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

Themes covered yesterday:

- 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

Themes covered yesterday:

- 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

Themes covered yesterday:

- 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

Themes covered yesterday:

- 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

Themes covered yesterday:

- 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

Themes covered yesterday:

- 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

Themes covered yesterday:

- 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

Themes covered yesterday:

- 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

Themes covered yesterday:

- 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

1 Motivating example: is the US an oligharchy?

2 OLS intuition

3 OLS formal form

4 Notation

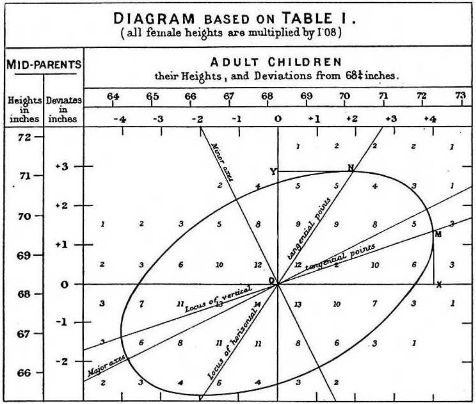
5 Omitted variable bias

6 Regression pitfalls

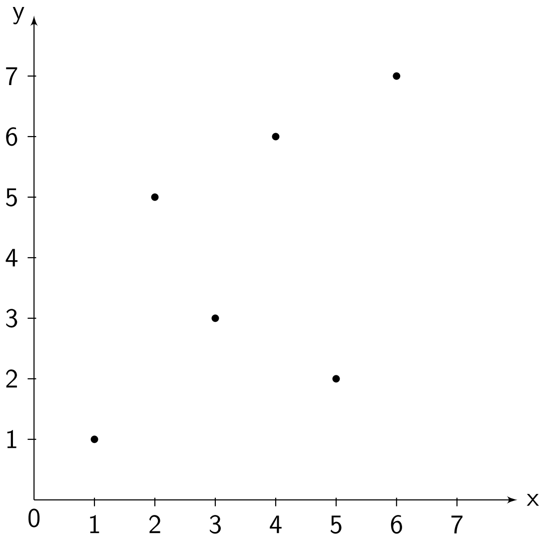
7 Gilens & Page

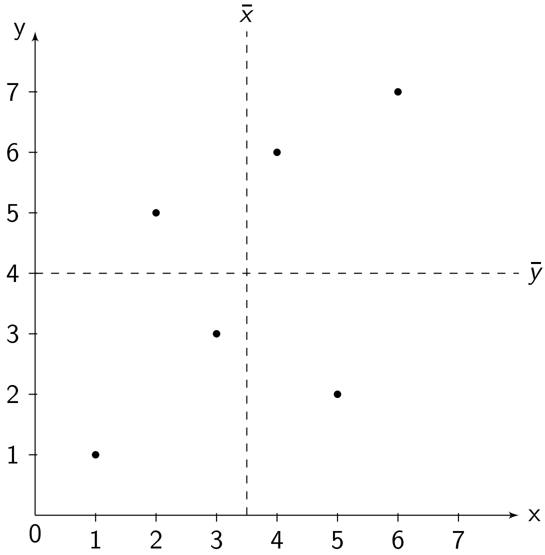
Ezra Klein om Gilens & Page

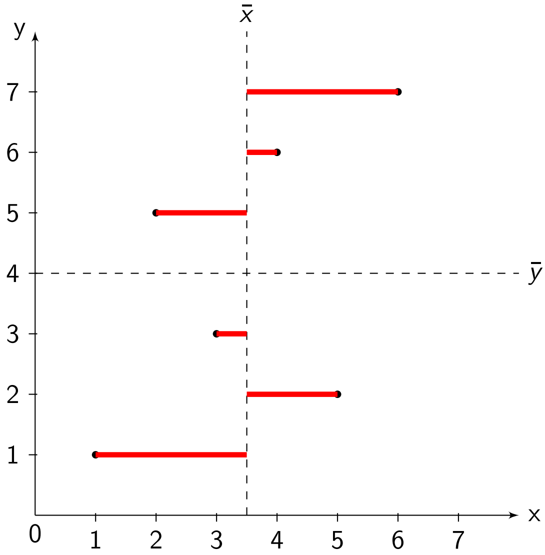
Galton, F. (1886). “Regression towards mediocrity in hereditary stature”. *The Journal of the Anthropological Institute of Great Britain and Ireland*. 15: 246–263

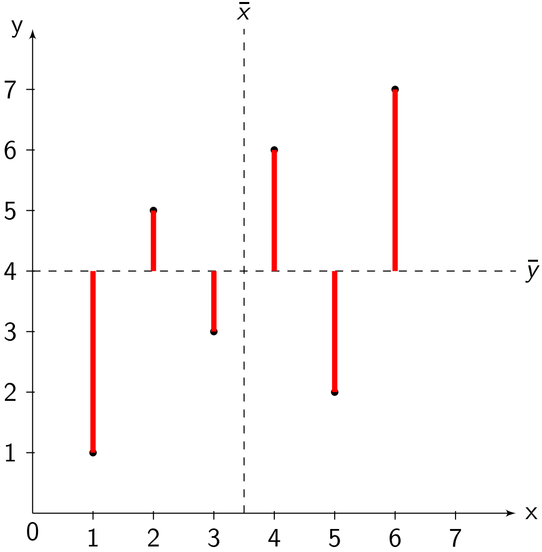


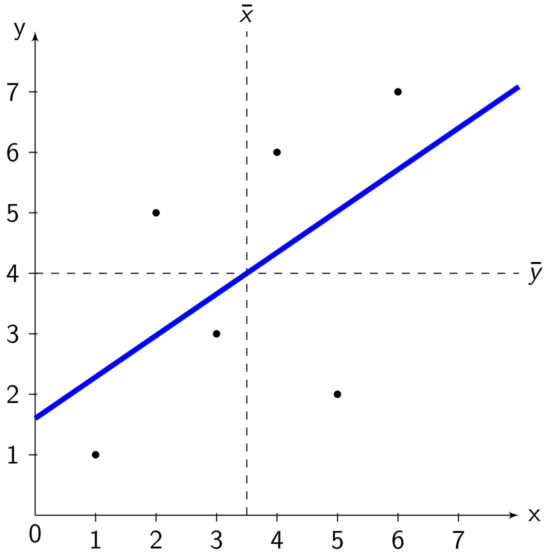

FROM THE DIRECTOR OF THE OTHERS
FEAR ALWAYS FINDS ITS VICTIM
ETHAN HAWKE
EMMA WATSON
REGRESSION
A FILM BY ALEJANDRO AMENÁBAR
COMING SOON


y
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
x


$\bar{x}$
y
7
6
5
4
3
2
1
0
1 2 3 4 5 6 7
x
$\bar{y}$


$\bar{x}$
y
7
6
5
4
3
2
1
0
1 2 3 4 5 6 7
x
$\bar{y}$


$\bar{x}$
y
7
6
5
4
3
2
1
$\bar{y}$
x
0
1
2
3
4
5
6
7


$\bar{x}$
y
7
6
5
4
3
2
1
$\bar{y}$
0
1
2
3
4
5
6
7
x

- Total Sum of Squares (TSS): $\sum_{i=1}^{n}(y_i - \bar{y})^2$
- TSS consists of two parts:
  - Explained Sum of Squares (ESS)
  - Residual Sum of Squares (RSS)
- $TSS = ESS + RSS$
- OLS estimates the line that minimizes RSS

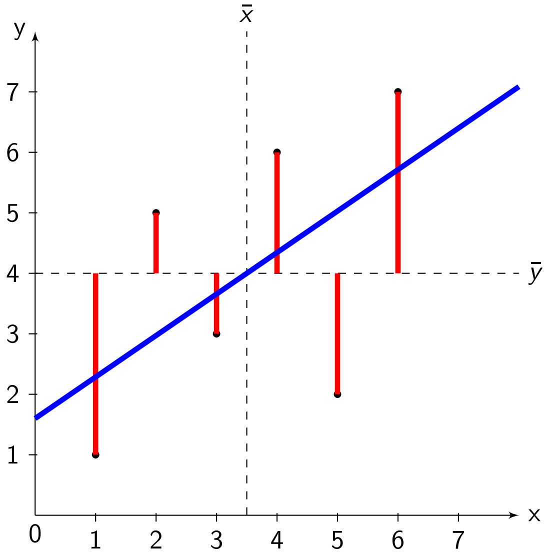
- Total Sum of Squares (TSS): $\sum_{i=1}^{n}(y_i - \bar{y})^2$
- TSS consists of two parts:
  - Explained Sum of Squares (ESS)
  - Residual Sum of Squares (RSS)
- $TSS = ESS + RSS$
- OLS estimates the line that minimizes RSS

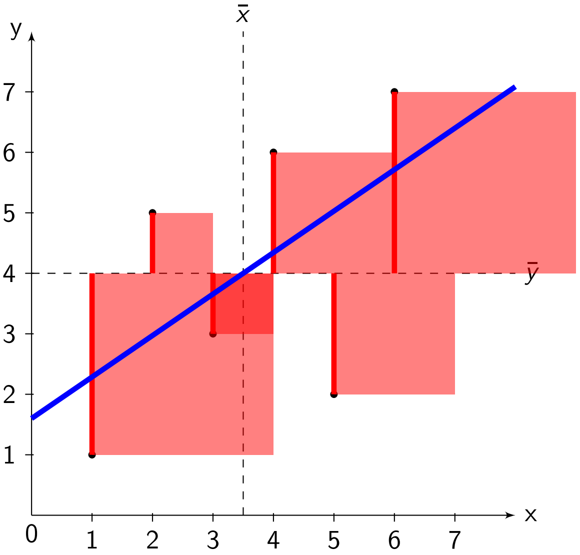
- Total Sum of Squares (TSS): $\sum_{i=1}^{n}(y_i - \bar{y})^2$
- TSS consists of two parts:
  - Explained Sum of Squares (ESS)
  - Residual Sum of Squares (RSS)
- $TSS = ESS + RSS$
- OLS estimates the line that minimizes RSS

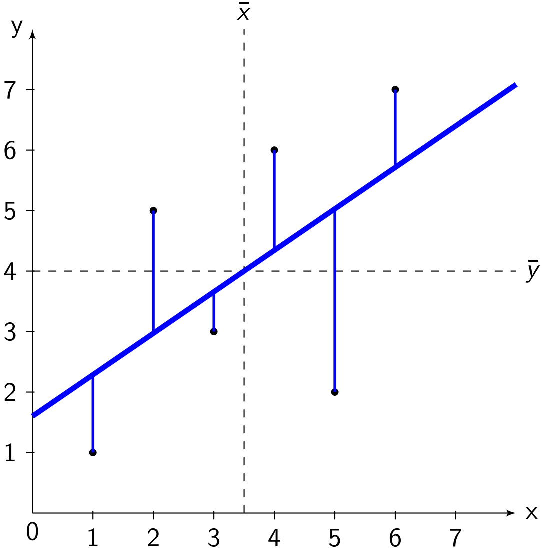
- Total Sum of Squares (TSS): $\sum_{i=1}^{n}(y_i - \bar{y})^2$
- TSS consists of two parts:
  - Explained Sum of Squares (ESS)
  - Residual Sum of Squares (RSS)
- $TSS = ESS + RSS$
- OLS estimates the line that minimizes RSS

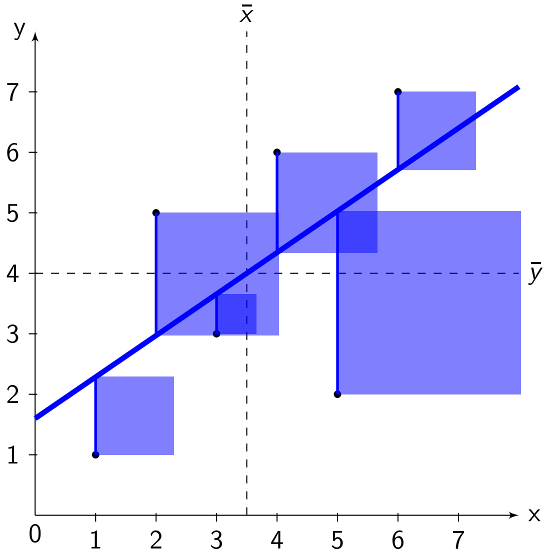
- Total Sum of Squares (TSS): $\sum_{i=1}^{n}(y_i - \bar{y})^2$
- TSS consists of two parts:
  - Explained Sum of Squares (ESS)
  - Residual Sum of Squares (RSS)
- $TSS = ESS + RSS$
- OLS estimates the line that minimizes RSS

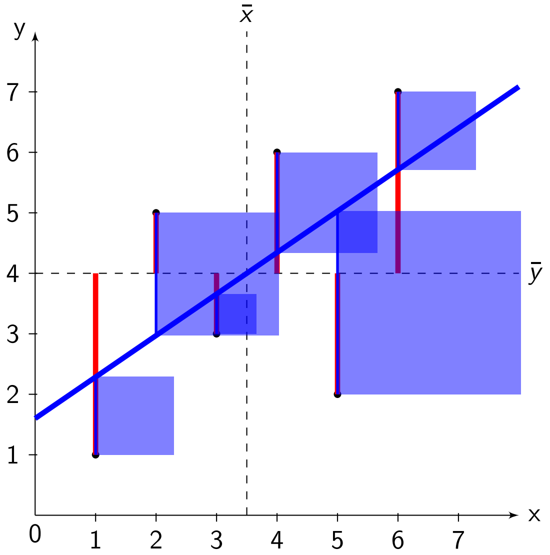
- Total Sum of Squares (TSS): $\sum_{i=1}^{n}(y_i - \bar{y})^2$
- TSS consists of two parts:
  - Explained Sum of Squares (ESS)
  - Residual Sum of Squares (RSS)
- $TSS = ESS + RSS$
- OLS estimates the line that minimizes RSS

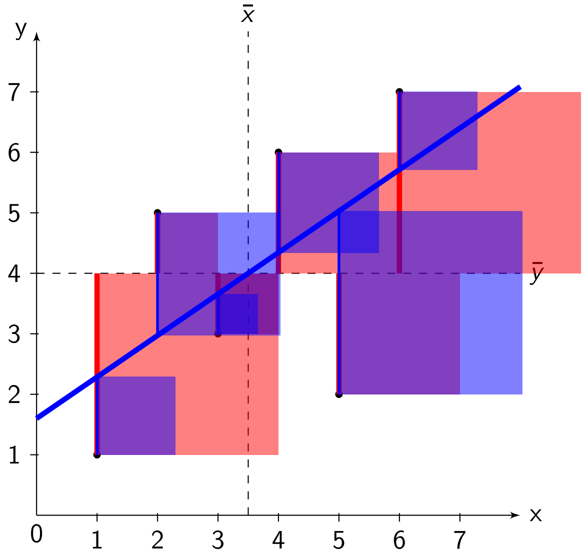

y
x
$\bar{x}$
$\bar{y}$
0
1
2
3
4
5
6
7

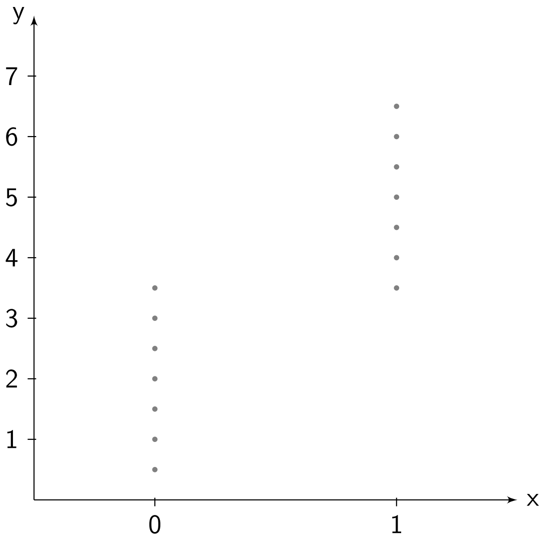

y
x
$\bar{x}$
$\bar{y}$
0
1
2
3
4
5
6
7
1
2
3
4
5
6
7

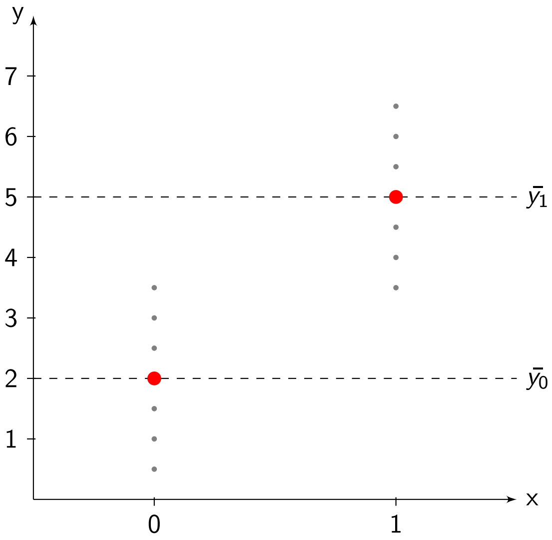

$\bar{x}$
y
7
6
5
4
3
2
1
$\bar{y}$
0
1
2
3
4
5
6
7
x

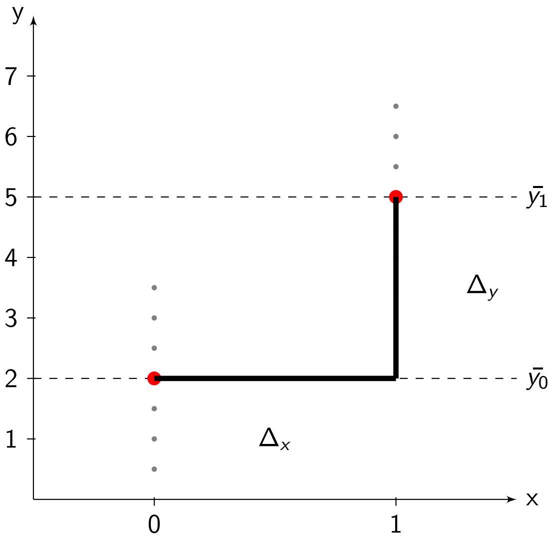

$\bar{x}$
y
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
x
$\bar{y}$

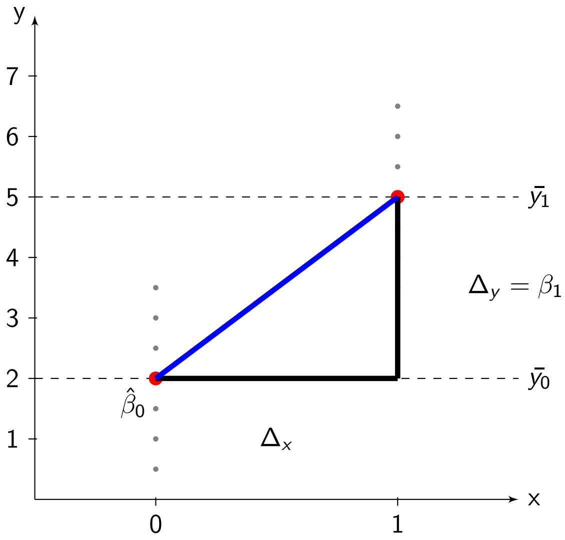

y
x
$\bar{x}$
$\bar{y}$
0
1
2
3
4
5
6
7
1
2
3
4
5
6
7

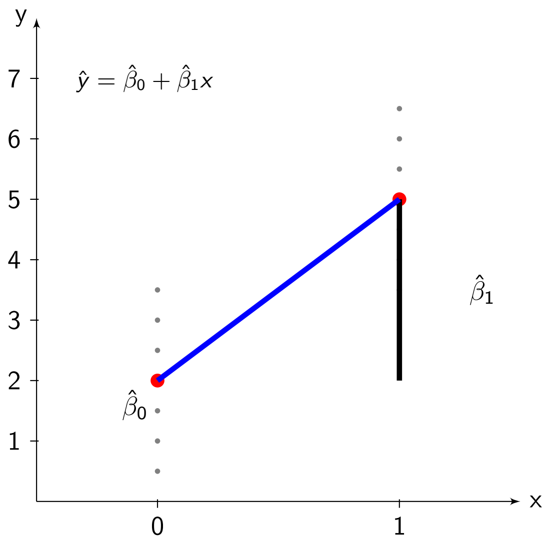

y
x
$\bar{x}$
$\bar{y}$
0
1
2
3
4
5
6
7
1
2
3
4
5
6
7

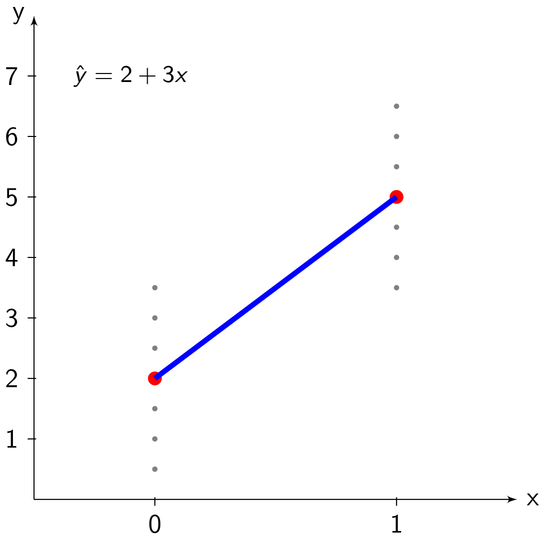

y
7
6
5
4
3
2
1
0
1
x

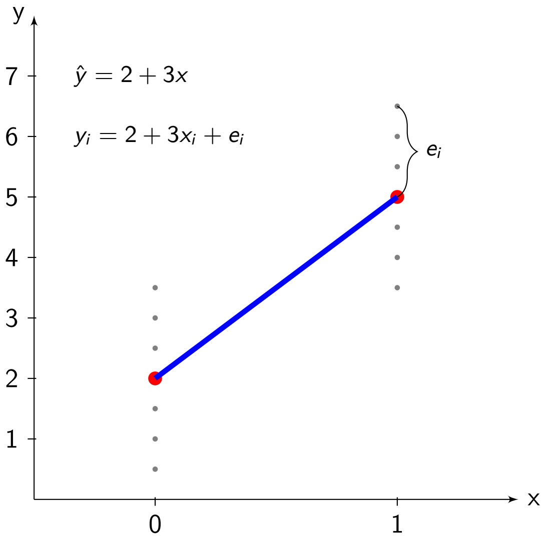

y
7
6
5
4
3
2
1
$\bar{y}_1$
$\bar{y}_0$
x
0
1


y
7
6
5
4
3
2
1
0
1
x
$\bar{y}_1$
$\bar{y}_0$
Δ_y
Δ_x


y
7
6
5
4
3
2
1
0
1
x
$\bar{y}_1$
$\bar{y}_0$
$\Delta_y = \beta_1$
$\hat{\beta}_0$
Δ_x


y
7
6
5
4
3
2
1
$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$
$\hat{\beta}_1$
$\hat{\beta}_0$
0
1
x


y
7
6
5
4
3
2
1
$\hat{y} = 2 + 3x$
0
1
x


y
7
6
5
4
3
2
1
0
1
x
$\hat{y} = 2 + 3x$
$y_i = 2 + 3x_i + e_i$
e_i

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

1. 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)

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

1. 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)

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

1. 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)

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

1. 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)

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

1. 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)

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

1. 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)

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

1. 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)

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

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

$\rightarrow$ OLS is Best Linear Unbiased Estimator (Gauss-Markov Theorem)

Simulated data, `simdat`:

y	x
2	1
8	3
11	5
14	6

R exercise:

→ how would you call the variable `x`?
→ how would you call only the bottom two rows?

Simulated data, `simdat`:

y	x
2	1
8	3
11	5
14	6

R exercise:

→ how would you call the variable `x`?
→ how would you call only the bottom two rows?

Simulated data, `simdat`:

y	x
2	1
8	3
11	5
14	6

R exercise:

→ how would you call the variable `x`?
→ how would you call only the bottom two rows?

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

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

and

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

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

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

and

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

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

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

and

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

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

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

and

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

Simulated data set, `simdat`:

y	x
2	1
8	3
11	5
14	6

→ what is $Cov(Y_i, X_i)$?
→ what is $Var(X_i)$?
→ what is β?

Simulated data set, `simdat`:

y	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?

Simulated data set, `simdat`:

y	x
2	1
8	3
11	5
14	6

→ what is $Cov(Y_i, X_i)$?
→ what is $Var(X_i)$?
→ what is β?

Simulated data set, `simdat`:

y	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?

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

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

then OLS solves

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

which implies

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

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

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

then OLS solves

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

which implies

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

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

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

then OLS solves

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

which implies

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

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

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

then OLS solves

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

which implies

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

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

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

then OLS solves

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

which implies

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

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

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

then OLS solves

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

which implies

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

Questions or comments?

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)

$$E[Y_i | P_i, A_i] \tag{5}$$

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

$$E[Y_i | P_i = 1, A_i] - E[Y_i | P_i = 0, A_i] = \beta \tag{6}$$

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

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

Hence,

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

Why do we get e_i there anyway?

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

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

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

Hence,

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

Why do we get e_i there anyway?

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

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

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

Hence,

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

Why do we get e_i there anyway?

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

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

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

Standard errors in model with K coefficients:

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

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

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

- $\downarrow \sigma_e$ and/or
- $\uparrow n$ and/or
- $\uparrow \sigma_{\tilde{X}_k}$

Standard errors in model with K coefficients:

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

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

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

- $\downarrow \sigma_e$ and/or
- $\uparrow n$ and/or
- $\uparrow \sigma_{\tilde{X}_k}$

Standard errors in model with K coefficients:

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

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

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

- $\downarrow \sigma_e$ and/or
- $\uparrow n$ and/or
- $\uparrow \sigma_{\tilde{X}_k}$

Standard errors in model with K coefficients:

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

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

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

- $\downarrow \sigma_e$ and/or
- $\uparrow n$ and/or
- $\uparrow \sigma_{\tilde{X}_k}$

Standard errors in model with K coefficients:

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

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

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

- $\downarrow \sigma_e$ and/or
- $\uparrow n$ and/or
- $\uparrow \sigma_{\tilde{X}_k}$

Standard errors in model with K coefficients:

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

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

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

- $\downarrow \sigma_e$ and/or
- $\uparrow n$ and/or
- $\uparrow \sigma_{\tilde{X}_k}$

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	.095	.086	.007	.003	.013
	(.055)	(.052)	(.034)	(.038)	(.039)	(.025)
Own SAT score ÷ 100		.048	.016		.033	.001
		(.009)	(.007)		(.007)	(.007)
Log parental income			.219			.190
			(.022)			(.023)

Short vs. long form:

$$Y_i = \alpha^l + \beta^l P_i + \gamma A_i + e_i^l \quad (12)$$

$$Y_i = \alpha^s + \beta^s P_i + e_i^s \quad (13)$$

$\rightarrow$ how different are β^l og β^l?

Short vs. long form:

$$Y_i = \alpha^l + \beta^l P_i + \gamma A_i + e_i^l \qquad (12)$$

$$Y_i = \alpha^s + \beta^s P_i + e_i^s \qquad (13)$$

$\rightarrow$ how different are β^l og β^l?

Short vs. long form:

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

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

$\rightarrow$ how different are β^l og β^l?

Short vs. long form:

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

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

$\rightarrow$ how different are β^l og β^l?

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

where π_1 is the coefficient of P_i on A_i:

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

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

where π_1 is the coefficient of P_i on A_i:

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

Typical regression pitfalls:

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

Typical regression pitfalls:

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

Typical regression pitfalls:

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

Typical regression pitfalls:

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

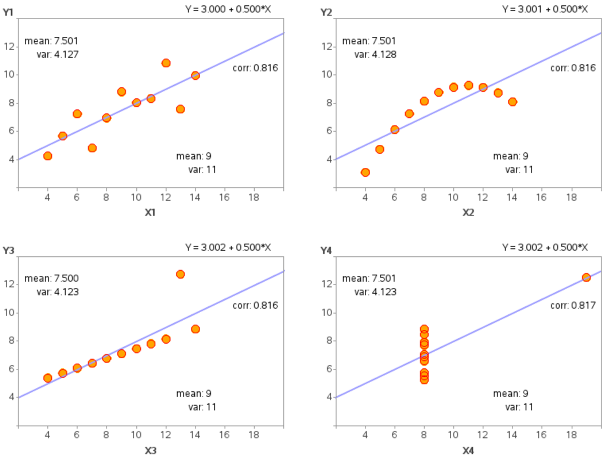
Typical regression pitfalls:

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

Typical regression pitfalls:

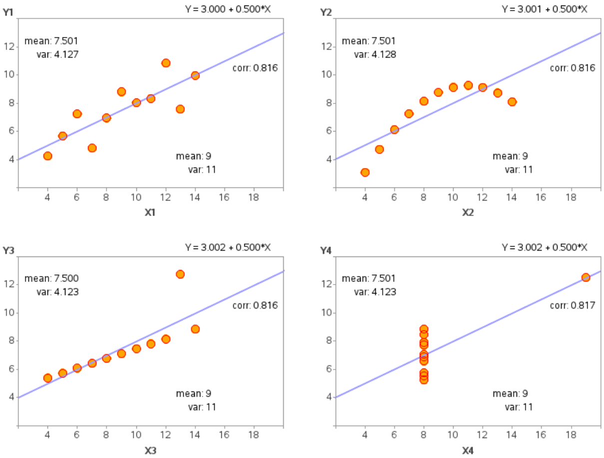
1. omitted variable bias (cf. above)
2. controlling for post-treatment variables
3. outliers
4. multicollinearity
5. non-linear functional form

On 3-5: cf. *Anscombe's Quartet*



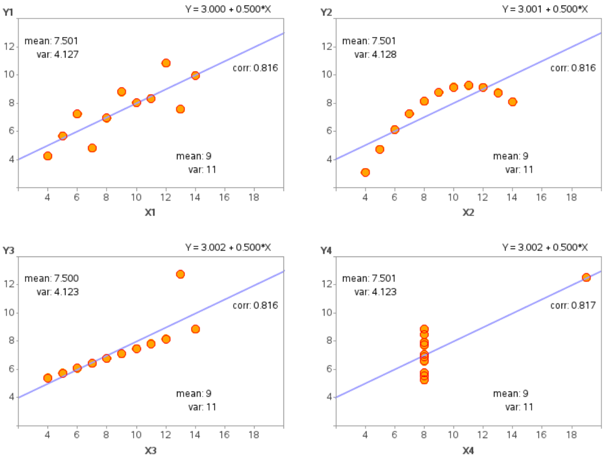
→ always look at the data!

On 3-5: cf. *Anscombe's Quartet*



→ always look at the data!

On 3-5: cf. *Anscombe's Quartet*



→ always look at the data!

In Gilens & Page, four traditions in democratic theory:

1. Majoritarian Electoral Democracy
2. Economic-Elite Domination
3. Majoritarian Pluralism
4. Biased Pluralism

In Gilens & Page, four traditions in democratic theory:

1. Majoritarian Electoral Democracy
2. Economic-Elite Domination
3. Majoritarian Pluralism
4. Biased Pluralism

In Gilens & Page, four traditions in democratic theory:

1. Majoritarian Electoral Democracy
2. Economic-Elite Domination
3. Majoritarian Pluralism
4. Biased Pluralism

In Gilens & Page, four traditions in democratic theory:

1. Majoritarian Electoral Democracy
2. Economic-Elite Domination
3. Majoritarian Pluralism
4. Biased Pluralism

In Gilens & Page, four traditions in democratic theory:

1. Majoritarian Electoral Democracy
2. Economic-Elite Domination
3. Majoritarian Pluralism
4. Biased Pluralism

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

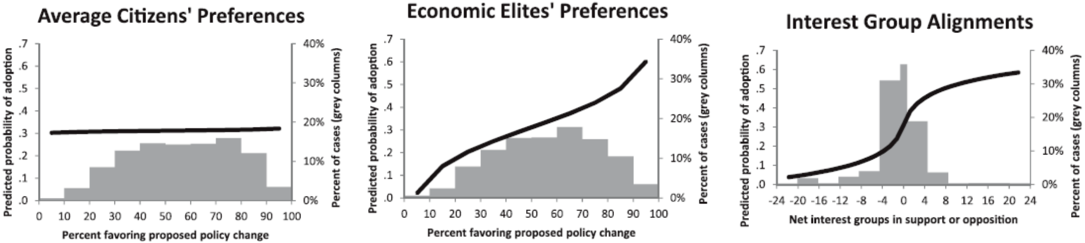
Theory (ideal type)	Sets of Actors: Average Citizens	Economic Elites	All Interest Groups	Mass Interest Groups	Business Interest Groups
Majoritarian Electoral Democracy	Y	n	n	n	n
Dominance by Economic Elites	y	Y	y	n	y
Majoritarian Pluralism	y	n	Y	Y	Y
Biased Pluralism	n	n	y	y	Y

n = little or no independent influence

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)



Critique of the oligarchy result: Bashir

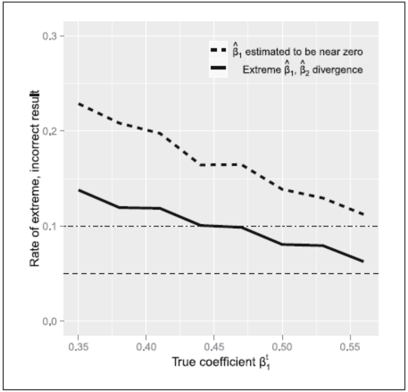


Figure 2. Simulation results. Whether one adopts a 90% or 95% confidence standard (horizontal lines), the authors' method too often produces extreme but erroneous results. The method frequently yields a value of $\hat{\beta}_1$ near zero even when β_1^t, the true coefficient value (*x* -axis), is much 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 $\hat{\beta}_1 \leqslant 0.05$. "Extreme $\hat{\beta}_1$, $\hat{\beta}_2$ divergence" entails both $\hat{\beta}_1 \leqslant 0.03$ and $\hat{\beta}_2 \geqslant 0.76$, with the latter significant at the 99.9% level.

Break for lunch