Identification of Simultaneous Equations System

Shreya Bhattacharya

Sargam Jain

Saish Nevrekar

Deepak Kumar

Econometric-II
Computer Project
19th April, 2013
Indira Gandhi Institute of Development Research

Objective

The objective of the study is to test whether the simultaneous equations in the model for sales and advertising relationships between filter and non-filter cigarette brands is identifiable or not.

Introduction to the Model

- The model tests for the success of the participating tobacco companies in targeting the youth in the US.
- Expenditures on cigarette advertisements in national newspapers, magazines and promotional billboards in USA were studied over a period of 10 years.
- The model focuses on those modes of advertisement which are closely targeted by the population above the age of 20 years.

Motivation

- Advertising is an important method of competition in industries that are highly concentrated, such as the cigarette industry(competition through increased sales with advertising and other marketing techniques)
- An important economic aspect of advertising is diminishing marginal product. It suggests that after some point, additions to advertising will result in ever smaller additions to consumption.
- Many of the recall studies are problematic (assumption: the direction of influence between advertising and smoking uptake and consumption is one-way).

Motivation

- Self-reports are often plagued by social desirability bias i.e. subjects often report what they believe the interviewer wishes to hear or what is the socially preferred option, rather than what is actually true.
- Many tobacco advertising exposure and recall studies suffer from a variety of methodological issues. These include misspecified variances, omitted interactions and paths, endogeneity, sample attrition and selection bias.

Simultaneous Equations System

- A system with two or more equations, where a variable explained in one equation appears as an explanatory variable in another.
- The structural form is as follows:
  $\beta' Y_t + \Gamma' X_t = \varepsilon_t$
- Also known as Behavioural Equations as they portray the structure of an economy or the behaviour of an economic agent.
- The **β'** and **Γ'** are known as structural parameters or coefficients.

Simultaneous Equations System

- Reduced Form is as follows:

  $\mathbf{Y_t = \pi' X_t + u_t}$
- It expresses an endogenous variable solely in terms of the predetermined variables and the stochastic disturbances.
- The reduced form coefficients are also known as **Impact or Short Run Multipliers.**

Question of Identification

- By identification we mean whether numerical estimates of the parameters of a structural equation can be obtained from the estimated reduced form coefficients.
- If this can be done, we say that the particular equation is identified else the equation is unidentified.
- A structural equation is said to be identified if and only if all its parameters are identified.
- A simultaneous equations model (system) is said to be identified if and only if all the structural equations in the system are identified.

Question of Identification

- The parameters of a structural equation are said to be identified if and only if they can be obtained in a unique way from the reduced form.
- There are two ways of identifying the equations:
- By a study of the linear combinations of the equations of the system, that is, from the structural form.
- By a study of the relation between π, β and Γ.

Conditions for Identification

- **RANK CONDITION:**

To derive a unique solution for β* from the reduced form model:

$r(\pi_{\Delta\Delta}) = r[\pi_{\Delta\Delta^*} \quad \pi_{\Delta\Delta 1}] = n^*$.

This condition is both a necessary as well as a sufficient condition.

- **ORDER CONDITION:**

The order condition is only a necessary condition: $m_{\Delta\Delta} \geq n^*$

The Model

- It consists of two demand equations for two competing groups of cigarette brands and two equations that describe the advertising relations of these groups of brands.
- Sales of the major filter cigarette brands have been aggregated to give one demand equation. Similarly, there is one demand equation for the major non-filter brands.

The Model- Variables used

- **Log ($Salesf_t$):** Logarithm of sales for filter cigarettes divided by population over age 20.
- **Log ($Salesnf_t$):** Logarithm of sales for non-filter cigarettes divided by population over age 20.
- **Log (Adf_t):** Logarithm of advertising dollar sales for filter cigarettes divided by population over age 20.
- **Log ($Adnf_t$):** Logarithm of advertising dollar sales for non-filter cigarettes divided by population over age 20.
- **Log (PDI_t):** Logarithm of disposable personal income divided by population over age 20 divided by consumer price index.
- **Log($Price_t$):** Logarithm of price per package of non-filter divided by consumer price.

The Model- Structural Form

- $-Log(Salesf_t)+0.Log\ (Salesnf_t)+\beta_1 Log(Adf_t)+\beta_2 Log(Adnf_t)+\gamma_1 Log(PDI_t)+\gamma_2 Log(Price_t)+\varepsilon_{1t}=0$
- $0.Log\ (Salesf_t)-Log(Salesnf_t)+\beta_3 Log(Adf_t)+\beta_4 Log(Adnf_t)+\gamma_4 Log(PDI_t)+\gamma_5 Log(Price_t)+\varepsilon_{2t}=0$
- $-Log\ (Salesf_{t)}+\beta_5 Log\ (Salesnf_t)+\beta_6 Log(Adf_t)+0.Log\ (Adnf_t)+0.\ Log\ (PDI_t)+0.\ Log\ (Price_t)+\varepsilon_{3t}=0$
- $\beta_7 Log\ (Salesf_{t)}-Log\ (Salesnf_t)+0.Log\ (Adf_t)+\beta_8 Log\ (Adnf_t)+0.\ Log\ (PDI_t)+0.\ Log\ (Price_t)+\varepsilon_{4t}=0$

The Model-Reduced Form

$$Log\ (Salesf_t) = \alpha_1\ Log(PDI_t) + \alpha_2\ Log(Price_t) + \varepsilon_{1t}$$

$$Log\ (Salesnf_t) = \alpha_4 Log(PDI_t) + \alpha_5 Log(Price_t) + \varepsilon_{2t}$$

$$Log\ (Adf_t) = \alpha_7\ Log\ (PDI_t) + \alpha_8\ Log(Price_t) + \varepsilon_{3t}$$

$$Log\ (Adnf_t) = \alpha_{10}\ Log\ (PDI_t) + \alpha_{11}\ Log(Price_t) + \varepsilon_{4t}$$

Identification using SAS

Commands Used in SAS:

```
proc syslin data= sasuser.data1 3sls;
endogenous salesf salesnf adf adnf;
instruments pdi price;
model salesf=adf adnf pdi price;
model salesnf=adf adnf pdi price;
model adf=salesf salesnf;
model adnf=salesf salesnf;
run;
```

Identification using SAS

The regression run in SAS displayed the following warning along with the estimation results:

"The model is not of full rank. Least Squares solutions for the parameters are not unique. Certain statistics will be misleading. A reported degree of freedom of 0 or B means the estimate is biased."

The variables in the unidentified equations can be expressed as linear combinations of other variables.

Identification using STATA

- The **checkreg3** command allows you to verify whether the estimation is meaningful by checking that the rank condition is satisfied for each of the N equations in the system.
- Unless the rank condition is satisfied for each equation in the system, the system is unidentified. Although unusual for a system to satisfy the order condition without satisfying the rank condition, it can occur.

Identification in STATA

```
. checkreg3 ( logsalesf  logadf logadnf logpdi logprice) ( logsalesnf  logadf lo
> gadnf logpdi logprice) ( logadf logsalesf logsalesnf) ( logadnf logsalesf logs
> alesnf)

Endogenous coefficients matrix

             logsalesf  logsalesnf      logadf     logadnf
 logsalesf          -1
logsalesnf           0          -1
    logadf          .5          .5          -1
   logadnf          .5          .5           0          -1

Exogenous coefficients matrix

              logpdi  logprice
 logsalesf        .5        .5
logsalesnf        .5        .5
    logadf         0         0
   logadnf         0         0


Eq 1 fails rank condition for identification
Eq 2 fails rank condition for identification
Eq 3 fails rank condition for identification
Eq 4 fails rank condition for identification

Rank deficiency: System is not identified
```

Conclusion

- From the above studied model on sales and advertising expenditure of cigarettes, it can be seen that each of the equations violate the Rank Condition necessary for identification as some of the parameters in each of the structural equations do not have unique solutions.
- In the simultaneous model specified above the following structural parameters remain unidentified:

$\boldsymbol{\beta_{1(adf)}}, \boldsymbol{\beta_{2(adnf)}}, \boldsymbol{\beta_{3(adf)}}, \boldsymbol{\beta_{4(adnf)}}, \boldsymbol{\gamma_{1(pdi)}}, \boldsymbol{\gamma_{2(price)}}, \boldsymbol{\gamma_{4(pdi)}}, \boldsymbol{\gamma_{5(prices)}}$

- Since, the model is not of full rank, the Least Squares solutions for the parameters are not unique. The estimates obtained will be misleading. There is possibility of biased estimates.

Bibliography

- Gujarati D. Basic Econometrics (4ed., MGH, 2004);
- Wooldridge J.M. Introductory econometrics (South-Western College Pub., 2003);
- Lecture Notes- Subrata Sarkar, IGIDR, Mumbai;
- Advertising Publications, Inc., "Costs of Cigarette Advertising: 1952-1959," Advertising Age, 31 (September 19, 1960), 126-127;
- "Costs of Cigarette Advertising: 1957-1965," Advertising Age, 37 (July 25, 1966), 56-58;
- U.S. Census Bureau, Statistical Abstract of United States, Washington, D.C.: Government Printing Office, 1965, 327; 360-1;
- http://data.worldbank.org/