Analysis of fit

Blain Morin

December 18, 2017

$$Y_i = \alpha_{j[i]} + \sum \beta_p X_{pi} + \epsilon_i, for hospitals (i = 1 to N)$$

$$\alpha_j = a + \sum b_k W_{kj} + u_j, for markets (j = 1 to J)$$

```
### load required packages
library(lme4)
```

```
## Loading required package: Matrix
```

```
library(tidyverse)
```

```
## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr
```

```
## Conflicts with tidy packages ----------------------------------------------
```

```
## expand(): tidyr, Matrix
## filter(): dplyr, stats
## lag():    dplyr, stats
```

```
### rstan requires having rtools installed
library(rstan)
```

```
## Warning: package 'rstan' was built under R version 3.4.3
```

```
## Loading required package: StanHeaders
```

```
## Warning: package 'StanHeaders' was built under R version 3.4.3
```

```
## rstan (Version 2.16.2, packaged: 2017-07-03 09:24:58 UTC, GitRev: 2e1f913d3ca3)
```

```
## For execution on a local, multicore CPU with excess RAM we recommend calling
## rstan_options(auto_write = TRUE)
## options(mc.cores = parallel::detectCores())
```

```
##
## Attaching package: 'rstan'
```

```
## The following object is masked from 'package:tidyr':
##
##     extract
```

```
###read in data
k12ReducedRG = read_csv("k12ReducedRG.csv")
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   episode = col_integer(),
```

```
##   Provider = col_character(),
##   hrr = col_character(),
##   avgagehrr = col_integer(),
##   `Rank for Variable dshpct` = col_integer(),
##   `Rank for Variable cmi` = col_integer(),
##   `Rank for Variable mdadjadmit` = col_integer(),
##   qstarrating = col_integer(),
##   urbanlocation = col_integer(),
##   joinnetwork = col_integer(),
##   jchaoaccredited = col_integer(),
##   qieffort = col_integer(),
##   reform = col_integer(),
##   mdaffiliation = col_integer(),
##   ownershipstatus = col_integer(),
##   hospitalbedsize = col_integer()
## )
```

```
## See spec(...) for full column specifications.
```

First, we create the null model with no fixed effects and a market random effect. This preliminary step allows us to partition the variance into hospital and market levels.

```
### This is the code for a null model with one random effect (hospital referral region)
m1 = lmer(episode ~ 1 + (1 | hrr) , data = k12ReducedRG)

summary(m1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: episode ~ 1 + (1 | hrr)
##    Data: k12ReducedRG
##
## REML criterion at convergence: 59231.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4593 -0.5832  0.0381  0.5955  6.8181
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  hrr      (Intercept) 1564856  1251
##  Residual             8795309  2966
## Number of obs: 3132, groups:  hrr, 304
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 18912.76      96.48     196
```

To check if this null model appoximates our data, we can use STAN to draw samples. We use the parameters found above to specify our STAN code.

```
###prepare data for STAN

m1_stan_data = list(episode = k12ReducedRG$episode,
             hrr = as.integer(as.factor(k12ReducedRG$hrr)),
             K = length(unique(k12ReducedRG$hrr)))

### Specify prior parameters
```

```
N = nrow(k12ReducedRG)

### Draw samples with stan
stan_null = stan(file = "m1.stan", data = m1_stan_data, iter = 2000, chains = 4)
```

```
## Warning: There were 1958 transitions after warmup that exceeded the maximum treedepth. Increase max_
## http://mc-stan.org/misc/warnings.html#maximum-treedepth-exceeded
```

```
## Warning: There were 4 chains where the estimated Bayesian Fraction of Missing Information was low. Se
## http://mc-stan.org/misc/warnings.html#bfmi-low
```

```
## Warning: Examine the pairs() plot to diagnose sampling problems
```

```
print(stan_null, pars = c("beta", "sigma_e", "sigma_w"),
      probs=c(0.025, 0.5, 0.975))
```

```
## Inference for Stan model: m1.
## 4 chains, each with iter=2000; warmup=1000; thin=1;
## post-warmup draws per chain=1000, total post-warmup draws=4000.
##
##             mean se_mean     sd     2.5%      50%    97.5% n_eff Rhat
## beta[1] 19008.61   37.39 128.72 18730.04 19037.47 19199.73    12 1.53
## sigma_e  3135.59   57.31 148.57  2913.07  3205.28  3334.60     7 2.45
## sigma_w   545.35  231.12 586.96     2.48    50.68  1352.75     6 2.83
##
## Samples were drawn using NUTS(diag_e) at Mon Dec 18 13:40:23 2017.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).
```

The mean intercept in the stan model (18987) is similar to the observed data model (18912).

Next, we add the hospital level predictors.

```
###rename columns
k12ReducedRG = k12ReducedRG %>%
  rename(cmirank = "Rank for Variable cmi",
         dshpctrank = "Rank for Variable dshpct",
         mdadjadmitrank = "Rank for Variable mdadjadmit")

m2 = lmer(episode ~ qstarrating +
        qieffort +
        jchaoaccredited +
        urbanlocation +
        mdaffiliation +
        ownershipstatus +
        hospitalbedsize +
        cmirank +
        dshpctrank +
        mdadjadmitrank +
        (1 | hrr),
        data = k12ReducedRG)

summary(m2)
```

```
## Linear mixed model fit by REML ['lmerMod']
```

```
## Formula:
## episode ~ qstarrating + qieffort + jchaoaccredited + urbanlocation +
##     mdaffiliation + ownershipstatus + hospitalbedsize + cmirank +
##     dshpctrank + mdadjadmitrank + (1 | hrr)
##    Data: k12ReducedRG
##
## REML criterion at convergence: 52030.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.4273 -0.5418 -0.0134  0.5431 10.7281
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  hrr      (Intercept)  689961   830.6
##  Residual             3006563  1733.9
## Number of obs: 2920, groups:  hrr, 304
##
## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)     14119.06     256.32   55.08
## qstarrating       -50.10      45.13   -1.11
## qieffort          108.84      82.86    1.31
## jchaoaccredited   373.81      95.31    3.92
## urbanlocation     656.51      95.18    6.90
## mdaffiliation     109.08      54.75    1.99
## ownershipstatus   335.50      58.98    5.69
## hospitalbedsize   565.36      71.54    7.90
## cmirank          1579.75      42.28   37.36
## dshpctrank        -27.77      34.93   -0.80
## mdadjadmitrank    -72.11      32.30   -2.23
##
## Correlation of Fixed Effects:
##            (Intr) qstrrt qiffrt jchccr urbnlc mdfflt ownrsh hsptlb cmirnk
## qstarrating -0.660
## qieffort    -0.076 -0.059
## jchaoccrdtd -0.179  0.001 -0.064
## urbanlocatn -0.086  0.020 -0.103 -0.022
## mdaffiliatn -0.181  0.099 -0.066  0.016 -0.024
## ownrshpstts -0.495  0.024  0.190 -0.108 -0.100  0.057
## hospitlbdsz -0.288  0.134 -0.157 -0.109 -0.156 -0.210 -0.002
## cmirank      0.210 -0.117 -0.138 -0.050 -0.262 -0.176 -0.056 -0.414
## dshpctrank  -0.277  0.216  0.082  0.005  0.088 -0.118  0.062 -0.194  0.095
## mdadjdmtrnk -0.260  0.017 -0.006  0.037  0.108 -0.115  0.117  0.001 -0.010
##             dshpct
## qstarrating
## qieffort
## jchaoccrdtd
## urbanlocatn
## mdaffiliatn
## ownrshpstts
## hospitlbdsz
## cmirank
## dshpctrank
```

```
## mdadjdmtrnk   0.007
```

Next, we check the model using STAN.

```
###change data to only complete cases
k12ReducedRG = k12ReducedRG[complete.cases(k12ReducedRG),]

###change N
N = nrow(k12ReducedRG)

### prepare data for STAN
m2_stan_data = list(episode = k12ReducedRG$episode,
                hrr = as.integer(as.factor(k12ReducedRG$hrr)),
                qstar = k12ReducedRG$qstarrating,
                qieffort = k12ReducedRG$qieffort,
                accredited = k12ReducedRG$jchaoaccredited,
                urban = k12ReducedRG$urbanlocation,
                mdaffiliation = k12ReducedRG$mdaffiliation,
                ownership = k12ReducedRG$ownershipstatus,
                bedsize = k12ReducedRG$hospitalbedsize,
                cmi = k12ReducedRG$cmirank,
                dsh = k12ReducedRG$dshpctrank,
                mdadjadmitrank = k12ReducedRG$mdadjadmitrank,
                K = length(unique(k12ReducedRG$hrr)),
                id = as.integer(as.factor(k12ReducedRG$Provider)))




### run STAN simulation
m2_stan = stan(file = "m2.stan", data = m2_stan_data, iter = 1000, chains = 2)
```

```
## Warning: There were 1000 transitions after warmup that exceeded the maximum treedepth. Increase max_
## http://mc-stan.org/misc/warnings.html#maximum-treedepth-exceeded
```

```
## Warning: Examine the pairs() plot to diagnose sampling problems
```

```
print(m2_stan, pars = c("beta", "sigma_e", "sigma_w"),
      probs=c(0.025, 0.5, 0.975))
```

```
## Inference for Stan model: m2.
## 2 chains, each with iter=1000; warmup=500; thin=1;
## post-warmup draws per chain=500, total post-warmup draws=1000.
##
##              mean se_mean      sd     2.5%      50%     97.5% n_eff Rhat
## beta[1]  19251.48   24.70   85.60 19059.61 19255.31 19403.53    12 1.10
## beta[2]     35.66   21.40   37.60   -10.57    21.71   121.77     3 2.53
## beta[3]   -141.96   70.90  101.48  -293.36  -158.04   110.68     2 1.60
## beta[4]    245.37  197.88  203.42   -32.88   239.57   556.85     1 4.13
## beta[5]   -114.83   66.81  137.65  -390.93   -71.71    97.42     4 2.36
## beta[6]   -112.49  183.18  194.72  -443.24  -100.73   162.15     1 3.87
## beta[7]    137.30  237.74  289.08  -143.15    15.92   750.13     1 4.51
## beta[8]    -69.54   55.30   66.31  -142.98   -95.21    85.12     1 2.13
## beta[9]     81.90   60.48  118.96  -220.59    36.66   289.46     4 1.57
## beta[10]   -61.41   24.65   62.60  -159.93   -71.01    89.94     6 1.67
## sigma_e   3014.89   32.67   58.78  2931.32  3007.04  3140.62     3 1.25
## sigma_w      1.05    0.54    0.54     0.48     0.98     1.81     1 9.94
```

```
##
## Samples were drawn using NUTS(diag_e) at Mon Dec 18 16:15:55 2017.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).
```

```r
summary(m2)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## episode ~ qstarrating + qieffort + jchaoaccredited + urbanlocation +
##     mdaffiliation + ownershipstatus + hospitalbedsize + cmirank +
##     dshpctrank + mdadjadmitrank + (1 | hrr)
##    Data: k12ReducedRG
##
## REML criterion at convergence: 52030.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.4273 -0.5418 -0.0134  0.5431 10.7281
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  hrr      (Intercept)  689961   830.6
##  Residual             3006563  1733.9
## Number of obs: 2920, groups:  hrr, 304
##
## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)     14119.06     256.32   55.08
## qstarrating       -50.10      45.13   -1.11
## qieffort          108.84      82.86    1.31
## jchaoaccredited   373.81      95.31    3.92
## urbanlocation     656.51      95.18    6.90
## mdaffiliation     109.08      54.75    1.99
## ownershipstatus   335.50      58.98    5.69
## hospitalbedsize   565.36      71.54    7.90
## cmirank          1579.75      42.28   37.36
## dshpctrank        -27.77      34.93   -0.80
## mdadjadmitrank    -72.11      32.30   -2.23
##
## Correlation of Fixed Effects:
##             (Intr) qstrrt qiffrt jchccr urbnlc mdfflt ownrsh hsptlb cmirnk
## qstarrating -0.660
## qieffort    -0.076 -0.059
## jchaoccrdtd -0.179  0.001 -0.064
## urbanlocatn -0.086  0.020 -0.103 -0.022
## mdaffiliatn -0.181  0.099 -0.066  0.016 -0.024
## ownrshpstts -0.495  0.024  0.190 -0.108 -0.100  0.057
## hospitlbdsz -0.288  0.134 -0.157 -0.109 -0.156 -0.210 -0.002
## cmirank      0.210 -0.117 -0.138 -0.050 -0.262 -0.176 -0.056 -0.414
## dshpctrank  -0.277  0.216  0.082  0.005  0.088 -0.118  0.062 -0.194  0.095
## mdadjdmtrnk -0.260  0.017 -0.006  0.037  0.108 -0.115  0.117  0.001 -0.010
##             dshpct
## qstarrating
```

```
## qieffort
## jchaoccrdtd
## urbanlocatn
## mdaffiliatn
## ownrshpstts
## hospitlbdsz
## cmirank
## dshpctrank
## mdadjdmtrnk  0.007
```