Measuring Upcoding, Costs and Profits

Lucas Do

August 2022

1 Part B

1.1 Identifying Upcoding

1.1.1 A basic Ordered Probit framework

Define the following 'correct coding' function:

$$C^0(\tilde{R}) = \begin{cases} 1 & -\infty < \tilde{R} < \gamma_1^0 \\ 2 & \gamma_1^0 \leq \tilde{R} < \gamma_2^0 \\ 3 & \gamma_2^0 \leq \tilde{R} < \gamma_3^0 \\ 4 & \gamma_3^0 \leq \tilde{R} < \gamma_4^0 \\ 5 & \gamma_4^0 \leq \tilde{R} < \infty \end{cases}$$

where $\tilde{R}$ perfectly encapsulates all of the patient's relevant health conditions from the doctor's perspective.

In practice, we observe the doctor's coding decision $C_i(\cdot)$. We also do not perfect observe $\tilde{R}$ but a noisy measure of R (e.g. risk score) of it:

$$R = \tilde{R} + \epsilon$$

We observe the patient's risk R but not her unobserved (to the econometrician, but observed by the doctor) characteristic $\epsilon \sim \mathcal{N}(0, \sigma^2)$. To quantify the probability of upcoding, suppose we know the thresholds γ_k and observe:

$$\gamma_2^0 \leq R < \gamma_3^0$$

But $C_i(R) = 5$. What is the probability of this NOT being a result of upcoding?

$$\begin{aligned} Pr[\text{not upcoding}] &= Pr[\tilde{R} \geq \gamma_4^0 | R] \\ &= Pr[R - \epsilon \geq \gamma_4^0 | R] \\ &= Pr[\epsilon \leq R - \gamma_4^0 | R] \\ &= \Phi\left(\frac{R - \gamma_4^0}{\sigma}\right) \end{aligned}$$

So the likelihood of upcoding is $1-\Phi\left(\frac{R-\gamma_4^0}{\sigma}\right)$. Note that $R-\gamma_4^0<0$ so without knowing σ, $Pr[\text{upcoding}]>\frac{1}{2}$ which checks out intuitively.

To compute this in practice, however, we need accurate estimates of (1) the thresholds γ_k^0 and (2) distribution/variance of ϵ. We also need the relationship between $\tilde{R}$ and R to be correctly specified.

1.1.2 Estimating γ_k^0

In theory, we can identify the true thresholds by observing decisions of a doctor for whom we know $C_i(\cdot)=C^0(\cdot)$, or equivalently $\gamma_k^i=\gamma_k^0$ for all k.

A doctor i is likely to be an upcoder if his coding function is such that the thresholds are lower for higher codes, e.g. $\gamma_k^i<\gamma_k^0$ for $k=4,5$. The thresholds reflect several things. For example, if EHR and profits are two factors affecting the thresholds, we can parametrize:

$$\gamma_k^i=\gamma_k^0-\beta_k 1[EHR_i=1]-\theta\Delta\pi_k+u_{ik} \tag{1}$$

where:

- $\beta_k>0$ indicates that having EHR helps lower the threshold for/cost of coding at the respective intensity level.
- θ reflects how much doctors care about profit incentives on average. W.l.o.g. we can specify $\theta_i=\theta+t_i$ and then put $t_i\Delta\pi_k$ into the error term.
- $\Delta\pi_k$ is the extra profit for coding the code corresponding to $[\gamma_k^0,\gamma_{k+1}^0)$ instead of $[\gamma_{k-1}^0,\gamma_k^0)$. There is variation over time in Part B payment schedules for this.
- u_i is a mean-zero doctor idiosyncratic shock. Uncorrelated with $\Delta\pi_k$, maybe also with $1[EHR_i=1]$.

Propose strategy:

1. For each doctor i, we can estimate γ_k^i with an ordered probit-type regression using claims-level data (just using over-time variation is not enough).
2. Run a 2SLS/IV regression for (1). If the model is correctly specified, the constant term gives us the 'true' threshold γ_k^0.

   The intuition is as follows: suppose deviation from the true threshold only arises from the mean-zero idiosyncratic shock u_{ik}. The constant term in this regression is just the average γ_k^i, which is precisely the true value because the shock is mean zero.
3. With estimates of $\hat{\gamma}_k^0$ we can classify which doctors are likely an upcoders. We can also assign, for any case of $C_i(R)$, the likelihood of it being an upcode, as in section 1.1.1, if we can somehow estimate the variance of unobserved patient characteristics σ.

e.g. if the measurement error is uncorrelated with R, we have:

$$\begin{aligned} var(R) &= var(\tilde{R}) + var(\epsilon) \\ \iff var(\epsilon) &= var(R) - var(\tilde{R}) > 0 \end{aligned}$$

Estimating $var(R)$ is straight forward and provides an upper bound for $\sigma = var(\epsilon)$. This provides us an upper bound for the probability of upcoding to accompany the lower bound of $\frac{1}{2}$ (see 1.1.1). We can calibrate our results at various level of $var(\tilde{R})$, unless we can estimate it credibly.

1.2 Measuring Profits

Measuring profits is straight forward for Part B data if we can tell which codes are upcodes and which are not (just use the Part B fee schedules).

After measuring profits we can use the variable in our random effect model.

2 New IT Data: Use to Supplement MIPS

We observe how well each attesting doctor fulfills different core and menu objectives and measures (there's quite a lot of variation in this). We also observe how much they get paid per year/stage (there is not much variation in this at the year level, but maybe more variation at the level of cumulative payment over the years).

We also have payment data at the hospital level, and for Medicare Advantage Organization Providers.

3 MIPS

A little iffy on using group-based scores.

3.1 Computing Revenues

The new QPP data sets can be used for measuring profits associated with QPP scores: associated with each doctor are a final score, a payment adjustment percentage, and an allowed charge. However, the data is de-identified, preventing us from linking the NPIs.

We can still use the data to create a mapping between final scores and payment adjustments, since I can't find the precise CMS formula besides the two thresholds for negative, neutral and positive adjustments each year

3.2 Computing costs and patient outcomes

To start, we could use quality scores as proxy for patient outcomes, and incentive payments as revenues. Although quality contributes to the final score, which outputs payments, there is still independent variation from the other categories to identify profit incentives.

We may be able to compute costs using a bounding exercise...

To fix ideas, we look at those that receive neutral adjustments at final score s_i that have near-max quality score q_i. Let $cost_i = cost(s_i)$ be the cost of achieving the respective final score via other categories, and let the threshold be $\bar{s}$. Since improving quality/patient health is no longer a channel through which payment could be increased, the doctor *could have* improved on the other categories, like IT/promoting interoperability, but chosen not to. So it must be the case that:

$$cost(\bar{s}) - cost(s_i) > \bar{payment}_i - 0$$

Similarly, those who max out on quality score q_i but have final scores s_i' above the threshold, it must be because the additional payment incentives exceeds the incremental cost:

$$payment(s_i') - \bar{payment}_i > cost(s_i') - cost(\bar{s})$$

Now if we pick final scores where we can assume the incremental costs from/to the threshold are the same, e.g. $cost(\bar{s}) - cost(s_i) = cost(s_i') - cost(\bar{s}) = \Delta cost$, then we can bound this change in cost:

$$payment(s_i') - \bar{payment}_i > \Delta cost > \bar{payment}_i$$

We can also do the same exercise by picking those near-max PI scores to estimate/bound the cost of providing extra units of patient health (via quality), and estimate a rich cost function describing the cost of increasing each of the components.