

450A Political Methodology I

Lecture 5: Regression with Two Regressors

1. Mechanics

Stanford University

Motivation for Multiple Regression

Two Cases

Mechanics of Multiple Regression

Why Do We Want More Than One Predictor?

- Summarize more information for descriptive inference
- Improve the fit and predictive power of our model
- Control for confounding factors for causal inference
- Model more complex non-linearities (e.g. $Y = \beta_0 + \beta_1 X + \beta_2 X^2 + u$)
- Model interactive effects (e.g. $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + u$)

Cigarette Smokers and Cigar or Pipe Smokers

Consider a stylized example based on Cochran (1968), who compared death rates of U.S. men who smoked only cigarettes with those who smoked cigars or pipes. Suppose we run a regression using:

- Y : Deaths per 1,000 Person-Years.
- X_1 : 0 if person smokes cigars or pipes; 1 if person smokes only cigarettes

We fit the regression and find:

$$\widehat{\text{Death Rate}} = 17 - 4 \text{ Cigarette Smoker}$$

What do we conclude?

- The average death rate is 17 deaths per 1,000 person-years for cigar or pipe smokers and 13 for cigarette smokers.
- A naive reading: cigarette smoking lowers the death rate by 4 deaths per 1,000 person-years.

When we “control” for age (in decades above age 55) we find:

$$\widehat{\text{Death Rate}} = 14 + 4 \text{ Cigarette Smoker} + 10 \text{ Age}$$

Why did the sign switch?

(Cochran's crude U.S. rates are 13.5 and 17.4. He adjusted for age by comparing men of similar age, not by regression; the second equation is illustrative.)

Parameters, Estimators, and Estimates

- $E[Y|X_1, X_2, \dots, X_k]$ is the **parameter(s)** of interest (i.e. the mean of the outcome variable conditional on the explanatory variables)
- $\hat{E}[Y|X_1, X_2, \dots, X_k]$ is the **estimator** of the parameter of interest.
- In linear regression models, this estimate will be a function:
 - A **line** in cases with a single X variable
 - A **plane** in cases with two X variables
 - A **hyperplane** in cases with more than two X variables
- We'll start with two simple cases:
 - Regression with one continuous and one dummy variable
 - Regression with two continuous variables

Motivation for Multiple Regression

Two Cases

Multiple Regression with a Binary Covariate

Multiple Regression with a Continuous Covariate

Mechanics of Multiple Regression

Regression with Two Explanatory Variables

Example: executive corruption and court independence, 168 countries in 2024.

- Variables of interest:
 - Y : Executive corruption: how often members of the executive take bribes or steal public funds (V-Dem index times 100: 0 = clean, 100 = most corrupt)
 - X_1 : High court independence: how often the top court rules independently of the government, from 0 = never to 4 = always (rated by country experts)
 - X_2 : Democracy (binary: V-Dem's electoral democracy index above 0.5) or income ($\log_{10}$ GDP per capita, continuous)
- With one predictor we ask: Does court independence (X_1) predict the level of corruption (Y)?
- With two predictors we ask questions like: Does court independence (X_1) predict the level of corruption (Y), once we “control” for democracy or income (X_2)?
- The rest of this lecture is designed to explain what is meant by “controlling for another variable” with linear regression.

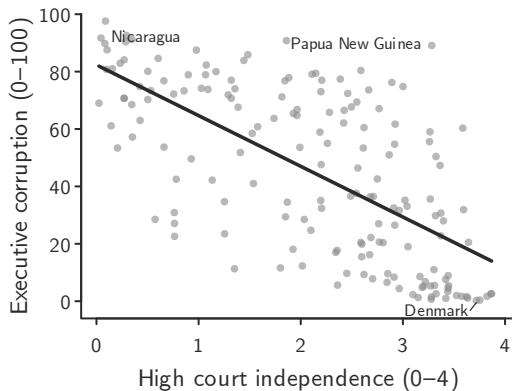
Source: V-Dem v16, 2024. World Bank WDI, 2024: GDP per capita, PPP.

Simple Regression of Corruption on Court Independence

- Let's look at the bivariate regression of Corruption on Court independence:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

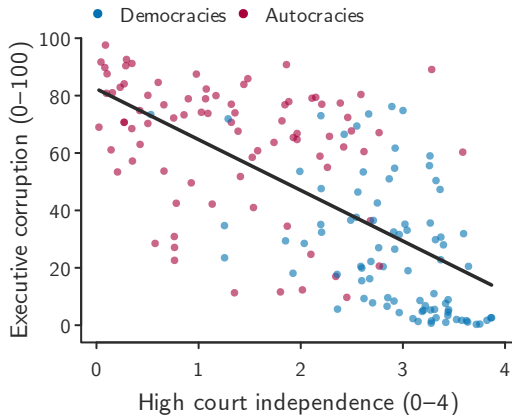
$$\widehat{Corr} = 82.3 - 17.7 \text{ Court}$$



Interpretation: Countries whose top court is rated one point more independent have, on average, 17.7 points less executive corruption. This compares countries. It does not show that independent courts reduce corruption.

Simple Regression of Corruption on Court Independence

- But we can use more information in our prediction equation.
- For example, some countries are democracies and others are not:
 - Democracies tend to have lower levels of corruption
 - Autocracies tend to have higher levels of corruption



Adding a Covariate

How do we do this? We can generalize the prediction equation:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i}$$

This implies that we want to predict y using the information we have about x_1 and x_2 , and we are assuming a linear functional form.

Notice that now we write X_{ji} where:

- $j = 1, \dots, k$ is the index for the explanatory variables
- $i = 1, \dots, n$ is the index for the observation
- we often omit i to avoid clutter

In words:

$$\widehat{Corruption} = \hat{\beta}_0 + \hat{\beta}_1 Court + \hat{\beta}_2 Democracy$$

Worksheet Q1

Suppose

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2, \quad (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2) = (78.3, -10.6, -21.5),$$

where y is executive corruption (0–100), x_1 is high court independence (0–4), and $x_2 \in \{0, 1\}$ indicates democracy.

- 1 For a fixed level of court independence x_1 , compute the predicted difference $\hat{y}(x_2 = 1) - \hat{y}(x_2 = 0)$ and interpret.
- 2 For a fixed regime type, compute the predicted change when x_1 increases by one unit.
- 3 Verify that the slope on court independence is the same in both groups and explain why.

We are fitting two lines with the **same slope** but **different intercepts**.

Regression of Corruption on Court Independence and Democracy

From R, we obtain estimates $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$:

Coefficients:

	Estimate
(Intercept)	78.3236
courtind	-10.6301
democracy	-21.5101

• Autocracies:

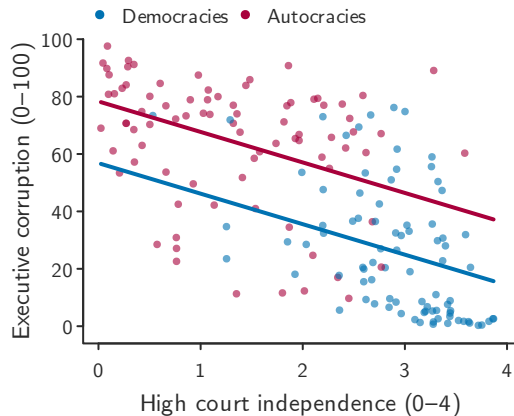
$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1$$

$$\hat{y} = 78.3 - 10.6 x_1$$

• Democracies:

$$\hat{y} = (\hat{\beta}_0 + \hat{\beta}_2) + \hat{\beta}_1 x_1$$

$$\hat{y} = 56.8 - 10.6 x_1$$

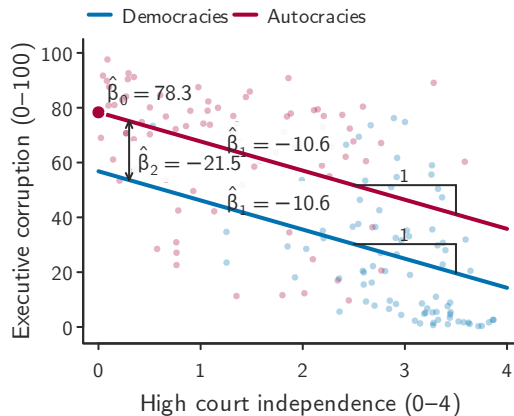


Regression of Corruption on Court Independence and Democracy

Our prediction equation is: $\hat{y} = 78.3 - 10.6 x_1 - 21.5 x_2$

Where do these quantities appear on the graph?

- $\hat{\beta}_0 = 78.3$ is the intercept for the prediction line for autocracies.
- $\hat{\beta}_1 = -10.6$ is the slope for both lines.
- $\hat{\beta}_2 = -21.5$ is the vertical distance between the lines for democracies and autocracies.
- Adding democracy shrinks the slope on court independence by about 40%, from -17.7 to -10.6 . Democracies have more independent courts (2.9 vs. 1.3 on average) and less corruption, so part of the simple slope reflects democracy.



Motivation for Multiple Regression

Two Cases

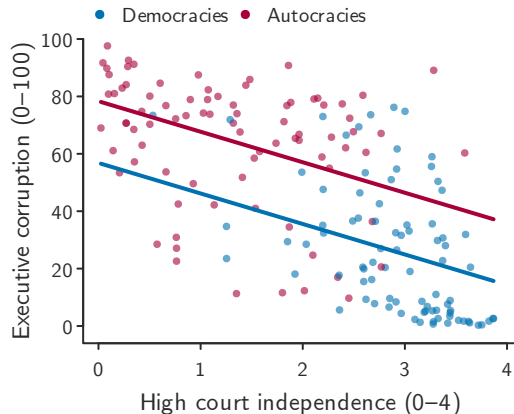
Multiple Regression with a Binary Covariate

Multiple Regression with a Continuous Covariate

Mechanics of Multiple Regression

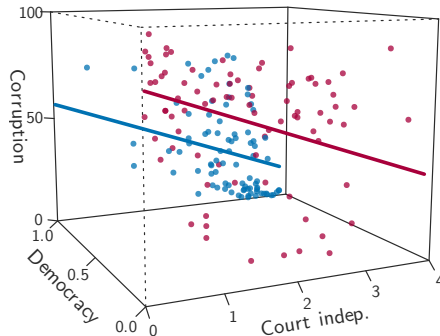
Fitting a regression plane

- We have considered an example of multiple regression with one **continuous** explanatory variable and one **binary** explanatory variable.
- This is easy to represent graphically in **two dimensions** because we can use colors to distinguish the two groups in the data.



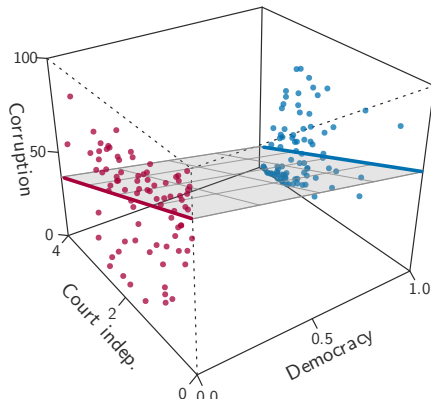
Regression of Corruption on Court Independence and Democracy

- These observations are actually located in a **three-dimensional** space.
- We can try to represent this using a **3D scatterplot**.
- In this view, we are looking at the data from the **court-independence side**; the two regression lines are drawn in the appropriate locations.



Regression of Corruption on Court Independence and Democracy

- We can also look at the 3D scatterplot from the **democracy** side.
- While the democracy variable is either 0 or 1, there is nothing in the prediction equation that requires this to be the case.
- In fact, the prediction equation defines a **regression plane** that connects the lines when $x_2 = 0$ and $x_2 = 1$.



Regression with two continuous variables

- Since we fit a regression plane to the data whenever we have two explanatory variables, it is easy to move to a case with **two continuous** explanatory variables.
- For example, we might want to use:
 - X_1 Court independence and X_2 $\log_{10}$ GDP per capita
 - Y Corruption

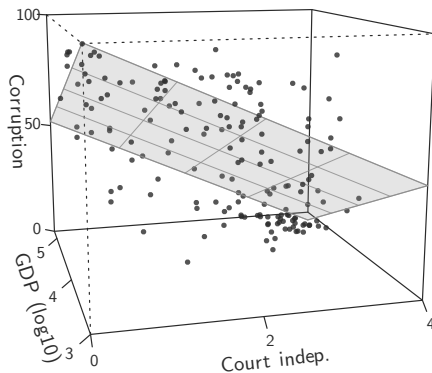
$$\widehat{\text{Corruption}} = \hat{\beta}_0 + \hat{\beta}_1 \text{Court} + \hat{\beta}_2 \log_{10} \text{GDP}$$

Regression of Corruption on Court Independence and GDP

- We can plot the points in a 3D scatterplot.
- R returns:
 - $\hat{\beta}_0 = 166.4$
 - $\hat{\beta}_1 = -13.95$ for Court independence
 - $\hat{\beta}_2 = -22.0$ for $\log_{10}$ GDP per capita

How does this look graphically?

- These estimates define a **regression plane** through the data.
- Turned so that court independence runs left to right, the plane slopes down in that direction too.



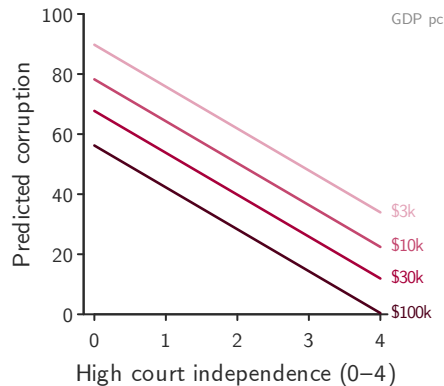
Interpreting a Continuous Covariate

- The coefficient estimates have a similar interpretation in this case as they did in the court independence and democracy example.
- For example, $\hat{\beta}_1 = -13.95$ represents our prediction of the difference in Corruption between two countries that differ by one point of court independence **but have the same GDP per capita**.
- The slope estimates have **partial effect** or **ceteris paribus** interpretations:

$$\frac{\partial \hat{y}}{\partial x_1} = \hat{\beta}_1, \quad \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$$

Interpreting a Continuous Covariate

- Again, we can think of this as defining a regression line for the relationship between Corruption and Court independence at every level of GDP per capita.
- All of these lines are parallel since they have the slope $\hat{\beta}_1 = -13.95$.
- The lines shift down as GDP per capita rises.



Worksheet Q2

With two continuous covariates, suppose

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2, \quad (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2) = (166.4, -13.95, -22.0),$$

where y is executive corruption (0–100), x_1 is high court independence (0–4), and x_2 is $\log_{10}$ GDP per capita.

- 1 Show that for any two units A and B ,

$$\hat{y}_A - \hat{y}_B = \hat{\beta}_1(x_{1A} - x_{1B}) + \hat{\beta}_2(x_{2A} - x_{2B}).$$

- 2 Using $(x_1, x_2) = (2.96, 4.48)$ for Chile and $(0.21, 4.38)$ for China, compute the predicted levels and their difference.
- 3 Explain the “partial effect” meaning of $\hat{\beta}_1$ and $\hat{\beta}_2$.

Motivation for Multiple Regression

Two Cases

Mechanics of Multiple Regression

- Derivation and Properties of OLS Estimator

- Analysis of Variance

- “Partialling Out” Interpretation

Deriving the Linear Least Squares Estimator

- In simple regression, we chose $(\hat{\beta}_0, \hat{\beta}_1)$ to minimize the sum of the squared residuals
- We use the same principle for picking $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)$ for regression with two regressors (x_i and z_i):

$$(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2) = \underset{\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\beta}_2}{\operatorname{argmin}} \sum_{i=1}^n (y_i - \tilde{\beta}_0 - x_i \tilde{\beta}_1 - z_i \tilde{\beta}_2)^2$$

- (The same works more generally for k regressors, but this is done more easily with matrices as we see later)

Deriving the Linear Least Squares Estimator

We want to minimize the following quantity with respect to $(\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\beta}_2)$:

$$S(\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\beta}_2) = \sum_{i=1}^n (y_i - \tilde{\beta}_0 - \tilde{\beta}_1 x_i - \tilde{\beta}_2 z_i)^2$$

How do we proceed?

- 1 Take the partial derivatives of S with respect to $\tilde{\beta}_0$, $\tilde{\beta}_1$ and $\tilde{\beta}_2$.
- 2 Set each of the partial derivatives to 0 to obtain the **first order conditions**.
- 3 Substitute $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ for $\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\beta}_2$ and solve for $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ to obtain the OLS estimator.

First Order Conditions

Setting the partial derivatives equal to zero leads to a system of 3 linear equations in 3 unknowns: $\hat{\beta}_0, \hat{\beta}_1$ and $\hat{\beta}_2$

$$\begin{aligned} -\frac{1}{2} \frac{\partial S}{\partial \tilde{\beta}_0} \Big|_{\tilde{\beta}=\hat{\beta}} &= \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i - \hat{\beta}_2 z_i) = 0 \\ -\frac{1}{2} \frac{\partial S}{\partial \tilde{\beta}_1} \Big|_{\tilde{\beta}=\hat{\beta}} &= \sum_{i=1}^n x_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i - \hat{\beta}_2 z_i) = 0 \\ -\frac{1}{2} \frac{\partial S}{\partial \tilde{\beta}_2} \Big|_{\tilde{\beta}=\hat{\beta}} &= \sum_{i=1}^n z_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i - \hat{\beta}_2 z_i) = 0 \end{aligned}$$

When will this linear system have a unique solution?

- More observations than predictors (i.e. $n > 2$)
- x and z are **linearly independent**, i.e.,
 - neither x nor z is a constant
 - x is not a linear function of z (or vice versa)
- Wooldridge calls this assumption **no perfect collinearity**

The OLS Estimator

The OLS estimators of $(\beta_0, \beta_1, \beta_2)$ can be written as

$$\begin{aligned}\hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x} - \hat{\beta}_2 \bar{z} \\ \hat{\beta}_1 &= \frac{\text{Cov}(x, y) \text{Var}(z) - \text{Cov}(z, y) \text{Cov}(x, z)}{\text{Var}(x) \text{Var}(z) - \text{Cov}(x, z)^2} \\ \hat{\beta}_2 &= \frac{\text{Cov}(z, y) \text{Var}(x) - \text{Cov}(x, y) \text{Cov}(z, x)}{\text{Var}(x) \text{Var}(z) - \text{Cov}(x, z)^2}\end{aligned}$$

where Var , Cov , Cor are sample moments.

For $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)$ to be well-defined we need:

$$\text{Var}(x) \text{Var}(z) \neq \text{Cov}(x, z)^2$$

Condition fails if:

- 1 If x or z is a constant ($\Rightarrow \text{Var}(x) \text{Var}(z) = \text{Cov}(x, z) = 0$)
- 2 One explanatory variable is an exact linear function of another ($\Rightarrow |\text{Cor}(x, z)| = 1 \Rightarrow \text{Var}(x) \text{Var}(z) = \text{Cov}(x, z)^2$)

Worksheet Q3 Analogous to the single variable case, show that the fitted values and residuals have three important properties:

- 1 $\sum_{i=1}^n \hat{u}_i = 0. \Rightarrow \bar{y} = \bar{\hat{y}}.$
- 2 The point $(\bar{y}, \bar{x}, \bar{z})$ is always on the regression plane.
- 3 $Cov(x, \hat{u}) = Cov(z, \hat{u}) = 0. \Rightarrow Cov(\hat{y}, \hat{u}) = 0.$

Motivation for Multiple Regression

Two Cases

Mechanics of Multiple Regression

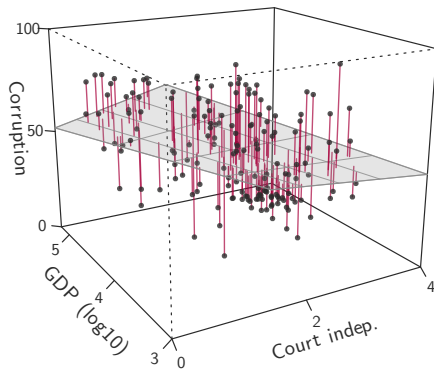
Derivation and Properties of OLS Estimator

Analysis of Variance

“Partialling Out” Interpretation

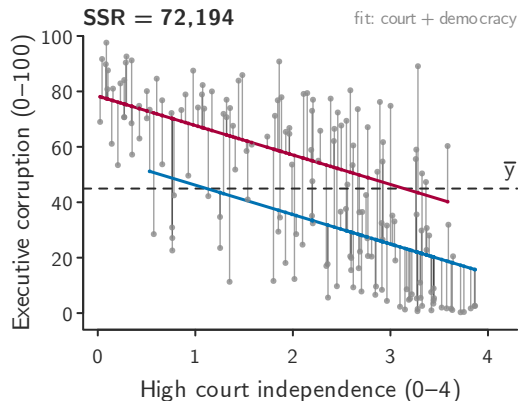
Fitted Values and Residuals

- We can define **fitted values** and **residuals** as we did in simple linear regression:
- Fitted Values:
$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i}$$
for $i = 1, \dots, n$
- Residuals:
$$\hat{u}_i = y_i - \hat{y}_i$$
for $i = 1, \dots, n$
- The residuals are the vertical distance between the observation and the regression plane.
- Corruption cannot fall below 0. Rich democracies such as Denmark (0.3) and Sweden (0.4) sit on this floor, about 7 to 8 points below the plane.



Analysis of Variance for Multiple Regression

- $\sum_i^n (y_i - \bar{y})^2 = SST$
(Total Sum of Squares)
- $\sum_i^n (\hat{y}_i - \bar{y})^2 = SSE$
(Explained Sum of Squares)
- $\sum_i^n (\hat{y}_i - y_i)^2 = \sum_i^n \hat{u}_i^2 = SSR$
(Residual Sum of Squares)



- $R^2 = SSE/SST$ never decreases when a regressor is added (same observations)
- Problematic as a criterion for **model selection**
- We may prefer a *parsimonious* model over a slightly worse fitting but much more complicated model; alternative: **Adjusted R^2**
- Penalizes model complexity: it multiplies $1 - R^2$ by $(n - 1)/(n - k - 1)$, which grows with k

$$\text{Adjusted } R^2 = 1 - (1 - R^2) \frac{n - 1}{n - k - 1}$$

Later we can show that Adjusted R^2 increases iff the absolute t -statistic on the added regressor exceeds 1.

- Drawbacks:
 - No longer guaranteed to be in $[0, 1]$
 - No longer directly interpretable as proportion of explained variance in y_i

```
> summary(lm(execorr100 ~ courtind + democracy, data = d))

Call:
lm(formula = execorr100 ~ courtind + democracy, data = d)

Residuals:
    Min       1Q   Median       3Q      Max
-52.662 -16.714   1.861  13.911  50.129

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   78.324     3.593   21.799 < 2e-16 ***
courtind     -10.630     2.113   -5.031 1.27e-06 ***
democracy     -21.510     4.672   -4.604 8.22e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 20.92 on 165 degrees of freedom
Multiple R-squared:  0.5039,    Adjusted R-squared:  0.4978
F-statistic: 83.78 on 2 and 165 DF,  p-value: < 2.2e-16
```

Worksheet Q4

Let $\hat{y}_i$ be fitted values from OLS with intercept and two regressors, and $\hat{u}_i = y_i - \hat{y}_i$.

- 1 Show: $SST = SSE + SSR$
- 2 Explain why R^2 is weakly increasing when a regressor is added

Later we can show that the adjusted R^2

$$\bar{R}^2 = 1 - (1 - R^2) \frac{n - 1}{n - k - 1}$$

increases iff the absolute t -statistic on the added regressor exceeds 1.

Motivation for Multiple Regression

Two Cases

Mechanics of Multiple Regression

Derivation and Properties of OLS Estimator

Analysis of Variance

“Partialling Out” Interpretation

“Partialling Out” Interpretation of the OLS Estimator

Assume $Y = \beta_0 + \beta_1 X + \beta_2 Z + u$. Another way to write the OLS estimator is:

$$\hat{\beta}_1 = \frac{\sum_i^n \hat{r}_{xz,i} y_i}{\sum_i^n \hat{r}_{xz,i}^2}$$

where $\hat{r}_{xz,i}$ are the residuals from the regression of X on Z :

$$x_i = \hat{\lambda} + \hat{\delta} z_i + \hat{r}_{xz,i}$$

- In other words, both of these regressions yield identical estimates $\hat{\beta}_1$ (but different fitted values $\tilde{y}$ and $\hat{y}$):

$$\tilde{y} = \hat{\gamma}_0 + \hat{\beta}_1 \hat{r}_{xz} \quad \text{and} \quad \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 z$$

- $\hat{\delta}$ is the slope of that regression; it is zero exactly when X and Z are uncorrelated in the sample. What is our estimator $\hat{\beta}_1$ if $\hat{\delta} = 0$ in our sample?

$$\hat{r}_{xz,i} = x_i - \hat{\lambda} = x_i - \bar{x} \quad \text{so} \quad \hat{\beta}_1 = \frac{\sum_i^n \hat{r}_{xz,i} y_i}{\sum_i^n \hat{r}_{xz,i}^2} = \frac{\sum_i^n (x_i - \bar{x}) y_i}{\sum_i^n (x_i - \bar{x})^2}$$

That is, same as the simple regression of Y on X alone.

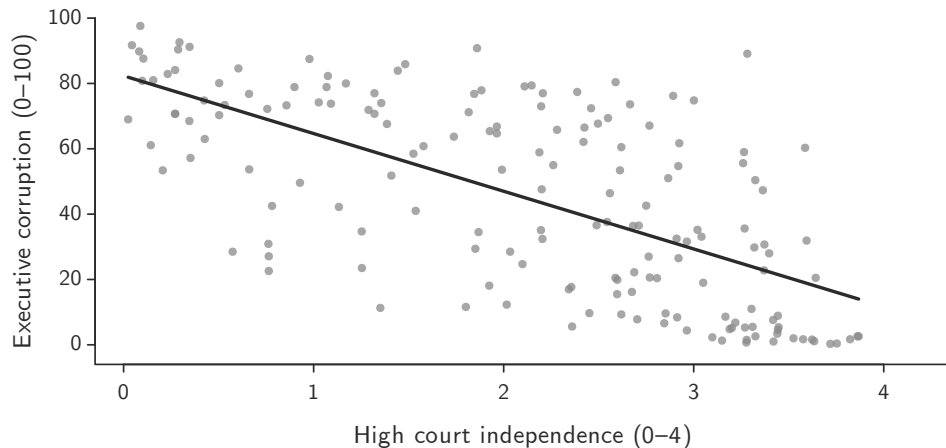
“Partialling Out” Interpretation of the OLS Estimator

Both of these regressions yield identical estimates $\hat{\beta}_1$ (but different fitted values $\tilde{y}$ and $\hat{y}$):

$$\tilde{y} = \hat{\gamma}_0 + \hat{\beta}_1 \hat{r}_{xz} \quad \text{and} \quad \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 z$$

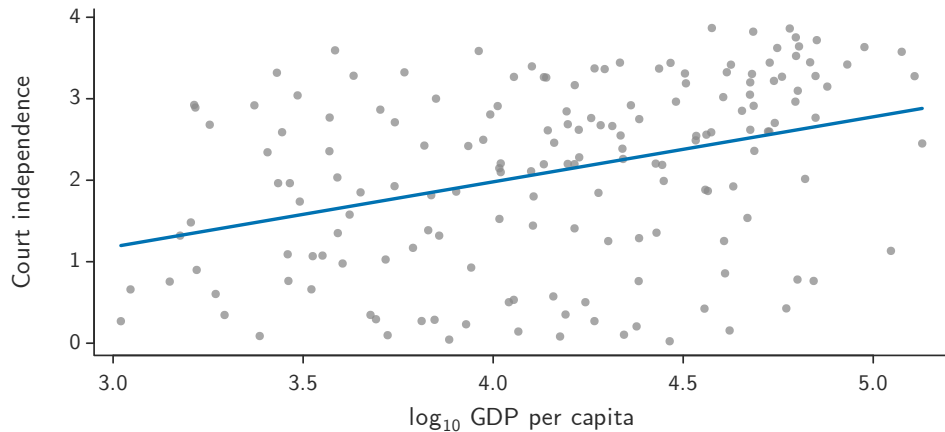
- Residuals $\hat{r}_{xz}$ are the part of X that is uncorrelated with Z . Put differently, $\hat{r}_{xz}$ is X , after its linear association with Z has been **partialled out** or netted out.
- We can use same equation with k explanatory variables; $\hat{r}_{xz}$ will then come from a regression of X on all the other explanatory variables.
- We will prove this later using matrix algebra.

Regression: Corruption on Court Independence



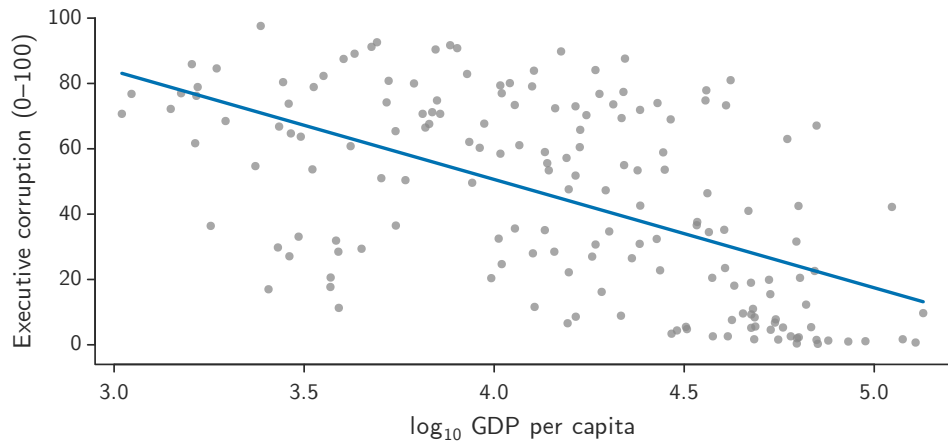
$\widehat{\text{Corr}} = 82.3 - 17.7 \text{ Court}$. But richer countries tend to have more independent courts, and richer countries are much less corrupt...

Regression: Court Independence on GDP



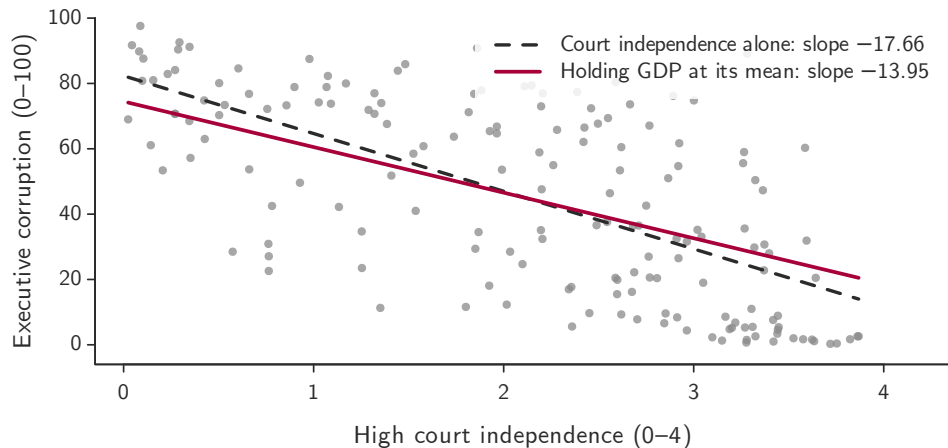
...like this...

Regression: Corruption on GDP



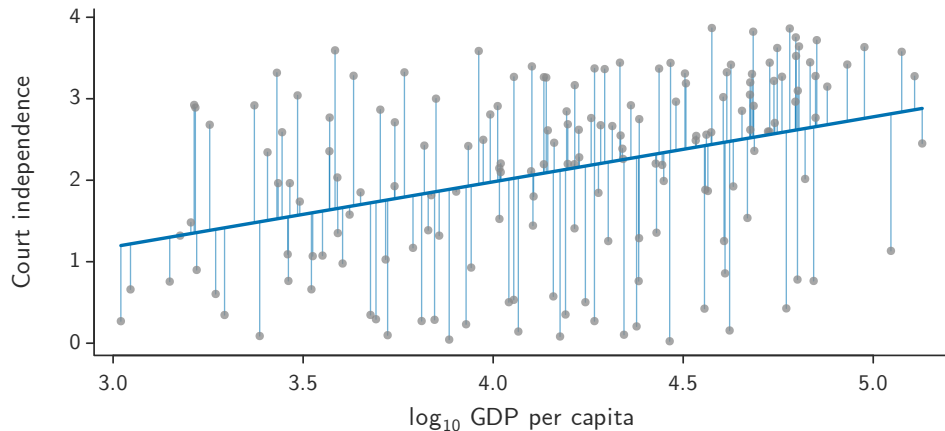
and like this.

Regression: Corruption on Court Independence



$\widehat{\text{Corr}} = 166.4 - 13.95 \text{ Court} - 22.0 \log_{10} \text{GDP}$. Holding GDP fixed, the slope shrinks from -17.7 to -13.95 . Court independence tracks income only loosely ($r = 0.37$), so income explains only a fifth of the slope.

Regression: Court Independence on GDP



$\widehat{\text{Court}} = -1.21 + 0.80 \log_{10} \text{GDP}$. The residuals from the regression of court independence on GDP measure the variation in court independence that is uncorrelated with GDP.

Regression: Corruption on Court Independence



$\widehat{\text{Corr}} = 44.94 - 13.95 \text{ Res: Court on GDP}$. Regressing corruption on these residuals gives the same slope as the plane.

Regression: Corruption on Court Independence

R Code

```
> summary(lm(execorr100 ~ courtind, data = d))

Call:
lm(formula = execorr100 ~ courtind, data = d)

Residuals:
    Min       1Q   Median       3Q      Max
-47.164 -17.940   0.949  15.506  64.745

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   82.327     3.692   22.30  <2e-16 ***
courtind     -17.664     1.546  -11.42  <2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Residual standard error: 22.15 on 166 degrees of freedom
Multiple R-squared: 0.4401,    Adjusted R-squared: 0.4367
F-statistic: 130.5 on 1 and 166 DF,  p-value: < 2.2e-16
```

Regression: Court Independence on GDP

R Code

```
> d$lgdp <- log10(d$gdppc) # log10 GDP per capita
> fit <- lm(courtind ~ lgdp, data = d)
> summary(fit)

Call:
lm(formula = courtind ~ lgdp, data = d)

Residuals:
    Min       1Q   Median       3Q      Max
-2.3267 -0.6259  0.1371  0.8021  1.9460

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.2123     0.6604  -1.836   0.0682 .
lgdp           0.7981     0.1572   5.078 1.01e-06 ***

Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Residual standard error: 1.035 on 166 degrees of freedom
Multiple R-squared:  0.1345,    Adjusted R-squared:  0.1292
F-statistic: 25.79 on 1 and 166 DF,  p-value: 1.015e-06

> d$res_court_gdp <- residuals(fit) # store residuals as a new variable
```

Regression: Corruption on Netted-Out Court Independence

R Code

```
> summary(lm(execorr100 ~ courtind + lgdp, data = d))
~~~~~
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  166.352      12.626   13.175 < 2e-16 ***
courtind      -13.952       1.469   -9.497 < 2e-16 ***
lgdp         -22.028       3.198   -6.889 1.13e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
~~~~~

> summary(lm(execorr100 ~ res_court_gdp, data = d))
~~~~~
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   44.942      1.994   22.534 < 2e-16 ***
res_court_gdp -13.952       1.939   -7.194 2.06e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
~~~~~
```