# Random Effect Model of Physician Decisions

Lucas

August 2022

## 1 With no random effects

Fix the time (i.e. year) and setting $s$ (e.g. Part D, Part B, MIPS). Physician $i$ uses treatment $j$ on patient $b$, if the utility of doing so exceeds that of the outside option, $j = 0$:

$$u^s_{ib,j} = h^s_{ij}(X_b)\beta^s + \pi^s_{ij}\theta^s + \eta^s_{ib,j} \geq 0 = u^s_{ib,0}$$

- $h^s_{ij}(\cdot)$: patient's health gain from treatment, which depends on their characteristics $X_b$. Parametrize $h^s_{ij}(x) = \tilde{h}^s_{ij} x\gamma^s$, where $\gamma^s$ scales the average health gain $\tilde{h}^s_{ij}$ based on the patient's characteristics and is estimatd in a previous step.
- $\pi^s_{ij}$: is the doctor's monetary payment from treating.
- $\eta^s_{ib,j}$: iid Type I Extreme Value errors.

Assuming patients who see the same doctor are homogenous, i.e. $X_b = \bar{X}^s_i$, the probability of $i$ choosing $j$ is given by the logit form:

$$\sigma(h^s_j, \pi^s_j, \bar{X}^s_i, \beta^s, \theta^s) = \frac{\exp(h^s_{ij}\beta^s + \pi^s_{ij}\theta^s)}{1 + \exp(h^s_{ij}\beta^s + \pi^s_{ij}\theta^s)}$$

where $h^s_{ij} = \tilde{h}^s_{ij}\bar{X}^s_i\gamma^s$. If $j \in \{0, 1\}$ i.e. treatment is binary, the model-implied number of treatments $Y^s_{ij}$ out of $B$ decisions is then distributed:

$$Y^s_{ij} \sim Binomial(\sigma(h^s_j, \pi^s_j, \bar{X}^s_i, \beta^s, \theta^s); B)$$

### 1.1 Prediction

If $h^s_{ij}$ and $\pi^s_{ij}$ for a given $(i, j)$ are independent across settings $s, s'$, this simple model predicts that any cross-setting relationships in a physician's behavior must come from cross-setting correlations in her patient mix interacted with the efficacy scale, $\bar{X}^s_i\gamma^s$.

- This could explain why prescription behaviors are correlated: the sets of patients to whom a doctor prescribe opioid and benzodiazepenes, respectively, are likely to be very similar.
- This does not explain why Part D and Part B behaviors are uncorrelated, since my prior is that a doctor's Part B patients are probably close in characteristics to her Part D patients.

# 2 Introducing random effects

Now let the taste coefficients vary by doctor. The probability of $i$ choosing $j$ conditional on a realization of $(\beta_i, \theta_i)$ is then given by:

$$\sigma(h_j^s, \pi_j^s, \bar{X}_i^s, \beta_i^s, \theta_i^s) = \frac{\exp(h_{ij}^s \beta_i^s + \pi_{ij}^s \theta_i^s)}{1 + \exp(h_{ij}^s \beta_i^s + \pi_{ij}^s \theta_i^s)}$$

$(\beta_i^s, \theta_i^s) \sim F^s$ are setting-specific random effects. Since they are unobserved, they need to be integrated out to obtain the unconditional probability:

$$\sigma(h_j^s, \pi_j^s, \bar{X}_i^s) = \int \frac{\exp(h_{ij}^s \beta + \pi_{ij}^s \theta)}{1 + \exp(h_{ij}^s \beta + \pi_{ij}^s \theta)} dF^s(\beta, \theta)$$

## 2.1 What we want

One way to explain a correlation between two settings is to look at the covariance matrix between $(\beta_i^s, \theta_i^s)$ and $(\beta_i^{s'}, \theta_i^{s'})$.

As we will see, the covariances are not identified without further parametrization. Intuitively, to identify the covariances we need to observe the same doctor making a treatment decision using $(\beta_i^s, \theta_i^s, \beta_i^{s'}, \theta_i^{s'})$ simultanously. Since our settings are mutually exclusive, this is infeasible.

## 2.2 Identification issues

To illustrate the identification issue, I plot whether the necessity conditions are satisfied for identification of models of various complexities. The procedure is as follows:

1. First, I create a synthetic aggregate data set by specifying a set of true fixed and random effect parameters, $\boldsymbol{\gamma_o}$, and drawing $(h_{ij}^s, \pi_j^s, \beta_i, \theta_i)$. Then I draw the number of treatments for each doctor using a Binomial distribution with the associated probability $\sigma_{ij}(\boldsymbol{\gamma_o})$.
2. I then perturb each parameter around the true value while keeping the other parameters at the true values. For each perturbation, I compute the value of the simulated likelihood function. Because I know the exact data generating processes, a necessary condition for identification is

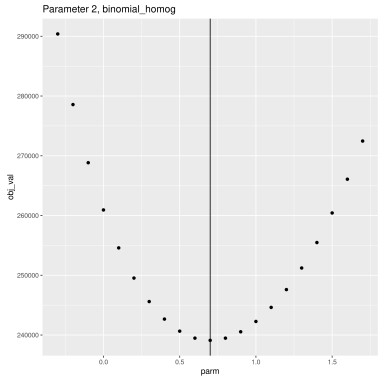


Figure 1: Necessary for identification

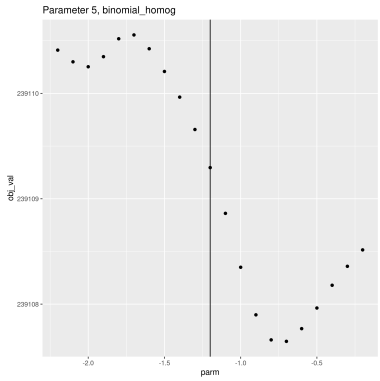


Figure 2: Not identified

that the negative likelihood function is minimized at the true value of the parameter.

The left plot is what we should expect to see if the parameter is identified. The right plot shows that the parameter cannot be identified from the data (because the negative likelihood function is not minimied at the true value).

I try various specifications of the random effects to see what kind of parameters are identified, and what are not.

- The main finding is that the covariance of any two parameters associated with different settings is never non-parametrically identified (see model (4), (5), (6)).
- The fixed (non-random) parameters are always identified.
- I make a suggestion on what we can do to proceed below.

| Model | Identified | Not Identified | Data |
|---|---|---|---|
| (1) $\beta_i = \bar{\beta} + \sigma u_i$ | All | None | S=1,T=1 |
| (2) $\beta_i = \bar{\beta} + u_i$<br>$\theta_i = \bar{\theta} + v_i$<br>$(u_i, v_i) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Lambda})$ | All | None | S=1,T=1 |
| (3) $\beta_i^s = \beta^s + u_i$<br>$\theta_i^s = \bar{\theta}^s + v_i$<br>$(u_i, v_i) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Lambda})$<br>time trends $\alpha^s$ | All | None | S=2,T=3 |
| (4) $\beta_i^s = \beta^s + u_i^s$<br>$(u_i^1, u_i^2) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Lambda})$<br>time trends $\alpha^s$ | Everything else | $cov(u_i^1, u_i^2)$ | S=2,T=3 |
| (5) $\beta_i^1 = \bar{\beta}^1 + u_i^1$<br>$\theta_i^2 = \bar{\theta}^2 + v_i^2$<br>$(u_i^1, v_i^2) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Lambda})$<br>time trends $\alpha_s$ | Everything else | $cov(u_i^1, v_i^2)$ | S=2,T=3 |
| (6) $\beta_i^1 = \beta + u_i^1$<br>$\beta_i^2 = \bar{\beta} + u_i^2$<br>$(u_i^1, u_i^2) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Lambda})$<br>common trend $\alpha$ | Everything else | $cov(u_i^1, u_i^2)$ | S=2,T=3 |
| (7) $\beta_i^1 = \bar{\beta}^1 + z_i\delta^1 + u_i^1$<br>$\theta_i^2 = \bar{\theta}^2 + z_i\delta^2 + v_i^2$<br>$(u_i^1, v_i^2) \sim \mathcal{N}(\mathbf{0}, diag(\sigma^1, \sigma^2))$<br>time trends $\alpha_s$ | All | None | S=2,T=3 |

### 2.2.1 Suggestion: Use Model (7)

In model (7) I add an observe doctor-invariant characteristic $z_i$ to as a component of heterogeneity in the taste parameters. This new interaction comes with a setting-specific coefficient $\delta^s$.

If $\delta^s$ and $\delta^{s'}$ have similar signs and magnitudes, this could explain a high correlation in the physician's behavior between $s$ and $s'$. If one of the $\delta$ is zero, then the correlation should be weaker.

An advantage to this approach is that we can test precisely what it is about the doctor that makes two settings correlate or not. A disadvantage is that we are assuming zero residual correlation due to unobserved characteristics, which would have been captured by the off-diagonal terms of the covariance matrix.