Lectures 08-09:
Paper overviews

PPHA 34600

Prof. Fiona Burlig

Harris School of Public Policy
University of Chicago

Lecture 08 recap

TL;DR:

1. Instrumental variables are very powerful
2. ...but they require extremely strong assumptions!
3. Hashtag no free lunch

An example: Conflict in Africa

Policy issue:

- Conflict is a big problem...
- ...but what causes it?
- How much support is there for economic channels?

Approach:

- (We're not actually evaluating a program here)
- We need a shock to economic conditions

→ We don't have randomization, so we use IV

- Instrument of choice: (changes in) rainfall

→ Do we believe this? Hold that thought...

Estimating treatment effects of growth on conflict

How does economic growth affect civil conflict (simplified)?

First stage:

$$growth_{it} = \alpha + \gamma \Delta R_{it} + \beta X_{it} + \eta_{it}$$

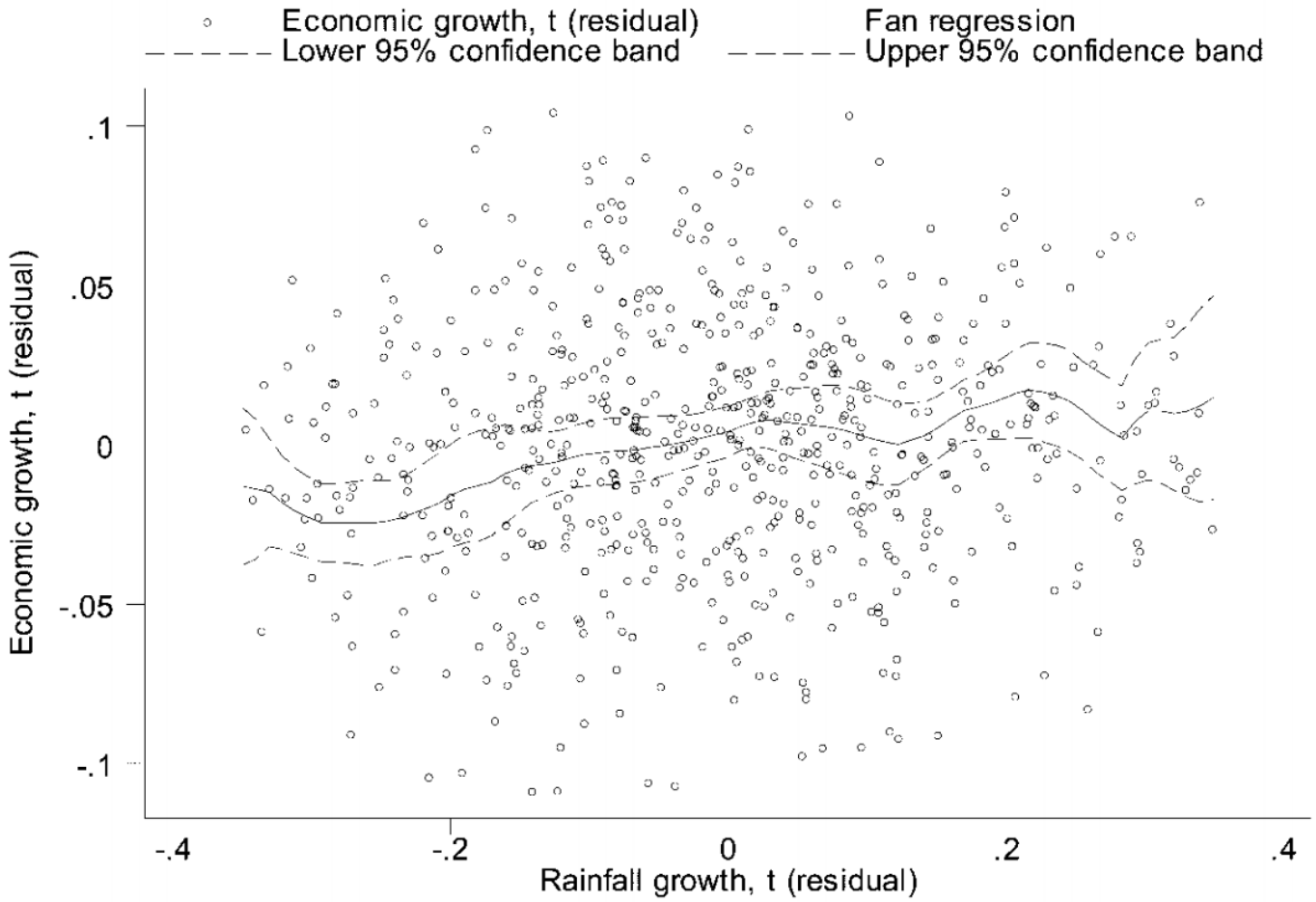
where

$growth_{it}$ is economic growth in country i, time t

$\Delta R_{it} = R_{i,t} - R_{i,t-1}$ is the change in rainfall

η_{it} is an error term

First stage (graphical)



First stage (tabular form)

RAINFALL AND ECONOMIC GROWTH (First-Stage)
Dependent Variable: Economic Growth Rate, t

EXPLANATORY VARIABLE	ORDINARY LEAST SQUARES				
	(1)	(2)	(3)	(4)	(5)
Growth in rainfall, t	.055***	.053***	.049***	.049***	.053***
	(.016)	(.017)	(.017)	(.018)	(.018)
Growth in rainfall, $t-1$	.034**	.032**	.028**	.028*	.037**
	(.013)	(.014)	(.014)	(.014)	(.015)
Growth in rainfall, $t+1$				.001	
				(.019)	
Growth in terms of trade, t					−.002
					(.023)
Log(GDP per capita), 1979		−.011			
		(.007)			
Democracy (Polity IV), $t-1$		.0000			
		(.0007)			
Ethnolinguistic fractionalization		.006			
		(.044)			
Religious fractionalization		.045			
		(.044)			
Oil-exporting country		.007			
		(.019)			
Log(mountainous)		.001			
		(.005)			
Log(national population), $t-1$		−.009			
		(.009)			
Country fixed effects	no	no	yes	yes	yes
Country-specific time trends	no	yes	yes	yes	yes
R^2	.02	.08	.13	.13	.16
Root mean square error	.07	.07	.07	.07	.06
Observations	743	743	743	743	661

NOTE.—Huber robust standard errors are in parentheses. Regression disturbance terms are clustered at the country level. A country-specific year time trend is included in all specifications (coefficient estimates not reported).

* Significantly different from zero at 90 percent confidence.

** Significantly different from zero at 95 percent confidence.

*** Significantly different from zero at 99 percent confidence.

Estimating treatment effects of growth on conflict

How does economic growth affect civil conflict (simplified)?

First stage:

$$growth_{it} = \alpha + \beta X_{it} + \gamma \Delta R_{it} + \eta_{it}$$

where
$growth_{it}$ is economic growth in country i, time t
$\Delta R_{it} = R_{i,t} - R_{i,t-1}$ is the change in rainfall
η_{it} is an error term

Second stage:

$$conflict_{it} + \alpha + \delta \widehat{growth}_{it} + \tau X_{it} + \eta_{it}$$

where
$\widehat{growth}_{it}$ is the fitted values from the first stage

Second stage (OLS only)

ECONOMIC GROWTH AND CIVIL CONFLICT

Explanatory Variable	Dependent Variable: Civil Conflict ≥25 Deaths				Dependent Variable: Civil Conflict ≥1,000 Deaths
	Probit (1)	OLS (2)	OLS (3)	OLS (4)	
Economic growth rate, t	−.37	−.33	−.21	−.21	
	(.26)	(.26)	(.20)	(.16)	
Economic growth rate, $t-1$	−.14	−.08	.01	.07	
	(.23)	(.24)	(.20)	(.16)	
Log(GDP per capita), 1979	−.067	−.041	.085		
	(.061)	(.050)	(.084)		
Democracy (Polity IV), $t-1$	.001	.001	.003		
	(.005)	(.005)	(.006)		
Ethnolinguistic fractionalization	.24	.23	.51		
	(.26)	(.27)	(.40)		
Religious fractionalization	−.29	−.24	.10		
	(.26)	(.24)	(.42)		
Oil-exporting country	.02	.05	−.16		
	(.21)	(.21)	(.20)		
Log(mountainous)	.077**	.076*	.057		
	(.041)	(.039)	(.060)		
Log(national population), $t-1$	.080	.068	.182*		
	(.051)	(.051)	(.086)		
Country fixed effects	no	no	no	yes	
Country-specific time trends	no	no	yes	yes	
R^2	…	.13	.53	.71	
Root mean square error	…	.42	.31	.25	
Observations	743	743	743	743	

NOTE.—Huber robust standard errors are in parentheses. Regression disturbance terms are clustered at the country level. Regression 1 presents marginal probit effects, evaluated at explanatory variable mean values. The instrumental variables for economic growth in regressions 5–7 are growth in rainfall, t and growth in rainfall, $t-1$. A country-specific year time trend is included in all specifications (coefficient estimates not reported), except for regressions 1 and 2, where a single linear time trend is included.

* Significantly different from zero at 90 percent confidence.

** Significantly different from zero at 95 percent confidence.

*** Significantly different from zero at 99 percent confidence.

Second stage (tabular form)

ECONOMIC GROWTH AND CIVIL CONFLICT

Explanatory Variable	Dependent Variable: Civil Conflict ≥25 Deaths						Dependent Variable: Civil Conflict ≥1,000 Deaths
	Probit (1)	OLS (2)	OLS (3)	OLS (4)	IV-2SLS (5)	IV-2SLS (6)	IV-2SLS (7)
Economic growth rate, t	−.37	−.33	−.21	−.21	−.41	−1.13	−1.48*
	(.26)	(.26)	(.20)	(.16)	(1.48)	(1.40)	(.82)
Economic growth rate, $t-1$	−.14	−.08	.01	.07	−2.25**	−2.55**	−.77
	(.23)	(.24)	(.20)	(.16)	(1.07)	(1.10)	(.70)
Log(GDP per capita), 1979	−.067	−.041	.085		.053		
	(.061)	(.050)	(.084)		(.098)		
Democracy (Polity IV), $t-1$	.001	.001	.003		.004		
	(.005)	(.005)	(.006)		(.006)		
Ethnolinguistic fractionalization	.24	.23	.51		.51		
	(.26)	(.27)	(.40)		(.39)		
Religious fractionalization	−.29	−.24	.10		.22		
	(.26)	(.24)	(.42)		(.44)		
Oil-exporting country	.02	.05	−.16		−.10		
	(.21)	(.21)	(.20)		(.22)		
Log(mountainous)	.077**	.076*	.057		.060		
	(.041)	(.039)	(.060)		(.058)		
Log(national population), $t-1$	.080	.068	.182*		.159*		
	(.051)	(.051)	(.086)		(.093)		
Country fixed effects	no	no	no	yes	no	yes	yes
Country-specific time trends	no	no	yes	yes	yes	yes	yes
R^2	…	.13	.53	.71	…	…	…
Root mean square error	…	.42	.31	.25	.36	.32	.24
Observations	743	743	743	743	743	743	743

NOTE.—Huber robust standard errors are in parentheses. Regression disturbance terms are clustered at the country level. Regression 1 presents marginal probit effects, evaluated at explanatory variable mean values. The instrumental variables for economic growth in regressions 5–7 are growth in rainfall, t and growth in rainfall, $t-1$. A country-specific year time trend is included in all specifications (coefficient estimates not reported), except for regressions 1 and 2, where a single linear time trend is included.

* Significantly different from zero at 90 percent confidence.

** Significantly different from zero at 95 percent confidence.

*** Significantly different from zero at 99 percent confidence.

Estimating the reduced form

How does (growth in) rainfall affect civil conflict (simplified)?

Reduced form:

$$conflict_{it} = \alpha + \theta \Delta R_{it} + \pi X_{it} + \eta_i$$

Reduced form (tabular)

RAINFALL AND CIVIL CONFLICT (Reduced-Form)

	DEPENDENT VARIABLE	
EXPLANATORY VARIABLE	Civil Conflict ≥25 Deaths (OLS) (1)	Civil Conflict ≥1,000 Deaths (OLS) (2)
Growth in rainfall, t	−.024	−.062**
	(.043)	(.030)
Growth in rainfall, $t-1$	−.122**	−.069**
	(.052)	(.032)
Country fixed effects	yes	yes
Country-specific time trends	yes	yes
R^2	.71	.70
Root mean square error	.25	.22
Observations	743	743

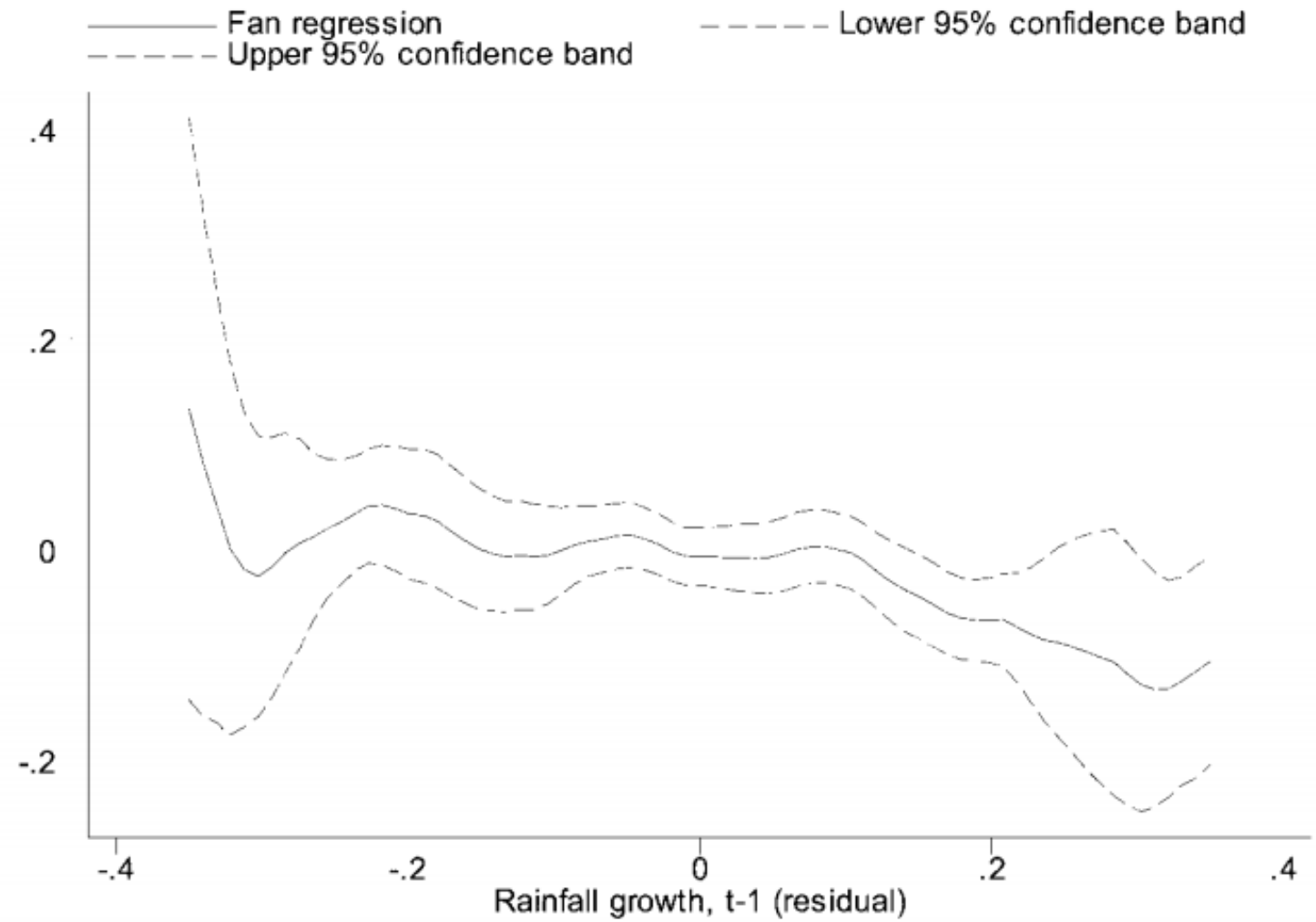
NOTE.—Huber robust standard errors are in parentheses. Regression disturbance terms are clustered at the country level. A country-specific year time trend is included in all specifications (coefficient estimates not reported).

* Significantly different from zero at 90 percent confidence.

** Significantly different from zero at 95 percent confidence.

*** Significantly different from zero at 99 percent confidence.

Reduced form (graphical)



The exclusion restriction is the key to any IV

You should always ask:
What is the exclusion restriction in this analysis saying?

The exclusion restriction is the key to any IV

You should always ask:
What is the exclusion restriction in this analysis saying?

Do we believe this? Why or why not?

Second stage (tabular form)

ECONOMIC GROWTH AND CIVIL CONFLICT

Explanatory Variable	Dependent Variable: Civil Conflict ≥25 Deaths						Dependent Variable: Civil Conflict ≥1,000 Deaths
	Probit (1)	OLS (2)	OLS (3)	OLS (4)	IV-2SLS (5)	IV-2SLS (6)	IV-2SLS (7)
Economic growth rate, t	−.37	−.33	−.21	−.21	−.41	−1.13	−1.48*
	(.26)	(.26)	(.20)	(.16)	(1.48)	(1.40)	(.82)
Economic growth rate, $t-1$	−.14	−.08	.01	.07	−2.25**	−2.55**	−.77
	(.23)	(.24)	(.20)	(.16)	(1.07)	(1.10)	(.70)
Log(GDP per capita), 1979	−.067	−.041	.085		.053		
	(.061)	(.050)	(.084)		(.098)		
Democracy (Polity IV), $t-1$	.001	.001	.003		.004		
	(.005)	(.005)	(.006)		(.006)		
Ethnolinguistic fractionalization	.24	.23	.51		.51		
	(.26)	(.27)	(.40)		(.39)		
Religious fractionalization	−.29	−.24	.10		.22		
	(.26)	(.24)	(.42)		(.44)		
Oil-exporting country	.02	.05	−.16		−.10		
	(.21)	(.21)	(.20)		(.22)		
Log(mountainous)	.077**	.076*	.057		.060		
	(.041)	(.039)	(.060)		(.058)		
Log(national population), $t-1$	.080	.068	.182*		.159*		
	(.051)	(.051)	(.086)		(.093)		
Country fixed effects	no	no	no	yes	no	yes	yes
Country-specific time trends	no	no	yes	yes	yes	yes	yes
R^2	…	.13	.53	.71	…	…	…
Root mean square error	…	.42	.31	.25	.36	.32	.24
Observations	743	743	743	743	743	743	743

NOTE.—Huber robust standard errors are in parentheses. Regression disturbance terms are clustered at the country level. Regression 1 presents marginal probit effects, evaluated at explanatory variable mean values. The instrumental variables for economic growth in regressions 5–7 are growth in rainfall, t and growth in rainfall, $t-1$. A country-specific year time trend is included in all specifications (coefficient estimates not reported), except for regressions 1 and 2, where a single linear time trend is included.

* Significantly different from zero at 90 percent confidence.

** Significantly different from zero at 95 percent confidence.

*** Significantly different from zero at 99 percent confidence.

Lecture 09 recap

TL;DR:

1. Instrumental variables are very powerful
2. With the right assumptions...
3. ...we can handle OVB and ME (and simultaneity)

An example: Early-life rainfall and health

Policy issue:

- Early-life shocks may be very important
- With bad harvests, kids may not get the proper nutrition

Approach:

- (We're not actually evaluating a program here)
- We want to estimate the effect of rainfall on health
- **Measurement of rainfall is poor in Indonesia**
- Instrument of choice: rainfall at weather stations $j \neq i$

Estimating the effects of rainfall on health

The authors will run a (simplified) version of:

$$Y_i = \tau Rainfall_i + \varepsilon_i$$

Where:

Y_i is a health outcome of interest

$Rainfall_i$ is rain in location i

- (They'll actually do this in a series of lags)

ε_i is an error term

Estimating the effects of rainfall on health

The authors will run a (simplified) version of:

$$Y_i = \tau Rainfall_i + \varepsilon_i$$

Where:

Y_i is a health outcome of interest

$Rainfall_i$ is rain in location i

- (They'll actually do this in a series of lags)

ε_i is an error term

A big concern

- $Rainfall_i$ is measured with error
- We are likely to understate the true effect
- **Solution:** $Z_i = Rainfall\ Nearby_i$!

First stage estimates

Dependent variable: Rainfall in birthyear and birthdistrict (deviation of log rainfall in birth district from log of 1953-1999 district mean rainfall)

	Women	Men
Birthyear/birthdistrict rainfall, 2nd-closest station	0.138 (0.024)***	0.120 (0.023)***
Birthyear/birthdistrict rainfall, 3rd-closest station	0.144 (0.039)***	0.158 (0.035)***
Birthyear/birthdistrict rainfall, 4th-closest station	0.088 (0.053)	0.081 (0.044)*
Birthyear/birthdistrict rainfall, 5th-closest station	0.125 (0.025)***	0.158 (0.039)***
Number of observations	4,615	4,277
R-squared	0.59	0.59
F-statistic: Joint significance of all four rainfall variables	31.61	28.80
P-value	0.000	0.000

Placebo test estimates

Coefficients (std. errors) in regression of outcome on *child's* birthyear rainfall.

	Women	Men
Mother's characteristics		
Completed grades of schooling	0.204	0.132
	(1.136)	(0.947)
	[2,447]	[2,258]
Currently alive (indicator)	0.084	0.029
	(0.083)	(0.108)
	[4,542]	[4,039]
Father's characteristics		
Completed grades of schooling	0.273	0.166
	(1.172)	(1.309)
	[2,810]	[2,621]
Currently alive (indicator)	0.010	-0.093
	(0.080)	(0.169)
	[4,541]	[4,040]

Placebo test estimates

	Women	Men
Self-rep. health status very good (indic.)	0.123	-0.115
	(0.099)	(0.078)
	[1,239]	[1,264]
Self-rep. health status poor/very poor (indic.)	0.090	0.106
	(0.154)	(0.134)
	[1,239]	[1,264]
Ln (lung capacity)	-0.067	0.008
	(0.034)*	(0.089)
	[1,195]	[1,130]
Height (cm.)	-1.165	3.054
	(1.660)	(2.017)
	[1,207]	[1,132]
Days absent due to illness (last 4 weeks)	0.669	3.075
	(0.688)	(1.505)*
	[1,240]	[1,261]
Completed grades of schooling	0.958	-1.441
	(1.274)	(1.947)
	[1,240]	[1,260]
Ln (expenditures per cap. in hh)	-0.193	-0.329
	(0.284)	(0.189)
	[1,240]	[1,264]
Asset index	-0.773	0.166
	(0.497)	(0.353)
	[1,240]	[1,264]
Ln (annual earnings)	0.202	-0.612
	(0.333)	(0.344)
	[631]	[1,142]

2SLS estimates

TABLE 2—EFFECT OF BIRTH YEAR RAINFALL ON ADULT OUTCOMES: WOMEN AND MEN BORN 1953–1974
(Instrumental variables estimates. Coefficients (standard errors) in regression of outcome on rainfall in individual's birth year and birth district. Instrumental variables for birth year/birth district rainfall are rainfall measured at second- through fifth-closest rainfall stations to respondent's birth district.)

	Women	Men
Self-reported health status very good (indicator)	0.101	−0.029
	(0.058)*	(0.072)
	[4,613]	[4,270]
Self-reported health status poor/very poor (indicator)	−0.192	−0.100
	(0.082)**	(0.098)
	[4,613]	[4,270]
Ln (lung capacity)	−0.044	−0.073
	(0.049)	(0.062)
	[4,454]	[3,907]
Height (centimeters)	2.832	0.998
	(0.821)***	(1.795)
	[4,495]	[3,924]
Days absent due to illness (last four weeks)	−1.175	0.515
	(0.831)	(0.779)
	[4,611]	[4,267]
Completed grades of schooling	1.086	−0.474
	(0.453)**	(1.490)
	[4,598]	[4,259]
Ln (expenditures per capita in household)	0.095	−0.274
	(0.204)	(0.301)
	[4,615]	[4,277]
Asset index	0.876	−0.279
	(0.324)**	(0.507)
	[4,613]	[4,276]
Ln (annual earnings)	0.065	−0.202
	(0.988)	(0.350)
	[2,332]	[3,963]

2SLS estimates

TABLE 3—EFFECT OF RAINFALL IN YEARS BEFORE AND AFTER BIRTH: WOMEN BORN 1953–1974
(Instrumental variables estimates. Rainfall in individual's birth year and birth district instrumented with rainfall measured at second- through fifth-closest rainfall stations to respondent's birth district.)

Dependent variable	Self-reported health status very good (indicator)	Self-reported health status poor/very poor (indicator)	Height (centimeters)	Completed grades of schooling	Asset index
Coefficient on rainfall in:					
Year −3	0.025	−0.114	1.505	−0.065	0.003
	(0.084)	(0.120)	(1.572)	(0.992)	(0.424)
Year −2	−0.037	−0.013	0.854	−0.852	−0.426
	(0.103)	(0.075)	(1.813)	(1.670)	(0.721)
Year −1	−0.080	−0.045	3.338	0.104	−0.380
	(0.123)	(0.088)	(2.155)	(1.332)	(0.530)
Year 0	0.090	−0.179	3.833	1.598	0.750
	(0.067)	(0.093)*	(1.420)**	(0.675)**	(0.399)*
Year 1	−0.008	−0.096	0.676	1.083	0.203
	(0.053)	(0.067)	(1.592)	(0.769)	(0.272)
Year 2	−0.041	−0.015	1.666	0.117	−0.229
	(0.043)	(0.068)	(0.984)	(0.840)	(0.452)
Year 3	−0.020	−0.104	1.996	−0.135	0.088
	(0.116)	(0.067)	(1.774)	(0.802)	(0.232)
Observations	4,613	4,613	4,495	4,598	4,613