Econometrics II

Lecture 5: Instrumental Variables Part II

Måns Söderbom*

13 April 2011

*Department of Economics, University of Gothenburg. Email: mans.soderbom@economics.gu.se. Web: www.economics.gu.se/soderbom, www.soderbom.net.

1. Introduction

In Lecture 4 we discussed the IV estimator in a modelling framework in which the causal effect of interest is assumed constant across individuals (e.g. the return to education is the same for everyone).

Assuming the causal effect to be constant (homogeneous) across individuals is clearly quite a restrictive.

In this lecture I will discuss interpretation of the IV estimator in a less restrictive framework, in which potential outcomes - and hence treatment effects - are allowed to be heterogeneous across individuals.

References for this lecture:

Angrist and Pischke (2009), Chapter 4.4-4.5.

2. IV with Heterogenous Potential Outcomes

Reference: Angrist-Pischke, Chapter 4.4.

2.0.1. LATE: Setting the scene

- Common features of the type of environment for which we may want to (and be able to do so) estimate the **Local Average Treatment Effect**, LATE:
  - The treatment status, from now on denoted D_i, depends on an underlying instrument Z_i.
  - The effect of Z_i on treatment is heterogeneous.
  - The effect of treatment D_i on the outcome variable of interest Y_i is also heterogeneous.
- Thus, the causal chain is as follows:

$$Z_i \to D_i \to Y_i,$$

  and we are primarily interested in the effect of treatment on outcomes; i.e. $D_i \to Y_i$.

- Define $Y_i(d, z)$ as the potential outcome of individual i, were this individual to have treatment status $D_i = d$ and instrument value $Z_i = z$. We focus on the case where both d and z can take two values, 0 or 1. That is, D_i and Z_i are dummy variables.

- Following Angrist-Pischke, we relate the exposition to a specific application, namely Angrist (1990), who looks at the effect of veteran status on earnings in the US.

Illustration: Angrist (1990; AER)

- **Context:** In the 1960s and 70s young men in the US were at risk of being drafted for military service in Vietnam. Fairness concerns led to the institution of a draft lottery in 1970 that was used to determine **priority** for conscription.
- In each year from 1970 to 1972, random sequence numbers were randomly assigned to each birth date in cohorts of 19-year-olds.
  - Men with lottery numbers below a cutoff were eligible for the draft
  - Men with lottery numbers above the cutoff were not.
- Many eligible men were exempted form service for health or other reasons.
- Others, who were not eligible, nevertheless volunteered for service.
- The smart idea: Use the lottery outcome as an instrument for veteran status, in an analysis of the causal effect of veteran status on earnings. What about relevance and validity?
  - While the lottery didn't completely determine veteran status, it certainly mattered: relevance.
  - The lottery outcome was random and seems reasonable to suppose that its only effect was on veteran status: validity.
- The instrument is thus defined as follows:

$$\begin{aligned} Z_i &= 1 \text{ if lottery implied individual } i \text{ would be draft eligible,} \\ Z_i &= 0 \text{ if lottery implied individual } i \text{ would not be draft eligible.} \end{aligned}$$

- The instrument affects treatment, which is this application amounts to entering the military service. The econometrician observes treatment status as follows:

$$\begin{aligned} D_i &= 1 \text{ if individual } i \text{ served in the Vietnam war (veteran),} \\ D_i &= 0 \text{ if individual } i \text{ did not serve in the Vietnam war (not veteran);} \end{aligned}$$

- Now define potential outcomes for D_i as D_{0i} and D_{1i}, respectively, where D_{0i} is the treatment status when $Z_i = 0$ and D_{1i} is the treatment status when $Z_i = 1$. We thus have:

$$\begin{aligned} D_{0i} &= 0 \text{ if individual } i \text{ would not serve in the military if not draft eligible} \\ D_{0i} &= 1 \text{ if individual } i \text{ would serve in the military even though not draft eligible} \\ D_{1i} &= 0 \text{ if individual } i \text{ would not serve in the military even though draft eligible} \\ D_{1i} &= 1 \text{ if individual } i \text{ would serve in the military if draft eligible.} \end{aligned}$$

- In view of this, the following way of categorizing types of individuals is useful (why will be clear later):

Compliers:	$D_{1i} = 1, D_{0i} = 0$
Never-takers:	$D_{1i} = 0, D_{0i} = 0$
Always-takers:	$D_{1i} = 1, D_{0i} = 1$
Defiers:	$D_{1i} = 0, D_{0i} = 1$

  Note that "defiers" are very odd cases - as we shall see, the basic LATE estimator assumes there are no defiers. In the present context, at least, it's hard to see why there might be defiers - right?

- The outcome variable of interest is earnings, and the main research question is whether veteran status causes earnings. The causal effect of veteran status, conditional on draft eligibility status, is defined as

$$Y_i(1, Z_i) - Y_i(0, Z_i).$$

- As usual, we can't identify individual treatment effects, because we don't observe all potential outcomes.

- Let's remind ourselves of what the OLS and IV estimators would look like in the present context.

2.0.2. Estimation by regression: OLS and IV

- If I use **OLS** to estimate a model of the following kind:

$$Y_i = \alpha + \theta D_i + \varepsilon_i,$$

  where α is a constant and ε_i a zero–mean residual, we know that θ^{OLS} is interpretable as an estimate of ATE and ATE_1, provided potential outcomes are independent of actual treatment status. That is, provided treatment is (as good as) randomly assigned. If that doesn't hold, OLS does not identify ATE or ATE_1. In the present context, it seems likely there are lots of unobservables correlated with veteran status, so the OLS estimator is hard to justify here.

- Suppose I were to use an **IV** estimator instead, with Z_i as a single instrument:

$$\begin{aligned} D_i &= \gamma + \phi Z_i + u_i \\ Y_i &= \alpha + \theta D_i + \varepsilon_i, \end{aligned}$$

  For now, don't worry about my reasons for doing this - just think about what I "would get" if I were to do this. Recall that in the **special case** where Z_i is a dummy variable, the IV estimator can be written simply as:

$$\theta^{IV} = \frac{E\left[Y_i|Z_i=1\right] - E\left[Y_i|Z_i=0\right]}{E\left[D_i|Z_i=1\right] - E\left[D_i|Z_i=0\right]},$$

  which is the **Wald estimator**. Inevitably, this is what I will get if I use IV to model earnings as a function of veteran status, while using draft eligibility status as an instrument for veteran status.

- But how should this quantity be **interpreted**? Does it estimate an average treatment effect?

- Yes, potentially.

2.0.3. LATE: A distinct evaluation parameter

- One common "evaluation parameter" estimated by means of instrumental variable techniques is the **Local Average Treatment Effect** (LATE).

- Suppose we are concerned that OLS doesn't identify ATE because there are unobserved differences between veterans and nonveterans (the standard endogeneity concern). We propose to use the draft lottery outcome as an instrument for veteran status. Suppose we are prepared to make **four assumptions** as follows:

Assumption A1: Independence between the potential outcomes $[Y_i(D_{1i},1), Y_i(D_{0i},0), D_{1i}, D_{0i}]$ and the instrument Z_i. That is, the instrument is as good as randomly assigned. This is what we mean by "exogenous" in the present context.

Assumption A2: Exclusion restriction. The potential outcomes $Y_i(d,z)$ is **only** a function of d; they are only affected by the instrument Z_i through the treatment variable D_i. Implies $Y_i(d,0) = Y_i(d,1)$.

Assumption A3: Relevance, first stage. $E[D_{1i} - D_{0i} \neq 0]$. The average causal effect of the instrument on veteran status is not zero.

Assumption A4: Monotonicity. $D_{1i} - D_{0i} \geq 0$ for all individuals (or vice versa). That is, no defiers.

- A1 states that the instrument is as good as randomly assigned. Draft eligibility was determined by a lottery, lending credibility to this assumption.

- A2 says that the instrument can have no direct effect on the outcome variable (earnings). May or may not hold in this case. (Why might it not hold?)

- A3 says that the instrument impacts on treatment - easy to check in practice.

- A4 says that any man who would serve if not draft eligible, would also serve if draft eligible. A reasonable assumption in this case it would seem.

- Under these assumptions the parameter you're estimating in the second stage of your IV procedure (the coefficient on veteran status, D_i) is interpretable as measuring the average effect of military service on earnings for *men who served because they were draft eligible, but who would not have served had they not been draft eligible.* That is, the average affect for the group of men whose treatment status can be changed by the instrument - the "compliers". Note that this group of people does not include volunteers (always-takers) or men who were exempted from service (never takers).

- The average effect for the compliers is a parameter called the LATE. Mathematically, we define the LATE as

$$LATE = E\left[Y_{1i} - Y_{0i} | D_{1i} - D_{0i} > 0\right],$$

  where $Y_{1i} - Y_{0i}$ denotes the difference in outcomes due to treatment, D_{1i} is the potential treatment status when the instrument $Z_i = 1$ and D_{0i} is the potential treatment status when $Z_i = 0$. Clearly $D_{1i} - D_{0i} > 0$ only applies for compliers.

- Under assumptions A1-A4 we can show that the Wald estimator **coincides** with the expression for $LATE$. In other words, IV identifies $LATE$, in this case.

- This is known as the **LATE Theorem**.

2.0.4. Analysis: Why Wald = LATE?

- Relate observed treatment status to potential treatment outcomes:

$$\begin{aligned} D_i &= D_{0i} + \left(D_{1i} - D_{0i}\right) Z_i, \\ D_i &= \pi_0 + \pi_{1i} Z_i + \xi_i, \end{aligned}$$

where $\pi_0 = E(D_{0i})$ and $\pi_{1i} = (D_{1i} - D_{0i})$ is the (note) **heterogeneous** causal effect of the instrument on D_i. Assumption A1 (independence) implies π_{1i} is interpretable as the causal effect of Z_i on treatment (compare this to the case where treatment is randomized). Assumption A4 (monotonicity) implies that $\pi_{1i} \geq 0$ for all i or $\pi_{1i} \leq 0$ for all i.

- Recall that the potential outcome of our main "dependent variable" is defined $Y_i(d,z)$. Assumption A2 (exclusion restriction) implies $Y_i(d,0) = Y_i(d,1)$, hence we can write the observed outcome:

$$\begin{aligned} Y_i &= Y_i(0,Z_i) + [Y_i(1,Z_i) - Y_i(0,Z_i)] D_i, \\ Y_i &= Y_{0i} + [Y_{1i} - Y_{0i}] D_i \\ Y_i &= \alpha_0 + \rho_i D_i + \eta_i, \end{aligned}$$

where $\alpha_0 = E(Y_{0i})$, $\rho_i = Y_{1i} - Y_{0i}$ is a **random coefficient** and η_i measures the discrepancy between $E(Y_{0i})$ and Y_{0i}. Note the heterogeneous causal effect of treatment (e.g. veteran status) on your outcome variable of interest (e.g. earnings).

- Now consider the formula for the Wald estimator:

$$\frac{E[Y_i|Z_i=1] - E[Y_i|Z_i=0]}{E[D_i|Z_i=1] - E[D_i|Z_i=0]}.$$

The exclusion restriction (A2) and independence (A1) assumptions imply

$$\begin{aligned} E[Y_i|Z_i=1] &= E[Y_{0i} + [Y_{1i} - Y_{0i}] D_{1i}|Z_i=1], \\ E[Y_i|Z_i=1] &= E[Y_{0i} + [Y_{1i} - Y_{0i}] D_{1i}]. \end{aligned}$$

By the same principles,

$$E[Y_i|Z_i=0] = E[Y_{0i} + [Y_{1i} - Y_{0i}] D_{0i}],$$

and so the numerator in the Wald estimator can be written

$$\begin{aligned}
\text{Wald-numerator} &= E\left[Y_i|Z_i=1\right]-E\left[Y_i|Z_i=0\right] \\
&= E\left[Y_{0i}+\left[Y_{1i}-Y_{0i}\right]D_{1i}\right]-E\left[Y_{0i}+\left[Y_{1i}-Y_{0i}\right]D_{0i}\right] \\
&= E\left[\left(Y_{1i}-Y_{0i}\right)\left(D_{1i}-D_{0i}\right)\right].
\end{aligned}$$

Now, monotonicity implies $(D_{1i}-D_{0i})$ is either equal to 1 or 0; hence

$$\begin{aligned}
\text{Wald-numerator} &= E\left[\left(Y_{1i}-Y_{0i}\right)\left(D_{1i}-D_{0i}\right)\right] \\
&= P\left(D_{1i}-D_{0i}>0\right)E\left[\left(Y_{1i}-Y_{0i}\right)|D_{1i}-D_{0i}>0\right].
\end{aligned}$$

The denominator of the Wald formula is

$$\text{Wald-denominator}=E\left[D_i|Z_i=1\right]-E\left[D_i|Z_i=0\right].$$

We can use exactly the same principles as for the numerator, and arrive at

$$\begin{aligned}
\text{Wald-denominator} &= P\left(D_{1i}-D_{0i}>0\right)E\left[\left(D_{1i}-D_{0i}\right)|D_{1i}-D_{0i}>0\right] \\
&= P\left(D_{1i}-D_{0i}>0\right).
\end{aligned}$$

Hence

$$\begin{aligned}
\frac{E\left[Y_i|Z_i=1\right]-E\left[Y_i|Z_i=0\right]}{E\left[D_i|Z_i=1\right]-E\left[D_i|Z_i=0\right]} &= E\left[\left(Y_{1i}-Y_{0i}\right)|D_{1i}-D_{0i}>0\right] \\
&= E\left[\rho_i|\pi_{1i}>0\right]
\end{aligned}$$

i.e. the IV estimator identifies $LATE$.

[Example: Simulating LATE. Appendix 2]

Summing up, we have seen how we can identify LATE - i.e. the average effect of treatment for the subpopulation of compliers. Compliers in the specific application referred to above are individuals who were induced by the draft lottery to serve in the military. Never-takers who would not serve irrespective of their lottery number, and always-takers, who would volunteer irrespective of their lottery number, clearly do not belong to this group.

Is LATE an economically interesting quantity? Perhaps. In any case, it may be that we cannot identify average treatment effects for the population because we cannot identify the average causal effect of treatment amongst the never-takers or the always-takers, if our instrument has no effect on individuals belonging to these groups.

3. Generalizing LATE

Reference: Section 4.5 in Angrist-Pischke

You have seen how the LATE theorem applies to stripped down model: a single dummy instrument is used to estimate the impact of a dummy treatment with no covariates. Whilst elegant, it has to be said this is a bit of a special case. (What about continuous variables? Or covariates?)

AP discuss three ways of extending the basic LATE framework:

1. $LATE$ with multiple instruments. The LATE is always closely connected to the underlying instrument, since whether someone is a complier likely depends on what the instrument is. *Different instruments will therefore identify different LATE:s.* If we have, say, two instruments with distinct complier groups and thus distinct LATEs, using 2SLS with both instruments simultaneously produces a linear combination of the instrument-specific LATEs. Whether or not that is interesting clearly depends on the context.

2. Covariates - where did the x-variables go? Clearly if have instruments that are randomly assigned, we don't really need to control for x-variables (as these will be orthogonal to the instrument anyway). However, the instrument may in fact covary with x-variables that also impact potential outcomes, in which case we should control for x-variables.

3. Variable treatment intensity. Our treatment variable, rather than being binary, can take on more than two values - e.g. years of schooling.

The econometric tool is still 2SLS. Let's have a closer look at these three settings.

3.1. LATE with multiple instruments

Consider a pair of dummy instruments, z_{1i} and z_{2i}. Assume these are mutually exclusive - i.e. together with a constant, they exhaust the information in the instrument set (e.g. $z_{1i} = 1$ if someone is born in May, June, July, August; $z_{2i} = 1$ if born in September or later in the year). We assume monotonicity holds for each of the instruments with a positive first stage. The author assumptions underlying the LATE theorem are assumed to hold as well.

This means we could obtain two different LATEs:

- If we were to use only z_{1i} we would get the LATE for the compliers associated with z_{1i} (e.g. those for whom being born in the May-August period **altered** their schooling; had these individuals not been born in this period, their schooling would differ (perhaps not the most intuitive example)).
- If we were to use only z_{2i} we would get the LATE for the compliers associated with z_{2i}.

These populations may not be the same; and if they have systematically different treatment effects, the two LATEs would differ. This is because the LATE is instrument-specific.

Suppose now we use both z_{1i} and z_{2i} in a 2SLS procedure. How should we interpret the estimated treatment effect?

Your econometric intuition at this stage suggests that you will get some kind of average of the two LATEs described above. This is indeed the case. Let

$$\rho_j = \frac{Cov\left(Y_i, z_{ji}\right)}{Cov\left(D_i, z_{ji}\right)}; j = 1, 2$$

denote the two IV estimands (the two LATEs). Now consider our 2SLS estimator based on a first-stage in which both z_{1i} and z_{2i} are included.

The first stage will give us

$$\hat{D}_i = \pi_{11} z_{1i} + \pi_{12} z_{2i},$$

where π_{11} and π_{12} are positive numbers (there should be a constant here too but its omission doesn't affect the point). By definition,

$$\rho_{2SLS} = \frac{Cov\left(Y_i, \hat{D}_i\right)}{Cov\left(\hat{D}_i, \hat{D}_i\right)} = \frac{Cov\left(Y_i, \hat{D}_i\right)}{Cov\left(D_i, \hat{D}_i\right)},$$

hence

$$\begin{aligned}
\rho_{2SLS} &= \frac{\pi_{11} Cov\left(Y_i, z_{1i}\right)}{Cov\left(\hat{D}_i, \hat{D}_i\right)} + \frac{\pi_{12} Cov\left(Y_i, z_{2i}\right)}{Cov\left(\hat{D}_i, \hat{D}_i\right)} \\
&= \frac{\pi_{11} Cov\left(D_i, z_{1i}\right)}{Cov\left(\hat{D}_i, \hat{D}_i\right)} \frac{Cov\left(Y_i, z_{1i}\right)}{Cov\left(D_i, z_{1i}\right)} \\
&\quad + \frac{\pi_{12} Cov\left(D_i, z_{2i}\right)}{Cov\left(\hat{D}_i, \hat{D}_i\right)} \frac{Cov\left(Y_i, z_{2i}\right)}{Cov\left(D_i, z_{2i}\right)} \\
\rho_{2SLS} &= \psi \rho_1 + (1 - \psi) \rho_2,
\end{aligned}$$

where

$$\psi = \frac{\pi_{11} Cov\left(D_i, z_{1i}\right)}{\pi_{11} Cov\left(D_i, z_{1i}\right) + \pi_{12} Cov\left(D_i, z_{2i}\right)}$$

is a weight bounded between zero and one.

Note that ψ depends on the relative strength of each instrument in the first stage. In other words, 2SLS is a weighted average of the causal effects for instrument-specific compliant subpopulations (the two LATEs).

3.2. Covariates in the heterogeneous effects model

The main reason covariates are included in a regression is that the conditional independence and exclusion restrictions underlying IV estimation may be more likely valid after conditioning on covariates.

For example, in the case of draft eligibility, older cohorts were more likely to be eligible be design, and because earnings covary with age (experience), draft eligibility is a valid instrument only conditional on year of birth.

IV estimation with covariates in the present framework may be justified by a conditional independence assumption of the following type:

$$Y_{i0}, Y_{i1}, D_{i0}, D_{i1} \text{ independent of } z_i | X_i,$$

i.e. think of the IV as being as good as randomly assigned, conditional on X_i.

- If we maintain the assumption that the causal effect of interest is constant, we are of course in very familiar territory: just include X_i as a control vector in the first and second stage of your 2SLS estimator.

- To relax the constant effects assumption, we might be prepared to specify the causal effect as

$$Y_{i1} - Y_{i0} = \rho\left(X\right),$$

i.e. as dependent on observable characteristics. This model can be estimated by adding interactions between D and X to the second stage and interactions between z and X in the first stage. You would then have more than one first-stage regression:

$$\begin{aligned}
D_i &= X_i'\pi_{00} + \pi_{01}z_i + z_iX_i'\pi_{02} + \xi_{0i} \\
D_iX_i^1 &= X_i'\pi_{00} + \pi_{01}z_i + z_iX_i'\pi_{02} + \xi_{0i} \\
&(...) \\
D_iX_i^K &= X_i'\pi_{00} + \pi_{01}z_i + z_iX_i'\pi_{02} + \xi_{0i}
\end{aligned}$$

assuming there are K covariates. The second stage would in this case be specified as

$$Y_i = \alpha' X_i + \rho_0 D_i + D_i X_i' \rho_1 + \eta_i,$$

which implies

$$\rho(X) = \rho_0 + X_i' \rho_1$$

- The heterogeneous effects model underlying the LATE theorem can also be extended to allow for covariates. Unfortunately, interpretation becomes less straightforward. We now have **covariate-specific** LATEs of the form

$$\lambda(X_i) \equiv E[Y_{1i} - Y_{0i} | X, D_{1i} > D_{0i}].$$

  If we work with a first-stage in which dummies for each value taken by the X-variable are used as predictors, the 2SLS estimator will produce a weighted average of these covariate-specific LATEs, where the weights will be higher for covariate values where the instrument creates more variation in fitted values.

- This estimator is not very attractive. Almost certainly, the first stage will contain too many instrument resulting in bias towards the OLS estimator. Moreover, interpretation of the 2SLS estimate of the causal effect is not straightforward (a weighted average of.... what exactly?).

- If we modify the first stage, perhaps using just X rather than dummy variables for each value of X, this may be an acceptable approximation. See Theorem 4.5.2 in AP for details.

4. Variable Treatment Intensity

Now consider a treatment variable that can take more than two values, e.g. schooling. Define potential outcomes as

$$Y_{si} \equiv f_i(S).$$

Suppose S_i takes on values in the set

$$\left\{0, 1, ..., \bar{S}\right\},$$

implying there are $\bar{S}$ causal effects, $Y_{si} - Y_{s-1,i}$. A linear causal model assumes these are the same for all s and i, "...obviously unrealistic assumptions" according to AP.

However, you can still justify using 2SLS and a linear specification of the form

$$Y_i = \alpha + \rho S_i + \eta_i.$$

It can be shown that ρ_{2SLS} in this case is a weighted average of unit causal effects, where the weights are determined by how the compliers are distributed over the range of S_i.

For example, returns to schooling estimated using quarter of birth come from shifts in the distribution of grades completed in **high school** (why?). Other instruments, such as distance to school, act elsewhere on the schooling distribution and therefore capture a different sort of return.

So, you see, it's all about interpretation!

PhD Programme: Econometrics II
Department of Economics, University of Gothenburg
Appendix: Lecture 5
Måns Söderbom

The Local Average Treatment Effect

1. Simulating LATE in Stata

Stata code:

```
clear
set seed 54687
set obs 20000

/* first, randomly assign the instrument  - say half-half */
ge z = uniform()>.5

/* then, generate never-takers (d00), always-takers (d11) and compliers
(d01), independent of z */

ge d00=(_n<=5000)
ge d11=(_n>5000 & _n<=10000)
ge d01=(_n>10000)

/* observed outcomes: always zero for never-takers, always one for
always-takers, depends on the IV for compliers */
ge D=d11+z*d01

/* now give the three groups different LATE. Without loss of
generality, assume within group homogeneity. */

ge late=-1 if d00==1
replace late=0 if d11==1
replace late=1 if d01==1

/* next generate potential outcomes y0,y1 */

ge y0=0.25*invnorm(uniform())
ge y1=y0+late

/* actual outcome depends on treatment status */
ge y = D*y1+(1-D)*y0

/* the average treatment effect is simply the sample mean of late */
sum late

/* OLS doesn't give you ATE or LATE */
reg y D

/* IV gives you the LATE for the compliers */
ivreg y (D=z)

exit
```

Results:

```
. /* the average treatment effect is simply the sample mean of late */
. sum late

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        late |     20000         .25    .8291769         -1          1

.
. /* OLS doesn't give you ATE or LATE */
. reg y D

      Source |       SS       df       MS              Number of obs =   20000
-------------+------------------------------           F(  1, 19998) = 6631.34
       Model |  1246.68658     1  1246.68658           Prob > F      =  0.0000
    Residual |  3759.60651 19998  .187999125           R-squared     =  0.2490
-------------+------------------------------           Adj R-squared =  0.2490
       Total |  5006.29309 19999  .250327171           Root MSE      =  .43359

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           D |   .4993491    .006132    81.43   0.000     .4873298    .5113684
       _cons |   .0005316   .0043511     0.12   0.903     -.007997    .0090602
------------------------------------------------------------------------------

.
. /* IV gives you the LATE for the compliers */
. ivreg y (D=z)

Instrumental variables (2SLS) regression

      Source |       SS       df       MS              Number of obs =   20000
-------------+------------------------------           F(  1, 19998) = 5048.40
       Model | -86.6843722     1 -86.6843722           Prob > F      =  0.0000
    Residual |  5092.97746 19998   .25467434           R-squared     =       .
-------------+------------------------------           Adj R-squared =       .
       Total |  5006.29309 19999  .250327171           Root MSE      =  .50465

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           D |   1.015767   .0142961    71.05   0.000     .9877453    1.043788
       _cons |  -.2594847   .0080341   -32.30   0.000    -.2752322   -.2437373
------------------------------------------------------------------------------
Instrumented:  D
Instruments:   z
------------------------------------------------------------------------------
```

Recall: Treatment effect is 1.0 for the compliers, 0.0 for the always-takers and -1.0 for the never-takers.