Sanjog Misra

Simon School of Business Administration
University of Rochester

SICS Berkeley
“Would patients take a more effective drug with higher risk instead of a less effective drug but safer drug?"

The paper models prescription drug choice as a risk-return trade-off.

- where agents are learning about the risks and returns.
- and there exists heterogeneity across agents in such trade-offs.
The “Utility" for a given patient at a given time (subscripts suppressed)

\[ U^j = f (e^j, s^j) + X^j \beta + \varepsilon^j \]

- \( e^j \) is the perceived "Effectiveness" of drug \( j \)
- \( s^j \) is the perceived (lack of) "Side effects" of drug \( j \)
- \( f \) is a CARA sub-utility function, particularly

\[ f (e, s) = -\exp (- (e + s)) \]

- \( X^j \) is a vector of preference shifters (including detailing)
- \( \varepsilon^j \) is a i.i.d. extreme value (Type I) shock
Summary: Decision Rules

- **Case 1:** New patient
  - Choose expected utility maximizing drug
    \[ E[U^j|\Omega] > E[U^k|\Omega] \quad \forall \ k \neq j \]

- **Case 2:** Existing patients
  - **Case 2a:** No switch
    \[ E[U^j|\Omega] > E[U^k|\Omega] - SC \]
  - **Case 2b:** Switched drugs where "Side effects" was cause
    \[ E[U^j|\Omega] < E[U^k|\Omega] - SC \quad \text{and} \quad s^j < E[s^k|\Omega] \quad \text{and} \quad s^j < e^j \]
  - **Case 2c:** Switched drugs where "Ineffectiveness" was cause
    \[ E[U^j|\Omega] < E[U^k|\Omega] - SC \quad \text{and} \quad e^j < E[e^k|\Omega] \quad \text{and} \quad e^j < s^j \]
  - **Case 2d:** Switched drugs where "Other causes" was cause
    \[ U^j < E[U^j'|\Omega] - SC \]
Summary: Learning and Heterogenenity

- Physicians learn in a standard Normal-Normal Bayesian framework
  - They learn about both "Effectiveness" and "Side effects"
- Physicians have (homogeneous) priors
- Informative signals come from detailing and patient feedback
  - Patient signals are correlated across drugs detailing signals are not
- Note: There is no uncertainty over patient heterogeneity only over the true "mean" drug quality.
Identification: How it works

- A main contribution of the paper is the decomposition of true mean quality into "Effectiveness" and "Side Effect" components.
- How do the authors manage this? One way to think about identification is as follows:
- If we had data, say, on \( e^j_h \) then \( \bar{E} = \frac{1}{H} \sum_h e^j_h \) and other measures follow similarly.
- In other words identifying \( \bar{E} \) and \( \bar{S} \) (and other related parameters) depends on how well one can nail down \( e^k_h \) and \( s^k_h \).
- The authors exploit the stated switching reasons to help identification. These data impose constraints on the \( e^k_h \) and \( s^k_h \).
- Additionally, the choice data also helps (weakly) by identifying \( (e^k_h + s^k_h) \)
- ... as does risk aversion (by helping identify variances)
- ... and the normality assumptions (point identification.)
Assume \( f(e^j_h, s^j_h) = \lim_{\gamma \to 0} - \exp(-\gamma (e^j_h + s^j_h)) = e^j_h + s^j_h \)
Results

- Very strong results with most effects significant.
- "First mover advantage" is huge
- Cialis most "effective" and has least "side effects"
- Detailing informative about effectiveness while patient visits are more informative about side effects.
- One detailing visit (with meal!) reduces most uncertainties, but persuasive effects persist.

Results consistent with "common wisdom" (WebMD!)
- Viagra and Levitra take 30 mins to work and last 4-5 hours
- Cialis works in 15 mints and can last 36 hours!
Comments & Issues

- **Sufficiency of Data**
  - Of the 13,619 visits there are only 929 (E) + 161 (S) informative switches.
  - Are there enough switches? Or is identification being achieved via functional form and parametric assumptions?

- **Model Assumptions**
  - *Risk aversion*:
    - Specification adopted implies that the CARA parameter $r = \frac{f''}{f'} = 1$
    - How might one justify this assumption? Is this nonlinearity essential?
  - *Switching Costs*
    - Can switching costs be independent of brands, of quality measures and marketing efforts in this category?

- **Heterogeneity vs. Learning**
  - There is no physician level heterogeneity (Priors are identical as are responses to marketing effects.)
  - Does this create a confound?
Overall a nice paper that attempts to decompose learning about multiple facets of a product.

More generally one can think of the it as using data from *consumer exit interviews* to help deconvolve otherwise unidentified effects.

Other applications might include

- Wireless carrier switches
- Bank account closures
- Job Quits

The paper also highlights the value of non-choice data.

- Any individual level data on preferences (or lack thereof) helps tease out heterogeneity better.
- Shameless promotion example: Survey data on preferences can help construct informative priors in learning models (Shin, Misra and Horsky 2007)