INSTRUMENTAL VARIABLES IN ACTION

LUISS Guido Carli
LIBERA UNIVERSITÀ INTERNAZIONALE DEGLI STUDI SOCIALI

Two Examples

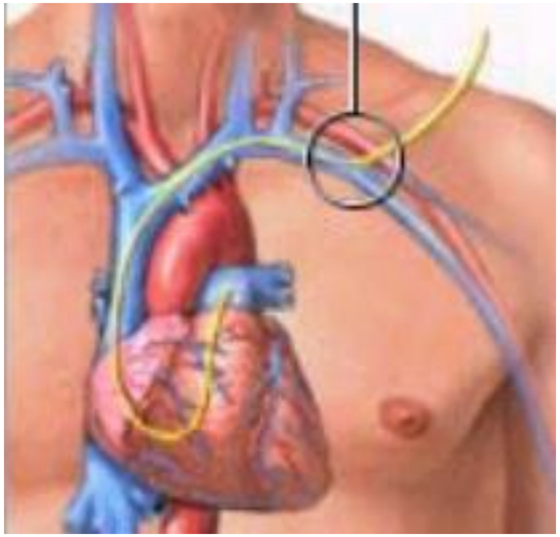
Swanz Ganz Cathether

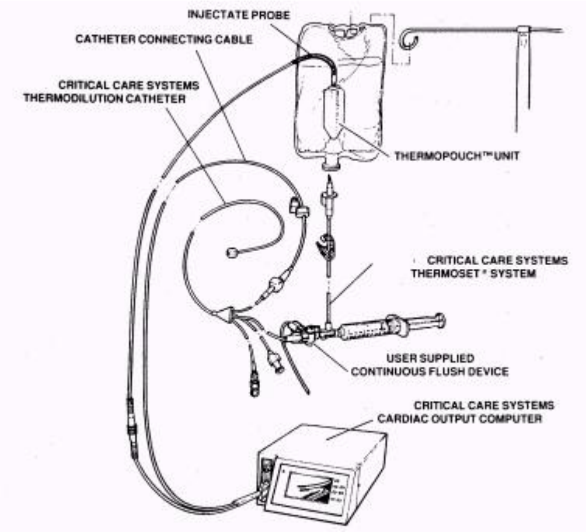
- Swan-Ganz catheter, which is a device used by intensive care unit (ICU) doctors to guide therapy

- A Swan-Ganz catheter is a long slender tube outfitted with sensors designed to measure hemodynamic pressures in the right side of the heart and in the pulmonary artery.

Swan Ganz Catheter

- Under sterile conditions, an ICU doctor will typically insert the catheter into the left subclavian vein (underneath the clavicle)





Swan-Ganz catheter

- A Swan-Ganz catheter is primarily used to measure pressure at different locations in the right side of the heart.
- These pressure measurements can provide important diagnostic information, such as whether the patient's heart valves are working, or whether the patient has pulmonary hypertension.
- Information gleaned from Swan-Ganz measurements is often used by ICU doctors to make decisions about treatment, such as whether to give the patient medications that affect the functioning of the heart.

- The procedure is done to evaluate how the blood moves (circulates) in people who have:
- Abnormal pressures in the heart arteries
- Burns
- Congenital heart disease
- Heart failure
- Kidney disease
- Leaky heart valves (valvular regurgitation)
- Shock

- Swan-Ganz catheterization can also be used to detect abnormal blood flow between two areas of the heart that are not normally connected.
- Conditions that can also be diagnosed or evaluated with Swan-Ganz catheterization include:
- Cardiac tamponade
- Pulmonary hypertension
- Restrictive cardiomyopathy

- It may also be done to monitor for complications of heart attack and to see how well certain heart medicines are working.

Swan-Ganz Catheter & Econometrics

- Connors et al. (1996), which examines the impact of Swan-Ganz catheterization on mortality outcomes among a population of patients admitted to the intensive care unit (ICU) at five prominent hospitals.
- Connors et al. (1996) reach the controversial conclusion that patients who receive Swan-Ganz catheterization during their first day in the ICU are **1.27 times more** likely to die within **180 days** of their admission.

- Before Connors et al. (1996), Gore et al. (1985) and Zion et al. (1990) also found found that catheterization increases mortality. Dalen (2001) criticized both studies because they did not control for clinically important differences between the patients who had catheters placed and those who did not.
- The Connors et al. (1996) study was conceived in part as a response to this criticism.

Dataset

The “outcome” variables are:

- surv7
- surv30
- surv60
- surv90
- surv120
- surv150
- surv180

$$surv7 = \begin{cases} 1 & \text{if patient is alive after 7 days} \\ 0 & otherwise \end{cases}$$

$$surv180 = \begin{cases} 1 & \text{if patient is alive after 180 days} \\ 0 & otherwise \end{cases}$$

Dataset

The main regressor is

- swang $= \begin{cases} 1 & \text{if patient received SG catheter} \\ 0 & otherwise \end{cases}$

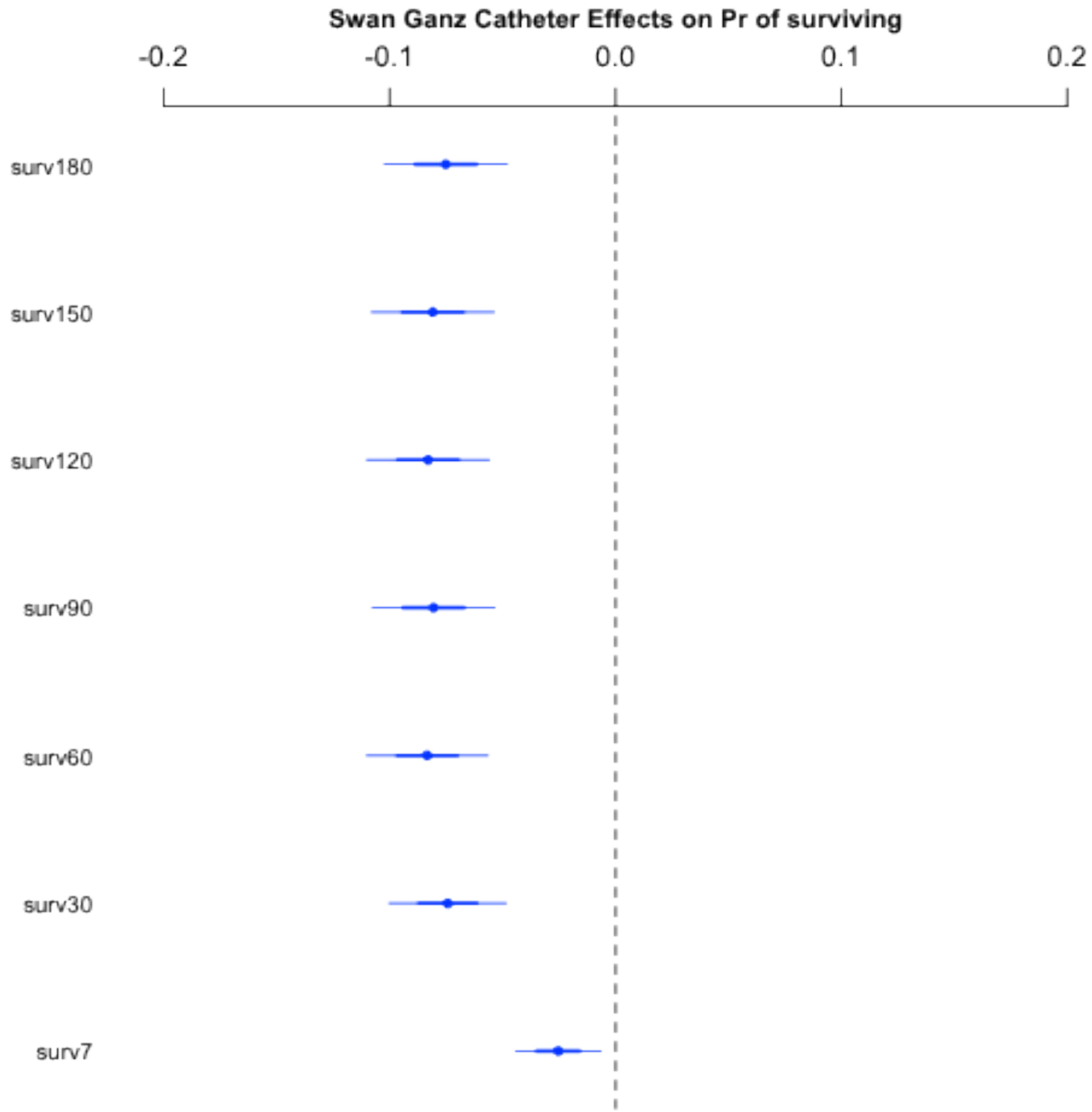
A first “stunning” regression

- We run the following regressions:

$$surv7_i = \beta_0 + \beta_1 swang_i + u_i$$
$$surv30_i = \beta_0 + \beta_1 swang_i + u_i$$
$$\vdots$$
$$surv180_i = \beta_0 + \beta_1 swang_i + u_i$$

- What is the interpretation of β_1 in each model?

Estimated Effects of Catheterization



Results

- Our simple analysis seems to confirm the results of the first generation studies
- Catheterization increases the probability of death by 5%-8%, depending on the time span.
- Catheterization also increases the probability of death at 7 days from admission...

There are however potential problems

Concerns

- There is big concern:
  - Omitted variables………

Omitted Variable

- Catheterization may depend on other personal characteristic that affect both catheterization decision and outcome
- Sicker patients are probably catheterized more frequently than relatively healtier patients

Adding Regressors

We have a wealth of information about the patients:

- *Personal characteristic*
  - *age, sex, education, race, education, income, etc.*
- *Categories of admission diagnosis*
  - respiratory, cardiovascular, neurological, gastrointestinal, renal, metabolic, hematologic, etc.
- *Categories of comorbidities illness:*
  - hearth failure, depression, cirrhosis, tumor, HIV, etc.
- Other information
  - estimate of the prob. of surviving 2 months

New Model

- We estimate:

$$\begin{aligned}
surv7_i &= \beta_0 + \beta_1 swang_i + x_i\gamma + u_i \\
surv30_i &= \beta_0 + \beta_1 swang_i + x_i\gamma + u_i \\
&\vdots \\
surv180_i &= \beta_0 + \beta_1 swang_i + x_i\gamma + u_i
\end{aligned}$$

Additional Problem

- Recall that the outcome variables are binary
  - surv7 is either 1 (patient is alive after 7 days) or 0
- The linear probability model could not be appropriate
  - We should use Probit and Logit to verify the robustness of the results
  - In this case we have to be careful in interpreting the results…

“Full” Linear Probability Model

$$\underset{}{surv7} = \underset{(0.0590)}{0.6306} - \underset{(0.0094)}{0.0150} \times swang + \ldots$$

$$surv30 = \underset{(0.0803)}{0.2709} - \underset{(0.0126)}{0.0587} \times swang + \ldots$$

$$surv60 = \underset{(0.0825)}{0.1572} - \underset{(0.0131)}{0.0648} \times swang + \ldots$$

$$surv90 = \underset{(0.0827)}{0.1309} - \underset{(0.0132)}{0.0625} \times swang + \ldots$$

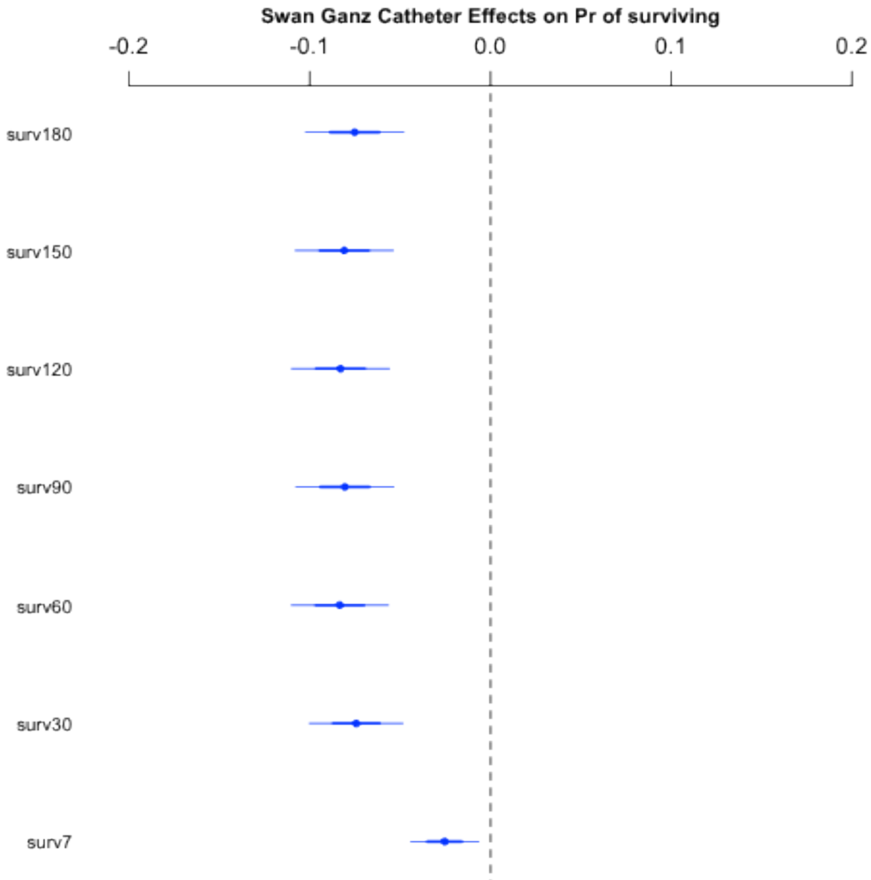
$$surv120 = \underset{(0.0830)}{0.0970} - \underset{(0.0133)}{0.0609} \times swang + \ldots$$

$$surv150 = \underset{(0.0829)}{0.0728} - \underset{(0.0133)}{0.0617} \times swang + \ldots$$

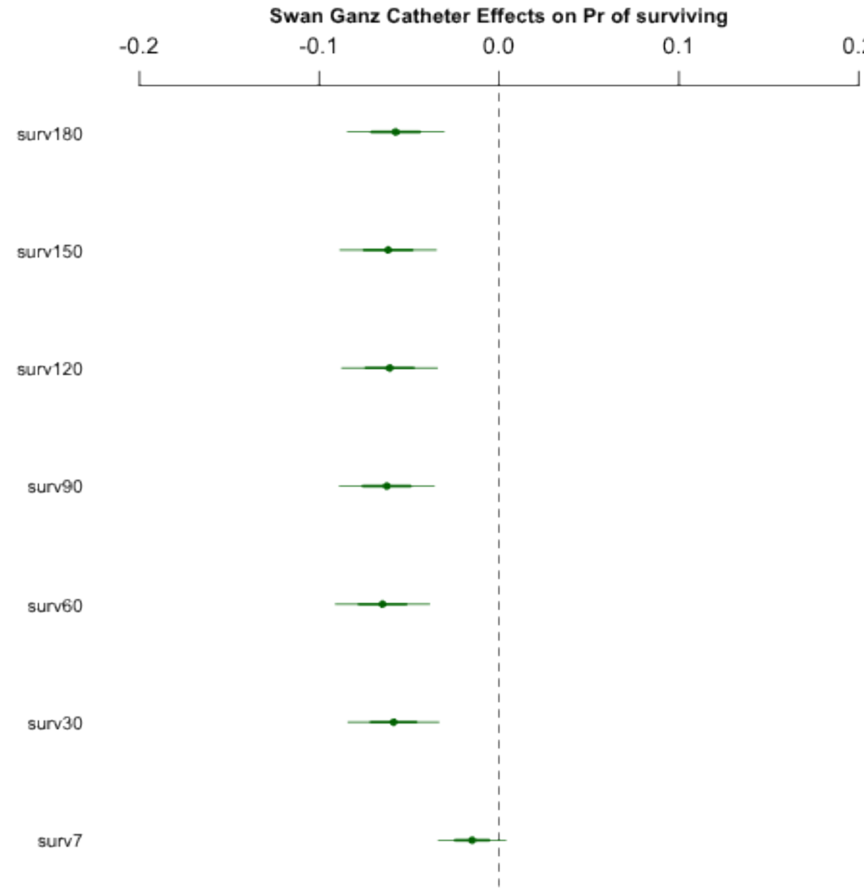
$$surv180 = \underset{(0.0829)}{0.0744} - \underset{(0.0134)}{0.0575} \times swang + \ldots$$

Comparison of the results

No controls



With controls



LUISS Guido Carli
LIBERA UNIVERSITÀ INTERNAZIONALE DEGLI STUDI SOCIALI

Comments

- When we add controls the effects tend to become smaller in magnitude, but they are still negative
- Except for surv7, they are still statistically different from 0
- The effect of catheterization ranges from 5 to 6 percentage points

- The unconditional probability of dying within 180 days from admission in ICU is

50.67%

- With a Swang-Ganz catheter

56.42%

- **56.42/ 50.67 = 1.11 more likely to die**

Probit and Logit

- We estimate now the following model:

$$\Pr(surv7_i = 1|swang_i, x_i) = \Phi(\beta_0 + \beta_1 swang_i + x_i'\gamma)$$

$$\Pr(surv30_i = 1|swang_i, x_i) = \Phi(\beta_0 + \beta_1 swang_i + x_i'\gamma)$$

$$\vdots$$

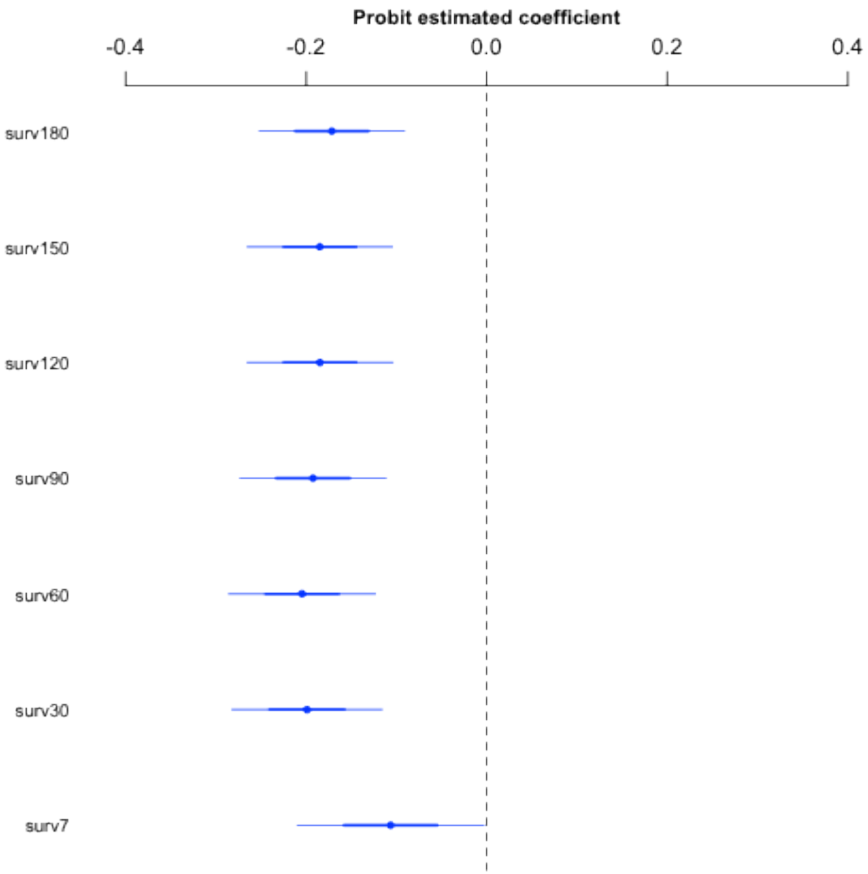
$$\Pr(surv180_i = 1|swang_i, x_i) = \Phi(\beta_0 + \beta_1 swang_i + x_i'\gamma)$$

- The interpretation of the coefficient is different..

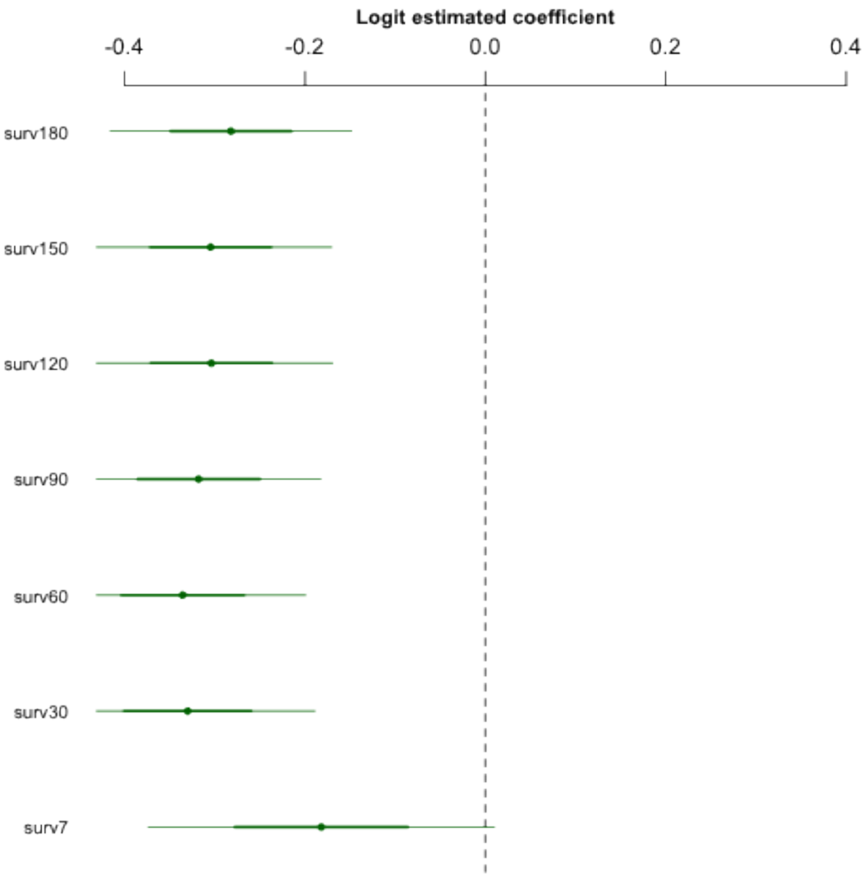
LUISS Guido Carli
LIBERA UNIVERSITÀ INTERNAZIONALE DEGLI STUDI SOCIALI

Probit/Logit Estimated Coefficients

Probit



Logit



LUISS Guido Carli
LIBERA UNIVERSITÀ INTERNAZIONALE DEGLI STUDI SOCIALI

Probit/Logit Effects

Probit - Effects

- Estimate

$$\Phi(\hat{\beta}_0 + \hat{\beta}_1 + \bar{x}'\hat{\gamma}) - \Phi(\hat{\beta}_0 + \bar{x}'\hat{\gamma})$$

Logit - Effects

- Estimate

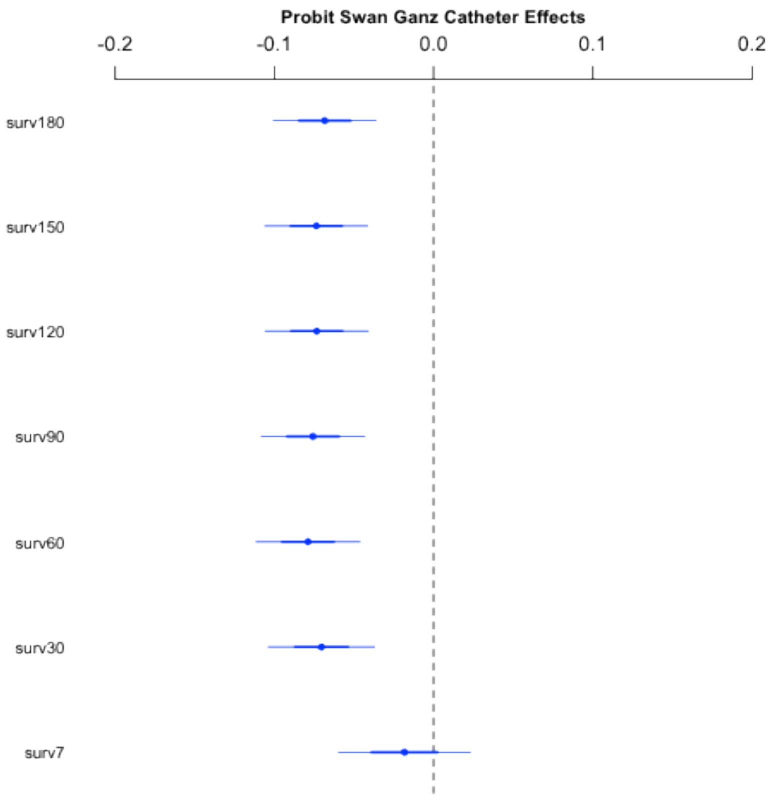
$$F(\hat{\beta}_0 + \hat{\beta}_1 + \bar{x}'\hat{\gamma}) - F(\hat{\beta}_0 + \bar{x}'\hat{\gamma})$$

where

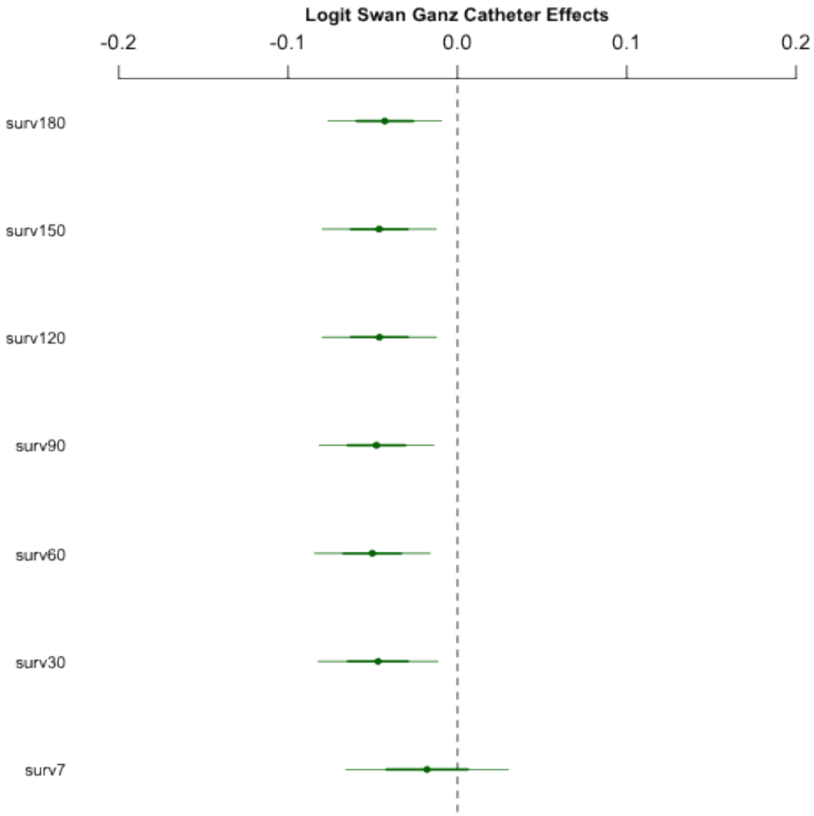
$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Probit/Logit Estimated Effects

Probit - Effects



Logit - Effects



Comments

- Using Probit and Logit instead of the Linear
- Probability model does not change the message:
  - Catheterization has a negative effect on the probability of surviving
  - The Probit effects are very close to the LPM effects
  - The Logit effects are somewhat smaller, yet still significantly different from 0

Unobserved Omitted Variable

- A threat to the internal validity of the study is the possibility that doctors do have information on health status of the patient

- This information may not be contained in the dataset

- If it is so, there is an 'unobserved' omitted variable that might render the interpretation of the effects as casual invalid

Instrument

- We could potentially solve the "endogeneity" caused by the unobserved omitted variable using instrumental variable technique.
- How do we find a valid instrument for this study?
- Think, think….

Two conditions for a valid instrument

For an instrumental variable (an "***instrument***") Z to be **valid**, it must satisfy two conditions:

- ***Instrument relevance***
- ***Instrument exogeneity***

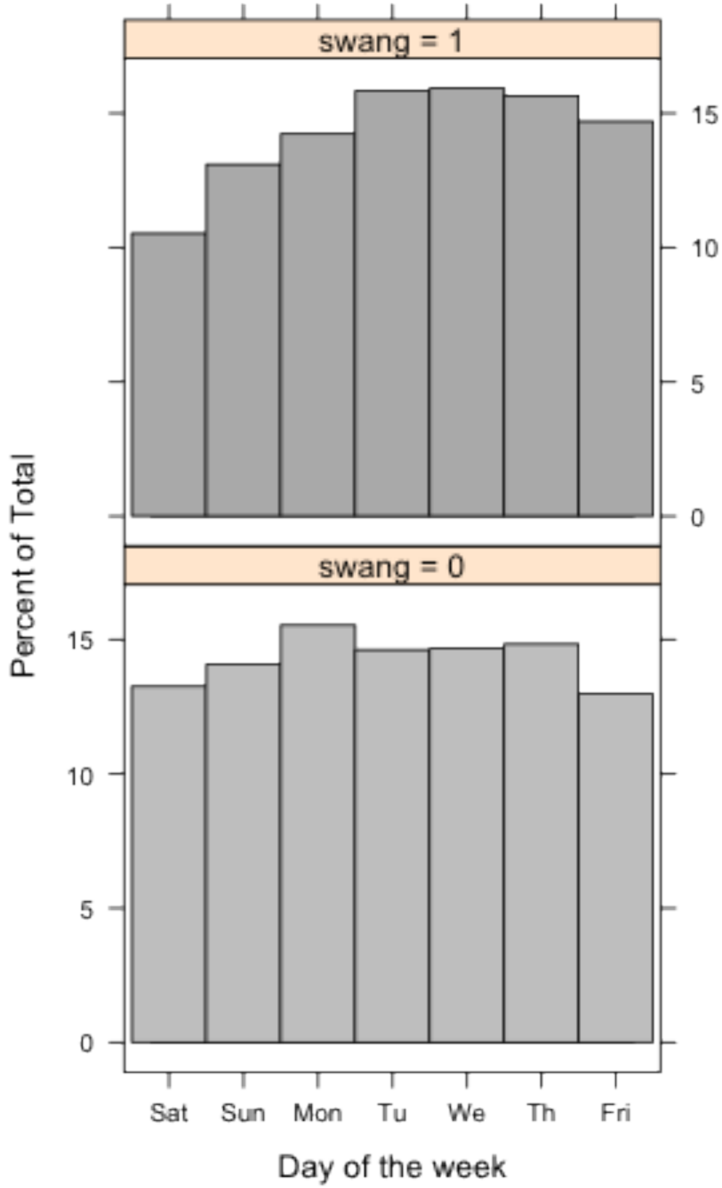
Instrumental Variable

- An instrument is a random variable Z
  - E[$Z'u$]=0 (exogenous)
  - It enters the first stage regression (relevance)
    - $surv7_i = \beta_0 + \underbrace{\beta_1 swang_i}_{\text{endogenous}} + \underbrace{x_i'\gamma}_{\text{exogenouss}} + u_i$

$$\underbrace{swang_i}_{\text{endogenous}} = \pi_0 + \pi_1 Z_i + \underbrace{x_i'\delta}_{\text{exogenouss}} + u_i$$

First Stage or Reduced Form

$$\pi_1 \neq 0$$



- There are more catheterization on weekdays:
  - ✓ Tuesday, Wednesday, Thuersday, Friday

Instrument

- We consider the following instrument
  - $$Z_i = \begin{cases} 1 & \text{if day of admission is Su, Sa, Mo} \\ 0 & otherwise \end{cases}$$
- Two questions:
  - Is it **exogenous**?
  - Is it **relevant**?

Exogeneity of day of the week

- Exogeneity means that the day of the week is uncorrelated with the unobserved health

- There is not *a priori* reason to believe that the day of the week is correlated with unobserved health conditions

- We **cannot** test this assumption since the model is exactly identified.

Relevance

- The instrument must be correlated (in the first stage sense) with *swang*

- *The idea is that*
  - *On weekends there are fewer ICU doctors on the premises and the probability of being catheterized is much lower;*
  - *There is evidence that staffing affects treatment decisions*

Testing relevance

First Stage

- Simple regression:

$$swang = \underset{(0.0076)}{0.3884} - \underset{(0.0141)}{0.0264} \times Z$$

- **First Stage**

$$swang = \underset{(0.0824)}{0.6175} - \underset{(0.0129)}{0.0271} \times Z + x_i'\hat{\delta}$$

Testing

- t-test:

$$\frac{0.0271}{0.0129} = 2.10 > 1.96$$

Statistically $\neq 0$ at 5%

TSLS

- Predict *swang*

$$\widehat{swang}_i = \underset{(0.0824)}{0.6175} - \underset{(0.0129)}{0.0271} \times Z_i + x_i'\hat{\delta}$$

- Second Stage(s)

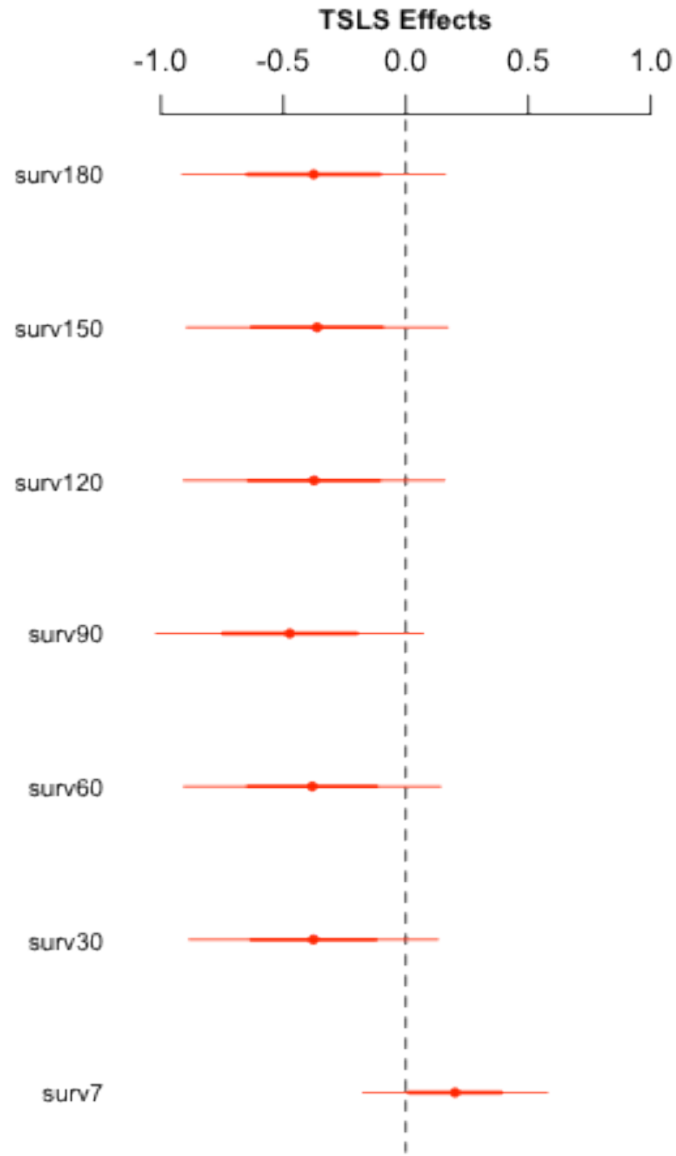
$$surv7_i = \beta_0 + \beta_1 \widehat{swang}_i + x_i\gamma + u_i$$

$$surv30_i = \beta_0 + \beta_1 \widehat{swang}_i + x_i\gamma + u_i$$

$$\vdots$$

$$surv180_i = \beta_0 + \beta_1 \widehat{swang}_i + x_i\gamma + u_i$$

TSLS Results



Comments

- TSLS:
  - The effects are estimated to be much (implausibly) larger
  - Catheterization seems to reduce mortality at shorter time horizon (7 days)
  - These effects are all statistically insignificant (not different from 0)

Caveats

- The first stage involves a binary variables
  - The predictions

$$\widehat{swang}_i = \underset{(0.0824)}{0.6175} - \underset{(0.0129)}{0.0271} \times Z_i + x_i'\hat{\delta}$$

  Can be outside the [0,1]

  - There are way to fix this problem by using a "probit instrumental variable" method.

- Weak instruments…..

Weak instruments

- The rule-of-thumb says that we should worry if the *Wald-stat* >*10m*

- What is the *Wald-stat* in our first stage?
  - *Hint 1: there is only one endogenous regressor*
  - *Hint 2: for testing only one coefficient the Wald-test is numerically equivalent to ….*

Effect of Education on Wages

- The objective is to estimate the "causal" effect of education on individual wages
- Very difficult to do with OLS because of omitted ability
- This is usually addressed by Instrumental Variables techniques

Data

- D. Card (1995), **"Using Geographic Variation in College Proximity to Estimate the Return to Schooling,"** in Aspects of Labour Market Behavior: Essays in Honour of John Vanderkamp. Ed. L.N. Christophides, E.K. Grant, and R. Swidinsky, 201-222. Toronto: University of Toronto Press.

Data

- Sample of 3009 workers from

  National Longitudinal Survey

  http://www.bls.gov/nls/

- Variables:
  - Hourly wage (**lwage**), age (**age**), race (**black**), experience (**exper**), geographic information (**smsa**), family background (**fatheduc, motheduc**), standardized test scores **(IQ, KWWW).**

Data

- MSA – Metropolitan Statistical Area
  - In the United States, a **metropolitan area** refers to a geographical region with a relatively high population density at its core and close economic ties throughout the area.
- KWW – Knowledge of the World of Work
  - A direct measure of “ability”
- IQ
  - Measure of intelligence.

Omitted Variable Bias

$$log(wage) = \beta_0 + \beta_1\, educ + u$$

- "Ability Bias"
  - Suppose that some individuals have an an unobserved characteristic ("ability") that enables them to earn higher wages at any level of education. If these individuals acquire higher than average schooling then the OLS estimate of β_1 will be upward-biased

Basic regression

```
> summary_rob(lm(lwage~educ, data=cd))

Dependent varaible:  lwage

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.57088    0.03914   142.4   <2e-16
educ         0.05209    0.00291    17.9   <2e-16
---
Heteroskadasticity robust standard errors used

Residual standard error: 0.4214 on 3008 degrees of freedom
Multiple R-squared:  0.09874, Adjusted R-squared:  0.09844
Wald-statistic: 320.5 on 1 and Inf DF,  p-value: < 2.2e-16
```

Return to education are estimated at 5.2%

Regression with controls

```
Dependent varaible:  lwage

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.6252589  0.0833168  55.514  < 2e-16
educ         0.0733229  0.0045343  16.171  < 2e-16
black       -0.1614265  0.0244866  -6.592 5.39e-11
south       -0.1104188  0.0177876  -6.208 6.40e-10
smsa         0.1621471  0.0180048   9.006  < 2e-16
exper        0.0885766  0.0080561  10.995  < 2e-16
expersq     -0.0023789  0.0003999  -5.949 3.13e-09
fatheduc    -0.0007076  0.0031278  -0.226    0.821
motheduc     0.0078193  0.0037065   2.110    0.035
---
Heteroskadasticity robust standard errors used

Residual standard error: 0.3773 on 2211 degrees of freedom
  (790 observations deleted due to missingness)
Multiple R-squared:  0.2663,      Adjusted R-squared:  0.2636
Wald-statistic: 836.8 on 8 and Inf DF,  p-value: < 2.2e-16
```

Return to education ≈ 7.3%

Ability

```
> summary_rob(lm(lwage~educ+black+south+smsa+exper+expersq+fatheduc+motheduc+IQ+KWW,data=cd))
Dependent varaible:  lwage

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.4909962  0.1266845  35.450  < 2e-16
educ         0.0628479  0.0066013   9.521  < 2e-16
black       -0.0963360  0.0364614  -2.642 0.008319
south       -0.0784613  0.0209866  -3.739 0.000192
smsa         0.1473359  0.0213596   6.898 7.59e-12
exper        0.0903101  0.0110791   8.151 7.21e-16
expersq     -0.0026815  0.0005713  -4.693 2.92e-06
fatheduc    -0.0046165  0.0037345  -1.236 0.216571
motheduc     0.0079514  0.0044815   1.774 0.076211
IQ           0.0018442  0.0008612   2.142 0.032384
KWW          0.0039044  0.0016718   2.335 0.019648
---
Heteroskadasticity robust standard errors used

Residual standard error: 0.3723 on 1593 degrees of freedom
  (1406 observations deleted due to missingness)
Multiple R-squared:  0.2162,	Adjusted R-squared:  0.2113
Wald-statistic: 472.1 on 10 and Inf DF,  p-value: < 2.2e-16
```

Return to education $\approx 6.28\%$

Retrun to education

- Returns to education are approximately **6%**
  - Confidence interval is

$$[0.050, 0.076]$$

Instruments

- `nearc2` and `nearc4` are indicator denoting whether the individual lives close to a 2 years and a 4 years college
- Are `nearc2` and `nearc4` valid instruments?
  - Relevance
  - Exogeneity

Relevance – First Stage Regressions

```
> summary_rob(lm(educ~nearc2+nearc4+black+south+smsa+exper+expersq+fatheduc+motheduc+IQ+KWW, data=cd))
Dependent varaible:  educ

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.886683   0.429152  23.038  < 2e-16
nearc2      -0.007681   0.077462  -0.099 0.921020
nearc4       0.231521   0.087863   2.635 0.008495
black        0.681182   0.140342   4.854 1.33e-06
south        0.129108   0.081699   1.580 0.114239
smsa         0.005987   0.094456   0.063 0.949470
exper       -0.553602   0.037193 -14.885  < 2e-16
expersq      0.012706   0.001953   6.504 1.04e-10
fatheduc     0.053552   0.014959   3.580 0.000354
motheduc     0.049008   0.017066   2.872 0.004137
IQ           0.022029   0.003291   6.694 3.01e-11
KWW          0.112311   0.005781  19.427  < 2e-16
---
Heteroskadasticity robust standard errors used

Residual standard error: 1.48 on 1592 degrees of freedom
  (1406 observations deleted due to missingness)
Multiple R-squared:  0.5762,           Adjusted R-squared:  0.5733
Wald-statistic:  3277 on 11 and Inf DF,  p-value: < 2.2e-16
```

```
wtest(lm1, testcoef=c("nearc2", "nearc4"))

Wald test

Null hypothesis:
nearc2 = 0
nearc4 = 0

 q        W    pvalue
 2 7.037288 0.0296396
```

The null hypothesis that nearc2 and nearc4 do **not enter** the first stage regression is rejected at 5%

TSLS

```
coeftest(ivreg(lwage~educ+black+south+smsa+exper+expersq+fatheduc+motheduc+IQ+KWW|
                  .+nearc2+nearc4-educ  , data=cd), vcov.=vcovHC)

              Estimate Std. Error t value  Pr(>|t|)    
(Intercept)  4.6448384  0.9213975  5.0411 5.157e-07 ***
educ         0.0474293  0.0916678  0.5174 0.6049458    
black       -0.0857098  0.0741675 -1.1556 0.2480081    
south       -0.0770074  0.0227986 -3.3777 0.0007485 ***
smsa         0.1484669  0.0223741  6.6357 4.415e-11 ***
exper        0.0817947  0.0519974  1.5731 0.1159051    
expersq     -0.0024872  0.0013017 -1.9107 0.0562248 .  
fatheduc    -0.0037760  0.0062892 -0.6004 0.5483230    
motheduc     0.0086968  0.0063199  1.3761 0.1689832    
IQ           0.0021872  0.0021802  1.0032 0.3158995    
KWW          0.0056362  0.0104880  0.5374 0.5910661    
---
```