We thank all reviewers for their reviews.

Most reviewers raise questions about the computation of weights. We note that we only show in the paper that the optimal weights can be computed, but it could be hard to find the closed form of the optimal weights. Here we provide some short clarifications and we will add more explanation in the next version of the paper to make it clear.

- For mean estimation, Lemma 3 and Claim 1 on page 3 show that finding the optimal weight vector c is equivalent to minimizing a function $v(h)$ of a single parameter $h \in [s_1, s_m]$. Optimizing this function of a single parameter can be simply done by setting the derivative to be 0 and considering locations where the derivative is not continuous. We will add more details to help readers.
- For ERM and quantile estimation, as stated in Line 194 on page 6 of the draft, we observe that minimizing the loss is similar to mean estimation and we can use the same method above to find the optimal weight vector.
- For linear regression with label privacy, Theorem 6 on page 7 states that the weight vector can be computed by minimizing a convex function. We can use a zero-order convex minimization algorithm to minimize this convex function.

Reviewer 1,3,4 point out that B is not defined in Section 3. We will add the definition.

Some response to other comments:

- Reviewer 2 “We need the number of data points of users public for weight computation”: You are correct. We will make it explicit in the paper.
- Reviewer 2 “The algorithm minimizes an upper bound on the variance of the estimator ”: You are correct. For example, for mean estimation, we are minimizing the variance of the estimators in the form as Algorithm 1.
- Reviewer 2 “This non-uniform weighting introduces some bias in the estimate”: In this work, we consider the case when user data is identical distributed. Even with non-uniform weights, for example in mean estimation, our estimator is unbiased. We agree with the reviewer that in a more complicated setting in which users have heterogeneous sample distributions, bias will be created. We will add discussions about such bias.
- Reviewer 3 “Line 142 s_h should be s_m”: You are correct.
- Reviewer 3 “Algorith 2, defintion of new loss l'”: l‘ is defined in line 179. We will make the defintion of l' clear in the pseudocode.
- Reviewer 3 “Why expected excess risk for Theorem 4”: The expected excess risk is defined for any convex Lipschitz loss functions. L2 norm in the mean estimation is a special case. We will add the discussion.
- Reviewer 3 “Figures on page 8 too small”: We will increase the size.
- Reviewer 3 “Lemma 3 proof, Case 1”: You are correct. We should swap the roles of p and q.

Finally, for related work, we thank reviewers for providing pointers to related work in label privacy and personalized/heterogeneous/group differential privacy. We will do a literature search from these pointers. We will add citations and provide relevant discussions.