Regression 1: Linear regression LQRPS

Frederik Hjorth fh@ifs.ku.dk fghjorth.github.io @fghjorth

Department of Political Science University of Copenhagen

February 7th, 2017

Motivating ex.

OLS intuition

- Classic view of quant polisci (Lijphart, Nørgaard
- Reioinder: Dahler-Larsen & Sylves
- The perestroika debate (Laitin, Flyvbjerg
- Ex. 1: the IE mode
- DA-RT
- Counter-DA-R1
- On The Run
- Intro to F

Motivating ex.

OLS intuition

- Classic view of quant polisci (Lijphart, Nørgaard)
- Rejoinder: Dahler-Larsen & Sylves
- The perestroika debate (Laitin, Flyvbjerg
- Ex. 1: the IE mode
- DA-RT
- Counter-DA-RT
- On The Run
- Intro to R

Motivating ex.

OLS intuition

- Classic view of quant polisci (Lijphart, Nørgaard)
- Rejoinder: Dahler-Larsen & Sylvest
- The perestroika debate (Laitin, Flyvbjerg
- Ex. 1: the IE mode
- DA-RT
- Counter-DA-RT
- On The Run
- Intro to R

Motivating ex.

- Classic view of quant polisci (Lijphart, Nørgaard)
- Rejoinder: Dahler-Larsen & Sylvest
- The perestroika debate (Laitin, Flyvbjerg)

OLS intuition

- Classic view of quant polisci (Lijphart, Nørgaard)
- Rejoinder: Dahler-Larsen & Sylvest
- The perestroika debate (Laitin, Flyvbjerg)
- Ex. 1: the IE model
- DA-RT
- Counter-DA-RT
- On The Run
- Intro to R

Motivating ex.

OLS intuition

- Classic view of quant polisci (Lijphart, Nørgaard)
- Rejoinder: Dahler-Larsen & Sylvest
- The perestroika debate (Laitin, Flyvbjerg)
- Ex. 1: the IE model
- DA-RT
- Counter-DA-R1
- On The Run
- Intro to R

Motivating ex.

OLS intuition

- Classic view of quant polisci (Lijphart, Nørgaard)
- Rejoinder: Dahler-Larsen & Sylvest
- The perestroika debate (Laitin, Flyvbjerg)
- Ex. 1: the IE model
- DA-RT
- Counter-DA-RT
- On The Run
- Intro to R

Motivating ex.

OLS intuition

- Classic view of quant polisci (Lijphart, Nørgaard)
- Rejoinder: Dahler-Larsen & Sylvest
- The perestroika debate (Laitin, Flyvbjerg)
- Ex. 1: the IE model
- DA-RT
- Counter-DA-RT
- On The Run
- Intro to R

Motivating ex.

Regression pitfalls

- Classic view of quant polisci (Lijphart, Nørgaard)
- Rejoinder: Dahler-Larsen & Sylvest
- The perestroika debate (Laitin, Flyvbjerg)
- Fx. 1: the IF model.
- DA-RT
- Counter-DA-RT
- On The Run
- Intro to R

Motivating ex.

Motivating example: is the US an oligharchy?

- 1 Motivating example: is the US an oligharchy?
- 2 OLS intuition

Background

OLS intuition

- Intuition
- 3 OLS formal form
 - Regression coefficients: bivariat
 - Regression coefficients: multivariat
- 4 Notation
 - Interpreting output
- 5 Omitted variable bias
- 6 Regression pitfalls
- 7 Gilens & Page
 - Critique of the oligarchy result: Bashi

Regression pitfalls

- 1 Motivating example: is the US an oligharchy?
- 2 OLS intuition

Background

OLS intuition

- Intuition
- 3 OLS formal form
 - Regression coefficients: bivariate
 - Regression coefficients: multivariate
- 4 Notation
 - Interpreting output
- 5 Omitted variable bias
- 6 Regression pitfalls
- 7 Gilens & Page
 - Critique of the oligarchy result: Bashi

- 1 Motivating example: is the US an oligharchy?
- 2 OLS intuition

Background

OLS intuition

- Intuition
- 3 OLS formal form
 - Regression coefficients: bivariate
 - Regression coefficients: multivariate
- 4 Notation
 - Interpreting output
- 5 Omitted variable bias
- 6 Regression pitfalls
- 7 Gilens & Page
 - Critique of the oligarchy result: Bashin

- 1 Motivating example: is the US an oligharchy?
- 2 OLS intuition

Background

OLS intuition

- Intuition
- 3 OLS formal form
 - Regression coefficients: bivariate
 - Regression coefficients: multivariate
- 4 Notation
 - Interpreting output
- 5 Omitted variable bias
- 6 Regression pitfalls
- 7 Gilens & Page
 - Critique of the oligarchy result: Bashi

- 1 Motivating example: is the US an oligharchy?
- 2 OLS intuition
 - Background

OLS intuition

- Intuition
- 3 OLS formal form
 - Regression coefficients: bivariate
 - Regression coefficients: multivariate
- 4 Notation
 - Interpreting output
- 5 Omitted variable bias
- 6 Regression pitfalls
- 7 Gilens & Page
 - Critique of the oligarchy result: Bashin

Motivating ex.

- Motivating example: is the US an oligharchy?
- OLS intuition
 - Background
 - Intuition
- OLS formal form
 - Regression coefficients: bivariate
 - Regression coefficients: multivariate
- Notation
 - Interpreting output
- Omitted variable bias
- Regression pitfalls
- Gilens & Page
 - Critique of the oligarchy result: Bashir

Regression pitfalls

- Motivating example: is the US an oligharchy?

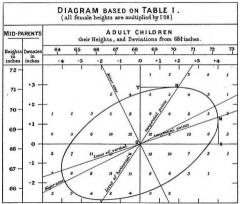
Ezra Klein om Gilens & Page

- 1 Motivating example: is the US an oligharchy?
- 2 OLS intuition
 - Backgroun
 - Intuitior
- 3 OLS formal form
- 4 Notatio
- 5 Omitted variable bia
- 6 Regression pitfall
- 7 Gilens & Page

- 1 Motivating example: is the US an oligharchy
- 2 OLS intuition
 - Background
 - a Intuition
- 3 OLS formal form
- 4 Notation
- 5 Omitted variable bia
- 6 Regression pitfall
- 7 Gilens & Pag

Galton, F. (1886). "Regression towards mediocrity in hereditary stature". The Journal of the Anthropological Institute of Great Britain and Ireland. 15: 246–263

Notation



OLS intuition

OLS formal form

Notation 000000 Omitted variable bias

Regression pitfalls Gilens & Page 00 00000

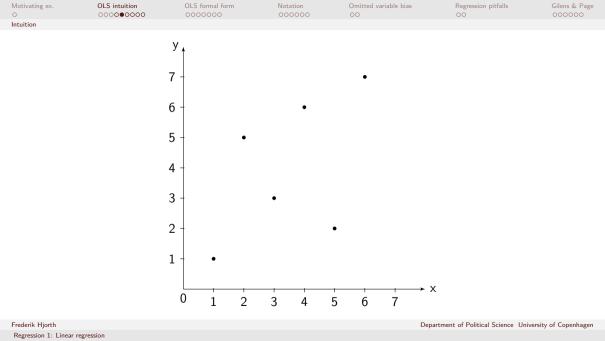
Background



- 1 Motivating example: is the US an oligharchy?
- 2 OLS intuition

Intuition

- Backgroup
- Intuition
- 3 OLS formal form
- 4 Notation
- 5 Omitted variable bia
- 6 Regression pitfal
- 7 Gilens & Pag



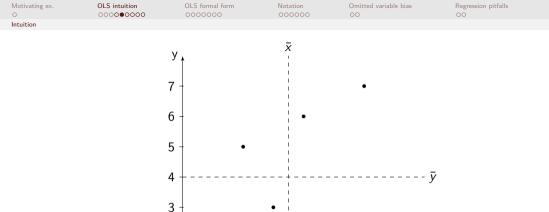
Notation

Omitted variable bias

Gilens & Page

OLS intuition

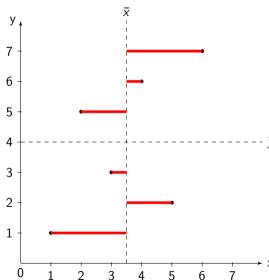
OLS formal form



2

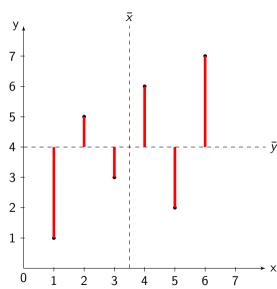
Gilens & Page



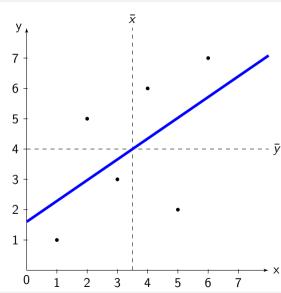


Frederik Hjorth Regression 1: Linear regression









Intuition

Motivating ex.

- Total Sum of Squares (TSS): $\sum_{i=1}^{n} (y_i \bar{y})^2$

OLS intuition

• Total Sum of Squares (TSS): $\sum_{i=1}^{n} (y_i - \bar{y})^2$

TSS consists of two parts:

OLS intuition

000000000

Motivating ex.

Intuition

Intuition

Motivating ex.

- Total Sum of Squares (TSS): $\sum_{i=1}^{n} (y_i \bar{y})^2$
- TSS consists of two parts:

OLS intuition

- Explained Sum of Squares (ESS)

- Total Sum of Squares (TSS): $\sum_{i=1}^{n} (y_i \bar{y})^2$
- TSS consists of two parts:

OLS intuition

- Explained Sum of Squares (ESS)
- Residual Sum of Squares (RSS)

Intuition

Motivating ex.

- Total Sum of Squares (TSS): $\sum_{i=1}^{n} (y_i \bar{y})^2$
- TSS consists of two parts:

OLS intuition

- Explained Sum of Squares (ESS)
- Residual Sum of Squares (RSS)
- TSS = ESS + RSS

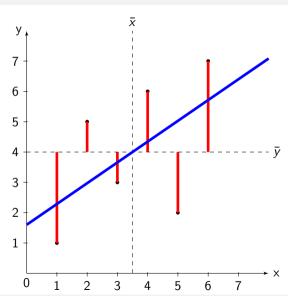
Intuition

Motivating ex.

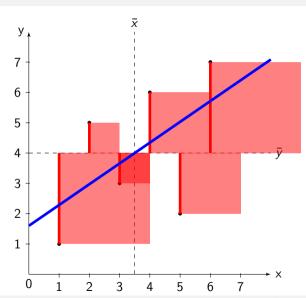
- Total Sum of Squares (TSS): $\sum_{i=1}^{n} (y_i \bar{y})^2$
- TSS consists of two parts:

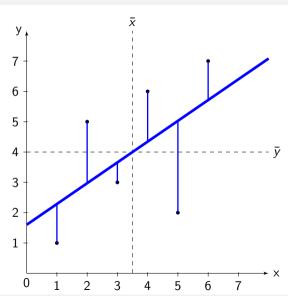
OLS intuition

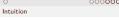
- Explained Sum of Squares (ESS)
- Residual Sum of Squares (RSS)
- TSS = ESS + RSS
- OLS estimates the line that minimizes RSS

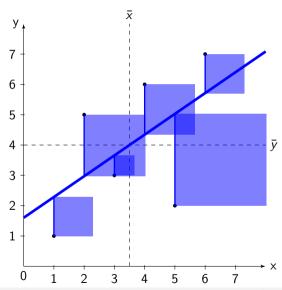


Intuition





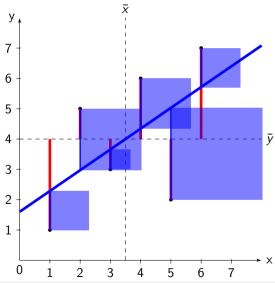




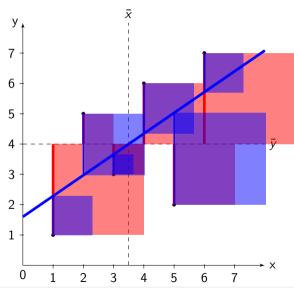
Omitted variable bias

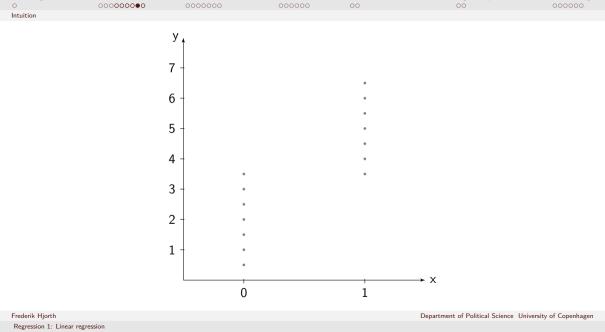
Regression pitfalls Gilens & Page 00 00000











Notation

Omitted variable bias

Regression pitfalls

Gilens & Page

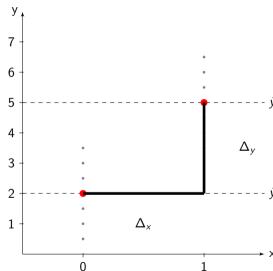
Motivating ex.

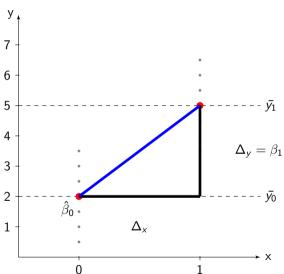
OLS intuition

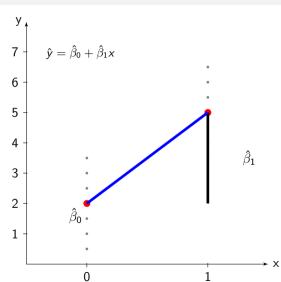
OLS formal form

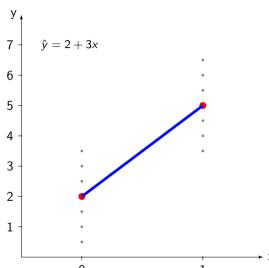
Frederik Hjorth Regression 1: Linear regression Department of Political Science University of Copenhagen

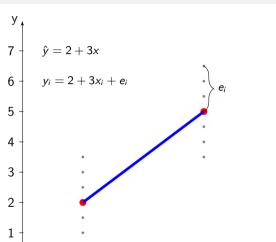
Gilens & Page











- data comes from random sample

Motivating ex.

- data comes from random sample

Intuition

- data comes from random sample
- true relationship is linear
- no perfect multicollinearity
- 4 independent variables are exogenous
- residuals are homoscedastic
- → OLS is Best Linear Unbiased Estimator (Gauss-Markov Theorem

Intuition

- 1 data comes from random sample
- 2 true relationship is linear

- 1 data comes from random sample
- 2 true relationship is linear
- 3 no perfect multicollinearity

Intuition

- 1 data comes from random sample
- 2 true relationship is linear
- 3 no perfect multicollinearity
- 4 independent variables are exogenous

Q: Why do we care? A: if these assumptions hold:

- data comes from random sample
- 2 true relationship is linear
- 3 no perfect multicollinearity
- 4 independent variables are exogenous
- 5 residuals are homoscedastic
- → OLS is Best Linear Unbiased Estimator (Gauss-Markov Theorem

Motivating ex.

Intuition

Q: Why do we care? A: if these assumptions hold:

- data comes from random sample
- 2 true relationship is linear

OLS intuition

00000000

- 3 no perfect multicollinearity
- 4 independent variables are exogenous
- 5 residuals are homoscedastic
- \rightarrow OLS is Best Linear Unbiased Estimator (Gauss-Markov Theorem)

Motivating ex.

Intuition

- 1 Motivating example: is the US an oligharchy?
- 2 OIS intuitio

3 OLS formal form

- Regression coefficients: hivariate
- Regression coefficients: multivariate
- 4 Notatio
- 5 Omitted variable bia
- 6 Regression pitfal
- 7 Gilens & Pag

- 1 Motivating example: is the US an oligharchy?
- 2 OIS intuition
- 3 OLS formal form
 - Regression coefficients: bivariate
 - Regression coefficients: multivariat
- 4 Notation
- 5 Omitted variable bia
- 6 Regression pitfall
- 7 Gilens & Page

Simulated data, simdat:

У	Х
2	1
8	3
11	5
14	6

R exercise

- \rightarrow how would vou call the variable x
- \rightarrow how would you call only the bottom two rows

Simulated data, simdat:

У	Х
2	1
8	3
11	5
14	6

R exercise:

- \rightarrow how would you call the variable x?
- ightarrow how would you call only the bottom two rows

Simulated data. simdat:

У	Х
2	1
8	3
11	5
14	6

R exercise:

- \rightarrow how would you call the variable x?
- \rightarrow how would you call only the bottom two rows?

Motivating ex.

 $b = \beta = \frac{Cov(Y_i, X_i)}{Var(X_i)}$

Motivating ex.

Gilens & Page

(1)

Gilens & Page

Regression coefficient formula (cf. MM p. 86n):

$$b = \beta = \frac{Cov(Y_i, X_i)}{Var(X_i)} \tag{1}$$

i.e.: the covariance btw. X and Y divided by the variance of X, where

$$Cov(Y_i, X_i) = \frac{\sum_{i=1}^{N} (Y_i - \bar{Y})(X_i - \bar{X})}{n - 1}$$
 (2)

Motivating ex.

Regression coefficients: bivariate

$$Var(X_i) = \frac{\sum_{i}^{N} (X_i - \bar{X})^2}{n - 1}$$
 (3)

Motivating ex.

Regression coefficient formula (cf. MM p. 86n):

$$b = \beta = \frac{Cov(Y_i, X_i)}{Var(X_i)} \tag{1}$$

i.e.: the covariance btw. X and Y divided by the variance of X, where

$$Cov(Y_i, X_i) = \frac{\sum_{i=1}^{N} (Y_i - \bar{Y})(X_i - \bar{X})}{n-1}$$
 (2)

$$Var(X_i) = \frac{\sum_{i}^{N} (X_i - \bar{X})^2}{n - 1}$$
 (3)

Motivating ex.

Regression coefficient formula (cf. MM p. 86n):

$$b = \beta = \frac{Cov(Y_i, X_i)}{Var(X_i)} \tag{1}$$

i.e.: the covariance btw. X and Y divided by the variance of X, where

$$Cov(Y_i, X_i) = \frac{\sum_{i=1}^{N} (Y_i - \bar{Y})(X_i - \bar{X})}{n-1}$$
 (2)

and

$$Var(X_i) = \frac{\sum_{i}^{N} (X_i - \bar{X})^2}{n - 1}$$
 (3)

Gilens & Page

Simulated data set, simdat:

У	X
2	1
8	3
11	5
14	6

Motivating ex.

Simulated data set, simdat:

У	X
2	1
8	3
11	5
14	6

 \rightarrow what is $Cov(Y_i, X_i)$?

Motivating ex.

Simulated data set, simdat:

У	X
2	1
8	3
11	5
14	6

- \rightarrow what is $Cov(Y_i, X_i)$?
- \rightarrow what is $Var(X_i)$?

Simulated data set, simdat:

У	X
2	1
8	3
11	5
14	6

- \rightarrow what is $Cov(Y_i, X_i)$?
- \rightarrow what is $Var(X_i)$?
- \rightarrow what is β ?

Motivating ex.

Gilens & Page

- 1 Motivating example: is the US an oligharchy?
- 2 OLS intuition
- 3 OLS formal form
 - Regression coefficients: bivariat
 - Regression coefficients: multivariate
- 4 Notatio
- 5 Omitted variable bia
- 6 Regression pitfall
- 7 Gilens & Page

Motivating ex.

More generally, consider a vector \mathbf{Y} and a matrix of X vectors \mathbf{X} , i.e.

$$Y = \mathbf{X}\widehat{\beta} + \mathbf{e}$$

then OLS solves

$$\arg\min\sum_{i=1}^n e_i^2$$

which implies

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T Y$$

Motivating ex.

More generally, consider a vector \mathbf{Y} and a matrix of X vectors \mathbf{X} , i.e.

$$Y = \mathbf{X}\widehat{eta} + \mathbf{e}$$

then OIS solves

$$\arg\min\sum_{i=1}^n e_i^2$$

which implies

$$\hat{\beta} = (\mathbf{X}^{\mathsf{T}} \mathbf{X})^{-1} \mathbf{X}^{\mathsf{T}} Y$$

Motivating ex.

More generally, consider a vector **Y** and a matrix of X vectors **X**, i.e.

$$Y = \mathbf{X}\widehat{eta} + \mathbf{e}$$

then OLS solves

$$\arg\min\sum_{i=1}^n e_i^2$$

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T Y$$

OLS intuition

Motivating ex.

More generally, consider a vector **Y** and a matrix of X vectors **X**, i.e.

$$Y = \mathbf{X}\widehat{eta} + \mathbf{e}$$

then OLS solves

$$\arg\min\sum_{i=1}^n e_i^2$$

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T Y$$

OLS intuition

Motivating ex.

More generally, consider a vector **Y** and a matrix of X vectors **X**, i.e.

$$Y = \mathbf{X}\widehat{eta} + \mathbf{e}$$

then OLS solves

$$\arg\min\sum_{i=1}^n e_i^2$$

which implies

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T Y$$

OLS intuition

Motivating ex.

More generally, consider a vector **Y** and a matrix of X vectors **X**, i.e.

$$Y = \mathbf{X}\widehat{eta} + \mathbf{e}$$

then OLS solves

$$\arg\min\sum_{i=1}^n e_i^2$$

which implies

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T Y$$

 Motivating ex.
 OLS intuition
 OLS formal form
 Notation
 Omitted variable bias
 Regression pitfalls
 Gilens & Page

 ○
 ○
 ○
 ○
 ○
 ○
 ○
 ○
 ○
 ○
 ○
 ○
 ○

Regression coefficients: multivariate

Questions or comments?

- 1 Motivating example: is the US an oligharchy?
- 2 OLS intuitio
- 3 OLS formal form
- 4 Notation
 - Interpreting output
- 5 Omitted variable bia
- 6 Regression pitfall
- 7 Gilens & Pag

Regression model w treatment variable P_i and control A_i :

$$Y_i = \alpha + \beta P_i + \gamma A_i + e_i \tag{4}$$

Alternative notation: CEF (Conditional Expectation Function)

$$E[Y_i|P_i,A_i] (5)$$

We can express coefficients as differences btw. CE's:

$$E[Y_i|P_i = 1, A_i] - E[Y_i|P_i = 0, A_i] = \beta$$
 (6)

Motivating ex.

By definition, the fitted Y_i , \hat{Y}_i , omits the error term:

$$\widehat{Y}_i = \alpha + \beta P_i + \gamma A_i \tag{7}$$

$$e_i = Y_i - \widehat{Y}_i = Y_i - \alpha + \beta P_i + \gamma A_i \tag{8}$$

OLS intuition

By definition, the fitted Y_i , \hat{Y}_i , omits the error term:

$$\widehat{Y}_i = \alpha + \beta P_i + \gamma A_i \tag{7}$$

Motivating ex.

$$e_i = Y_i - \widehat{Y}_i = Y_i - \alpha + \beta P_i + \gamma A_i \tag{8}$$

OLS intuition

By definition, the fitted Y_i , \hat{Y}_i , omits the error term:

$$\widehat{Y}_i = \alpha + \beta P_i + \gamma A_i \tag{7}$$

Hence.

Motivating ex-

$$e_i = Y_i - \widehat{Y}_i = Y_i - \alpha + \beta P_i + \gamma A_i \tag{8}$$

Why do we get e_i there anyway?

OLS intuition

- omitted variables
- measurement error
- fundamental random variation (MM: 'serendipitous variation')

OLS intuition

Controls can also be categorical (e.g. combinations of accepting schools) or continuous (e.g. SAT score)

$$ln(Y_i) = \alpha + \beta P_i + \sum_{j=1}^{150} \gamma_j GROUP_{ji} + \delta_1 SAT_i + \delta_2 PI_i + e_i \qquad (9)$$

OLS intuition

$$Y_i = \alpha + \sum_{k=1}^K \beta_k X_{ki} + \gamma A_i + e_i$$
 (10)

$$SE(\widehat{\beta}_k) = \frac{\sigma_e}{\sqrt{n}} \times \frac{1}{\sigma_{\tilde{\chi}_k}}$$
 (11)

Implication: to get smaller s.e.'s (ightarrow precise estimates), you need

- $\downarrow \sigma_e$ and/or
- ↑ n and/or
- ↑ σ_X

OLS intuition

$$Y_i = \alpha + \sum_{k=1}^K \beta_k X_{ki} + \gamma A_i + e_i$$
 (10)

$$SE(\widehat{\beta}_k) = \frac{\sigma_e}{\sqrt{n}} \times \frac{1}{\sigma_{\tilde{X}_k}}$$
 (11)

OLS intuition

$$Y_i = \alpha + \sum_{k=1}^K \beta_k X_{ki} + \gamma A_i + e_i$$
 (10)

$$SE(\widehat{\beta}_k) = \frac{\sigma_e}{\sqrt{n}} \times \frac{1}{\sigma_{\tilde{\chi}_k}}$$
 (11)

Implication: to get smaller s.e.'s (\rightarrow precise estimates), you need

$$Y_i = \alpha + \sum_{k=1}^K \beta_k X_{ki} + \gamma A_i + e_i$$
 (10)

$$SE(\widehat{\beta}_k) = \frac{\sigma_e}{\sqrt{n}} \times \frac{1}{\sigma_{\tilde{\chi}_k}}$$
 (11)

Implication: to get smaller s.e.'s (\rightarrow precise estimates), you need

- $\downarrow \sigma_e$ and/or

$$Y_i = \alpha + \sum_{k=1}^K \beta_k X_{ki} + \gamma A_i + e_i$$
 (10)

$$SE(\widehat{\beta}_k) = \frac{\sigma_e}{\sqrt{n}} \times \frac{1}{\sigma_{\tilde{X}_k}}$$
 (11)

Implication: to get smaller s.e.'s (\rightarrow precise estimates), you need

- $\downarrow \sigma_e$ and/or
- $\uparrow n$ and/or

$$Y_i = \alpha + \sum_{k=1}^K \beta_k X_{ki} + \gamma A_i + e_i$$
 (10)

$$SE(\widehat{\beta}_k) = \frac{\sigma_e}{\sqrt{n}} \times \frac{1}{\sigma_{\tilde{X}_k}}$$
 (11)

Implication: to get smaller s.e.'s (\rightarrow precise estimates), you need

- $\downarrow \sigma_e$ and/or
- $\uparrow n$ and/or
- $\uparrow \sigma_{\tilde{\mathbf{x}}}$

- 1 Motivating example: is the US an oligharchy?
- 2 OLS intuitio
- 3 OLS formal form
- 4 Notation
 - Interpreting output
- 5 Omitted variable bia
- 6 Regression pitfall
- 7 Gilens & Pag

Regression 63

TABLE 2.2
Private school effects: Barron's matches

	No selection controls			Selection controls		
	(1)	(2)	(3)	(4)	(5)	(6)
Private school	.135 (.055)	.095 (.052)	.086 (.034)	.007 (.038)	.003 (.039)	.013
Own SAT score ÷ 100		.048 (.009)	.016 (.007)		.033 (.007)	.001 (.007)
Log parental income			.219			.190

(.022)

- 1 Motivating example: is the US an oligharchy?
- 2 OLS intuition
- 3 OLS formal form
- 4 Notatio
- 5 Omitted variable bias
- 6 Regression pitfall
- 7 Gilens & Pag

Regression pitfalls

$$Y_i = \alpha' + \beta' P_i + \gamma A_i + e_i' \tag{12}$$

$$Y_i = \alpha^s + \beta^s P_i + e_i^s \tag{13}$$

Regression pitfalls

Short vs. long form:

$$Y_i = \alpha^l + \beta^l P_i + \gamma A_i + e_i^l \tag{12}$$

$$Y_i = \alpha^s + \beta^s P_i + e_i^s \tag{13}$$

Regression pitfalls

Short vs. long form:

$$Y_i = \alpha^l + \beta^l P_i + \gamma A_i + e_i^l \tag{12}$$

$$Y_i = \alpha^s + \beta^s P_i + e_i^s \tag{13}$$

OLS intuition

Short vs. long form:

$$Y_i = \alpha' + \beta' P_i + \gamma A_i + e_i' \tag{12}$$

•0

$$Y_i = \alpha^s + \beta^s P_i + e_i^s \tag{13}$$

 \rightarrow how different are β^I og β^I ?

$$\beta^{s} - \beta^{l} = \pi_{1} \times \gamma \tag{14}$$

where T is the coefficient of P. on A.

$$A_i = \pi_0 + \pi_1 P_i + u_i \tag{15}$$

$$\beta^s - \beta^l = \pi_1 \times \gamma \tag{14}$$

where π_1 is the coefficient of P_i on A_i :

$$A_i = \pi_0 + \pi_1 P_i + u_i \tag{15}$$

- 1 Motivating example: is the US an oligharchy?
- 2 OLS intuitio
- 3 OLS formal form
- 4 Notatio
- 5 Omitted variable bia
- 6 Regression pitfalls
- 7 Gilens & Pag

OLS intuition

- omitted variable bias (cf. above)
- controlling for post-treatment variables
- Outliers
- 4 multicollinearity
- non-linear functional form

Motivating ex.

- omitted variable bias (cf. above)
- 2 controlling for post-treatment variables
- Outliers
- 4 multicollinearity
- non-linear functional form

OLS intuition

- omitted variable bias (cf. above)
- 2 controlling for post-treatment variables
- Outliers
- 4 multicollinearity
- non-linear functional form

OLS intuition

- omitted variable bias (cf. above)
- 2 controlling for post-treatment variables
- outliers
- 4 multicollinearity
- non-linear functional form

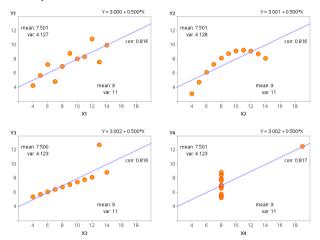
OLS intuition

- omitted variable bias (cf. above)
- 2 controlling for post-treatment variables
- Outliers
- multicollinearity
- 6 non-linear functional form

OLS intuition

- omitted variable bias (cf. above)
- 2 controlling for post-treatment variables
- outliers
- 4 multicollinearity
- 6 non-linear functional form

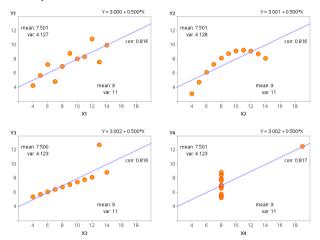
On 3-5: cf. Anscombe's Quartet



→ always look at the data!

Motivating ex.

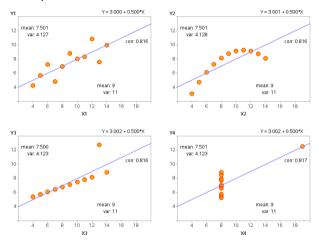
On 3-5: cf. Anscombe's Quartet



→ always look at the data!

Motivating ex.

On 3-5: cf. Anscombe's Quartet



 \rightarrow always look at the data!

Motivating ex.

- 1 Motivating example: is the US an oligharchy?
- 2 OLS intuitio
- 3 OLS formal form
- 4 Notation
- 5 Omitted variable bia
- 6 Regression pitfall
- 7 Gilens & Page
 - Critique of the oligarchy result: Bashi

- Majoritarian Electoral Democracy
- Economic-Elite Domination
- Majoritarian Pluralism
- A Biased Pluralism

Motivating ex.

- Majoritarian Electoral Democracy
- Majoritarian Pluralism

OLS intuition

- Majoritarian Electoral Democracy
- Economic-Elite Domination
- Majoritarian Pluralism

OLS intuition

- Majoritarian Electoral Democracy
- Economic-Elite Domination
- Majoritarian Pluralism

OLS intuition

- Majoritarian Electoral Democracy
- Economic-Elite Domination
- Majoritarian Pluralism

OLS intuition

Biased Pluralism

Table 1
Theoretical predictions concerning the independent influence of sets of actors upon policy outcomes

	Sets of Actors						
Theory (ideal type)	Average Citizens	Economic Elites	All Interest Groups	Mass Interest Groups	Business Interest Groups		
Majoritarian Electoral Democracy	Υ	n	n	n	n		
Dominance by Economic Elites	У	Y	у	n	У		
Majoritarian Pluralism	У	n	Υ	Y	Υ		
Biased Pluralism	ń	n	У	у	Υ		

n = little or no independent influence

OLS intuition

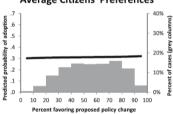
y = some independent influence

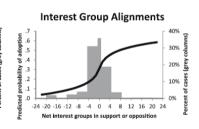
Y = substantial independent influence

Figure 1 Predicted probability of policy adoption (dark lines, left axes) by policy disposition; the distribution of preferences (gray columns, right axes)

ight axes)

Average Citizens' Preferences





Critique of the oligarchy result: Bashir

- Gilens & Page
 - Critique of the oligarchy result: Bashir

OLS intuition OLS formal form Notation Omitted variable bias 00000000 000000 000000 00

Critique of the oligarchy result: Bashir

Motivating ex.

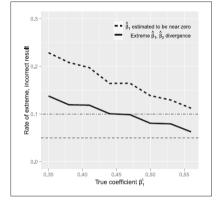


Figure 2. Simulation results. Whether one adopts a 90% or 95% confidence standard (horizontal lines), the authors' method too often produces extreme but erroneous results. The method frequently yields a value of $\hat{\beta}_1$ near zero even when β_1^1 , the true coefficient value (x-axis), is much higher. The two ways to interpret the study's main result are listed in the legend. Near or "essentially zero" is defined by the authors to be $\hat{\beta}_1 \approx 0.05$. "Extreme $\hat{\beta}_1$, $\hat{\beta}_2$ indergence" entails both $\hat{\beta}_1 \approx 0.03$ and $\hat{\beta}_2 \approx 0.76$, with the latter significant at the 99.9% level.

Regression pitfalls

Omitted variable bias

Notation

Break for lunch

Frederik Hjorth

Motivating ex.

OLS intuition

OLS formal form

Department of Political Science University of Copenhagen

Regression pitfalls