Simulation 3

Mengqi Liu

Jul 17， 2023

Setting

- d: number of nodes
- $s0$: expected number of edges
- graph_type: ER
    - ER: Erdős-Rényi Graph randomly choose one from all graphs with d nodes and $s0$ edges + random orientation
- n: number of samples, n=inf mimics population risk

- sem_type: gauss, exp, gumbel, uniform, logistic, poisson
- loss_type: l2, logistic
- lambda1: penalty for sparsity
- Here I use MCMC to simulate for 20 times.
- losses:
    - $h(W) = 0$: constraint function
    - loss_est: loss decided by loss_type
    - loss_l1: loss_est + l1 penalty
    - obj_new: + quadratic penalty of h
    - obj_dual: dual function pertaining to obj_new with Lagrange multiplier α

Results

Gains during learning

- The algorithm runs too slow when lambda1 is too small and the estimates are too coarse when lambda1 is too large.
- When considering discrete distributions like poisson and logistic, we may choose loss_type to be 'logistic' or 'poisson' to achieve better performance (in terms of fdr, tpr, fpr, shd, nnz and hamming distance). However, the performance is not ideal (details in later).

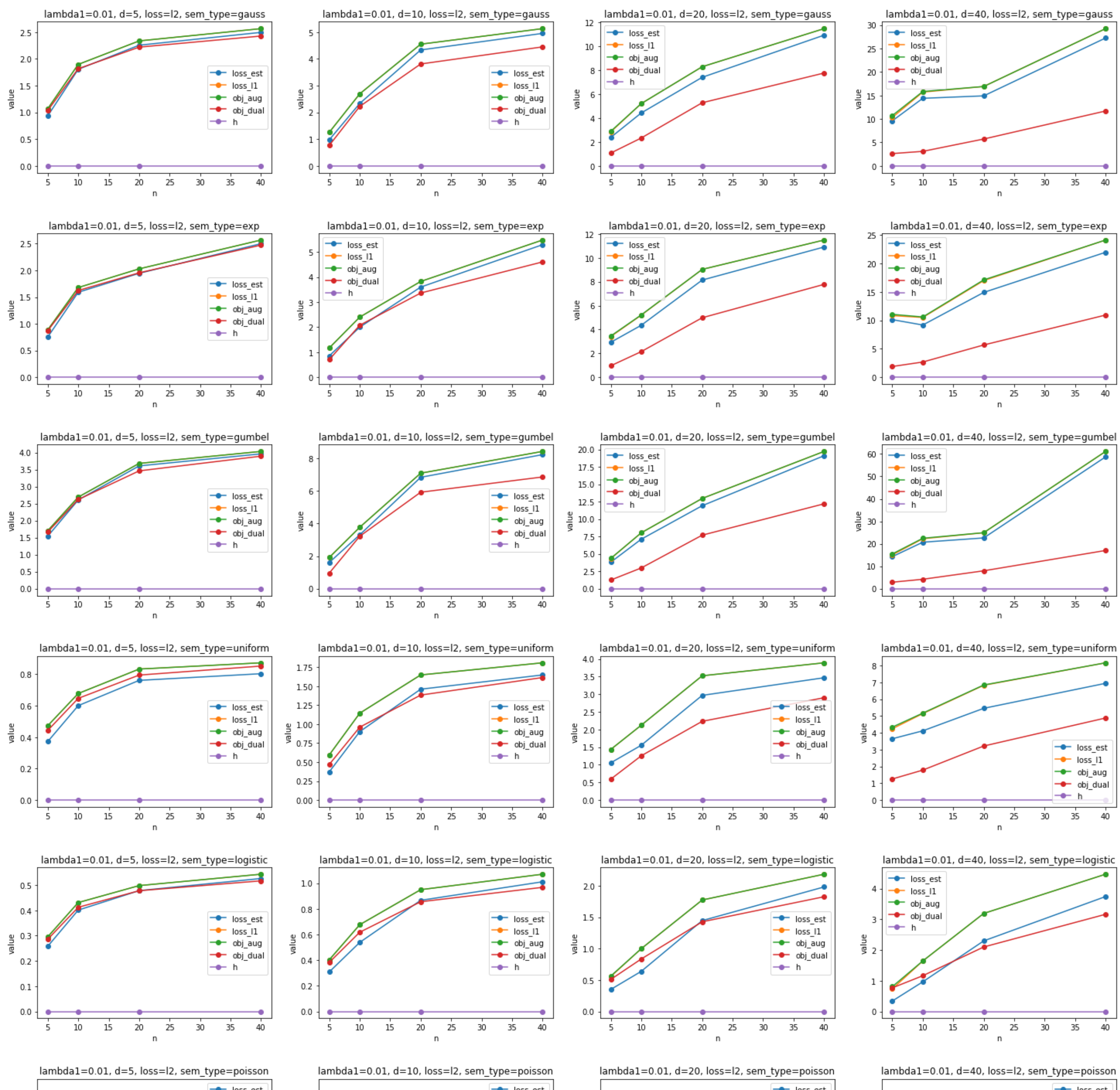
Experiment 1

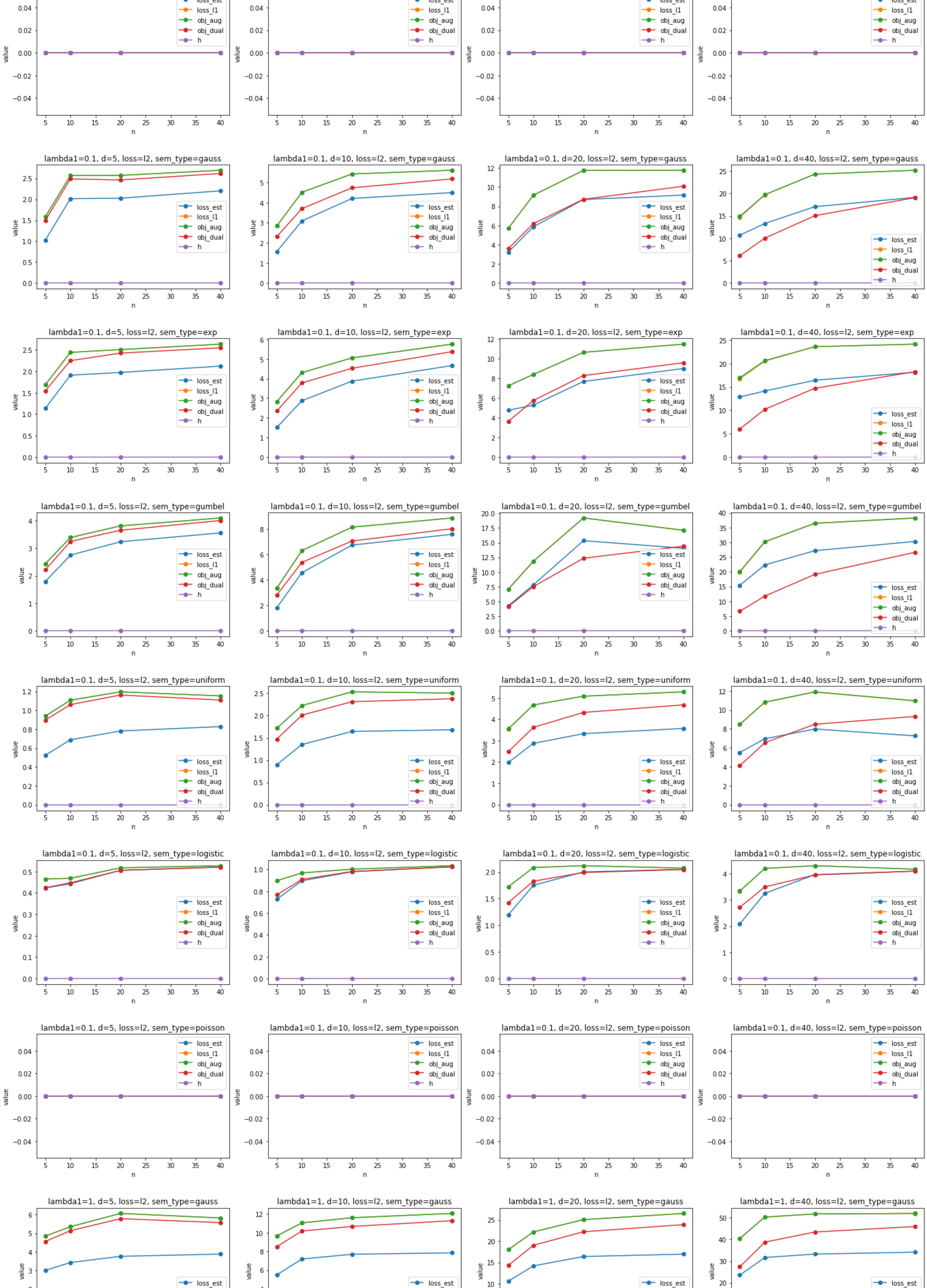
Here I choose $d \in \{5, 10, 20, 40\}$, $s0 = d$, lambda1$\in \{0.01, 0.1, 1, 5\}$ and compute losses with different $n \in \{2, 5, 7, 10, 15, 20\}$.

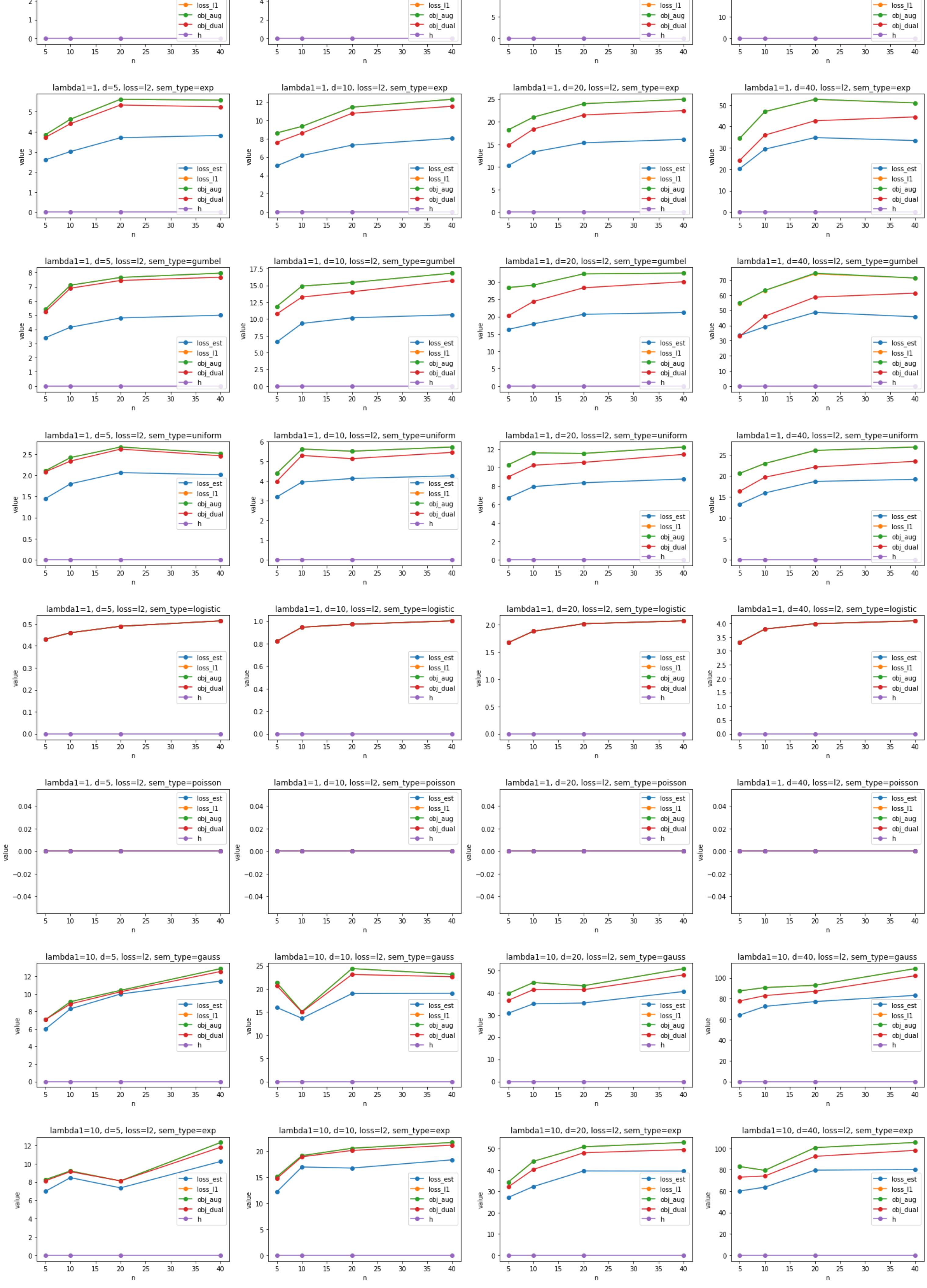
Gains:

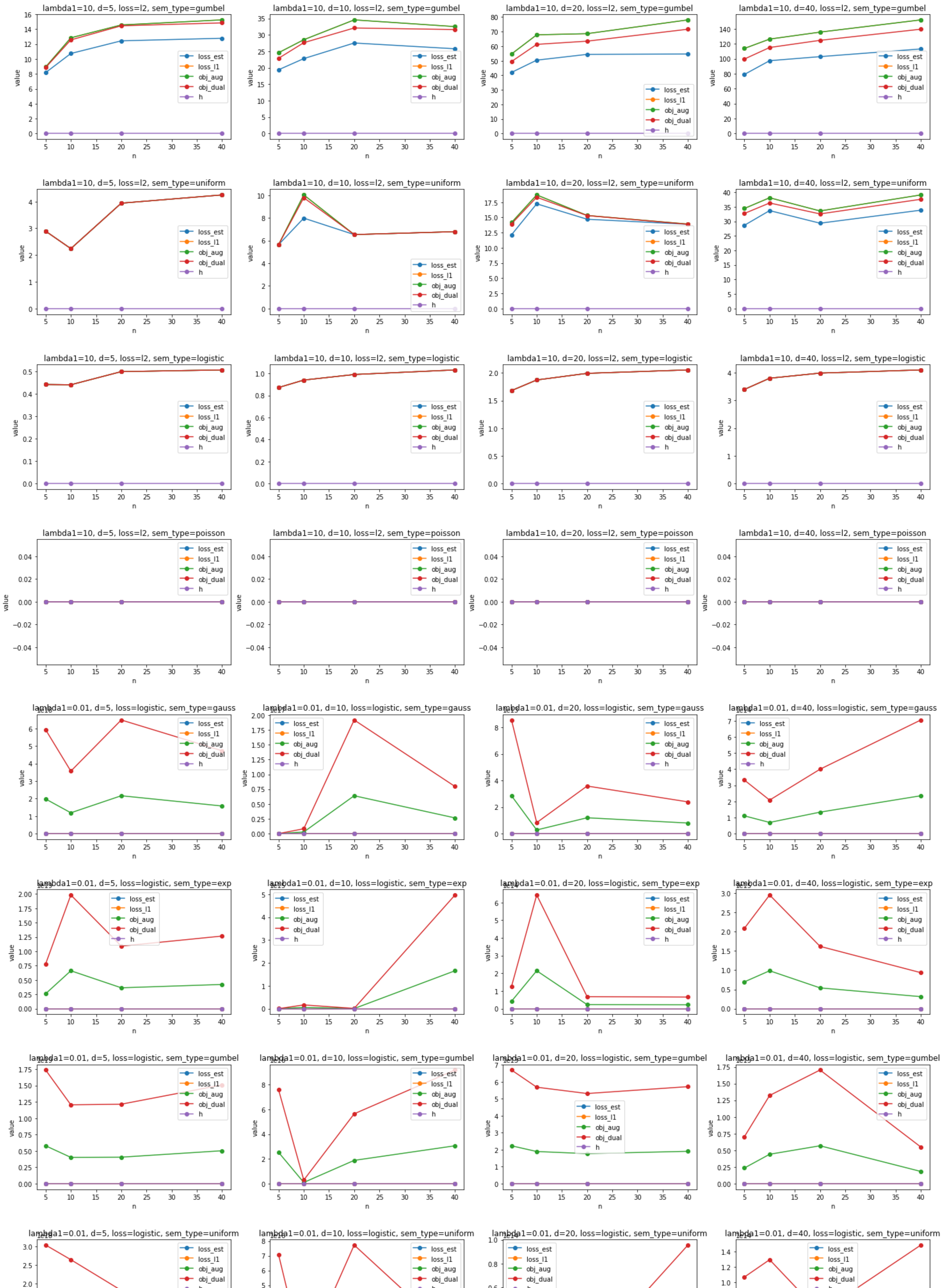
- obj_aug $\approx$ loss_l1 (h=0, constaint function is satisfied)
- obj_dual $\gtrapprox$ loss_est

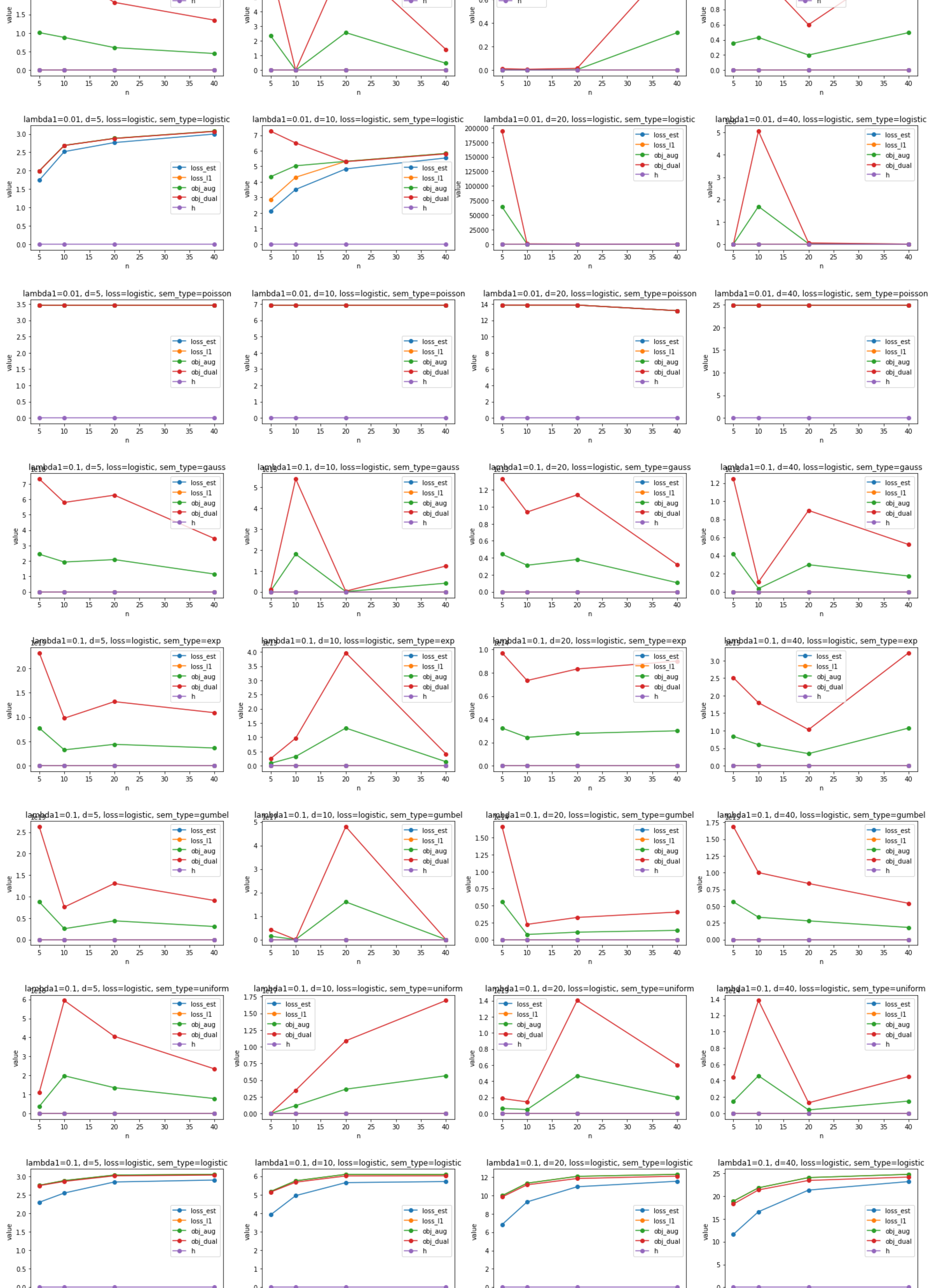
- When loss_type = l2,
  - Generally obj_aug $>$ obj_dual $\rightarrow$ **not strong duality**
  - Value of obj_aug-obj_dual will increase as d increases, but basically does not increase as lambda1 increases $\rightarrow$ **The optimal solution is robust to lambda1, but not to d**
  - The loss_est keeps increasing when lambda1 increases $\rightarrow$ **Small lambda1 may lead to better optimal solution as well as sparse optimal solution**
  - The losses keep decreasing when n decreases $\rightarrow$ **risk of overfitting**
  - Although the learning curves are normal for discrete distributions (poisson and logistic), but we should always be aware that the algorithm with loss_type = l2 performs very poorly under these two distributions
- When loss_type = logistic,
  - Curves are only normal for logistic distribution when lambda1 is not too large or too small $\rightarrow$ **logistic loss is only suitable for logistic distribution and is sensitive to the choose of lambda1**

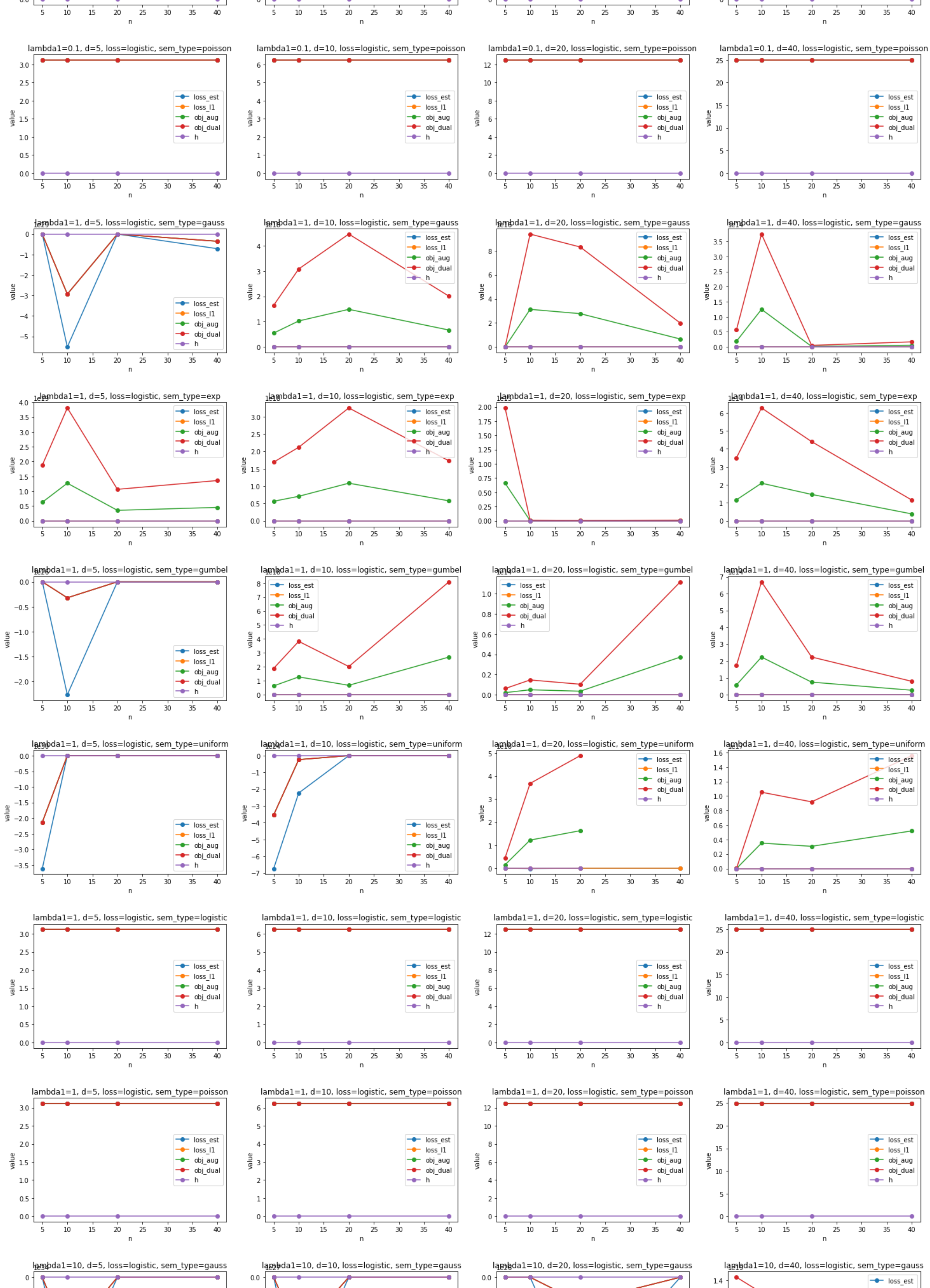



lambda1=0.1, d=5, loss=l2, sem_type=gauss
lambda1=0.1, d=10, loss=l2, sem_type=gauss
lambda1=0.1, d=20, loss=l2, sem_type=gauss
lambda1=0.1, d=40, loss=l2, sem_type=gauss
lambda1=0.1, d=5, loss=l2, sem_type=exp
lambda1=0.1, d=10, loss=l2, sem_type=exp
lambda1=0.1, d=20, loss=l2, sem_type=exp
lambda1=0.1, d=40, loss=l2, sem_type=exp
lambda1=0.1, d=5, loss=l2, sem_type=gumbel
lambda1=0.1, d=10, loss=l2, sem_type=gumbel
lambda1=0.1, d=20, loss=l2, sem_type=gumbel
lambda1=0.1, d=40, loss=l2, sem_type=gumbel
lambda1=0.1, d=5, loss=l2, sem_type=uniform
lambda1=0.1, d=10, loss=l2, sem_type=uniform
lambda1=0.1, d=20, loss=l2, sem_type=uniform
lambda1=0.1, d=40, loss=l2, sem_type=uniform
lambda1=0.1, d=5, loss=l2, sem_type=logistic
lambda1=0.1, d=10, loss=l2, sem_type=logistic
lambda1=0.1, d=20, loss=l2, sem_type=logistic
lambda1=0.1, d=40, loss=l2, sem_type=logistic
lambda1=0.1, d=5, loss=l2, sem_type=poisson
lambda1=0.1, d=10, loss=l2, sem_type=poisson
lambda1=0.1, d=20, loss=l2, sem_type=poisson
lambda1=0.1, d=40, loss=l2, sem_type=poisson
lambda1=1, d=5, loss=l2, sem_type=gauss
lambda1=1, d=10, loss=l2, sem_type=gauss
lambda1=1, d=20, loss=l2, sem_type=gauss
lambda1=1, d=40, loss=l2, sem_type=gauss
loss_est
loss_l1
obj_aug
obj_dual
h
value
n

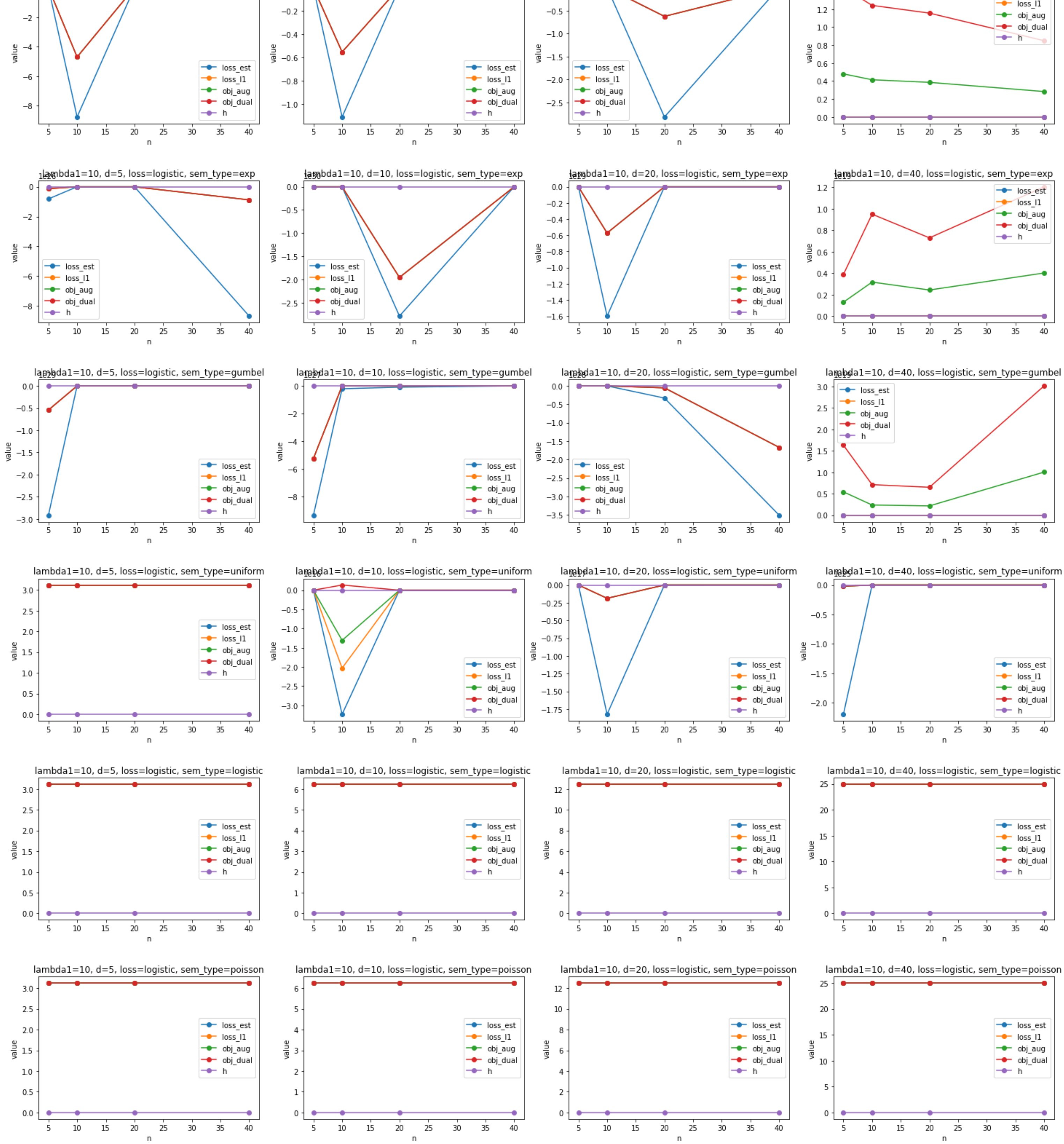

lambda1=1, d=5, loss=l2, sem_type=exp
lambda1=1, d=10, loss=l2, sem_type=exp
lambda1=1, d=20, loss=l2, sem_type=exp
lambda1=1, d=40, loss=l2, sem_type=exp
lambda1=1, d=5, loss=l2, sem_type=gumbel
lambda1=1, d=10, loss=l2, sem_type=gumbel
lambda1=1, d=20, loss=l2, sem_type=gumbel
lambda1=1, d=40, loss=l2, sem_type=gumbel
lambda1=1, d=5, loss=l2, sem_type=uniform
lambda1=1, d=10, loss=l2, sem_type=uniform
lambda1=1, d=20, loss=l2, sem_type=uniform
lambda1=1, d=40, loss=l2, sem_type=uniform
lambda1=1, d=5, loss=l2, sem_type=logistic
lambda1=1, d=10, loss=l2, sem_type=logistic
lambda1=1, d=20, loss=l2, sem_type=logistic
lambda1=1, d=40, loss=l2, sem_type=logistic
lambda1=1, d=5, loss=l2, sem_type=poisson
lambda1=1, d=10, loss=l2, sem_type=poisson
lambda1=1, d=20, loss=l2, sem_type=poisson
lambda1=1, d=40, loss=l2, sem_type=poisson
lambda1=10, d=5, loss=l2, sem_type=gauss
lambda1=10, d=10, loss=l2, sem_type=gauss
lambda1=10, d=20, loss=l2, sem_type=gauss
lambda1=10, d=40, loss=l2, sem_type=gauss
lambda1=10, d=5, loss=l2, sem_type=exp
lambda1=10, d=10, loss=l2, sem_type=exp
lambda1=10, d=20, loss=l2, sem_type=exp
lambda1=10, d=40, loss=l2, sem_type=exp
loss_est
loss_l1
obj_aug
obj_dual
h
value
n


lambda1=10, d=5, loss=l2, sem_type=gumbel
lambda1=10, d=10, loss=l2, sem_type=gumbel
lambda1=10, d=20, loss=l2, sem_type=gumbel
lambda1=10, d=40, loss=l2, sem_type=gumbel
lambda1=10, d=5, loss=l2, sem_type=uniform
lambda1=10, d=10, loss=l2, sem_type=uniform
lambda1=10, d=20, loss=l2, sem_type=uniform
lambda1=10, d=40, loss=l2, sem_type=uniform
lambda1=10, d=5, loss=l2, sem_type=logistic
lambda1=10, d=10, loss=l2, sem_type=logistic
lambda1=10, d=20, loss=l2, sem_type=logistic
lambda1=10, d=40, loss=l2, sem_type=logistic
lambda1=10, d=5, loss=l2, sem_type=poisson
lambda1=10, d=10, loss=l2, sem_type=poisson
lambda1=10, d=20, loss=l2, sem_type=poisson
lambda1=10, d=40, loss=l2, sem_type=poisson
lambda1=0.01, d=5, loss=logistic, sem_type=gauss
lambda1=0.01, d=10, loss=logistic, sem_type=gauss
lambda1=0.01, d=20, loss=logistic, sem_type=gauss
lambda1=0.01, d=40, loss=logistic, sem_type=gauss
lambda1=0.01, d=5, loss=logistic, sem_type=exp
lambda1=0.01, d=10, loss=logistic, sem_type=exp
lambda1=0.01, d=20, loss=logistic, sem_type=exp
lambda1=0.01, d=40, loss=logistic, sem_type=exp
lambda1=0.01, d=5, loss=logistic, sem_type=gumbel
lambda1=0.01, d=10, loss=logistic, sem_type=gumbel
lambda1=0.01, d=20, loss=logistic, sem_type=gumbel
lambda1=0.01, d=40, loss=logistic, sem_type=gumbel
lambda1=0.01, d=5, loss=logistic, sem_type=uniform
lambda1=0.01, d=10, loss=logistic, sem_type=uniform
lambda1=0.01, d=20, loss=logistic, sem_type=uniform
lambda1=0.01, d=40, loss=logistic, sem_type=uniform
loss_est
loss_l1
obj_aug
obj_dual
h
value
n


lambda1=0.01, d=5, loss=logistic, sem_type=logistic
lambda1=0.01, d=10, loss=logistic, sem_type=logistic
lambda1=0.01, d=20, loss=logistic, sem_type=logistic
lambda1=0.01, d=40, loss=logistic, sem_type=logistic
lambda1=0.01, d=5, loss=logistic, sem_type=poisson
lambda1=0.01, d=10, loss=logistic, sem_type=poisson
lambda1=0.01, d=20, loss=logistic, sem_type=poisson
lambda1=0.01, d=40, loss=logistic, sem_type=poisson
lambda1=0.1, d=5, loss=logistic, sem_type=gauss
lambda1=0.1, d=10, loss=logistic, sem_type=gauss
lambda1=0.1, d=20, loss=logistic, sem_type=gauss
lambda1=0.1, d=40, loss=logistic, sem_type=gauss
lambda1=0.1, d=5, loss=logistic, sem_type=exp
lambda1=0.1, d=10, loss=logistic, sem_type=exp
lambda1=0.1, d=20, loss=logistic, sem_type=exp
lambda1=0.1, d=40, loss=logistic, sem_type=exp
lambda1=0.1, d=5, loss=logistic, sem_type=gumbel
lambda1=0.1, d=10, loss=logistic, sem_type=gumbel
lambda1=0.1, d=20, loss=logistic, sem_type=gumbel
lambda1=0.1, d=40, loss=logistic, sem_type=gumbel
lambda1=0.1, d=5, loss=logistic, sem_type=uniform
lambda1=0.1, d=10, loss=logistic, sem_type=uniform
lambda1=0.1, d=20, loss=logistic, sem_type=uniform
lambda1=0.1, d=40, loss=logistic, sem_type=uniform
lambda1=0.1, d=5, loss=logistic, sem_type=logistic
lambda1=0.1, d=10, loss=logistic, sem_type=logistic
lambda1=0.1, d=20, loss=logistic, sem_type=logistic
lambda1=0.1, d=40, loss=logistic, sem_type=logistic
value
n
loss_est
loss_l1
obj_aug
obj_dual
h

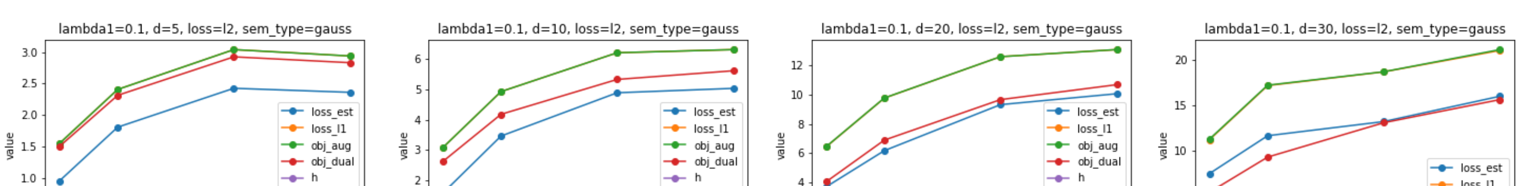


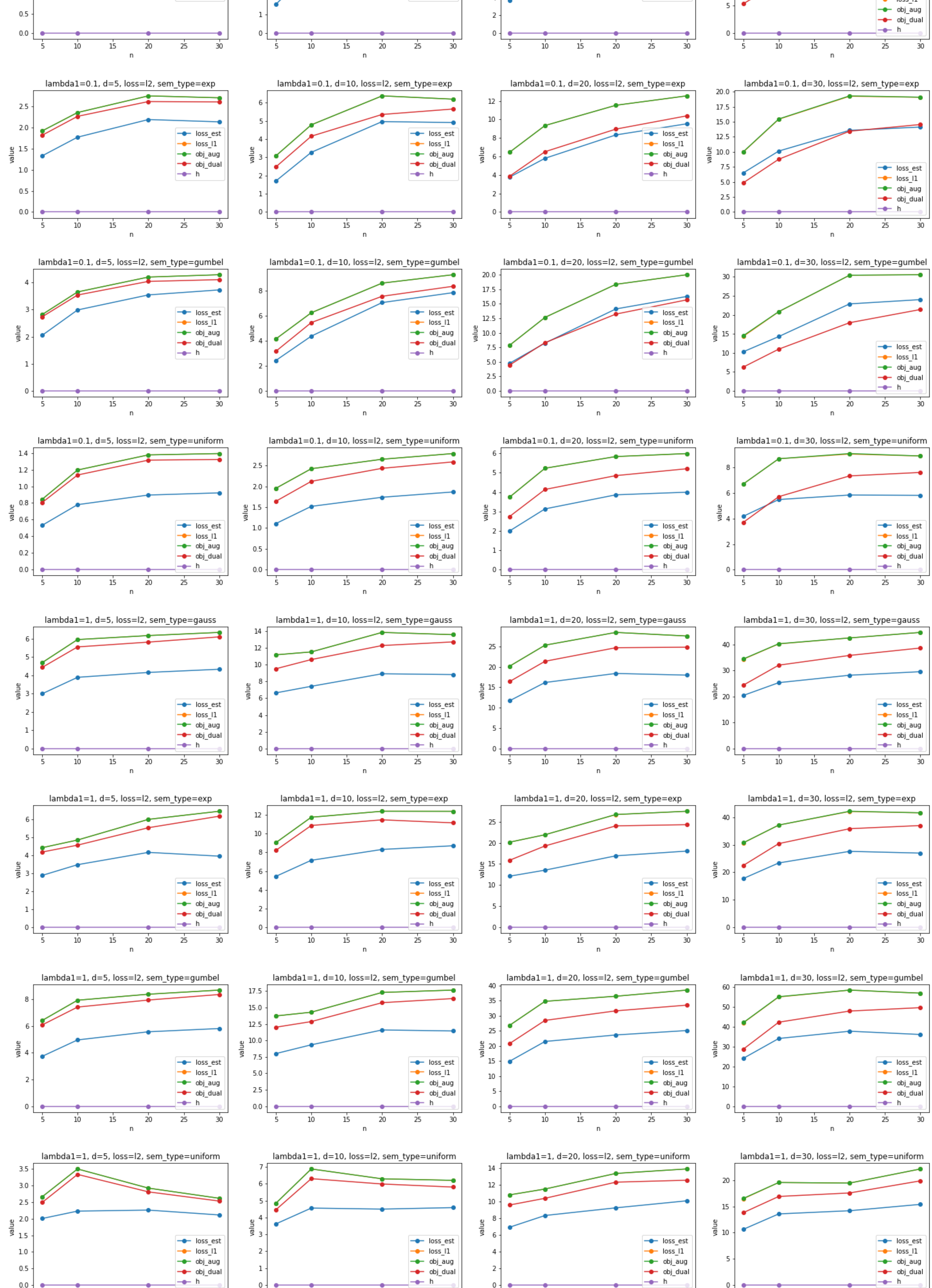
lambda1=0.1, d=5, loss=logistic, sem_type=poisson
lambda1=0.1, d=10, loss=logistic, sem_type=poisson
lambda1=0.1, d=20, loss=logistic, sem_type=poisson
lambda1=0.1, d=40, loss=logistic, sem_type=poisson
lambda1=1, d=5, loss=logistic, sem_type=gauss
lambda1=1, d=10, loss=logistic, sem_type=gauss
lambda1=1, d=20, loss=logistic, sem_type=gauss
lambda1=1, d=40, loss=logistic, sem_type=gauss
lambda1=1, d=5, loss=logistic, sem_type=exp
lambda1=1, d=10, loss=logistic, sem_type=exp
lambda1=1, d=20, loss=logistic, sem_type=exp
lambda1=1, d=40, loss=logistic, sem_type=exp
lambda1=1, d=5, loss=logistic, sem_type=gumbel
lambda1=1, d=10, loss=logistic, sem_type=gumbel
lambda1=1, d=20, loss=logistic, sem_type=gumbel
lambda1=1, d=40, loss=logistic, sem_type=gumbel
lambda1=1, d=5, loss=logistic, sem_type=uniform
lambda1=1, d=10, loss=logistic, sem_type=uniform
lambda1=1, d=20, loss=logistic, sem_type=uniform
lambda1=1, d=40, loss=logistic, sem_type=uniform
lambda1=1, d=5, loss=logistic, sem_type=logistic
lambda1=1, d=10, loss=logistic, sem_type=logistic
lambda1=1, d=20, loss=logistic, sem_type=logistic
lambda1=1, d=40, loss=logistic, sem_type=logistic
lambda1=1, d=5, loss=logistic, sem_type=poisson
lambda1=1, d=10, loss=logistic, sem_type=poisson
lambda1=1, d=20, loss=logistic, sem_type=poisson
lambda1=1, d=40, loss=logistic, sem_type=poisson
lambda1=10, d=5, loss=logistic, sem_type=gauss
lambda1=10, d=10, loss=logistic, sem_type=gauss
lambda1=10, d=20, loss=logistic, sem_type=gauss
lambda1=10, d=40, loss=logistic, sem_type=gauss
loss_est
loss_l1
obj_aug
obj_dual
h
value
n

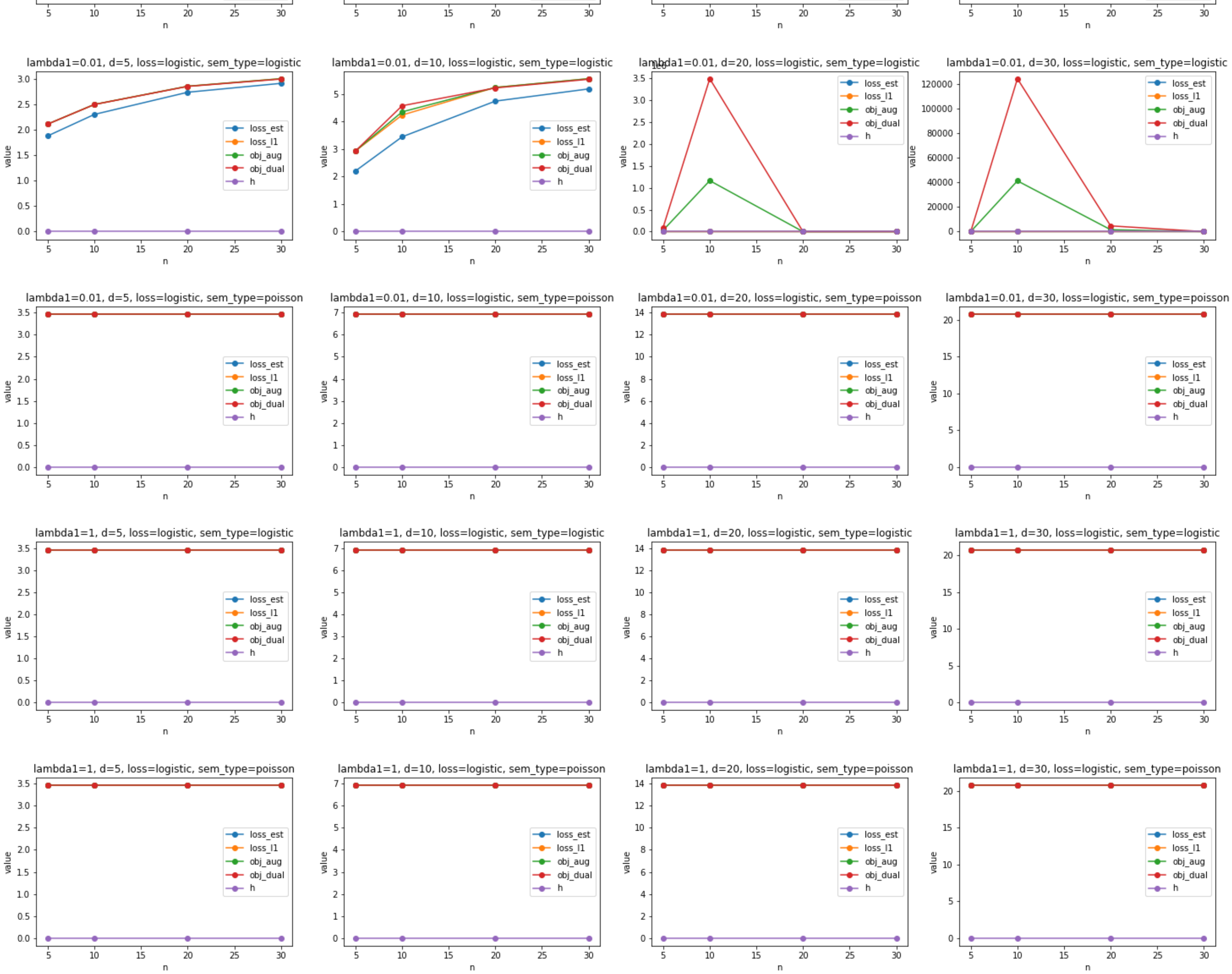


Experiment 2

Here I choose $d \in \{5, 10, 20, 30\}$, $s0 = d$, lambda1$\in \{0.1, 1\}$ for continuous distributions with loss_type = 'l2' and lambda1$\in \{0.01, 1\}$ for discrete distributions with loss_type = 'logistic', and then compute losses with different $n \in \{5, 10, 20, 30\}$. The results are basically consistent with the conclusions in Experiment 1.




lambda1=0.1, d=5, loss=l2, sem_type=exp
lambda1=0.1, d=10, loss=l2, sem_type=exp
lambda1=0.1, d=20, loss=l2, sem_type=exp
lambda1=0.1, d=30, loss=l2, sem_type=exp
lambda1=0.1, d=5, loss=l2, sem_type=gumbel
lambda1=0.1, d=10, loss=l2, sem_type=gumbel
lambda1=0.1, d=20, loss=l2, sem_type=gumbel
lambda1=0.1, d=30, loss=l2, sem_type=gumbel
lambda1=0.1, d=5, loss=l2, sem_type=uniform
lambda1=0.1, d=10, loss=l2, sem_type=uniform
lambda1=0.1, d=20, loss=l2, sem_type=uniform
lambda1=0.1, d=30, loss=l2, sem_type=uniform
lambda1=1, d=5, loss=l2, sem_type=gauss
lambda1=1, d=10, loss=l2, sem_type=gauss
lambda1=1, d=20, loss=l2, sem_type=gauss
lambda1=1, d=30, loss=l2, sem_type=gauss
lambda1=1, d=5, loss=l2, sem_type=exp
lambda1=1, d=10, loss=l2, sem_type=exp
lambda1=1, d=20, loss=l2, sem_type=exp
lambda1=1, d=30, loss=l2, sem_type=exp
lambda1=1, d=5, loss=l2, sem_type=gumbel
lambda1=1, d=10, loss=l2, sem_type=gumbel
lambda1=1, d=20, loss=l2, sem_type=gumbel
lambda1=1, d=30, loss=l2, sem_type=gumbel
lambda1=1, d=5, loss=l2, sem_type=uniform
lambda1=1, d=10, loss=l2, sem_type=uniform
lambda1=1, d=20, loss=l2, sem_type=uniform
lambda1=1, d=30, loss=l2, sem_type=uniform
value
n
loss_est
loss_l1
obj_aug
obj_dual
h



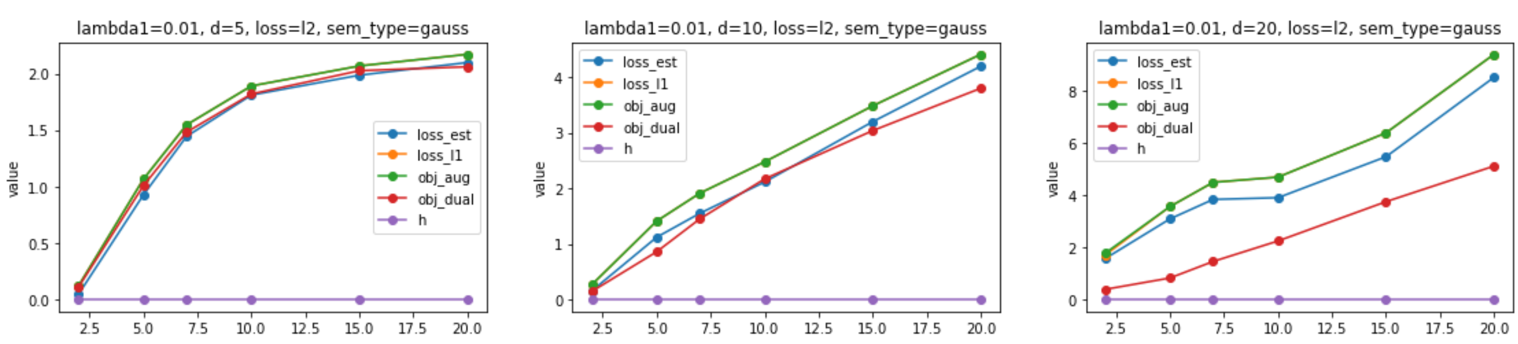
Experiment 3

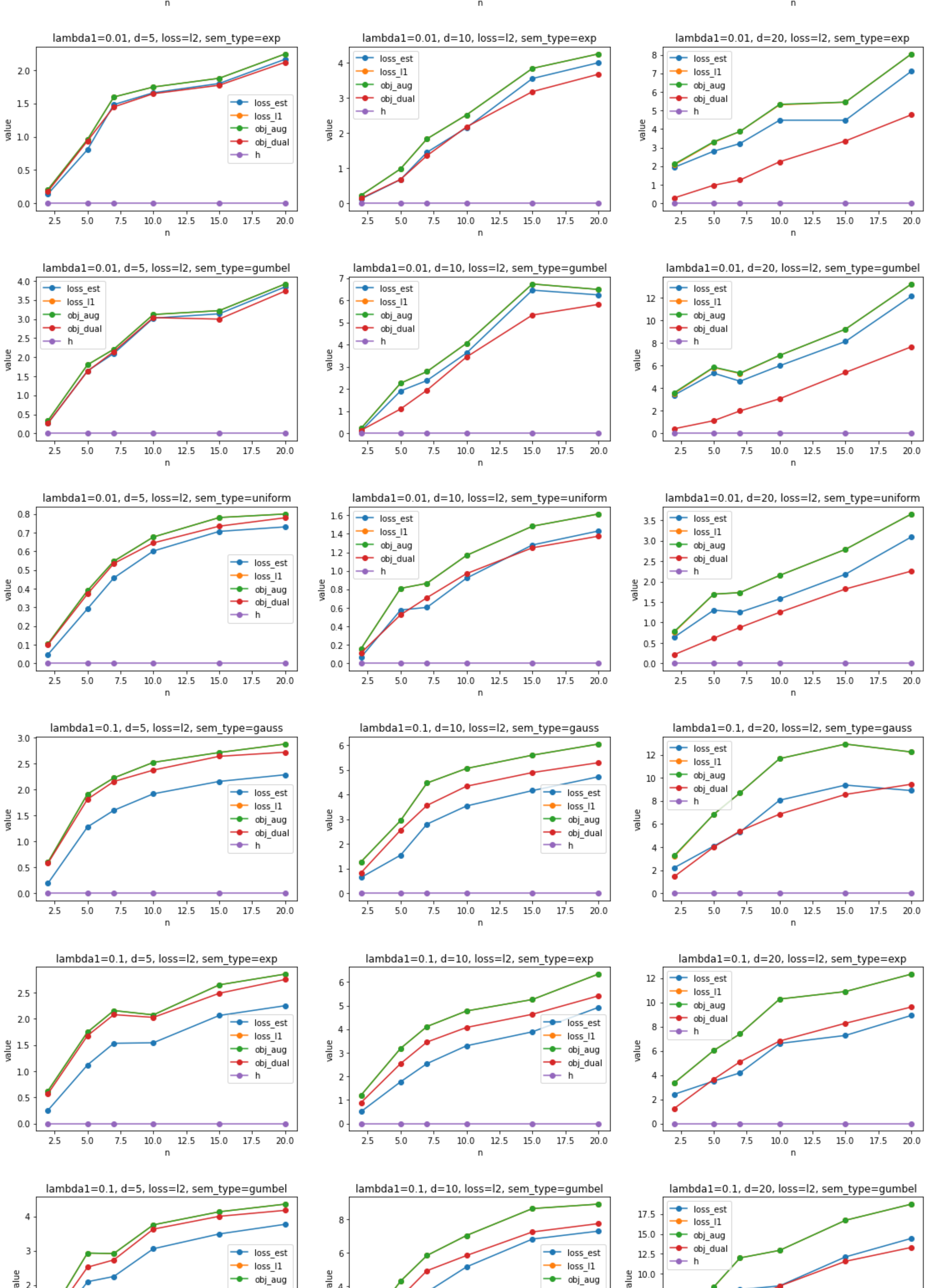
Here I choose $d \in \{5, 10, 20\}$, $s0 = d$, lambda1$\in \{0.01, 0.1, 1, 5\}$, loss_type = 'l2' for continuous distributions and loss_type = 'logistic' for discrete distributions , and then compute losses with different $n \in \{2, 5, 7, 10, 15, 20\}$.

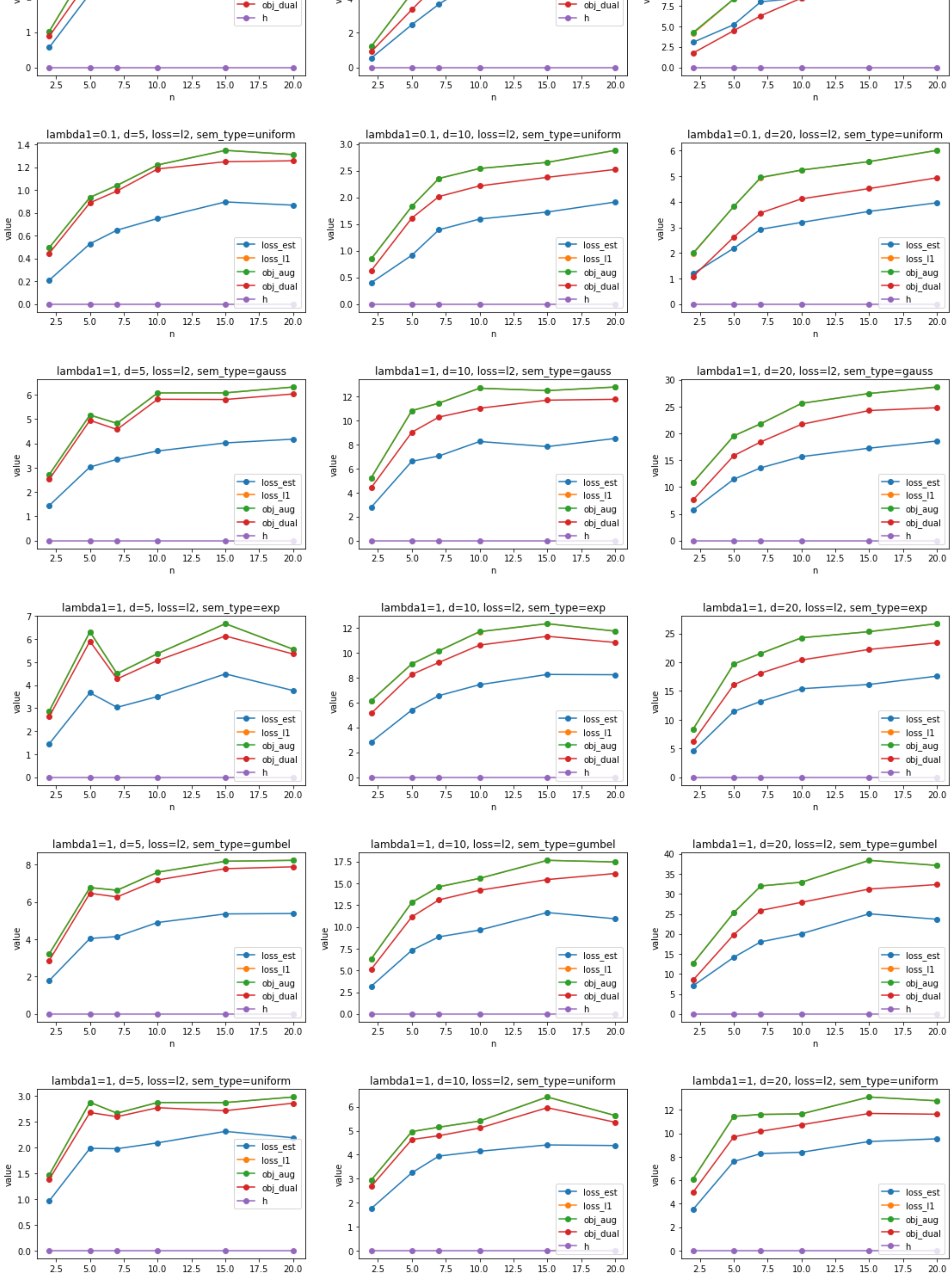
The purpose of doing this experiment is to observe the performance of the losses when n and d are relatively small.

Gains

- When loss_type = 'l2' and distributions are continuous, losses will approach 0.



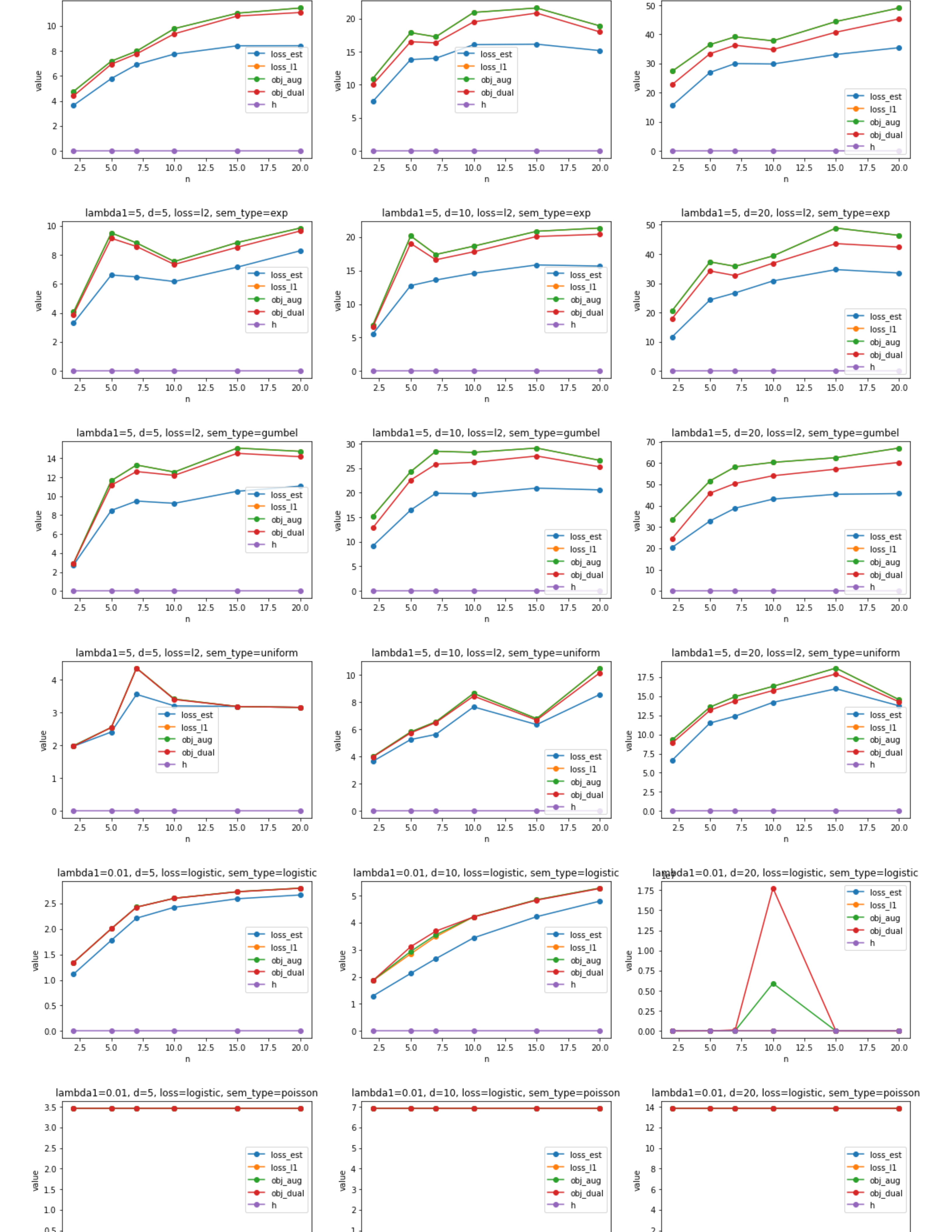

lambda1=0.01, d=5, loss=l2, sem_type=exp
lambda1=0.01, d=10, loss=l2, sem_type=exp
lambda1=0.01, d=20, loss=l2, sem_type=exp
lambda1=0.01, d=5, loss=l2, sem_type=gumbel
lambda1=0.01, d=10, loss=l2, sem_type=gumbel
lambda1=0.01, d=20, loss=l2, sem_type=gumbel
lambda1=0.01, d=5, loss=l2, sem_type=uniform
lambda1=0.01, d=10, loss=l2, sem_type=uniform
lambda1=0.01, d=20, loss=l2, sem_type=uniform
lambda1=0.1, d=5, loss=l2, sem_type=gauss
lambda1=0.1, d=10, loss=l2, sem_type=gauss
lambda1=0.1, d=20, loss=l2, sem_type=gauss
lambda1=0.1, d=5, loss=l2, sem_type=exp
lambda1=0.1, d=10, loss=l2, sem_type=exp
lambda1=0.1, d=20, loss=l2, sem_type=exp
lambda1=0.1, d=5, loss=l2, sem_type=gumbel
lambda1=0.1, d=10, loss=l2, sem_type=gumbel
lambda1=0.1, d=20, loss=l2, sem_type=gumbel
loss_est
loss_l1
obj_aug
obj_dual
h
value
n



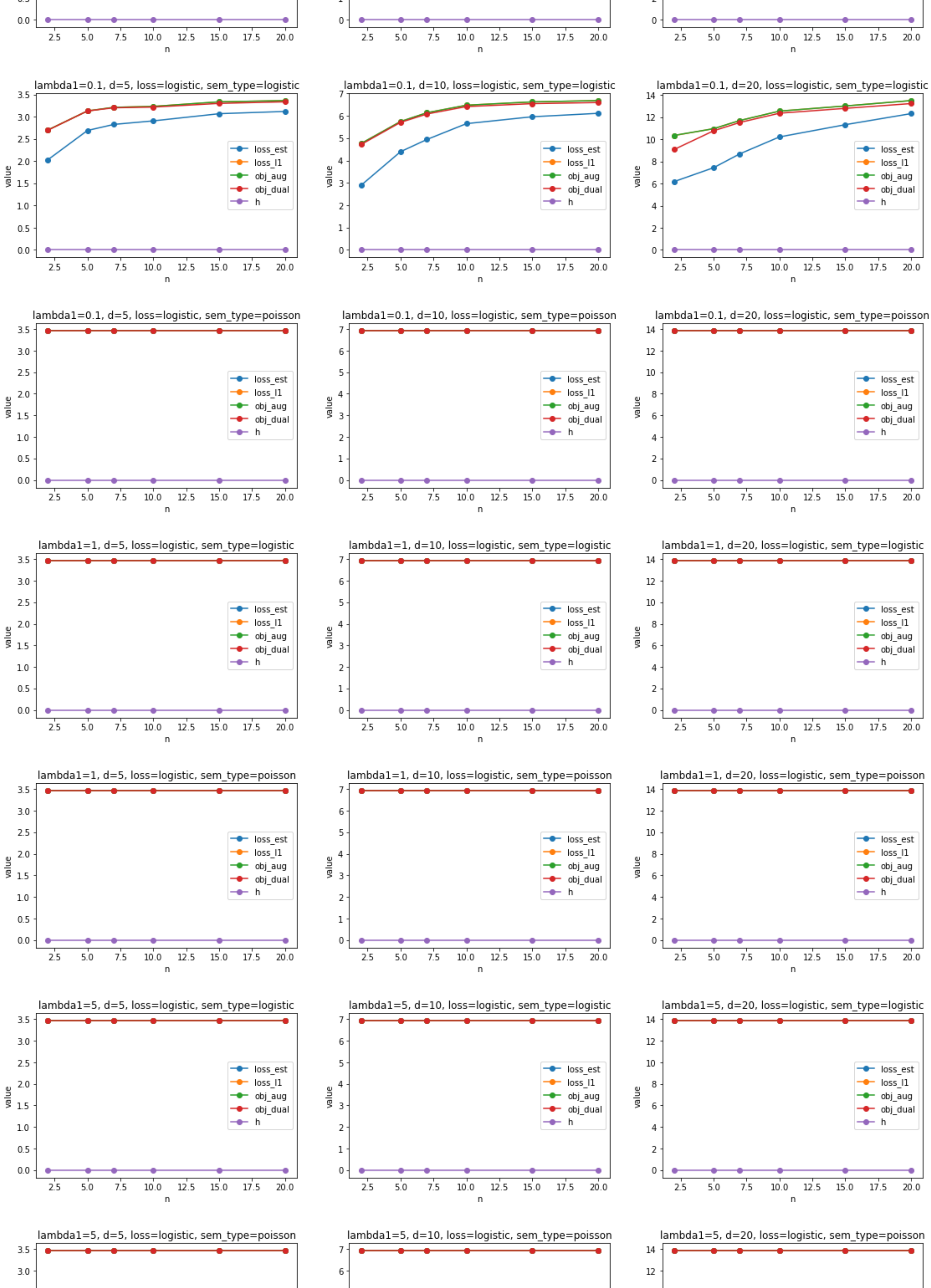
lambda1=5, d=5, loss=l2, sem_type=gauss

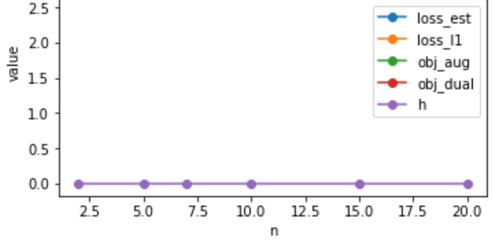
lambda1=5, d=10, loss=l2, sem_type=gauss

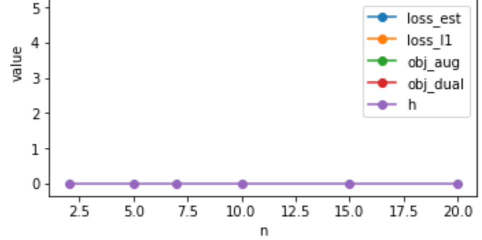
lambda1=5, d=20, loss=l2, sem_type=gauss

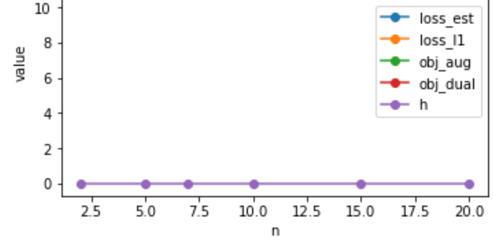


lambda1=5, d=5, loss=l2, sem_type=exp
lambda1=5, d=10, loss=l2, sem_type=exp
lambda1=5, d=20, loss=l2, sem_type=exp
lambda1=5, d=5, loss=l2, sem_type=gumbel
lambda1=5, d=10, loss=l2, sem_type=gumbel
lambda1=5, d=20, loss=l2, sem_type=gumbel
lambda1=5, d=5, loss=l2, sem_type=uniform
lambda1=5, d=10, loss=l2, sem_type=uniform
lambda1=5, d=20, loss=l2, sem_type=uniform
lambda1=0.01, d=5, loss=logistic, sem_type=logistic
lambda1=0.01, d=10, loss=logistic, sem_type=logistic
lambda1=0.01, d=20, loss=logistic, sem_type=logistic
lambda1=0.01, d=5, loss=logistic, sem_type=poisson
lambda1=0.01, d=10, loss=logistic, sem_type=poisson
lambda1=0.01, d=20, loss=logistic, sem_type=poisson
loss_est
loss_l1
obj_aug
obj_dual
h
value
n


lambda1=0.1, d=5, loss=logistic, sem_type=logistic
lambda1=0.1, d=10, loss=logistic, sem_type=logistic
lambda1=0.1, d=20, loss=logistic, sem_type=logistic
lambda1=0.1, d=5, loss=logistic, sem_type=poisson
lambda1=0.1, d=10, loss=logistic, sem_type=poisson
lambda1=0.1, d=20, loss=logistic, sem_type=poisson
lambda1=1, d=5, loss=logistic, sem_type=logistic
lambda1=1, d=10, loss=logistic, sem_type=logistic
lambda1=1, d=20, loss=logistic, sem_type=logistic
lambda1=1, d=5, loss=logistic, sem_type=poisson
lambda1=1, d=10, loss=logistic, sem_type=poisson
lambda1=1, d=20, loss=logistic, sem_type=poisson
lambda1=5, d=5, loss=logistic, sem_type=logistic
lambda1=5, d=10, loss=logistic, sem_type=logistic
lambda1=5, d=20, loss=logistic, sem_type=logistic
lambda1=5, d=5, loss=logistic, sem_type=poisson
lambda1=5, d=10, loss=logistic, sem_type=poisson
lambda1=5, d=20, loss=logistic, sem_type=poisson
loss_est
loss_l1
obj_aug
obj_dual
h
value
n


loss_est
loss_l1
obj_aug
obj_dual
h
value
n


loss_est
loss_l1
obj_aug
obj_dual
h
value
n


loss_est
loss_l1
obj_aug
obj_dual
h
value
n