A Model of Physician Behavior

Lucas Do

July 2022

1 Questions to ask

1. Basically, I have a random effect model. Tried to estimate by MLE using simulated data. The fixed effects are identified but the random effects are nowhere near identified even when using a stripped down model despite me knowing the DGP.

   So I'm trying to understand why my model isn't identified, what other variation do I need or what sort of model restrictions.

   (Can show plots of what should be global/local minima. They are as expected for the FEs but not for the REs.)

2. Only have two options so may not be able to pin down random effect parameters in the traditional BLP sense, but I do have something else in the data that may help identification. Not sure if my arguments are cogent though.

2 Starting from Basics

2.1 A logit model

Let's start with a simple logit model, with no random effects. Fix setting s and time t. Doctor i's utility from choosing treatment j for patient b is:

$$u_{ib,j} = h_j(X_b)\beta + \pi_j\theta + \epsilon_{ib,j} \tag{1}$$

where

- h_j is patient's health gain from treatment, which deends on their characteristics X_b.
- π_j is the doctor's monetary payment from treating.
- $\eta_{ib,j}$ are iid Type I Extreme Value errors.

For simplicity, we assume i has homogeneous patients so that $h_j(X_b) = h_j(\bar{X}_i) = h_{ji}$. Also assume binary treatment decisions and normalize the utility of not treating to zero (plus a $\eta_{ib,0}$ shock). (Question: I probably need more than 2 to identify random effects?). Then we have the usual logit probability of treatment j being chosen:

$$\Pr[i \text{ chooses } j] = \sigma_{ij} = \frac{\exp(h_j\beta + \pi_j\theta)}{1 + \exp(h_j\beta + \pi_j\theta)}$$

Deviation of predicted probability σ_{ij} from actual probability s_{ij} arises from sampling errors in $\epsilon_{ib,j}$. Since treatment is binary, we do observe the probability of the outside option as well, and we can invert the share to estimate the following via OLS:

$$\log(s_{ij}) - \log(s_{i0}) = h_j\beta + \pi_j\theta$$

An instrumental variable approach can readily accommodate the case where either h_j or π_j is endogenous.

An implication of this model is that every doctor facing the same h_j and θ_j will have the exact same probability of treatment. Furthermore, the profit-elasticity of treatment decision is:

$$\frac{\partial s_{ij}}{\partial \pi_j} = \theta s_{ij}(1 - s_{ij})$$

such that every doctor with the same probability of treatment will respond the same to a unit increase in profit incentive (or change in patient's health gain health). If we pool settings together to estimate this equation, it is implied that these properties are also shared across settings.

What identifies (β, θ) is the shift in observed (relative) probability of treatment as h_j and π_j vary.

2.2 Random effects: simplest case

Now imagine doctors are heterogeneous in their taste parameters, so that:

$$\beta_i = \bar{\beta} + Z_i\beta^o + w_i^\beta$$

$$\theta_i = \bar{\theta} + Z_i\theta^o + w_i^\theta$$

Z_i is a vector of observed doctor attributes, and w_i are unobserved components of heterogeneity. This gives the following probability of treatment conditional of (w_i^β, w_i^θ):

$$\begin{aligned}\sigma_{ij|\cdot} &= \frac{\exp(h_j\bar{\beta} + \pi_j\bar{\theta} + Z_ih_j\beta^o + Z_i\pi_j\theta^o + h_jw_i^\beta + w_i^\theta)}{1 + \exp(h_j\bar{\beta} + \pi_j\bar{\theta} + Z_ih_j\beta^o + Z_i\pi_j\theta^o + h_jw_i^\beta + \pi_jw_i^\theta)} \\ &= \frac{\exp(\delta_j + Z_ih_j\beta^o + Z_i\pi_j\theta^o + h_jw_i^\beta + \pi_jw_i^\theta)}{1 + \exp(\delta_j + Z_ih_j\beta^o + Z_i\pi_j\theta^o + h_jw_i^\beta + \pi_jw_i^\theta)}\end{aligned}$$

where $\delta_j = h_j\bar{\beta}+\pi_j\bar{\theta}$ captures the mean utility of treatment. This is identified by the average shift in observed probability of treatment as h_j and π_j vary. (β^o, θ^o) are identified by how doctors with different Z_i differentially deviate from this average shift. And what identifies $var(\begin{pmatrix} w_i^\beta \\ w_i^\theta \end{pmatrix})$ is the degree of residual variation in this shift across doctors.

Since the random effects (w_i^β, w_i^θ) are not known, the unconditional probability of treatment is calculated by integrating over their joint distribution, which gives us an 'average' probability across doctors in the same setting-year:

$$\sigma_{ij} = \int \frac{\exp(\delta_j + zh_j\beta^o + zi\pi_j\theta^o + h_j w^\beta + \pi_j w^\theta)}{1+\exp(\delta_j + zh_j\beta^o + z\pi_j\theta^o + h_j w^\beta + \pi_j w^\theta)} dF(w,z)$$

We can also express the probability as:

$$s_{ij} = \int_{A_j} s_{ij|\cdot} dP^*(Z, w, \eta)$$

where A_j is the set of doctors who choose to treat at least once and $P^*(\cdot)$ denotes the population distribution of doctors, characterized by (Z, w, η). The integral is weighted by the percentage of times each doctor i treats, which may depend on consumer characteristics so will need a model for that too...but we are interested in doctors' intrinsic/taste qualities.

The unconditional profit elasticity is the weighted average of individual-specific elasticities:

$$\frac{\partial \sigma_{ij}}{\partial \pi_j} = \int_{w,z} (\bar{\theta} + w^\theta)\sigma_{ij|w}(1-\sigma_{ij|w}) dF(w,z)$$

Identification argument for the variances: Recall that the logit model would predict every doctor with similar existing probability will respond the same to a change in (π_j, h_j). The degree to which doctors differ in their responses to changes in (π_j, h_j) will be driven by $F(w,z)$ and pin down $var(w)$. The covariance term is identified by observing that doctors who are on average more responsive to changes in π_j are also more/less responsive on average to changes in h_j. OR are these variations indistinguishable from sampling errors in the logit term?

Now if $F(w,z)$ is setting-specific, that is, $F(w_s, z_s) \neq F(w_{s'}, z_s)$, then estimating the model for each setting identifies each $var(w^s)$ separately.

2.3 Random effects: what I really want

Basically, I want to model correlation *across* settings. I have around 4-5 settings. **It is observed that doctor behavior is correlated across some settings but not others, and I am trying to rationalize that with a unifying model that links the settings together** (and we are looking to test a theory of physician altruism).

Concern: The 'products' (i.e. treatments) and choice sets are mutually exclusive across settings, aside from the fact I conceptualize them all as 'treatment' versus 'no treatment'.

If unrestricted, the random effects I want to specify are:

$$\beta_i^s = \bar{\beta} + Z_i\beta^o + w_i^{\beta,s}$$

$$\theta_i^s = \bar{\theta} + Z_i\theta^o + w_i^{\theta,s}$$

for each s.

For tractability I assume that $(w_i^{\beta,s}, w_i^{\theta,s})_s$ are jointly distributed $\mathcal{N}(\mathbf{0}, \mathbf{\Lambda})$. My goal is to estimate $\mathbf{\Lambda}$.

The diagonal elements could probably be estimated by just running mixed effect logit separately on each setting and using the identification argument in the previous section and marginal properties of joint normals. But I'm really most interested in the off-diagonal elements. I think it's possible to do this non-parametrically[1], but my simulations say otherwise. I think the fact that the settings are mutually exclusive have to do with it, somehow.

I can also correlate these decisions through observed factors but Z_i is not adequate for explaining the correlations I want. Since Z_i is setting-invariant, it only explains why some the propensity to treat is correlated between some settings. It does not explains why it is are uncorrelated between others.

$\implies$ Need Z_i^s that appears only in some settings but not others? Pool all settings together in one equation and use setting dummies. However, does this make sense when the treatment variables are different across settings?

What variation is needed to identify he off-diagonal covariances? Either need data variation or further model restrictions.

2.3.1 Does this make sense?

Seems like another way to generate correlation is to introduce a theory of treatment substitution patterns between settings i.e. if doctors are gaining a lot from setting A they might treat in setting B less. So we will use doctors who appear in both A and B. We haven't made it explicit here...

Features of the data that I think are potentially exploitable.

1. In one set of settings S_1, the utility of treatment j_1 looks like this:

$$u_{ib,j_1} = h_{j_1}\beta_i + \epsilon_{ib,j_1}$$

2. In another set of settings S_2, the utility of treatment j_2 looks like this:

$$u_{ib,j_2} = \pi_{j_2}\theta_i + \epsilon_{ib,j_2}$$

[1] Intuitively, consider the case where $\pi_{j,s}$ increases. $\bar{\theta}$ is identified from a weighted average of the mean observed shifts across settings. $var(w_i^\theta)_s$ is identified from the setting-specific dispersion across doctors in response to the change. Identification of $cov(w_i^{\theta,s}, w_i^{\theta,s'})$ requires observing whether doctors who are on average more responsive to changes in $\pi_{j,s}$ are on average more/less responsive to changes in $\pi_{j,s'}$

3. And in a third set of settings S_3, the utility of treatment j_3 looks like this:

$$u_{ib,j_3} = h_{j_3}\beta_i + \pi_{j_3}\theta_i + \epsilon_{ib,j_3}$$

2.4 Question for IO Research Group

1. Estimation algorithm. Should I pool together all doctors to estimate the unconditional σ_{ij} or integrate over w and estimate it for each doctor?

2. Nested: search over covariance. for each covariance, search over variance. for each of this variance, solve for fixed effect. then solve for variance then solve for covariance.

Explanation for correlations:

1. Correlations in patient characteristics between some settings but not others!

Variation in choice sets?