Week 12: DID

JJ Chen

April 3, 2015

Today's Plan

- Basic DID recap
- An DID example: Nunn and Qian (2011)
- Other DID extensions

Potential Outcome DID

- Two groups of units: $\text{TREAT}_i = 0$, $\text{TREAT}_i = 1$
- Two periods: $\text{POST}_t = 0 = before$, $\text{POST}_t = 1 = after$
- Treatment group receives treatment in post period: $\text{TP}_{it} = \text{TREAT}_i \times \text{POST}_t$
- Two potential outcomes: Y_{0it}, Y_{1it}
  - Observed outcomes: $\text{Y}_{it} = \text{Y}_{0it} \times (1 - \text{TP}_{it}) + \text{Y}_{1it} \times \text{TP}_{it}$
- ATET: $\delta^{ATET} = \text{E}(\text{Y}_{1it} - \text{Y}_{0it} \,|\, \text{TP}_{it} = 1)$
  - $\implies \delta^{ATET} = \text{E}(\text{Y}_{1i,after} - \text{Y}_{0i,after} \,|\, \text{TREAT} = 1)$. Why? So what?
  - $\implies \delta^{ATET} = \text{E}(\text{Y}_{1a} - \text{Y}_{0a} \,|\, \text{T} = 1)$

Potential Outcome DID

- ATET: $\delta^{ATET} = \mathrm{E}(\mathrm{Y}_{1a} - \mathrm{Y}_{0a} \mid \mathrm{T} = 1)$
- DID gives ATET?

$$\begin{aligned}
\delta^{DID} &= \mathrm{E}(\mathrm{Y}_{a} - \mathrm{Y}_{b} \mid \mathrm{T} = 1) - \mathrm{E}(\mathrm{Y}_{a} - \mathrm{Y}_{b} \mid \mathrm{T} = 0) \\
&= \mathrm{E}(\mathrm{Y}_{1a} - \mathrm{Y}_{0b} \mid \mathrm{T} = 1) - \mathrm{E}(\mathrm{Y}_{0a} - \mathrm{Y}_{0b} \mid \mathrm{T} = 0) \\
&= \mathrm{E}(\mathrm{Y}_{1a} - \mathrm{Y}_{0b} \mid \mathrm{T} = 1) - \mathrm{E}(\mathrm{Y}_{0a} - \mathrm{Y}_{0b} \mid \mathrm{T} = 1) \\
&= \mathrm{E}(\mathrm{Y}_{1a} - \mathrm{Y}_{0a} \mid \mathrm{T} = 1) \\
&= \delta^{ATET}
\end{aligned}$$

- ID assumption for ATET:
  $\mathrm{E}(\mathrm{Y}_{0a} - \mathrm{Y}_{0b} \mid \mathrm{T} = 1) = \mathrm{E}(\mathrm{Y}_{0a} - \mathrm{Y}_{0b} \mid \mathrm{T} = 0)$
  - Or: $\mathrm{Y}_{0a} - \mathrm{Y}_{0b} \perp\!\!\!\perp \mathrm{T}$
  - Can be extended to selection-on-observable: $\mathrm{Y}_{0a} - \mathrm{Y}_{0b} \perp\!\!\!\perp \mathrm{T} \mid \mathrm{X}$

Regressions DID

- Simple: $\text{Y}_{it} = \alpha + \beta \text{TREAT}_i + \gamma \text{POST}_t + \delta \text{TP}_{it} + e_{it}$
- Multiple groups and periods:

$$\text{Y}_{it} = \alpha + \sum_{k=1}^{K} \beta_k \text{TREAT}_{ki} + \sum_{j=1}^{J} \gamma_j \text{PERIOD}_{jt} + \delta \text{TP}_{it} + e_{it}$$

  - Or: $\text{Y}_{it} = \alpha + \beta_i + \gamma_t + \delta \text{TP}_{it} + e_{it}$

- Allow group specific trends: $\text{Y}_{it} = \alpha + \sum_{k=1}^{K} \beta_k \text{TREAT}_{ki} +$

$$\sum_{j=1}^{J} \gamma_j \text{PERIOD}_{jt} + \sum_{k=1}^{K} \theta_k (\text{TREAT}_{ki} \times t) + \delta \text{TP}_{it} + e_{it}$$

Introduction

- Nunn and Qian (2011) examine the impact of potato on Old World population and urbanization
- Source of identification:
  - regional variation in suitability for planting potatoes
  - time variation in introduction of potato from Americans
  - a typical DID research design
- What're the differences here?

Empirical Implications

- $$\mathrm{Y}_{it} = \beta(\ln \mathrm{PotatoArea}_i \times \mathrm{I}_t^{Post}) + \sum_{j=1100}^{1900} \mathrm{X}_i' \mathrm{I}_{jt} \boldsymbol{\phi}_j + \sum_c \gamma_c \mathrm{I}_{ci} + \sum_{j=1100}^{1900} \rho_j \mathrm{I}_{jt} + e_{it}$$
  - Time periods include 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1750, 1800, 1850, and 1900
  - There are countries fixed effects and time fixed effects
  - Post periods are 1750, 1800, 1850, and 1900
- $$\mathrm{Y}_{it} = \sum_{j=1100}^{1900} \beta_j(\ln \mathrm{PotatoArea}_i \times \mathrm{I}_{jt}^{Post}) + \sum_{j=1100}^{1900} \mathrm{X}_i' \mathrm{I}_{jt} \boldsymbol{\phi}_j + \sum_c \gamma_c \mathrm{I}_{ci} + \sum_{j=1100}^{1900} \rho_j \mathrm{I}_{jt} + e_{it}$$

Outcome Variable

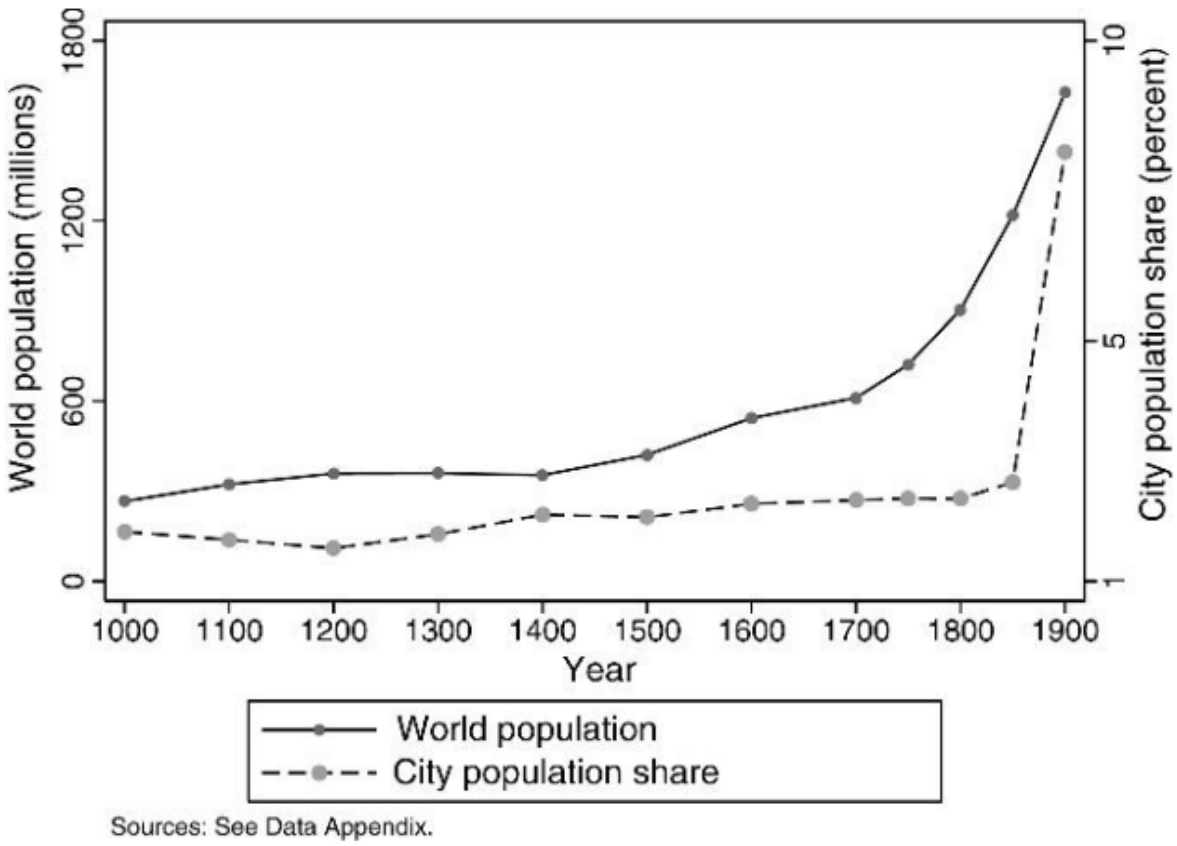


FIGURE I
Growth in World Population and Urbanization, 1000–1900

Variations in Suitatble Potato Area: Map

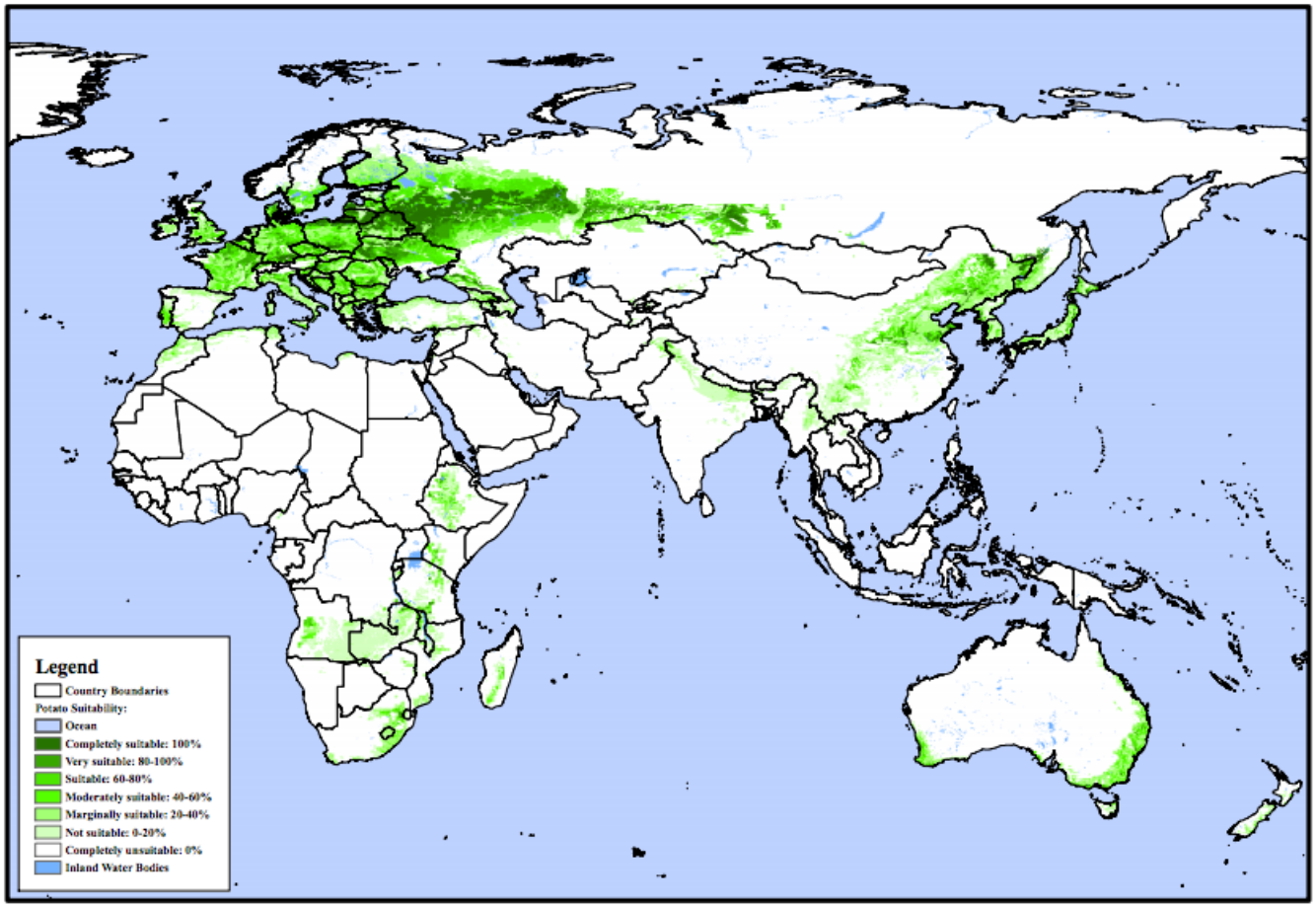


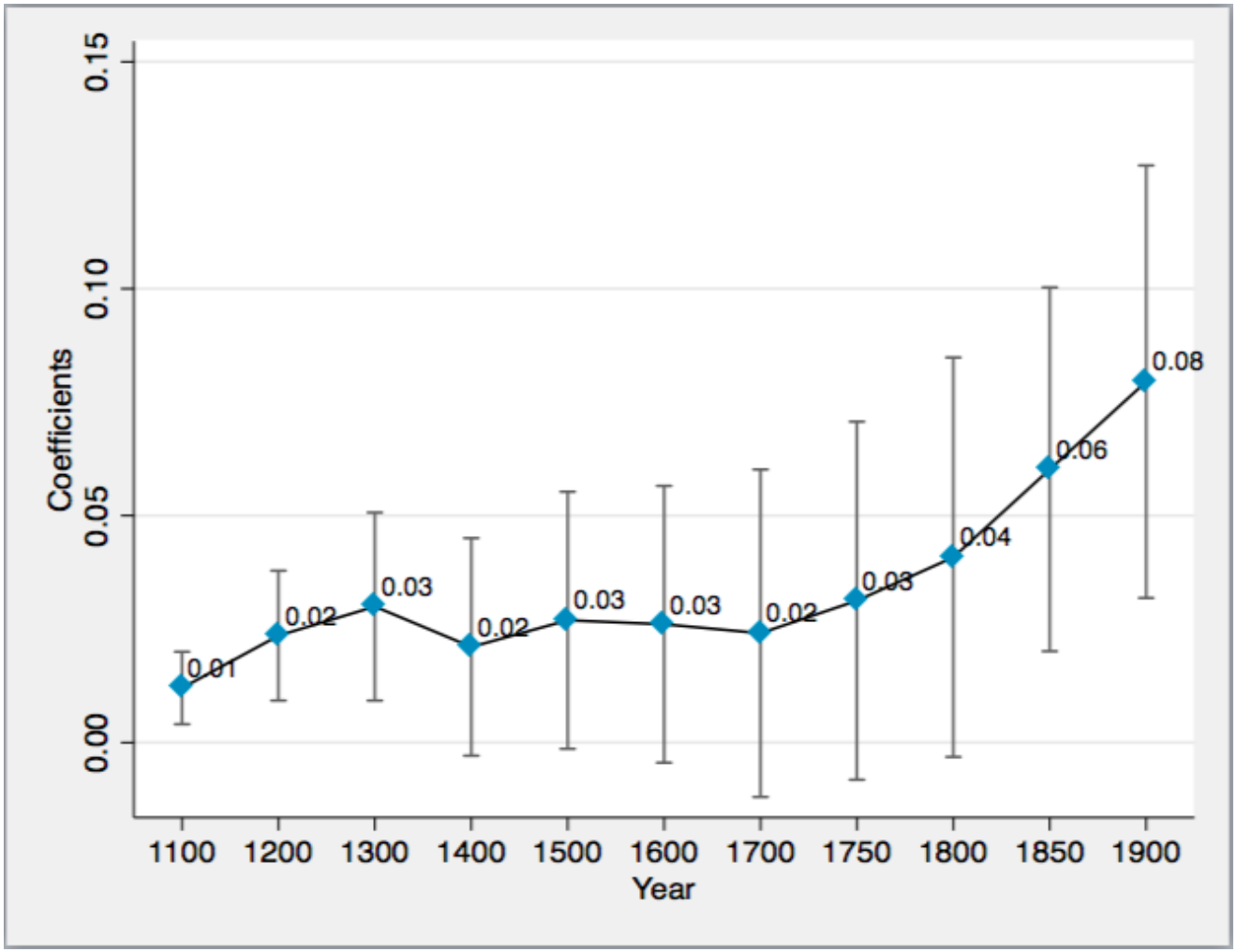
Figure 2: Average potato suitability among Old World countries, measured at the grid-cell level.

Flexible Estimation Result: Table 2

	ln Total Population			City Population Share		
	(1)	(2)	(3)	(4)	(5)	(6)
ln Suitable Area × 1100	0.013***	0.011***	0.012***	-0.002	-0.001	-0.001
	(0.003)	(0.003)	(0.004)	(0.001)	(0.001)	(0.001)
ln Suitable Area × 1200	0.029***	0.024***	0.024***	-0.001	-0.001	-0.001
	(0.005)	(0.005)	(0.007)	(0.001)	(0.001)	(0.001)
ln Suitable Area × 1300	0.039***	0.031***	0.030***	-0.000	-0.001	0.001
	(0.007)	(0.007)	(0.010)	(0.001)	(0.001)	(0.001)
ln Suitable Area × 1400	0.019**	0.004	0.021*	0.001	0.000	0.001
	(0.008)	(0.008)	(0.012)	(0.001)	(0.002)	(0.001)
ln Suitable Area × 1500	0.034***	0.014	0.027*	0.000	-0.000	0.001
	(0.009)	(0.010)	(0.014)	(0.001)	(0.001)	(0.001)
ln Suitable Area × 1600	0.041***	0.021**	0.026*	0.000	-0.001	-0.000
	(0.009)	(0.011)	(0.015)	(0.001)	(0.003)	(0.003)
ln Suitable Area × 1700	0.043***	0.018	0.024	0.002**	0.002	0.002
	(0.012)	(0.013)	(0.018)	(0.001)	(0.001)	(0.002)
ln Suitable Area × 1750	0.055***	0.030**	0.031	0.001	0.001	0.001
	(0.012)	(0.014)	(0.020)	(0.001)	(0.001)	(0.002)
ln Suitable Area × 1800	0.073***	0.048***	0.041*	0.002**	0.002	0.002
	(0.014)	(0.015)	(0.022)	(0.001)	(0.001)	(0.002)
ln Suitable Area × 1850	0.095***	0.069***	0.060***	0.002**	0.002	0.003*
	(0.015)	(0.017)	(0.020)	(0.001)	(0.001)	(0.002)
ln Suitable Area × 1900	0.121***	0.092***	0.080***	0.012***	0.012***	0.010***
	(0.017)	(0.021)	(0.024)	(0.002)	(0.002)	(0.003)
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
ln Old World Crops Area	No	Yes	Yes	No	Yes	Yes
ln Elevation	No	No	Yes	No	No	Yes
ln Ruggedness	No	No	Yes	No	No	Yes
R-square	0.989	0.989	0.990	0.418	0.421	0.455
Observations	1552.000	1552.000	1552.000	1552.000	1552.000	1552.000
F Stat for Joint Signficance 1750-1900	17.882	13.595	4.198	8.017	8.817	4.892

Method: areg. Standard errors clusted at isocode in parenthesis, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Flexible Estimation Result: Figure IVa



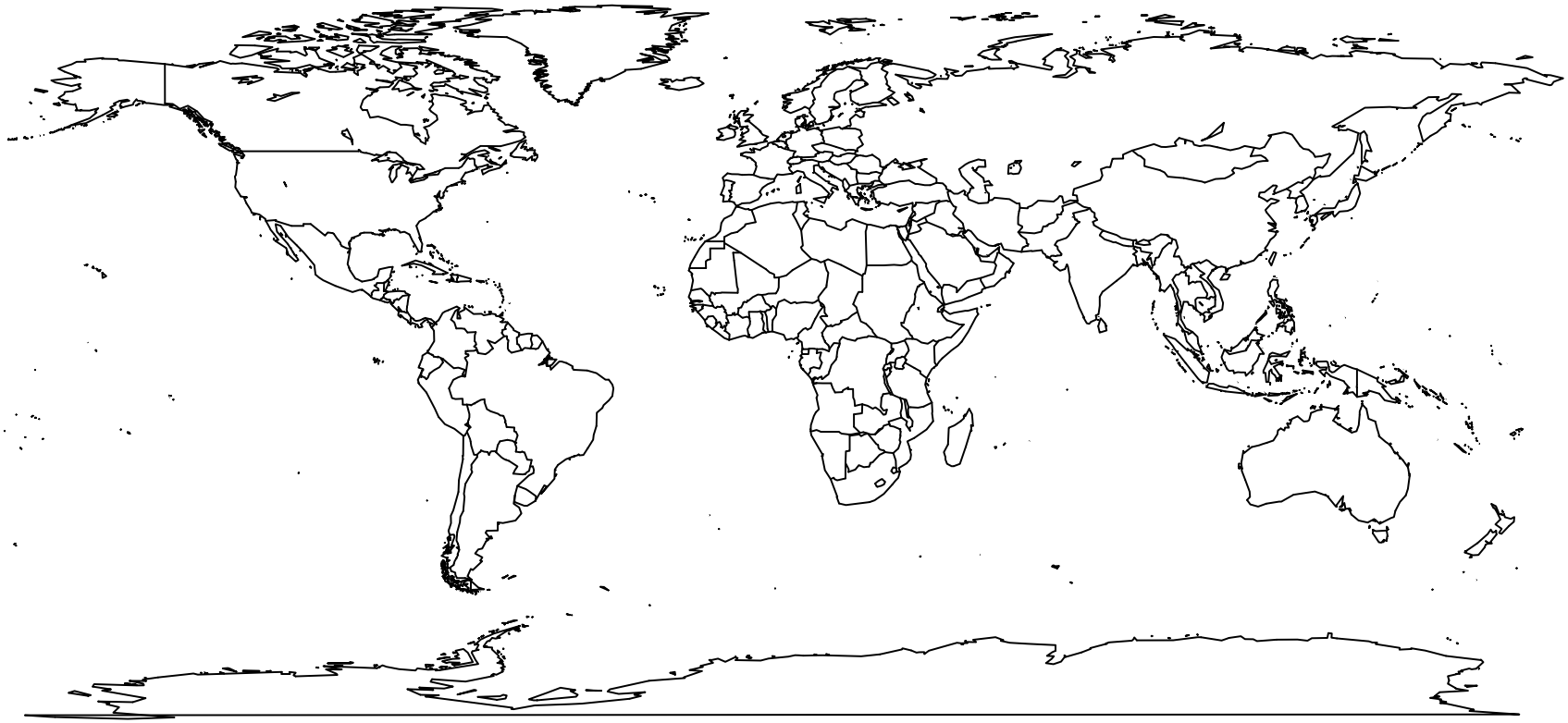
Reasons

- Many empirical economics papers, especially those involve a DID type of research design, have maps that illustrate spatial variations in variables of interest.
- Let's look at three R packages:`maps`, `ggmap` and `leaflet`
  - `maps` and `ggmap` are useful for creating static maps
  - `leaflet` is good for dynamic maps (See another html file for examples)
  - Often times static maps would be enough for us
- Ultimately, our job is to explain variations, but knowing a little map-making could be beneficial maybe
- There are many other software for making maps. I also suggest QGIS.

maps: Calling Base Maps I

```r
# install.packages("maps")
library(maps)
map("world")
```

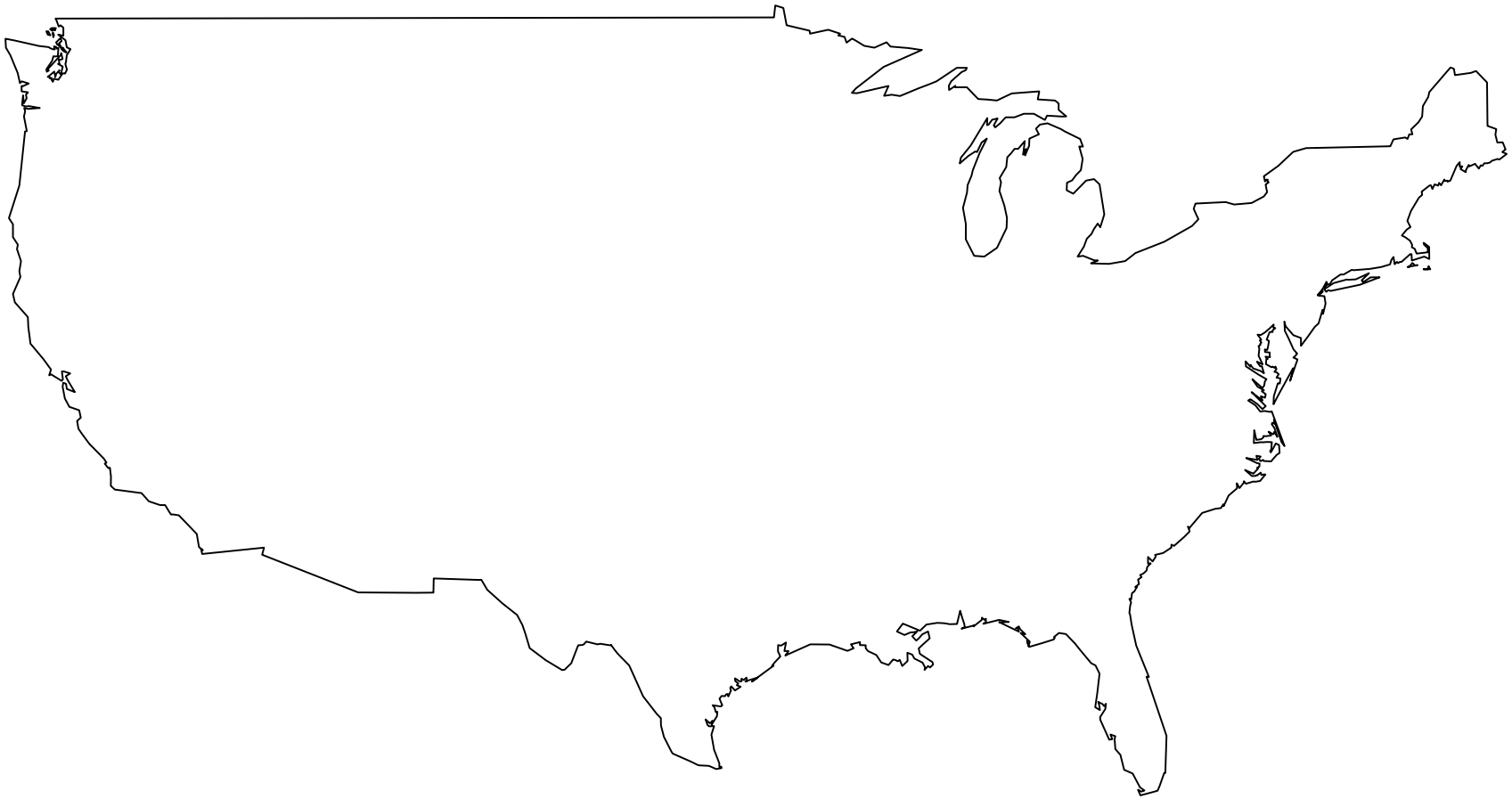
maps: Calling Base Maps II



maps: Calling Base Maps I

```
map("usa")
```

maps: Calling Base Maps II



maps: Calling Base Maps I

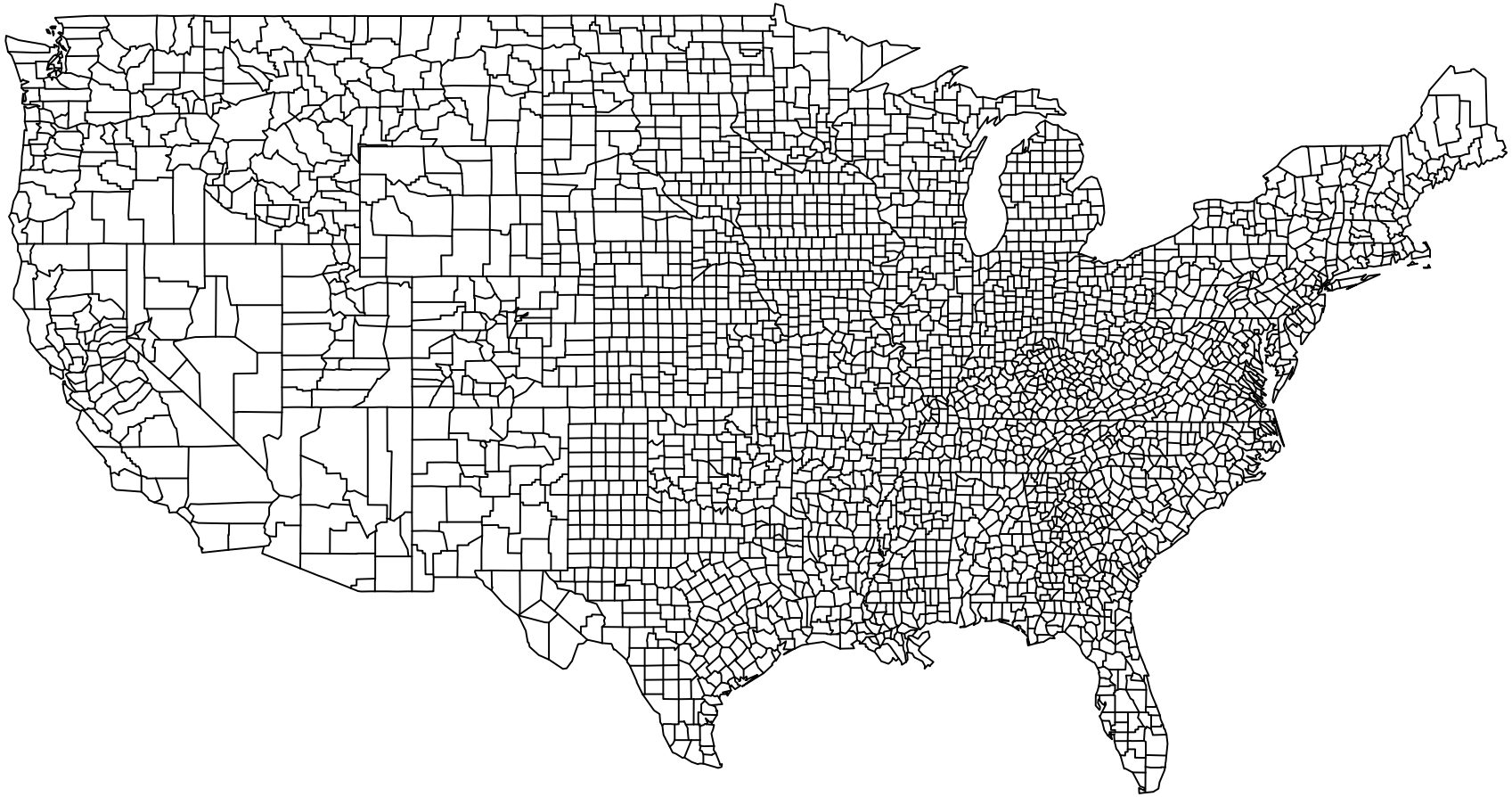
```
map("state")
```

maps: Calling Base Maps II

maps: Calling Base Maps I

```
map("county")
```

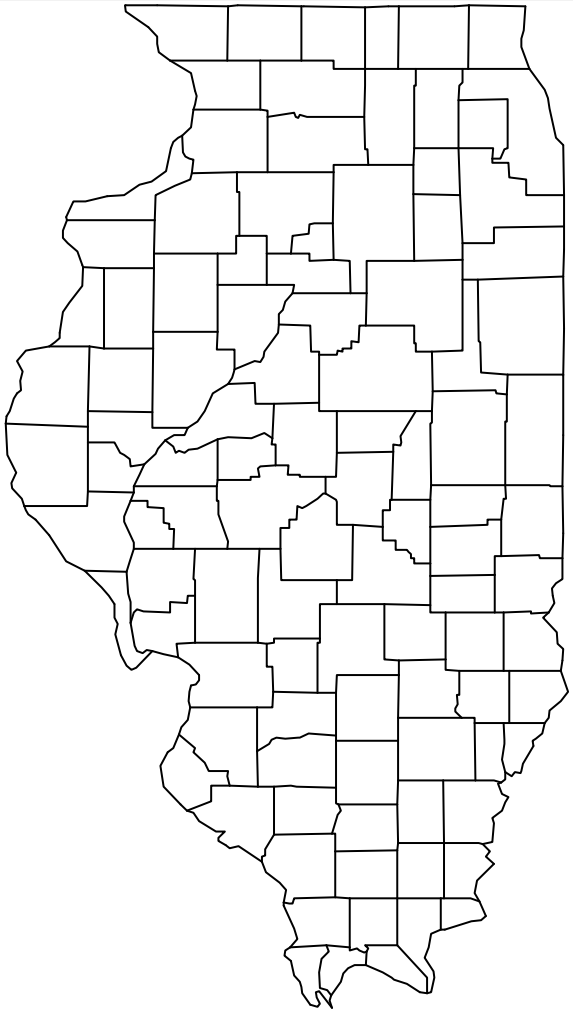
maps: Calling Base Maps II



maps: Calling Base Maps I

```
map("county", "illinois")
```

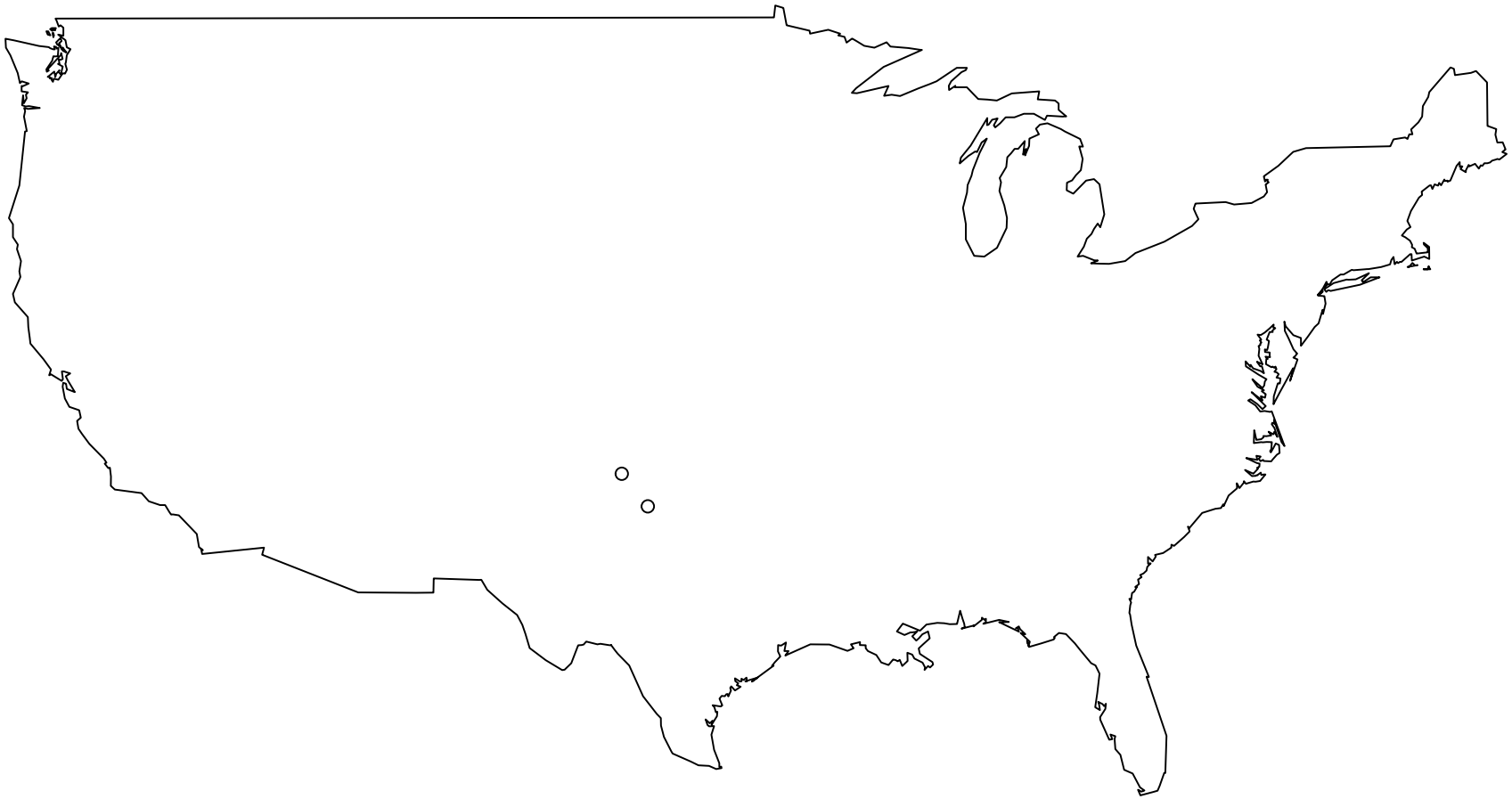
maps: Calling Base Maps II



maps: Dot Distribution Map I

```
map("usa")
# Add two silly points; point 1: long = -100, lat = 34
points(x = c(-100, -101), y = c(34, 35))
```

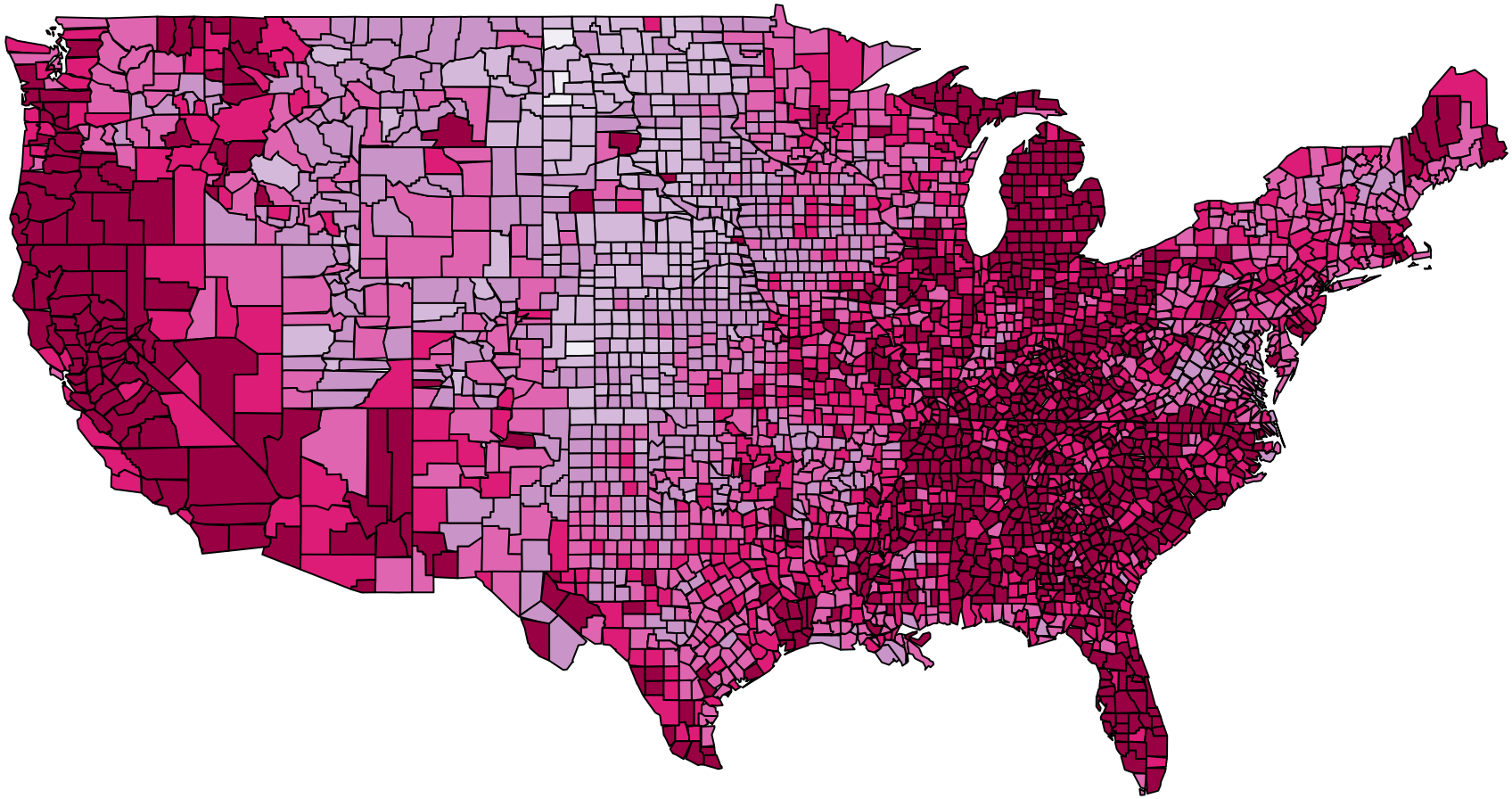
maps: Dot Distribution Map II



maps: Choropleth Map I

```
data(unemp)
data(county.fips)
colors = c("#F1EEF6", "#D4B9DA", "#C994C7",
           "#DF65B0", "#DD1C77", "#980043")
unemp$colorBuckets <-
    as.numeric(cut(unemp$unemp,
                   c(0, 2, 4, 6, 8, 10, 100)))
cnty.fips <- county.fips$fips[
          match(map("county", plot=FALSE)$names,
          county.fips$polyname)]
colorsmatched <-
    unemp$colorBuckets[match(cnty.fips, unemp$fips)]
map("county", col = colors[colorsmatched], fill = TRUE)
```

maps: Choropleth Map II



ggmap: Base Map I

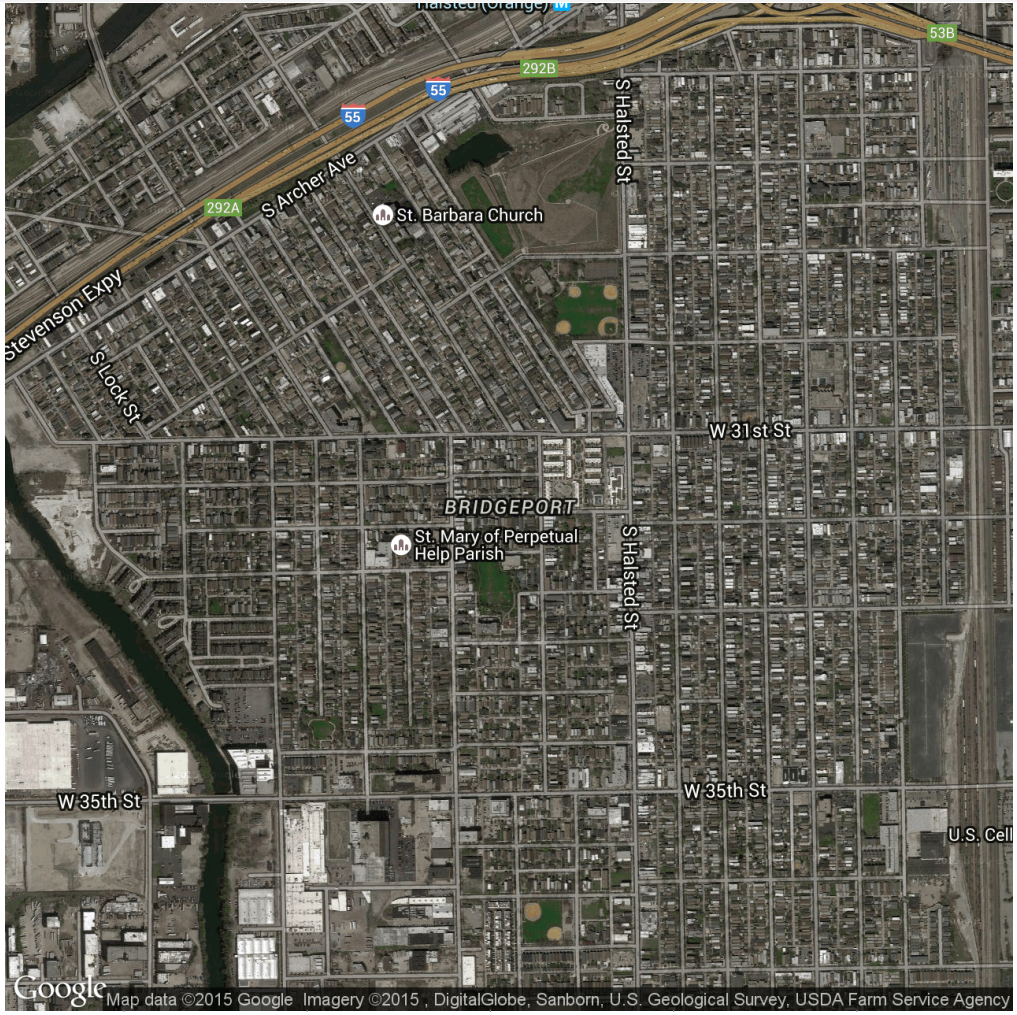
```
# install.packages("ggmap")
library(ggmap)
qmap("united states", zoom = 4)
```

ggmap: Base Map II

ggmap: Base Map I

```
qmap("bridgeport, Chicago", zoom = 15, maptype = "hybrid")
```

ggmap: Base Map II



ggmap: Dot Distribution Map I

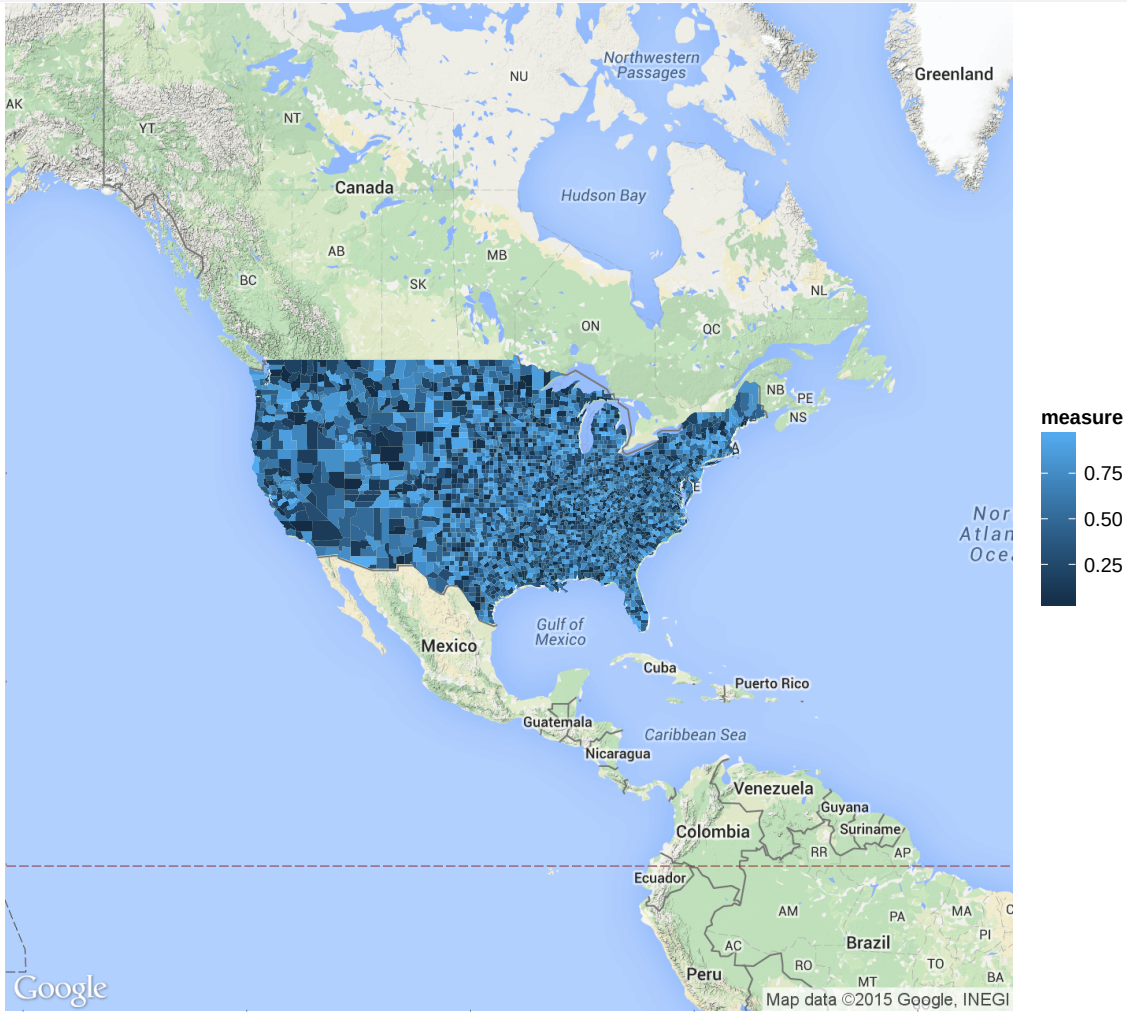
```
# generate some points with random longitude and latitude
points = data.frame(lon = rnorm(100, mean = -100, sd = 2),
                    lat = rnorm(100, mean = 40, sd = 2))
qmap("united states", zoom = 4) +
    geom_point(data = points)
```

ggmap: Dot Distribution Map II

ggmap: Choropleth Maps I

```
county <- fortify(map_data("county"))
number.of.county <- max(county$group)
measure <- data.frame(group = 1:number.of.county,
                      measure = runif(number.of.county))
county <- merge(county, measure, by = "group")
qmap("united states", zoom = 3) +
    geom_polygon(data = county,
    aes(x= long, y = lat, group = group, fill = measure))
```

ggmap: Choropleth Maps II



DID Extensions

- Wald-DID: combining DID with IV
  - In LATE framework, both RF and 1st estimates should be unbiased (required by the independent assumption), thus DID can be used to construct unbiased RF and 1st estimates
  - Require two common trend assumptions
  - Next week we will cover a paper by Duflo that illustrates this type of design
- Back-door DID: DID is not limited by panel data
  - The time dimension can be replaced by other covariate that refine the comparison group
  - Also another DDD paper next week
- Front-door DID: It's possible to even use post-treatment variable to define groups

Wald-DID Example: Bleakley and Chin (2004)

- Why Language Skill?
  - Arguably one big difference between immigrants and natives
  - Important for understanding assimilation, earning gap
- Empirical Challenge
  - Language skill is endogenous
  - Immigrants who speak good English might also that other better skills that affect earnings
  - Much like ability bias
- Language skill is hard to measure
  - Simple regression strategy might lead to upward bias and attenuation bias

Identification Strategy: Instrumental Variable

- Immigrants' **age at arrival** provides possibly exogenous variation of language skills

1. Is there a first stage?
   - Kids pick up language easier than teenage and adults
2. Is age-at-arrival as good as randomly assigned?
   - Probably not. Parents' plan, expectations, resources.
3. Is age-at-arrival satisfied exclusion restriction?
   - Probably not. Lots of things can happen for a few more years in U.S.

Identification Strategy: Extra Controls for Nonlanguage Effects

- Bring in another control group: immigrants from English-speaking countries
  - Control for secular age-at-arrival effect
- Simple illustration:

Immigrants	COO	Age-at-arrival	Earning (age 30)
Jamarco	Jamaica	5	12
Javina	Jamaica	15	10
Mateo	Mexico	5	10
Milana	Mexico	15	9

- Reduced-form relationship: (10-9) - (12-10) = -1.

Wald Estimator

- A simple Wald estimator

$$\rho_{IV} = \frac{\left(\bar{\mathrm{Y}}_{A=1,N=1} - \bar{\mathrm{Y}}_{A=0,N=1}\right) - \left(\bar{\mathrm{Y}}_{A=1,N=0} - \bar{\mathrm{Y}}_{A=0,N=0}\right)}{\left(\bar{\mathrm{X}}_{A=1,N=1} - \bar{\mathrm{X}}_{A=0,N=1}\right) - \left(\bar{\mathrm{X}}_{A=1,N=0} - \bar{\mathrm{X}}_{A=0,N=0}\right)}$$

  - A: dummy for having arrived young
  - N: dummy for having been born in a non-English-speaking countries
  - Y: log wage
  - X: a measure of English-language skills (1 = poor, 2 = well, 3 = very well)

Regressions

- First stage regression:

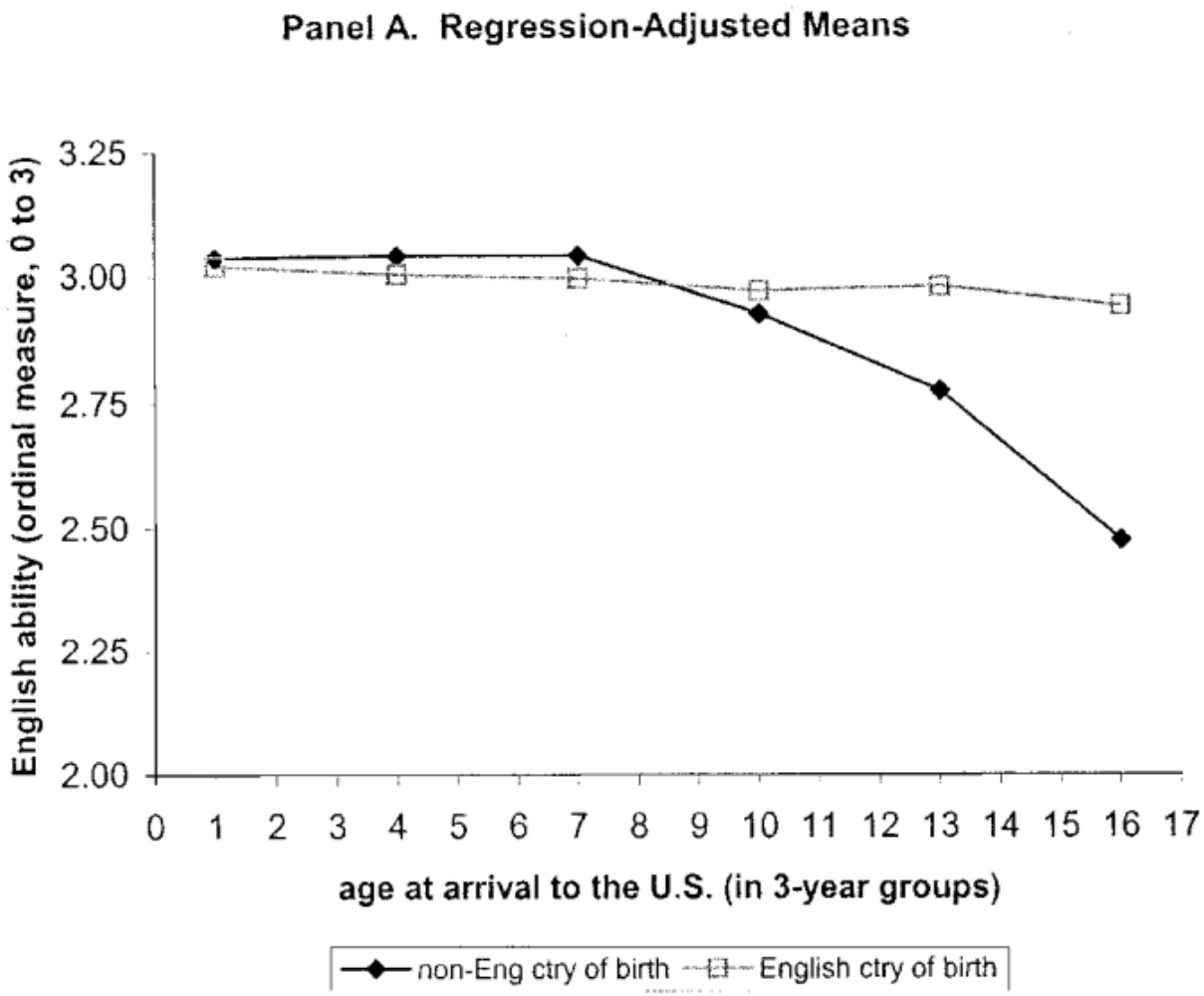
$$\mathrm{X}_{ija} = \alpha + \rho \mathrm{Z}_{ija} + \delta \mathrm{A}_a + \gamma \mathrm{N}_j + \mathrm{W}'\boldsymbol{\phi} + e_{ija}$$

- Reduced form regression:

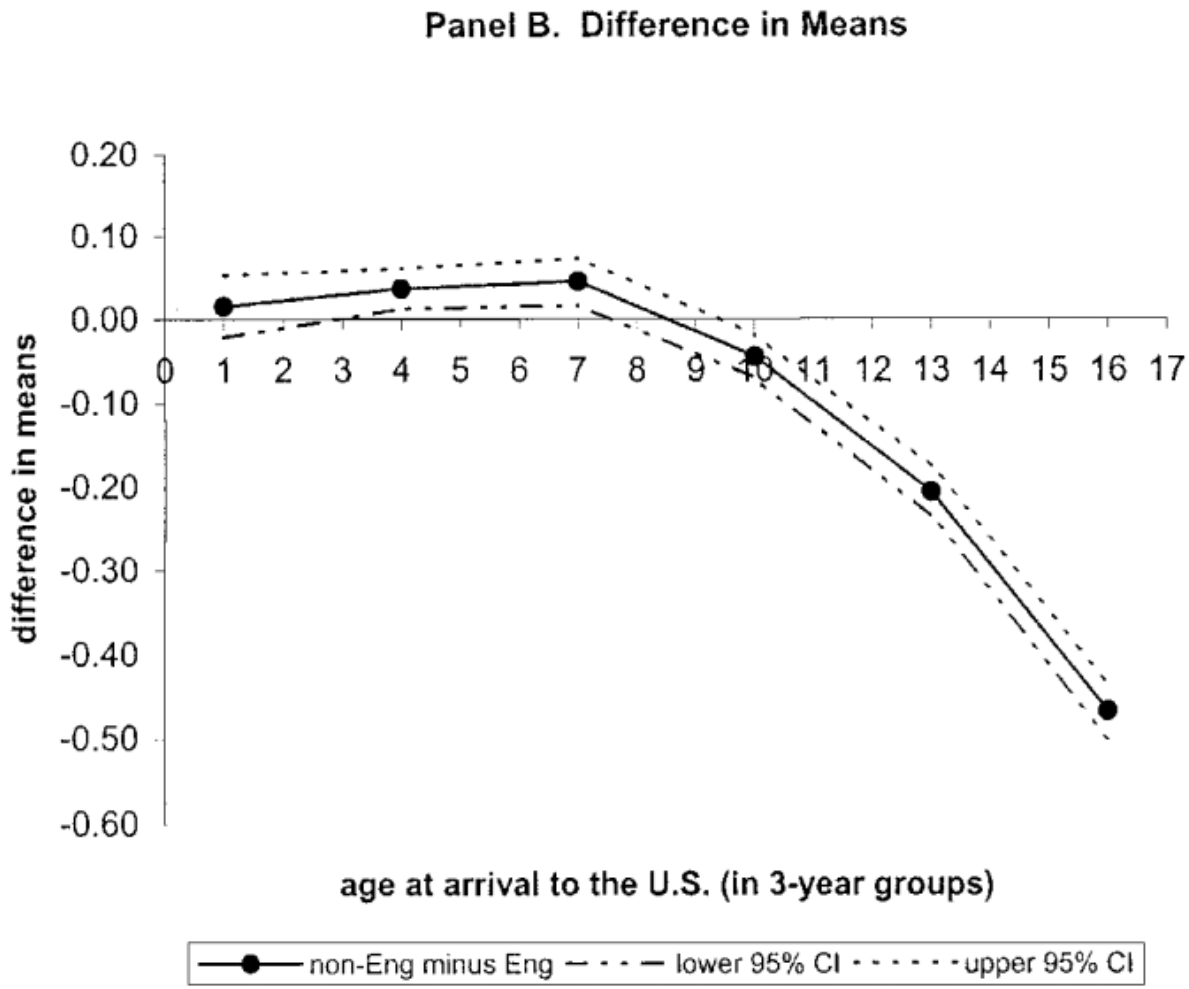
$$\mathrm{Y}_{ija} = \alpha + \rho \mathrm{Z}_{ija} + \delta \mathrm{A}_a + \gamma \mathrm{N}_j + \mathrm{W}'\boldsymbol{\phi} + e_{ija}$$

- Instrument Z_{ija} is an interaction between A_{ija} and N_j.

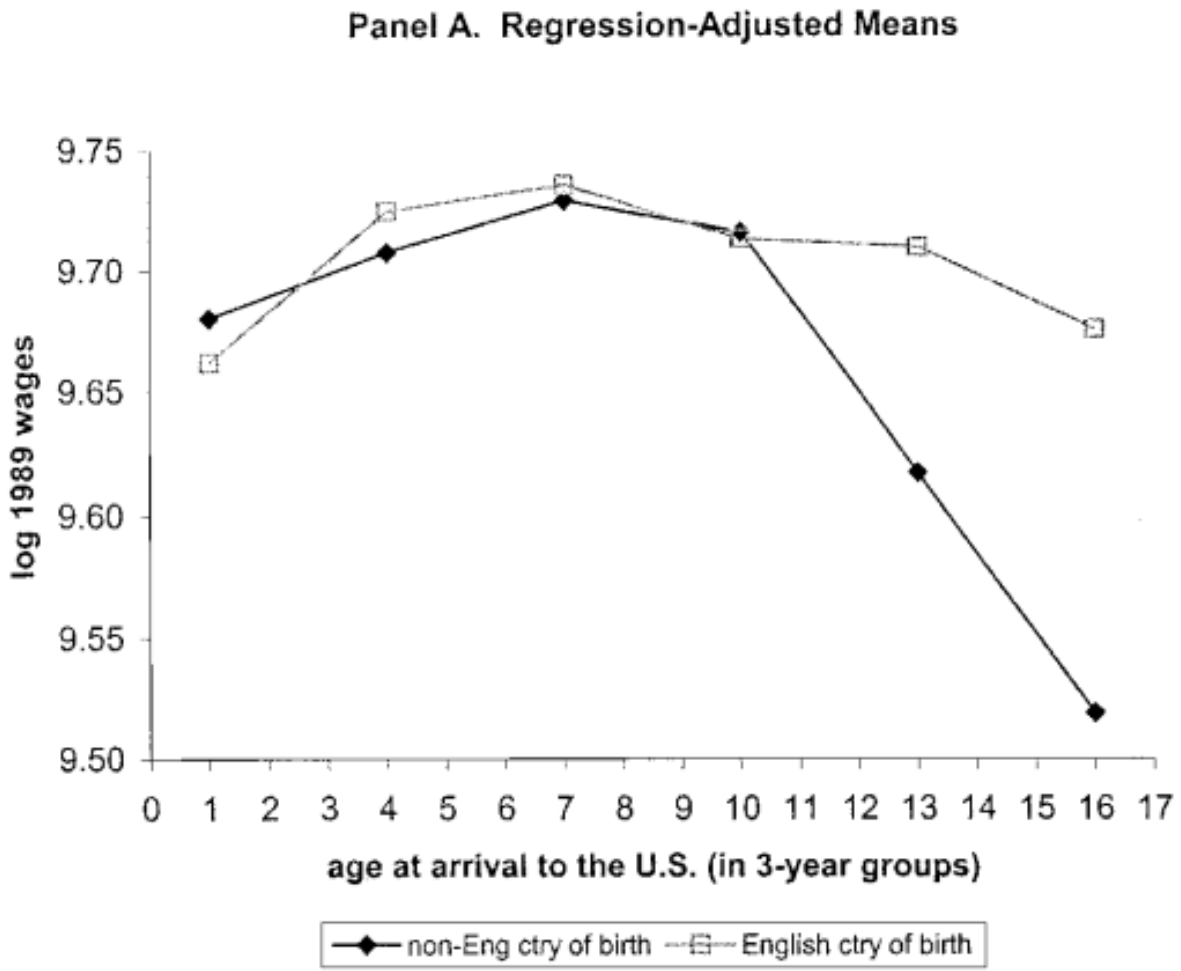
First Stage


Panel A. Regression-Adjusted Means
English ability (ordinal measure, 0 to 3)
3.25
3.00
2.75
2.50
2.25
2.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
age at arrival to the U.S. (in 3-year groups)
non-Eng ctry of birth
English ctry of birth

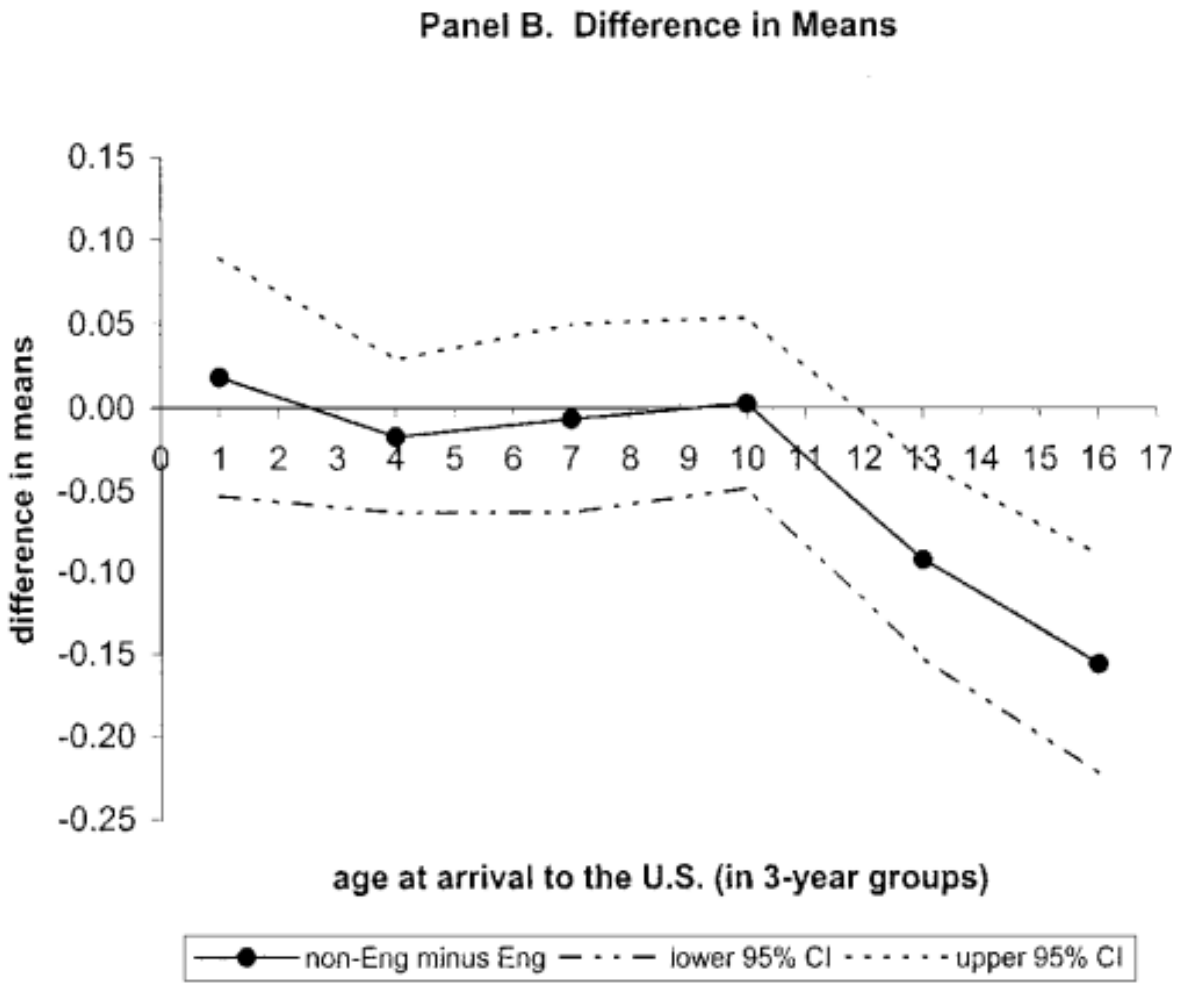
First Stage: Difference



Reduced Form for Wages



Reduced Form for Wages: Difference



First Stage and Reduced From Estimates

- Control for age, gender, race, and Hispanic

	ln Total Population			City Population Share		
	(1)	(2)	(3)	(4)	(5)	(6)
ln Suitable Area × 1100	0.013***	0.011***	0.012***	-0.002	-0.001	-0.001
	(0.003)	(0.003)	(0.004)	(0.001)	(0.001)	(0.001)
ln Suitable Area × 1200	0.029***	0.024***	0.024***	-0.001	-0.001	-0.001
	(0.005)	(0.005)	(0.007)	(0.001)	(0.001)	(0.001)
ln Suitable Area × 1300	0.039***	0.031***	0.030***	-0.000	-0.001	0.001
	(0.007)	(0.007)	(0.010)	(0.001)	(0.001)	(0.001)
ln Suitable Area × 1400	0.019**	0.004	0.021*	0.001	0.000	0.001
	(0.008)	(0.008)	(0.012)	(0.001)	(0.002)	(0.001)
ln Suitable Area × 1500	0.034***	0.014	0.027*	0.000	-0.000	0.001
	(0.009)	(0.010)	(0.014)	(0.001)	(0.001)	(0.001)
ln Suitable Area × 1600	0.041***	0.021**	0.026*	0.000	-0.001	-0.000
	(0.009)	(0.011)	(0.015)	(0.001)	(0.003)	(0.003)
ln Suitable Area × 1700	0.043***	0.018	0.024	0.002**	0.002	0.002
	(0.012)	(0.013)	(0.018)	(0.001)	(0.001)	(0.002)
ln Suitable Area × 1750	0.055***	0.030**	0.031	0.001	0.001	0.001
	(0.012)	(0.014)	(0.020)	(0.001)	(0.001)	(0.002)
ln Suitable Area × 1800	0.073***	0.048***	0.041*	0.002**	0.002	0.002
	(0.014)	(0.015)	(0.022)	(0.001)	(0.001)	(0.002)
ln Suitable Area × 1850	0.095***	0.069***	0.060***	0.002**	0.002	0.003*
	(0.015)	(0.017)	(0.020)	(0.001)	(0.001)	(0.002)
ln Suitable Area × 1900	0.121***	0.092***	0.080***	0.012***	0.012***	0.010***
	(0.017)	(0.021)	(0.024)	(0.002)	(0.002)	(0.003)
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
ln Old World Crops Area	No	Yes	Yes	No	Yes	Yes
ln Elevation	No	No	Yes	No	No	Yes
ln Ruggedness	No	No	Yes	No	No	Yes
R-square	0.989	0.989	0.990	0.418	0.421	0.455
Observations	1552.000	1552.000	1552.000	1552.000	1552.000	1552.000

Back-door DID Example: Madrian (1994)

- Hypothesis: Is There Evidence of Job-lock?
  - Does health insurance distort job mobility?
  - Do employees really feel constrained by health insurance?
- Having health insurance increases the cost of quitting.
- Why is this important?
  - Health insurance is one of the most important job amenities
  - Evaluating health care system reforms
  - Match-specific component of productivity

Identification

- Ideally, we want to have two groups of workers who are similar in all respects except for their insurance status.
  - If we find difference in mobility, it can be attributed to the insurance.
- In reality, native comparison has selection bias.
- Use three cost factors to refine comparison groups:
  - Having other health insurance
  - Having larger family size
  - Having pregnant wives

Cost Factor 1: A DID Comparison

- Implications and required assumptions
  - $M_{11} - M_{01} > 0$
  - $(M_{11} - M_{01)} - M_{11} - M_{01} > 0$

	Employer-provided health insurance	
	No	Yes
No other HI	M_{00}	M_{01}
Other HI	M_{10}	M_{11}

Empirical Implementation

$$\begin{aligned} \textit{Probability of Changing Jobs} &= \Phi\Big(\beta_0 + \beta_1 \times \textit{Health Insurance} + \beta_2 \times \textit{Cost Factor} \\ &\qquad + \beta_3 \times \textit{Health Insurance} \times \textit{Cost Factor} + \mathbf{z}'\gamma\Big), \\ &= \Phi(A_i), \end{aligned}$$

- Φ is the standard normal cumulative density function
- $\boldsymbol{z}$ is a vector of observable demographic characteristics
- β_0: mobility rate for people who have no insurance
- β_1: marginal impact of having own employment-based HI on mobility
- β_2: marginal impact of having other source of HI on mobility
- β_3: extra impact of having both sources of HI on mobility

Cost Factor 1: Having Other HI

TABLE III

EFFECT OF HEALTH INSURANCE ON THE TURNOVER PROBABILITY OF MARRIED MEN (STANDARD ERRORS IN PARENTHESES)

	Simple probit						RE probit	
	(1)		(2)		(3)		(4)	
A. Coefficient estimates								
Union	−.357	(.0842)	−.345	(.0861)	−.342	(.0878)	−.287	(.1054)
Black	−.031	(.0874)	−.022	(.0893)	−.041	(.0898)	−.032	(.0750)
Education	−.019	(.0139)	−.007	(.0142)	−.007	(.0143)	−.006	(.0122)
Experience	−.018	(.0037)	−.016	(.0038)	−.016	(.0038)	−.014	(.0050)
Log hourly wage	−.164	(.0619)	−.080	(.0639)	−.078	(.0644)	−.067	(.0570)
Months b/t interviews	.071	(.0256)	.074	(.0281)	.077	(.0282)	—	—
Health insurance (β_1)	—	—	−.626	(.0696)	−.715	(.0950)	−.586	(.1694)
Other health ins. (β_2)	—	—	—	—	−.039	(.1075)	−.029	(.0852)
HI × other HI (β_3)	—	—	—	—	.211	(.1339)	.167	(.1106)
σ_θ	—	—	—	—	—	—	.536	(.3693)
Log likelihood	−1040.55		−997.05		−994.73		−996.01	
B. Magnitude of job-lock								
Test 1: $\beta_2 + \beta_3 > 0$	—	—	—	—	.171	[.017]	.138	[.031]
$\hat{\beta}_2 + \hat{\beta}_3$ [p-value]								
Test 2: $\beta_3 > 0$	—	—	—	—	.211	[.058]	.167	[.066]
$\hat{\beta}_3$ [p-value]								
Degree of job-lock	—		—		26% to 30%		25% to 28%	

Cost Factor 1: Mobility Matrix

Predicted turnover probabilities	Employer-provided health insurance			
	No		Yes	
No other HI	.256	(.032)	.085	(.012)
Other HI	.244	(.032)	.115	(.017)
Estimates of job-lock				
a. Row difference among those with HI		26.0%	(13.8)	
b. Simple difference-in-difference		31.1%	(17.7)	
c. Adjusted difference-in-difference		29.6%	(13.8)	

- a: $\frac{0.115-0.085}{0.115} \approx 0.260.$
- b: $0.26 - \frac{0.244-0.256}{0.244} \approx 0.26 - (-5.1) = 0.311.$
- c: $\frac{0.115-0.081}{0.115} \approx 0.296$, where $0.081 = \frac{0.085}{1+0.051}$.

References I

Bleakley, Hoyt, and Aimee Chin. 2004. "Language Skills and Earnings: Evidence from Childhood Immigrants*." *Review of Economics and Statistics* 86 (2). MIT Press: 481–96.

Madrian, Brigitte C. 1994. "Employment-Based Health Insurance and Job Mobility: Is There Evidence of Job-Lock?" *The Quarterly Journal of Economics* 109 (1). MIT Press: 27–54.

Nunn, Nathan, and Nancy Qian. 2011. "The Potato's Contribution to Population and Urbanization: Evidence from a Historical Experiment*." *The Quarterly Journal of Economics* 126 (2). Oxford University Press: 593–650.