A normalization approach to discrete choice models in willingness-to-pay space

Ricardo A Daziano
Cornell University, CEE
July 2013
Willingness to pay (WTP)

- Marginal rate of substitution between a qualitative attribute and money
- Amount of money that a consumer is WTP for a qualitative improvement (time savings, energy-efficiency gains)
- DCM: Let $V_{ij}$ be the deterministic utility of individual $i$ and alternative $j$
- $q_{ij}$ is a vector of $K$ qualitative attributes
- $I_i$ is income of individual $i$, then

$$WTP_{q_{kij}} = \frac{\partial V_{ij}}{\partial q_{kij}} \frac{\partial V_{ij}}{\partial I_i}$$
WTP estimation

Consider the linear specification $V_{ij} = q'_{ij}\beta - \lambda p_{ij}$

$\lambda$ is the marginal utility of income ($= -\beta_p$)

Meaningful function with a clear economic interpretation (cf. marginal utilities $\beta$)

The WTP point estimation problem reduces to inference on parameter ratios
WTP inference

- Most studies (in transportation) report only WTP point estimates, without any measure of uncertainty
- Providing confidence sets for WTP measures is relevant
- WTP confidence sets are necessary for comparing the results of competing models
- and for determining whether there is consumer heterogeneity
- **Constructing WTP confidence sets does not have a clear answer even for the simplest case of a parameter ratio**
Inference on parameter ratios

- The ratio of the marginal utilities – which are asymptotically normal – is **locally almost unidentified**
- Conventional techniques, such as the delta method, break down even for large samples (Brownstone, 2001; Bolduc et al. 2010)
  - CIs too narrow
- Frequentist methods (Krinsky Robb - Bootstrapping, Simulation) exhibit several problems (see Czajkowski and Carson, 2013)
  - Poor empirical coverage
  - Not really compatible with weak identification
- **Fieller-CIs**: inverting Walt-type tests (asymptotic solution)
The problem of WTP heterogeneity distributions

- Previous literature focused on heterogeneity distributions
- Random parameters in the context of preference heterogeneity (mixed logit)
- Several authors have analyzed this problem from a frequentist perspective (Algers et al., 1998; Revelt and Train, 1998; Hensher and Greene, 2003)
- Heterogeneity distributions that may be unbounded or may lack finite moments
- Partial solutions:
  1. Fixing the cost parameter
  2. Assuming lognormally distributed parameters
  3. **Reparameterization of the parameter space** (Train and Sonnier, 2004)
From preference to WTP space

- Following Train and Weeks (2005)

\[ V_{ij} = q'_{ij} \beta - \lambda p_{ij} = \lambda q'_{ij} \left( \frac{1}{\lambda} \right) \beta - \lambda p_{ij} = \lambda q'_{ij} \text{WTP} - \lambda p_{ij} \]

- The reparameterization transforms the space parameter to WTP space (with one dimension being the MUI)
- It also transforms the specification of utility from a linear to a nonlinear problem
- The added nonlinearity may explain losses in fit (Train and Weeks, 2005)
- Empirical applications show that the reparameterized models produce more realistic estimates of willingness-to-pay (Sonnier et al., 2007; Scarpa et al., 2008; Balcombe et al., 2009)
This paper: revisiting normalization

- DCM: choice probabilities are invariant to location and scale
- Scale must be normalized to ensure identification of the model parameters
- Dominant approach: fixing some element of the variance of the cumulative distribution of the error term of the model in differences
- Implicitly assume a standard distribution by normalizing a nuisance parameter to one
- However, **any parameter of the model can be normalized**
- Normalizing any $\beta$ is standard in semiparametric estimators of binary outcome (logistic) regressions (Ichimura, 1993; Klein and Spady, 1993)
A normalization approach to WTP space

- Idea: normalizing the marginal utility of income

\[
U_{ij} = \lambda (q'_{ij} WTP - p_{ij}) + \varepsilon_{ij} \Leftrightarrow
\]

\[
CS_{ij} = \frac{U_{ij}}{\lambda} = q'_{ij} WTP - p_{ij} + \frac{\varepsilon_{ij}}{\lambda}
\]

- Consumer surplus \((CS_{ij})\) model (cf. Jedidi et al., 2003)

\[
CS_{ij} = q'_{ij} WTP - p_{ij} + \varepsilon_{ij}
\]

- Since price \(p_{ij}\) is a continuous variable, the normalization of \(\lambda\)
  ensures identification of the model (Klein and Spady, 1993)

- Under this normalization, scale of \(\varepsilon_i\) is identified (up to the
  normalized parameter) and can be thus estimated
Equivalence of both approaches

- Working in preference or willingness-to-pay space leads to estimators that are equivalent
- Invariance property of the frequentist MLE
- In GEV models normalization of $\mu$ (Train and Weeks, 2005) or $\lambda$ produces no difference in the MLE implementation

$$U_{ij} = \lambda q'_{ij} \text{WTP} - \lambda p_{ij} + \varepsilon_{ij} = \lambda (q'_{ij} \text{WTP} - p_{ij}) + \varepsilon_{ij}$$

- Implementation of both reparameterization approaches is not the same for **probit models**
- This fact can be exemplified using both frequentist and Bayes estimators
A multinomial probit model for consumer-surplus maximization

\[ CS_i = Q_i WTP - p_i + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \Sigma) \]

- Location set by working with the model in differences (as usual)
- All parameters in the equation above are normalized (no need for further normalizations)
- All nuisance parameters in the reduced-dimension covariance \( \Delta_j \Sigma \Delta'_j \) can be estimated
- Frequentist estimator: MSLE using the GHK simulator (without normalizing the first element in the Cholesky root)
Bayes estimator

- If the dependent variable were observable we could use an ordinary regression
- \( CS_i \) is treated as a latent variable and we use the DCM framework
- In a Bayesian setting, the parameter space is augmented treating latent variables as additional parameters to estimate
- Subject to the feasibility region imposed by the choice indicators, \( CS_i \) has a truncated normal distribution
- This distribution can be used to sample realizations of the consumer surplus
- Parameters of the model are estimated using a Bayesian regression!
Bayes estimator cont’d 1

- Linear regression with a **partially observed** dependent variable

\[
\Delta_j CS_i + \Delta_j p_i = \Delta_j Q_i WTP + \Delta_j \epsilon_i, \quad \Delta_j \epsilon_i \sim \mathcal{N}(0_{J-1 \times 1}, \Delta_j \Sigma \Delta_j')
\]

- Let \( \mathbf{L} \) be the Cholesky root of \((\Delta_j \Sigma \Delta_j')^{-1}\), then

\[
\mathbf{L}' \Delta_j CS_i + \mathbf{L}' \Delta_j p_i = \mathbf{L}' \Delta_j Q_i WTP + \mathbf{L}' \Delta_j \epsilon_i, \quad \mathbf{L}' \Delta_j \epsilon_i \sim \mathcal{N}(0, \mathbf{I}_{(J-1 \times J-1)})
\]

- Data augmentation step still needed for producing observations of \(CS_i\) (truncated normally distributed vector)
- \(WTP\) estimated using Bayesian OLS
- Elements of \(\mathbf{L}\) sampled from an inverted Wishart distribution
Bayes estimator cont’d 2

- Start at any given point \( \text{CS}_i^{(0)}, \theta^{(0)} = \text{WTP}^{(0)} \), and \( \Sigma^{(0)-1} = L^{(0)}L^{(0)\prime} \) in the parameter space.

- For \( g \in \{1, \ldots, G\} \)
  1. \( \Delta_j \text{CS}_i^{(g)} \sim \mathcal{N}(\Delta_jQ_i\theta^{(g-1)} - \Delta_j p_{ij}, \Delta_j \Sigma^{(g-1)} \Delta_j')1(\Delta_{jj'}\text{CS}_{ij'} < 0, \forall j' \neq j) \)
  2. \( \text{WTP}^{(g)} \sim \mathcal{N}\left((\tilde{\nu}_\theta^{-1} \tilde{\theta} + (L^{(g-1)'} \Delta_j Q)')(L^{(g-1)'} \Delta_j Q)^{-1}\times(\tilde{\nu}_\theta^{-1} + Q' \Delta_j L^{(g-1)}(L^{(g-1)'} \Delta_j Q'))^{-1}\right) \)
  3. \( \Sigma^{(g)} \sim IW(\tilde{\nu} + N, \tilde{\Sigma} + \sum_{i=1}^{N} \Delta_j (\text{CS}_i^{(g)} + p - Q_i \theta^{(g)})(\text{CS}_i^{(g)} + p - Q_i \theta^{(g)}')) \Delta_j' \)

- Because the long-run distribution of the iterative process is the true posterior of interest, each sample \( \theta^{(g)} \) can be treated as a random draw of the joint posterior.
Automakers are engineering low-carbon alternatives to the standard internal combustion engine and introducing into the market a mix of different technologies.

Internal combustion engines are highly inefficient (tank-to-wheel energy efficiency is about 15%; engine loss is 76%).

Tank-to-wheel efficiency of 100% electric vehicles: $\sim 85\%$

Automakers need to respond to efficiency standards.

Consumer willingness-to-pay is key to inform pricing strategies in new and early markets.
Data: vehicle purchase decisions in California (Train & Hudson, 2000)

- Each of the 500 respondents in the survey answered up to 15 experimental choice situations between three vehicle alternatives (unlabeled)
- The sample contains a total of 7437 observations
Attributes

1. Energy source [gas internal combustion, electric, gas-electric hybrid]
2. Purchase price [thousands of US dollars]
3. Operating cost [US dollars per month]
4. Driving range for electric vehicles [hundreds of miles]
5. Performance [high, medium, low]
6. Body type [mini car, small car, mid car, large car, small SUV, mid SUV, large SUV, compact PU, full PU, minivan]
DCE

<table>
<thead>
<tr>
<th>Vehicle Type</th>
<th>Vehicle A</th>
<th>Vehicle B</th>
<th>Vehicle C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engine Type</td>
<td>Large SUV</td>
<td>Mini Car</td>
<td>Compact PU</td>
</tr>
<tr>
<td>Performance</td>
<td>Gasoline</td>
<td>Electric</td>
<td>Hybrid</td>
</tr>
<tr>
<td>Top Speed 80MPH</td>
<td>0-60: 16 Seconds</td>
<td>0-60: 16 Seconds</td>
<td>0-60: 16 Seconds</td>
</tr>
</tbody>
</table>

**Total Purchase Price**

- $36,298
- $16,594
- $33,025

**Operating Cost**

- $56.70/mo (less routine maintenance)
- $7.85/mo
- $29.10/mo

**Range (in miles)**

- 300 - 500
- 140 - 150
- 400 - 700

*Total Purchase Price is the amount customers can expect to pay for the vehicle new at the dealership

**Figure:** Sample of a choice situation presented to respondents of the stated preference experiment. Source: Train and Hudson (2000).
Stylized facts

- Average WTP for improvements in range (at 100 miles): 108 $/mile (median: 105 $/mile)
- 95% CI: [83, 134]
- At 100 miles the marginal cost of producing batteries with an additional mile of range is 160 $/mile
  - Given the current lower bound of the cost of lithium-ion batteries (475 [$/kWh])
- The cost of range improvements are greater than the amount most people are willing to pay
- No wonder BEV sales have been so low even with subsidies
- More in a companion paper (Daziano, RESEN 2013)
Vehicle choice in Canada – SP, 2002

- 1877 pseudoindivduals
- Four vehicle types:
  1. **Standard gasoline vehicle** (SGV): operating on gasoline or diesel,
  2. **Alternative fuel vehicle** (AFV): natural-gas vehicle,
  3. **Hybrid vehicle** (HEV): gasoline-electric, and
  4. **Hydrogen fuel cell vehicle** (HFC).

Attributes:

1. **Capital Cost** \( cc \): purchase price
2. **Operating Cost** \( fc \): fuel cost
3. **Fuel Available** \( fa \): percentage of stations selling the proper fuel type
4. **Express Lane Access** \( exp \): whether or not the vehicle would be granted express lane access
5. **Power** \( pow \): power compared to their current vehicle
WTP estimates

Results

<table>
<thead>
<tr>
<th></th>
<th>Point est.</th>
<th>2.5%</th>
<th>5%</th>
<th>95%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduction in Fuel Cost [$/-month]</td>
<td>94.78</td>
<td>91.49</td>
<td>94.44</td>
<td>95.51</td>
<td>96.36</td>
</tr>
<tr>
<td>Increase in Fuel Availability [$/1-% inc.]</td>
<td>156.18</td>
<td>75.37</td>
<td>85.71</td>
<td>309.09</td>
<td>376.60</td>
</tr>
<tr>
<td>Access to Express lane [$]</td>
<td>1797.75</td>
<td>74.63</td>
<td>317.46</td>
<td>4909.09</td>
<td>6382.98</td>
</tr>
<tr>
<td>Increase in Power [$/1-% inc.]</td>
<td>307.87</td>
<td>100.00</td>
<td>125.40</td>
<td>701.82</td>
<td>885.11</td>
</tr>
</tbody>
</table>

For example, due to operating cost savings, consumers would be willing to pay about $6,000 for buying the electric version of the Ford Focus (105 MPGe vs. 40 MPG)

The actual difference in price is $16,000 (tight CI)

Some of the CI are wide (Access to Express Lane)
Summary

1. Consumer valuation is a relevant input for welfare analysis, policy-making, and marketing strategies.
2. Concept of willingness-to-pay (WTP) for qualitative improvements.
3. Economics: marginal rate of substitution between quality of a discrete good and money.
4. Statistics: if utility is linear, WTP is a parameter ratio (nonlinear transformation).
5. Inference on nonlinear transformations of the model parameters encounters statistical problems.
6. Parameter ratios are often weakly identified and the asymptotic distribution of the ratio lacks finite moments.
Conclusions

- Normalizing scale can be done by fixing the MUI.
- This alternative normalization recasts the parameter space from preferences to WTP.
- Fixing the MUI is technically the same as working with the reparameterization of Train and Weeks for GEV models (and GEV kernels).
- For probit models, fixing the MUI is simpler than the reparameterization.
- The Bayes probit estimator is relatively simpler in WTP space.
Next steps

- Non-linear specification of utility
  1. Frequentist methods only apply to parameter ratios
  2. Postprocessing is feasible to any transformation of the parameters
- Other Bayesian methods for WTP inference (test inversion)
- Random consumer heterogeneity and inference on consumer surplus (and choice probabilities)
Thank you!