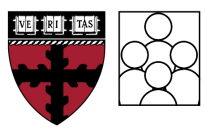
CS208: Applied Privacy for Data Science
Machine Learning under DP

James Honaker & Salil Vadhan

School of Engineering & Applied Sciences
Harvard University

April 12, 2019

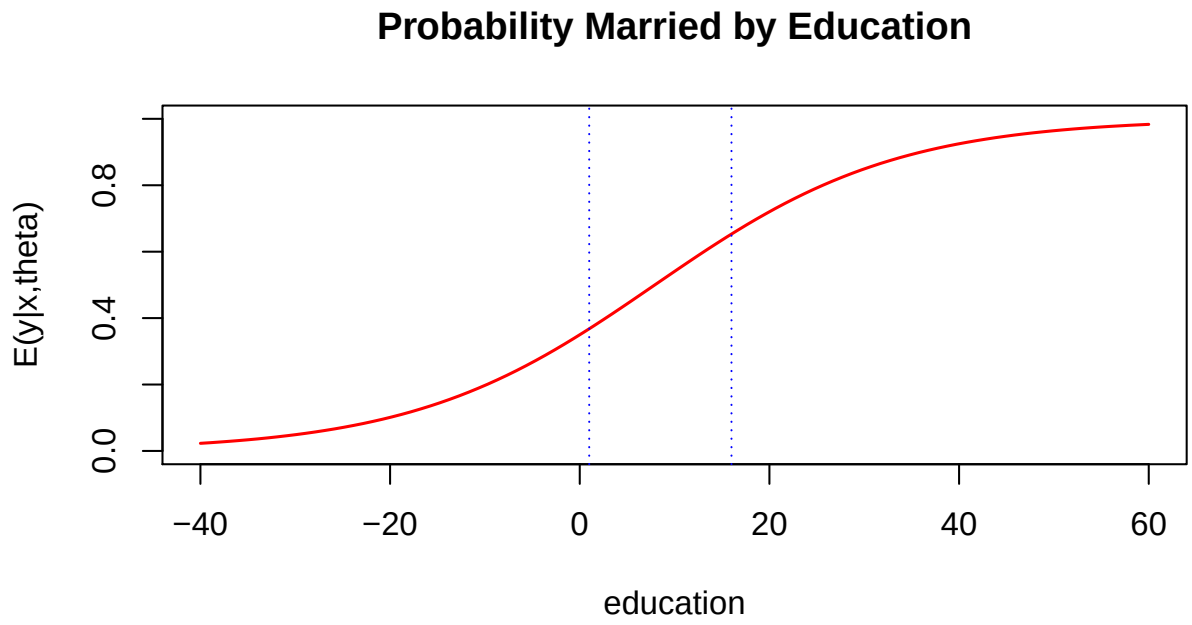


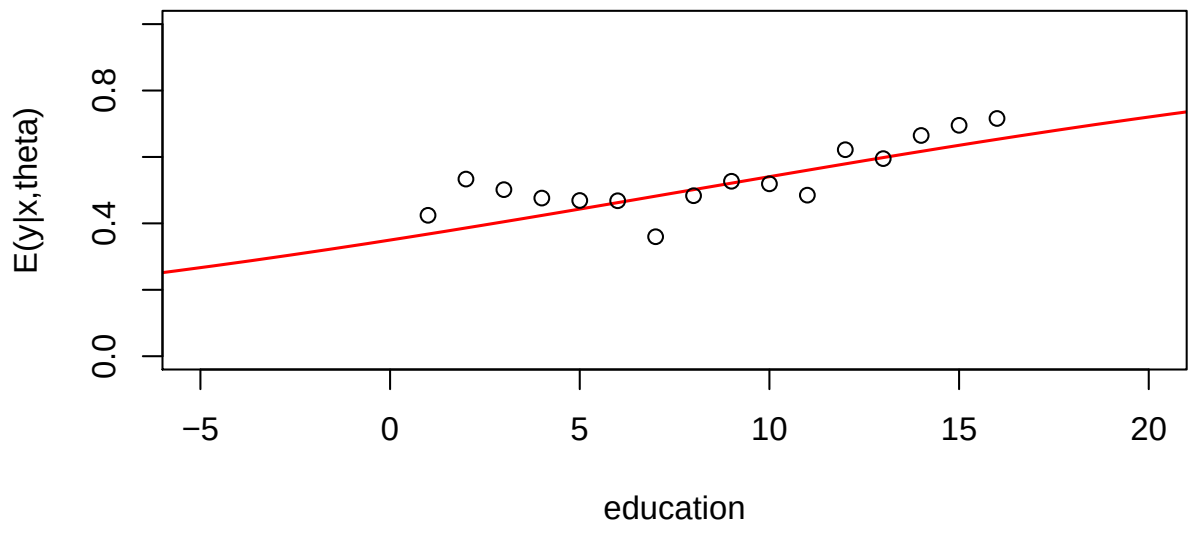
CRCS Center for Research on Computation and Society
at Harvard John A. Paulson School of Engineering and Applied Sciences

Logit Model

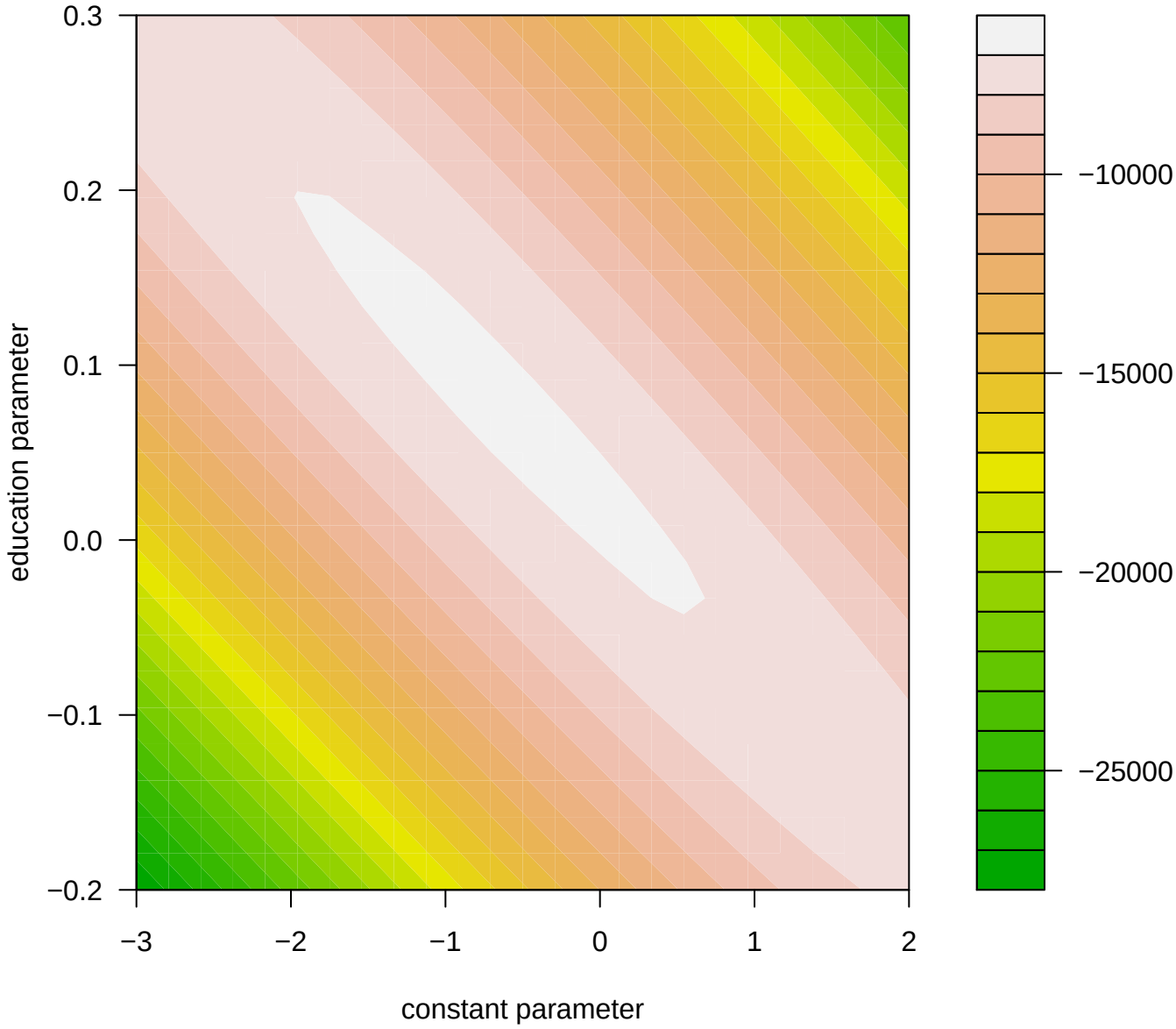
$$logL(y|x,\theta) = \sum_{i=1}^{N} y_i \log(\pi_i) + (1-y_i)\log(1-\pi_i),$$

$$\pi_i = \frac{1}{1+e^{-\beta_0-\beta_1 x_i}}.$$


Probability Married by Education
E(y|x,theta)
0.0
0.4
0.8
−40
−20
0
20
40
60
education


E(y|x,theta)
0.0
0.4
0.8
−5
0
5
10
15
20
education

logLikelihood surface



Algorithm 1 Differentially private SGD (Outline)

Input: Examples $\{x_1, \ldots, x_N\}$, loss function $\mathcal{L}(\theta) = \frac{1}{N} \sum_i \mathcal{L}(\theta, x_i)$. Parameters: learning rate η_t, noise scale σ, group size L, gradient norm bound C.

Initialize θ_0 randomly

for $t \in [T]$ **do**

Take a random sample L_t with sampling probability L/N

Compute gradient

For each $i \in L_t$, compute $\mathbf{g}_t(x_i) \leftarrow \nabla_{\theta_t} \mathcal{L}(\theta_t, x_i)$

Clip gradient

$\bar{\mathbf{g}}_t(x_i) \leftarrow \mathbf{g}_t(x_i) / \max\left(1, \frac{\|\mathbf{g}_t(x_i)\|_2}{C}\right)$

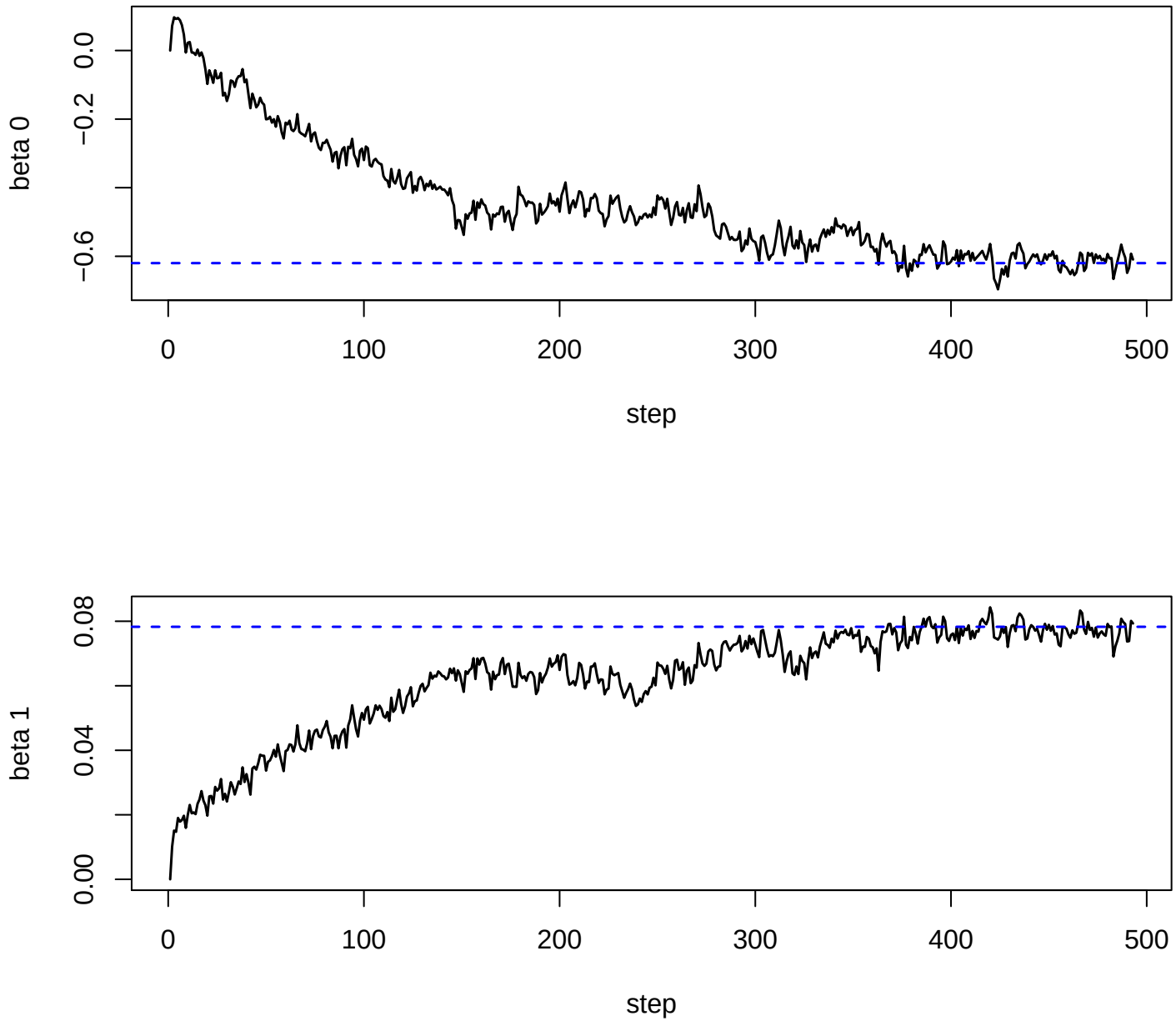
Add noise

$\tilde{\mathbf{g}}_t \leftarrow \frac{1}{L}\left(\sum_i \bar{\mathbf{g}}_t(x_i) + \mathcal{N}(0, \sigma^2 C^2 \mathbf{I})\right)$

Descent

$\theta_{t+1} \leftarrow \theta_t - \eta_t \tilde{\mathbf{g}}_t$

Output θ_T and compute the overall privacy cost (ε, δ) using a privacy accounting method.


beta 0
0.0
−0.2
−0.6
0
100
200
300
400
500
step
beta 1
0.08
0.04
0.00
0
100
200
300
400
500
step

Installing tensorflow/privacy

```
## Tensorflow privacy install
git clone -n
https://github.com/tensorflow/privacy.git
cd privacy
# There appears to be a breaking commit Thursday
afternoon!
git checkout
e8113a03658c40eeb8ad6d9cbb3d165848b9d68a
pip install tensorflow
pip install -e .
```