Supplementary Materials

Profiling Dialysis Facilities for Adverse Recurrent Events

Jason P. Estes[1], **Yanjun Chen**[2], **Damla Şentürk**[3], **Connie M. Rhee**[4], **Esra Kürüm**[5], **Amy S. You**[4], **Elani Streja**[4], **Kamyar Kalantar-Zadeh**[4], **and Danh V. Nguyen**[6] *

[1]Research, Pratt & Whitney, East Hartford, CT 06042, U.S.A.
[2]Institute for Clinical and Translational Science, University of California, Irvine, CA 92687, U.S.A.
[3]Department of Biostatistics, University of California, Los Angeles, CA 90095, U.S.A.
[4]Harold Simmons Center for Chronic Disease Research and Epidemiology, University of California Irvine School of Medicine, Orange, CA 92868, U.S.A.
[5]Department of Statistics, University of California, Riverside, CA 92521, U.S.A.
[6]Department of Medicine, University of California Irvine, Orange, CA 92868, U.S.A.
email: danhvn1@uci.edu

*Corresponding author. Email: danhvn1@uci.edu

1 Example and R Codes

1.1 Example Datasets

This supplementary materials document provides a tutorial on fitting the Poisson and negative binomial profiling models with flagging/identification of outlier facilities (providers) based on nominal p-values. Sample datasets analogous to the Poisson and negative binomial models described in the main paper with 1,000 providers and a tutorial on the implementation along with R codes are provided here. The example (simulated) dataset containing 1,000 providers and 2 patient case-mix along with R functions and documentation for fitting the poisson and negative binomial models can be downloaded at

`http://faculty.sites.uci.edu/nguyenlab/supplement/`.

(Note that the input dataset has been sorted by provider ID's, namely fid.)

Load the datasets in R:

```
load("sampledata.RData")
ls()
[1] "ds.negb" "ds.pois"
```

Each dataset is a data frame with variables: facility ID's (fid), outcome count (y), follow-up time (t), and case-mix covariates, named in the sample data as z1 and z2:

```
head(ds.pois)
  fid y    t z1          z2
1   1 0 2.64  0  1.11651293
2   1 1 2.64  0  0.08436492
3   1 0 2.64  0  1.69733507
4   1 5 2.64  0  0.47171038
5   1 0 0.88  0  1.38912344
6   1 1 0.44  1 -0.64603955

dim(ds.pois)
[1] 78953     5
```

The sample dataset illustrating the Poisson model fit has 78,953 patients in 1,000 facilities.

Similarly, the sample data illustrating the NB model fit is a data frame with 76,946 patients in 1,000 faciliities.

```
head(ds.negb)
  fid y    t z1          z2
1   1 0 0.44  0 -0.08319949
2   1 1 2.64  0  0.84072866
3   1 1 0.88  0  0.86435909
4   1 2 2.64  1 -0.97581369
5   1 0 2.64  0  0.62968255
6   1 0 0.44  1 -1.45998774

dim(ds.negb)
[1] 76946     5
```

1.2 Fitting High-Dimensional Poisson and NB Profiling Models

The following usages fit the Poisson and NB models, respectively. Other optional input arguments are available to specify the initial facility effect values, case-mix parameter values, the number of resampling iterations for hypothesis testing of the equality of facility effects and the median facility, convergence, and maximum number of iterations..

```
# Source needed functions for model fitting:
source("model_fit.R")

# Fit Poisson model:
case.mix.vars <- c("z1","z2")
 fit.pois <- fit.poisson(ds.pois, case.mix.vars)

# Fit NB model:
fit.negb <- fit.nb(ds.negb, case.mix.vars)
```

The first input argument is the dataset name and second is the vector of patient case-mix covariate names. The offset is taken to be the log of the follow-up time, $\log(t)$. See the function headers for a more detailed documentation of required and optional input arguments.

1.3 Output Objects from Fitted Models

The output object is a list with seven elements. Elements are:

```
> names(fit.pois)
[1] "coefficients"    "fac.effect.ests" "ser.ests"        "p.vals"
[5] "od.est"          "ser.labels"      "num.iter"
```

The `coefficients` contains the case-mix parameter estimates ($r \times 1 = 2 \times 1$ in example), `fac.effect.ests` contains the facility effect estimates ($I \times 1 = 1000 \times 1$ in example), `ser.ests` contains the facility SER estimates ($I \times 1$), `p.vals` contains the facility nominal p-values in testing equality of the facility effect and median effect, `od.est` contains the over-dispersion estimate, `ser.labels` contains the classification ("B","ND","W") for each facility ($I \times 1$), and `num.iterations` contains the number of steps needed for convergence.

For the Poisson data, the (partial) results of the model fit are:

```
> print.summary(fit.pois)

   ser.ests p.vals fac.effect.ests ser.labels
1 0.5727896      0      -0.5679981          B
2 0.5807295      0      -0.5542314          B
3 0.6101956      0      -0.5047370          B
4 0.5792725      0      -0.5567435          B
5 0.6141551      0      -0.4982689          B
6 0.5371637      0      -0.6322136          B

Number of Facilities Flagged:
B:    348
ND:   281
W:    371

Coefficient Estimates:
z1:  0.5088418
z2: -0.5003711

Overdispersion Estimate:  0.994329061853391
```

The output structure for the NB model fit is the same:

```
> print.summary(fit.negb)

   ser.ests p.vals fac.effect.ests ser.labels
1 0.6128734  0.000      -0.4992100          B
2 0.6108527  0.000      -0.5025126          B
3 0.6081557  0.000      -0.5069374          B
4 0.6659223  0.004      -0.4161954          B
5 0.5284035  0.000      -0.6475082          B
6 0.6097260  0.000      -0.5043586          B
```

```
Number of Facilities Flagged:
B:    266
ND:   440
W:    294

Coefficient Estimates:
z1:  0.5014649
z2: -0.5019854

Overdispersion Estimate:  2.95848387762615
```

2 Standard Error Estimates

We conducted a supplementary simulation study to examine whether the standard error estimates based on the square root of the diagonal elements of the inverse of the observed information (SE_{OI}) matrix at convergence targets the true variability of the estimators, $\widehat{\beta}_r$, i.e. $SD_r \equiv \{Var(\widehat{\beta}_r)\}^{0.5}$. Because the alternating one-step Newton-Raphson estimation algorithm is not classical MLEs, these SE estimates generally may not target true variability of the estimators (SD). Thus, we also examined bootstrap SE (SE_{boot}) estimates as an alternative. Following the simulation setting described in the paper, we generated 500 Monte Carlo (MC) data sets and for each simulated dataset, 200 bootstrap samples were taken based on resampling facilities (to preserve correlation within facilities).

The results are summarized in Figure S1. The "true" SD's are estimates based on the sample SD over 500 MC simulation runs. Averages over the 500 simulated datasets shows that SE_{OI} do not target SD. In fact, the true SD is outside $\overline{SE}_{\text{OI}} \pm 1.96 \times SD_{SE_{\text{OI}}}$, where $\overline{SE}_{\text{OI}}$ and $SD_{SE_{\text{OI}}}$ are the mean and the standard deviation over the 500 MC runs, respectively. However, bootstrap SE estimates (SE_{boot} target SD well; given in Figure S1 are bootstrap mean plus/minus 1.96 times the standard deviation SE_{boot} over the 500 MC runs.

3 Excluded Extremely Low Information Facilities

As indicated in Section 4 of the paper, our analysis excluded 3.3% (213) of facilities with < 10 patients. Among these, $\sim 37\%$ had 1 or 2 patients and $\sim 56\%$ had 5 or fewer patients. Although reliable SER estimation is challenging in this sparse data context, for a descriptive summary of these facilities, we "estimated" the SER for these excluded facilities by taking the ratio of the observed event counts divided by the expected counts using the denominator in equation (3) for each of these 213 facilities. Figure S2 displays the distribution of SER for these very low volume providers with an average SER of 1.039 and standard deviation (SD) of 0.523. Compared to the analysis of the 6,188 included facilities, we note that the SER distribution of facilities flagged as "not different" from the national average had an average SER of 1.01 (SD 0.11). Thus, although these excluded facilities had SER average similar to the ND facilities, the SD was nearly 5 times larger.

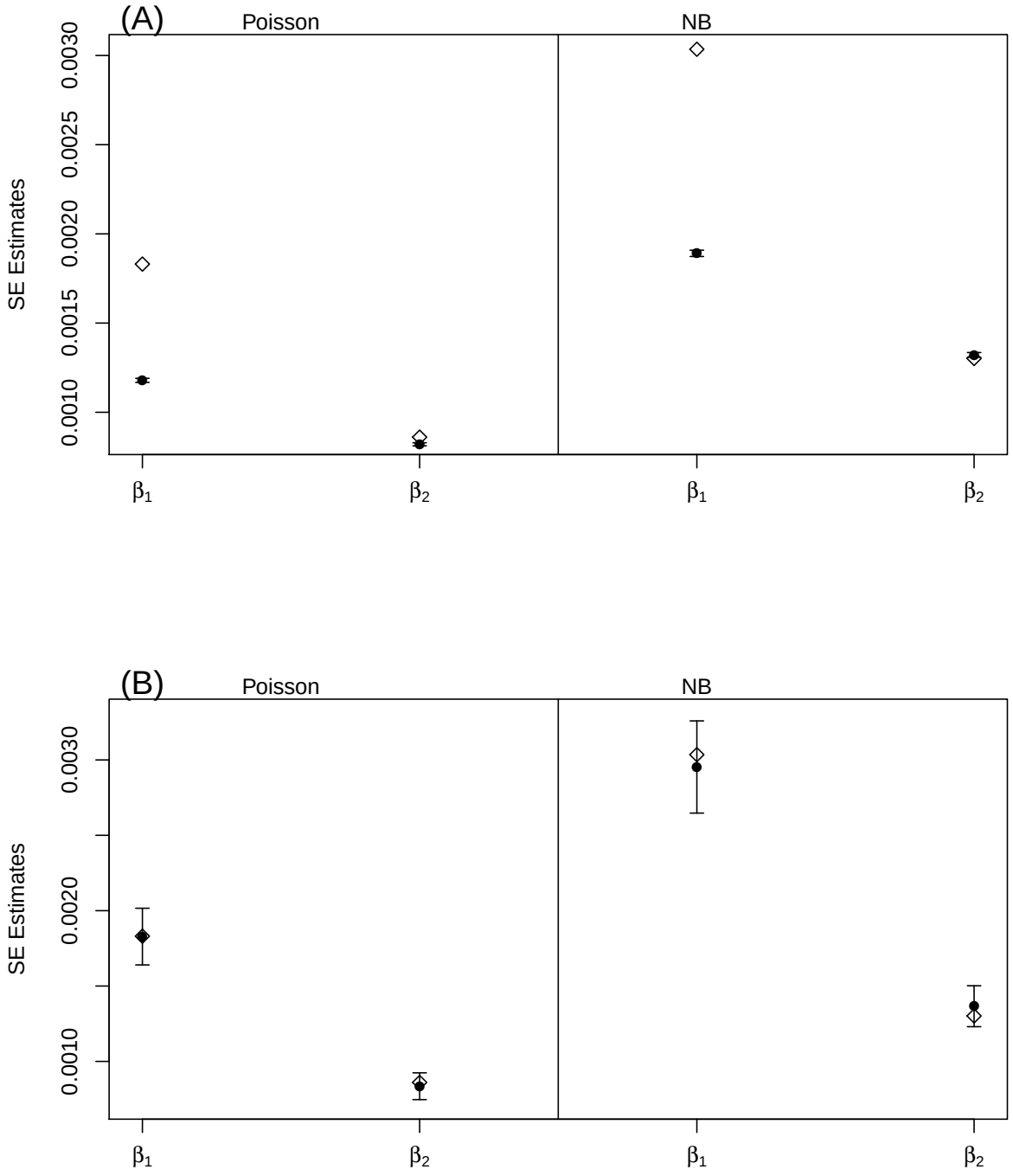

Figure S1: (A) Standard error estimates based on observed information matrix (SE_{OI}) and (B) based on bootstrap SE estimates (SE_{boot}), resampling facilities. Given are results for over 500 Monte Carlo datasets and for each simulated dataset, 200 bootstrap samples were taken (mean = black circle, mean $\pm 1.96 SE_{\text{boot}}$; Estimate of true SD = diamond based on 500 Monte Carlo replicates).

Table S1: Comparison of dialysis facility flagging between Poisson and negative binomial regression models for 6,188 facilities based on empirical null distribution.

	Poisson model			
Negative binomial model	Better	Not different	Worse	Total
Better	220	280	0	500 (8.1%)
Not different	75	5121	116	5312 (85.8%)
Worse	0	198	178	376 (6.1%)
Total	295 (4.8%)	5599 (90.5%)	294 (4.8%)	6188

Table S2: Results of identifying extreme dialysis facilities using the negative binomial model among all facilities and by facility size based on empirical null distribution (small: 10-55, medium: 56-96, and large: 97-560 patients).

Facility size	Worse		Not different		Better	
Small	66	3.2%	1943	93.8%	62	3.0%
Medium	167	8.1%	1759	85.8%	124	6.0%
Large	267	12.9%	1610	77.9%	190	9.2%
Overall	500	8.1%	5312	85.8%	376	6.1%

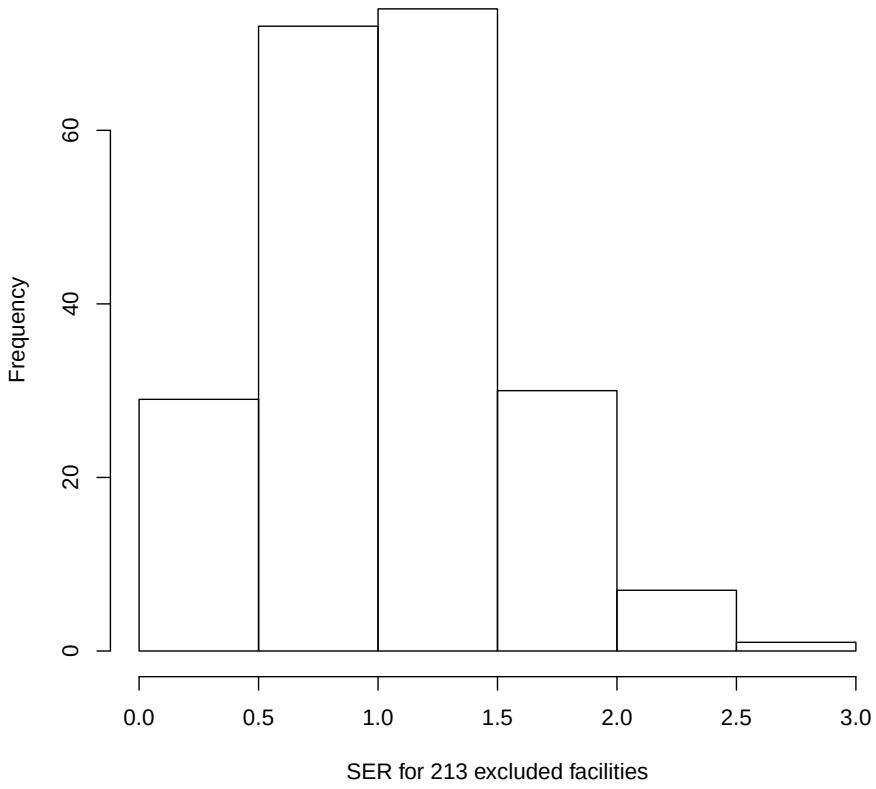

Figure S2: Distribution of "SER" for 213 excluded facilities (mean 1.039, standard deviation 0.523).

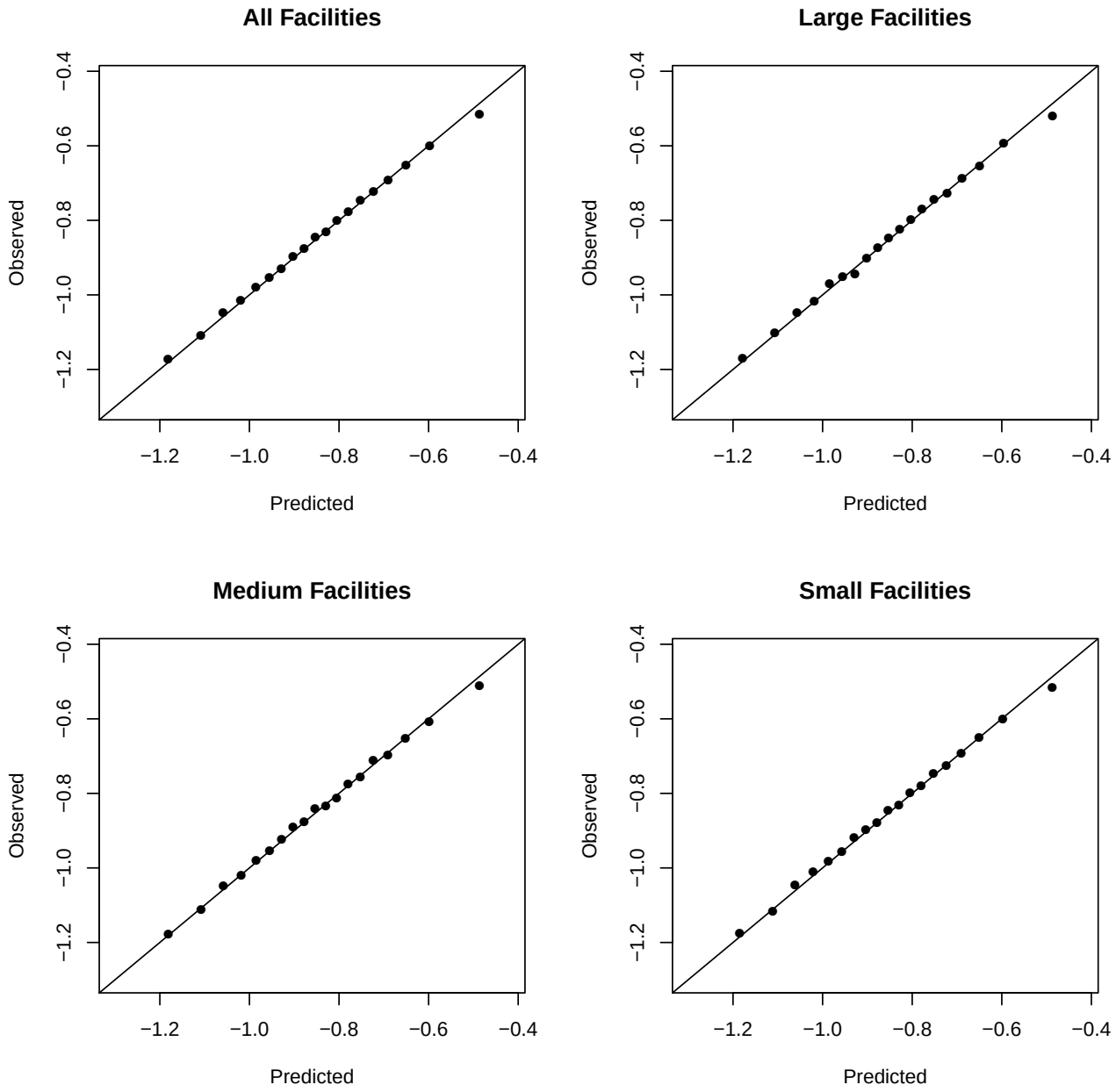

Figure S3: Observed and predicted/estimated rates on the log scale. Plotted are observed vs. predicted (log) rates in 20 bins based on model-based predicted values (observed vs. average predicted values in each bin).