

COMPLEMENTARITY ANALYSIS IN MULTINOMIAL MODELS: THE GENTZKOW COMMAND

Yunrong Li & Ricardo Mora

SWUFE & UC3M

Madrid, Oct 2017

Outline

- 1 Getzkow (2007)
- 2 Case Study: social vs. internet interactions
- 3 The `gentzkow` command in Stata
- 4 Results

Getzkow (2004): Valuing new goods in a model with complementarity: Online newspapers, *The American Economic Review*, 2007

Online vs. printed news

- Do online newspapers crowd out consumption of printed newspapers?
 - Print and online papers are substitutes.
 - The welfare benefits of the online newspaper outweigh its costs. The newspaper firm (the Washington Post) gains with advertising.
 - The potential firm's gain of charging a positive price for online content disappears with advertising
- Methodologically:
 - Many important economic questions hinge on the extent to which new goods either crowd out or complement consumption of existing products.
 - Standard methods for studying new goods rule out complementarity by assumption, so their applicability to these questions has been limited.
 - Getzkow develops a new model that relaxes this restriction.

Gentzkow's Multinomial Model

$$u_0 = 0$$

$$u_A = \beta_A^0 + \beta_A x - \alpha_A p_A + v_A$$

$$u_B = \beta_B^0 + \beta_B x - \alpha_B p_B + v_B$$

$$u_{AB} = u_A + u_B + \Gamma$$

Definition

Alternatives A and B are complements (substitutes) if and only if $\text{sign} \left(\frac{\partial Q^A}{\partial p_B} \right) < (>) 0$

Theorem

Alternatives A and B are complements (substitutes) if and only if $\Gamma > (<) 0$

Remarks

- Getzkow's pattern of complementarity depends on parameter Γ : it describes by how much utility increases when both alternatives are chosen together
 - If the price of one alternative decreases, then the desirability of the other alternative indirectly increases through Γ
- The complementarity pattern is unrelated to the IIA assumption
- The model imposes a rank condition on the variance-covariance matrix in a typical multinomial model that can be tested using a LM test: $v_{AB} = v_A + v_B$
- If this condition is not satisfied, then the sign of Γ does not necessarily reflect the complementarity pattern

Gentzkow's analysis in the unrestricted multinomial model

$$u_0 = 0$$

$$u_A = \beta_A^0 + \beta_A x - \alpha_A p_A + v_A$$

$$u_B = \beta_B^0 + \beta_B x - \alpha_B p_B + v_B$$

$$u_{AB} = \beta_{AB}^0 + \beta_{AB} x + -\alpha_A p_A - \alpha_B p_B + v_{AB}$$

- Complementarity is individual specific: it depends on x and on $(v_{AB} - v_A - v_B)$. Take

$$\delta(x) = (\beta_{AB}^0 - \beta_A^0 - \beta_B^0) + (\beta_{AB} - \beta_A - \beta_B)x + (v_{AB} - v_A - v_B).$$

- No link between $\delta(x)$ and $E\left(\frac{\partial Q^A}{\partial p_B}\right)$
- We need to numerically estimate $E\left(\frac{\partial Q^A}{\partial p_B}\right)$
- We can condition on x : $E\left(\frac{\partial Q^A}{\partial p_B} | x\right)$

Case Study: social vs. internet interactions

The data

- We use data from the 2010 national baseline survey of the China Family Panel Studies (CFPS)
- The sample is representative of 95% of the Chinese population in 2010
- Sample: children who use internet between 10 and 15 years old (885 observations)
- Variables of interest:
 - *socialout*: the child does at least one of the following activities with friends at least once a month: going to karaoke, going to disco, dining out, and playing at a internet bar.
 - *socialnet*: one if child says doing social interactions on internet is important for him/her.
- Prices:
 - for real social activities: pocket money per month,
 - for internet social activities: whether parents use a popular online social interaction application (QQ)

Our goals

- In our application:
 - To study whether internet social interactions crowd out real social interactions
 - We estimate the general multinomial model and compute cross-price elasticities
 - In particular, we want to see the role of parents income and child age
- We need to develop a command for Stata

The `gentzkow` command in Stata

What the `ado` file does

- First, it estimates a multinomial model
- Then, it estimates the partial derivative of the demand of one of the alternatives when a price of the other alternative has an infinitesimal change
 - it is an integral over the utility shocks space: we implement it using a numerical Monte Carlo approximation
 - this object is independent of the alternative **AND** the price chosen
 - but interpretation is not clear when no price is included in the specification
- Partial derivatives can be computed for different values of controls, x
- Standard errors are computed using the bootstrap
- In the current implementation, `mlogit`, `mprobit`, and `gsem` are available to estimate the multinomial model

Inputs

- Required:
 - list of variables: at the very least the two binary choices, possibly expanded with controls x and prices p
- Optional:
 - number of replications in numerical integration
 - number of bootstrapped samples
 - seed to control random number generation
 - expected obs. per cell under uniform assumption: this is the parameter we use to control the accuracy of the Monte Carlo computation of the integral
 - number of quadrature points (only with `mprobit` and `gsem`)
 - control variable whose value changes, and different evaluation values
 - model: `mlogit`, `mprobit`, `gsem`, and all options related to these commands, including constraints
 - whether multinomial model estimates and elasticities are displayed

Syntax

```

syntax varlist(min=2) [if] [in] [fw pw iw], ///
  [
    ///
    Replications(integer 500) /// // # replications in integration
    Bootstraps(integer 100) /// // # bootstrapped samples
    SEED(real 1966) /// // seed of random number generation
    Ncells(integer 15) /// // expected obs. per cell
    INTPoints(integer 15) /// // # of quadrature points
    Conditioning(varname) /// // evaluation variable
    VALues(numlist) /// // values for evaluation
    ATVALues(name) /// // matrix name of control values for evaluation
    Model(string) /// // mlogit, mprobit, gsem (default is mlogit)
    NOCONSTant /// // suppress constant term
    CONSTraints(numlist) /// // list of linear constraints
    Level(real 95) /// // set confidence level; default is level(95)
    Detail /// // displays multinomial model estimates
    CROSS /// // only shows cross elasticities
    MAXimize(string) /// // string containing maximize_options
    TWOlevel(varname) /// // variable for level random effects (gsem)
  * ]

```

Example 1 (I)

```
gentzkow socialout socialnet `prices' `lista_x0' Dp*, ///  
    constraint(1/`i') ///  
    r(5000) b(200) n(50) ///  
    max("iterate(50)") ///  
    cross
```

- constrained multinomial logit
- bootstrap standard errors using 200 simulations
- numerical integration using 5000 simulations, with around 50 observations within each cell to compute frequencies
- only displays cross-price elasticity

Example 1 (II)

```
(running mlogit on estimation sample)
```

```
Bootstrap replications (200)
```

```

-----|----- 1 -----|----- 2 -----|----- 3 -----|----- 4 -----|----- 5
..... 50
..... 100
..... 150
..... 200

```

```
Warning: only 191 converged bootstraps
```

```
Cross-price elasticity
```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
_cons	.0337866	.0115226	2.93	0.003	.0112027	.0563706

Example 2 (I)

```

gentszkow socialout socialnet `prices' `lista_x0' Dp*, ///
          constraint(1/`i') ///
          r(5000) b(200) n(50) ///
          max("iterate(50)") ///
          c(age) val(10(2)14) ///
          cross
test _b[r1]-_b[r3]=0

```

- computes expectation conditional on age: 10,12,14
- only displays cross-elasticities
- tests whether the expectation is equal at age 10 than at age 14

Example 2 (II)

```
(running mlogit on estimation sample)
```

```
Bootstrap replications (200)
```

```

-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
      1     2     3     4     5
..... 50
..... 100
..... 150
..... 200

```

```
Cross-price elasticity
```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
r1	.0049281	.0062625	0.79	0.431	-.007346 .0172023
r2	.0234086	.0091993	2.54	0.011	.0053784 .0414389
r3	.0366767	.0132364	2.77	0.006	.0107338 .0626196

```
( 1)  r1 - r3 = 0
```

```

      chi2( 1) =      7.78
      Prob > chi2 =      0.0053

```

Results

	<i>Cross elast.</i>	<i>Std.Err.</i>	<i>Own price: Real life</i>	<i>Std.Err.</i>	<i>Own price: Internet</i>	<i>Std.Err.</i>
Model,Elasticities	1	2	3	4	5	6
Unconditional	0.008	(0.008)	0.114 ***	(0.009)	0.147 ***	(0.010)
Uncond., Province	0.033 **	(0.014)	0.128 ***	(0.022)	0.180 ***	(0.018)
Conditional	0.017	(0.013)	0.111 ***	(0.015)	0.166 ***	(0.017)
Cond., Province	0.034 ***	(0.012)	0.126 ***	(0.022)	0.183 ***	(0.017)
Cond., Major Prov.	0.033 ***	(0.011)	0.120 ***	(0.021)	0.186 ***	(0.015)

Note: Estimated cross-price elasticities and own-price elasticities of our unrestricted multinomial model using the multinomial logit model option are reported in the table. "Cross elast." refers to the change in the child's probability of participating in real life social interactions with respect to an infinitesimal change in a factor that positively affects the child's net utility of doing internet social interactions but does not affect the child's net utility of doing real life social interactions. Equivalently, "Cross elast." also refers to the change in the child's probability of participating in internet social interactions with respect to an infinitesimal change in a factor that positively affects the child's net utility of doing real life social interactions but does not affect the child's net utility of doing internet social interactions. "Std.Err." refers to the standard error. "Own price: Real life" refers to the change in the child's probability of participating in real life social interactions with respect to an infinitesimal change in a factor that positively affects the child's net utility of doing real life social interactions but does not affect the child's net utility of doing internet social interactions. "Own price: Internet" refers to the change in the child's probability of participating in internet social interactions with respect to an infinitesimal change in a factor that positively affects the child's net utility of doing internet social interactions but does not affect the child's net utility of doing real life social interactions.

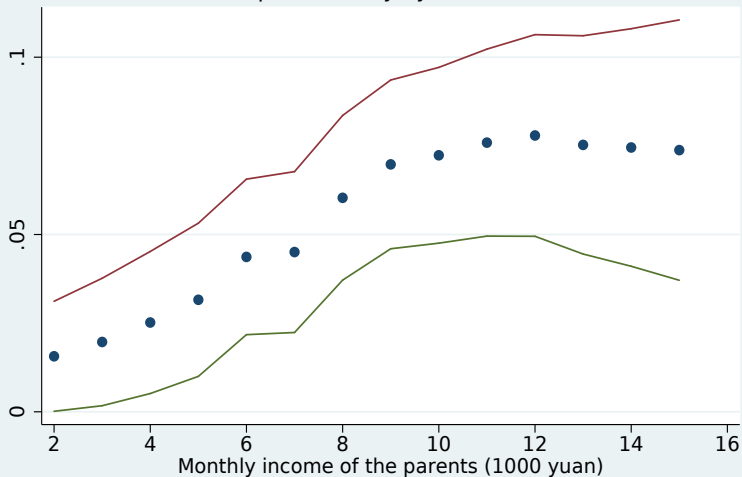
Note: "Prices" affect utility positively.

	<i>P-value of test</i>	<i>Value</i>	<i>Cross elast.</i>	<i>Std.Err.</i>	<i>Value</i>	<i>Cross elast.</i>	<i>Std.Err.</i>
A. Price variables	1	2	3	4	5	6	7
Pocket money of the child in log	0.067	0	0.018 ***	(0.007)	6	0.044 ***	(0.018)
Both parents use QQ every-day	0.112	0	0.030 ***	(0.011)	1	0.019 *	(0.011)
B. Child characteristics							
Gender	0.583	0	0.020 **	(0.010)	1	0.045 ***	(0.014)
Age in years	0.014	10	0.008	(0.006)	15	0.045 ***	(0.016)
C. Household characteristics							
Urban	0.797	0	0.020 **	(0.010)	1	0.041 ***	(0.014)

Note: Estimated cross-price elasticities of our unrestricted multinomial model using the multinomial logit model option are reported in the table. The multinomial logit model is estimated with price variables, all the other controls, and major province dummies. Cross-price elasticities are estimated for a series of values of each of the controls. "P-value of test" refers to the p-value of the test on the joint equality of all the cross-price elasticities based on different values of the controls. "Value" refers to different values of the control variable. "Cross elast." refers to the change in the child's probability of participating in real life social interactions with respect to an infinitesimal change in a factor that positively affects the child's net utility of doing online social interactions but does not affect the child's net utility of doing real life social interactions. Equivalently, "Partial derivative" also refers to the change in the child's probability of participating in online social interactions with respect to an infinitesimal change in a factor that positively affects the child's net utility of doing real life social interactions but does not affect the child's net utility of doing online social interactions. "Std.Err." refers to the standard error.

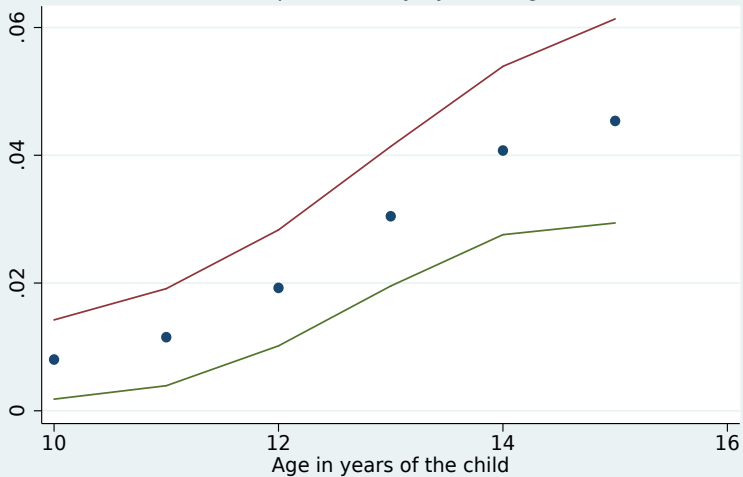
Internet and traditional social interactions

Complementarity by income level



Internet and traditional social interactions

Complementarity by child age



Extensions

- coefficient's display should be improved
- only cross-elasticity results posted as results, all results stored in matrix
- panel data implementation not checked out: Monte Carlo using gsem
- testing the Gentzkow model using LM test

Thank you