Simulated Risk Model

James T. Bang

September 5, 2022

Data Generation and Estimation Equations

Data Generation

The data for observed patient outcomes of readmission, home health care (HHC) assignment, patient demographics, prior hospitalization, diagnosis, and payer type come from a single hospital in the READI study (Weiss, et al., 2019).
The challenge is simulating the patient's latent "risk" for the adverse outcome of readmission. We propose that this is not the same as the *ex post* ***likelihood*** of readmission. We also propose that patients with higher risk of adverse outcomes - including readmission - will be more likely to receive HHC assignments. Hence, there will be a bias in the using the observedlikelihood to measure risk, and therefore a bias in the observed effectiveness of the intervention of assigning HHC. To simulate this bias, we randomly assign ten percent of the patients who received HHC assignment but were not eventually readmitted a positive readmission risk value. The remaining risk values remain the same as the observed readmission outcome.

Estimation Equations

The equations to estimate the *observed* rate of readmission and home health assignment, respectively, are:

$$HHC = \beta_{10} + \beta_{11} Age + \beta_{12} Prior + \sum_{r=1}^{R} \rho_{1r} Race_r + \sum_{h=1}^{H} \theta_{1h} Ethnicity + \sum_{p=1}^{P} \pi_{1p} Paytype_p$$
$$+ \sum_{m=1}^{M} \delta_{1m} Diagnosis + \alpha_{1j} + \epsilon_{1i},$$
$$Risk = \beta_{20} + \beta_{21} Age + \beta_{22} Prior + \sum_{r=1}^{R} \rho_{2r} Race_r + \sum_{h=1}^{H} \theta_{2h} Ethnicity + \sum_{p=1}^{P} \pi_{2p} Paytype_p$$
$$+ \sum_{m=1}^{M} \delta_{2m} Diagnosis + \alpha_{2j} + \epsilon_{2i},$$

and

$$Readmission = \beta_{30} + \beta_{31} Age + \beta_{32} Prior + \beta_{33} HHC + \sum_{r=1}^{R} \rho_{3r} Race_r + \sum_{h=1}^{H} \theta_{3h} Ethnicity$$
$$+ \sum_{p=1}^{P} \pi_{3p} Paytype_p + \sum_{m=1}^{M} \delta_{3m} Diagnosis + \alpha_{3j} + \epsilon_{3i},$$

where α_j represents the fixed effect for hospital *j*.

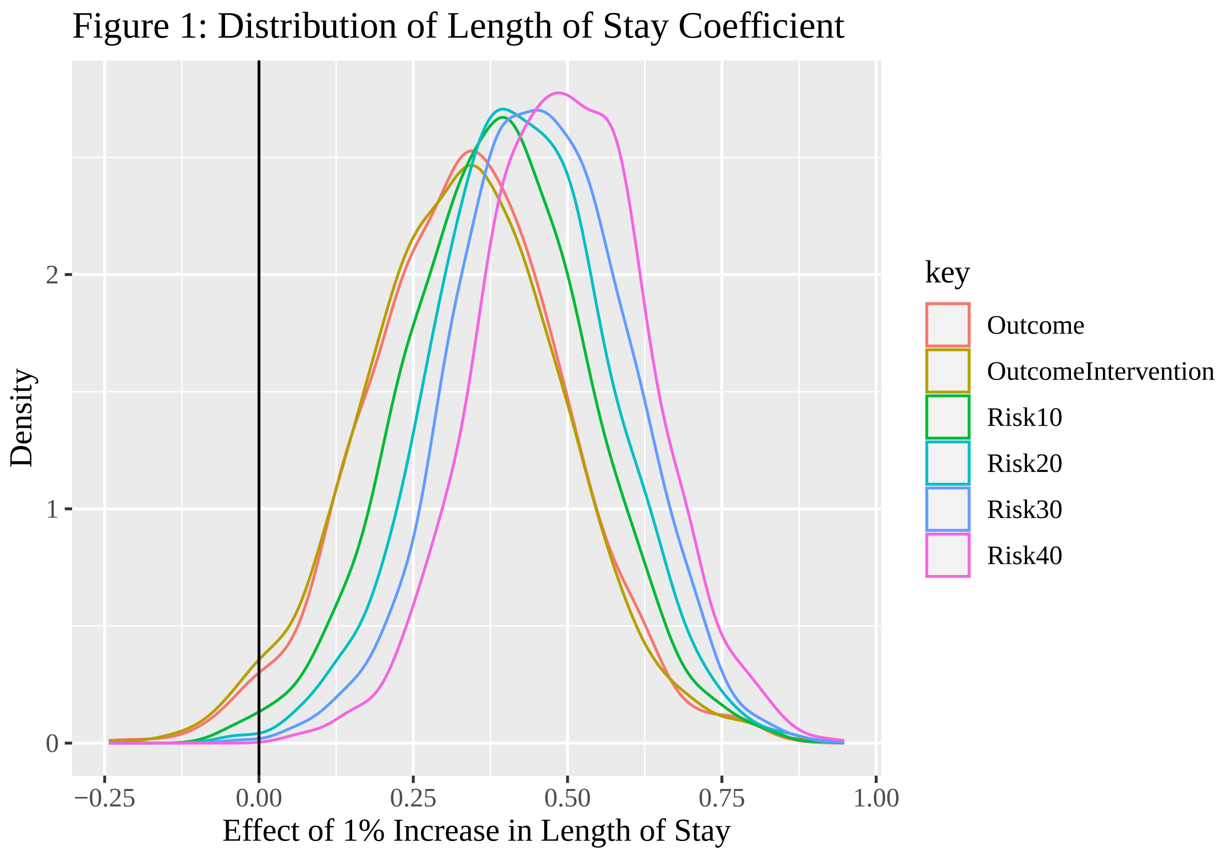
Simulation Results.

We simulated the model 1000 times using 1000 observations randomly selected from the dataset described above in each iteration. We present summary statistics for the impact of age on *risk* of readmission and the *likelihood* of readmission in Table 1.

Table 1:

	Dependent variable:		
	hh	read30_01	
	(1)	(2)	(3)
medsugptSurgical	0.163	−0.530*	−0.534*
	(0.250)	(0.318)	(0.319)
icustayICU Stay	0.343	−0.243	−0.247
	(0.245)	(0.305)	(0.305)
genderFemale	0.227	−0.073	−0.076
	(0.224)	(0.241)	(0.242)
raceAsian	0.110	0.176	0.176
	(0.286)	(0.312)	(0.313)
raceBlack	−0.297	0.013	0.015
	(0.312)	(0.323)	(0.323)
ethnicityHispanic	−0.353	0.401	0.403
	(0.738)	(0.412)	(0.414)
maritalsMarried	−0.174	−0.094	−0.094
	(0.217)	(0.243)	(0.243)
mortriskModerate	0.352	0.292	0.291
	(0.317)	(0.342)	(0.343)
mortriskMajor	0.387	0.433	0.429
	(0.396)	(0.434)	(0.435)
severityModerate	0.372	0.240	0.238
	(0.360)	(0.366)	(0.366)
severityMajor	0.624	0.448	0.443
	(0.434)	(0.444)	(0.446)
age_cats45-54	0.486	0.119	0.116
	(0.456)	(0.394)	(0.395)
age_cats55-64	0.848**	0.020	0.015
	(0.417)	(0.369)	(0.371)
age_cats65-74	1.240***	0.026	0.017
	(0.408)	(0.388)	(0.390)
age_cats75+	1.767***	0.049	0.034
	(0.423)	(0.392)	(0.398)
lnlos	0.861***	0.337**	0.331**
	(0.151)	(0.159)	(0.163)
hhHHC			0.061
			(0.302)
Constant	−4.973***	−3.385***	−3.376***
	(0.872)	(0.684)	(0.689)
Observations	1,000	1,000	1,000
Log Likelihood	−335.269	−315.279	−315.020
Akaike Inf. Crit.	738.538	698.559	700.040
Note:			*p<0.1; **p<0.05; ***p<0.01

Figure 1 presents a kernel density plot of the distributions of the two outcomes.




```
## <ggproto object: Class ScaleDiscrete, Scale, gg>
##     aesthetics: fill
##     axis_order: function
##     break_info: function
##     break_positions: function
##     breaks: waiver
##     call: call
##     clone: function
##     dimension: function
##     drop: TRUE
##     expand: waiver
##     get_breaks: function
##     get_breaks_minor: function
##     get_labels: function
##     get_limits: function
##     guide: legend
##     is_discrete: function
##     is_empty: function
##     labels: Outcome Outcome with Intervention Risk - 10% Bias Risk - ...
##     limits: NULL
##     make_sec_title: function
##     make_title: function
##     map: function
##     map_df: function
##     n.breaks.cache: NULL
```

```
##     na.translate: TRUE
##     na.value: grey50
##     name: waiver
##     palette: function
##     palette.cache: NULL
##     position: left
##     range: <ggproto object: Class RangeDiscrete, Range, gg>
##         range: NULL
##         reset: function
##         train: function
##         super:  <ggproto object: Class RangeDiscrete, Range, gg>
##     rescale: function
##     reset: function
##     scale_name: hue
##     train: function
##     train_df: function
##     transform: function
##     transform_df: function
##     super:  <ggproto object: Class ScaleDiscrete, Scale, gg>
```