Physician Models of Prescribing and Topcoding

Lucas

July 2022

1 Chandra-Staiger-style Model

I use ‘treat’ as a shortcut for ‘opioid prescribing’ or ’top-coding’. The change in doctor i’s latent utility, W_{ibtr}, from treating patient b in year t and setting $r \in \{\text{Part B, Part D}\}$ depends on the patient’s perceived utility change U_{ibtr} from being treated, the profit π_{ibtr} accrued to the doctor, and a mean-zero shock ϵ_{ibtr} unknown to the doctor himself which reflects randomness the in treatment outcomes.

$$W_{ibtr} = U_{ibtr}\theta_{it} + \pi_{ibtr}\gamma_{it} + \epsilon_{ibtr} \tag{1}$$

θ_{it} and γ_{it} govern how much the doctor cares about the patient and his own profit, respectively. The objective is to estimate the altruism parameter θ_{it} for each doctor i, which may or may not be time-invariant.

The patient’s utility change U_{ibtr} depends on her observed characteristics X_b and those observed by the doctor but not the econometrician, η_{ibtr}. Each doctor has a different perception β_{itr} of how X_b affects the patient’s outcome:

$$U_{ibtr} = X_b\beta_{itr} + \eta_{ibtr} \tag{2}$$

Let $Y_{ibtr} \in \{0, 1\}$ denote the decision whether to treat. The doctor treats the patient iff the expected utility change exceeds a threshold, τ_{itr}. The ex-ante probability of treating a patient b is: :

$$\begin{aligned} Pr\{Y_{ibtr} = 1\} &= Pr\{E[W_{ibtr}] \geq \tau_{itr}\} \\ &= Pr\{(X_b\beta_{itr} + \eta_{ibtr})\theta_{it} + \pi_{ibtr}\gamma_{it} \geq \tau_{itr}\} \\ &= Pr\{-\eta_{ibtr} \leq \pi_{ibtr}\frac{\gamma_{it}}{\theta_{it}} + X_b\beta_{itr} - \frac{\tau_{itr}}{\theta_{it}}\} \end{aligned} \tag{3}$$

By assuming that $-\eta_{ibtr}$ is iid across (i, b, t, r), I can estimate the parameters using a single index model, such as conditional logit. With τ_{itr} being doctor-year-setting fixed effect, $(\frac{\gamma_{it}}{\theta_{it}}, \beta_{itr})$ are identified as long as there is variation in the doctor’s per-transaction profit and his patients’ characteristics within a given year and setting.

1.1 Identifying θ_{it}

Another restriction needs to be made to identify θ_{it} separately from the ratios. Two possible solutions:

1. Let $\gamma_{it} = 1 - \theta_{it}$. More altruistic doctors care less about profit. I will go with this.

2. Fix $\gamma_{it} = 1$. All doctors care the same amount about profit.

1.2 Identification with Aggregate Data

Estimating the above decision model requires observing the profit π_{ibtr} and patient characteristics X_b for every transaction. However, the currently available data are aggregated at the doctor-year-level. I observe how many times treatment occurs that year, as well as the case mix $E[X_b|i,t,r]$.

Let's aggregate up the model. Doctor i's expected number of treatments T_{itr}, given the set $\mathcal{B}_{itr}$ of N_{itr} patients, is:

$$E[T_{itr}|\mathcal{B}_{itr}] = \sum_{b=1}^{N_{itr}} E[1\{Y_{ibtr} = 1\}] = \sum_{b=1}^{N_{itr}} Pr\{\eta_{ibtr} \leq \pi_{ibtr}\frac{1-\theta_{it}}{\theta_{it}} + X_b\beta_{itr} - \frac{\tau_{itr}}{\theta_{it}}\} \tag{4}$$

Directly incorporating the case mix $E[X_b|i,t,r]$ into the aggregation is difficult because of the problem is non-linear. There are some potential solutions:

1.2.1 Simulating Heterogeneity in Patient Characteristics

Although I do not observe individual X_b, I can simulate $\mathcal{B}_{itr}$ for each doctor by stratifying X_b into bins and drawing patient characteristics from a multinomial distribution with parameter $\mathbf{p}_i = E[X_b|i,t,r]$. Then I can compute $\pi_{ibtr} = \pi_{itr}(X_b)$ and match the observed total treatment occurrences with the model-implied total, using a minimum-distance or GMM estimator. Each year-setting pair will form a separate key or moment condition for estimation.

1.2.2 Assuming Homogeneity in Patient Characteristics

Without simulations, I have to account for the fact that I do not observe patient-specific data by:

1. Assuming that $\pi_{ibtr} = \pi_{itr}$ (and known). In a given year and setting, doctor i's profit from treatment is identical for all patients. A candidate is $\pi_{itr} = \pi_{itr}(E[X_b|i,t,r]) =: \bar{\pi}_{itr}$ i.e. profit from the 'average' patient.

2. Assuming that all patients are observably identical to the 'average' patient, $X_b = E[X_b|i,t,r] =: \bar{X}_{itr}$. Alternatively, I could specify $U_{ibtr} = \eta_{ibtr}$ without an observable component. Here, I choose the former.

Then the expected average number of treatments given N_{itr} observably identical patients is:

$$\begin{aligned}\frac{E[T_{itr}|N_{itr}]}{N_{itr}} &= E\left[\frac{T_{itr}}{N_{itr}}|N_{itr}\right] = \frac{1}{N_{itr}}\sum_{b=1}^{N_{itr}} Pr\{\eta_{ibtr} \leq \bar{\pi}_{itr}\frac{1-\theta_{it}}{\theta_{it}} + \bar{X}_{itr}\beta_{itr} - \frac{\tau_{itr}}{\theta_{it}}\} \\ &= \frac{N_{itr}}{N_{itr}} \cdot Pr\{\eta_{ibtr} \leq \bar{\pi}_{itr}\frac{1-\theta_{it}}{\theta_{it}} + \bar{X}_{itr}\beta_{itr} - \frac{\tau_{itr}}{\theta_{it}}\} \\ &= Pr\{\eta_{ibtr} \leq \bar{\pi}_{itr}\frac{1-\theta_{it}}{\theta_{it}} + \bar{X}_{itr}\beta_{itr} - \frac{\tau_{itr}}{\theta_{it}}\}\end{aligned} \tag{5}$$

Since $\frac{\tau_{itr}}{\theta_{it}}$ soaks up all patient-invariant effects, nothing is identified here without further restrictions:

1. At the very least, I need to specify coarser-level fixed effects, e.g. $\tau_{itr} = \tau_{it}$. Then I can exploit within-doctor variation in profit across settings in a given year to estimate $\frac{1-\theta_{it}}{\theta_{it}}$ and recover τ_{it}.

2. Similarly, β_{itr} need to be at a coarser level than $\bar{X}_{itr}$. If I assume $\beta_{itr} = \beta_{it}$, it can be identified using variation in the patient mix across settings.

If the parameters are at the (i, t) level, however, they are only identified for doctors who appear in both Part D and Part B that year. And since there are only two potential settings per doctor-year, the estimates may be highly imprecise.

Therefore, *without transaction-level data*, it is better to specify all parameters at the doctor-level and use variation both across settings and years to estimate $(\theta_i, \beta_i, \tau_i)$ by matching the observed averages with the model-implied averages, with a GMM or minimum-distance estimator:

$$E\left[\frac{T_{itr}}{N_{itr}}|N_{itr}\right] = Pr\{\eta_{ibtr} \leq \bar{\pi}_{itr}\frac{1-\theta_i}{\theta_i} + \bar{X}_{itr}\beta_i - \frac{\tau_i}{\theta_i}\} \tag{6}$$

1.3 Takeaways

1. If we don't trust simulations, transaction-level data allow us to identify the parameters at a finer level. Several of our candidate predictors of θ vary at the doctor-year level (e.g. group membership), so recovering θ_{it} is desirable. But given that there are many more doctors than years, recovering θ at the doctor-level is likely not too bad.

2. According to the model, altruism is identified by variation in per-transaction profit. Do we have good measures for this?