Instrumental variables

Mauricio Romero

Instrumental variables

Basic idea

Two stage least squares

Weak instruments

Practical IV Tips

Example

Heterogeneity and the LATE

Imperfect Compliance

Re-cap

Instrumental variables

When is IV used?

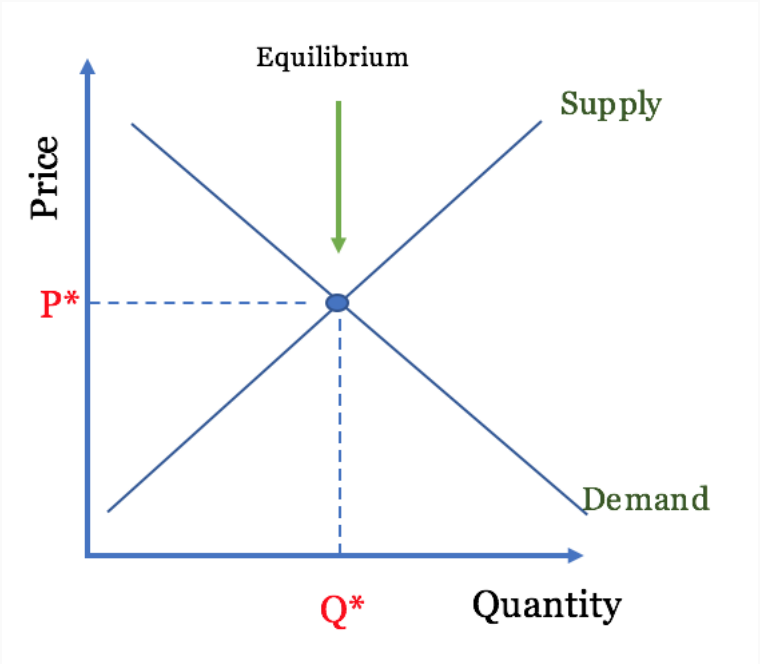
Instrumental variables methods are typically used to address the following kinds of problems

1. Omitted variable bias
2. Measurement error
3. Simultaneity bias
4. Reverse causality
5. Randomized control trials with noncompliance

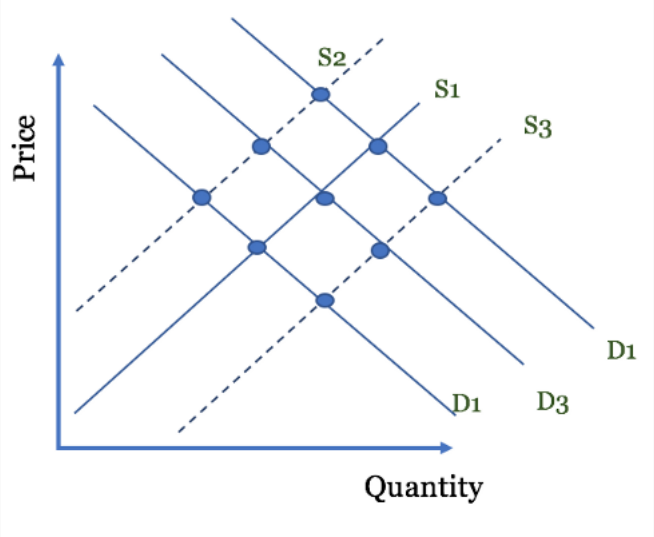
Demand and supply curves are unidentified

- You can't simple look at correlations between price and quantity to get elasticity of demand

- The pairs of quantity and price are equilibrium values and therefore don't reflect the demand or the supply curve


Equilibrium
Supply
Price
P*
Demand
Q*
Quantity

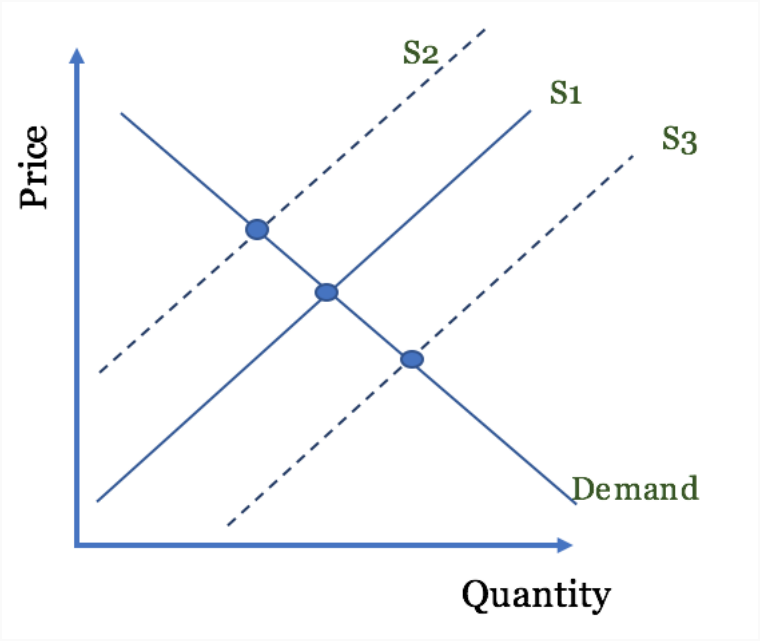

Price
Quantity


S2
S1
S3
Price
D1
D1
D3
Quantity

Solution:

- Something that shifts only one of the curves and traces the other

- This is called an "instrument"


S2
S1
S3
Price
Demand
Quantity

For now, assume constant treatment effects

- Constant treatment effects (i.e., β is constant across all individual units)
  - Constant treatment effects is the traditional econometric pedagogy when first learning instrumental variables
  - Identical to assuming that ATE=ATT=ATU because constant treatment effects assumes $\beta_i = \beta_{-i} = \beta$ for all units

Our causal model: Returns to schooling again

$$Y = \alpha + \delta S + \underbrace{\gamma A + \nu}_{\varepsilon}$$

where Y is log earnings, S is years of schooling, A is unobserved ability, and ε is the total error (ν is an idiosyncratic shock)

- Problem: S is correlated with the error term ε

Our causal model: Returns to schooling again

$$Y = \alpha + \delta S + \underbrace{\gamma A + \nu}_{\varepsilon}$$

where Y is log earnings, S is years of schooling, A is unobserved ability, and ε is the total error (ν is an idiosyncratic shock)

- Problem: S is correlated with the error term ε
- Suppose there is a variable, Z_i correlated with S_i and uncorrelated with A and ν

How can IV be used to obtain consistent estimates?

$$Cov(Y, Z) = Cov(\alpha + \delta S + \gamma A + \nu, Z)$$

How can IV be used to obtain consistent estimates?

$$
\begin{aligned}
Cov(Y, Z) &= Cov(\alpha + \delta S + \gamma A + \nu, Z) \\
&= E[(\alpha + \delta S + \gamma A + \nu)Z] - E[\alpha + \delta S + \gamma A + \nu]E[Z]
\end{aligned}
$$

How can IV be used to obtain consistent estimates?

$$
\begin{aligned}
Cov(Y, Z) &= Cov(\alpha + \delta S + \gamma A + \nu, Z) \\
&= E[(\alpha + \delta S + \gamma A + \nu)Z] - E[\alpha + \delta S + \gamma A + \nu]E[Z] \\
&= \{\alpha E(Z) - \alpha E(Z)\} + \delta\{E(SZ) - E(S)E(Z)\}
\end{aligned}
$$

How can IV be used to obtain consistent estimates?

$$
\begin{aligned}
Cov(Y, Z) &= Cov(\alpha + \delta S + \gamma A + \nu, Z) \\
&= E[(\alpha + \delta S + \gamma A + \nu)Z] - E[\alpha + \delta S + \gamma A + \nu]E[Z] \\
&= \{\alpha E(Z) - \alpha E(Z)\} + \delta\{E(SZ) - E(S)E(Z)\} \\
&\quad +\gamma\{E(AZ) - E(A)E(Z)\} + E(\nu Z) - E(\nu)E(Z)
\end{aligned}
$$

How can IV be used to obtain consistent estimates?

$$
\begin{aligned}
Cov(Y, Z) &= Cov(\alpha + \delta S + \gamma A + \nu, Z) \\
&= E[(\alpha + \delta S + \gamma A + \nu)Z] - E[\alpha + \delta S + \gamma A + \nu]E[Z] \\
&= \{\alpha E(Z) - \alpha E(Z)\} + \delta\{E(SZ) - E(S)E(Z)\} \\
&\quad +\gamma\{E(AZ) - E(A)E(Z)\} + E(\nu Z) - E(\nu)E(Z) \\
Cov(Y, Z) &= \delta Cov(S, Z) + \gamma Cov(A, Z) + Cov(\nu, Z)
\end{aligned}
$$

How can IV be used to obtain consistent estimates?

$$
\begin{aligned}
Cov(Y, Z) &= Cov(\alpha + \delta S + \gamma A + \nu, Z) \\
&= E[(\alpha + \delta S + \gamma A + \nu)Z] - E[\alpha + \delta S + \gamma A + \nu]E[Z] \\
&= \{\alpha E(Z) - \alpha E(Z)\} + \delta\{E(SZ) - E(S)E(Z)\} \\
&\quad + \gamma\{E(AZ) - E(A)E(Z)\} + E(\nu Z) - E(\nu)E(Z) \\
Cov(Y, Z) &= \delta Cov(S, Z) + \gamma Cov(A, Z) + Cov(\nu, Z)
\end{aligned}
$$

- $Cov(S, Z) \neq 0$: Instrument is **relevant** or "first stage" exists. S and Z are correlated

How can IV be used to obtain consistent estimates?

$$
\begin{aligned}
Cov(Y,Z) &= Cov(\alpha + \delta S + \gamma A + \nu, Z) \\
&= E[(\alpha + \delta S + \gamma A + \nu)Z] - E[\alpha + \delta S + \gamma A + \nu]E[Z] \\
&= \{\alpha E(Z) - \alpha E(Z)\} + \delta\{E(SZ) - E(S)E(Z)\} \\
&\quad +\gamma\{E(AZ) - E(A)E(Z)\} + E(\nu Z) - E(\nu)E(Z) \\
Cov(Y,Z) &= \delta Cov(S,Z) + \gamma Cov(A,Z) + Cov(\nu,Z)
\end{aligned}
$$

- $Cov(S,Z) \neq 0$: Instrument is **relevant** or "first stage" exists. S and Z are correlated

$$
\frac{Cov(Y,Z)}{Cov(S,Z)} = \delta + \gamma \frac{Cov(A,Z)}{Cov(S,Z)} + \frac{Cov(\nu,Z)}{Cov(S,Z)}
$$

How can IV be used to obtain consistent estimates?

$$\frac{Cov(Y,Z)}{Cov(S,Z)} = \delta + \gamma \frac{Cov(A,Z)}{Cov(S,Z)} + \frac{Cov(\nu,Z)}{Cov(S,Z)}$$

How can IV be used to obtain consistent estimates?

$$\frac{Cov(Y,Z)}{Cov(S,Z)} = \delta + \gamma\frac{Cov(A,Z)}{Cov(S,Z)} + \frac{Cov(\nu,Z)}{Cov(S,Z)}$$

- $Cov(A,Z) = Cov(\nu,Z) = 0$: Instrument is **valid** or "exclusion restriction" (Z only affects Y through S)

$$\frac{Cov(Y,Z)}{Cov(S,Z)} = \delta + \underbrace{\gamma\frac{Cov(A,Z)}{Cov(S,Z)} + \frac{Cov(\nu,Z)}{Cov(S,Z)}}_{0}$$

$$\frac{Cov(Y,Z)}{Cov(S,Z)} = \delta$$

Instrumental variables

Basic idea

Two stage least squares

Weak instruments

Practical IV Tips

Example

Heterogeneity and the LATE

Imperfect Compliance

Re-cap

Two-stage least squares

- The two-stage least squares estimator was developed by Theil (1953) and Basman (1957) independently

- While IV is a research design, 2SLS is a specific estimator.

- Others include LIML, the Wald estimator, jacknive IV, two sample IV, and more

Two-stage least squares concepts

- Causal model. Sometimes called the structural model:

$$Y_i = \alpha + \delta S_i + \eta_i$$

- First-stage regression. Gets the name because of two-stage least squares:

$$S_i = \gamma + \rho Z_i + \zeta_i$$

- Second-stage regression. Notice the fitted values, $\widehat{S}$:

$$Y_i = \beta + \delta \widehat{S}_i + \nu_i$$

- Reduced form a regression of Y onto the instrument:

$$Y_i = \psi + \pi Z_i + \varepsilon_i$$

where $Cov(Z, \eta_i) = 0$ (instrument is valid) and $\rho \neq 0$ (instrument is relevant).

Two-stage least squares: Estimator I

$$
\begin{aligned}
\delta_{IV} &= \frac{Cov(Y, Z)}{Cov(S, Z)} \\
&= \frac{\frac{Cov(Z,Y)}{Var(Z)}}{\frac{Cov(Z,S)}{Var(Z)}}
\end{aligned}
$$

Two-stage least squares: Estimator I

$$\begin{aligned}\delta_{IV} &= \frac{Cov(Y,Z)}{Cov(S,Z)} \\ &= \frac{\frac{Cov(Z,Y)}{Var(Z)}}{\frac{Cov(Z,S)}{Var(Z)}}\end{aligned}$$

Calculate the ratio of "reduced form" (π) to "first stage" coefficient (ρ):

$$\widehat{\delta}_{2sls} = \frac{\frac{Cov(Z,Y)}{Var(Z)}}{\frac{Cov(Z,S)}{Var(Z)}} = \frac{\widehat{\pi}}{\widehat{\rho}}$$

Two-stage least squares: Estimator II

Recall

$$S_i = \gamma + \rho Z_i + \zeta_i$$

Then

$$\widehat{S} = \widehat{\gamma} + \widehat{\rho} Z$$

$$\widehat{\delta}_{2sls} = \frac{Cov(\widehat{S}, Y)}{Var(\widehat{S})}$$

Two-stage least squares: Estimator II

Recall

$$S_i = \gamma + \rho Z_i + \zeta_i$$

Then

$$\widehat{S} = \widehat{\gamma} + \widehat{\rho} Z$$

$$\begin{aligned} \widehat{\delta}_{2sls} &= \frac{Cov(\widehat{S}, Y)}{Var(\widehat{S})} \\ &= \frac{Cov(\widehat{\gamma} + \widehat{\rho} Z, Y)}{Var(\widehat{\gamma} + \widehat{\rho} Z)} \end{aligned}$$

Two-stage least squares: Estimator II

Recall

$$S_i = \gamma + \rho Z_i + \zeta_i$$

Then

$$\widehat{S} = \widehat{\gamma} + \widehat{\rho} Z$$

$$\begin{aligned}
\widehat{\delta}_{2sls} &= \frac{Cov(\widehat{S}, Y)}{Var(\widehat{S})} \\
&= \frac{Cov(\widehat{\gamma} + \widehat{\rho} Z, Y)}{Var(\widehat{\gamma} + \widehat{\rho} Z)} \\
&= \frac{\widehat{\rho} Cov(Z, Y)}{\widehat{\rho}^2 Var(Z)}
\end{aligned}$$

Two-stage least squares: Estimator II

Recall

$$S_i = \gamma + \rho Z_i + \zeta_i$$

Then

$$\widehat{S} = \widehat{\gamma} + \widehat{\rho} Z$$

$$\begin{aligned}
\widehat{\delta}_{2sls} &= \frac{Cov(\widehat{S}, Y)}{Var(\widehat{S})} \\
&= \frac{Cov(\widehat{\gamma} + \widehat{\rho} Z, Y)}{Var(\widehat{\gamma} + \widehat{\rho} Z)} \\
&= \frac{\widehat{\rho} Cov(Z, Y)}{\widehat{\rho}^2 Var(Z)} \\
&= \frac{Cov(Z, Y)}{\widehat{\rho} Var(Z)}
\end{aligned}$$

Two-stage least squares: Estimator II

$$\widehat{\delta}_{2sls} = \frac{Cov(Z, Y)}{\widehat{\rho} Var(Z)}$$

Two-stage least squares: Estimator II

$$\widehat{\delta}_{2sls} = \frac{Cov(Z,Y)}{\widehat{\rho} Var(Z)}$$

Rewrite $\widehat{\rho}$ as

$$\widehat{\rho} = \frac{Cov(Z,S)}{Var(Z)}$$

Then

$$\begin{aligned}\widehat{\delta}_{2sls} &= \frac{Cov(Z,Y)}{\frac{Cov(Z,S)}{Var(Z)} Var(Z)} \\ &= \frac{\frac{Cov(Z,Y)}{Var(Z)}}{\frac{Cov(Z,S)}{Var(Z)}}\end{aligned}$$

Intuition of 2SLS

- Two stage least squares is nice because in addition to being an estimator, there's also great intuition contained in it which you can use as a device for thinking about IV more generally.

- The intuition is that 2SLS estimator replaces S with the fitted values of S (i.e., $\widehat{S}$) from the first stage regression of S onto Z and all other covariates.

- By using the fitted values of the endogenous regressor from the first stage regression, our regression now uses *only* the exogenous variation in the regressor due to the instrumental variable itself

Intuition of IV in 2SLS

- . . . but think about it – that variation was there before, but was just a subset of all the variation in the regressor
- Endogenous variable has pieces that are as good as random, and IV finds them
- Instrumental variables therefore reduces the variation in the data, but that variation which is left is *exogenous*

Estimation with software

- One manual way is just to estimate the reduced form and first stage coefficients and take the ratio of the respective coefficients on Z
  - While it is always a good idea to run these two regressions, don't compute your IV estimate this way
  - Often the case that a pattern of missing data will differ between Y and S
  - What is the standard error of δ_{IV} in this case?
- Another is to find the fitted values values of S and run the regression with them
  - While it is always a good idea to run this regression too (the first stage), don't compute your IV estimate this way
  - The standard errors from the second stage regression are also wrong

Estimation with software

- Estimate this in Stata using ivregress 2sls.

- Estimate this in R using ivreg in the AER package or using felm in the lfe package

Instrumental variables

Basic idea

Two stage least squares

Weak instruments

Practical IV Tips

Example

Heterogeneity and the LATE

Imperfect Compliance

Re-cap

Instrumental variables

Weak instruments

- A weak instrument is one that is not strongly correlated with the endogenous variable in the first stage
- This can happen if the two variables are independent or the sample is small
- If you have a weak instrument, the cure ends up being worse than the disease

Weak instruments

Back to our causal model

$$\delta_{iv} = \frac{Cov(Y, Z)}{Cov(S, Z)} = \delta + \gamma \frac{Cov(A, Z)}{Cov(S, Z)} + \frac{Cov(\nu, Z)}{Cov(S, Z)}$$

If $Cov(S, Z)$ is small, then it "blows-up" the $\gamma(Cov(A, Z) + Cov(\eta, Z))$ term

Instrumental variables

Basic idea

Two stage least squares

Weak instruments

Practical IV Tips

Example

Heterogeneity and the LATE

Imperfect Compliance

Re-cap

Look at the reduced form

1. Look at the reduced form

   - The reduced form is estimated with OLS and is therefore unbiased

   - If you can't see the causal relationship of interest in the reduced form, it is probably not there

Report the first stage

2. Report the first stage (preferably in the same table as your main results)
    - Does it make sense?
    - Do the coefficients have the right magnitude and sign?
    - Please make beautiful IV tables – you'll be celebrated across the land if you do
3. Report OLS – you said it was biased, but we want to still see it

Instrumental variables

Basic idea

Two stage least squares

Weak instruments

Practical IV Tips

Example

Heterogeneity and the LATE

Imperfect Compliance

Re-cap

Instrumental variables

Example

Time Use and Labor Productivity: The Returns to Sleep

Matthew Gibson and Jeffrey Shrader

Posted Online December 19, 2018
https://doi.org/10.1162/rest_a_00746

Review of Economics and Statistics
Volume 100 | Issue 5 | December 2018
p.783-798

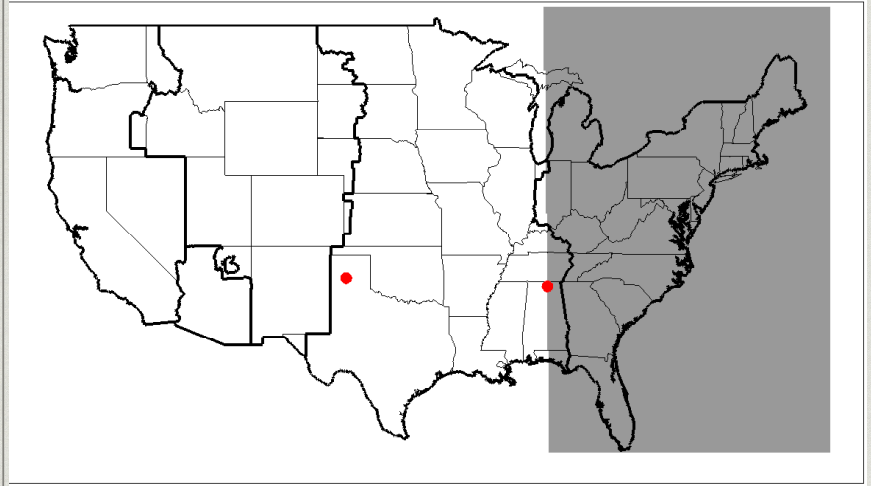
Question

How does sleep affect productivity and wages?

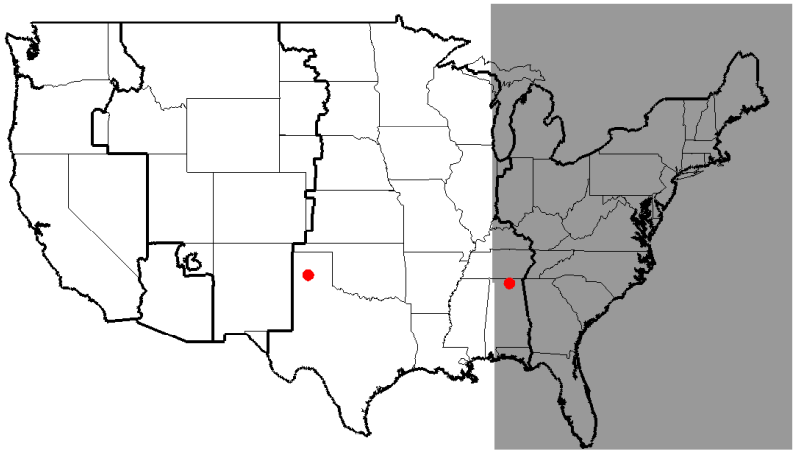
Sunset time instrument

- We require exogenous variation in sleep
- Claims:
  - Earlier local sunset time causes longer sleep
  - Does not co-vary with unobserved determinants of wages

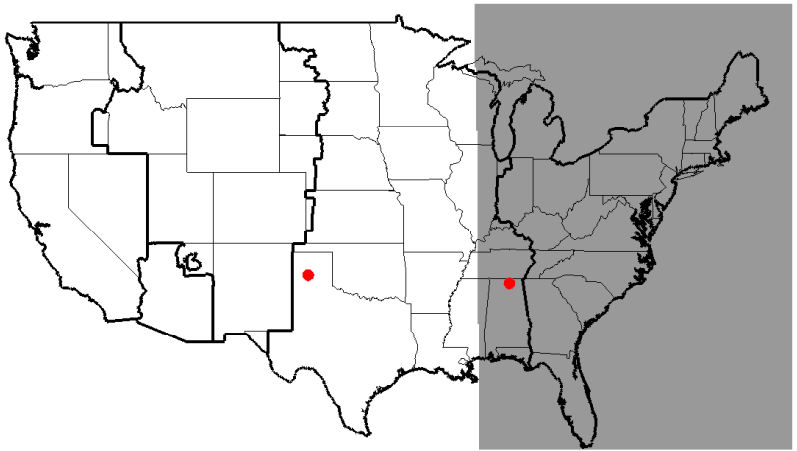
Sep-23 23:44



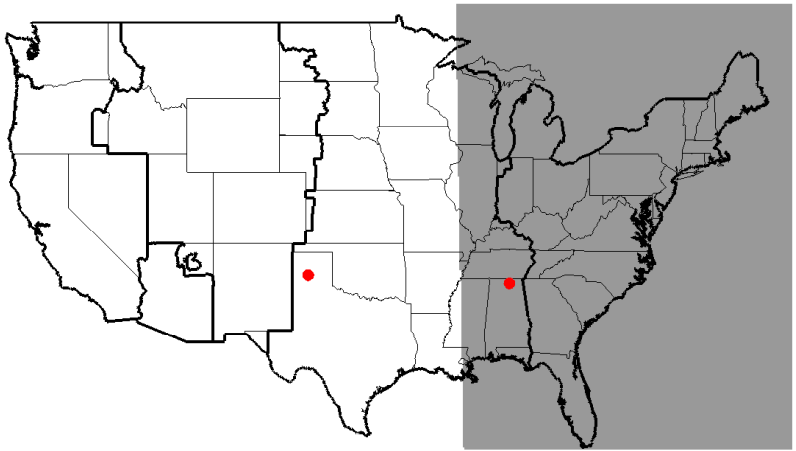
Sep-23 23:49



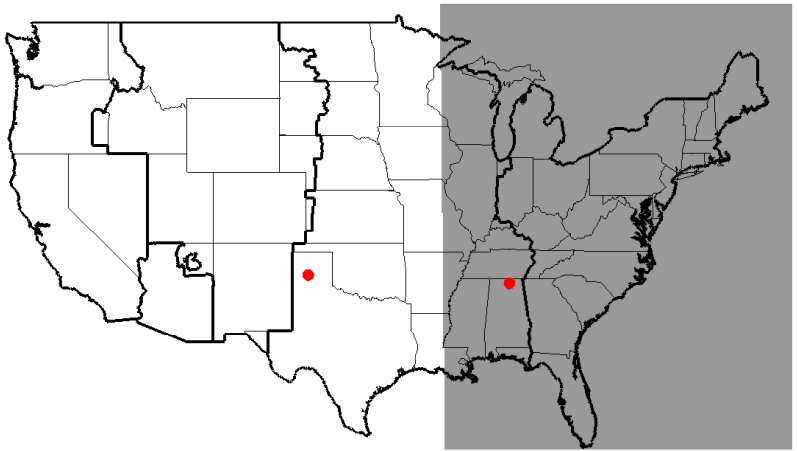
Sep-23 23:54



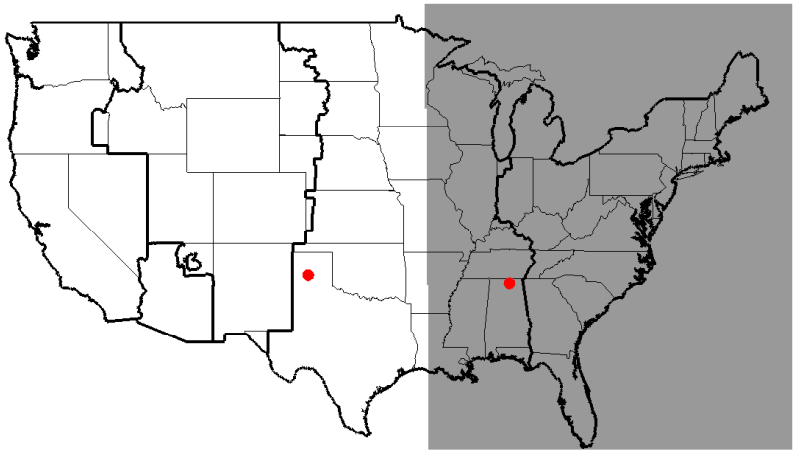
Sep-23 23:59



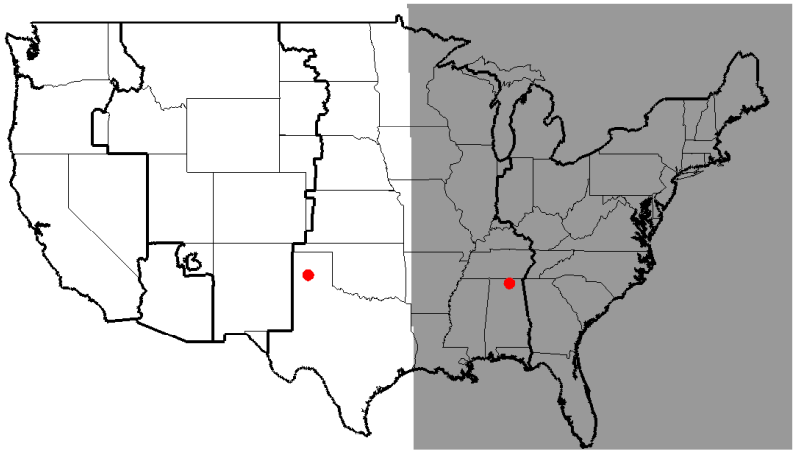
Sep-24 00:04



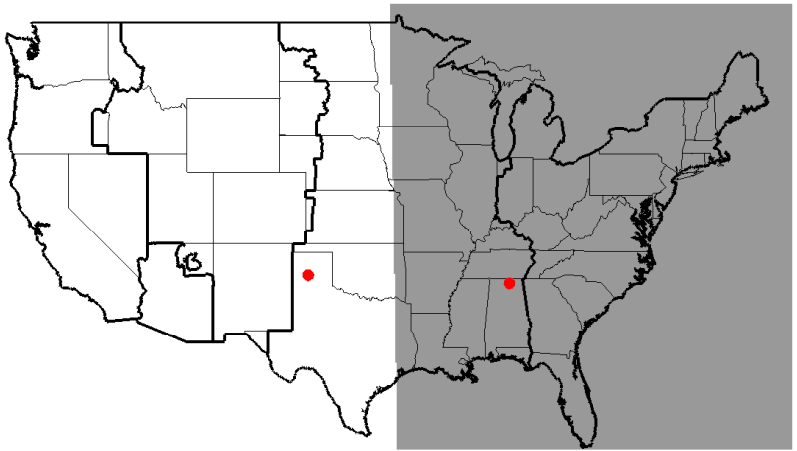
Sep-24 00:09



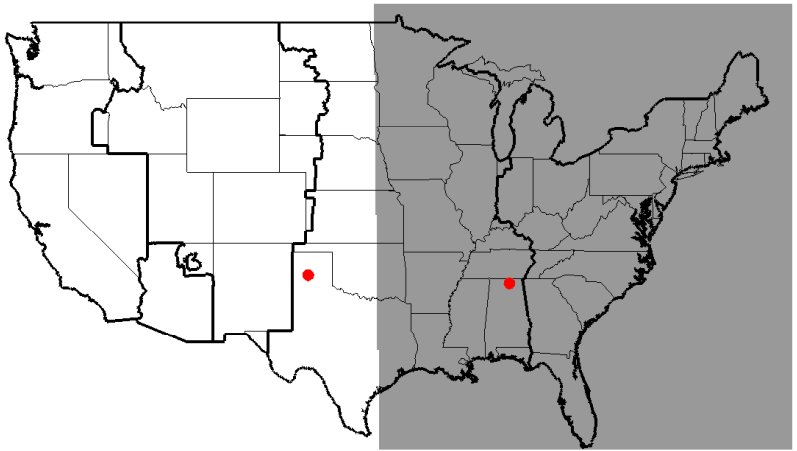
Sep-24 00:14



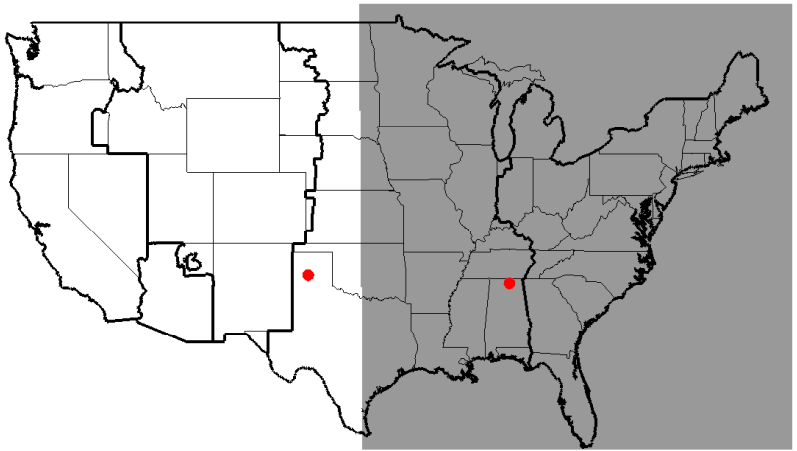
Sep-24 00:19



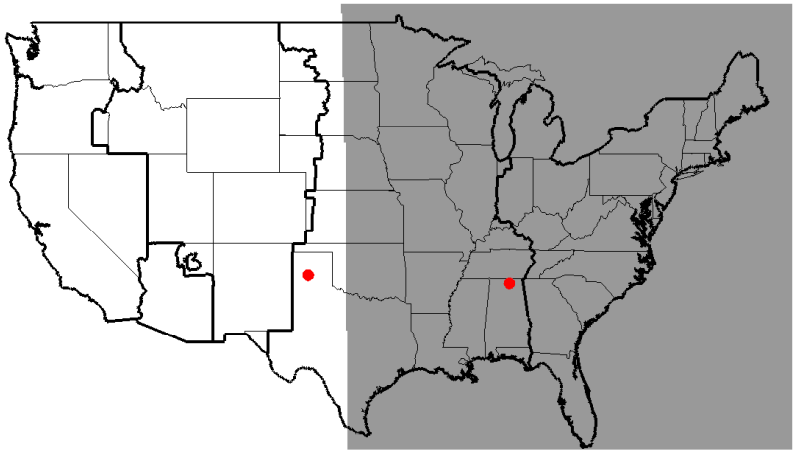
Sep-24 00:24



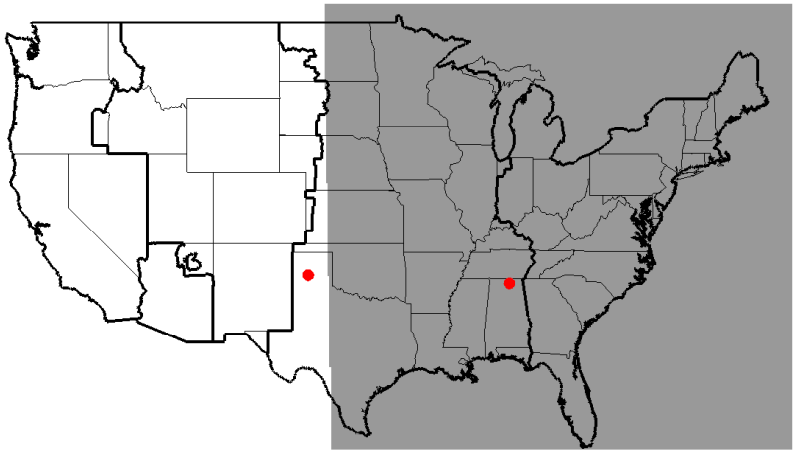
Sep-24 00:29



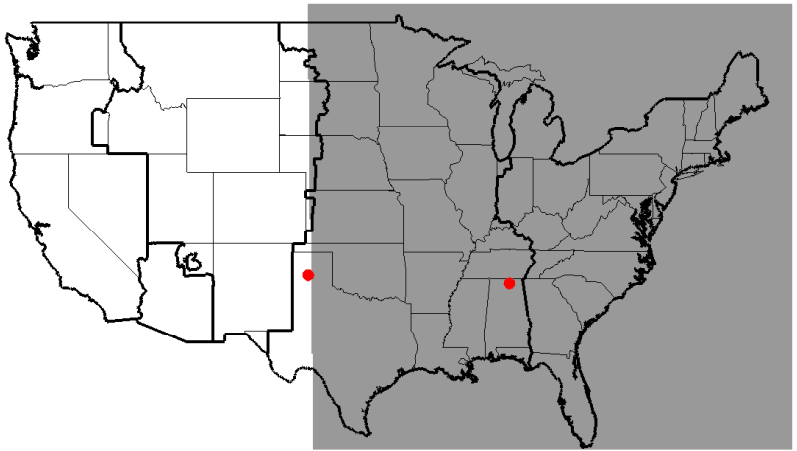
Sep-24 00:34



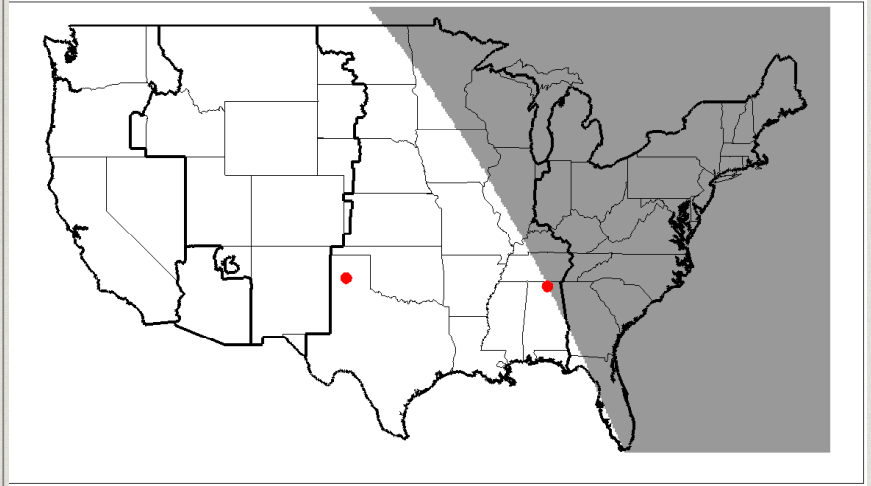
Sep-24 00:39



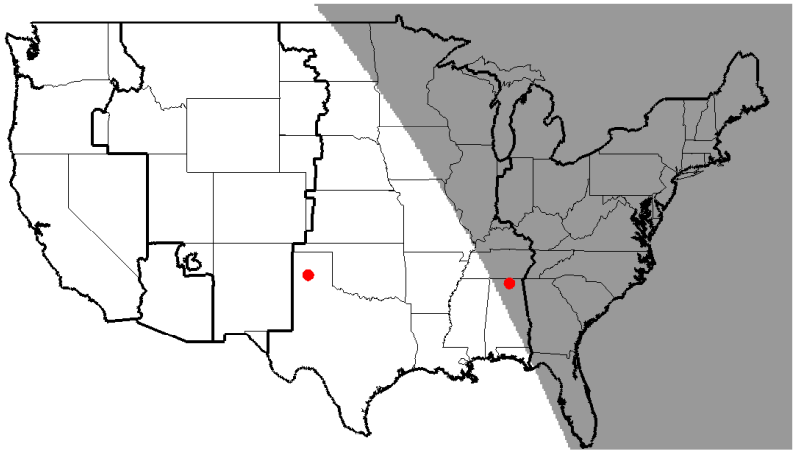
Sep-24 00:44



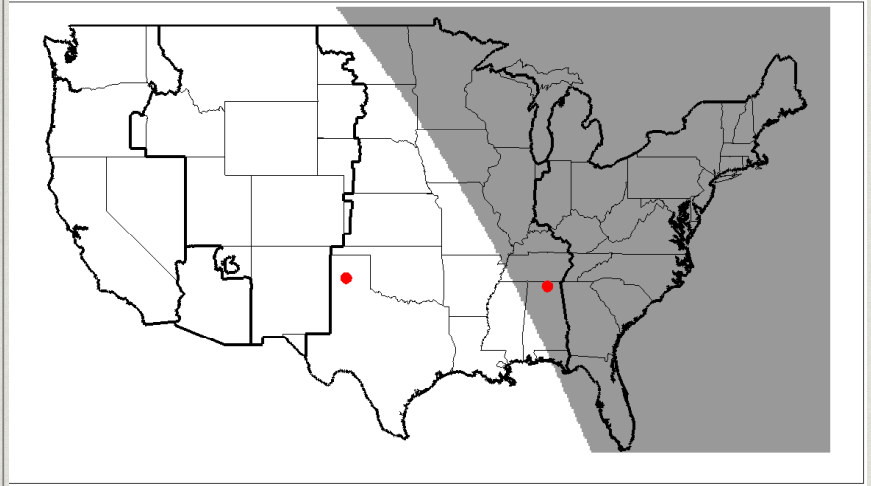
Dec-21 22:36



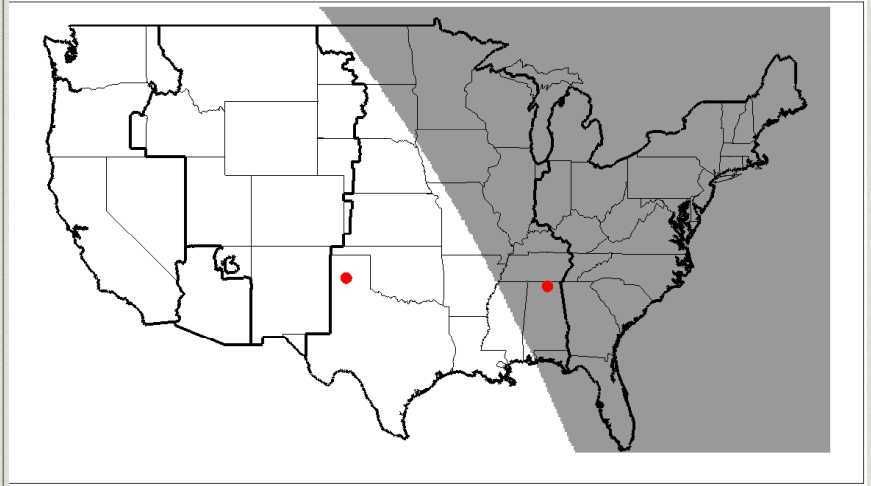
Dec-21 22:41



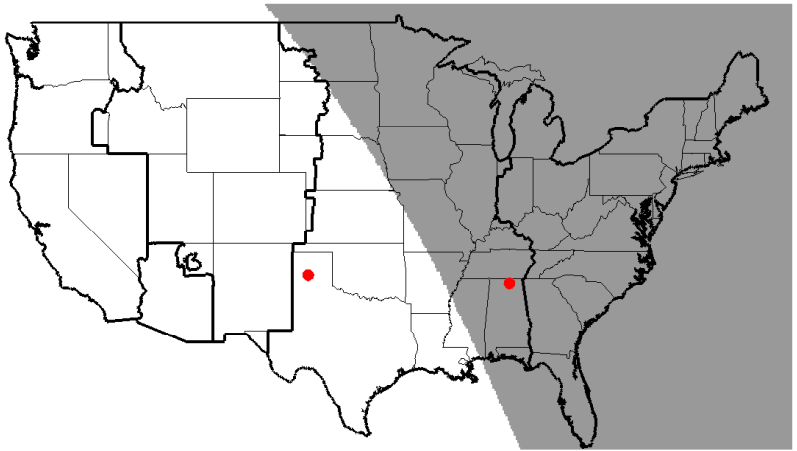
Dec-21 22:46



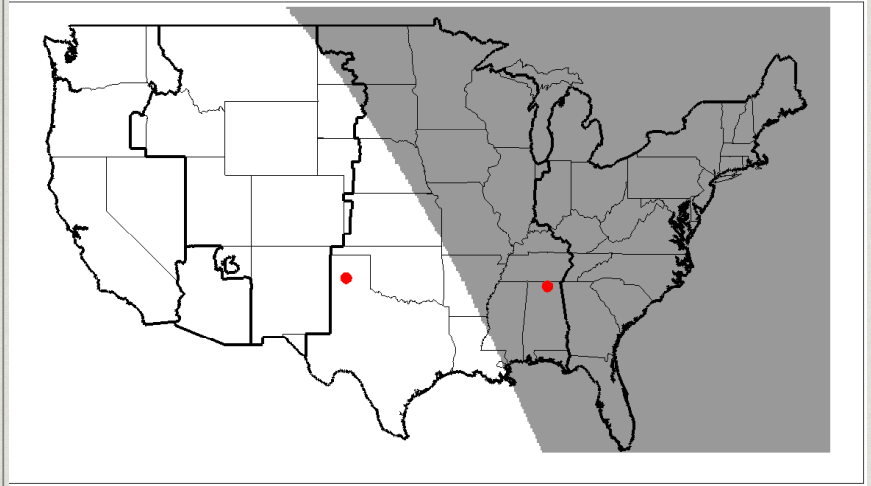
Dec-21 22:51



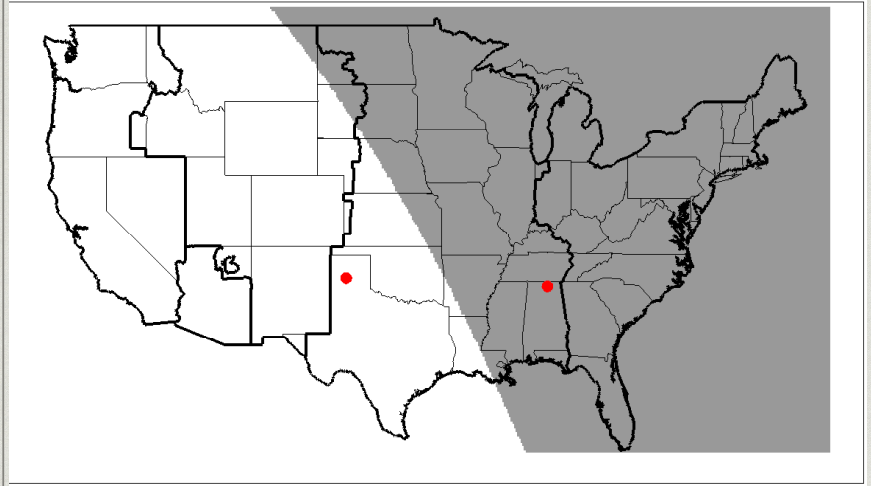
Dec-21 22:56



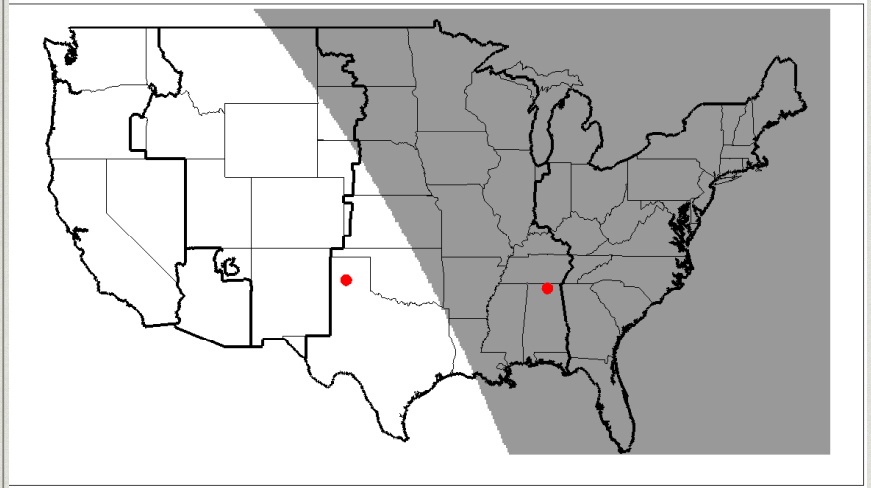
Dec-21 23:01



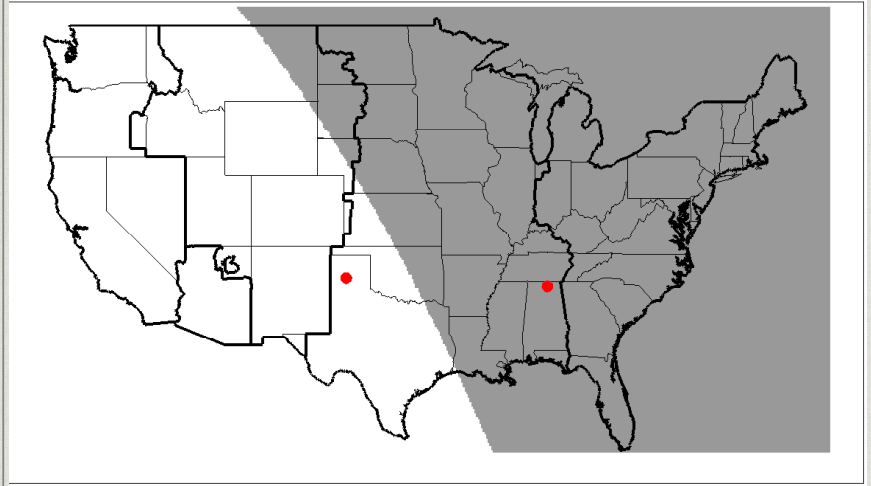
Dec-21 23:06



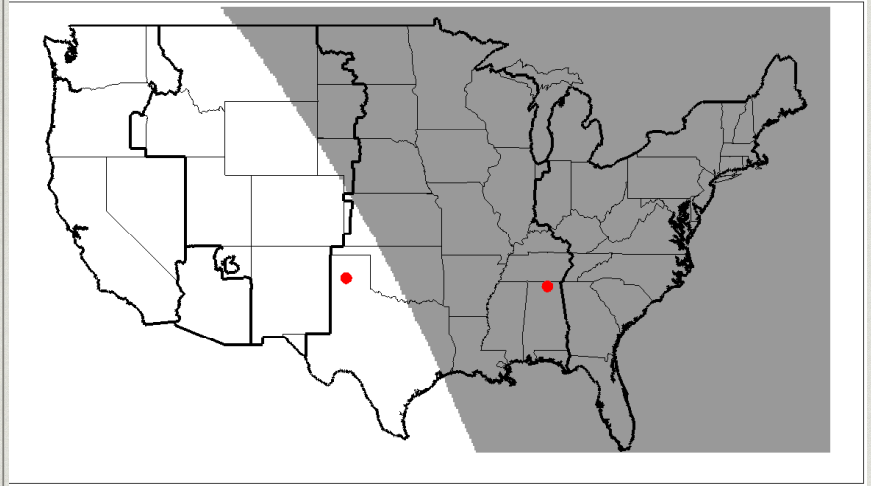
Dec-21 23:11



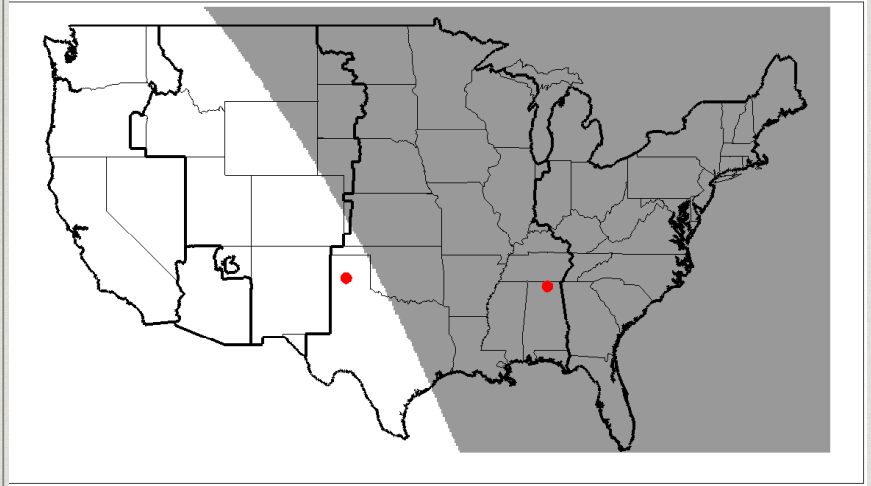
Dec-21 23:16



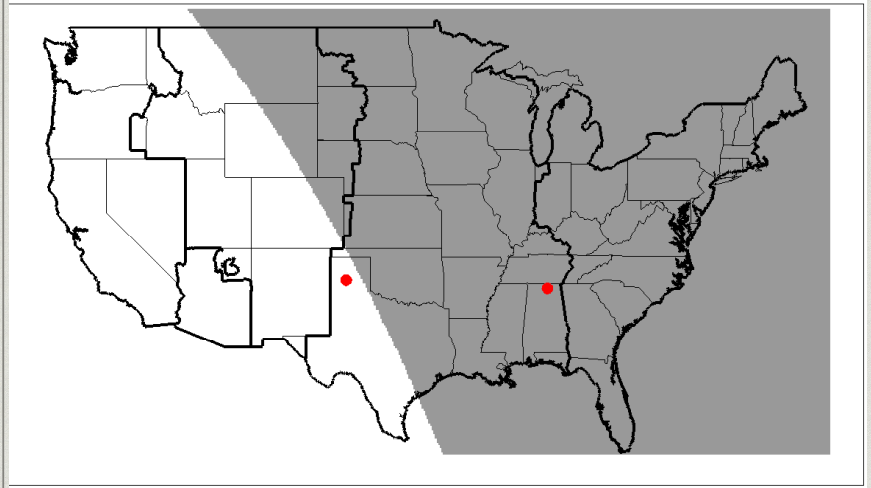
Dec-21 23:21



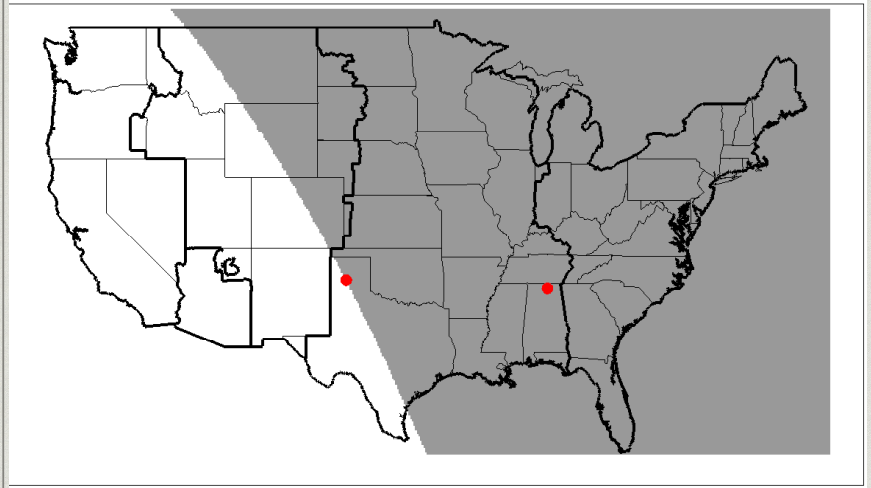
Dec-21 23:26



Dec-21 23:31

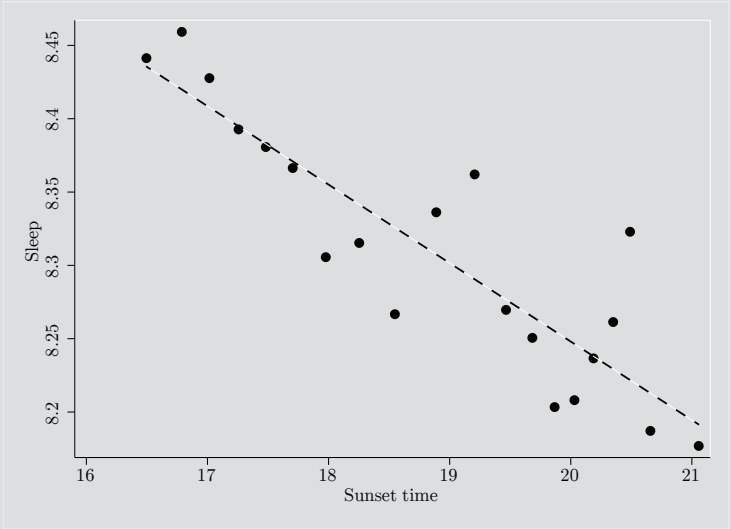


Dec-21 23:36



Sunset time instrument: relevance

- Earlier sunset leads to earlier bedtimes (Roenneberg et al 2007)
- Work start times do not respond to later sunrise (Hamermesh et al 2008)
- Therefore earlier sunset increases sleep duration
  - Maximum US sunset time difference ~1 hr; induces ~35 min weekly sleep difference



RAW SLEEP AND SUNSET TIME

SOURCE: ATUS

Sunset time instrument: validity

- Uncorrelated with daylight duration
- US time zones first implemented 1883
- Designed around scientific concerns

Sunset time instrument: validity

- Optimal sorting would vary seasonally
- No sorting incentives (more on this later)
- No observed sorting (more on this later)

Table III: Linear ATUS Estimates

	First stage Sleep	Reduced form ln(earnings)	2SLS ln(earnings)	OLS ln(earnings)
Sunset time	-24.1***	-0.0085***		
	(2.39)	(0.0019)		
Sleep			0.00035***	-0.000041***
			(0.000085)	(0.0000031)
Individual controls	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Time controls	Yes	Yes	Yes	Yes
Occupation	Yes	Yes	Yes	Yes
Observations	71947	71947	71947	71947
Adjusted R^2	0.123	0.410	0.284	0.411
F-stat on IV	101.83			

LINK TO HOURLY WORKERS

Instrumental variables

Basic idea

Two stage least squares

Weak instruments

Practical IV Tips

Example

Heterogeneity and the LATE

Imperfect Compliance

Re-cap

Heterogenous Treatment Effects

- Up to this point, the causal effect was the same for all individuals
  - Constant treatment effects where $Y_i^1 - Y_i^0 = \delta$ for all i units
- Let's now try to understand what instrumental variables estimation is measuring if treatment effects are *heterogenous*
  - $Y_i^1 - Y_i^0 = \delta_i$ which varies across the population

Why do we care about heterogeneity?

- Heterogeneity, it turns out, makes life interesting and challenging
- What parameter did we even estimate using IV when there are heterogenous treatment effects?
- There are two issues here:
    1. Internal validity: Does the design successfully uncover causal effects for the population that we are studying?
    2. External validity: Does the study's results inform us about different populations?

Potential outcome notation

"Potential treatment status" (D^j) versus "observed" treatment status (D)

- $D_i^1 = i$'s treatment status when $Z_i = 1$
- $D_i^0 = i$'s treatment status when $Z_i = 0$

We'll represent outcomes as a function of both treatment status and instrument status. In other words, $Y_i(D_i = 0, Z_i = 1)$ is represented as $Y_i(0, 1)$

Switching equation

Move from potential treatment status to observed treatment status

$$\begin{aligned} D_i &= D_i^0 + (D_i^1 - D_i^0)Z_i \\ &= \pi_{0i} + \pi_{1i}Z_i + \zeta_i \end{aligned}$$

$$\begin{aligned} \pi_{0i} &= E[D_i^0] \\ \pi_{1i} &= (D_i^1 - D_i^0)\text{is the heterogenous causal effect of the IV on } D_i. \\ E[\pi_{1i}] &= \text{The average causal effect of } Z_i \text{ on } D_i \end{aligned}$$

Identifying assumptions under heterogenous treatment effects

1. Stable Unit Treatment Value Assumption (SUTVA)
2. Random Assignment
3. Exclusion Restriction
4. Nonzero First Stage
5. Monotonicity

Stable Unit Treatment Value Assumption (SUTVA)

Stable Unit Treatment Value Assumption (SUTVA)
If $Z_i = Z_i'$, then $D_i(\mathbf{Z}) = D_i(\mathbf{Z}')$
If $Z_i = Z_i'$ and $D_i = D_i'$, then $Y_i(\mathbf{D},\mathbf{Z}) = Y_i(\mathbf{D'},\mathbf{Z'})$

- Potential outcomes for each person i are unrelated to the treatment status of other individuals

- The instrument must not be related to treatment status of other individuals

Independence assumption

Independence assumption (e.g., "as good as random assignment")
$\{Y_i(D_i^1, 1), Y_i(D_i^0, 0), D_i^1, D_i^0\} \perp\!\!\!\perp Z_i$

- The IV is independent of the vector of potential outcomes and potential treatment assignments (i.e. "as good as randomly assigned")

- It's all about the *randomness* of the instrument, in other words, not the instrument's effect

Independence

Independence means that the first stage measures the causal effect of Z_i on D_i:

$$\begin{aligned} E[D_i|Z_i = 1] - E[D_i|Z_i = 0] &= E[D_i^1|Z_i = 1] - E[D_i^0|Z_i = 0] \\ &= E[D_i^1 - D_i^0] \end{aligned}$$

Independence

The independence assumption is sufficient for a causal interpretation of the reduced form:

$$\begin{aligned} E[Y_i|Z_i=1]-E[Y_i|Z_i=0] &= E[Y_i(D_i^1,1)|Z_i=1] \\ &\quad -E[Y_i(D_i^0,0)|Z_i=0] \\ &= E[Y_i(D_i^1,1)]-E[Y_i(D_i^0,0)] \end{aligned}$$

Exclusion Restriction

Exclusion Restriction
Y(D,Z) = Y(D,Z') for all **Z**, **Z'**, and for all **D**

- Any effect of Z on Y must be via the effect of Z on D. In other words, $Y_i(D_i, Z_i)$ is a function of D only. Or formally:

$$Y_i(D_i, 0) = Y_i(D_i, 1) \text{ for } D = 0, 1$$

- Sometimes called the "only through" assumption because you're assuming the effect of Z on Y is "only through" its effect on D.

Exclusion restriction

- Use the exclusion restriction to define potential outcomes indexed only by the treatment status:

$$\begin{aligned} Y_i^1 &= Y_i(1,1) = Y_i(1,0) \\ Y_i^0 &= Y_i(0,1) = Y_i(0,0) \end{aligned}$$

- Rewrite the switching equation:

$$\begin{aligned} Y_i &= Y_i(0,Z_i) + [Y_i(1,Z_i) - Y_i(0,Z_i)]D_i \\ Y_i &= Y_i^0 + [Y_i^1 - Y_i^0]D_i \end{aligned}$$

- Random coefficients notation for this is:

$$Y_i = \alpha_0 + \delta_i D_i$$
with $\alpha_0 = E[Y_i^0]$ and $\delta_i = Y_i^1 - Y_i^0$

Spotting violations of exclusion is a sport

Watch the gears turn:

- We are interested in causal effect of military service on earnings, and so use draft number are instrument for military service.
- Draft number is generated by a random number generator. Therefore independence is met as draft number is independent of potential outcomes and potential treatment status.
- But, people with higher draft numbers evade draft by investing in schooling. Earnings change for reasons other than military service. Exclusion is violated
- In other words, random lottery numbers (independence) do not imply that the exclusion restriction is satisfied

Strong first stage

Nonzero Average Causal Effect of Z on D
$E[D_i^1 - D_i^0] \neq 0$

- D^1 means instrument is turned on, and D^0 means it is turned off. We need treatment to change when instrument changes.
- Z has to have some statistically significant effect on the average probability of treatment
- First two children of the same gender makes you more likely to have a third.
- Finally – a testable assumption. We have data on Z and D

Monotonicity

Monotonicity
Either $\pi_{1i} \geq 0$ for all i or $\pi_{1i} \leq 0$ for all $i = 1, \ldots, N$

- Recall that π_{1i} is the reduced form causal effect of the instrumental variable on an individual i's treatment status.
- Monotonicity requires that the instrumental variable (weakly) operate in the same direction on all individual units.
- In other words, while the instrument may have no effect on some people, all those who are affected are affected *in the same direction* (i.e., positively or negatively, but not both).
- Without monotonicity, IV estimators are not guaranteed to estimate a weighted average of the underlying causal effects of the affected group, $Y_i^1 - Y_i^0$.

Force yourself to think of monotonicity violations

- In the quarter of birth example for schooling, this assumption may not be satisfied

- Being born in the 4^{th} quarter (which typically increases schooling) may have reduced schooling for some because their school enrollment was held back by their parents

Local average treatment effect

If all 1-5 assumptions are satisfied, then IV estimates the **local average treatment effect (LATE)** of D on Y:

$$\delta_{IV,LATE} = \frac{\text{Effect of Z on Y}}{\text{Effect of Z on D}}$$

Estimand

Instrumental variables (IV) estimand:

$$\begin{aligned}\delta_{IV,LATE} &= \frac{E[Y_i(D_i^1, 1) - Y_i(D_i^0, 0)]}{E[D_i^1 - D_i^0]} \\ &= E[(Y_i^1 - Y_i^0)|D_i^1 - D_i^0 = 1]\end{aligned}$$

Local Average Treatment Effect

- The LATE parameters is the average causal effect of D on Y for those whose treatment status was changed by the instrument, Z
- For example, IV estimates the average effect of military service on earnings for the subpopulation who enrolled in military service because of the draft but would not have served otherwise.
- LATE does not tell us what the causal effect of military service was for patriots (volunteers) or those who were exempted from military service for medical reasons

LATE cont.

- We have reviewed the properties of IV with heterogenous treatment effects using a very simple dummy endogenous variable, dummy IV, and no additional controls example.

- The intuition of LATE generalizes to most cases where we have continuous endogenous variables and instruments, and additional control variables.

LATE and subpopulations

The instrument partitions any population into 4 distinct groups:

1. Compliers: The subpopulation with $D_i^1 = 1$ and $D_i^0 = 0$. Their treatment status is affected by the instrument in the "correct direction".

2. Always takers: The subpopulation with $D_i^1 = D_i^0 = 1$. They always take the treatment independently of Z.

3. Never takers: The subpopulation with $D_i^1 = D_i^0 = 0$. They never take the treatment independently of Z.

4. Defiers: The subpopulation with $D_i^1 = 0$ and $D_i^0 = 1$. Their treatment status is affected by the instrument in the "wrong direction".

Subpopulations of soldieres

Examples of subpopulations:

1. Compliers: I only enrolled in the military because I was drafted otherwise I wouldn't have served
2. Always takers: My family have always served, so I serve regardless of whether I am drafted
3. Never takers: I'm a contentious objector so under no circumstances will I serve, even if drafted
4. Defiers: When I was drafted, I dodged. But had I not been drafted, I would have served. I can't make up my mind.

Never-Takers

$D_i^1 - D_i^0 = 0$

$Y_i(0,1) - Y_i(0,0) = 0$

By Exclusion Restriction, causal effect of Z on Y is zero.

Defier

$D_i^1 - D_i^0 = -1$

$Y_i(0,1) - Y_i(1,0) = Y_i(0) - Y_i(1)$

By Monotonicity, no one in this group

Complier

$D_i^1 - D_i^0 = 1$

$Y_i(1,1) - Y_i(0,0) = Y_i(1) - Y_i(0)$

Average Treatment Effect among Compliers

Always-taker

$D_i^1 - D_i^0 = 0$

$Y_i(1,1) - Y_i(1,0) = 0$

By Exclusion Restriction, causal effect of Z on Y is zero.

Monotonicity Ensures that there are no defiers

- Why is it important to not have defiers?
    - If there were defiers, effects on compliers could be (partly) canceled out by opposite effects on defiers
    - One could then observe a reduced form which is close to zero even though treatment effects are positive for everyone (but the compliers are pushed in one direction by the instrument and the defiers in the other direction)
- Monotonicity assumes there are no defiers

What Does IV (Not) Estimate?

- As said, with all 5 assumptions satisfied, IV estimates the average treatment effect for *compliers*, or LATE
- Without further assumptions (e.g., constant causal effects), LATE is not informative about effects on never-takers or always-takers because the instrument does not affect their treatment status
- So what? Well, it matters because in most applications, we would be mostly interested in estimating the average treatment effect on the whole population:

$$ATE = E[Y_i^1 - Y_i^0]$$

- But that's not possible usually with IV

Summarizing

- The potential outcomes framework gives a more subtle interpretation of what IV is measuring
    - In the constant coefficients world, IV measures δ which is "the" causal effect of D_i on Y_i, and assumed to be the same for all i units
    - In the random coefficients world, IV measures instead an *average* of heterogeneous causal effects across a particular population – $E[\delta_i]$ for some group of i units
    - IV, therefore, measures the *local average treatment effect* or LATE parameter, which is the average of causal effects across the subpopulation of *compliers*, or those units whose covariate of interest, D_i, is influenced by the instrument.

Instrumental variables

Basic idea

Two stage least squares

Weak instruments

Practical IV Tips

Example

Heterogeneity and the LATE

Imperfect Compliance

Re-cap

Compliance

- Most RCTs examine programs that will not be universally adopted when offered
- **Compliance**: Whether someone takes the treatment when they are offered
  - In the case of a new type of farming practice there will be farmers offered the practice who do not adopt it
  - In the case of a job training program there will be entrepreneurs who choose not to attend the training
  - These are the non-compliers
- Let C_i be the compliance status of individual i
  - If she chooses to accept the program then $C_i = 1$
  - If not then $C_i = 0$
- Two impacts: intention-to-treat (ITT) and treatment-on-the-treated (ToT)

Intention to Treat Effect

- The Intention to Treat Effect

$$ITT = E(Y_{1i}|T_i = 1) - E(Y_{0i}|T_i = 0)$$

- The ITT essentially ignores non-compliance: estimates the effect of 'intending to treat' some units, regardless of how many take up the treatment
- Partially sidesteps compliance issues: Focus on treatment/ignore compliance
- It cannot completely escape compliance because a decreasing compliance rate will push the ITT towards zero

Treatment Effect on the Treated

- The Treatment Effect on the Treated

$$ToT = E(Y_{1i}|T_i = 1; C_i = 1) - E(Y_{0i}|T_i = 0; C_i = 1)$$

- This is the treatment effect on those who actually choose to accept the treatment
- **The counterfactual is those who would have accepted the treatment if they had been offered it**
- Non-compliance drives down the ITT relative to the ToT
- If a program has no spillover effect (i.e., non-compliers in the treatment area receive no indirect effect from the treatment taking place around them), the treatment effect on the non-compliers is 0

Treatment Effect on the Treated

There is a mechanical relationship between the ITT and ToT

$$ITT = E(Y_{1i}|T_i = 1) - E(Y_{0i}|T_i = 0)$$

Treatment Effect on the Treated

There is a mechanical relationship between the ITT and ToT

$$\begin{aligned} ITT &= E(Y_{1i}|T_i = 1) - E(Y_{0i}|T_i = 0) \\ &= cE(Y_{1i}|T_i = 1; C_i = 1) + (1-c)E(Y_{1i}|T_i = 1; C_i = 0) - \\ &\quad cE(Y_{1i}|T_i = 0; C_i = 1) - (1-c)E(Y_{1i}|T_i = 0; C_i = 0) \end{aligned}$$

Treatment Effect on the Treated

There is a mechanical relationship between the ITT and ToT

$$
\begin{aligned}
ITT &= E(Y_{1i}|T_i = 1) - E(Y_{0i}|T_i = 0) \\
&= cE(Y_{1i}|T_i = 1; C_i = 1) + (1-c)E(Y_{1i}|T_i = 1; C_i = 0) - \\
&\quad cE(Y_{1i}|T_i = 0; C_i = 1) - (1-c)E(Y_{1i}|T_i = 0; C_i = 0) \\
&= c\left[E(Y_{1i}|T_i = 1; C_i = 1) - E(Y_{0i}|T_i = 0; C_i = 1)\right] + \\
&\quad (1-c)\underbrace{\left[E(Y_{1i}|T_i = 1; C_i = 0) - E(Y_{1i}|T_i = 0; C_i = 0)\right]}_{\text{zero because of no spillovers (i.e., exclusion restriction)}}
\end{aligned}
$$

Treatment Effect on the Treated

There is a mechanical relationship between the ITT and ToT

$$
\begin{aligned}
ITT &= E(Y_{1i}|T_i = 1) - E(Y_{0i}|T_i = 0) \\
&= cE(Y_{1i}|T_i = 1; C_i = 1) + (1-c)E(Y_{1i}|T_i = 1; C_i = 0) - \\
&\quad cE(Y_{1i}|T_i = 0; C_i = 1) - (1-c)E(Y_{1i}|T_i = 0; C_i = 0) \\
&= c\left[E(Y_{1i}|T_i = 1; C_i = 1) - E(Y_{0i}|T_i = 0; C_i = 1)\right] + \\
&\quad (1-c)\underbrace{\left[E(Y_{1i}|T_i = 1; C_i = 0) - E(Y_{1i}|T_i = 0; C_i = 0)\right]}_{\text{zero because of no spillovers (i.e., exclusion restriction)}} \\
&= ToT * c
\end{aligned}
$$

where c is the compliance rate

Treatment Effect on the Treated

- Compliance is not typically observed in the control group!
- Estimator for the ToT given above cannot be estimated with standard data
- If we are willing to assume that there is no interference with the control group then we can back out the ToT as ITT/c
- Standard empirical way of estimating ToT effects is to instrument actual receipt of treatment with being offered treatment
  - Run a regression with compliance as the endogeneous variable
  - Being in the treatment group is the instrument
- In a regression without any other control variables, this instrumented ToT will be exactly the ITT blown up by the inverse of the compliance rate.

Treatment Effect on the Treated

- Think of the compliance as an endogenous variable with an IV (treatment)

$$\begin{aligned} C_i &= \gamma_0 + \gamma_1 T_i + \mu_i \\ Y_i &= \beta_0 + \beta_1 \widehat{C}_i + \varepsilon \end{aligned}$$

- $\widehat{C}_i = c$ (i.e, the compliance rate in the treatment group)
- Regressing the outcome on the treated yields the ITT
- Thus, $\widehat{\beta}_1 = \frac{ITT}{c} = ToT$

Two-sided non-compliance

- Two-sided non-compliance:
  - Individuals in the control who get treated
  - Individuals in the treatment who do not comply
- In this context it is natural to think of the treatment simply as something that boosts compliance, and not as the actual receipt of the treatment itself
  - Email encouraging people to do something
  - Facebook/Twitter/Google ads with information
- We can of course continue to estimate a kind of ITT in this context (difference between the group offered the "treatment" and the group not offered)
- What we are estimating with the ITT is the **impact of the intervention that changes compliance and not the impact of the treatment itself**
- If the compliance rate in the treatment and control groups is the same, we have no experiment at all!

Local Average Treatment Effects

- The treatment and compliance possibilities define four possible cells (assuming that the purpose of the treatment is to increase compliance):

  1. Always takers: $c_i(T_i = 0) = c_i(T_i = 1) = 1$
  2. Never takers: $c_i(T_i = 0) = c_i(T_i = 1) = 0$
  3. Compliers: $c_i(T_i = 0) = 0$ and $c_i(T_i = 1) = 1$
  4. Defiers: $c_i(T_i = 0) = 1$ and $c_i(T_i = 1) = 0$

Local Average Treatment Effects

- We are not typically interested in the impact of the intervention to boost compliance, but rather the impact of the treatment itself
- We can instrument compliance with offering the treatment: In some ways this is simply a standard implementation of the instrumented TOT
- Abadie and Imbens: Cannot use this instrument to understand the effect of the promotion on "always compliers" nor on of the "defiers"
- In other words, the instrument has no first-stage for groups that were going to comply or not comply in the absence of the promotion
- What we estimate with this technique is the Local Average Treatment Effect
- This is the impact of the actual treatment (rather than the promotion), estimated only upon those types who were induced to comply by the promotion

Local Average Treatment Effects

- The analogy to the estimation of the ToT with one-sided non-compliance would be to inflate the ITT estimated from a Randomized Promotion design by the differential compliance between the treatment and control
- But in order to do this we must make an assumption stronger than the "Non-Interference and Exclusion Restriction" assumptions already laid out
- We must add a Monotonicity assumption known as the "No Defiers" assumption
- The reason we need this assumption is precisely the heterogeneity of impacts
- Under homogeneous impacts and the exclusion restriction, if the fraction of Defiers and Compliers in the sample were equal, we could not have a treatment effect

"No Defiers" assumption

- We set the fraction of defiers to zero: $\pi_D = 0$
- Then
  - Fraction of always takers: $\pi_{AT} = E[C_i(T_i = 0)]$
  - Fraction of never takers: $\pi_{NT} = 1 - E[C_i(T_i = 1)]$
  - Fraction of compliers: $\pi_C = E[C_i(T_i = 1) - C_i(T_i = 0)]$

Local Average Treatment Effects

- We can write the ITT as a weighted average of outcomes in the four cells:

$$
\begin{aligned}
ITT \quad = \quad & \pi_C E[Y_i(T_i = 1; C_i = 1) - Y_i(T_i = 0; C_i = 1)] + \\
& \pi_{AT} \underbrace{E[Y_i(T_i = 1; C_i = 1) - Y_i(T_i = 0; C_i = 1)]}_{\text{zero by the exclusion restriction (no spillovers)}} + \\
& \pi_{NT} \underbrace{E[Y_i(T_i = 1; C_i = 0) - Y_i(T_i = 0; C_i = 0)]}_{\text{zero by the exclusion restriction (no spillovers)}} + \\
& \underbrace{\pi_D E[Y(T_i = 1; C_i = 0) - Y_i(T_i = 1; C_i = 0)]}_{\text{zero by the no defier assumption}}
\end{aligned}
$$

Local Average Treatment Effects

- We can write the ITT as a weighted average of outcomes in the four cells:

$$
\begin{aligned}
ITT &= \pi_C E[Y_i(T_i=1;C_i=1) - Y_i(T_i=0;C_i=1)] + \\
&\quad \pi_{AT} \underbrace{E[Y_i(T_i=1;C_i=1) - Y_i(T_i=0;C_i=1)]}_{\text{zero by the exclusion restriction (no spillovers)}} + \\
&\quad \pi_{NT} \underbrace{E[Y_i(T_i=1;C_i=0) - Y_i(T_i=0;C_i=0)]}_{\text{zero by the exclusion restriction (no spillovers)}} + \\
&\quad \underbrace{\pi_D E[Y(T_i=1;C_i=0) - Y_i(T_i=1;C_i=0)]}_{\text{zero by the no defier assumption}} \\
&= \pi_C E[Y_i(T_i=1;C_i=1) - Y_i(T_i=0;C_i=1)]
\end{aligned}
$$

Local Average Treatment Effects

- Problem: I can tell the fraction of compliers in the population (by comparing uptake rates in treatment relative to the control) but I cannot tell which individuals were induced to comply by the promotion
- This LATE may be interesting (if the promotion is a real policy that is being considered, or is based off of price variation that we will really observe)
- ...Or may be completely artificial (if the promotion induces a group to comply that would never comply in the native implementation of the program)

Local Average Treatment Effects

Imbens: Reporting the local average treatment effect, solely, or in combination with bounds or point estimates for the overall average based on additional assumptions, is thus emphatically not motivated by a claim that the local average treatment effect is the sole or primary effect of interest. Rather, it is motivated by a sober assessment that estimates for other subpopulations do not have the same internal validity, and by an attempt to clarify what can be learned from the data in the absence of identification of the population average effect

When are LATEs what we want to measure?

- A price randomization where the implementer is considering a range of different prices/subsidies and the experiment includes the relevant range (this design still only gives the marginal impact and not the impact on the whole client pool)
- Eligibility randomization 'on the bubble' where the question answered is the impact of extending access on the eligibility margin
- A randomized promotion intervention where a technology is universally available but not widely adopted, so the relevant policy question is the effect of expanding uptake through adoption-enhancing interventions

When are LATEs NOT what we want to measure?

- A price incentive randomization in a context where the product is provided by the private market and the prices in the study will never be observed in reality
- Randomized promotion campaign where the promotion is so expensive that it yields a group of beneficiaries who would never take the product in reality
- LATE or IV based on variation that selects an odd sample that is not representative of the implementation population (e.g., Deaton's earthquakes example)

Instrumental variables

Likely source of OLS bias? Exclusion restriction? First stage?

- Vietnam draft lottery
- Job Training Partnership Act (JTPA) randomized trial
- Ocean weather
- Rainfall
- Slope of land

When is IV used?

Instrumental variables methods are typically used to address the following kinds of problems

1. Omitted variable bias
2. Measurement error
3. Simultaneity bias
4. Reverse causality
5. Randomized control trials with noncompliance

Instrumental variables

- Instrumental variables offers some hope at recovering the causal effect of D on Y
- The best instruments come from deep knowledge of institutional details (Angrist and Krueger 1991)
- Certain types of natural experiments can be the source of such opportunities and may be useful