Micro Data, Panel Data, and Hybrids

Phil Haile
Yale University

January 2020

Micro Data

Micro Data (in discrete choice setting) = observation of *individual consumer characteristics* d_i *matched to individual consumer choices* q_i

Note:

- observing individual consumer choices without matched consumer characeteristics offers *no new information* relative to market-level data
- observing consumer characteristics not matched to individual choices is the market-level case in BLP95, Nevo, etc.

Micro Data in Practice

Examples

- classic McFadden work on individual consumer choice of mode of transport, reflecting consumer location, employment, etc.
- choice of residential location based on family demographics and demographics of neighborhood (see papers by Pat Bayer, e.g.)
- choice of hospital based on location, type of illness, insurer, etc. (see papers by Kate Ho, e.g.)
- choice of newspaper subscription based on ideological match of consumer to newspaper (see Gentzkow-Shapiro).

Gains from Micro Data?

Relative to Market Level Data

Micro data links household choices directly to household characteristics, e.g.,

- large family$\longrightarrow$like station wagon or minivan
- rich$\longrightarrow$less sensitive to price

With market level data, we learn about the *marginal* distributions $F_D(d_i), F_Q(q_i)$ of demographics and choices

With micro data, we learn about their *joint* distribution $F_{DQ}(d_i, q_i)$

We'll see that this can allow us to relax reliance on instruments.

BLP Style Preferences

$u_{ijmt} = x_{jmt}\beta_{im} + \xi_{jmt} + \epsilon_{ijmt}$, where

- $t \in$ indexes time
- $m \in \{1, \ldots, M\}$ indexes geographic markets
- $j \in \{0, 1, \ldots, J\}$ indexes goods
- $i \in \{1, \ldots, N\}$ indexes consumers
- observables $x_{jmt} \in \mathbb{R}^G$ include price
- consumer heterogeneity
  - $\epsilon_{imt} = (\epsilon_{i0mt}, \ldots, \epsilon_{iJmt})$, i.i.d. type 1 EV
  - $\beta_{im} = \beta_0 + \gamma d_{im} + \sigma\zeta_{im}$, where d_{im} are demographics, ζ_{im} is random vector $\sim \Phi(\cdot)$ iid across consumers

Note: consumer tastes β_{im} labeled as constant across time. This is not essential, and is an assumption with bite only if one has a consumer panel (otherwise, i does not define a fixed consumer across time anyway).

Discrete Choice

Rewrite

$$u_{ijmt} = \delta_{jmt} + x_{jmt}\left(\gamma d_{imt} + \sigma \zeta_{im}\right) + \epsilon_{ijmt}$$

where

$$\delta_{jmt} = x_{jmt}\beta_0 + \xi_{jmt}$$

The probability that consumer i in market m chooses good j in period t takes the form

$$\int_{\mathbb{R}^G} \frac{\exp(\delta_{jmt} + x_{jmt}\left(\gamma d_{imt} + \sigma \zeta_{im}\right))}{1 + \sum \exp(\delta_{kmt} + x_{kmt}\left(\gamma d_{imt} + \sigma \zeta_{im}\right))} d\Phi\left(\zeta_{im}\right).$$

An Estimation Approach for Micro Data

$$\int_{\mathbb{R}^G} \frac{\exp(\delta_{jmt} + x_{jmt}(\gamma d_{imt} + \sigma\zeta_{im}))}{1 + \sum \exp(\delta_{kmt} + x_{kmt}(\gamma d_{imt} + \sigma\zeta_{im}))} d\Phi(\zeta_{im})$$

Replacing j with the index of i's chosen good, this becomes the likelihood contribution of each observation as a function of the "parameters" (δ, γ, σ). (would need to simulate from Φ).

One *might* estimate all parameters by MSM, with moments:

- score of likelihoood wrt (δ, γ, σ)
- orthogonality conditions $E\left[z'\left(\delta - x_{jmt}\beta_0\right)\right] = 0$

But use of simulated likelihood/score often performs poorly (see, e.g., Train, 2003). Common to replace the score with "micro moments" of the joint distn of characteristics of csrs and their chosen-products—often using aggregation to avoid tiny cells for sample moments (example below).

Estimation with Micro Data

Notice that: micro data reduces the role of orthogonality conditions. If we dropped the IV moments, we may still be able to estimate (δ, γ, σ), which includes the "nonlinear parameters" governing substitution patterns

- Berry-Haile (2020) show this formally for nonparametric model: identification of demand with IVs only for prices
- similar benefits with other data types:
  - hybrids (market-level data + limited micro data)
  - consumer panels (e.g., Handel (2013)
  - ranked choices (e.g., BLP (2004))

  Similar estimation approaches for all of these cases too: MSM estimation using a mix of micro-moments, aggregated moments, and IV moments.

Valuing New Goods

Some Big Economic Questions

- how big are the welfare gains from innovation?
- what share of these gains are captured by the innovator?
- how big are the negative externalities ("business stealing") on other producers?
- in the case of innovation by an incumbant, how big is the negative effect of innovation on itself ("cannibalization")?
- how does innovation alter competition/market power?
- do the resulting incentives lead to too much or too little innovation?

Petrin (2002): A case study of the minivan, combining market level data and micro data.

Combining Market and Micro Data

1. Market level data (like BLP, Nevo)
    - car characteristics and market shares (same as BLP)
    - all of U.S., 1982-1993
    - household demographics for representative sample of U.S. population (CES)
2. "Micro data"
    - "CEX" : Extended Consumer Expenditure Study
    - demographics <u>and</u> new car purchases for smaller sample: 30,000 households
        - too few households to use these data alone
        - for example, only about 2700 purchases; many cars never purchased in this sample.

MSM Estimation with Market-Micro Hybrid Data

- BLP moments, including inversion of mkt shares + supply side
- "micro moments":

$$
\begin{aligned}
&E\left[i \text{ buys minivan} \middle| y_i \in \text{``bin'' } b\right] \qquad b = 1, 2, 3 \\
&E\left[fs_i \middle| \text{buy minivan}\right] \\
&E\left[fs_i \middle| \text{buy station wagon}\right] \\
&E\left[fs_i \middle| \text{S.U.V.}\right] \\
&E\left[fs_i \middle| \text{full size passenger van}\right]
\end{aligned}
$$

  - the demand model predicts these
  - expected difference between prediction and sample mean is zero at the true parameter vector

Note: not the only possible type of micro moments; but averages conditional on coarse partition are good choices when survey is small (e.g., here, 120 purchasers of minivans) or when individual market shares very small.

Counterfactual Simulation: Equilibrium with No Minicvan

Benefits (or harm) of the minivan introduction...

- to innovating firm (Chrysler)
- to imitating firms (almost all others eventually imitate)
- to firms with competing products (e.g., station wagon)
- to consumers who buy minivan
- to consumers who buy other cars at reduced price

...quantified by comparing the estimated model's predictions with vs. without minivan

Note: In general, to compare counterfactual world to real world, we want to compare the model predictions for these two worlds. Comparing actual data to counterfactual predictions would mix sampling error (and any specification error) with the effects of the counterfactual change.

Results

- big winners: Chrysler and consumers (even those not buying a minivan)
- losers: Ford, GM, at least in first few years

Here a potentially important maintained assumption is that in the absence of the minivan nothing else would have changed (e.g.,no change in entry/exit of products or firms). Relaxing this assumption would require endogenizing entry/exit etc. Steve's part of the course will move in this direction.

SCHOOL CHOICE

Shool Vouchers

- **supporters**: choice among public schools system creates only mild incentives. Privatization (e.g., allowing use of voucher at private schools) incentivizes entry and innovation, creates the threat of failure, and causes bad schools to exit. Going even farther and letting parents pay extra for higher quality will intensify supply response
- **detractors**: free market and profit motive inconsistent with good provision of quasi-public good. Allowing parents to pay for for better schools will perpetuate inequality, in effect reducing access of poor to the best schools

What happens in practice? Can targeted vouchers for poor students improve outcomes or mitigate adverse distributional effects?

Nielson 2020: School Choice and Vouchers in Chile

Outcomes (e.g., achievement, exposure to good schools) under school choice policies reflect

- family preferences, housing location
- price sensitivity (to the exent that private schools relevant)
- incentives for schools, e.g., to improve quality

What is the quantitative impact of each of these demand factors and supply factors?

Figure 1: Voucher Size

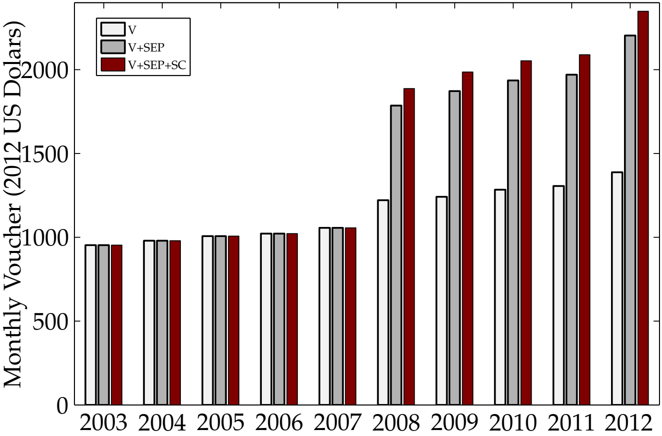
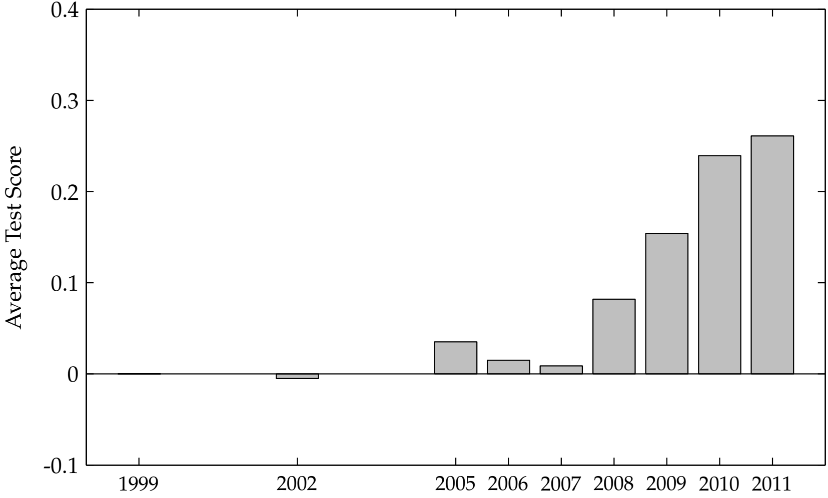


Figure 2: Average Test Scores



Estimates/Counterfactuals

From demand estimates using micro data:

- distance and price important factors, especially for families in which mother has lower education
- similar preference for school quality among all family types
- counterfactual simulations: how much of the gains to poor students' scores result from...
  1. *better choices (demand)*: poor students select better schools when price not a deterrent
  2. *school improvements (supply):* schools improve in order to compete for students who now carry a more valuable voucher
  3. *entry (supply)*

  Results: similar contributions from each. And supply-side incentives seem to matter: #2 largest where schools' profit-maximizing quality choices increases most after voucher-induced demand shifts.