Measurement of Consumer Welfare

Aviv Nevo

University of Pennsylvania and NBER

Fall 2018

Introduction

- A common use of empirical demand models is to compute consumer welfare
- We will focus on welfare gains from the introduction of new goods
- The methods can be used more broadly:
    - other events: e.g., mergers, regulation
    - CPI
- In this lecture we will cover
    - Hausman (96): valuation of new goods using demand in product space
    - consumer welfare in DC models

Hausman, "Valuation of New Goods Under Perfect and Imperfect Competition" (NBER Volume, 1996)

- Suggests a method to compute the value of new goods under perfect and imperfect competition
- Looks at the value of a new brand of cereal – Apple Cinnamon Cheerios
- Basic idea:
  - Estimate demand
  - Compute "virtual price" – the price that sets demand to zero
  - Use the virtual price to compute a welfare measure (essentially integrate under the demand curve)
  - Under imperfect competition need to compute the effect of the new good on prices of other products. This is done by simulating the new equilibrium

Data

Monthly (weekly) scanner data for RTE cereal in 7 cities over 137 weeks

Note: the frequency of the data. Also no advertising data.

Multi-level Demand Model

- Lowest level (demand for brand within segment): AIDS

$$s_{jt} = \alpha_j + \beta_j \ln(y_{gt}/\pi_{gt}) + \sum_{k=1}^{J_g} \gamma_{jk} \ln(p_{kt}) + \varepsilon_{jt}$$

where,

  - s_{jt} dollar sales share of product j out of total segment expenditure
    %item y_{gt} overall per capita segment expenditure
  - π_{gt} segment level price index
  - p_{kt} price of product k in market t.

π_{gt} (segment price index) is either Stone logarithmic price index

$$\pi_{gt} = \sum_{k=1}^{J_g} s_{kt} \ln(p_{kt})$$

or

$$\pi_{gt} = \alpha_0 + \sum_{k=1}^{J_g} \alpha_k p_k + \frac{1}{2} \sum_{j=1}^{J_g} \sum_{k=1}^{J_g} \gamma_{kj} \ln(p_k) \ln(p_j).$$

Multi-level Demand Model

- Middle level (demand for segments)

$$\ln(q_{gt}) = \alpha_g + \beta_g \ln(Y_{Rt}) + \sum_{k=1}^{G} \delta_k \ln(\pi_{kt}) + \varepsilon_{gt}$$

  where

  - q_{gt} quantity sold of products in the segment g in market t
  - Y_{Rt} total category (e.g., cereal) expenditure
  - π_{kt} segment price indices

Multi-level Demand Model

- Top level (demand for cereal)

$$\ln(Q_t) = \beta_0 + \beta_1 \ln(I_t) + \beta_2 \ln \pi_t + Z_t\delta + \varepsilon_t$$

where

- Q_t overall consumption of the category in market t
- I_t real income
- π_t price index for the category
- Z_t demand shifters

Estimation

- Done from the bottom level up;
- IV: for bottom and middle level prices in other cities.

Table 5.6: overall elasticities for family segment

Table 5.6 **Overall Elasticities for Family Segment of RTE Cereal**

	Cheerios	Honey-Nut Cheerios	Apple-Cinnamon Cheerios	Corn Flakes	Kellogg's Raisin Bran	Rice Krispies	Frosted Mini-Wheats	Frosted Wheat Squares	Post Raisin Bran
Cheerios	−1.92572	0.01210	0.04306	−0.02798	0.03380	−0.20642	0.23990	0.18758	−0.51019
	(0.05499)	(0.04639)	(0.07505)	(0.06123)	(0.05836)	(0.07398)	(0.06455)	(0.10703)	(0.14309)
Honey-Nut Cheerios	0.03154	−1.98037	0.21247	−0.21316	0.07136	0.00079	−0.05929	0.32712	−0.16719
	(0.03080)	(0.05808)	(0.06808)	(0.04805)	(0.04861)	(0.05199)	(0.06752)	(0.12496)	(0.11643)
Apple-Cinnamon Cheerios	0.01747	0.08317	−2.17304	−0.04561	0.05287	−0.00824	−0.04682	−0.14074	−0.03304
	(0.01919)	(0.02690)	(0.07525)	(0.03144)	(0.03224)	(0.03111)	(0.04591)	(0.08462)	(0.08000)
Corn Flakes	0.07484	−0.13069	−0.02343	−2.16585	0.15311	−0.01918	0.03460	0.13556	−0.03062
	(0.03008)	(0.03850)	(0.06503)	(0.06155)	(0.04759)	(0.04555)	(0.06405)	(0.10926)	(0.11573)
Kellogg's Raisin Bran	0.03995	0.06155	0.12056	0.07455	−2.06965	−0.28837	0.36331	0.46661	−0.60598
	(0.03184)	(0.04109)	(0.07011)	(0.05064)	(0.07614)	(0.05456)	(0.06673)	(0.11558)	(0.13005)
Rice Krispies	−0.02457	0.08459	0.07548	−0.00219	−0.21300	−2.17246	0.07967	−0.15285	0.47670
	(0.03109)	(0.03368)	(0.05384)	(0.04071)	(0.04308)	(0.06354)	(0.04854)	(0.07886)	(0.11284)
Frosted Mini-Wheats	0.10797	−0.04239	−0.06872	−0.03001	0.24504	−0.00943	−2.55178	0.78352	−0.09987
	(0.02567)	(0.04189)	(0.06978)	(0.04629)	(0.04735)	(0.04162)	(0.11603)	(0.16839)	(0.11360)
Frosted Wheat Squares	0.01315	0.03020	−0.03440	0.00473	0.05064	−0.02772	0.12664	−3.17781	−0.06489
	(0.00656)	(0.01217)	(0.02015)	(0.01216)	(0.01274)	(0.01045)	(0.02682)	(0.13863)	(0.03082)
Post Raisin Bran	−0.02239	0.04018	0.07738	0.06288	−0.16016	0.26985	0.04499	−0.14035	−2.62151
	(0.02908)	(0.03840)	(0.06837)	(0.04415)	(0.04953)	(0.04521)	(0.06495)	(0.11447)	(0.15447)

Welfare

- Value of AC-Cheerios
- Under perfect competition approx. $78.1 million per year for the US
- Imperfect competition: needs to simulate the world without AC Cheerios
    - assumes Nash Bertrand
    - ignores effects on competition
    - finds approx $66.8 million per year;
- Extrapolates to an overall bias in the CPI 20%-25% bias.

Comments

- Most economists find these numbers too high
  - are they really?
- Questions about the analysis
  - IVs (advertsing)
  - computation of Nash equilibrium (has small effect)

Consumer Welfare Using the Discrete Choice Model

- Assume the indirect utility is given by

$$u_{ijt} = x_{jt}\beta_i + \alpha_i p_{jt} + \xi_{jt} + \varepsilon_{ijt}$$

  ε_{ijt} i.i.d. extreme value

- The *inclusive value* (or social surplus) from a subset $A \subseteq \{1, 2, ..., J\}$ of alternatives:

relative to outside good being 0

$$\omega_{iAt} = \ln\left(\sum_{j \in A} \exp\left\{x_{jt}\ \beta_i + \alpha_i\ p_{jt} + \xi_{jt}\right\}\right)$$

for exp util need Euler constant

- The expected utility from A prior to observing $(\varepsilon_{i0t}, ...\varepsilon_{iJt})$, knowing choice will maximize utility after observing shocks.
- Note
  - If no hetero ($\beta_i = \beta$, $\alpha_i = \alpha$) IV captures average utility in the population;
  - with hetero need to integrate over it
  - if utility linear in price convert to dollars by dividing by α_i
  - with income effects conversion to dollars done by simulation

Applications

- Trajtenberg (JPE, 1989) estimates a (nested) Logit model and uses it to measure the benefits from the introduction of CT scanners
  - does not control for endogeneity (pre BLP) so gets positive price coefficient
  - needs to do "hedonic" correction in order to do welfare
- Petrin (JPE, 2003) uses the BLP data to repeat the Trajtenberg exercise for the introduction of mini-vans
  - e.g. mkt share by income gp
  - adds micro moments to BLP estimates good thing to add
  - predictions of model with micro moments more plausible
  - attributes this to "micro data appear to free the model from a heavy dependence on the idiosyncratic logit "taste" error

Table 5: RC estimates

TABLE 5
RANDOM COEFFICIENT PARAMETER ESTIMATES

	Random Coefficients (γ's)	
Variable	Uses No Microdata (1)	Uses CEX Microdata (2)
Constant	1.46	3.23
	(.87)*	(.72)**
Horsepower/weight	.10	4.43
	(14.15)	(1.60)**
Size	.14	.46
	(8.60)	(1.07)
Air conditioning standard	.95	.01
	(.55)*	(.78)
Miles/dollar	.04	2.58
	(1.22)	(.14)**
Front wheel drive	1.61	4.42
	(.78)**	(.79)**
γ_{mi}	.97	.57
	(2.62)	(.10)**
γ_{sw}	3.43	.28
	(5.39)	(.09)**
γ_{su}	.59	.31
	(2.84)	(.09)**
γ_{pv}	4.24	.42
	(32.23)	(.21)**

Table 8: welfare estimates

TABLE 8
AVERAGE COMPENSATING VARIATION CONDITIONAL ON MINIVAN PURCHASE, 1984: 1982–84 CPI-ADJUSTED DOLLARS

	OLS Logit	Instrumental Variable Logit	Random Coefficients	Random Coefficients and Microdata
Compensating variation:				
Median	9,573	5,130	1,217	783
Mean	13,652	7,414	3,171	1,247
Welfare change from difference in:				
Observed characteristics ($\delta_j+\mu_{ij}$)	−81,469	−44,249	−820	851
Logit Error (ϵ_{ij})	95,121	51,663	3,991	396
Income of minivan purchasers:				
Estimate from model	23,728	23,728	99,018	36,091
Difference from actual (CEX)	−15,748	−15,748	59,542	−3,385

Discussion

- The micro moments clearly improve the estimates and help pin down the non-linear parameters
- What is driving the change in welfare?
- One option
  - welfare is an order statistic
  - by adding another option we increase the number of draws
  - hence (mechanically) increase welfare
  - as we increase the variance of the RC we put less and less weight on this effect

A different take

- The analysis has 2 steps
  1. Simulate the world without \with minivans (depending on the starting point)
  2. Summarize the simulated \observed prices and quantities into a welfare measure
- Both steps require a model
- If we observe pre- and post- introduction data might avoid step 1
  - does not isolate the effect of the introduction
- Logit model fails (miserably) in the first step, but can deal with the second
  - just to be clear: heterogeneity is important
  - NOT advocating for the Logit model
  - just trying to be clear where it fails

Red-bus-Blue-bus problem Debreu (1960)

Indep of Irr. Assump.

- Originally, used to show the IIA problem of Logit
- Worst case scenario for Logit
- Consumers choose between driving car to work or (red) bus
  - working at home not an option
  - decision of whether to work does not depend on transportation
- Half the consumers choose a car and half choose the red bus
- Artificially introduce a new option: a blue bus
  - consumers color blind
  - no price or service changes
- In reality half the consumers choose car, rest split between the two color buses
- Consumer welfare has not changed

Example (cont)

Suppose we want to use the Logit model to analyze consumer welfare generated by the introduction of the blue bus

$$u_{ijt} = \xi_{jt} + \varepsilon_{ijt}$$

$t = 0$			$t = 1$			
	observed		predicted		observed	
option	share	ξ_{j0}	share	ξ_{j1}	share	ξ_{j1}
car	0.5					
red bus	0.5					
blue bus	–					
welfare						

Example (cont)

$$u_{ijt} = \xi_{jt} + \varepsilon_{ijt}$$

$t = 0$			$t = 1$			
	observed		predicted		observed	
option	share	ξ_{j0}	share	ξ_{j1}	share	ξ_{j1}
car	0.5	0				
red bus	0.5	0				
blue bus	–	–				
welfare	$\ln(2)$					

normalizing $\xi_{car0} = 0$, therefore $\xi_{bus0} = 0$

Example (cont)

$$u_{ijt} = \xi_{jt} + \varepsilon_{ijt}$$

$t=0$			$t=1$			
	observed		predicted		observed	
option	share	ξ_{j0}	share	ξ_{j1}	share	ξ_{j1}
car	0.5	0	0.33	0		
red bus	0.5	0	0.33	0		
blue bus	–	–	0.33	0		
welfare	$\ln(2)$		$\ln(3)$			

If nothing changed, one might be tempted to hold ξ_{jt} fixed.
This is the usual result: with predicted shares Logit gives gains

Example (cont)

$$u_{ijt} = \xi_{jt} + \varepsilon_{ijt}$$

$t = 0$			$t = 1$			
	observed		predicted		observed	
option	share	ξ_{j0}	share	ξ_{j1}	share	ξ_{j1}
car	0.5	0	0.33	0	0.5	
red bus	0.5	0	0.33	0	0.25	
blue bus	–	–	0.33	0	0.25	
welfare	$\ln(2)$		$\ln(3)$			

Suppose we observed actual shares

Example (cont)

$$u_{ijt} = \xi_{jt} + \varepsilon_{ijt}$$

$t = 0$			$t = 1$			
	observed		predicted		observed	
option	share	ξ_{j0}	share	ξ_{j1}	share	ξ_{j1}
car	0.5	0	0.33	0	0.5	0
red bus	0.5	0	0.33	0	0.25	$\ln(0.5)$
blue bus	–	–	0.33	0	0.25	$\ln(0.5)$
welfare	$\ln(2)$		$\ln(3)$		$\ln(2)$	

logit fails to predict, doesn't understand there's correlation

To rationalize observed shares we need to let ξ_{jt} vary
What exactly did we mean when we introduced blue bus?

Generalizing from the example

- In the example, the Logit model fails in the first step
- Holds more generally,
    - with Logit, expected utility is $\ln(1/s_{0t})$
    - since s_{0t} did not change in the observed data the Logit model predicted no welfare gain
    - Monte Carlo results in Berry and Pakes (2007) give similar answer
        - find that pure characteristics model matters for the estimated elasticities (and mean utilities) but not the welfare numbers
        - conclude: "the fact that the contraction fits the shares exactly means that the extra gain from the logit errors is offset by lower δ's, and this roughly counteracts the problems generated for welfare measurement by the model with tastes for products."

Generalizing from the example

- With more heterogeneity. Logit will get second step wrong
  - difference with RC

$$\ln\left(\frac{1}{s_{0,t}}\right) - \ln\left(\frac{1}{s_{0,t-1}}\right) = \ln\left(\frac{s_{0,t-1}}{s_{0,t}}\right) = \ln\left(\frac{\int s_{i,0,t-1} dP_\tau(\tau)}{\int s_{i,0,t} dP_\tau(\tau)}\right)$$

and

$$\int \left[\ln\left(\frac{1}{s_{i,0,t}}\right) - \ln\left(\frac{1}{s_{i,0,t-1}}\right)\right] dP_\tau(\tau) = \int \ln\left(\frac{s_{i,0,t-1}}{s_{i,0,t}}\right) dP_\tau(\tau)$$

  - the difference depends on the change in the heterogeneity in the probability of choosing the outside option, $s_{i,0,t}$
  - difference can be positive or negative

Final comments

- The key in the above example is that ξ_{jt} was allowed to change to fit the data.
- This works when we see data pre and post (allows us to tell how we should change ξ_{jt})
- What if we do not not have data for the counterfactual?
  - have a model of how ξ_{jt} is determined
  - make an assumption about how ξ_{jt} changes
  - bound the effects
- Nevo (ReStat, 2003) uses the latter approach to compute price indexes based on estimated demand systems