Environmental and Development Economics
Module 4 - WTP for Environmental Quality in LMICs

Raahil Madhok
UMN Apllied Economics

2024-09-19

Lecture 6

WTP Estimation: Revealed Preference Models

Housekeeping

- Research Proposal: you do NOT have to carry out analysis
- First draft: Oct 3rd
- Replication: Oct 21st
  - Start soon

Today

- **Guiding question:** what is the WTP for environmental quality in LMICs?

- Today's focus: How do we even measure this?
  - Behavioral models
  - Choice probabilities
  - Estimation
  - Application to WTP

- Next time: why is WTP so low in developing countries?

Estimating Non-market Value of Goods and Services

- **Stated Preference:**
  - Hypothetical data to estimate **ex-ante** WTP
  - Contingent valuation, contingent behavior
  - Respondents directly asked about WTP (phone, mail, etc.)
  - Issue: hypothetical bias and strategic bias
- **Revealed preferences**
  - Behavioral data to estimate **ex-post** WTP
  - Travel cost, averting behavior, hedonic price
  - Pro: based on **actual** choices
  - Con (?): need a **behavioral model** in which to analyze choices

Behavioral Models

- Our goal is to define a behavioral process generating agent's choice
- Let agent face choice among set of options (products, actions, etc.)
  - What leads agent to choose one?
- Let some factors (x) be observed by researcher, and some not (ϵ)
- Factors relate to agent choice through a **behavioral process**:

$$y = h(x, \epsilon)$$

- $h(x, \epsilon)$ determines the selected choice based on these factors

Probabilistic Choice

- Since ϵ unobserved, the outcome is probabilistic
- **Choice probability** of choosing y derived by assuming $\epsilon \sim f(\epsilon)$

$$P(y|x) = \text{Prob}(\epsilon \quad \text{s.t.} \quad h(x, \epsilon) = y)$$

- This probability is much more tractable i.e., fully characterized by $f(\epsilon)$
- We can rewrite this as the expected value of an indicator function:

$$P(y|x) = \int I[h(x, \epsilon) = y] f(\epsilon) d\epsilon$$

- where $I[h(x, \epsilon) = y] = 1$ when statement in brackets is true
- Integral of behavioral process indicator over all possible values of unobservables

Evaluating the Choice Probability

- Method 1: Closed-form solution
  - For certain specifications of h and f, $P(y|x)$ calculated via a formula
  - Main example is Logit

- Method 2: Simulation
  - Some specifications of h and f have no closed form solution
  - We simulate $P(y|x)$ by taking random draws from f and taking the average

Evaluating the Choice Probability: Closed Form

First, define the behavioral model

- Consider a binary model where agent considers whether to take an action
- Agent gets utility (+/-) from the action:

$$U = \beta' x + \epsilon$$

- U: Utility of taking an action; $\beta' x$: Observed component of utility.
- ϵ: Unobserved component; lets assume **logistic distribution**.
- Take action if utility is positive:

$$P = \int I[\beta' x + \epsilon > 0] f(\epsilon) d\epsilon$$

Closed Form Solution

$$
\begin{aligned}
P &= \int I[\beta' x + \epsilon > 0] f(\epsilon) d\epsilon \\
&= \int_{\epsilon > -\beta' x} f(\epsilon) d\epsilon \\
&= 1 - F(-\beta' x) \\
&= 1 - \frac{1}{1 + e^{\beta' x}} \\
&= \frac{e^{\beta' x}}{1 + e^{\beta' x}}
\end{aligned}
$$

Evaluating the Choice Probability: Simulation

Recall: $P(y|x) = \int I[h(x,\epsilon) = y] f(\epsilon) d\epsilon$

- **Step 1:** Take draw of ϵ from $f(\epsilon)$
  - Label it ϵ^1, denoting first draw
- **Step 2:** Check whether $h(x, \epsilon^1) = y$
  - If so, create $I^1 = 1$, otherwise $I^1 = 0$
- **Step 3:** Repeat this R times and collect I^r for $r = 1, ..., R$
- **Step 4:** Average **simulated** probability: $P(i|x) = \frac{1}{R}\sum_{r=1}^{R} I^r$
  - Proportion of times that draws of ϵ, when combined with x, yield y

Lets define a realistic behavioural model

Grounded in economic theory

Random Utility Model

Behavioural model:

- Agent, n faces choice among J alternatives
  - Choice set must be 1) mutually exclusive, 2) exhaustive, 3) finite
- Utility from j is: U_{nj} for $j = 1, ..., J$
- Utility known to agent but not the researcher
- Agent chooses j that maximizes utility:

$$\text{Choose i iff} \quad U_{ni} > U_{nj} \quad \forall j \neq i$$

Random Utility Model

The Researcher

- Observes attributes of alternatives, $x_{nj} \quad \forall j$
- Observed attributes of agent, s_n
- Let $V_{nj} = V(x_{nj}, s_n) \quad \forall j$ (representative utility)
- **V can depend on parameters unknown to researcher**
  - But can be structurally estimated...

Random Utility Model

- Since $V_{nj} \neq U_{nj}$, we can write:

$$U_{nj} = V_{nj} + \epsilon_{nj}$$

  - where joint density of random vector $\epsilon_n = (\epsilon_{n1}, ..., \epsilon_{nJ}) = f(\epsilon_n)$

- Then the probability than agent n chooses alternative i is:

$$\begin{aligned} P_{ni} &= P(U_{ni} > U_{nj}, \forall j \neq i) \\ &= P(V_{ni} + \epsilon_{ni} > V_{nj} + \epsilon_{nj}, \forall j \neq i) \\ &= P(\epsilon_{nj} - \epsilon_{ni} < V_{ni} - V_{nj}, \forall j \neq i) \leftarrow \text{CDF!} \\ &= \int I(\epsilon_{nj} - \epsilon_{ni} < V_{ni} - V_{nj}) f(\epsilon) d\epsilon \end{aligned}$$

- Different DCMs obtained from different choices of $f(\epsilon)$

Logit

- Assumes ϵ_{nj} distributed iid extreme value

$$\text{PDF:} \quad f(\epsilon_{nj}) = e^{-\epsilon_{nj}} e^{-e^{-\epsilon_{nj}}}, \quad \text{CDF:} \quad F(\epsilon_{nj}) = 1 - e^{-e^{\epsilon_{nj}}}$$

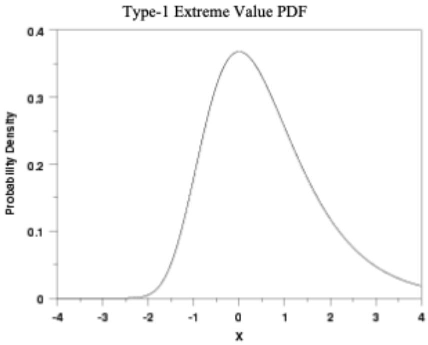
- Note: if ϵ's are iid EV, then $\epsilon^*_{nji} = \epsilon_{nj} - \epsilon_{ni}$ is distributed logistic:

$$F(\epsilon^*_{nji}) = \frac{e^{\epsilon^*_{nji}}}{1 + \epsilon^*_{nji}}$$

Aside: Extreme Value Distribution

Useful Properties

1. Mathematical tractability
   - Closed form choice probability
2. Long right tail
   - Unobservables can drive unexpected choices
3. IIA property
4. Connection to logistic regression
5. Easy to (structurally) estimate



Logit

- The probability that agent n chooses alternative i:

$$
\begin{aligned}
P_{ni} &= P(V_{ni} + \epsilon_{ni} > V_{nj} + \epsilon_{nj} \quad \forall j \neq i) \\
&= P(\epsilon_{nj} < \epsilon_{ni} + V_{ni} - V_{nj} \quad \forall j \neq i) \leftarrow \text{invoke iid assumption!} \\
&= \int \left(\prod_{j \neq i} e^{-e^{-(\epsilon_{ni} + V_{ni} - V_{nj})}} \right) e^{-\epsilon_{ni}} e^{-e^{-\epsilon_{ni}}} \, d\epsilon_{ni} \\
&= \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}}
\end{aligned}
$$

- **Preview:** If V is linear in parameters, $V_{nj} = \beta' x_{nj}$, then: $P_{ni} = \frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}}$

Independence of Irrelevant Alternatives

- Crucial assumption in logit model
- Relative odds of choosing b/w 2 i and k does not change when j introduced
- For any two alternatives i and k, the ratio of logit probabilities is:

$$\frac{P_{ni}}{P_{nk}} = \frac{e^{V_{ni}} / \sum_j e^{V_{nj}}}{e^{V_{nk}} / \sum_j e^{V_{nj}}} = \frac{e^{V_{ni}}}{e^{V_{nk}}} = e^{V_{ni} - V_{nk}}$$

- Ratio does not depend on any alternatives other than i and k!
  - Denominators cancel

Discussion of IIA Assumption

- Advantages

- Disadvantages

5 minute break

Estimation: Maximum Likelihood Structural Estimation

- Suppose we have have data on N agent's choices
- The likelihood of observing the choices made by each individual is:

$$L(\beta) = \prod_{n=1}^{N} \prod_{i} (P_{ni})^{y_{ni}}$$

- where $y_{ni} = 1$ if alternative i is chosen, and 0 otherwise
- Log-likelihood is easier to maximize:

$$LL(\beta) = \sum_{n=1}^{N} \sum_{i} y_{ni} \log(P_{ni})$$

- ML estimator is the β that maximizes this function

Method of Moments

- FOC of $LL(\beta) \leftrightarrow \frac{\partial LL(\beta)}{\partial \beta} = 0$
- This can be rewritten as:

$$\frac{1}{N}\sum_{n=1}^{N}\sum_{i} y_{ni}x_{ni} = \frac{1}{N}\sum_{n=1}^{N}\sum_{i} P_{ni}x_{ni}$$

- LHS: mean x over choices by sampled agents
- RHS: mean x over **predicted** choices by sampled agents
- β found s.t. predicted mean of each explanatory variable equals observed mean
  - Model reproduces observed averages in the data
  - Also known as **matching on the first moment**

Limitations

- Cannot represent random taste variation
- Implies proportional substitution across alternatives (IIA)

What if we want to estimate more moments of β?

Specifying Taste Variation

- Supposed agent now has the utility: $U_{nj} = \beta_n{}' x_{nj} + \epsilon_{nj}$
  - where $\beta_n \sim f_n(\beta|\Theta)$ and ϵ_{nj} still iid EV
- If we observed β_n, then we are back to logit: $P_{ni}(\beta_n) = \frac{e^{\beta_n' x_{ni}}}{\sum_j e^{\beta_n' x_{nj}}}$
- Instead, we integrate logit over all possible parameters of β_n:

$$P_{ni} = \int \underbrace{\left(\frac{e^{\beta_n' x_{ni}}}{\sum_j e^{\beta_n' x_{nj}}} \right)}_{\text{logit probability}} \underbrace{f(\beta|\Theta) d(\beta)}_{\text{mixing distribution}} \qquad \leftarrow \text{no closed form solution!}$$

- Mixing distribution usually normal: $f(\beta|\Theta) = f(\beta|\mu, \sigma)$
- Note: P_{ni} is a function of Θ, not β! (β is integrated out)

Aside: Properties of Mixed Logit

- Relaxes IIA assumption!
- Ratio of mixed logit probabilities, P_{ni}/P_{nj} depends on all data
  - Denominators are inside the integral and do not cancel
- Enables correlation in unobserved factors across alternatives
  - Flexible substitution patterns

Steps for Simulating Mixed Logit Probability

1. Draw a value β^r from $f(\beta|\Theta)$, where r is which draw
2. Calculate the logit formula $P_{ni}(\beta^r)$ using the drawn value
3. Repeat steps 1 and 2 for R draws.
4. Average the results to obtain the simulated probability:

$$\hat{P}_{ni} = \frac{1}{R}\sum_{r=1}^{R} P_{ni}(\beta^r)$$

Simulated Log-Likelihood Estimation

- As before, insert simulated $\hat{P}_{ni}$ into log-likelihood:

$$SLL = \sum_{n=1}^{N} \sum_{j=1}^{J} y_{nj} \ln \hat{P}_{ni}$$

- where $y_{nj} = 1$ if n chose j and zero otherwise
- Maximum simulated likelihood estimator (MSLE) is the Θ that maximizes SLL

This recovers taste distribution (μ, σ) that matches observed choices to predicted choice probabilities as close as possible

Lets do a simple application

What is the WTP for biodiversity?

WTP for Species Diversity in India

- **Context:** people travel to "hotspots" to see pretty birds
  - Birds are a good proxy for overall biodiversity
- **Data:** individual level data from citizen science app (eBird)
  - location, alternative sites, income, **travel cost**, **species richness**
- **Behavioral model:** choose site that maximizes utility
- **Question:** what cost will you incur to $\uparrow$ utility from biodiversity?
  - How does your "price" change for an additional unit of biodiversity? (MWTP)

Deriving WTP in Practice

- Agent i's utility from site j at time t is:

$$\begin{aligned} U_{ijt} &= V_{ijt} + \epsilon_{ijt} \\ &= \underbrace{\beta_1}_{\text{MU Income}} (y_i - c_{ijt}) + \underbrace{(\beta_2 + \eta_i)}_{\text{MU EQ}} e_{jt} + \ldots + \epsilon_{ijt} \end{aligned}$$

- If stay home: $U_{i0t} = \beta_1 y_i + \epsilon_{i0t}$
- Cost s.t. i indifferent b/w travelling and avoiding cost:

$$c_{ijt}* = \frac{1}{\beta_1}[(\beta_2 + \eta_i)e_{jt} + \beta_3 x_{ijt} + \ldots + (\epsilon_{ijt} - \epsilon_{i0t})]$$

- MWTP for one more species, assuming $\eta_i \sim N(0, \sigma)$:

$$MWTP = \boxed{\frac{\partial c_{ijt}*}{\partial d_{jt}} = \frac{\hat{\beta}_2}{\hat{\beta}_1} + \cancel{\frac{\hat{\eta}_i}{\hat{\beta}_1}}}$$

Implementation

- Data
  - attributes of choice and alternatives for each **choice occasion**
  - Variable *choice=1* if chosen, otherwise 0
  - Data can get huge, and simulation takes forever
  - solution: restrict choice set
- Stata
  - clogit: conditional logit (fast)
  - mixlogit: mixed logit estimation (slow)
- R
  - *mlogit* library
  - Exercises and applications: https://cran.r-project.org/web/packages/mlogit/vignettes/c5.mxl.html

Aside

- **If you want to get involved in this project, let me know**
- Goal: estimate WTP for species diversity in India

–>

–> –>