



Stockholm
University

Pär Nyman

Methods in Political Science
8 April 2026

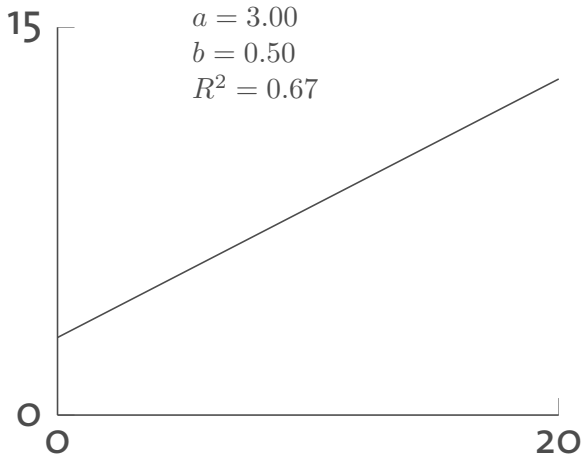
Additional topics

Non-linearities, outliers and conditional relationships

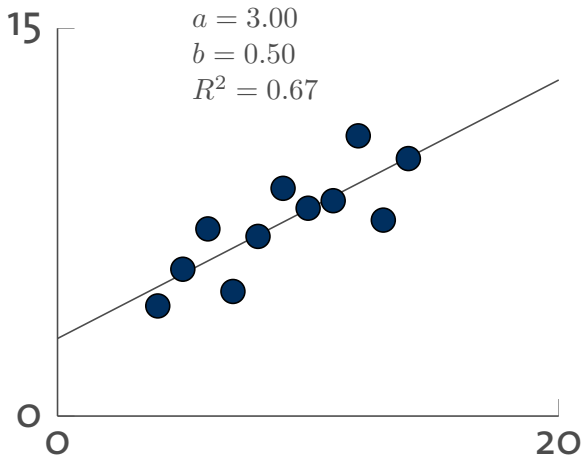
Introduction

- Tendency to see graphs as simple – suitable for those uncomfortable with numbers – and regression tables as serious stuff.
- This view is misguided for several reasons.
- One is that regression results can be misleading when the model does not align with what the data look like.

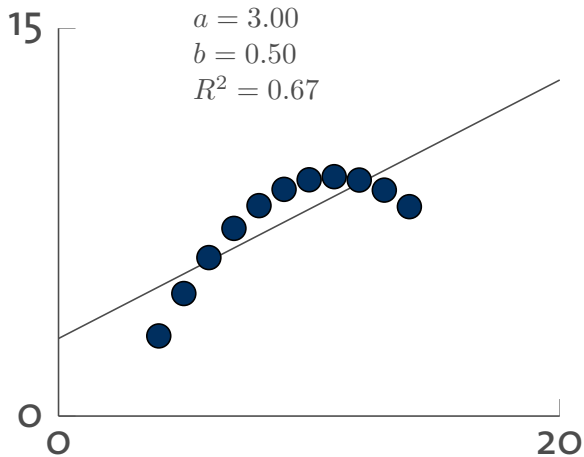
Introduction



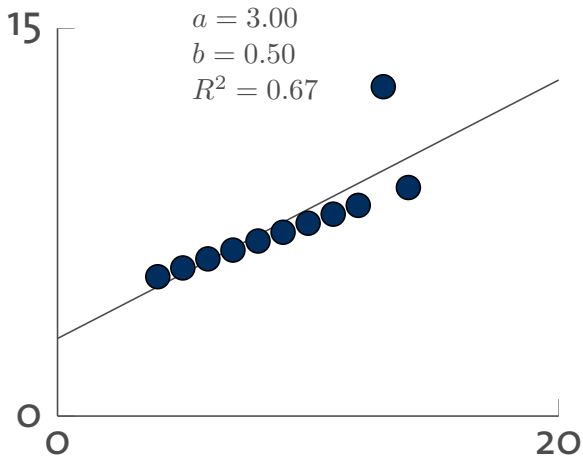
Introduction



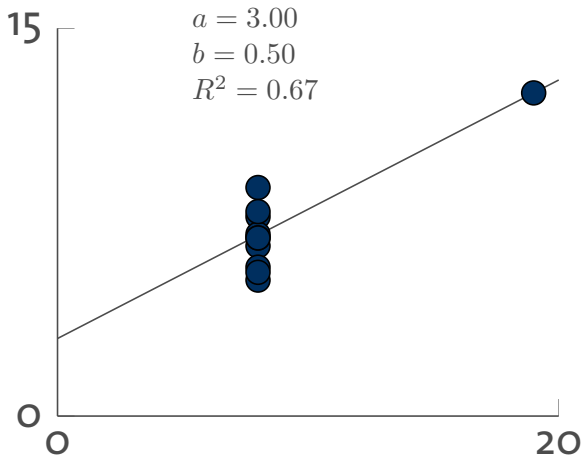
Introduction



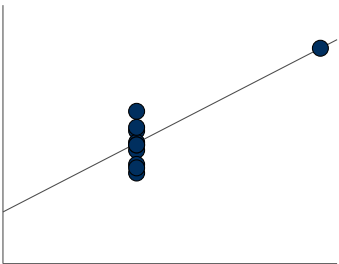
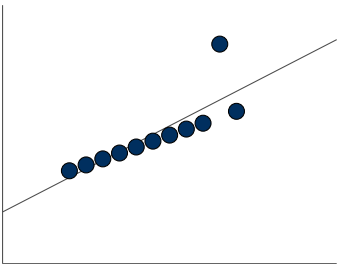
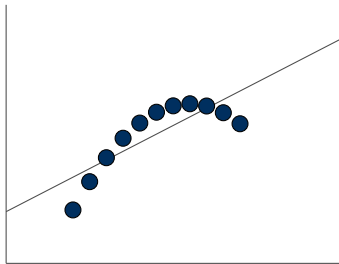
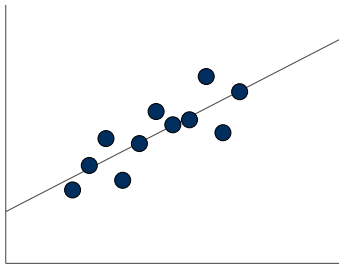
Introduction



Introduction



Introduction



Introduction

Lessons

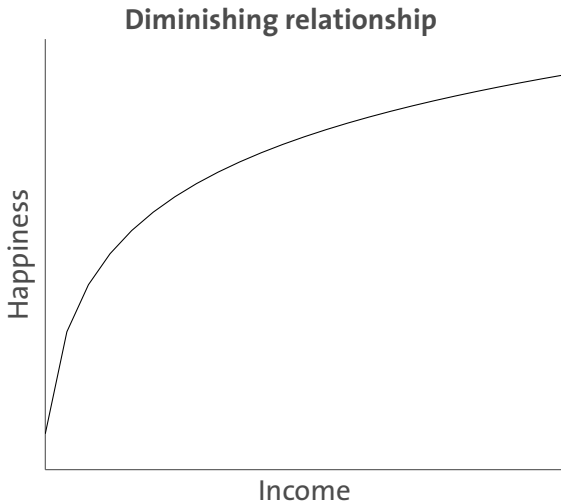
- Graphs are important.
- Misleading regression results when there is:
 - Non-linearity.
 - Outliers.

Non-linearity

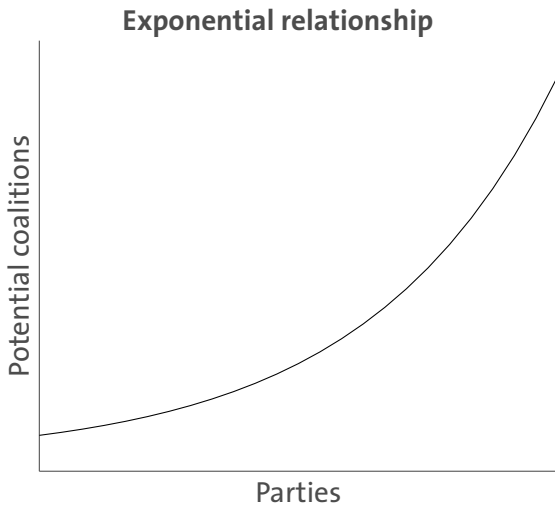
The problem

- The linear model assumes that a one-unit increase in x always corresponds to the same change in y (i.e., β).
- This assumption is often violated.
- Sometimes we expect specific non-linearities ...

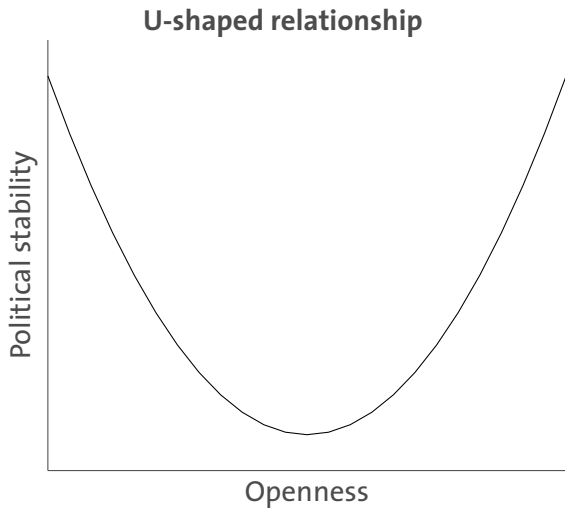
Non-linearity



Non-linearity



Non-linearity



Non-linearity

...we don't have to expect something this clear-cut.

- Better predictions?
- Better controls?
- Test the plausibility of the linear model.

Non-linearity

How can we accomplish this?

- The linear model is our framework.
- Solution: add non-linear transformations of the independent variable.
 - Log variables.
 - Quantiles and other groups.
 - Polynomials.

Non-linearity

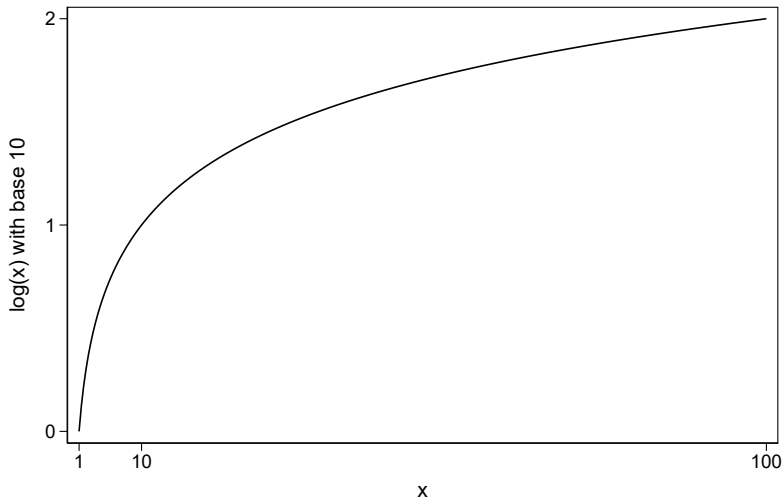
Log-transformation

x	$\log_{10}(x)$
1	0
10	1
100	2
1000	3
10 000	4

x	$\log_2(x)$
1	0
2	1
4	2
8	3
16	4

Non-linearity

Log-transformation



Non-linearity

Log-transformation

- It is common to model diminishing relationships by replacing x with its logarithm $\log(x)$.
- Strong assumption with attractive interpretation: we analyze the response to *relative* instead of *absolute* changes.
- Exact interpretation depends on base of log.
 - $\log_2$: If x increases with 100 percent, y increases with β .
 - $\log_{10}$: If x increases with factor 10, y increases with β .
- Same functional form regardless of log base.

Non-linearity

Groups

- We divide the data into groups based on x and add a dummy per group (e.g., income decile fixed effects).
 - Quantiles.
 - Other groups (often natural categories).
- No assumption about functional form, but $\hat{y}$ constant within groups.

Non-linearity

Polynomial

- We add higher-order terms (x^2 , x^3 , etc.) to the model.
- Weak assumptions (but difficult to formulate).
- Very difficult to interpret coefficients if higher than second-order.

Non-linearity

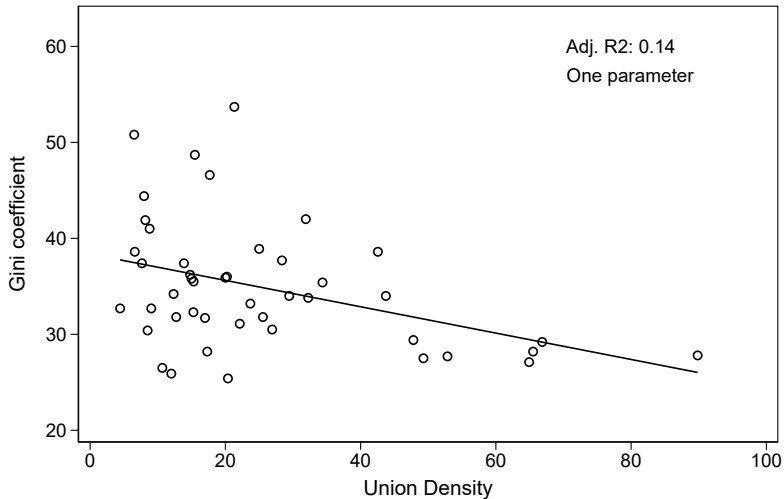
Why not choose the most flexible specification?

- Statistical precision and overfitting.
- Interpretability and presentation.
- Conventions are convincing.

Conclusion: data and purpose should determine model choice.

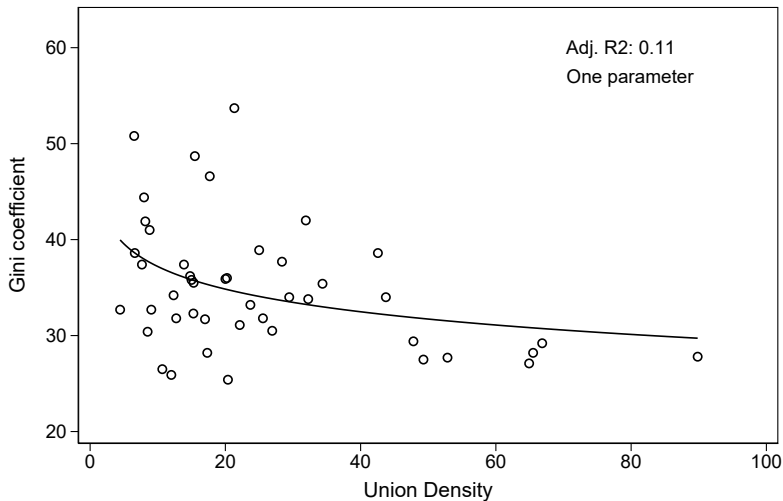
Non-linearity

Linear



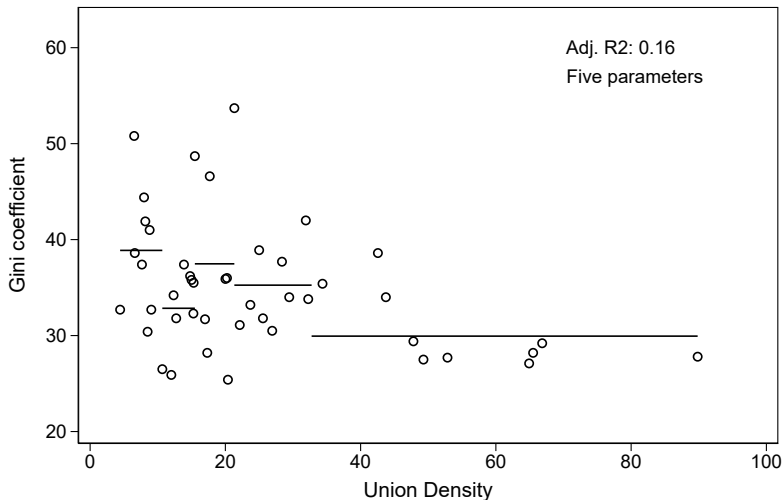
Non-linearity

Logged density



Non-linearity

Quintile groups (five quantile groups)



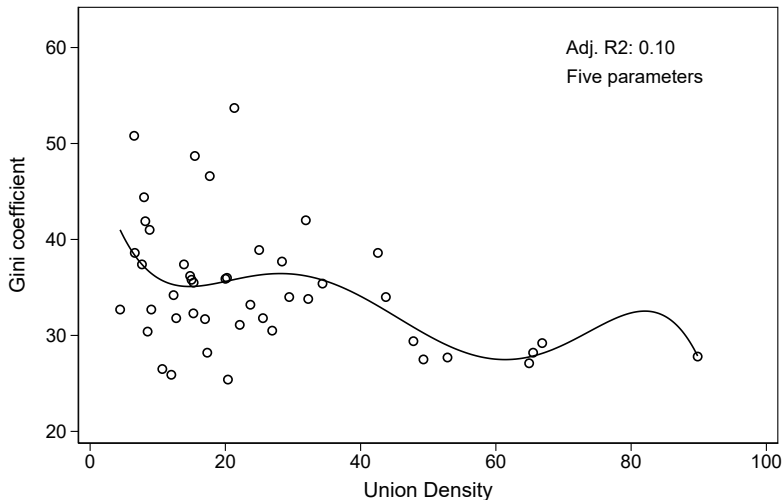
Non-linearity

Third-order polynomial



Non-linearity

Fifth-order polynomial



Non-linearity

Binary dependent variables

- If the dependent variable is binary, the linear model is unlikely to be the correct functional form. E.g., linear models can make predictions outside the 0–1 interval.
- Yet, for most purposes, the linear model usually works fine.
- The expected/predicted value in can be interpreted as the probability of a “positive” outcome.
- Logit/probit are alternatives which we do not cover.

Outliers

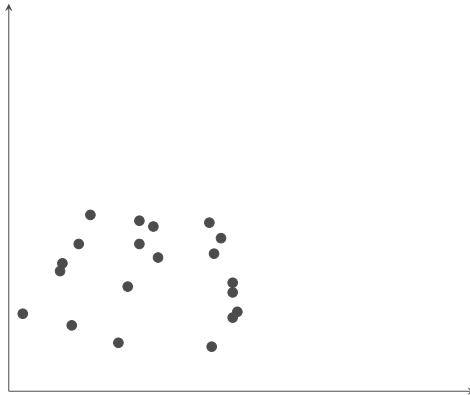
- Outliers are observations with extreme values on one or a combination of variables.
- They can have disproportionately strong influence on the regression results.

Outliers

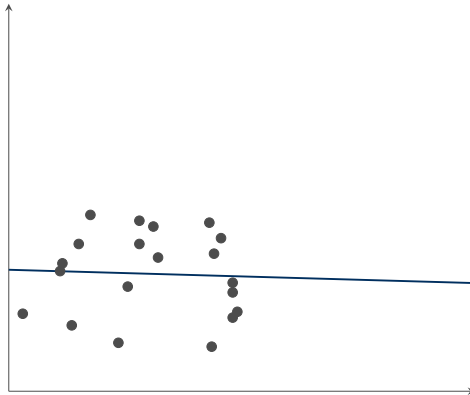
Three kinds of “outlierness”

- Observations with *leverage* have extreme values on the independent variables.
- Observations with *discrepancy* have extreme values on the dependent variable (given x).
- Those with both leverage and discrepancy are *influential observations*.

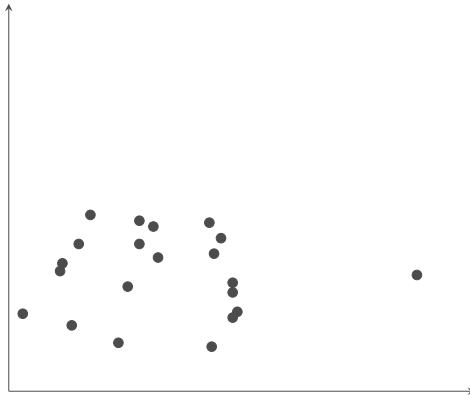
Outliers



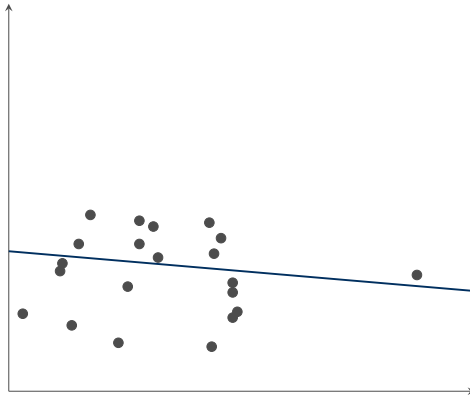
Outliers



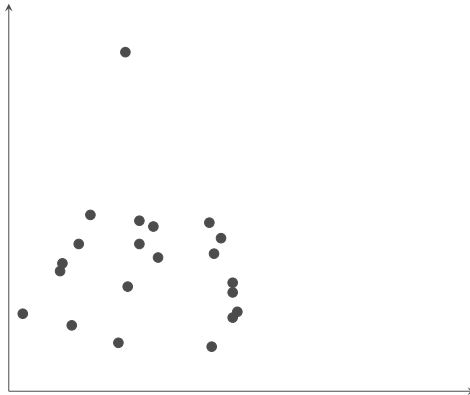
Outliers



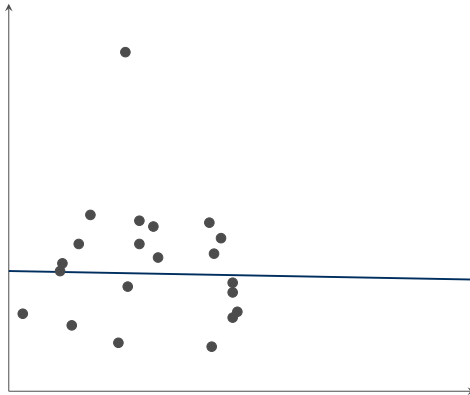
Outliers



Outliers



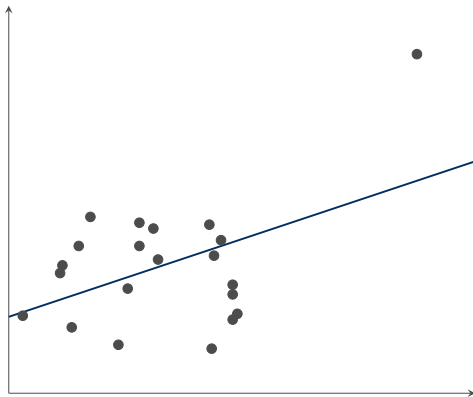
Outliers



Outliers



Outliers



Outliers

Identifying outliers

- Graphs!
- `dfbeta` (or `dfbetas`, which is a rescaled version of `dfbeta`) says how much b would change if an observation is excluded.
- Exclude countries, years or similar depending on setting.

Outliers

- If measurement error, remove and explain.
- If not, present results with and without.
- You should also take a look at the obs which are driving your results. Representative?
- There are also other estimators than OLS (see robust regression), although they are not covered by this course.

Conditional relationships

- Unless we specify otherwise, regression models assume that the effect of an independent variable is constant across all observations.
- If effects differ substantially across groups we want to capture it in our models.
 - Theoretically interesting (main point).
 - Better predictions.
 - Better controls.
- We can do that by splitting the sample, and thereby estimating separate effects for the two samples, or by including an interaction term.

Conditional relationships

Interpreting (any) interaction models

- If we suspect that the effect of x is conditioned on z we could estimate this equation: $\hat{y} = a_1 + b_1x + b_2xz + b_3z$
- We rewrite to highlight that the effect of x is $b_1 + b_2z$:
$$\hat{y} = a_1 + (b_1 + b_2z)x + b_3z$$
- b_1 is the effect of x when $z = 0$.
- b_2 is how much the effect of x increases if z increases by one.

Conditional relationships

Interpreting interaction models with a dichotomous variable

- When z is binary (0/1), there is one effect when $z = 0$ (b_1) and another effect when $z = 1$ ($b_1 + b_2$).
- This is equivalent to the split sample.
 - Unless there are other variables in the model.
 - The split sample does not get a coefficient/test for the difference between the two effects, while the interaction model does not get a test/coefficient for the effect when $z = 1$, but all this can be calculated.
- Can be visualized as two regression lines with different slopes and intercepts.
- The interpretation is more difficult when z is continuous.

Conditional relationships

Facilitate interpretation of continuous interactions

- Still holds that b_1 is the effect of x when $z = 0$ and that b_2 says how sensitive the effect is to changes in z .
- Center z around its mean
 - = Subtract the variable mean, so the new mean is zero.
 - Then b_1 shows the effect of x when z is at its mean.
- Marginal effects graphs
 - Show the expected change in y when x increases by 1 over different values of z .
 - Are often combined with a histogram of z

Conditional relationships

Remember the multiple comparisons problem

- Most studies are under-powered, and the situation is much worse for subgroups and the differences between them.
 - The precision for a sub-group is almost always worse than for the main sample.
 - The precision for the difference between two groups is rarely sufficient.
 - There are many subgroups and possible comparisons. Be aware of multiple comparisons!
- IMO, the discipline is too fond of heterogeneity analyzes.

Conditional relationships

Marginal effects graph in Stata (QoG)

```
ssc install marhis  
gen ln_gdp = wdi_gdpcapcur  
label define demo 0 "Non-democracy" 1 "Democracy"  
label values br_dem demo
```

```
reg wvs_polint c.vdem_polyarchy##c.ln_gdp  
marhis vdem_polyarchy, mar(ln_gdp)
```

```
reg wvs_polint i.br_dem##c.ln_gdp  
marhis ln_gdp, cate(br_dem)
```

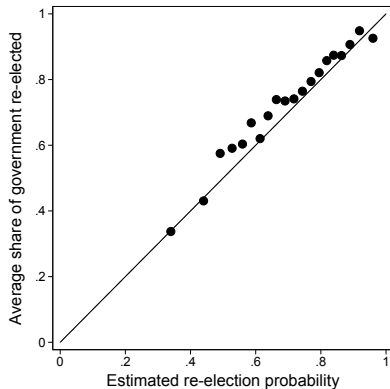
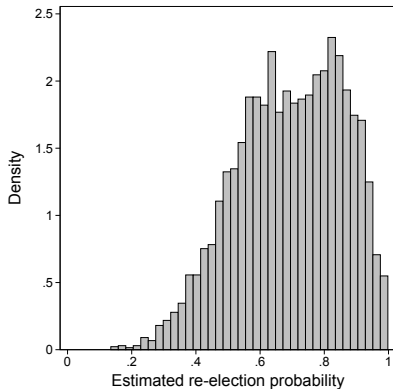
Empirical application

How do governments adjust tax policy when expecting a close election?

- Myopia: They cut taxes regardless of ideology.
- Moderation: They move towards the center.
 - Left-wing incumbents cut taxes.
 - Right-wing incumbents raise taxes.
 - Compared to what they would otherwise do.

Our selling point: a new dataset on estimated re-election prospects for Swedish local governments.

Empirical application



Empirical application

Modeling considerations

- Using fixed effects on municipality-term and year we control for circumstances specific to each government as well as the trends in the national polls. The variation we use is the interplay between them. Plausibly exogenous?
- If we code left-wing so that it is 0 on average (approximately -0.5 for right and 0.5 for left) we can interpret b_1 as myopia and b_2 as moderation.

$$Tax_{t+1} = b_0 + b_1 \times competitiveness + b_2 \times left * competitiveness$$

Empirical application

Table 1: Effects of electoral competitiveness on taxation

	(1)	(2)	(3)
Competitiveness (myopia)	-0.199** (0.077)	-0.094 (0.073)	-0.072 (0.073)
Left-wing $\times$ Competitiveness (moderation)	-0.541** (0.233)	-0.723*** (0.213)	-0.858*** (0.230)
Observations	3,750	3,749	3,749
Within-period R^2	0.006	0.165	0.175
Mean of dep. variable	21.4	21.4	21.4
Municipality-term and year FE	Yes	Yes	Yes
Additional controls	No	Yes	Yes
Additional interactions	No	No	Yes

The left-wing incumbent indicator is centered around its mean, which means that the effect of competitiveness can be interpreted as the opportunistic policy response of an average (rather than a right-wing) government. The constitutive term for the left-wing incumbent indicator is absorbed due to the inclusion of the election period (municipality $\times$ election term) fixed effects, and is thus not estimable. Standard errors in parentheses, clustered by election period.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For full output, see Table S3.

Empirical application

Running 280 regressions removing one municipality at a time

