Differential Privacy Logistic regression study

Maurice Lubetzki

This study is based on the mortality data set and on the theory of the paper:

<https://pdfs.semanticscholar.org/a188/d2ac0d10bdd4d4a04c92cdc76523e11c155c.pdf>

We are using L1 regularization and performing a logistic regression Differentially private versus Non-Differentially private.

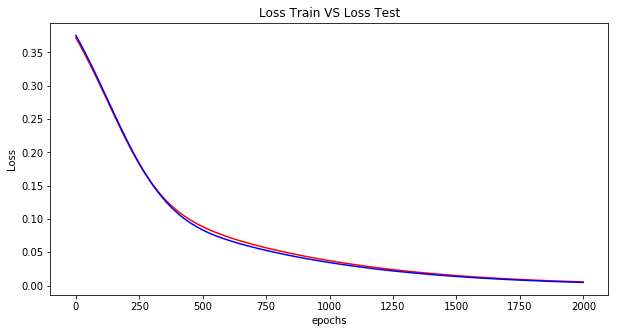
DP1 = Only add some noise to the classifier

DP2 = We add some noise to the Loss function we want to optimize

**Loss Function Study**

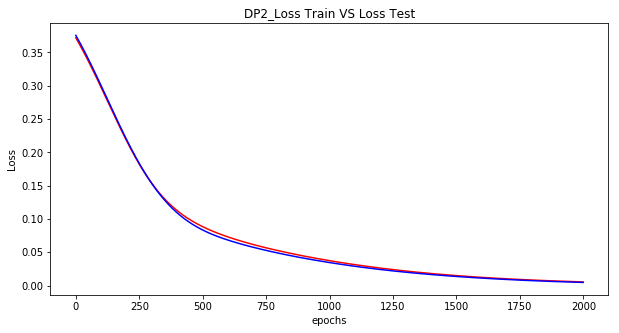
Without Differential privacy and with learning rate = 0.01

Evolution of the Loss function evaluated on the training set ( in Red ) VS test set ( In Blue) a



With differential privacy DP2, learning rate = 0.01

Evolution of the DP2- Loss function evaluated on the training set (in Red) VS test set (in Blue) at each iteration.

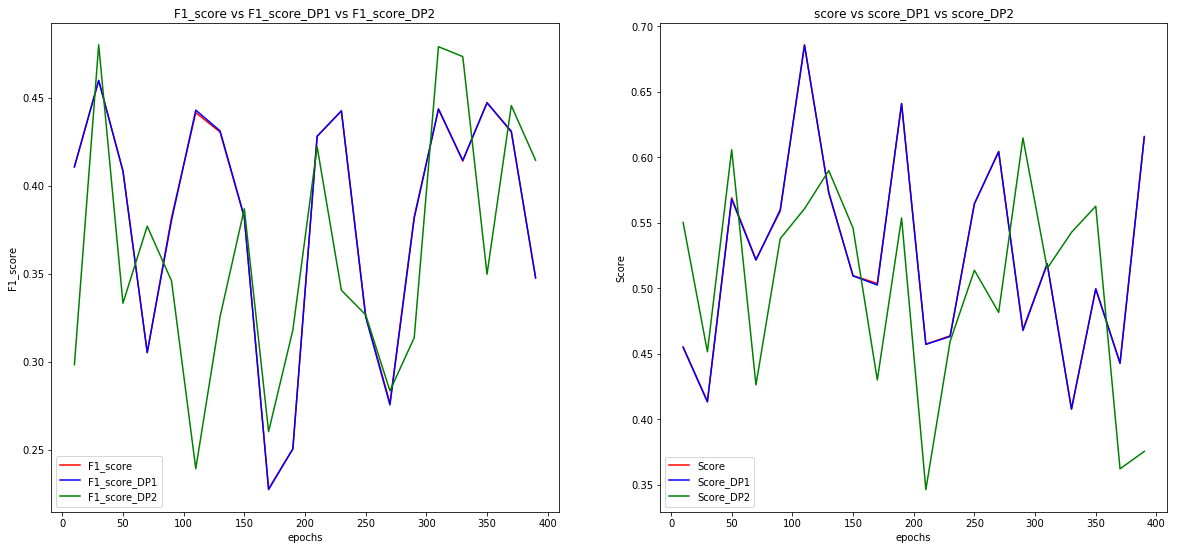


Loss function curves overlap perfectly between test and train for both DP and non-DP.

**F1-score and Accuracy study**:

Learning rate 0.01 privacy budget: 0.5

Epochs = iteration

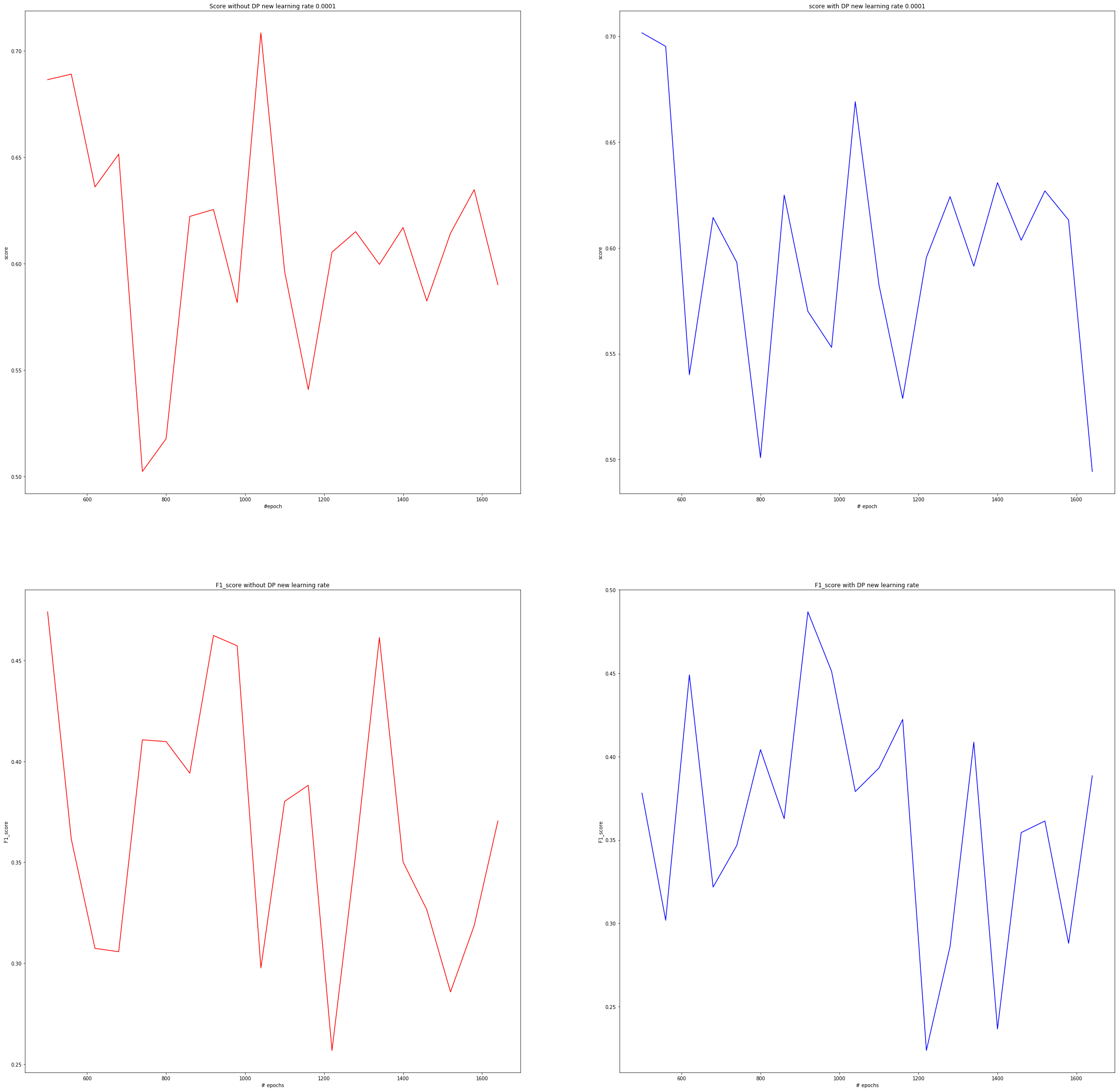


DP1 and non-DP curves overlap totally. In addition, DP2 and DP1 have approximately the same shape although there’s a slight increasing or decreasing phase between the two cases.

We can conclude that the behavior of the two DPs are almost the same whether it’s with the score metric or F1-score.

Same Analysis but with a different learning rate, number of iterations and privacy budget, only on F1-score.

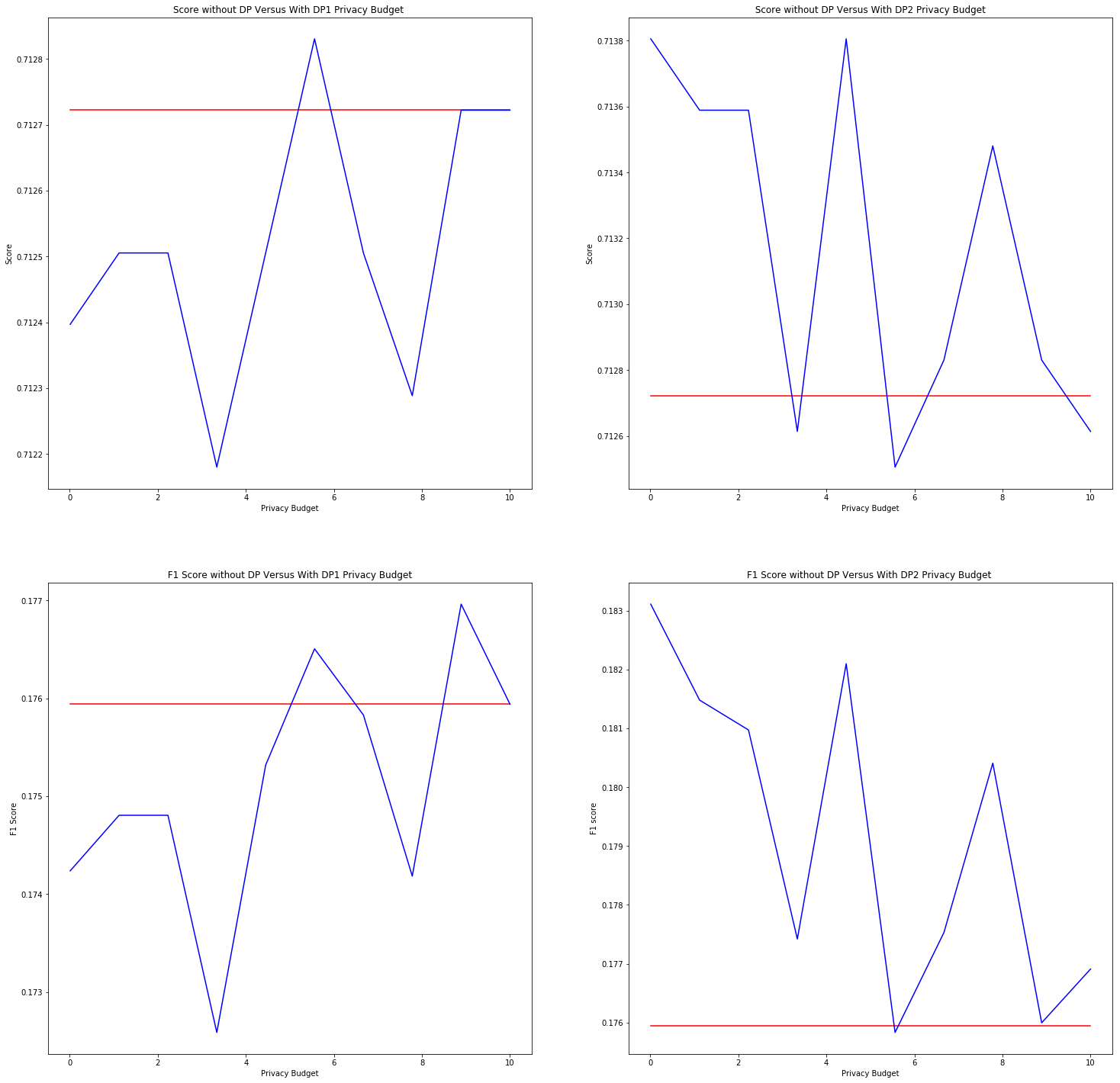
Learning rate 0.0001 / epoch = 500 – 1700/ Eps-0 DP 0.5 / DP1



The only remark we can make here is that F1 scores fluctuates as much with and without DP.

**Study performances of the model over the Privacy Budget:**

iterations = 150 learning rate 0.01

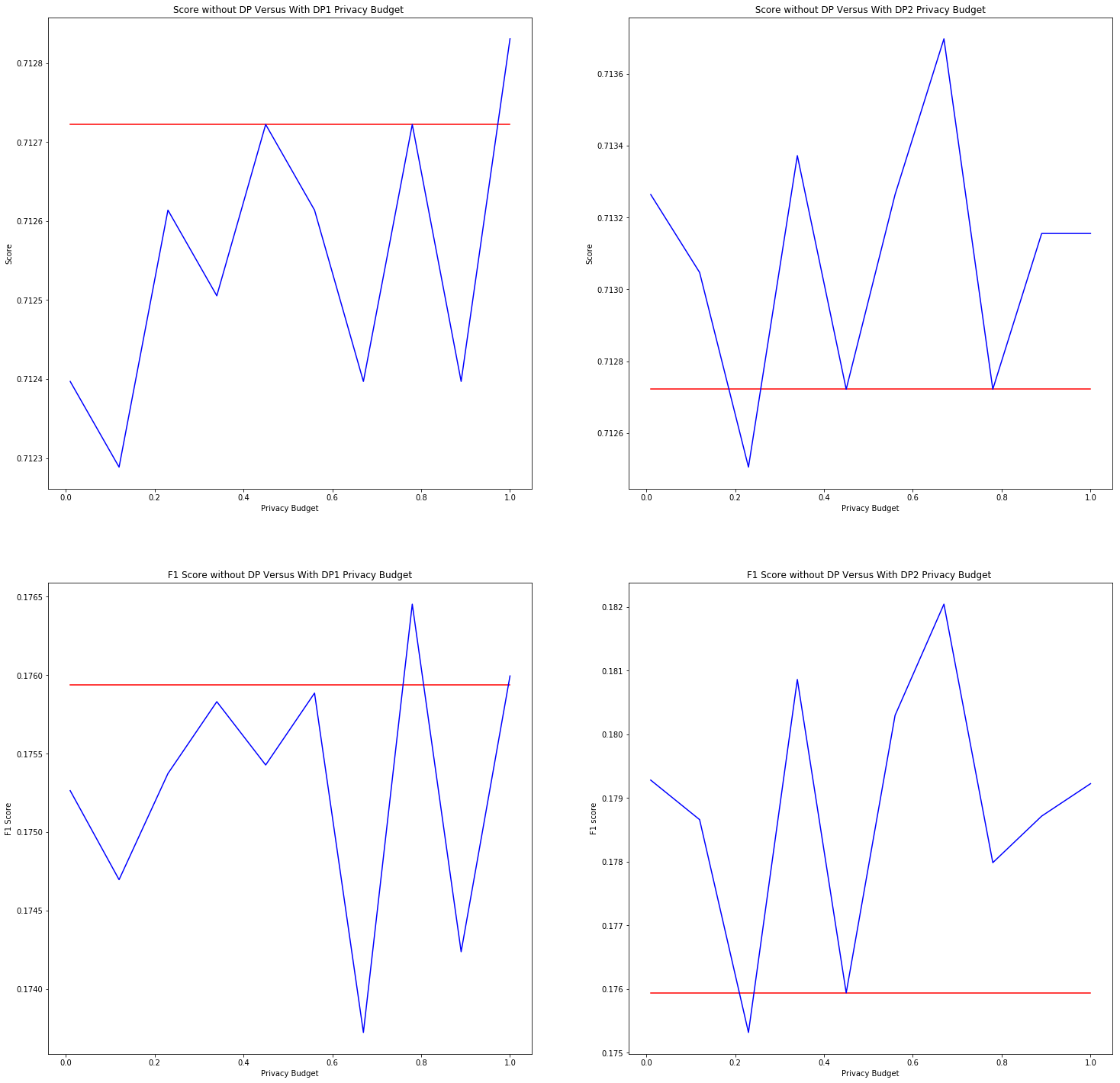


Overall, DP2 performances are always outperforming non-DP. Whereas, DP1 is mostly underperforming non-DP.

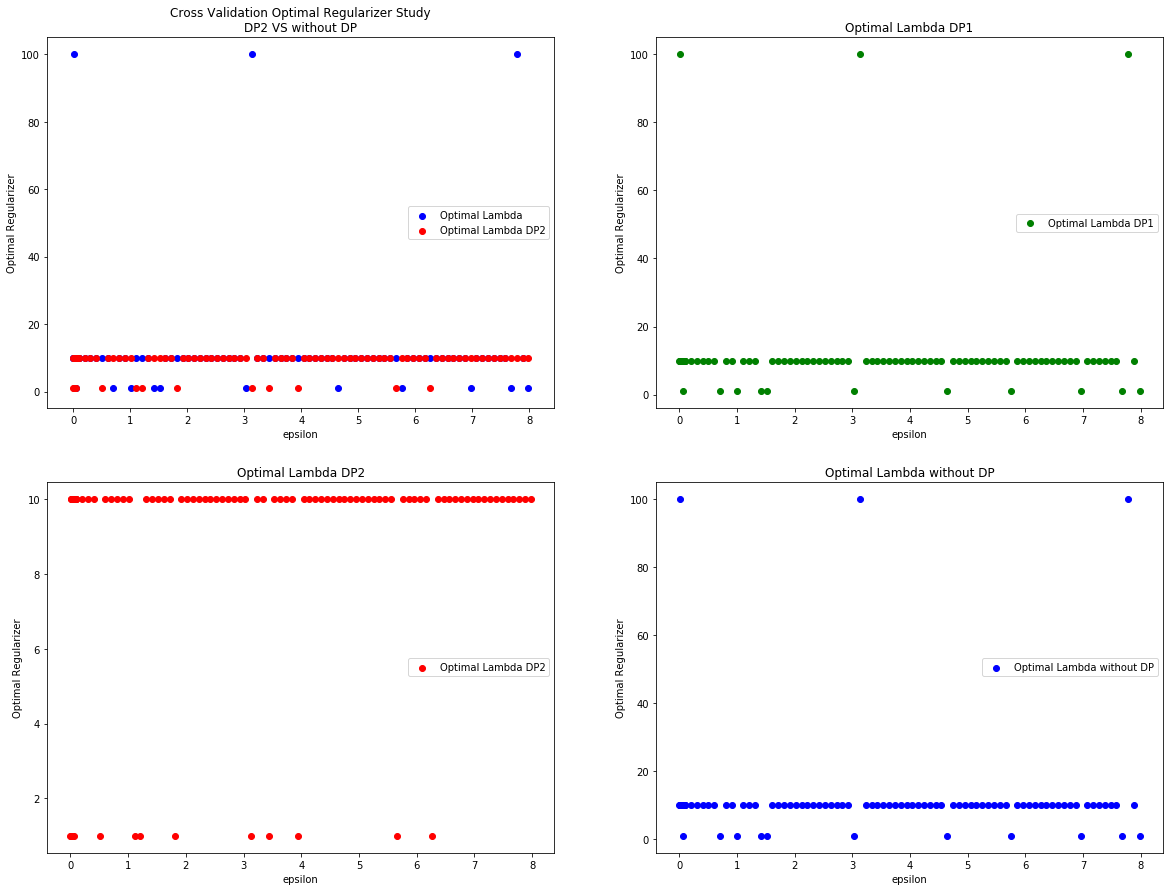
DP2 loss function offsets negative coefficients updates within the weight while performing gradient descent. It is an another regularizer.

Thus, DP2 is preferred over DP1 with this model.

epoch = 150 learning rate 0.5



Cross Validation Study over set of regularizer: [ 0.1, 1, 10, 100, 1000]

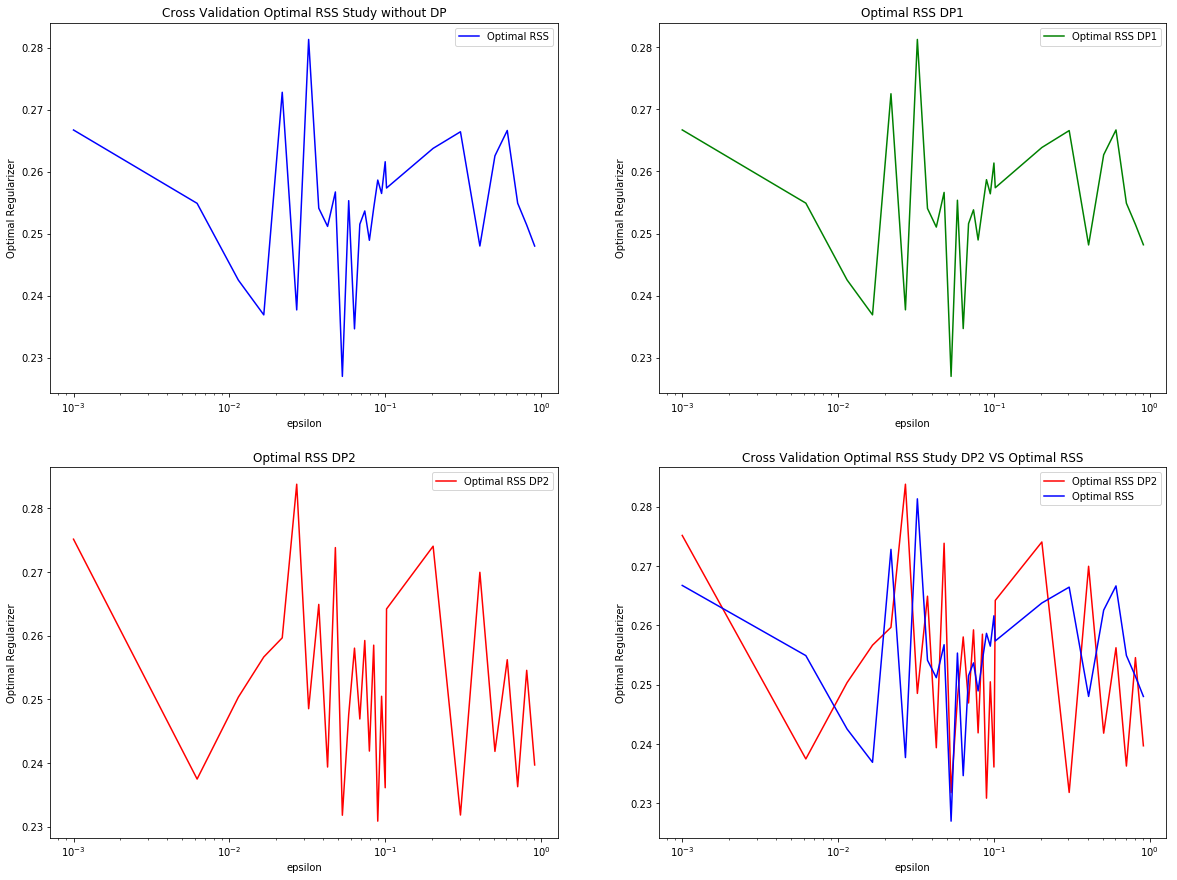


Usually, without DP and DP1 are behaving the same way regardless the privacy budget.

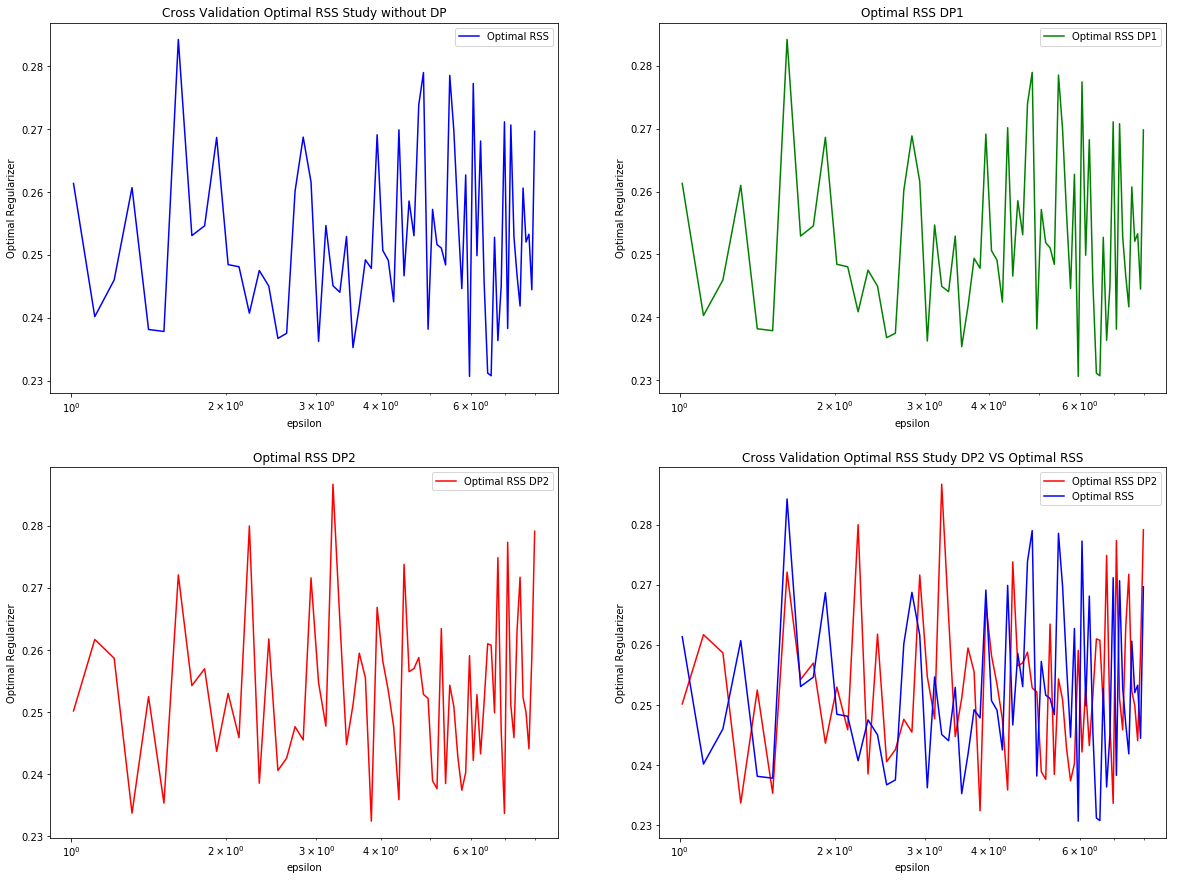
Without DP and DP2 are either the same or DP2’s have its optimal lambda less than the lambda without DP.

RSS study:

Restriction of privacy budget between 0.001 and 1.



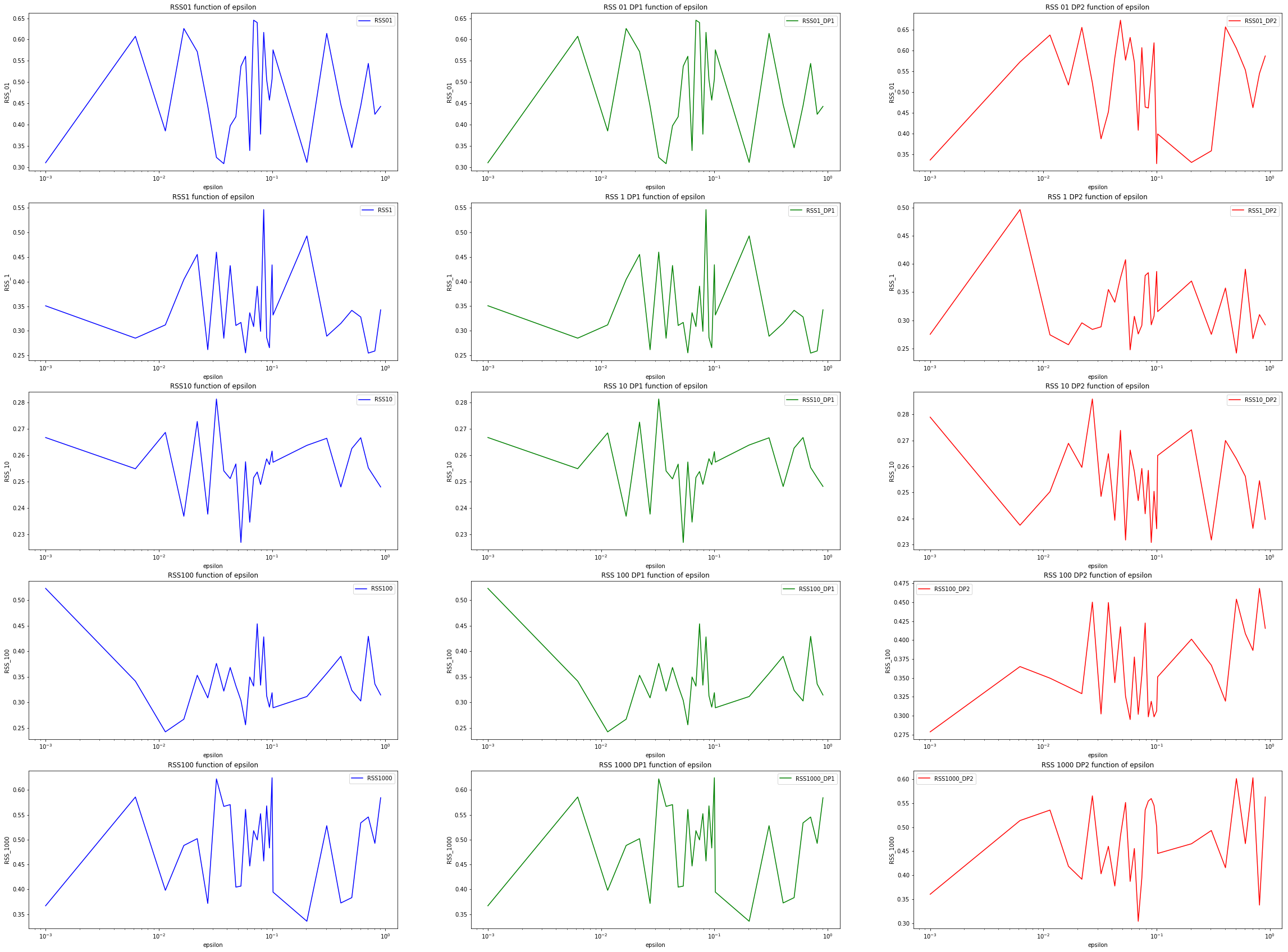
Same pattern without DP and DP2. Again



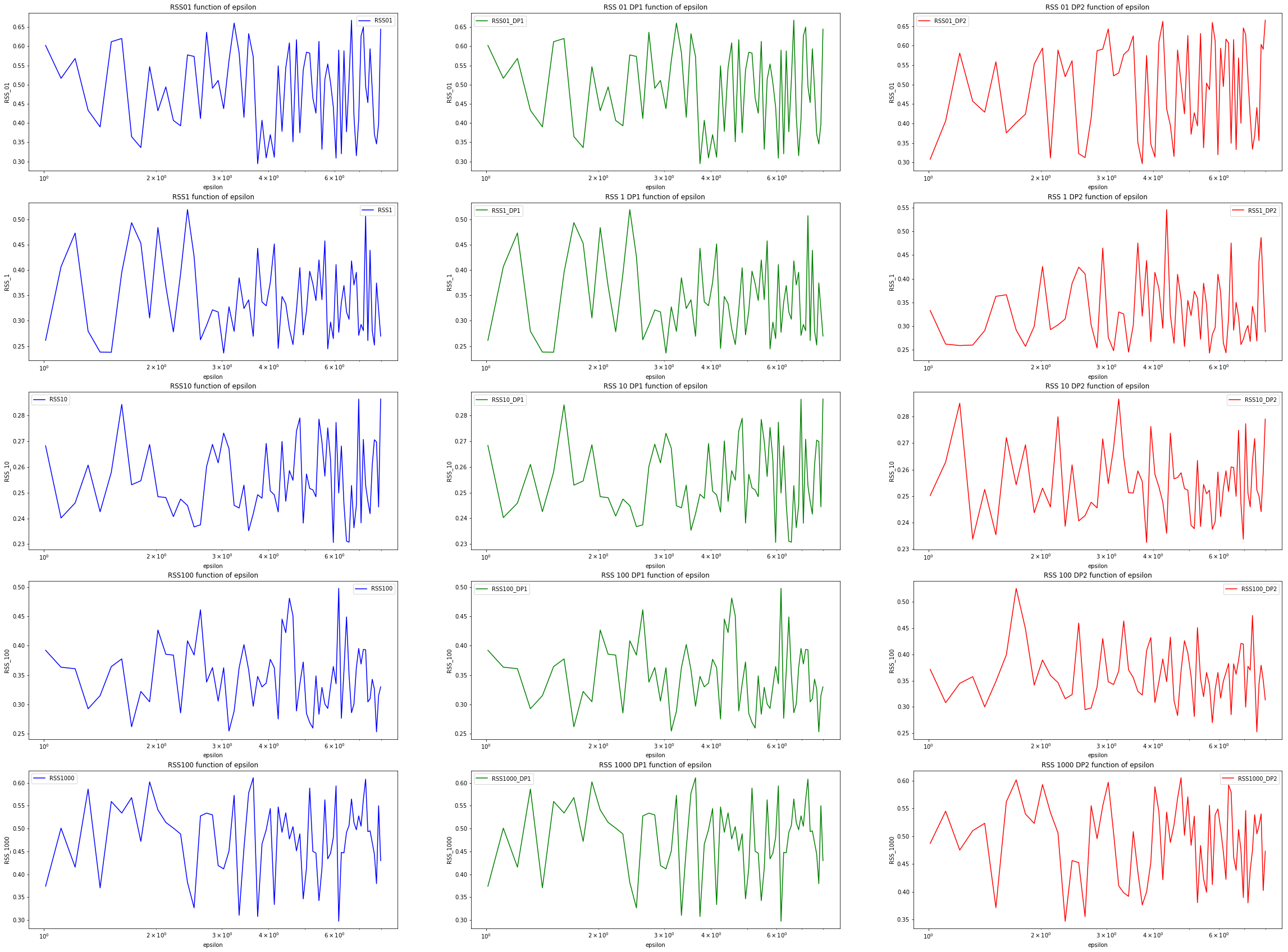
RSS function of privaty budget for different lambdas

DP1 and non DP have the exact same behavior.

Eps between 0.001 and 1. 0.001 and 0.01 sees their RSS fluctuates a lot. RSS(privacy budget ) is a very sensitive function, no robustness depicted here.



Same RSS study but for eps (a.k.a privacy budget) between 1 and 8

Again, the Less noise we are implementing (less privacy), the more fluctuation we have.