Endogeneity in Empirical Corporate Finance

Endogeneity in Empirical Corporate Finance

- Endogenous problem is an important issue in the fields of corporate governance and corporate finance.

Roberts, Michael R. and Whited, Toni
M., Endogeneity in Empirical Corporate Finance (October 5, 2012). Simon School Working Paper No. FR 11-29.

Introduction to Endogeneity

What is "endogeneity"?

- Loosely define as a correlation between the explanatory variables and the error in a regression.

Problem of "endogeneity"

- Leads to biased and inconsistent parameter estimate that make *reliable inference* virtually **impossible**

Introduction to Endogeneity

The first step to address endogeneity

- Which variables are endogenous
- Why they are endogenous

Sources of endogeneity

- Omitted variables
- Simultaneity
- Measurement error

Ways to address endogeneity problem

Omitted Variables

Omitted variables: variables that should be included in the (vector of) explanatory variables, but for various reasons are not.

For example, executive compensation may depend on executives’ ability, which are difficult to quantify and observe.

Omitted Variables

Assume the true economic relation is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \gamma\omega + u$$

- ω is an unobservable explanatory variables and γ is its coefficient.

The estimable regression is

- $v = \gamma\omega + u$ is the composite error term.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + v$$

Omitted Variables

- If the omitted variable ω is correlated with any of the explanatory variables ($x_1, \dots, x_k$), the composite error term *v* is correlated with the explanatory variables. Thus all elements of β are **inconsistent estimates**.

$$\text{plim}\widehat{\beta}_j = \beta_j + \gamma \frac{cov(x_j, \omega)}{var(X_j)}, j = 1, \dots, k$$

Corporate finance assume that all of the other explanatory variables are partially *uncorrelated* with the omitted variable

Omitted Variables

Economic meanings of estimator

$$\text{plim}\widehat{\beta}_j = \beta_j + \gamma\frac{cov(x_j,\omega)}{var(X_j)}, j = 1, \ldots, k$$

- If $\omega$and $X_j$are *uncorrelated* ($cov(x_j,\omega)$ = 0), then OLS estimator is **consistent**.
- If $\omega$and $X_j$are *correlated*($cov(x_j,\omega)$>0), then OLS estimator is **inconsistent**.
- If γ and $\frac{cov(x_j,\omega)}{var(x_j)}$ have the same (different) sign, the bias is positive (negative).

Omitted Variables: An Example

Firm size is a determinant in CEO compensation (Core, Guay, and Larcker, 2008)

Large firms are more difficult to manage and then require more skilled manager (Gabaix and Landier, 2008).

Firm size is *endogenous* because managerial ability, which is *unobservable* (is in the error term) and is *correlated* with firm size.

Simultaneity

Simultaneity bias occurs when *y* and one or more of the *x*'s are determined in equilibrium

$$x_k \rightarrow y \quad \text{or} \quad y \rightarrow x_k$$

For example, when examining the relation between *antitakeover provision* and *firm value*

- Antitakeover provisions leads to a loss in firm value
- Managers of low-value firm adopt antitakeover provisions in order to entrench themselves

Simultaneity

Assume *y* and *x* are determined jointly as follow:

$$y = \beta x + v$$
$$x = \ \alpha y + v$$

- where *u* is uncorrelated with *v*.

Thus $\hat{\beta} = \frac{cov(x,y)}{var(x)} = \frac{cov(x,\beta x+v\ )}{var(x)} = \beta + \frac{cov(x,u)}{var(x)}$

- where $\frac{cov(x,u)}{var(x)} = \frac{\alpha(1-\alpha\beta)var(u)}{\alpha^2 var(u)+var(v)}$

Measurement Error

Most empirical studies in corporate finance use proxies for unobservable or difficult to quantify variables. Any ***difference*** between the true variable of interest and the proxy leads to **measurement error**.

- Record variables incorrectly
- Conceptual differences between proxies and their unobservable counterparts

Measurement error leads to regression error

Measurement Error

Measurement error in the dependent variable

Consider a main variable of interest in capital structure theories should be "market leverage ratio" (Fischer, Heinkel, and Zechner, 1989)

- Ratio of the market value of debt to the market value of firm (debt plus equity)
- Market value of debt is difficult to measure because most debt is privately held by banks and other financial institutions; and most of public debt is infrequently traded.

Measurement Error

- Often use book debt value as a proxy; creates a wedge between the empirical measure and the true economic measure.

Total compensation (including stock options) for executives can also be difficult to measure

- Stock options often vest over time and are valued by an approximation, e.g. Black-Scholes (Core, Guay, and Larcker, 2008)

Measurement Error

Consider the following model

$$y^* = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + u$$

- where y^* is an unobservable measure and y is the observable version of or proxy for y^*. The difference between the two is $w = y - y^*$. The estimable model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + v$$

- where $v = w + u$ is the composite error

OLS estimates **might be inconsistent**, which are similar to those of an omitted variable

Measurement Error

Measurement error in the independent variables

Assume the estimation model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k^* + u$$

- where x_k^* is an unobservable measure and x_k is it observable proxy. Define measure error to be w = $x_k - x_k^*$. The estimable model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + v$$

- where $v = u - \beta_k x_k$ is the composite error term.

Measurement Error

- Since ω is uncorrelated with each x_j, OLS will produce **consistent estimates**.
- When the measurement error appears in the independent variables, **it will affect the variance of the error term**, which changes from $var(u) = \sigma_u^2$ to $var(u - \beta_k\omega) = \sigma_u^2 + \beta_k^2\sigma_\omega^2 - 2\beta_k\sigma_{u\omega}$.

Ways to Address Endogeneity Problem

Rely on a clear source of **exogeneous** variation for identifying the co-effects of interest

- Instrumental variables
- Difference-in-difference estimator
- Regression discontinuity design*

Rely more heavily on **modeling** assumption

- Panel data methods, e.g., fixed and random effects
- Matching methods
- Measurement error methods.

Instrumental Variables

Basic econometric framework

Recall the single equation linear model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + u$$

- Suppose one regressor, x_k is correlated with u, then only the estimate of β_k is biased.

An instrument, *z*, is a variable that satisfies two conditions that we refer to as the *relevance* and *exclusion* conditions

Instrumental Variables

*Relevance*condition

- Requires that the partial correlation between the instrument and the endogenous variable not be zero; It requires the coefficient γ in the regression does not equal zero.

*Exclusion*condition

- $cov(z, u) = 0$, it implies that the only role that the instrument *z* plays in influencing the outcome y is through its affect on the endogenous variable x_k

$$x_k = \alpha_0 + \alpha_1 x_1 + \cdots + \alpha_{k-1} x_{k-1} + \gamma z + v$$

Instrumental Variables

There is noting restricting the number of *instruments* to just one

- From an asymptotic efficiency perspective, more instruments is better.
- From a finite sample perspective, more instruments is not necessarily better and can even exacerbate the bias inherent in **2SLS**

There is noting restricting the number of *endogenous variable* to just one.

Instrumental Variables

Two-stage least squares (2SLS)

Estimate the predicted variable, $\hat{x}_k$, by regressing the endogenous variable x_k on all of the exogenous variables

- Controls $x_1, \ldots, x_{k-1}$ and instruments $z_1, \ldots, z_m$

Replace the endogenous variable x_k with its **predicted values** from the first stage $\hat{x}_k$, and regress the outcome variable y on all of the control variables and $\hat{x}_k$.

Instrumental Variables

This two-step procedure can be done all at once; but not separately estimate the first and second stages.

- Generate regressor, $\hat{x}_k$ in the second stage itself will induce estimation error. This estimation error must be taken into account when computing the standard error of its, and the other explanatory variables' coefficients.

Instrumental Variables

Where do valid instruments come from?

The only way to find a good instruments is to understand the economic of the question.

Institutional changes (or policy changes)

Instrumental Variables

Bennedsen et al. (2007): study CEO succession in family firms; whether replacing an outgoing CEO with a family member hurts firm performance?

- Dependent variable: firm performance
- Explanatory variable: family CEO succession

Sources of endogeneity

- The characteristics of the firm and family that cause it to choose a family CEO may also cause the change in performance **(omitted variable)**

Instrumental Variables

Bennedsen et al. (2007) choose an IV approach to isolate *exogenous variation* in the CEO succession decisions.

The instrument, *z*, Bennedsen et al. (2007) choose is **the gender of the first-born child of a departing CEO**.

- Do not affect firm performance
- Might be correlated with family CEO succession

Instrumental Variables

Results of Bennedsen et al. (2007)

- CEO with boy-first families are significantly more likely to appoint a family CEO in the first stage regression (*relevance condition* is satisfied)
- In the second stage regression, the IV estimates of the *negative effect* of in-family CEO succession are much larger than the OLS estimates.

Lagged Instruments

The use of lagged dependent variables and lagged endogenous variables has become widespread in corporate finance.

- Use lagged values of the dependent variable

Difference-in-Differences

Difference-in-Differences (DD) estimators are used to recover the treatment effects stemming from sharp changes in the economic environment, government policy, or institutional environment.

- Single cross-sectional differences after treatment
- Single time-series differences after treatment
- Double difference estimator: Difference in Differences

Difference-in-Differences

Single cross-sectional differences after treatment

- Compare the **post-treatment outcomes** of the treatment and control groups.

Example:

- Estimate the effect of *state antitakeover laws* on leverage by examine one year of data after the law passage (Garvey and Hanka, 1999).

Difference-in-Differences

- Compare the leverage ratios of firms in states that passed the law (**treatment group**) and did not pass the law (**control group**):

  $$y = \beta_0 + \beta_1 d + u$$

  where *y* is leverage, and *d* is the treatment assignment indicator equal to one if the firm is incorporated in a state that passed the antitakeover law and zero otherwise.
- The ***difference* between treatment and control group** averages is β_1

Difference-in-Differences

Single time-series differences before and after treatment

- Compare the outcome ***after*** the onset of the treatment with the outcome ***before*** the onset of treatment for just those subjects that are treaded.

Example

- Examine the impact of *deregulation* of the French banking industry on the behavior of the borrowers and banks (Bertrand, Scholar, and Thesmar, 2007).

Difference-in-Differences

- One *before*and one *after*treatment: comparison can be accomplished with a two period panel regression using only subject affected by the event

$$y = \beta_0 + \beta_1 p + u$$

  where *p* equals one if the observation is made after treatment onset (i.e. post-treatment) and zero otherwise (i.e. pre-treatment)

- The ***treatment effect*** is β_1.

Difference-in-Differences

Double difference estimator: difference-in differences (DD)

- Combine cross-sectional/time-series estimators to take advantage of both estimators' strengths
- Consider firm-year panel dataset in which there are two time periods, one before and one after the onset of treatment, and only some of the subjects are treated.

Difference-in-Differences

For example, Arizona passed antitakeover legislation in 1987 at which time Connecticut had not passed (Bertrand and Mullainathan, 2003).

- 1986: pre-treatment period
- 1987: post-treatment period
- Firms registered in Arizona: treatment group
- Firms registered in Connecticut: control group

Difference-in-Differences

$$y = \beta_0 + \beta_1 d \times p + \beta_2 d + \beta_3 p + u$$

- where *d* is the *treatment* assignment variable equal to one if a firm is registered in Arizona, zero if registered in Connecticut; *p* is the *post-treatment* indicator equal to one in 1987 and zero in 1986.
- β_1 represents the DD estimate.

$$E(y|d = 1, p = 1) = \beta_0 + \beta_1 + \beta_2 + \beta_3$$
$$E(y|d = 1, p = 0) = \beta_0 + \beta_2$$
$$E(y|d = 0, p = 1) = \beta_0 + \beta_3$$
$$E(y|d = 0, p = 0) = \beta_0$$

Difference-in-Differences

		1987	1986	
		Post-Treatment	Pre-Treatment	Difference
AZ:	Treatment	$\beta_0 + \beta_1 + \beta_2 + \beta_3$	$\beta_0 + \beta_2$	$\beta_1 + \beta_3$
CT:	Control	$\beta_0 + \beta_3$	β_0	β_3
	Difference	$\beta_1 + \beta_2$	β_2	β_1

$\beta_1 + \beta_3$: Difference between 1987 and 1986 in AZ

β_3: Difference between 1987 and 1986 in CT

$\beta_1 + \beta_2$: Difference between AZ and CT in 1987

β_2: Difference between AZ and CT in 1986

β_1: Difference-in-Differences (double difference estimator)

Panel Data Methods

Assume a panel regression model as follow:

$$y = \beta_0 + \beta_1 x_{it} + u_{it}$$

- The error term, u_{it} can be decomposed as $u_{it} = c_i + e_{it}$. c_i can be interpreted as capturing the aggregate effect of all of the unobservable, time-invariant explanatory variables for y_{it}
- If c_i and x_{it} are correlated, then c_i is referred to as a **"fixed effect."**
- If c_i and x_{it} are NOT correlated, then c_i is referred to as a **"random effect."**

Panel Data Methods

- In the former case (c_i and x_{it} are correlated), endogeneity is obviously a concern
- In the later, endogeneity is not a concern; however, the computation of standard error is affected.

Fixed effects

- Industry fixed effect
- Year fixed effect
- Firm fixed effect

Petersen's clustering-adjusted standard error

Ownership Structure around the World

LLS (1999): better (poor) investors protection is associated with dispersed (concentrated) ownership structure

The classification of countries based on the legal rules for protecting minority shareholder may be ***endogenous***.

- Countries with economically and politically powerful controlling shareholders may *enact laws* that entrench such shareholders and reduce minority rights.

Ownership Structure around the World

Way to solve ***endogenous***

- Use *the origin of the commercial laws* instead of the legal rules of investor protecting.
- Common law (英美法) **vs** Civil law (成文法)
- Both historically predetermined and highly correlated with shareholder protection
  - Historically predetermined means that both cannot be changed in the current (X → Y but not Y → X).
  - Common law is associated with good legal protection on minor shareholders, while Civil law is not.

The Fraction of Widely Held Firms around the World

	N	Sample of Large Firms: 20% Definition of Control	Sample of Large Firms: 10% Definition of Control	Sample of Medium Firms: 20% Definition of Control	Sample of Medium Firms: 10% Definition of Control
Panel A: Legal Origin					
Common law	9	0.4778	0.3833	0.4552	0.2000
Civil law	18	0.3083	0.1694	0.1278	0.0611
Diff (*t-statistic*)		1.4779	2.1900	3.9725	2.3526

- Countries with Common (Civil) law have higher (lower) fraction of widely held firms
- Results are robustness across different criteria of control (20%, 10%) and different sample selection (large, medium)

Family Ownership & Firm Value

An *alternative* explanation for the performance difference is that families in poorly performing firms (or foreseeing poor performance) are more likely to sell their shares and exit the firm.

- Use **instrumental variable** (IV) regressions (two-stage least squares) to examine the nature of causality between family ownership and firm performance.
- Results still hold.

Endogeneity Concern

Two-stage Least Squares

The first-stage: regress family ownership on the *natural log of total assets*, the *square of the natural log of total assets*, and *monthly stock return volatility.*

- The *predicted value of family firms* is defined as instruments.

The second-stage: regress performance measures on the predicted value of family firms, and other control variables.

	Return on Assets (Using EBITDA)	Return on Assets (Using Net Income)	Tobin's q
Intercept	– 0.041	– 0.108	– 0.013
	(0.66)	(1.81)	(0.12)
Predicted value of family firm	0.137	0.184	1.882
	(2.22)	(3.12)	(2.12)
Officer and directors ownership (less family)	0.205	0.290	4.200
	(1.28)	(1.52)	(1.67)
Unaffiliated blockholders	– 0.010	– 0.004	– 0.177
	(0.98)	(0.47)	(1.51)
Outside directors	0.129	0.186	1.722
	(1.81)	(2.70)	(1.78)
CEO equity-based pay	0.047	0.041	0.428
	(2.61)	(2.79)	(2.61)
Long-term assets/sales	0.152	– 0.018	– 0.532
	(5.88)	(0.81)	(1.79)
$(\text{Long-term assets/sales})^2$	– 0.010	0.010	0.115
	(0.63)	(0.88)	(0.82)
Operating income/sales			0.004
			(1.90)
R&D/long-term assets	0.032	– 0.003	0.564
	(2.01)	(0.14)	(3.10)
R&D dummy variable	– 0.024	– 0.023	– 0.021
	(2.44)	(2.39)	(0.15)
Advertising expenses/long-term assets	0.015	0.008	0.017
	(6.00)	(2.49)	(0.43)
Advertising expense dummy variable	0.009	0.004	0.027
	(0.99)	(0.48)	(0.26)
Capital expenditures/long-term assets	0.096	0.026	0.766
	(2.82)	(0.79)	(1.76)
Adjusted R square	0.401	0.199	0.353