| \ln |\mathcal{X}|}{T^2} \exp \left( - \frac{\varepsilon NT^3}{128 \ln |\mathcal{X}|} \right),$$  

fixing the accuracy constraint $A(\varepsilon, N) \leq \alpha$. The budget constraint is $(\varepsilon^2 - 1)EN \leq B$, like the previous example.

For the various parameters, suppose we want $\mathcal{X} = \{0, 1\}^5$ so $|\mathcal{X}| = 2^5$ and accuracy $T = 0.2$ for 20% error. Further, we want to answer 1000 queries, so $|\mathcal{X}| = 1000$. For the budget, suppose the individuals remain worried about their health insurance premiums, so $E = 254.8$. If the budget $B = 2.9 \times 10^8$, the constraints are satisfiable—take $\varepsilon = 0.2$, $N = 5 \times 10^9$, when each participant is paid $(\varepsilon^2 - 1)E = 56.4$. In Section 8, we will see a different version of MWEM with better costs.
6 The true cost of privacy

Now that we have a method for estimating the cost of a private study, we can compare this cost to that of an equivalent, non-private study. Differential privacy is often cast as adding additional noise to protect privacy, requiring a larger sample to achieve the same accuracy. Hence, one would expect private studies to be more expensive than equivalent non-private studies.

While this is true if individuals are paid the same in both cases, differential privacy has a real advantage: it can bound the harm to individuals, whereas—as numerous de-anonymization attacks have shown—it is difficult for non-private studies to make any guarantees about privacy. When an individual participates in a non-private study, it is very possible that their information can be completely recovered from published results—non-private studies do not guarantee that re-identification happens with probability less than one.

Thus, to calculate the cost for the non-private study, we consider a hypothetical world where non-private study participants are compensated for their worst-case expected cost $W$, i.e., their cost for having their data published in the clear. We will use our model to find the cost of the private study.

Note the difference between $W$ and $E$: the former measures the harm to the individual of revealing private data, while the latter measures the harm to the individual from running the study, even without the individual’s participation. Naturally, $W$ is typically much higher than $E$—the expected harm from having private data directly published is usually greater than the expected harm of not even contributing private information.\(^2\)

Consider the medical study from Section 5. For the non-private study with $N'$ individuals, we directly release the sample mean $g(D) = \frac{1}{N'} \sum X_i$. Thus, we do not have to bound the error from Laplace noise; instead, all the error is due to the sample mean deviating from the population mean. First, we calculate the sample size $N'$ a non-private study needs in order to achieve a level of accuracy that is equivalent to that of the private study. This will determine the minimum budget that is needed for the non-private study.

We use the following bound:

**Theorem 7** (Chernoff Lower Bound). Suppose \{\{X_i\}\} are $N$ independent, identically distributed $0/1$ random variables with mean $\mu \leq 1/4$ and sample mean $Y = \frac{1}{N} \sum X_i$. For $T \in \{0, 1\}$,

$$\Pr[|Y - \mu| \geq T] \geq \frac{1}{2} \exp\left(-\frac{2NT^2}{\mu}\right).$$

Intuitively, the standard Chernoff bound says that the probability of the sample mean deviating from the population mean is at most some value, but this bound might be very weak: the true probability of deviation could be much lower. The lower bound theorem says that the probability of deviation is also at least some value. This will allow us to lower bound the number of individuals needed to achieve a desired accuracy.

**Theorem 8.** Given a target error $T \geq 0$ and target accuracy $\alpha > 0$, the private medical study will be cheaper than the non-private medical study if the following condition holds.

$$\frac{T}{2} \leq \ln \left(1 + \frac{W \ln \frac{1}{\alpha}}{96E \ln \frac{1}{\alpha}}\right)$$

\(^2\)While $E$ may include costs for the bad event “an adversary guesses my data and publishes it in the clear,” which would lead to cost $W$, this cost should be weighted by the (very low) probability of disclosure if the individual does not even contribute her data.

**Proof.** First, we derive a lower bound on the sample size $N'$ that is necessary for the non-private study to have low error. Since we want the deviation probability to be at most $\alpha$ for all $\mu$, in particular the sample size must be large enough guarantee this error for $\mu = 1/4$. Theorem 7 gives

$$\alpha \geq \Pr[|Y - \mu| \geq T] \geq \frac{1}{2} e^{-2NT^2/\mu},$$

which is equivalent to

$$N' \geq \frac{1}{8T^2} \ln \frac{1}{2\alpha},$$

Thus, the minimum budget for the non-private study is at least $B' = N'W = \frac{W \ln \frac{1}{2\alpha}}{96E \ln \frac{1}{\alpha}}$, so by Equation (8), the private study will be cheaper than the non-private study if

$$\frac{T}{2} \leq \ln \left(1 + \frac{W \ln \frac{1}{\alpha}}{96E \ln \frac{1}{\alpha}}\right).$$

Concretely, recall we wanted to estimate the mean of a population to accuracy $T = 0.005$, with failure probability at most $\alpha = 0.05$. The non-private study is compensating each individual with their worst-case cost, which is at least the increase in insurance premium to the rate for smokers. Using publicly available data for premiums, smokers pay an additional $1274 per year compared to nonsmokers. For comparison with the private study, smokers believe their insurance premium will increase by $1274$ if their data is published, so $W$ is at least $1274$. For the private study, individuals continue to have base cost of $E = 254.8$. Plugging into Equation (9), we find that the private study is cheaper.

Let us compare the size and costs of the two studies, for the case when $\mu = 1/4$. Plugging into Equation (10), we get $N' \geq 11513$. As expected, the non-private study needs far fewer people to achieve the same accuracy compared to the private study ($N = 2 \times 10^9$), since no noise is added. However, the total cost for non-private study would be $B' = 1274 \times 11513 \approx 1.5 \times 10^7$, significantly more than the budget required for the private study ($B = 4.5 \times 10^5$).

If both private and non-private studies have the same budget, the private study can buy more participants to further improve its accuracy. Thus, this private study is more accurate and cheaper (and more private!) than the non-private study.

7 Case Study: Educational statistics

In this section, we will show how to apply our model to a more complex example. Suppose a university department wants to conduct statistical analysis on academic data from its students. This data is sensitive—it contains information like grades and class year. At many universities, access to such data is tightly controlled by an institutional review board, charged with protecting the privacy of the students. Since all requests need to go through this review, a simple and rigorous method of ensuring privacy would make the process more efficient. At first glance, differential privacy seems to fit the bill. What does our model say about this situation?

Concretely, suppose that each student’s record contains class year (4 possible values), grade point average (rounded to the nearest letter grade, so 5 possible values), years declared in the major (100 possible combinations). The total space of possible values is $|X| = 2 \times 10^9$. In this scenario, the administrators already have this data: the students may not have a meaningful choice to participate or not in the study—they may have already graduated, and it might be difficult to track them down to compensate them. So far, our model
assumes the individual has a real choice about whether to participate or not. To get around this problem, we could imagine running a thought experiment to calculate how much we would have had to pay each student to incentivize them to participate if they really had a choice. If this amount is small, then the expected harm to any student is not very high, and we can run the study.

Recall that our model assumes the study will be run, and compensates individuals their marginal increase in expected harm when participating. In general, the increase in cost for participating in a study compared to not running the study at all can be high.\textsuperscript{13} Since our study will not be run if the harm to students is high, should students demand more compensation?

The answer turns out to be no. We first estimate the costs, and then decide whether or not to run the study. If we do, we ask students if they want to participate in exchange for compensation (in our thought experiment). The point is that when the students are given the (hypothetical) choice, we have already decided that the study will happen. Thus, they should be compensated for their marginal harm in participating.

For a worked out instance, suppose we decide to use the Laplace mechanism to answer one counting query on this data. The first step is calculating the accuracy. Unlike the medical study example, we assume our sample contains the entire population since we have data for all the students. For a statistic g and a database D, the Laplace mechanism returns g(D) with noise v(e,N) drawn from Lap(1/N\varepsilon). We bound the probability of the error being too high:

\[ A(e,N) := \Pr[|g(D) + v(e,N) - g(D)| > T] \leq \alpha. \]

By the tail bound on Laplace noise (Theorem 6), we know

\[ A(e,N) = \Pr[|v(e,N)| > T] \leq \exp(-T\varepsilon). \]

Let the tolerable additive error be T = 0.05, or 5%. We want this to be satisfied with probability at least 95%, so \( \alpha = 0.05 \). Taking the constraint \( A(e,N) \leq \alpha \) and plugging in, we require

\[ A(e,N) = \exp(-0.05 \cdot N\varepsilon) \leq 0.05. \]

For the budget side of the problem, we first need to estimate E, the expected harm from not participating in the study. For a concrete instance, let us suppose that students are worried about employers finding out that they did not take an important course, and withdrawing their offer. If students do not participate, it is hard to see how employers could learn this information—most employers do not ask for a full list of courses that a student has put that publishes user y’s information, suppose \( \Pr[M(D') \in S] = 0 \). Under \( \varepsilon \)-differential privacy, M can never output s, on any database. However, under \( (\varepsilon, \delta) \)-differential privacy, M is free to output s with probability up to \( \delta \), when fed any neighboring database D. In particular, if \( D = D' \cup x \setminus y \), this means that if user x replaces any individual y in the database, with \( \delta \) probability the mechanism provides no useful guarantee: her information may be publicly released.

**Modeling the parameter \( \delta \)** By considering non-private mechanisms that satisfy \( (\varepsilon, \delta) \)-privacy, we can upper bound \( \delta \). For example, for a database with N records and for \( \delta = 1/N \), the mechanism that randomly outputs a record from the database is \( (0, \delta) \)-private. This mechanism is not intuitively private, so we need \( \delta \ll 1/N \) for any reasonable guarantee.

For a more careful method of picking this parameter, we can model the costs associated with different levels of \( \delta \). The first step is to bound the increase in expected cost for participating in an \( (\varepsilon, \delta) \)-private mechanism. We need to assume a bound on an individual’s cost if their data is publicly revealed, since with probability \( \delta \) the mechanism may do just that.

**Proposition 10.** Let M be an \( (\varepsilon, \delta) \)-private mechanism with output range \( \mathcal{R} \), and let \( f \) be a non-negative cost function over \( \mathcal{R} \). Let \( W = \max_{e \in \mathcal{R}} f(e) \). Then, for neighboring databases D, D',

\[ E[f(M(D))] \leq e^\varepsilon E[f(M(D'))] + \delta W. \]

We can now incorporate the \( \delta \) parameter into our model.
DEFINITION 11 ((ε, δ)-PRIVATE ANALYST MODEL). An
(ε, δ)-private analyst is an analyst whose accuracy function \(A\) is a
function of \(N, \delta\).

DEFINITION 12 ((ε, δ)-PRIVATE INDIVIDUAL MODEL). An
(ε, δ)-private individual is an individual with a worst-case cost
\(W\), which measures the cost of publicly revealing the individual’s
private information. The individual wants to be compensated
for her worst-case marginal cost of participating under these
assumptions, or \(e^\epsilon E + \delta W - E = (e^\epsilon - 1)E + \delta W\).

Since \((\epsilon, \delta)\)-privacy is weaker than pure \(\epsilon\)-privacy, why is this
a useful notion of privacy? It turns out that in many cases, \((\epsilon, \delta)\)-
private algorithms are more accurate than their pure privacy coun-
terparts. Let us consider such an example.

Revisiting MWEM In Section 5, we analyzed the cost of
MWEM for a medical study application. We will now revisit that
example with an \((\epsilon, \delta)\)-private version of MWEM. The setting re-
 mains the same: we wish to answer a large number of counting
queries to good accuracy, while preserving privacy.

The main difference is the accuracy, due to Hardt and Roth-
blum [14]. Suppose the space of records is \(\mathcal{X}\) and we want to
answer a set of counting queries \(C\) to additive accuracy \(T\) with
probability at least \(1 - \beta\). The \((\epsilon, \delta)\)-private MWEM has accuracy
\[
T = \frac{8(\ln |\mathcal{X}| \ln(1/\delta))^{1/4} \ln(1/2)}{\epsilon 2 \beta^2}.
\]
We define our accuracy measure \(A(\epsilon, N)\) to be the failure probabil-
ity \(\beta\). Solving this, means
\[
A(\epsilon, N) := \beta = \frac{32|\mathcal{X}| \ln |\mathcal{X}|}{T^2} \exp \left(-\frac{\epsilon NT^2}{8(\ln |\mathcal{X}| \ln(1/\delta))^{1/2}}\right).
\]
If \(\alpha\) is the target accuracy, the accuracy constraint is \(A(\epsilon, N) \leq \alpha\).

For the budget constraint, suppose we are in the same medical
study setting as before. Individuals are worried about their health
insurance premium increasing, and have base cost \(E = 254.8\). Next,
we estimate the worst-case expected cost \(W\). We do not know
exactly how much cost an individual will face if their record is
published—the cost may be very high, due to the sensitive nature
of the data. So, we take a conservative upper bound on this cost:
\(W = 10^9\). Let the total budget be \(B = 8 \times 10^7\), so that the budget
constraint is \((e^\epsilon - 1)EN + \delta W N \leq B\). For the other parameters,
suppose the records are bit strings: \(\mathcal{X} = \{0, 1\}^6\) and \(|\mathcal{X}| = 2^6\).

We want to answer \(|C| = 1000\) queries, to 5% error (\(T = 0.05\)),
with probability at least 95\% (\(\alpha = 0.05\)).

Plugging in the numbers, we find that the accuracy and budget
constraints can both be satisfied, for \(\epsilon = 0.1\), \(\delta = 10^{-5}\), and \(N = 2.8 \times 10^6\).
Each individual is compensated \((e^\epsilon - 1)E + \delta M = 27\),
for a total cost of \(7.56 \times 10^7\). This cost is below the cost of
the \(\epsilon\)-private MWEM \((7.6 \times 10^7\) rather than \(2.9 \times 10^8\), with much
better accuracy (\(T = 0.05\) rather than \(T = 0.2\)), and a larger space
of possible records (\(|\mathcal{X}| = 2^7\) rather than \(|\mathcal{X}| = 2^5\).

9 Discussion

In this section, we discuss potential refinements to our model.

Why all the complexity? Compared to earlier threat models
from the differential privacy literature, our model may seem overly
complex: Dwork’s original definition [6] had only one parameter
(\(\epsilon\)), while our model requires practitioners to choose a number of
different parameters \((\alpha, A_M(\epsilon, N), B, \ldots)\). So, at first glance, the
original model seems preferable. However, we argue that this com-
plexity is not an artifact of our model, it is present in the real world:
the individuals really do have to consider the possible consequences
of participating in the study, the researchers really do require a cer-
tain accuracy, etc. The original model blends these considerations
into a single number, \(\epsilon\), but as a result, \(\epsilon\) is a rather abstract quan-
tity that is difficult to choose. Since our proposed model is more
detailed, it makes several of these choices explicit, and it forces
the user to think quantitatively about how a private study would affect
events in the real world. This is not a bug, it is a feature!

Possible refinements The key challenge in designing any
model is to balance complexity and accuracy. Our model is in-
tended to produce reasonable suggestions for \(\epsilon\) in most situations
while keeping only the essential parameters. Below, we review
some areas where our model could be refined or generalized.

Collusion: Though the individuals in our model have the choice of
participating, we assume the study will happen regardless of their
choice. However, this may not hold if individuals collude. In an ex-
treme case, all individuals could collectively opt out, perhaps mak-
ing a study impossible to run. While widespread collusion could be
problematic, it may be possible to incorporate assumptions about
the size of limited coalitions into our model.

Large \(\epsilon\): As \(\epsilon\) increases, our model predicts that the individ-
ual’s marginal expected harm increases without bound. This is
unreasonable—there should be some maximum cost for participat-
ing in a study (perhaps the worst case cost \(W\)). The cost curve could
be refined to be more accurate for small and large \(\epsilon\).

Judgment under uncertainty: Research in psychology has iden-
tified a number of cognitive biases that limit how well people can
reason about uncertain events [15]. Thus, if the consequences of
participation are uncertain, the individuals might under- or overes-
timate their expected cost (in some experiments, people willingly
give up private data for as little as a dollar [31]). To account for this,
our model could be combined with models from psychology.

Risk and loss aversion: Individuals tend to be averse to risk and
to losses [25]; for instance, losing a thousand dollars tends to cause
individuals more pain than winning the same amount would bring
them pleasure. To account for this, it may be necessary to increase
the compensation offered to the individuals.

Alternative mechanisms: When our model shows that a mecha-
nism is infeasible, it is not clear whether the problem lies with the
mechanism, or whether the desired task is impossible under differ-
ential privacy. Often, there are multiple ways to answer a given
question privately. For example, answering range queries which
can be done in ways that are more efficient than the straightforward
approach, e.g., using wavelets [27].

10 Related work

There is by now a vast literature on differential privacy, which we
do not attempt to survey. We direct the interested reader to an ex-
cellent survey by Dwork [6].

The question of how to set \(\epsilon\) has been present since the intro-
duction of differential privacy. Indeed, in early work on differential
privacy, Dwork [6] indicates that the value of \(\epsilon\), in economic terms
or otherwise, is a “social question.” Since then, few works have
taken an in-depth look at this question. Works applying differential
privacy have used a variety of choices for \(\epsilon\), mostly ranging from
0.01–10 , with little or no convincing justification.

The most detailed discussion of \(\epsilon\) we are aware of is due to Lee
and Clifton [16]. They consider what \(\epsilon\) means for a hypotheti-
cal adversary, who is trying to discover whether an individual has
participated in a database or not. The core idea is to model the adversary as a Bayesian agent, maintaining a belief about whether the individual is in the database or not. After observing the output of a private mechanism, he updates his belief depending on whether the outcome was more or less likely if the individual had participated.

As Lee and Clifton show, $\epsilon$ controls how much an adversary’s belief can change, so it is possible to derive a bound on $\epsilon$ in order for the adversary’s belief to remain below a given threshold. We share the goal of Lee and Clifton of deriving a bound for $\epsilon$, and we improve on their work. First, the “bad event” they consider is the adversary discovering an individual’s participation in a study. However, by itself, this knowledge might be relatively harmless—indeed, a main goal of differential privacy was to consider harm beyond reidentification.

Second, and more seriously, the adversary’s Bayesian updates (as functions of the private output of the mechanism) are themselves differentially private: the distribution over his posterior beliefs conditioned on the output of the mechanism is nearly unchanged regardless of a particular agent’s participation. Therefore, it is not clear that the adversary’s belief updates should concern an agent thinking about participating in the study, since her decision has little effect on this update.

Related to our paper, there are several papers investigating (and each proposing different models for) how rational agents should evaluate their costs for differential privacy [26, 12, 22, 4, 18]. We adopt the simplest and most conservative of these approaches, advocated by Nissim, et al. [22], and assume that agents costs are upper bounded by a linear function of $\epsilon$.

Alternatively, privacy (quantified by $\epsilon$) can be thought of as a fungible commodity, with the price discovered by a market. Li, et al. [17] consider how to set arbitrage-free prices for queries. Another line of papers [12, 11, 18, 2, 5, 23] consider how to discover the value of $\epsilon$ via an auction, when $\epsilon$ is set to be the largest value that the data analyst can afford.

11 Conclusion
We have proposed a simple economic model that enables users of differential privacy to choose the key parameters $\epsilon$ and $\delta$ in a principled way, based on quantities that can be estimated in practice. To the best of our knowledge, this is the first comprehensive model of its kind. We have applied our model in two case studies, and we have used it to explore the surprising observation that a private study can be cheaper than a non-private study with the same accuracy. We have discussed ways in which our model could be refined, but even in its current form the model provides useful guidance for practical applications of differential privacy.

12 References