CSSS 569 · Visualizing Data

VISUALIZING INFERENCE

Christopher Adolph

Department of Political Science

and

Center for Statistics and the Social Sciences

University of Washington, Seattle

Lessons so far...

Principles for effective visual display

How to avoid cognitive pitfalls in designing visuals

Basics of R graphics

Application of the above to Exploratory Data Analysis (EDA)

Most resources on scientific visualization stop with EDA,
or using graphics to understand the data

But quantitative social science emphasizes *modeling* data

Next step: Designing effective visuals for understanding *models*

Roadmap for Visualizing Inference

1. Obtaining Quantities of Interest from models
2. Introduction to the `tile` package
3. Graphical approaches to model inference
4. Graphical approaches to model robustness
5. Visualizing interactive models

What determines cross-national inflation performance?

Source: Adolph, *BBC*, Ch. 3

Method: Time series cross-section regression with compositional covariates

Who votes in American elections?

Source: King, Tomz, and Wittenberg (2000)

Method: Logistic regression

How do Chinese leaders gain power?

Source: Shih, Adolph, and Liu (2012)

Method: Bayesian model of partially observed ranks

When do governments choose liberal or conservative central bankers?

Source: Adolph, *BBC*, Ch. 8

Method: Zero-inflated compositional data model

How long do central bankers stay in office?

Source: Adolph, *BBC*, Ch. 9

Method: Cox proportional hazards model

Presenting Estimated Models in Social Science

Most empirical work in social science is regression model-driven,
with a focus on conditional expectation

Our regression models are

- full of covariates
- often non-linear
- usually involve interactions and transformations

If there is anything we need to visualize well, it is our models

Yet we often just print off tables of parameter estimates

Limits readers' *and analysts'* understanding of the results

Presenting Estimated (Causal) Models in Social Science

What if you work in a causal inference framework?

Still a great need for visualization:

- to show robustness across different techniques
- to show differences across quantities of interest (e.g., ATE vs. ATT)
- to show variation across different kinds of subjects (variation in LATEs/LATTs)

What Most Discussions of Statistical Graphics Leave Out

Tufte's books have had a huge impact on information visualization

However, they have two important limits:

Modeling Most examples are either exploratory or very simple models;

Social scientists want cutting edge applications

Tools Need to translate aesthetic guidelines into software

Social scientists are unlikely to do this on their own –
and shouldn't have to!

Goals for Visualizing Model Inference

Key need Ready-to-use techniques to visually present model results:

- for many variables

Goals for Visualizing Model Inference

Key need Ready-to-use techniques to visually present model results:

- for many variables
- for many robustness checks

Goals for Visualizing Model Inference

Key need Ready-to-use techniques to visually present model results:

- for many variables
- for many robustness checks
- showing uncertainty

Goals for Visualizing Model Inference

Key need Ready-to-use techniques to visually present model results:

- for many variables
- for many robustness checks
- showing uncertainty
- without accidental extrapolation

Goals for Visualizing Model Inference

Key need Ready-to-use techniques to visually present model results:

- for many variables
- for many robustness checks
- showing uncertainty
- without accidental extrapolation
- for an audience without deep statistical knowledge

"Tufte without Tears"

Voting Example (Logit Model)

We will explore a simple dataset using a simple model of voting

People either vote ($Vote_i = 1$), or they don't ($Vote_i = 0$)

Many factors could influence turn-out; we focus on age and education

National Election Survey (2000) data

"Did you vote in 2000 election?"

vote00	age	hsdeg	coldeg
1	49	1	0
0	35	1	0
1	57	1	0
1	63	1	0
1	40	1	0
1	77	0	0
0	43	1	0
1	47	1	1
1	26	1	1
1	48	1	0
...			

Logit of Decision to Vote, 2000 Presidential NES

	est.	s.e.	*p*-value
Age	0.074	0.017	0.000
Age^2	-0.0004	0.0002	0.009
High School Grad	1.168	0.178	0.000
College Grad	1.085	0.131	0.000
Constant	-3.05	0.418	0.000

Age enters as a quadratic to allow the probability of voting to first rise and eventually fall over the life course

Results look sensible, but what do they mean?

Which has the bigger effect, age or education?

What is the probability a specific person will vote?

An Alternative to Eye-glazing Tables

1. Estimate your model as normal; treat the output as an intermediate step

An Alternative to Eye-glazing Tables

1. Estimate your model as normal; treat the output as an intermediate step
2. Translate your model results back into the scale of the response variable
   - Modeling war? Show the change in probability of war associated with x
   - Modeling counts of crimes committed? Show how those counts vary with x
   - Unemployment rate time series? Show how a change in x shifts the unemployment rate over the following t years

An Alternative to Eye-glazing Tables

1. Estimate your model as normal; treat the output as an intermediate step
2. Translate your model results back into the scale of the response variable
    - Modeling war? Show the change in probability of war associated with x
    - Modeling counts of crimes committed? Show how those counts vary with x
    - Unemployment rate time series? Show how a change in x shifts the unemployment rate over the following t years
3. Calculate or simulate the uncertainty in these final quantities of interest

An Alternative to Eye-glazing Tables

1. Estimate your model as normal; treat the output as an intermediate step
2. Translate your model results back into the scale of the response variable
   - Modeling war? Show the change in probability of war associated with x
   - Modeling counts of crimes committed? Show how those counts vary with x
   - Unemployment rate time series? Show how a change in x shifts the unemployment rate over the following t years
3. Calculate or simulate the uncertainty in these final quantities of interest
4. Present visually as many scenarios calculated from the model as needed

A bit more formally... (KTW, 2000, *AJPS*)

We want to know the behavior of $E(y|x)$ as we vary x

In non-linear models with multiple regressors, this gets tricky

The effect of x_1 depends on all the other x's and $\hat{\beta}$'s

Generally, we will need to make a set of "counterfactual" assumptions:
$x_1 = a, \qquad x_2 = b, \qquad x_3 = c, \quad \ldots$

- Choose $a, b, c, \ldots$ to match a particular counterfactual case of interest *or*
- Hold all but one of the x's at their mean values (or other baseline, such as the factual values by case), then systematically vary the remaining x

The same trick works if we are after differences in y related to changes in x, such as

$$E(y_{scen2} - y_{scen1}|x_{scen1}, x_{scen2}) \qquad \text{or} \qquad E(y_{scen2}/y_{scen1}|x_{scen1}, x_{scen2})$$

Calculating quantities of interest

Our goal to obtain "quantities of interest," like

- Expected Values: $E(y|x_c)$
- Differences: $E(y_{c2} - y_{c1}|x_{c1}, x_{c2})$
- Risk Ratios: $E(y_{c2}/y_{c1}|x_{c1}, x_{c2})$
- or any other function of the above

for some counterfactual x_c's.

For our Voting example, that's easy – just plug x_c into the systematic component:

$$E(y|x_c) = logit^{-1}(x_c\beta)$$

Our goal to obtain "quantities of interest," like

- Expected Values: $E(y|x_c)$
- Differences: $E(y_{c2} - y_{c1}|x_{c1}, x_{c2})$
- Risk Ratios: $E(y_{c2}/y_{c1}|x_{c1}, x_{c2})$
- or any other function of the above

for some counterfactual x_c's.

For our Voting example, that's easy – just plug x_c into the systematic component:

$$E(y|x_c) = logit^{-1}(x_c\beta) = \frac{1}{1 + \exp(-x_c\beta)}$$

Our goal to obtain "quantities of interest," like

- Expected Values: $E(y|x_c)$
- Differences: $E(y_{c2} - y_{c1}|x_{c1}, x_{c2})$
- Risk Ratios: $E(y_{c2}/y_{c1}|x_{c1}, x_{c2})$
- or any other function of the above

for some counterfactual x_c's.

For our Voting example, that's easy – just plug x_c into the systematic component:

$$E(y|x_c) = logit^{-1}(x_c\beta) = \frac{1}{1 + \exp(-x_c\beta)} = Pr(y|x_c)$$

Getting confidence intervals is harder, but there are several options:

- For maximum likelihood models,
  simulate the response conditional on the regressors

  These simulations can easily be summarized as CIs: sort them and take percentiles

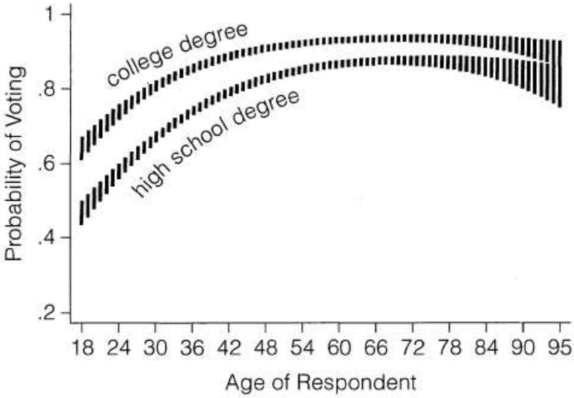
  See King, Tomz, and Wittenberg, 2000, *American Journal of Political Science*, and the `Zelig` or `simcf` packages for `R` or `Clarify` for Stata.

- For Bayesian models, usual model output *is* a set of posterior draws

  See Andrew Gelman and Jennifer Hill, 2006, *Data Analysis Using Hierarchical/ Multilevel Models*, Cambridge UP.

Once we have the quantities of interest and confidence intervals,
we're ready to make some graphs...but how?

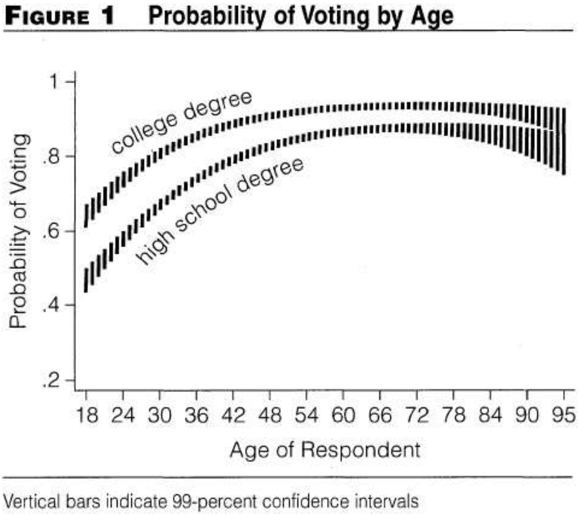
Figure 1 Probability of Voting by Age



Vertical bars indicate 99-percent confidence intervals

Here is the graph that King, Tomz, and Wittenberg created for this model

How would we make this ourselves?

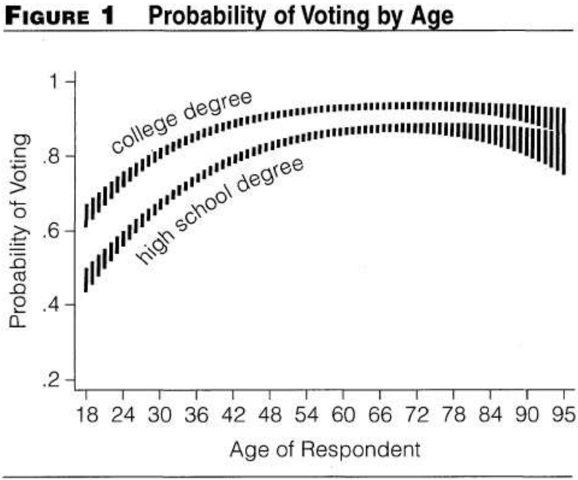


Here is the graph that King, Tomz, and Wittenberg created for this model

How would we make this ourselves?

Could use the default graphics in `Zelig` or `Clarify` (limiting, not as nice as the above)

Could do it by hand (tedious)



Vertical bars indicate 99-percent confidence intervals

We'll return to this example later and develop tools for making plots like this

Note we don't have to always present model inference in this format

Beginners often fixate on plots like this, with a continuous covariate on the *x*-axis, but there are other options...

American Interest Rate Policy

Recall this example from my own work on central banking (*Bankers, Bureaucrats, and Central Bank Politics*, Cambridge U.P., 2013, Ch. 4)

Federal Reserve Open Market Committee (FOMC) sets interest rates 10×/year

Members of the FOMC vote on the Chair's proposed interest rate

Dissenting voters signal whether they would like a higher or lower rate

Dissents are rare but may be symptomatic of how the actual rate gets chosen

American Interest Rate Policy

Recall this example from my own work on central banking (*Bankers, Bureaucrats, and Central Bank Politics*, Cambridge U.P., 2013, Ch. 4)

Federal Reserve Open Market Committee (FOMC) sets interest rates 10$\times$/year

Members of the FOMC vote on the Chair's proposed interest rate

Dissenting voters signal whether they would like a higher or lower rate

Dissents are rare but may be symptomatic of how the actual rate gets chosen

Many factors could influence interest rate votes:

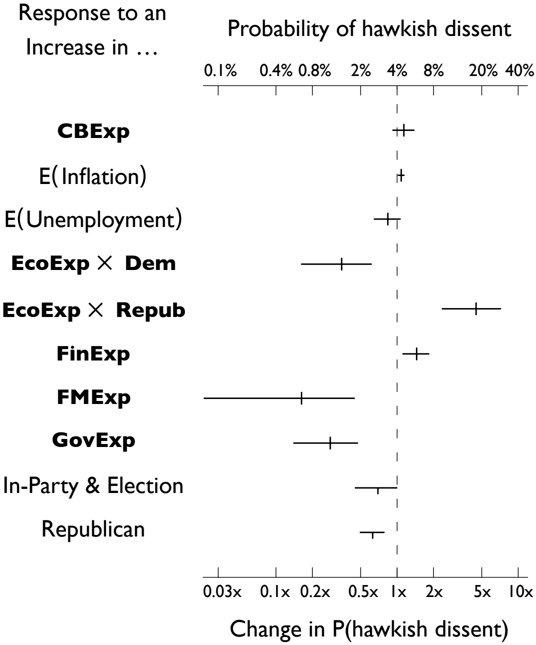
Individual	Career background Appointing party Interactions of above
Economy	Expected inflation Expected unemployment
Politics	Election cycles

Response variable: FOMC Votes (1 = ease, 2 = accept, 3 = tighten)					
EVs	**param.**	**s.e.**	**EVs**	**param.**	**s.e.**
FinExp	−0.021	(0.146)	E(Inflation)	0.019	(0.015)
GovExp	−0.753	(0.188)	E(Unemployment)	−0.035	(0.022)
FMExp	−1.039	(0.324)	In-Party, election year	−0.182	(0.103)
CBExp	−0.142	(0.141)	Republican	−0.485	(0.102)
EcoExp × Repub	0.934	(0.281)	Constant	2.490	(0.148)
EcoExp × Dem	−0.826	(0.202)	Cutpoint (τ)	3.745	(0.067)
N	2957		ln likelihood	−871.68	

Table 1: **Problematic presentation: FOMC member dissenting votes—Ordered probit parameters.** Estimated ordered probit parameters, with standard errors in parentheses, from the regression of a $j = 3$ category variable on a set of explanatory variables (EVs). Although such nonlinear models are often summarized by tables like this one, especially in the social sciences, it is difficult to discern the effects of the EVs listed at right on the probability of each of the j outcomes. Because the career variables *XXX*Exp are logically constrained to a unit sum, even some of the signs are misleading. The usual quantities of interest for an ordered probit model are not the parameters ($\boldsymbol{\beta}$ and τ), but estimates of $\Pr(y_j|\mathbf{x}_c, \boldsymbol{\beta}, \tau)$ for hypothetical levels of the EVs $\mathbf{x}_c$, which I plot in Figure 1.

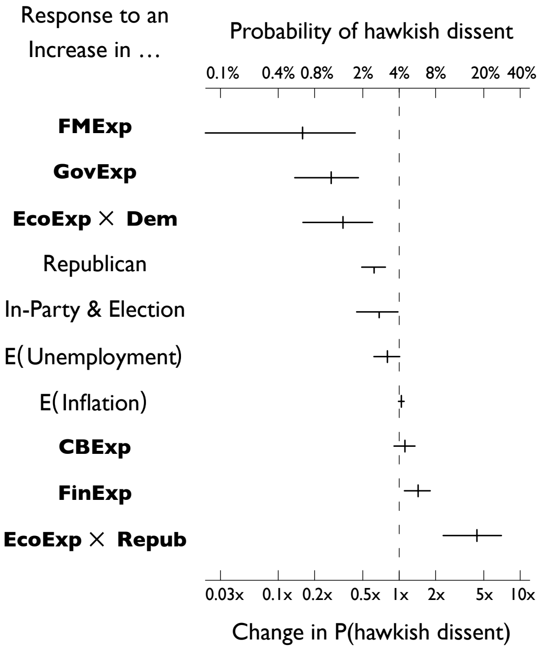
And the hopelessness of directly interpreting ordered probit on compositional covariates

Instead, we used a *dotplot* of probabilities simulated for a series of interesting scenarios



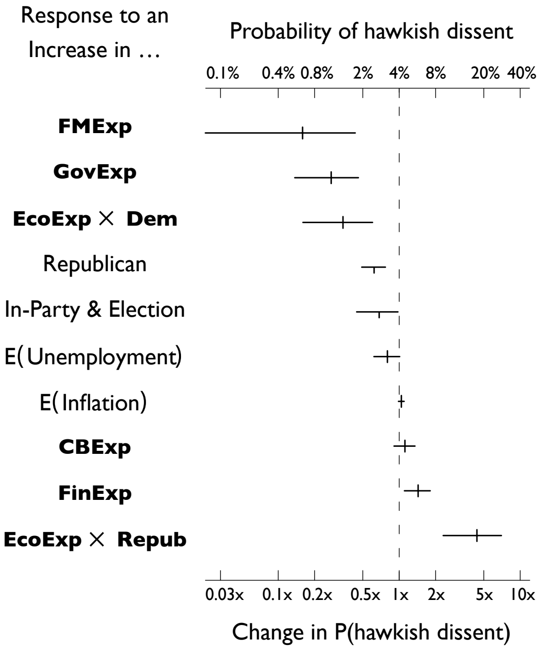
Instead, we used a *dotplot* of probabilities simulated for a series of interesting scenarios

...which we carefully sorted to produce a more readable *diagonalized* presentation



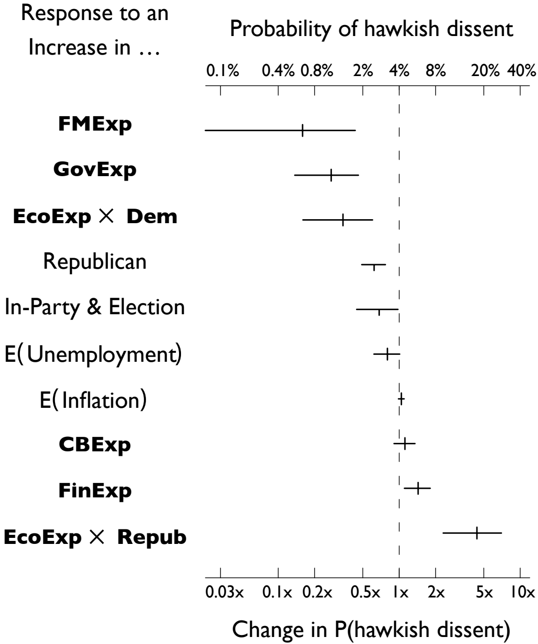
And explained the results in terms of those scenarios, uniting the text of our report with the figure:

"The average central banker dissents in favor of tighter interest rates 4% of the time. In contrast, former treasury officials in the FOMC dissent 0.6% of the time, with a 95% CI from 0.05% to 2%."



And explained the results in terms of those scenarios, uniting the text of our report with the figure:

"Other former bureaucrats issue hawkish dissents 1% of the time [95% CI: 0.5 to 2.0], all else equal."



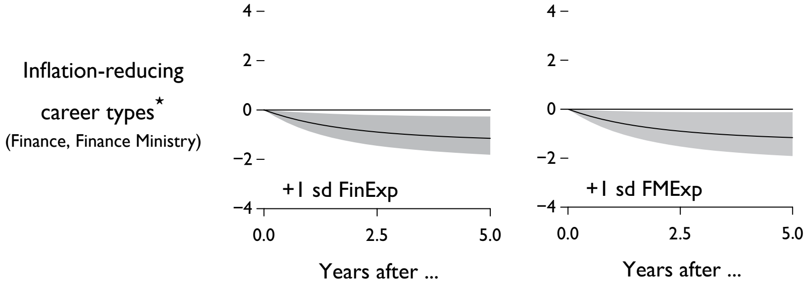
Comparative Inflation Performance

Now instead of studying individual central bankers in the United States, we study a panel of 20 central banks across the industrialized world (pre-Euro data)

We ask what effect the average career composition of the central bank policy board has on inflation

Comparative Inflation Performance

Change in inflation, over time, from changing career composition of the central bank

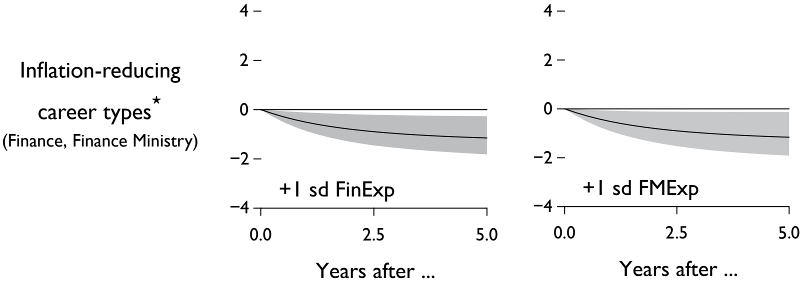


Now instead of studying individual central bankers in the United States, we study a panel of 20 central banks across the industrialized world (pre-Euro data)

We ask what effect the average career composition of the central bank policy board has on inflation

Comparative Inflation Performance

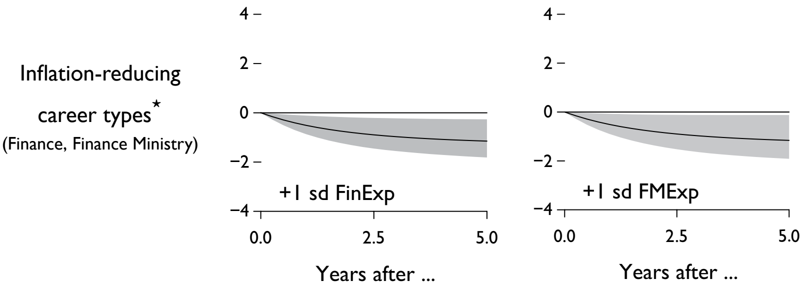
Change in inflation, over time, from changing career composition of the central bank



We imagine a central bank that initially has central bankers with typical career experience (i.e., the global average in each category)

Then, we imagine raising experience in one category (say finance, or FinExp), and use the model to predict how inflation will change over the next 5 years

Change in inflation, over time, from changing career composition of the central bank



Why not just show a coefficient for each career category?

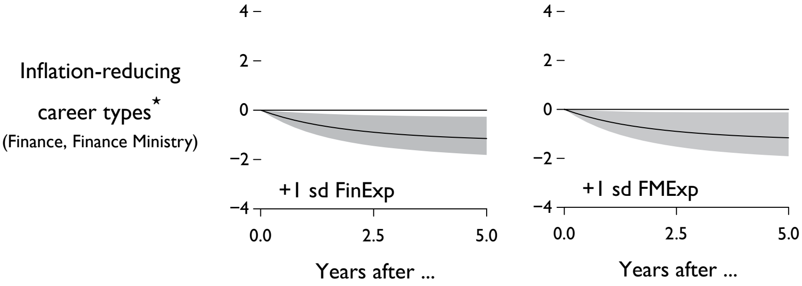
Two reasons to show the first difference in inflation over time:

1. Raising FinExp means lowering the other categories,

so effects are blended across coefficients

Comparative Inflation Performance

Change in inflation, over time, from changing career composition of the central bank



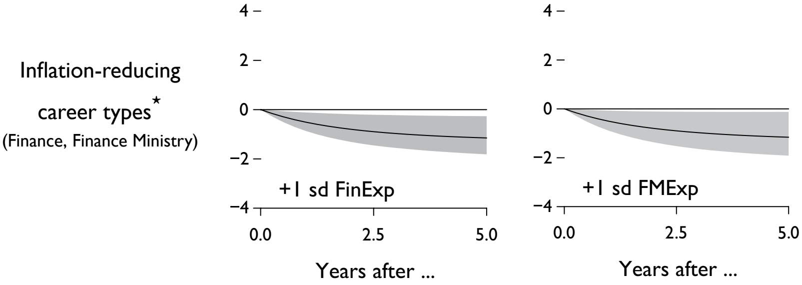
Why not just show a coefficient for each career category?

Two reasons to show the first difference in inflation over time:

2. Effects in time series models build over time;
coefficients show (somewhat arbitrary) first period effects

Comparative Inflation Performance

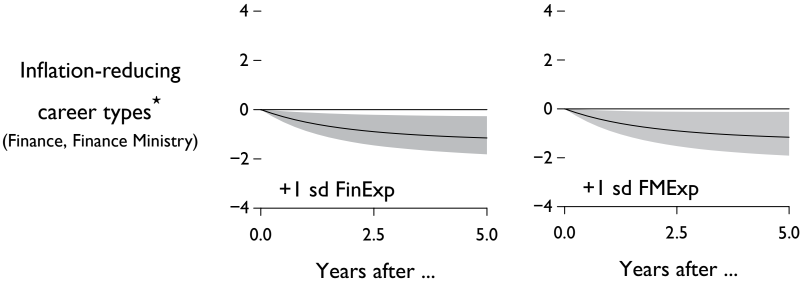
Change in inflation, over time, from changing career composition of the central bank



We simply iterate the KTW simulation algorithm over 5 periods,
computing for each period the difference from inflation under the average board

I used `ldvsimfd()` in the `simcf` package for `R`;
see my course on Panel Data Analysis offered at Essex Summer School

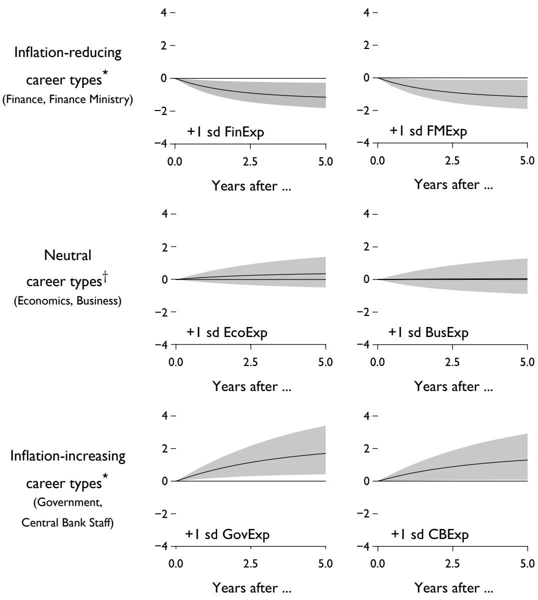
Change in inflation, over time, from changing career composition of the central bank



In the plot above, I show two different scenarios iterated over time: increasing finance experience, or increasing finance ministry experience

Both produce significant reductions in inflation compared to the baseline, and mostly converge to new equilibria after 5 years

Change in inflation, over time, from changing career composition of the central bank



Once we've explained the model, simulation method, and a single plot in our graphic, we can expand to multiple displays

The plot at left replaces an eye-glazing, opaque, and (because of compositional constraints) *misleading* table of regression coefficients

Table 3.7. *Log inflation regressed on central banker characteristics, twenty countries, 1973 to 2000, quarterly.*

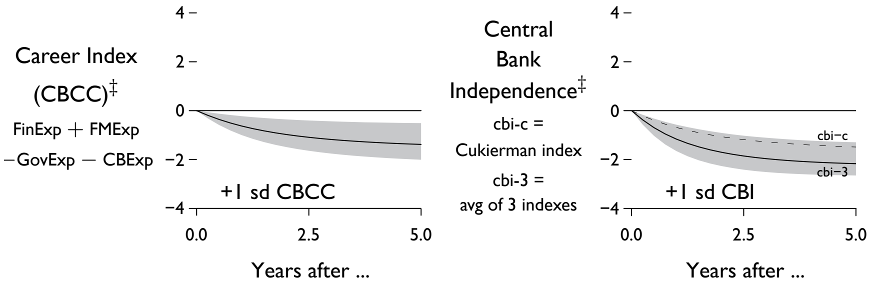
Variable	Expected Sign	DV: ln(Inflation) 1	2	3	4
$\text{FinExp}_{j,t-2}$	$-$	-0.14			-0.09
		(0.08)			(0.07)
$\text{FMExp}_{j,t-2}$	$-/+$	-0.08			-0.13
		(0.06)			(0.06)
$\text{CBExp}_{j,t-2}$	$+/-$	0.12			0.12
		(0.05)			(0.05)
$\text{GovExp}_{j,t-2}$	$+$	0.23			0.19
		(0.08)			(0.08)
$\text{CBI}_{j,t-2}$	$-$	-0.91	-0.92	-0.90	-0.94
		(0.30)	(0.29)	(0.29)	(0.30)
$\text{CBCC}^{\text{med}}_{j,t-2}$	$-$		-0.09	-0.03	
			(0.03)	(0.07)	
$\text{CBI}_{j,t-2} \times \text{CBCC}^{\text{med}}_{j,t-2}$	$-$			-0.12	
				(0.15)	
$(\text{Imports}/\text{GDP})_{j,t-2}$	$-$	-0.02	0.02	0.05	-0.25
		(0.26)	(0.25)	(0.26)	(0.26)
$\%\text{EcDegree}_{j,t-2}$	$-$				0.04
					(0.06)
$\ln \pi_{j,t-1}$		0.97	0.97	0.97	0.96
		(0.04)	(0.04)	(0.04)	(0.04)
$\ln \pi_{j,t-2}$		-0.03	-0.03	-0.03	-0.01
		(0.04)	(0.04)	(0.04)	(0.04)

People often ask, "What if the journal insists on a table instead of the figure?"

In my experience, no one prefers this table to the graph

Give them both, focus your write-up on the graphic, and make sure the graphic explains everything you wanted to get from the table

Change in inflation, over time, from changing career composition of the central bank



No tradeoffs: The small multiple graphs are more accessible to a broad audience and more useful to specialists than a table

You can always include the table as an appendix for those who want to "look under the hood," but cast your argument in terms of the graphics

Wanted: an easy-to-use R package that

1. takes as input the *output* of estimated statistical models
2. makes a variety of plots for model interpretation
3. plots "triples" (lower, estimate, upper) from estimated models well
4. lays out these plots in a tiled arrangement (small multiples)
5. takes care of axes, titles, and other fussy details

Wanted: an easy-to-use R package that

1. takes as input the *output* of estimated statistical models
2. makes a variety of plots for model interpretation
3. plots "triples" (lower, estimate, upper) from estimated models well
4. lays out these plots in a tiled arrangement (small multiples)
5. takes care of axes, titles, and other fussy details

With considerable work, one could

- coerce R's basic graphics to do this badly
- or get `ggplot2` or `lattice` to do this fairly well for a specific case

But an easy-to-use, general solution is lacking

The `tile` package

My answer is the `tile` package, written using `R`'s `grid` graphics

Some basic tile graphic types:

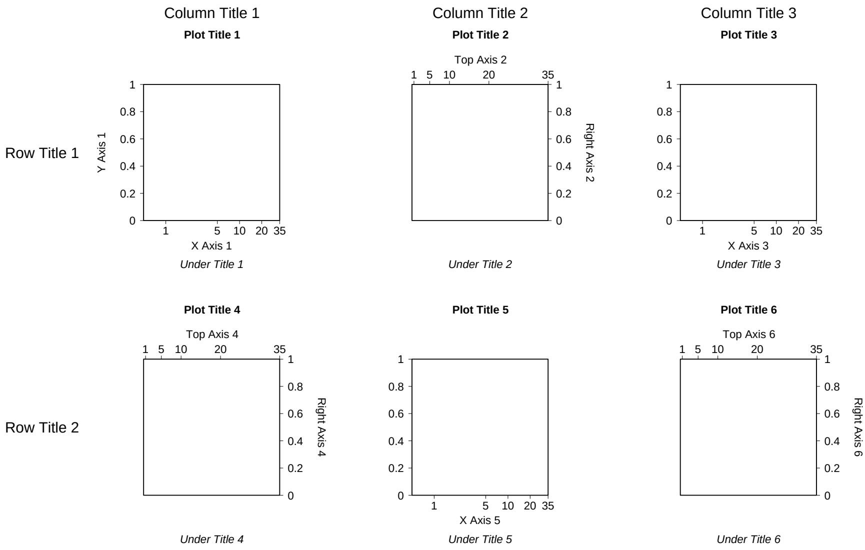
scatter	Scatterplots with fits, CIs, and extrapolation checking
lineplot	Line plots with fits, CIs, and extrapolation checking
ropeladder	Dot plots with CIs and extrapolation checking

Each can take as input draws from the posterior of a regression model

A call to a tile function makes a multiplot layout:

ideal for small multiples of model parameters

An example `tile` layout, minus traces



Three steps to make tile plots

1. **Create data traces.** Each trace contains the data and graphical parameters needed to plot a single set of graphical elements to one or more plots.

Three steps to make tile plots

1. **Create data traces.** Each trace contains the data and graphical parameters needed to plot a single set of graphical elements to one or more plots.
   - Could be a set of points, *or* text labels, *or* lines, *or* a polygon

Three steps to make tile plots

1. **Create data traces.** Each trace contains the data and graphical parameters needed to plot a single set of graphical elements to one or more plots.
    - Could be a set of points, *or* text labels, *or* lines, *or* a polygon
    - Could be a set of points *and* symbols, colors, labels, fit line, CIs, and/or extrapolation limits

Three steps to make tile plots

1. **Create data traces.** Each trace contains the data and graphical parameters needed to plot a single set of graphical elements to one or more plots.
   - Could be a set of points, *or* text labels, *or* lines, *or* a polygon
   - Could be a set of points *and* symbols, colors, labels, fit line, CIs, and/or extrapolation limits
   - Could be the data for a dotchart, with labels for each line

Three steps to make tile plots

1. **Create data traces.** Each trace contains the data and graphical parameters needed to plot a single set of graphical elements to one or more plots.
   - Could be a set of points, *or* text labels, *or* lines, *or* a polygon
   - Could be a set of points *and* symbols, colors, labels, fit line, CIs, and/or extrapolation limits
   - Could be the data for a dotchart, with labels for each line
   - Could be the marginal data for a rug
   - **All** annotation must happen in this step

Three steps to make tile plots

1. **Create data traces.** Each trace contains the data and graphical parameters needed to plot a single set of graphical elements to one or more plots.
   - Could be a set of points, *or* text labels, *or* lines, *or* a polygon
   - Could be a set of points *and* symbols, colors, labels, fit line, CIs, and/or extrapolation limits
   - Could be the data for a dotchart, with labels for each line
   - Could be the marginal data for a rug
   - **All** annotation must happen in this step
   - Primitive traces: `linesTile()`, `pointsile()`, `polygonTile()`, `polylinesTile()`, `textTile()`, and `circleTile()`
   - Complex traces: `lineplot()`, `scatter()`, `ropeladder()`, and `rugTile()`

Trace functions in tile

Primitive trace functions:

`linesTile`	Plot a set of connected line segments
`pointsTile`	Plot a set of points
`polygonTile`	Plot a shaded region
`polylinesTile`	Plot a set of unconnected line segments
`circleTile`	Plot a set of circles
`textTile`	Plot text labels

Complex traces for model or data exploration:

`lineplot`	Plot lines with confidence intervals, extrapolation warnings
`ropeladder`	Plot dotplots with confidence intervals, extrapolation warnings, and shaded ranges
`rugTile`	Plot marginal data rugs to axes of plots
`scatter`	Plot scatterplots with text and symbol markers, fit lines, and confidence intervals

Three steps to make tile plots

1. **Create data traces.** Each trace contains the data and graphical parameters needed to plot a single set of graphical elements to one or more plots.

2. **Plot the data traces.** Using the `tile()` function, simultaneously plot all traces to all plots.

Three steps to make tile plots

1. **Create data traces.** Each trace contains the data and graphical parameters needed to plot a single set of graphical elements to one or more plots.

2. **Plot the data traces.** Using the `tile()` function, simultaneously plot all traces to all plots.

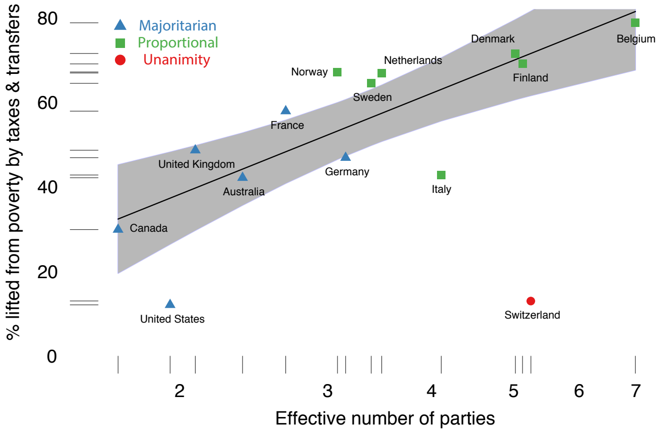
   - This is the step where the scaffolding gets made: axes and titles
   - Set up the rows and columns of plots
   - Titles of plots, axes, rows of plots, columns of plots, etc.
   - Set up axis limits, ticks, tick labels, logging of axes

Three steps to make tile plots

1. **Create data traces.** Each trace contains the data and graphical parameters needed to plot a single set of graphical elements to one or more plots.

2. **Plot the data traces.** Using the `tile()` function, simultaneously plot all traces to all plots.

3. **Examine output and revise.** Look at the graph made in step 2, and tweak the input parameters for steps 1 and 2 to make a better graph.



Let's make this plot

CODE EXAMPLE

`inequalityScatter.R`

Plot simulations of QoI

Generally, we want to plot triples: lower, estimate, upper

We could do this for specific **discrete scenarios**, e.g.

Pr(Voting) given five distinct sets of x's

Recommended plot: **Dotplot** with confidence interval lines

Plot simulations of QoI

Generally, we want to plot triples: lower, estimate, upper

We could do this for specific **discrete scenarios**, e.g.

Pr(Voting) given five distinct sets of x's

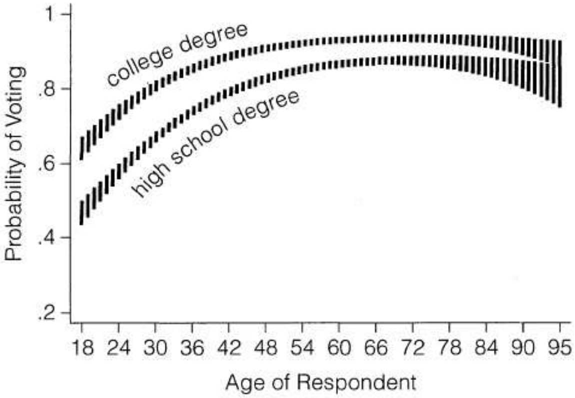
Recommended plot: **Dotplot** with confidence interval lines

Or for a continuous **stream of scenarios**, e.g.,

Hold all but Age constant, then calculate Pr(Voting) at every level of Age

Recommended plot: **Lineplot** with shaded confidence intervals

FIGURE 1 Probability of Voting by Age

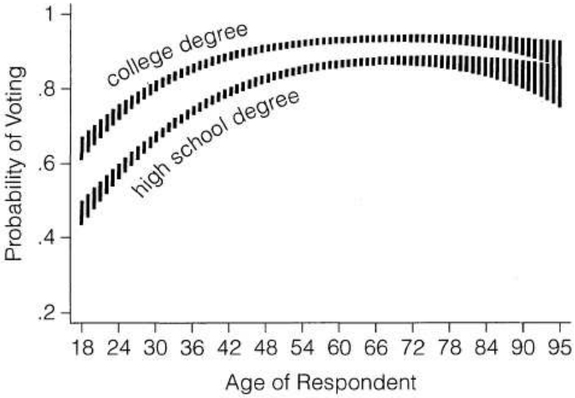


Vertical bars indicate 99-percent confidence intervals

This example is obviously superior to the table of logit coefficients

But is there anything wrong or missing here?

Figure 1 Probability of Voting by Age



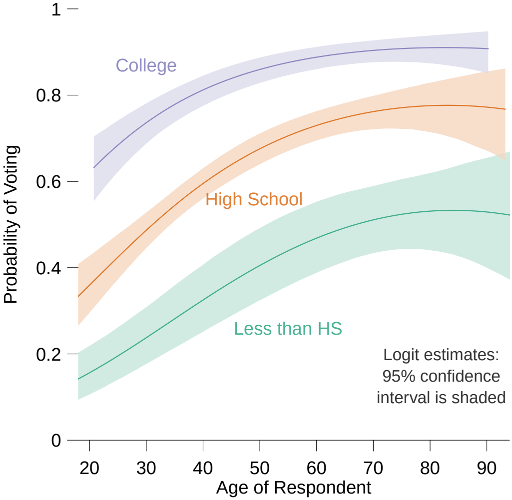
Vertical bars indicate 99-percent confidence intervals

This example is obviously superior to the table of logit coefficients

But is there anything wrong or missing here?

18 year old college grads?!

And what about high school dropouts?



Here is the graphic redrawn in tile

`tile` helps us systematize plotting model results, and helps avoid unwanted extrapolation by limiting results to the convex hull

CODE EXAMPLE

`votingLineplots.R`

Voting Example Redux

Next step: learn to simulate and plot first differences and relative risks

We could do this with our current example.
    E.g., hold age fixed and compute the change in Pr(Vote)
    given an increase in education

But for pedagogical reasons, it will be more useful to add an additional covariate

We now add to our voting model whether the respondent was married

Theory: Marriage should increase voting by increasing concern
for a variety of public goods, or by forming ties to a local community, etc.

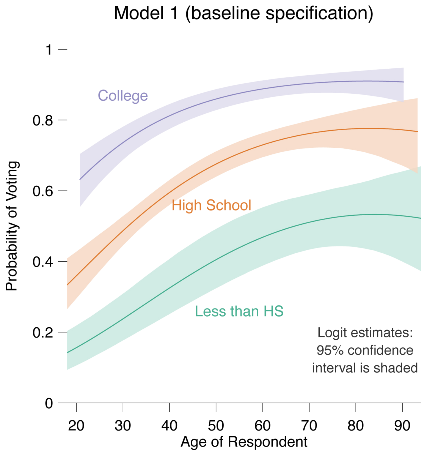
How would this competing model normally be presented?

Logit of Decision to Vote, 2000 Presidential NES

	M1	M2
Age	0.075	0.061
	(0.017)	(0.017)
Age^2	-0.0004	-0.0003
	(0.0002)	(0.0002)
High School Grad	1.124	1.099
	(0.180)	(0.181)
College Grad	1.080	1.053
	(0.131)	(0.132)
Married		0.373
		(0.110)
Constant	-3.019	-2.866
	(0.418)	(0.421)
log likelihood	-1101.370	-1099.283
N	1783	1783

Comparing Logistic Regression Models

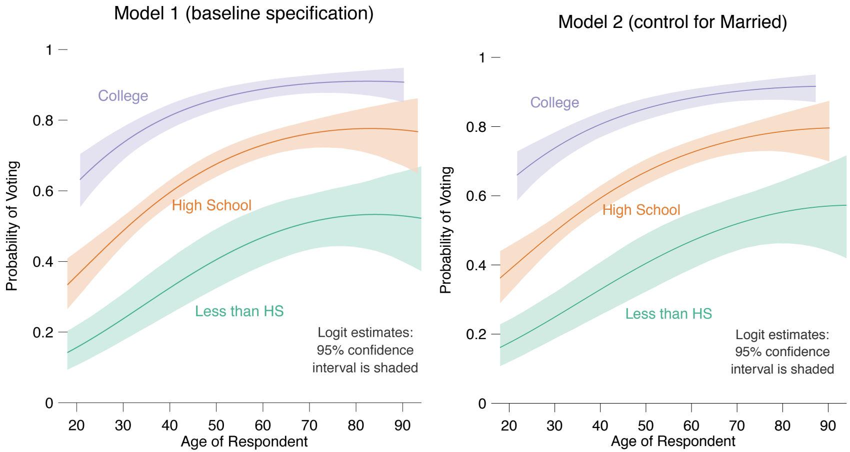
But we can also compare our results in an intelligible way.



Effects of Age and Education haven't discernably changed

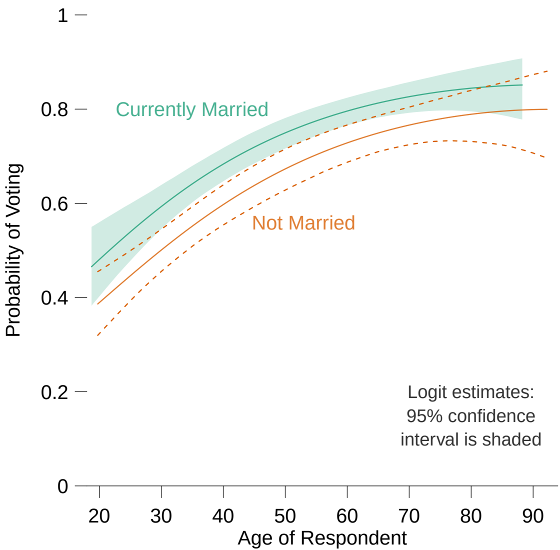
Comparing Logistic Regression Models

But we can also compare our results in an intelligible way.



Our first attempt to show model robustness – we'll find more efficient ways

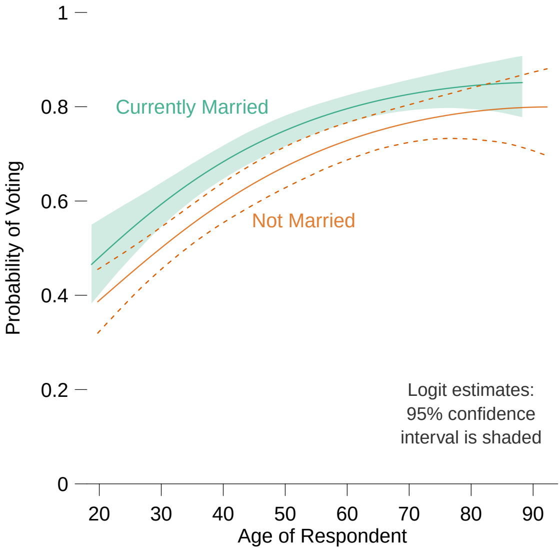
A common misconception about confidence intervals



Ceteris paribus, marriage has a moderate effect.

Is this effect statistically significant?

A common misconception about confidence intervals



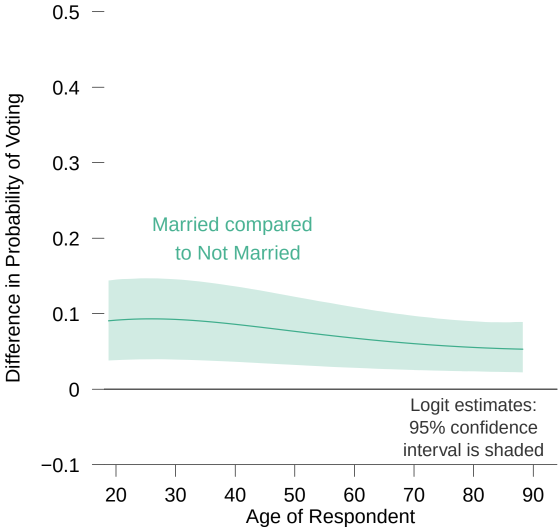
Ceteris paribus, marriage has a moderate effect.

Is this effect statistically significant?

Overlapping CIs of expected values doesn't always imply an insignificant difference

This is not necessarily the case in "close calls"

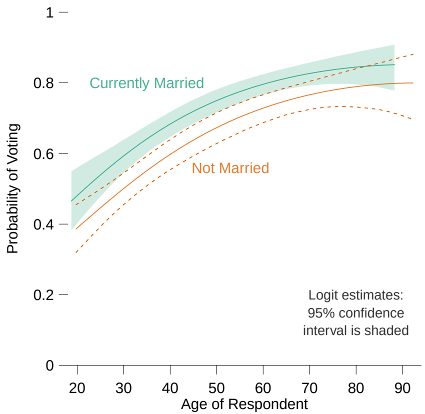
A common misconception about confidence intervals

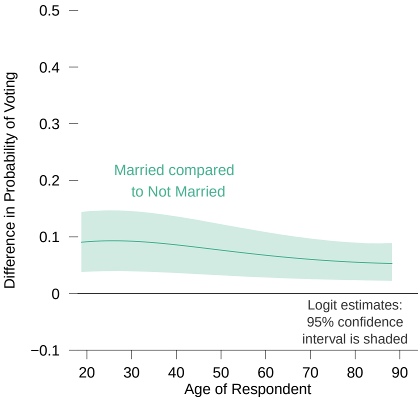


The right way to assess statistical significance: simulate the CI of the first difference (or relative risk) *directly*

This first difference is always bounded away from zero, hence always significant

Avoid mistakenly rejecting significant first differences

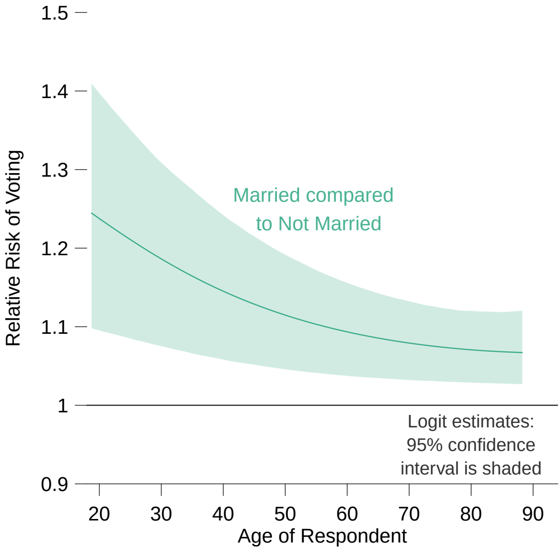




Expected values estimate both difference & *location*;
demanding a more detailed estimate from the model increases uncertainty

First differences and relative risks estimate the difference only, so they have slightly tighter confidence intervals

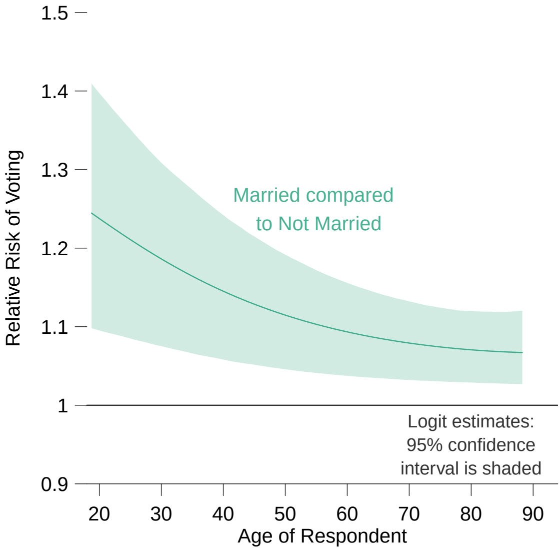
Relative risk plots



Consider showing relative risks instead of (or in support of) first differences

Relative risks show "how many times more likely" a categorical outcome is under the counterfactual

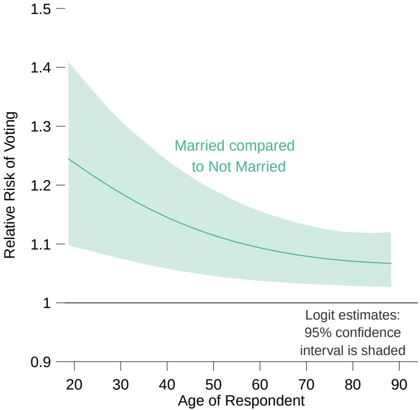
Relative risk plots



For continuous outcomes, RR shows how many times bigger the outcome is under the counterfactual

As with first differences, relative risk should be simulated directly to calculate CIs correctly

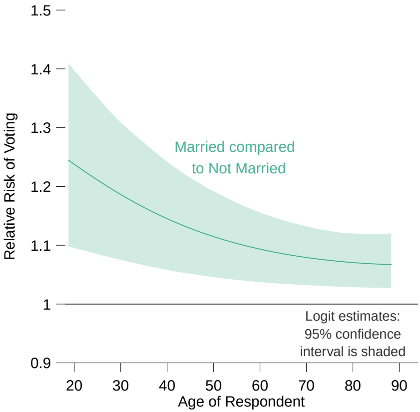
Setting up before-and-after scenarios



Setting up counterfactuals for FDs or RRs is tricky, as we will see in the code

Here I set before and after *age* to the same value (which varied across the plot) but I set Married to different values (0 before, 1 after)

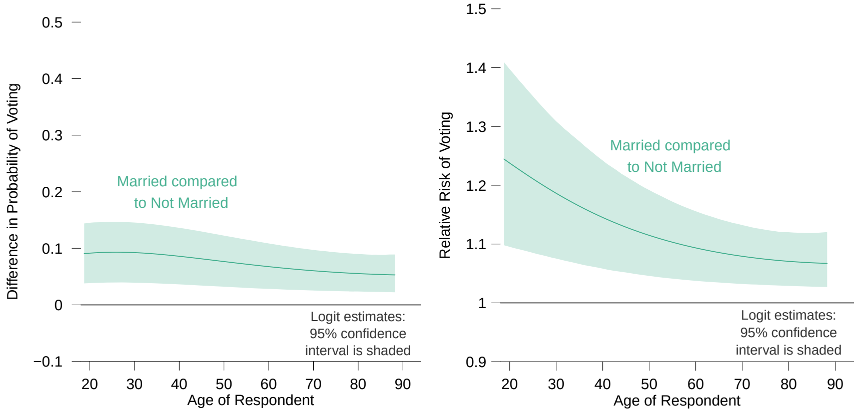
Setting up before-and-after scenarios



Take care in selecting the before and after values of *all* covariates

Most common place to make mistakes, with *huge* substantive consequences

Setting up before-and-after scenarios



RETURN TO CODE EXAMPLE

`votingLineplots.R`

Robustness Checks

So far, we've presenting conditional expectations & differences from regressions

But are we confident that these were the "right" estimates?

The language of inference usually assumes we

- correctly specified our model
- correctly measured our variables
- chose the right probability model
- don't have influential outliers, etc.

Robustness Checks

So far, we've presenting conditional expectations & differences from regressions

But are we confident that these were the "right" estimates?

The language of inference usually assumes we

- correctly specified our model
- correctly measured our variables
- chose the right probability model
- don't have influential outliers, etc.

We're never completely sure these assumptions hold.

Most people present one model, and argue it was the best choice

Sometimes, a few alternatives are displayed

The race of the variables

	Model 1	Model 2	Model 3	Model 4	Model 5
My variable of interest, x_1	X.XX (X.XX)	X.XX (X.XX)	X.XX (X.XX)	X.XX (X.XX)	
A control I "need"	X.XX (X.XX)	X.XX (X.XX)	X.XX (X.XX)	X.XX (X.XX)	X.XX (X.XX)
A control I "need"	X.XX (X.XX)	X.XX (X.XX)	X.XX (X.XX)	X.XX (X.XX)	X.XX (X.XX)
A candidate control		X.XX (X.XX)		X.XX (X.XX)	
A candidate control			X.XX (X.XX)	X.XX (X.XX)	
Alternate measure of x_1					X.XX (X.XX)

Robustness Checks

Problems with the approach above?

1. Lots of space to show a few permutations of the model

   Most space wasted or devoted to ancillary info

Robustness Checks

Problems with the approach above?

1. Lots of space to show a few permutations of the model

   Most space wasted or devoted to ancillary info

2. What if we're really interested in $E(y|X)$, not $\hat{\beta}$?

   E.g., because of nonlinearities, interactions, scale differences, etc.

Robustness Checks

Problems with the approach above?

1. Lots of space to show a few permutations of the model

   Most space wasted or devoted to ancillary info

2. What if we're really interested in $E(y|X)$, not $\hat{\beta}$?

   E.g., because of nonlinearities, interactions, scale differences, etc.

3. The selection of permutations is *ad hoc*.

Robustness Checks

Problems with the approach above?

1. Lots of space to show a few permutations of the model

   Most space wasted or devoted to ancillary info

2. What if we're really interested in $E(y|X)$, not $\hat{\beta}$?

   E.g., because of nonlinearities, interactions, scale differences, etc.

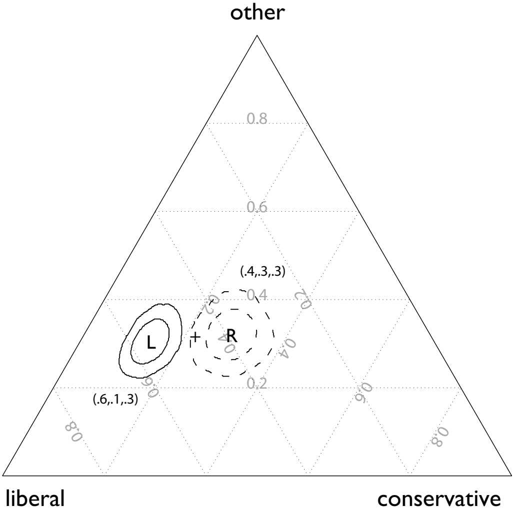
3. The selection of permutations is *ad hoc*.

We'll try to fix 1 & 2.

Objection 3 is harder, but worth thinking about.

Robustness Checks: An algorithm

1. Identify a relation of interest between a concept $\mathcal{X}$ and a concept $\mathcal{Y}$
2. Choose:
   - a measure of $\mathcal{X}$, denoted X,
   - a measure of $\mathcal{Y}$, denoted y,
   - a set of confounders, Z,
   - a functional form, $g(\cdot)$
   - a probability model of y, $f(\cdot)$
3. Estimate the probability model $y \sim f(\mu, \alpha)$, $\mu = g(\text{vec}(X, Z), \beta)$.
4. Simulate the quantity of interest such as $E(y|X)$, $E(y_2 - y_1|X_1, X_2)$, or $E(y_2/y_1|X_1, X_2)$ to obtain a point estimate and confidence interval.
5. Repeat 2–4, changing at each iteration one of the choices in step 2.
6. Compile the results in a variant of the dot plot called a *ropeladder*.

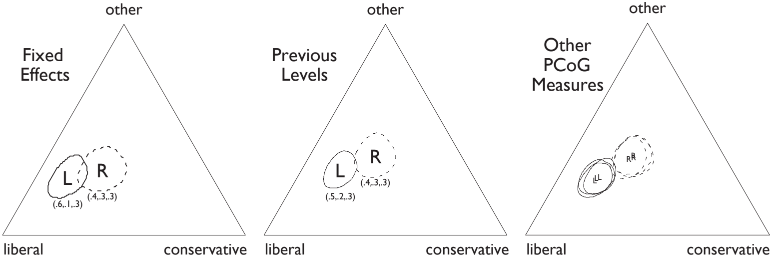


In **Gallery 7**, we reviewed a compositional data model from Ch. 8 of *Bankers, Bureaucrats, and Central Bank Politics*

We used a ternary plot to show the career composition of appointed central bankers depends on the partisanship of the appointing government

How would we show robustness under alternative specifications?

One tool for robustness: Small Multiples & Backup Dancers

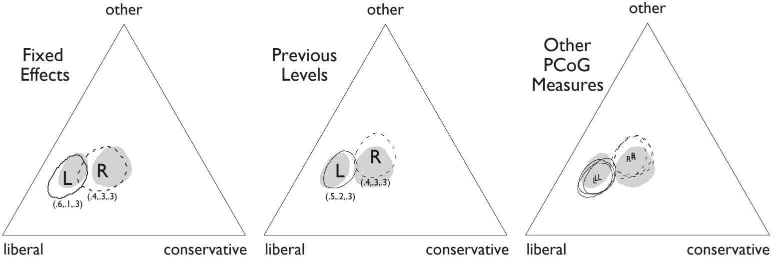


Once people understand ternary plots,
they will immediately absorb a small, simplified version

Each of these small multiples shows our result under a different model

The similarity of each plot is immediately obvious here

One tool for robustness: Small Multiples & Backup Dancers



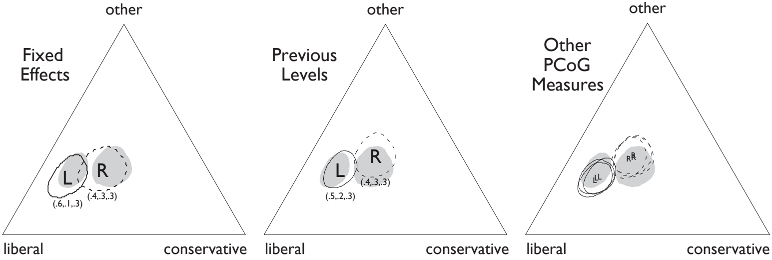
Once people understand ternary plots,
they will immediately absorb a small, simplified version

Each of these small multiples shows our result under a different model

The similarity of each plot is immediately obvious here

If not, putting the original plot in gray in the background helps:
Amanda Cox call these “backup dancers”

One tool for robustness: Small Multiples & Backup Dancers

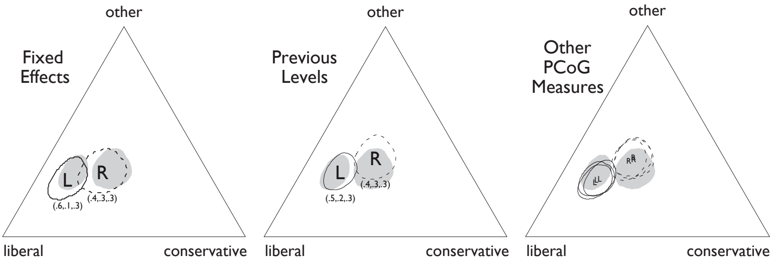


But in this case, I need *lots* of robustness checks

Because of the multiple equations, my statistical model is so demanding it's hard to include many regressors at once

If I try them one at a time, I would fill pages with triangle plots

One tool for robustness: Small Multiples & Backup Dancers

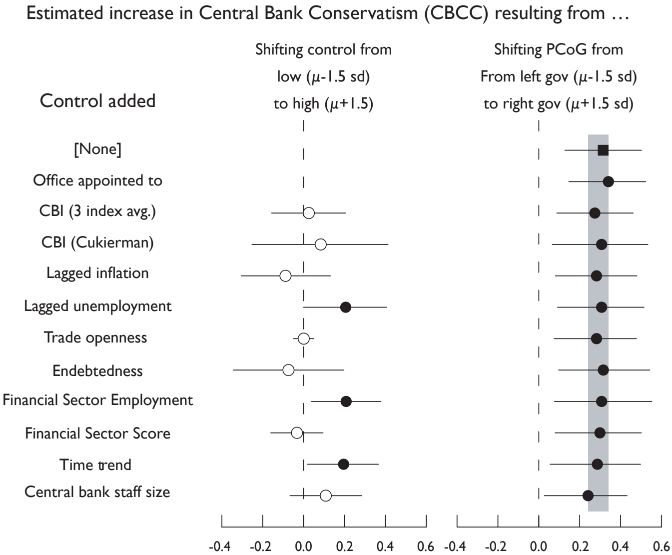


However, the horizontal dimension is the substantively important one:
the one that affects affects economic outcomes

So I create a new QoI: Central Banker Career Conservatism (CBCC)

CBCC = Conservative Experience - Liberal Experience

And use my model to predict changes in CBCC and plot them on a *ropeladder*


Estimated increase in Central Bank Conservatism (CBCC) resulting from …
Shifting control from
low (μ-1.5 sd)
to high (μ+1.5)
Shifting PCoG from
From left gov (μ-1.5 sd)
to right gov (μ+1.5 sd)
Control added
[None]
Office appointed to
CBI (3 index avg.)
CBI (Cukierman)
Lagged inflation
Lagged unemployment
Trade openness
Endebtedness
Financial Sector Employment
Financial Sector Score
Time trend
Central bank staff size
-0.4 -0.2 0.0 0.2 0.4 0.6
-0.4 -0.2 0.0 0.2 0.4 0.6

Anatomy of a ropeladder plot

I call this a **ropeladder** plot.

The column of dots shows the relationship between *y* and a specific *X* under different model assumptions

Each entry corresponds to a different assumption about the specification, or the measures, or the estimation method, etc.

Anatomy of a ropeladder plot

I call this a **ropeladder** plot.

The column of dots shows the relationship between y and a specific X under different model assumptions

Each entry corresponds to a different assumption about the specification, or the measures, or the estimation method, etc.

If all the dots line up, with narrow, similar CIs, we say the finding is robust, and reflects the data under a range of reasonable assumptions

If the ropeladder is "blowing in the wind", we may be skeptical of the finding.
It depends on model assumptions that may be controversial

Anatomy of a ropeladder plot

I call this a **ropeladder** plot.

The column of dots shows the relationship between y and a specific X under different model assumptions

Each entry corresponds to a different assumption about the specification, or the measures, or the estimation method, etc.

If all the dots line up, with narrow, similar CIs, we say the finding is robust, and reflects the data under a range of reasonable assumptions

If the ropeladder is “blowing in the wind”, we may be skeptical of the finding.
It depends on model assumptions that may be controversial

The shaded gray box shows the full range of the point estimates for the QoI.

Narrow is better.

Why ropeladders?

1. Anticipate objections on model assumptions, and have concrete answers.

   Avoid: "I ran it that other way, and it came out the 'same."'

   Instead: "I ran it that other way, and look – it made no *substantive* or *statistical* difference worth speaking of."

   Or: "...it makes *this* much difference."

Why ropeladders?

1. Anticipate objections on model assumptions, and have concrete answers.

   Avoid: "I ran it that other way, and it came out the 'same."'

   Instead: "I ran it that other way, and look – it made no *substantive* or *statistical* difference worth speaking of."

   Or: "...it makes *this* much difference."

2. Investigate robustness more thoroughly

   Traditional tabular presentation would have run to 7 pages, making comparison hard and discouraging a thorough search

Why ropeladders?

1. Anticipate objections on model assumptions, and have concrete answers.

   Avoid: "I ran it that other way, and it came out the 'same."'

   Instead: "I ran it that other way, and look – it made no *substantive* or *statistical* difference worth speaking of."

   Or: "...it makes *this* much difference."

2. Investigate robustness more thoroughly

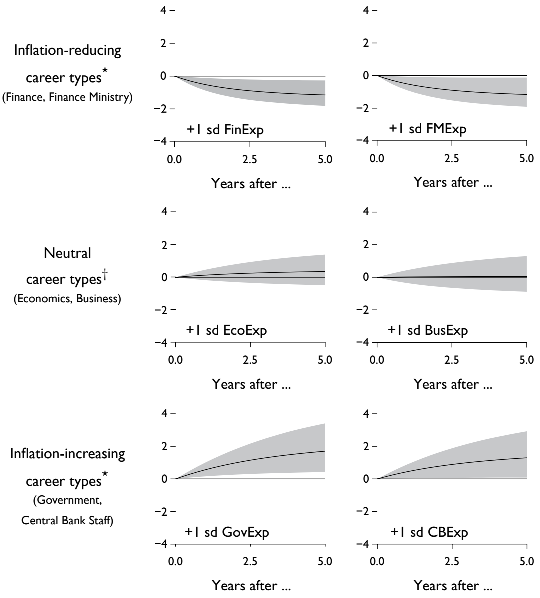
   Traditional tabular presentation would have run to 7 pages, making comparison hard and discouraging a thorough search

3. Find patterns of model sensitivity

   Two seemingly unrelated changes in specification had the same effect. (Unemployment and Financial Sector Size)

   Turned out to be a missing third covariate (time trend)

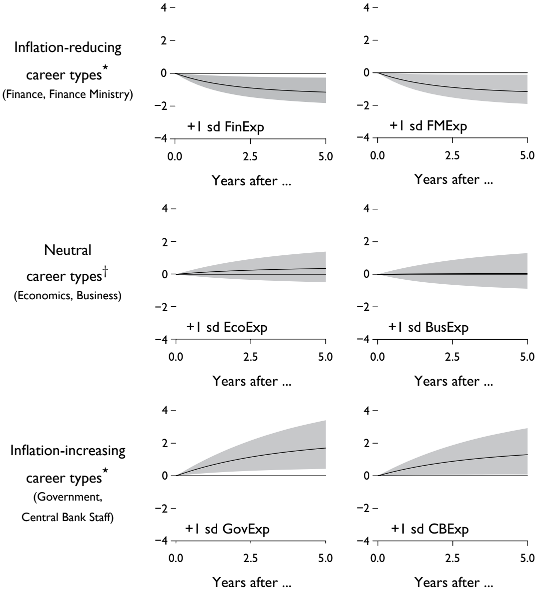
Change in inflation, over time, from changing career composition of the central bank



Recall the TSCS model of inflation performance

How would we show robustness here?

Change in inflation, over time, from changing career composition of the central bank



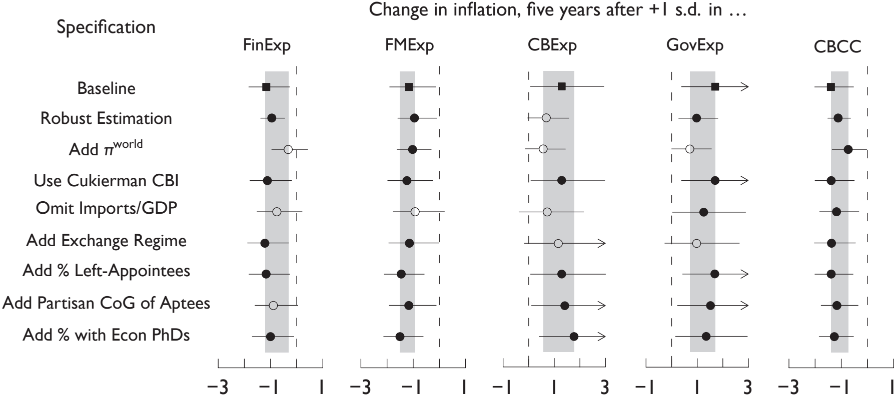
Recall the TSCS model of inflation performance

How would we show robustness here?

Once we understand the dynamics over time, we can simplify our presentation

What if we isolate the 5 year mark, and compare the estimated effects of covariate on inflation at that point under different models?

Robustness for several QoIs at once

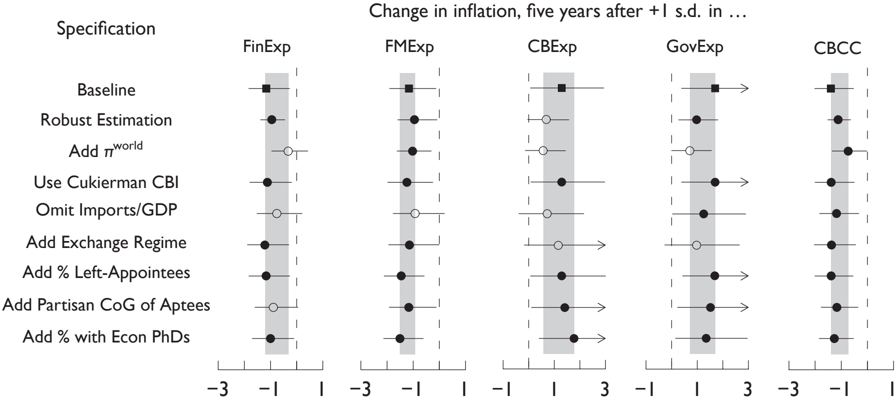


Each ropeladder, or column, shows the effect of a different variable on the response

That is, reading *across* shows the results from a single model

Reading *down* shows the results for a single question across different models

Robustness for several QoIs at once

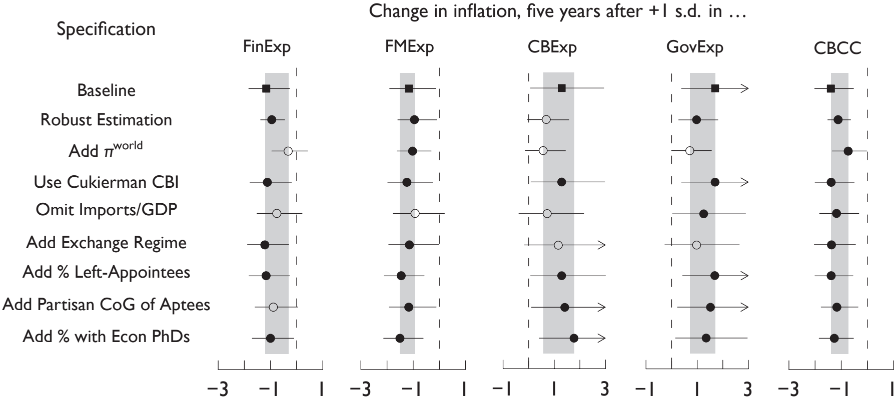


Arrows indicate confidence intervals that extend outside the plot

Choosing our own plotting area using `limits=` is critical for ropeladders

Focus on the area with the point estimates and on any problematic CIs

Robustness for several QoIs at once



To write up robustness, show this graphic and relegate tables to the appendix

You can be specific about the nature of robustness (no “hand-waving”),

yet still write up 8 robustness checks on 5 covariates in 2 pages total

Our tools are flexible

From the first few examples,
you might think lineplots are for model inference
and dotplots (ropeladders) are for model robustness

But these tools are *flexible* and reward creativity

In the following examples,
I use dotplots made with `ropeladder()` to explore models,
then use lineplots to explore robustness

When simulation is the only option: Chinese leadership

Shih, Adolph, and Liu investigate the advancement of elite Chinese leaders in the Reform Period (1982–2002)

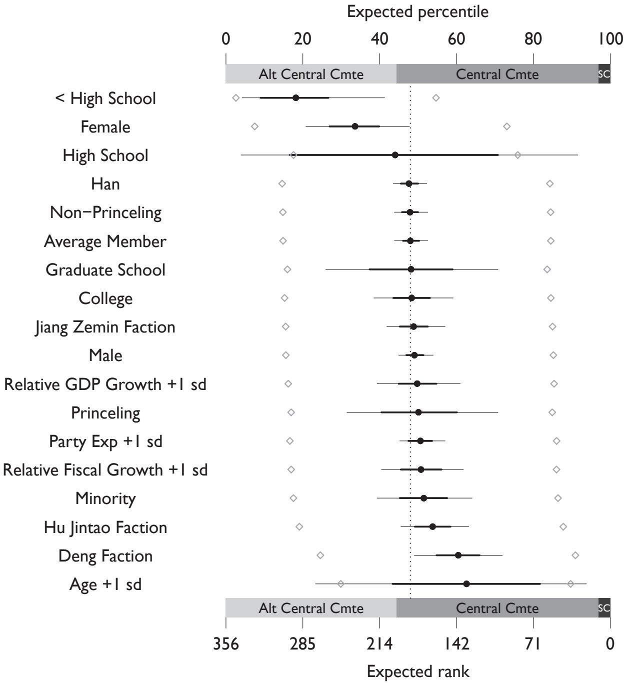
Explain (partially observed) ranks of the top 300 to 500 Chinese Communist Party leaders as a function of:

Demographics	*age, sex, ethnicity*
Education	*level of degree*
Performance	*provincial growth, revenue*
Faction	*birth, school, career, and family ties to top leaders*

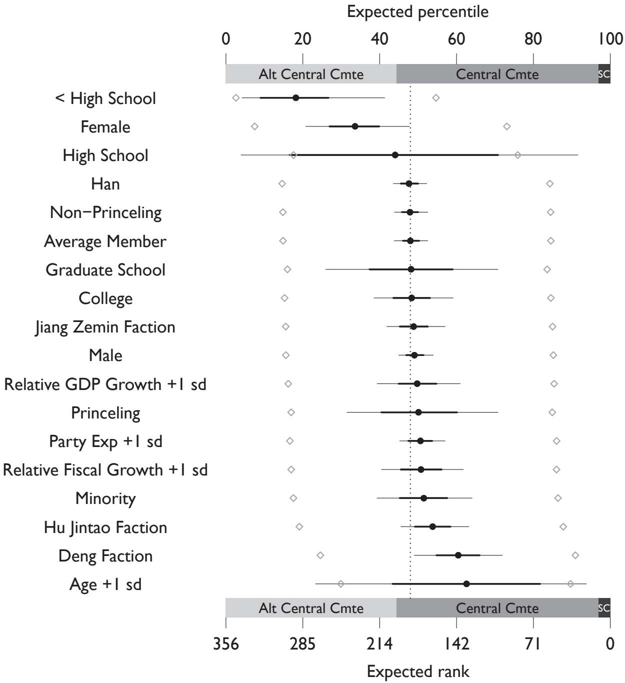
Bayesian model of partially observed ranks of CCP officials

Model parameters difficult to interpret: on a latent scale
and individual effects are conditioned on all other ranked members

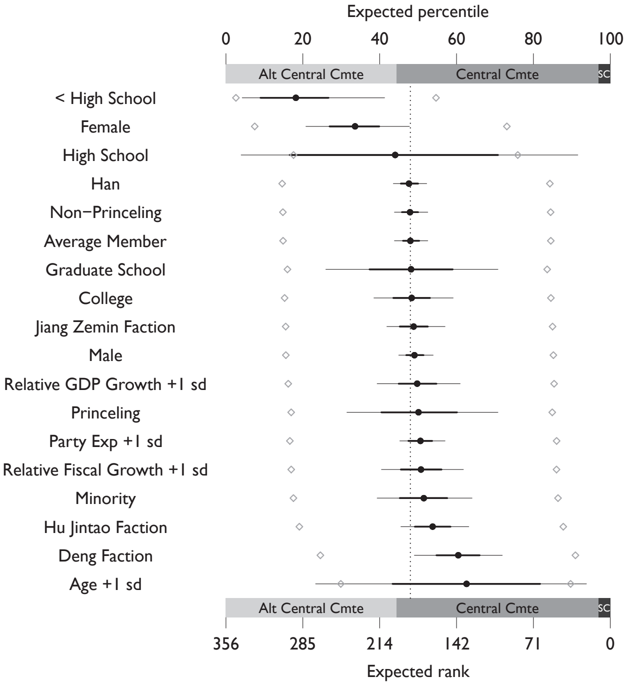
Only solution:
Simulate ranks of hypothetical officials as if placed in the observed hierarchy



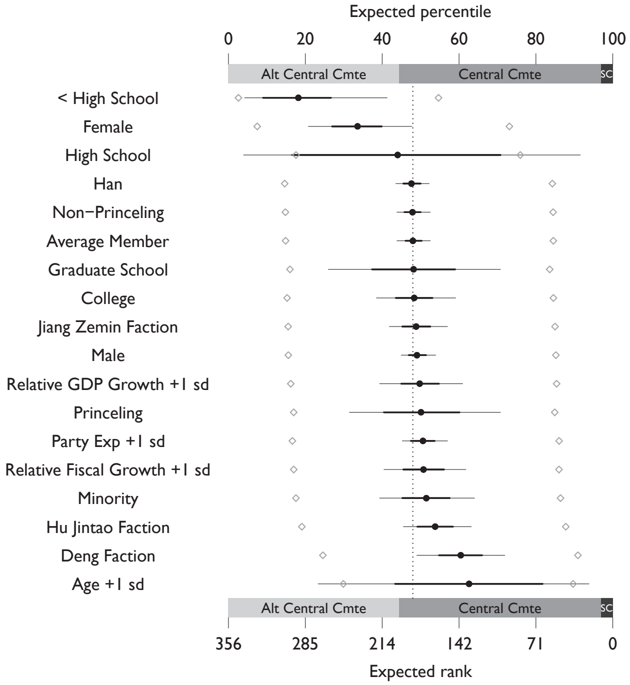
Black circles show expected ranks for otherwise average Chinese officials with the characteristic listed at left



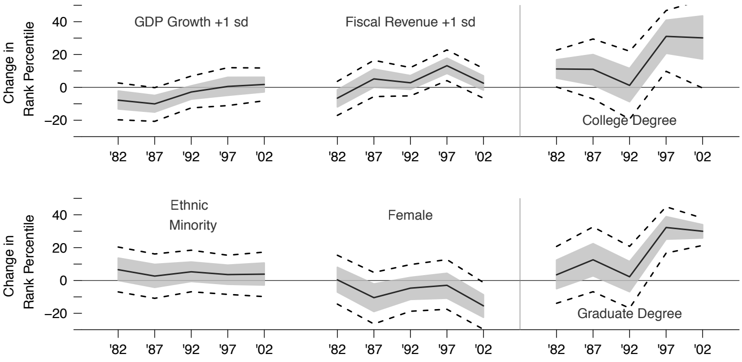
Thick black horizontal lines are 1 std error bars, and thin lines are 95% CIs



Gray triangles are officials with random effects at ± 1 sd; how much unmeasured factors matter

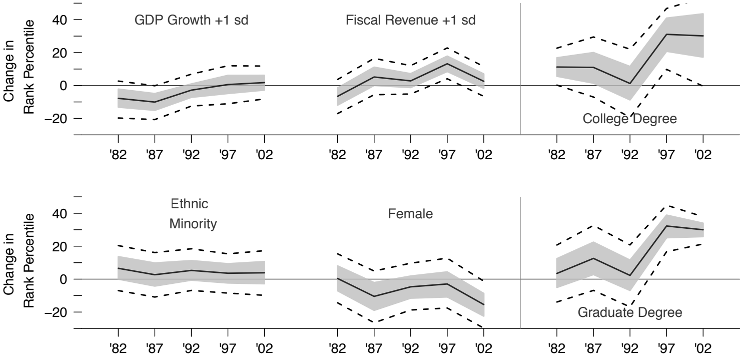


It helps to sort rows of the plot from smallest to largest effect (diagonalization)



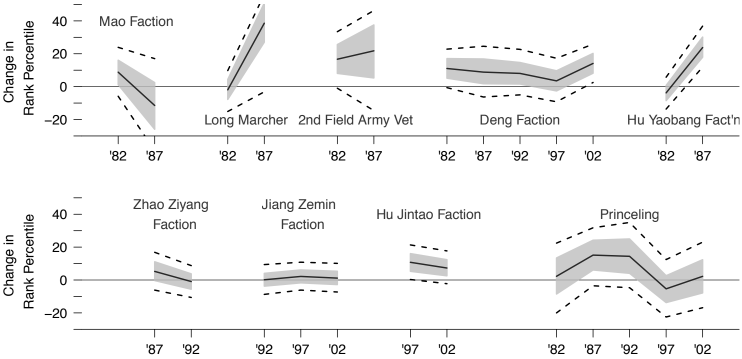
We re-estimate the model *separately* for each year,
leading to a large number of results with varying sets of covariates

A complex lineplot helps organize these results and facilitate comparisons



Note that these results are now *first differences*:

the expected percentile change in rank for an otherwise average official who gains the characteristic noted



Over time, officials' economic performance never matters, but factions often do

Runs counter to the conventional wisdom that meritocratic selection of officials lies behind Chinese economic success

Chinese Officials Robustness

Our findings were controversial: countered the widely accepted belief that Chinese officials are rewarded for economic performance

Critics asked for lots of alternative specifications to probe our results

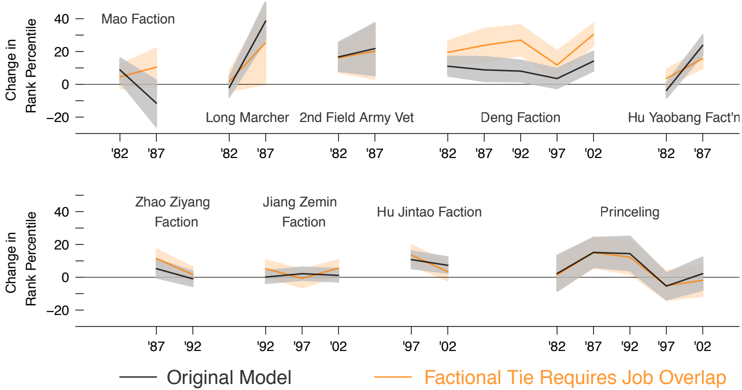
We used `tile` to show exactly what difference these robustness checks made using overlapping lineplots

We provide detailed one-to-one comparisons of our model with each alternative, for a lengthy appendix...

...And a single page summary for the printed article collecting all robustness checks

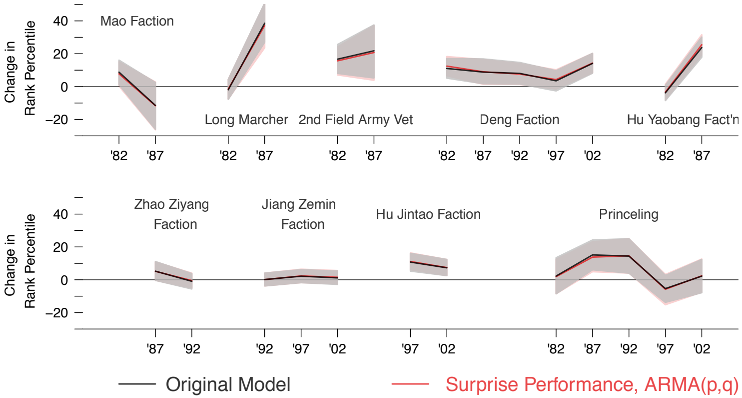
Some critics worried that our measures of faction were too sensitive, so we considered a more specific alternative

This didn't salvage the conventional wisdom on growth...



But did (unsuprisingly) strengthen our factional results

(Specific measures pick up the strongest ties)

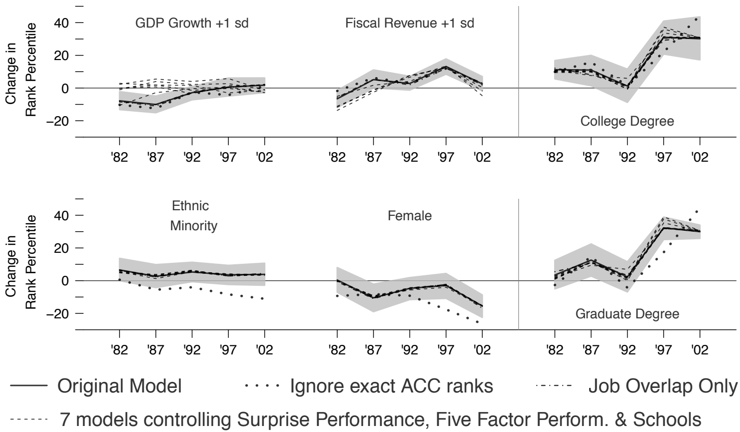


Other critics worried about endogeneity or selection effects flowing from political power to economic performance

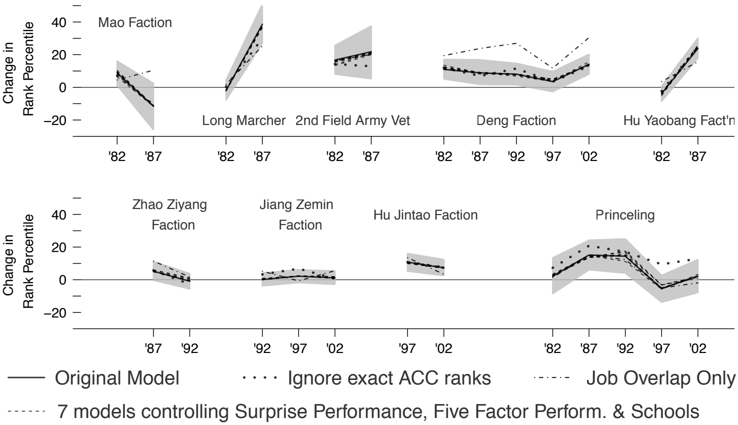
We used measures of unexpected growth to zero in on an official's own performance in office – which still nets zero political benefit

The above summarizes results combined from 2 versions of a model applied over 5 periods, each with 5 multiply imputed datasets (50 models)

But it still takes many pages to show all our robustness checks.
Is there a more efficient way to show that our results stay essentially the same?



In our printed article, we show only this plot,
which overlays the full array of robustness checks



Conveys hundreds of separate findings in a compact, readable form

No knowledge of Bayesian methods or partial rank coefficients required!

More uses of ropeladder traces for model exploration

We will discuss implementation of ropeladders – for robustness and general model inference – shortly

But first, let's explore three more uses of ropeladder dotplots that show off the full range of features of these traces:

- Exploring interactive models using differences-in-differences
- Grouping variables and interactions for easier comprehension and explanation
- Grouping categorical responses to multinomial models

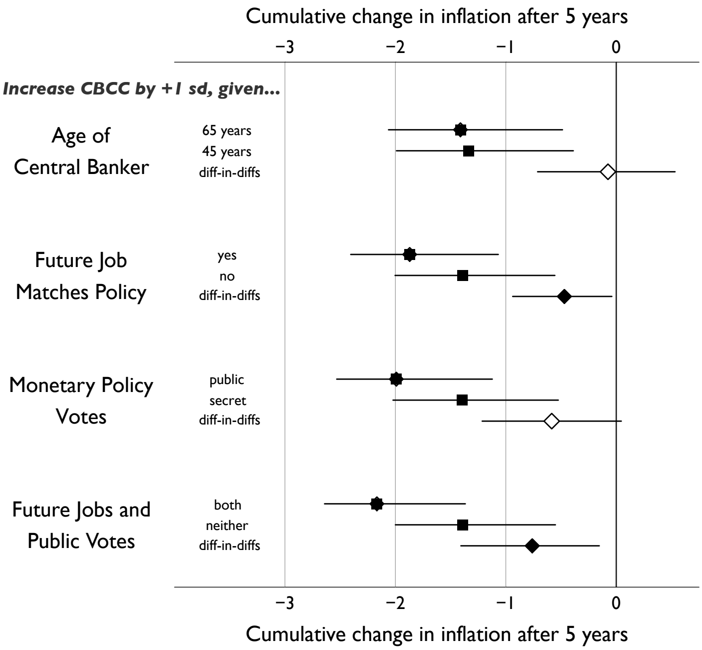
Remember, ropeladders are flexible – surely the *most* flexible way to present models

Be willing to experiment to make your model easier to explain

Recall our comparative inflation example

Central banker careers $\rightarrow$ Inflation performance

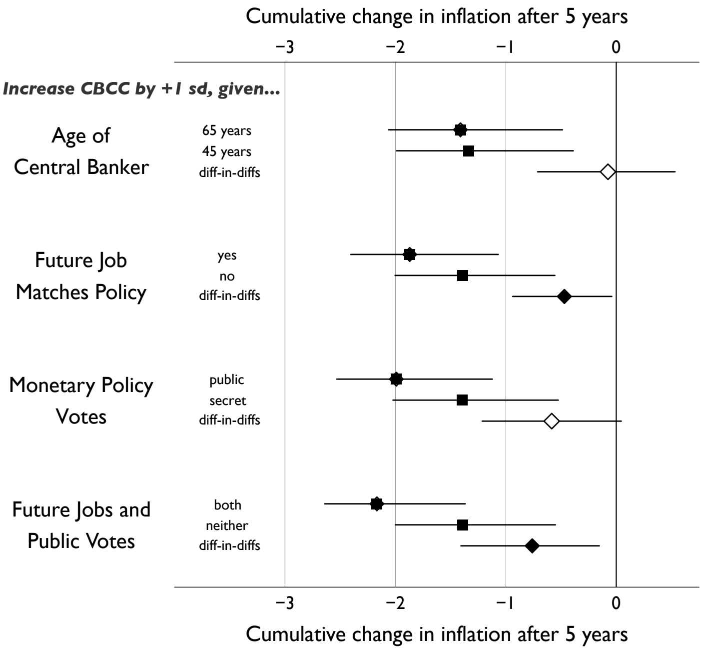
But is this a result of socialization or incentives?



Recall our comparative inflation example

Central banker careers → Inflation performance

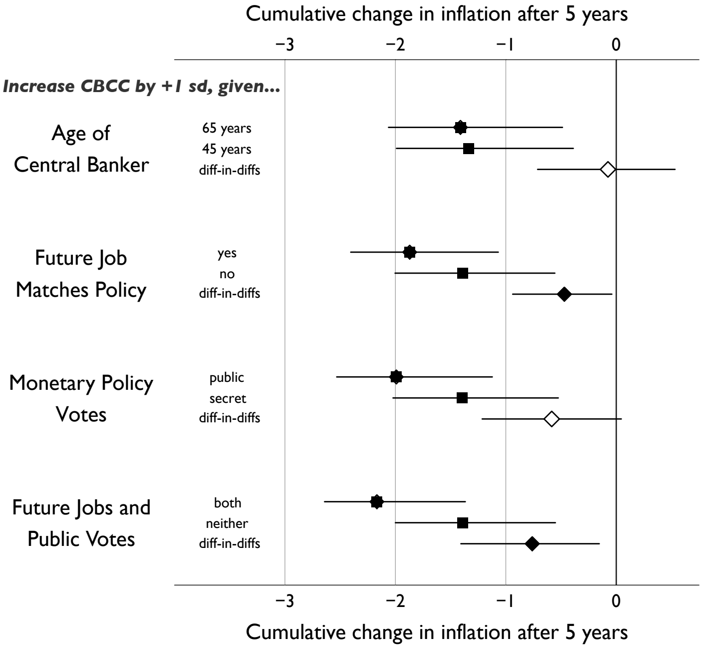
But is this a result of socialization or incentives?



Interactions may reveal the sources of career effects

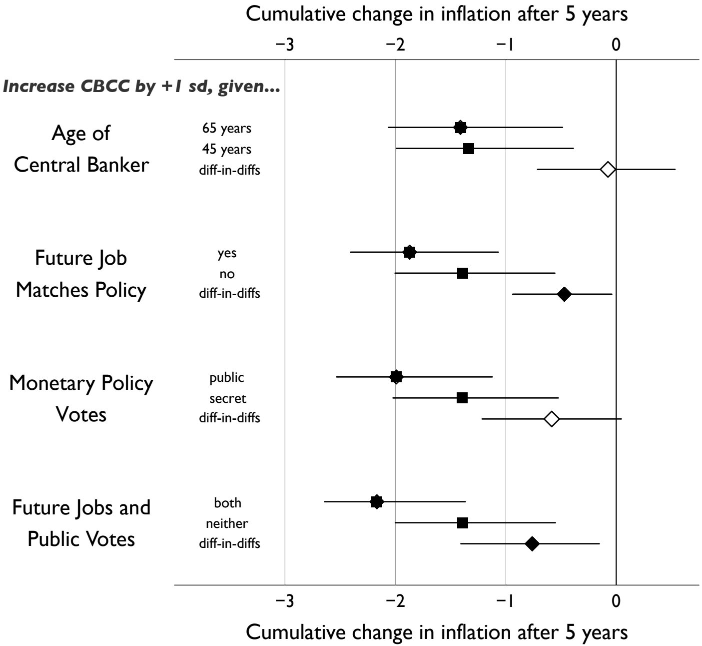
I interact "conservative" careers with:

- Age
- Future Job Awarded
- Public Voting



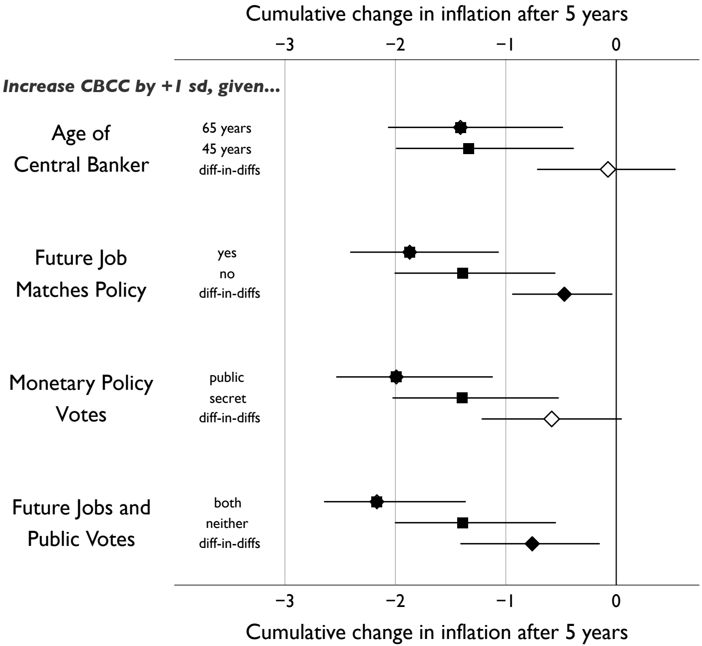
Focus on the "Age of Central Banker" trio of results

This model interacts Career Conservatism (CBCC) with central banker age



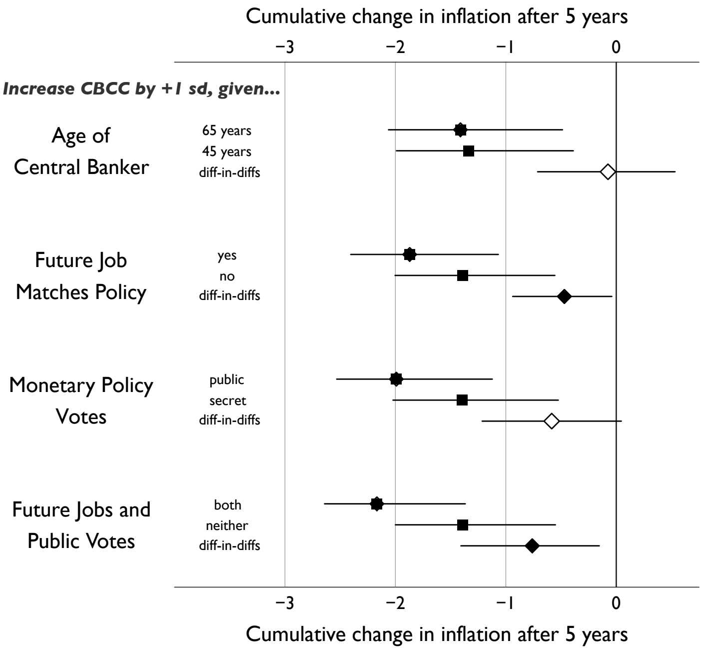
We simulate the effect of +1 sd CBCC given either 65 year old officials or 45 year old officials

We are especially interested in the difference of the first differences across these scenarios

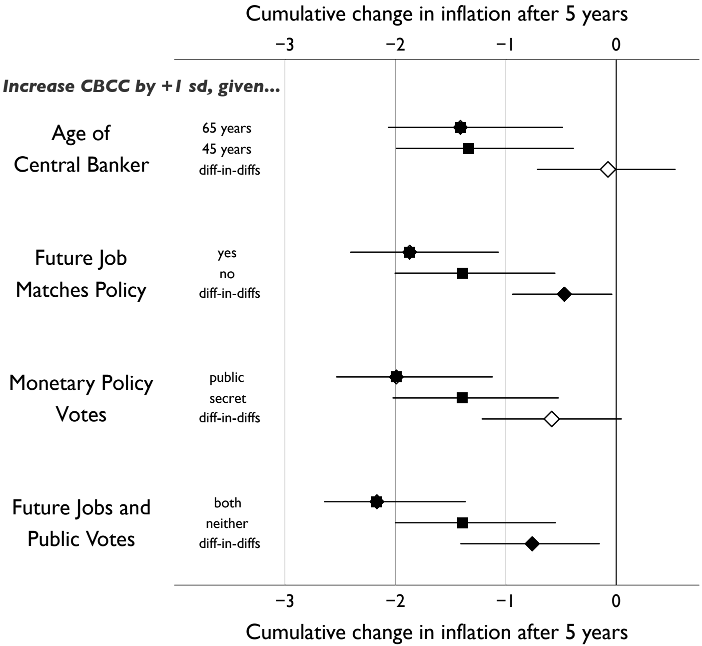


We use the shape of symbols to suggest the "building up" of the full effect for 65 year olds

While open vs. filled indicates significance

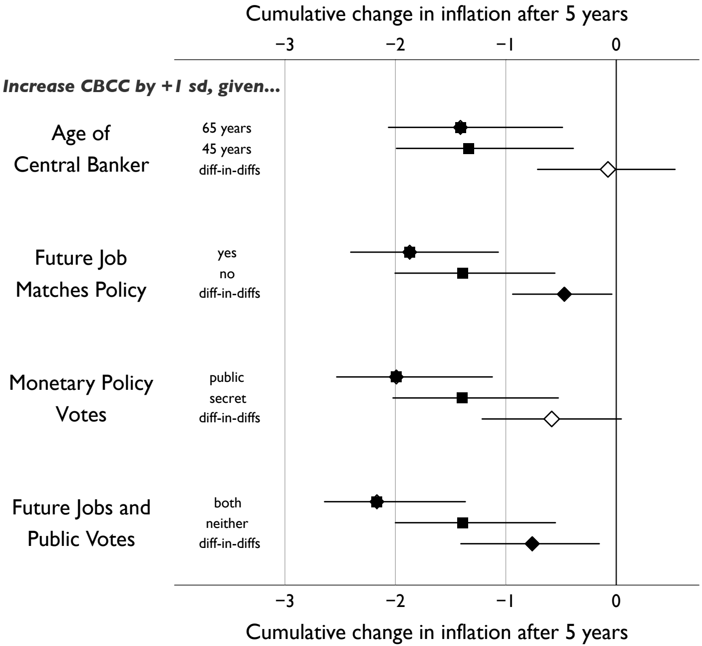


Looking at the whole plot, we find conservatism has bigger inflation-fighting effects when central bankers end up taking jobs in the financial sector

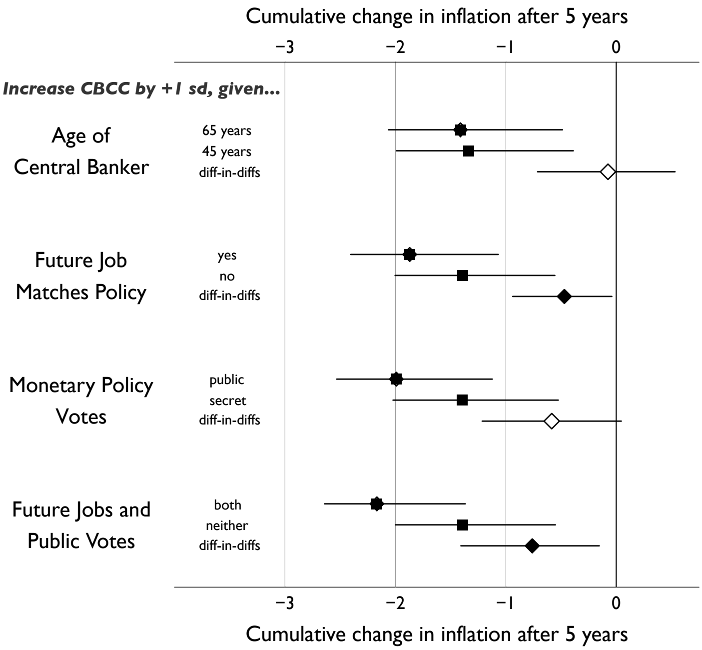


Conservatism may be stronger when votes are public, but this is not quite significant

Age is a wash



Ropeladders can explore interactive effects by working through each combination of values for the interacted covariates



2 keys to success:

Find the minimum number of combinations needed to tell the story

Clear organization – designed *around* the clearest write-up of the result

Chapter 9 of *BBC* explores correlates of central banker tenure in 20 industrialized countries using a Cox proportional hazards model

Covariate

Age
Career types
Economic performance
Change in government
Performance $\times$ Party

Last is most interesting: *are central bankers graded on a partisan curve, with the Left penalizing unemployment and the Right inflation?*

Chapter 9 of *BBC* explores correlates of central banker tenure in 20 industrialized countries using a Cox proportional hazards model

Covariate

- Age
- Career types
- Economic performance
- Change in government
- Performance × Party

Last is most interesting: *are central bankers graded on a partisan curve, with the Left penalizing unemployment and the Right inflation?*

Table 9.1. *Cox proportional hazards estimates of central banker tenure.*

Covariate	Hazard ratio	95% CI lower	95% CI upper
Age > 75	5.78	2.28	14.68
$70 <$ Age ≤ 75	3.48	2.32	5.22
$65 <$ Age ≤ 70	2.01	1.24	3.27
Other Government Experience	1.86	0.82	4.23
Abs diff in PCoG, appt party vs. current	1.67	1.24	2.25
Financial Experience	1.40	0.83	2.38
Finance Ministry Experience	1.34	0.71	2.52
Current PCoG × Inflation	1.05	1.00	1.11
Unemployment	1.04	1.00	1.08
Inflation	1.04	1.01	1.07
Current PCoG × Unemployment	0.95	0.89	1.02
Central Bank Staff Experience	0.90	0.62	1.30
Economics Experience	0.87	0.52	1.43
Current Partisan Center of Gravity (PCoG)	0.86	0.41	1.82
N	10,863	349 individuals	
log likelihood	−1229.4	LR test $p < 10^{-9}$	

Entries are hazard ratios (exponentiated coefficients) and their associated 95 percent confidence intervals. Hazard ratios greater than one indicate factors making retirement/dismissal *more* likely. Confidence intervals are calculated using standard errors clustered by country; significant results are those with lower and upper bounds on the same side of 1.00.

The table is actually fairly interpretable, *except*:

The career covariates are compositional, so their effects are blended

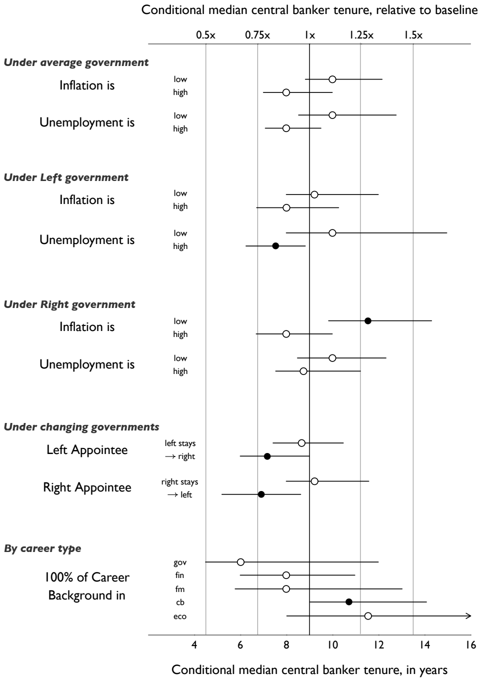
The interaction terms are hard to mentally combine, and it's impossible to get CIs without a computer to help

...so maybe this isn't *that* interpretable

Table 9.1. *Cox proportional hazards estimates of central banker tenure.*

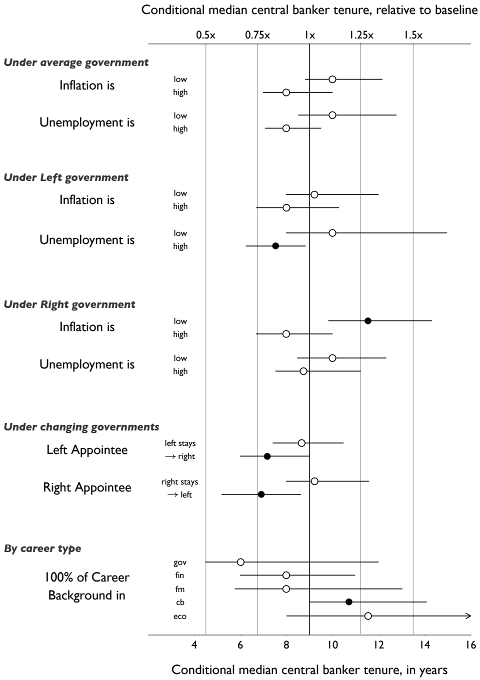
Covariate	Hazard ratio	95% CI lower	95% CI upper
Age > 75	5.78	2.28	14.68
$70 <$ Age ≤ 75	3.48	2.32	5.22
$65 <$ Age ≤ 70	2.01	1.24	3.27
Other Government Experience	1.86	0.82	4.23
Abs diff in PCoG, appt party vs. current	1.67	1.24	2.25
Financial Experience	1.40	0.83	2.38
Finance Ministry Experience	1.34	0.71	2.52
Current PCoG $\times$ Inflation	1.05	1.00	1.11
Unemployment	1.04	1.00	1.08
Inflation	1.04	1.01	1.07
Current PCoG $\times$ Unemployment	0.95	0.89	1.02
Central Bank Staff Experience	0.90	0.62	1.30
Economics Experience	0.87	0.52	1.43
Current Partisan Center of Gravity (PCoG)	0.86	0.41	1.82
N	10,863	349 individuals	
log likelihood	-1229.4	LR test $p < 10^{-9}$	

Entries are hazard ratios (exponentiated coefficients) and their associated 95 percent confidence intervals. Hazard ratios greater than one indicate factors making retirement/dismissal *more* likely. Confidence intervals are calculated using standard errors clustered by country; significant results are those with lower and upper bounds on the same side of 1.00.



We can replace the entire table with a complex dotplot

(Aside: It's okay to provide handouts of really large plots – they don't display on LCD projectors well)



Instead of thinking, "What covariates do I plot," ask:

"What is the minimum set of scenarios that will explore the full model space"

The key is picking out counterfactuals that explore effects of both inflation and unemployment under each type of government *and* under each possible change in government

Ropeladder Robustness Example: 1960 US crime rates

My apologies – this example isn't particularly substantively interesting or sharp

We have data from each of the 50 US states on crime rates in 1960

And a variety of covariates as seen on the next slide

We will fit a set of models with the same specification but different estimators

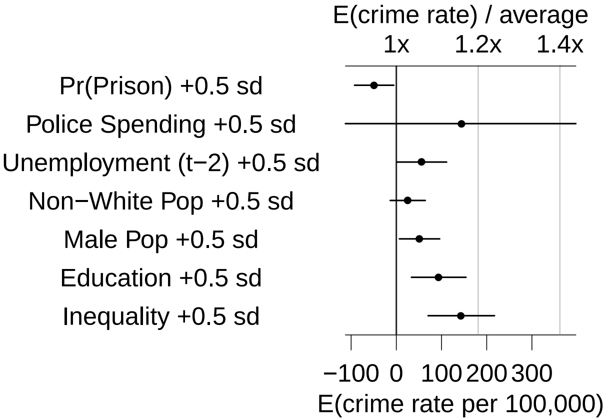
We will then consider several ropeladder-based presentations of robustness

Kitchen sink models of 1960 US crime rates

	Linear	Robust	Poisson	Neg Bin
Constant	-28820.91	-17784.56	-19.08	-15.43
	(10199.82)	(8158.71)	(1.77)	(7.81)
% males aged 14–24	1156.49	2480.55	1.1	1.53
	(522.98)	(418.32)	(0.1)	(0.4)
Southern state	0.97	138.11	0.06	0.06
	(141.49)	(113.18)	(0.02)	(0.11)
Mean education (yrs)	1802.64	1413.62	1.84	2.11
	(590.84)	(472.61)	(0.11)	(0.45)
Police spending 1960	897.54	422.45	0.81	0.76
	(813.8)	(650.95)	(0.15)	(0.62)
Police spending 1959	6.66	651.14	0.01	0.01
	(823.35)	(658.59)	(0.15)	(0.63)
Labor participation	143.91	2235.29	0.63	0.62
	(727.79)	(582.15)	(0.13)	(0.56)
Males per 1000	94.71	-3469.7	-1.46	-2.3
	(1943.8)	(1554.82)	(0.36)	(1.49)
State population	-79.39	-138.58	-0.08	-0.07

Kitchen sink models of 1960 US crime rates, continued

	continued			
	Linear	Robust	Poisson	Neg Bin
Nonwhites per 1000	61.25	32.47	0.11	0.11
	(47.85)	(38.28)	(0.01)	(0.04)
Unem, males 14–24	-325.65	-444.95	-0.18	-0.18
	(336.46)	(269.13)	(0.06)	(0.26)
Unem, males 35–39	475.14	895.28	0.39	0.46
	(239.62)	(191.67)	(0.04)	(0.18)
Gross state product, pc	282.31	-196.44	0.69	0.64
	(420.2)	(336.11)	(0.08)	(0.32)
Income inequality	1461.68	943.27	1.68	1.56
	(386.64)	(309.27)	(0.07)	(0.3)
Pr(imprisonment)	-226.39	-443.28	-0.29	-0.31
	(103.39)	(82.7)	(0.02)	(0.08)
E(time in prison)	-69.91	-294.41	-0.16	-0.27
	(184.13)	(147.29)	(0.03)	(0.14)

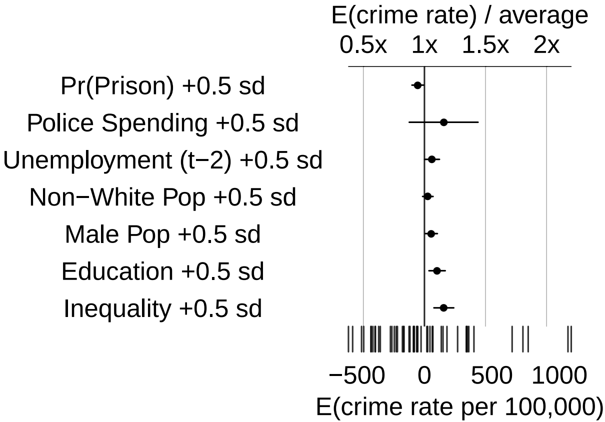


A simple inference dotplot with an extra axis showing relative risk

CODE EXAMPLE

`crimeRopeladders.r`

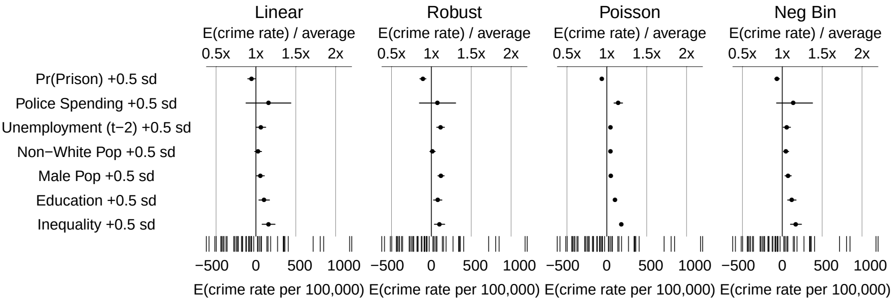
Ropeladder Robustness Example: US Crime



An inference dotplot with a marginal plot of the data

The data vary more widely than the first differences, stretching the plot

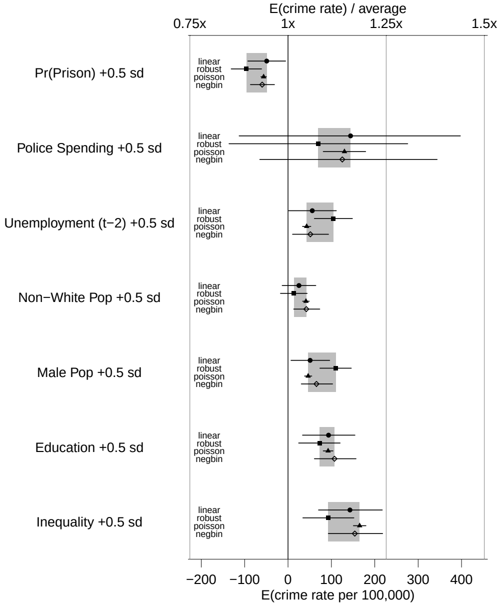
Ropeladder Robustness Example: US Crime



Side-by-side inference dotplots

The focus here is comparisons across covariates within models

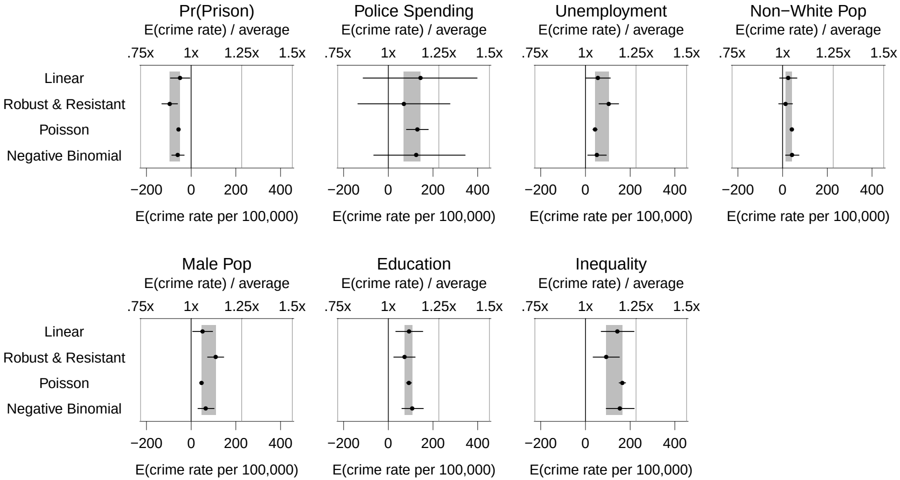
Ropeladder Robustness Example: US Crime



A superplot of ropeladders

Equal focus on comparisons across covariates and across models

Ropeladder Robustness Example: US Crime



Side-by-side robustness ropeladders –

focus is now on comparisons across models, not variables

How Do I Visualize Interactions of Covariates?

To effectively visualize interactive specifications, you need:

1. A strategy for constructing counterfactuals that survey the model space

2. An algorithm that assembles logically coherent counterfactuals
   and correctly computes QoIs and their CIs

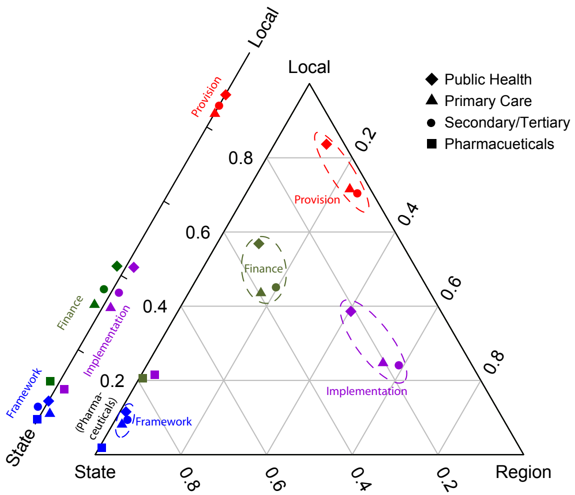
What you don't need: special machinery to calculate "marginal effects"

A generic counterfactual and simulation package that can use model formulas will correctly compute EVs, FDs, and RRs of the outcome variable

`simcf` does this – that's basically why it exists

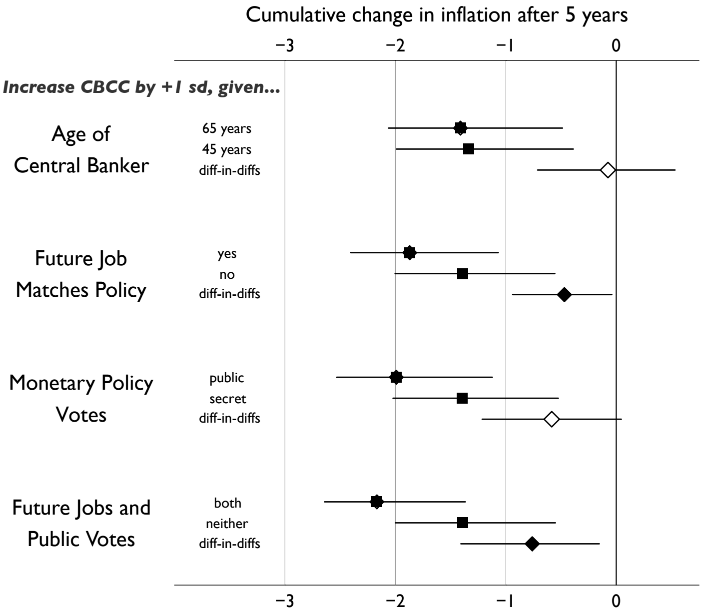
Strategies for Visualizing Interactive Covariates

Interaction	Counterfactual Strategy	Plot
discrete with discrete	one cf for each combination of values (full factorial)	`ropeladder`
continuous with discrete	choose combinations of representative values	`ropeladder`
	or	
	combine a continuum with each discrete value	`lineplot`
continuous with continuous	choose combinations of represented values	`ropeladder`
	or	
	combine a continuum with each discrete value	`lineplot`
	or	
	combine a continuum with a continuum	3D functional boxplots



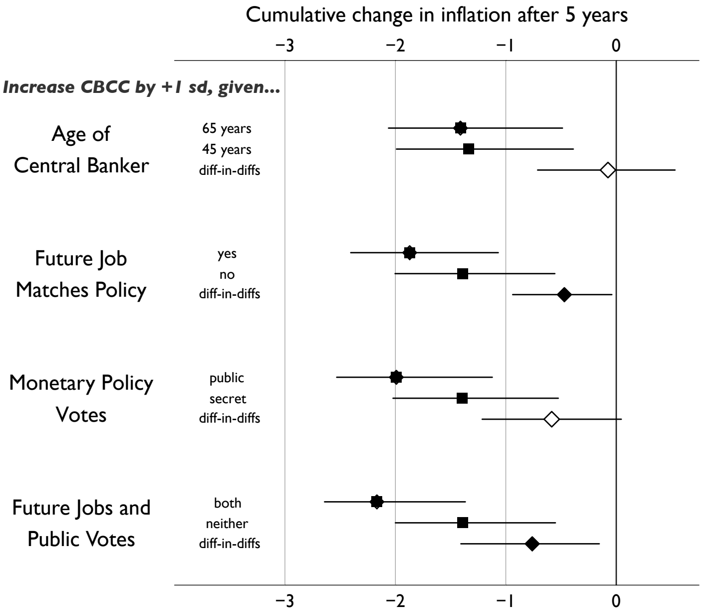
Let's convert this to a ropeladder ON THE WHITEBOARD

Continuous × Discrete Interactions: Ropeladders

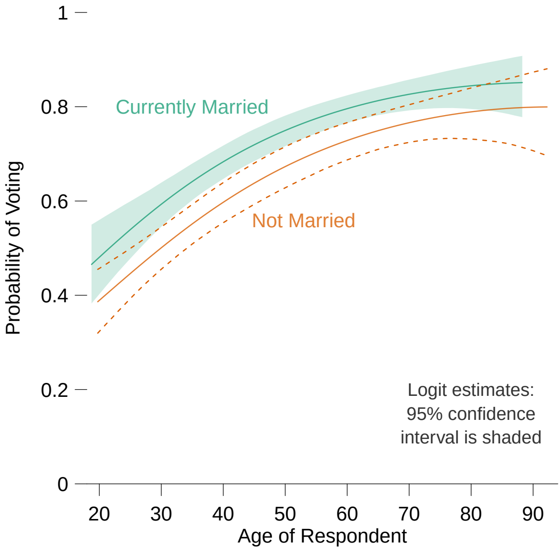


In this model of inflation, a continuous variable – central banker conservatism, CBCC – is interacted with binary variables like whether public votes are taken

Continuous × Discrete Interactions: Ropeladders



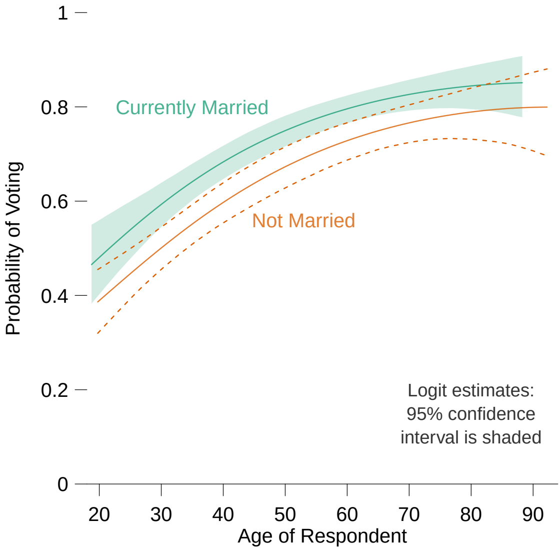
I selected a specific change in my continuous variable and computed counterfactuals under each value of my discrete variable



Let's return to our voting example, where voting was a function of Age, Age^2, Education, and Marriage Status

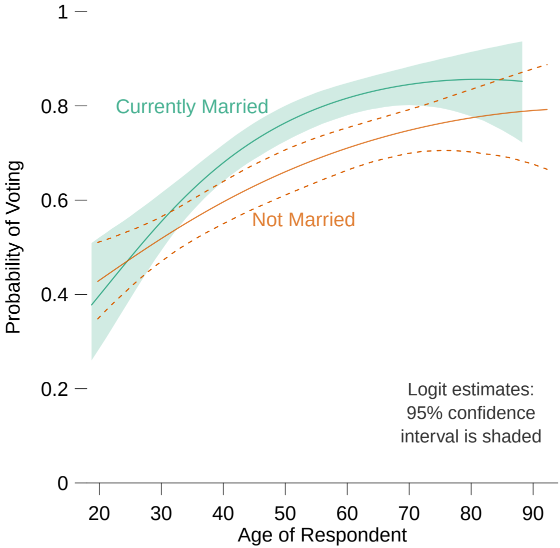
Suppose we add the interactions Age $\times$ Married and Age^2 $\times$ Married

Continuous × Discrete Interactions: Lineplots



Warning: I have no theoretical reason to do so, and the model fit suggests this is an overspecified model. We just want an example of *how* to do this

Continuous × Discrete Interactions: Lineplots



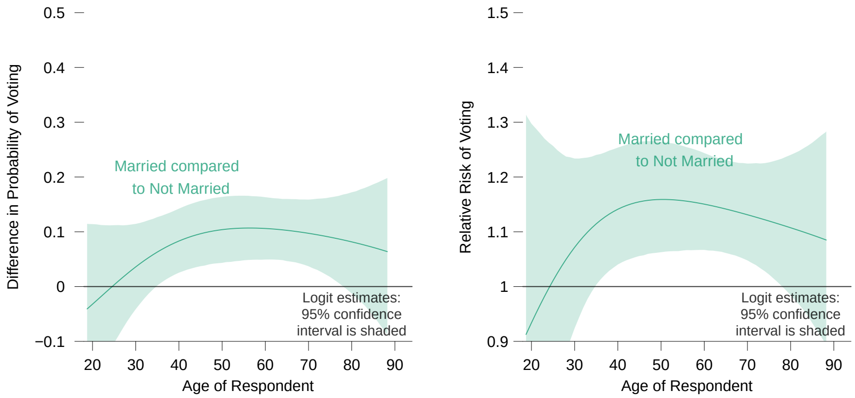
To get the new plots from `votingLineplots.R`, I need to change only *one line*:

```
model2 <-
vote00 ~ age + I(age^2)
 + hsdeg + coldeg
 + marriedo
```

changes to

```
model2 <-
vote00 ~ age*marriedo
 + I(age^2)*marriedo
 + hsdeg + coldeg
```

Continuous × Discrete Interactions: Lineplots

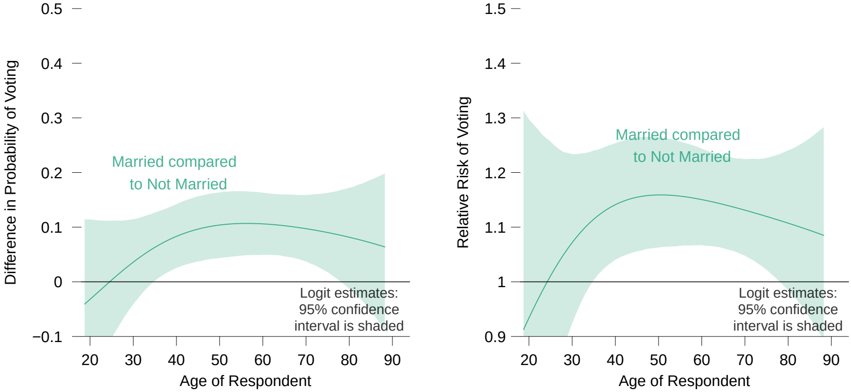


`simcf` **takes care of the rest –**

it will correctly set up interactions and combines their uncertainty into the QoIs

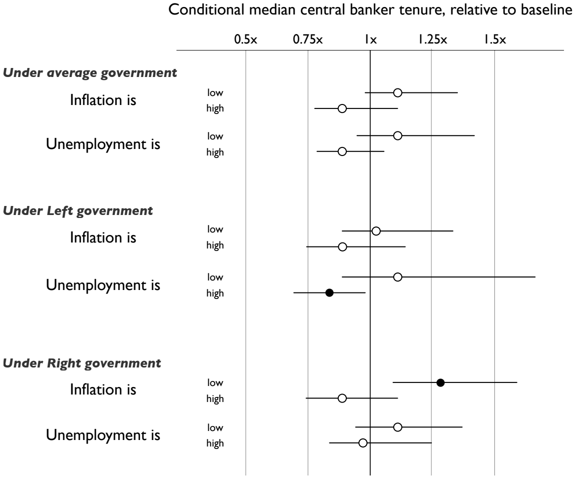
The first differences and relative risks above *are* your marginal effect plots

Continuous × Discrete Interactions: Lineplots



The results suggest this interaction wasn't a great idea...

Continuous × Continuous Interactions: Ropeladders



We've already dealt with a continuous × continuous interaction:

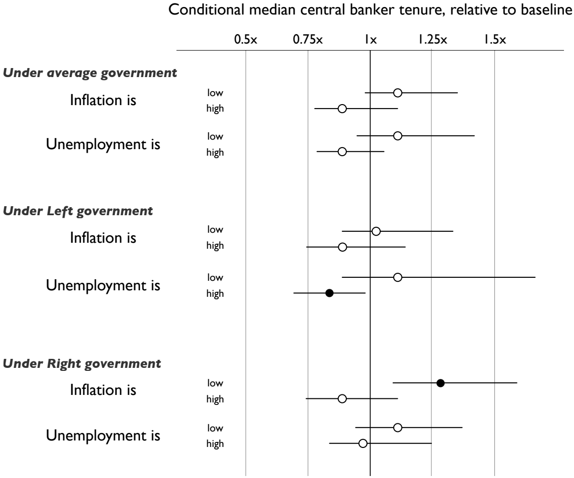
Central banker tenure depended on:

Inflation × Party CoG

and

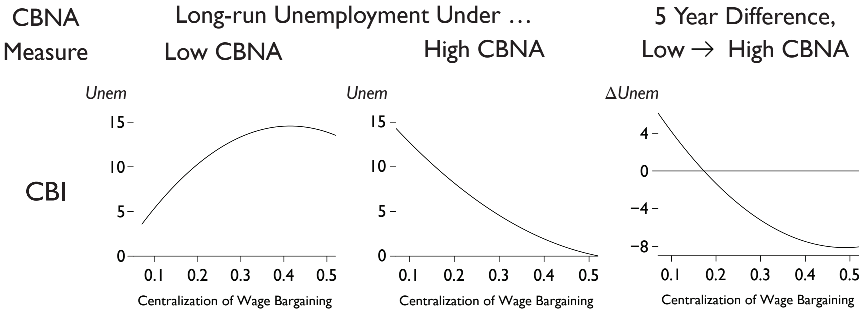
Unemployment × Party CoG

Continuous × Continuous Interactions: Ropeladders



I simply examined every combination of high, low, and average partisanship with high and low inflation or unemployment

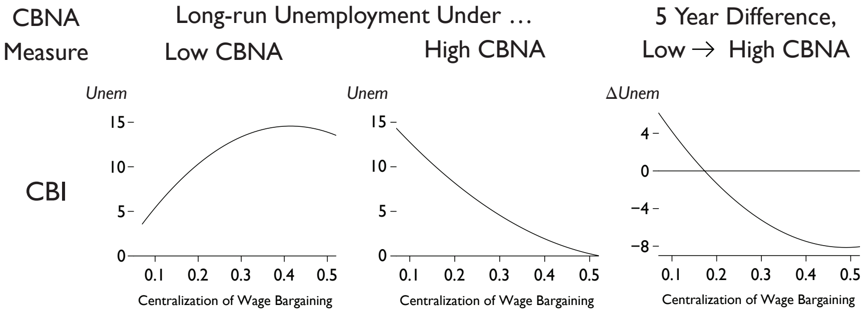
Grouping and labeling the dotplot helps catalog the combinations



In Chapter 6 of *Bankers, Bureaucrats, and Central Bank Politics*, I consider the interactive effects of central bank "nonaccommodation" (autonomous conservatism) and wage bargaining centralization on unemployment

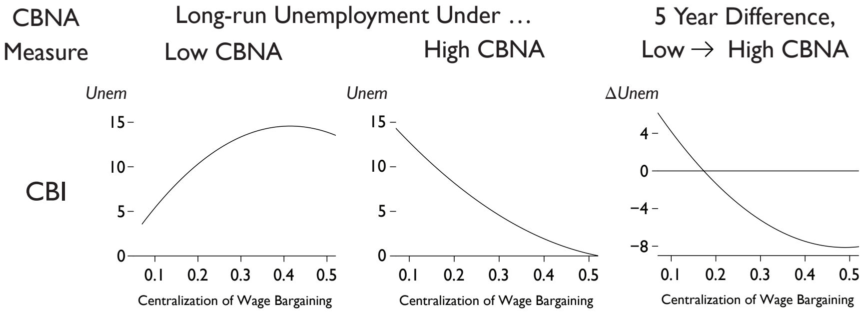
I build on and test a complex literature positing interactive, nonlinear effects

Continuous × Continuous Interactions: Lineplots



I investigate how different measures of nonaccommodation affect the results

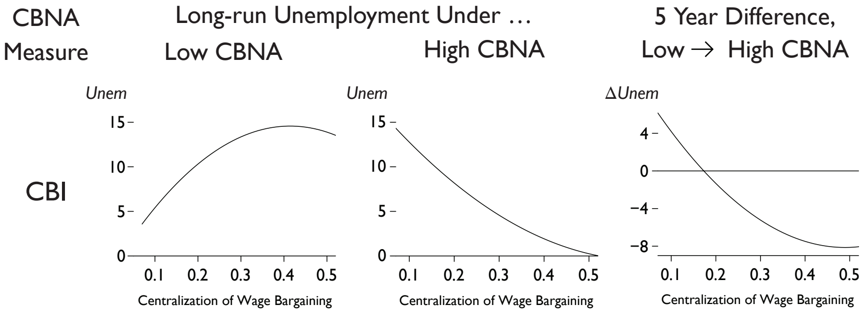
I start with a crude "independence only" measure



The left and middle show expected unemployment across the continuum of CWB for two different levels of CBNA

The right plot shows the first difference in unemployment given a change in CBNA at each level of CWB

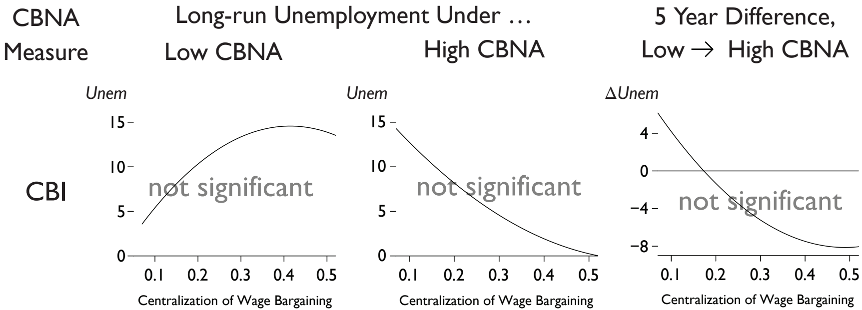
Continuous × Continuous Interactions: Lineplots



This is an intuitive measure of the wage-bargaining-conditional effect of nonaccommodation

`simcf` can produce this, with the right syntax

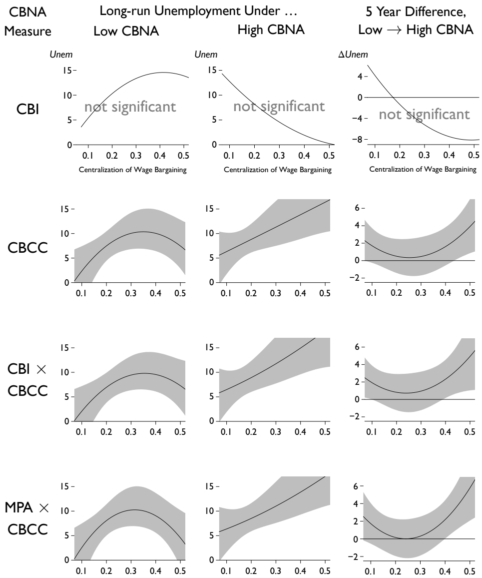
Note we've also iterated over time, so you would use `ldvsimfd()`



Why no CIs? Because they would fill the whole plot!

I could make the plot area bigger, but that would make comparison hard

Continuous × Continuous Interactions: Lineplots

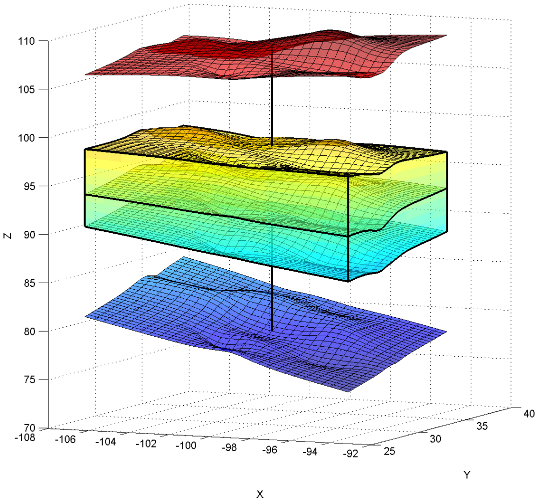


The real goal here is a robustness exercise

Measures of CBNA incorporating career conservatism produce similar and generally more precise results, alone or in combination with different measures of autonomy

Yet another approach to showing robustness – one that emphasizes similarity of fits *and* CIs for conditional relationships

Continuous × Continuous Interactions: 3D Boxplots?



I was long a skeptic of including confidence volumes in 3D plots

This example made me a believer

If it is *really* important to see smooth variation in 2 interacting continuous covariates at the same time, investigate *functional boxplots*

Source: Ying Sun and Marc G. Genton. 2011. "Functional boxplots." *JCGS* 20:2)

Strategies for Interactions of 3 Continuous Variables

Any set of interactions involving 2 or fewer continuous variables can be addressed with the above methods

What if you have 3 continuous variables? Some strategies:

Strategy 1 – Ropeladder

	x_1	x_2	x_3
	high	high	high
	high	high	low
values	high	low	high
used	high	low	low
in	low	high	high
counterfactuals	low	high	low
	low	low	high
	low	low	low

Descriptive names for these combinations essential for presentation in this case

Strategies for Interactions of 3 Continuous Variables

Any set of interactions involving 2 or fewer continuous variables can be addressed with the above methods

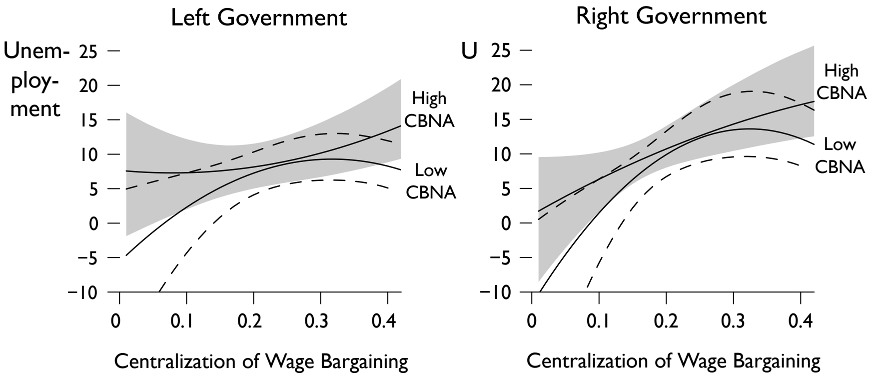
What if you have 3 continuous variables? Some strategies:

Strategy 2 – Lineplots (overlapping and/or side-by-side)

	X_1	X_2	X_3
values		high	high
used	contin-	high	low
in	uum	low	high
counterfactuals		low	low

Use tiling of lineplots to your advantage in this case

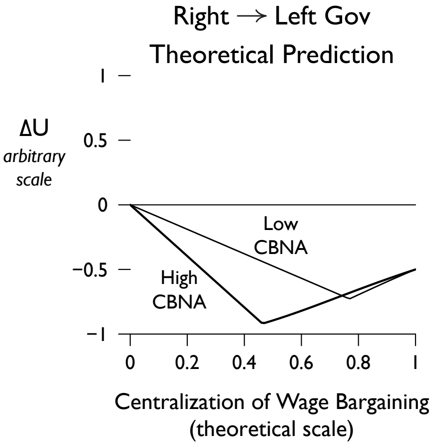
Strategies for Interactions of 3 Continuous Variables

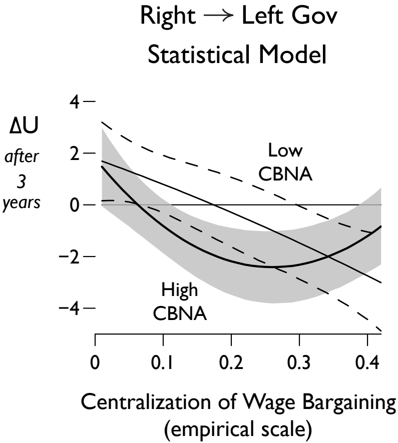


In Chapter 7 of *BBC*, I add a third interactive term to wage bargaining centralization and central bank nonaccommodation: partisanship of government

The theory and model is complex, but graphically, I just plot 4 traces instead of 2

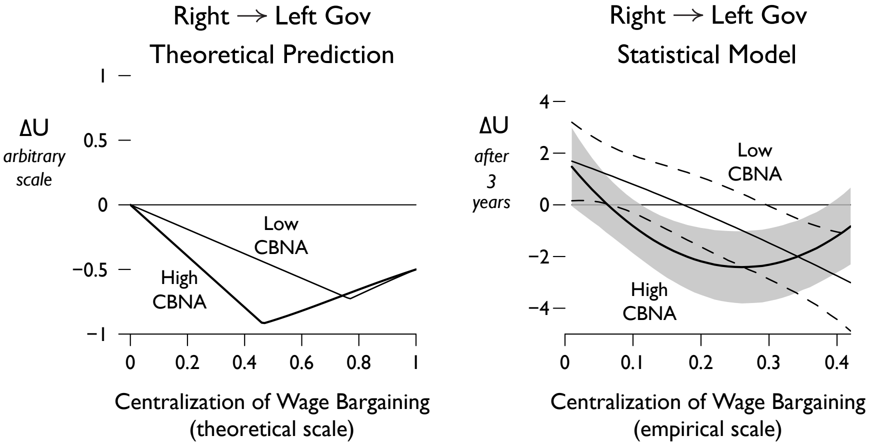
Strategies for Interactions of 3 Continuous Variables





If we plot first differences across partisanship, we're back to a 2 trace plot, but with a separate continuum of first difference for each level of CBNA

Strategies for Interactions of 3 Continuous Variables



For complex models it helps to have a theory and to show it in the same format
– both to justify and to explain the empirical result

Strategies for Interactions of 3 Continuous Variables

Any set of interactions involving 2 or fewer continuous variables can be addressed with the above methods

What if you have 3 continuous variables? Some strategies:

Strategy 3 – Side-by-side 2D boxplots?

	X_1	X_2	X_3
values used in counterfactuals	contin-uum	contin-uum	high low

I've never tackled this problem – but this is the strategy I'd use

Lessons for practice of data analysis

Simulation + Graphics can summarize complex models for a broad audience

You might even find something you missed as an analyst

And even for fancy or complex models, we can and should show uncertainty

Payoff to programming: this is hard the first few times, but gets easier, especially with `simcf`+`tile`

Your code is re-usable and encourages more ambitious modeling