Markdown and universal writing

Writing up research is a complicated, messy process!

Itty bitty pieces

Data

Statistical results

Fieldwork

Interviews

Analysis

Figures

Images

Tables

Citations

Your actual words

Each of these comes from a different place!

Two general approaches for this mess

The Office model

Put everything in one document

The Engineering model

Embrace the bittiness and compile it all at the end

The Office model

Everything lives in one `.docx` file

Drag images in

Copy/paste stats from R or Stata

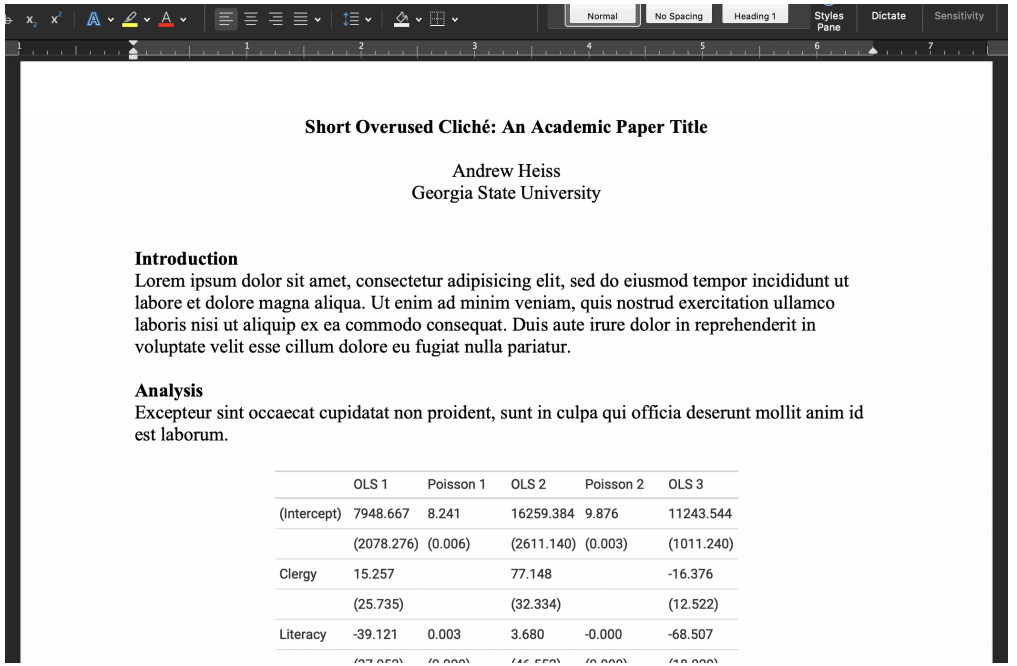
Connect Word to Zotero or Endnote

Track versions with filenames:

`ms.docx`, `ms2_final.docx`,

`ms2_final_final.docx`

Final output = `.docx` file



Short Overused Cliché: An Academic Paper Title

Andrew Heiss
Georgia State University

Introduction
Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur.

Analysis
Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

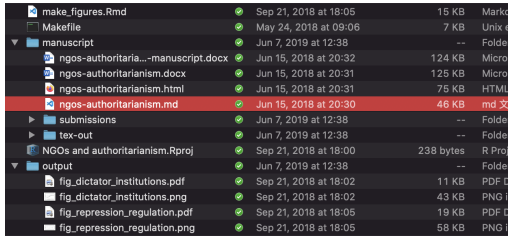
	OLS 1	Poisson 1	OLS 2	Poisson 2	OLS 3
(Intercept)	7948.667	8.241	16259.384	9.876	11243.544
	(2078.276)	(0.006)	(2611.140)	(0.003)	(1011.240)
Clergy	15.257		77.148		-16.376
	(25.735)		(32.334)		(12.522)
Literacy	-39.121	0.003	3.680	-0.000	-68.507

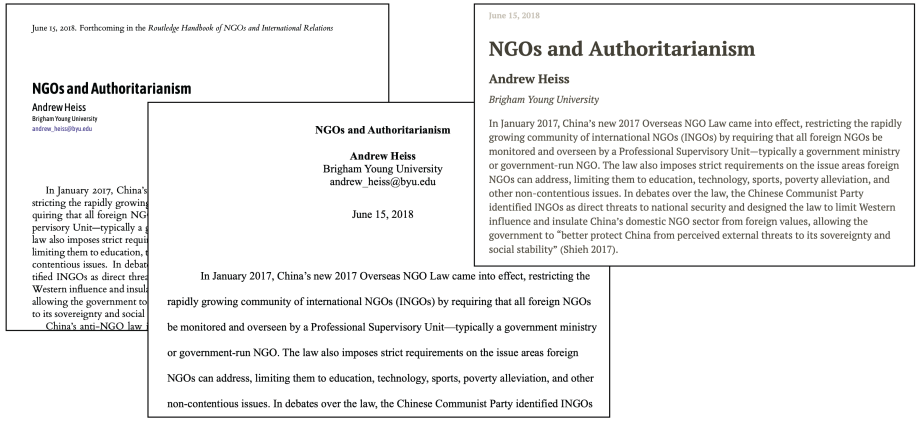
The Engineering model

Everything lives separately and is combined in the end

- Type text in a plain text document
- Import images automatically
- Import stats automatically from R scripts (`.R` or `.qmd`) or `.do` files
- Store citations in reference manager
- Track versions with git

Final output = whatever you want (Word, PDF, HTML)





No one right way!

The Office model can be clunky and you'll inevitably forget to update figures, tables, results, etc.

BUT

The whole world runs on Word

Even if you're a strict Engineering person, you'll still collaborate with Office people!

Coauthors will work in Word, advisers will give comments and track changes in Word, journals will demand final Word files

Reproducibility

The Engineering model is definitely more fiddly

BUT

There's less cognitive load!

No need to copy/paste new results,
add updated figures, reformat citation, etc.

There's a record of everything you do!

Your findings are reproducible by anyone (and yourself!)

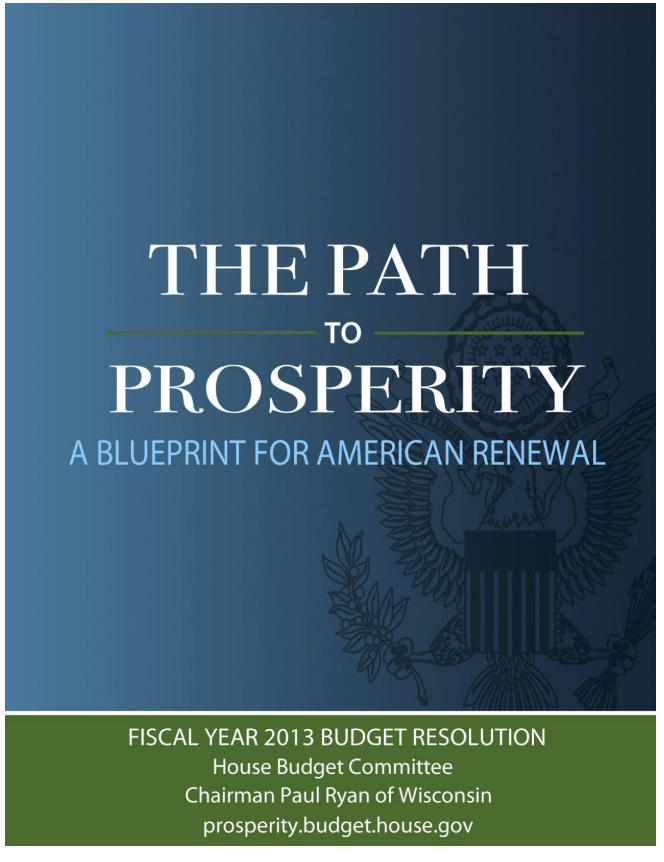
Austerity and Excel

Growth in a Time of Debt
Carmen M. Reinhart and Kenneth S. Rogoff
NBER Working Paper No. 15639
January 2010, Revised January 2010
JEL No. E2,E3,E6,F3,F4,N10

ABSTRACT

We study economic growth and inflation at different levels of government and external debt. Our analysis is based on new data on forty-four countries spanning about two hundred years. The dataset incorporates over 3,700 annual observations covering a wide range of political systems, institutions, exchange rate arrangements, and historic circumstances. Our main findings are: First, the relationship between government debt and real GDP growth is weak for debt/GDP ratios below a threshold of 90 percent of GDP. Above 90 percent, median growth rates fall by one percent, and average growth falls considerably more. We find that the threshold for public debt is similar in advanced and emerging economies. Second, emerging markets face lower thresholds for external debt (public and private)—which is usually denominated in a foreign currency. When external debt reaches 60 percent of GDP, annual growth declines by about two percent; for higher levels, growth rates are roughly cut in half. Third, there is no apparent contemporaneous link between inflation and public debt levels for the advanced countries as a group (some countries, such as the United States, have experienced higher inflation when debt/GDP is high). The story is entirely different for emerging markets, where inflation rises sharply as debt increases.

Debt:GDP ratio
90%+ → −0.1% growth



Paul Ryan's 2013 House budget resolution

Austerity and Excel

Thomas Herndon

Over time, another problem emerged: Other researchers, using seemingly comparable data on debt and growth, couldn't replicate the Reinhart-Rogoff results. They typically found some correlation between high debt and slow growth — but nothing that looked like a tipping point at 90 percent or, indeed, any particular level of debt.

Finally, Ms. Reinhart and Mr. Rogoff allowed researchers at the University of Massachusetts to look at their original spreadsheet — and the mystery of the irreproducible results was solved. First, they omitted some data; second, they used unusual and highly questionable statistical procedures; and finally, yes, they made an Excel coding error. Correct these oddities and errors, and you get what other researchers have found: some correlation between high debt and slow growth, with no indication of which is causing which, but no sign at all of that 90 percent "threshold."

From Paul Krugman, "The Excel Depression"

Austerity and Excel

Table 1. Real GDP Growth as the Level of Government Debt Varies:
Selected Advanced Economies, 1790-2009
(annual percent change)

Country	Period	Central (Federal) government debt/ GDP			
		Below 30 percent	30 to 60 percent	60 to 90 percent	90 percent and above
Australia	1902-2009	3.1	4.1	2.3	4.6
Austria	1880-2009	4.3	3.0	2.3	n.a.
Belgium	1835-2009	3.0	2.6	2.1	3.3
Canada	1925-2009	2.0	4.5	3.0	2.2
Denmark	1880-2009	3.1	1.7	2.4	n.a.
Finland	1913-2009	3.2	3.0	4.3	1.9
France	1880-2009	4.9	2.7	2.8	2.3
Germany	1880-2009	3.6	0.9	n.a.	n.a.
Greece	1884-2009	4.0	***0.3***	***4.8***	2.5
Ireland	1949-2009	4.4	4.5	4.0	2.4
Italy	1880-2009	***5.4***	***4.9***	1.9	0.7
Japan	1885-2009	4.9	3.7	3.9	0.7
Netherlands	1880-2009	4.0	2.8	2.4	2.0
New Zealand	1932-2009	2.5	2.9	3.9	**3.6**
Norway	1880-2009	2.9	4.4	n.a.	n.a.
Portugal	1851-2009	4.8	2.5	1.4	n.a.
Spain	1850-2009	***1.6***	3.3	***1.3***	2.2
Sweden	1880-2009	2.9	2.9	2.7	n.a.
United Kingdom	1830-2009	2.5	2.2	2.1	1.8
United States	1790-2009	4.0	3.4	3.3	***-1.8***
Average		**3.7**	**3.0**	**3.4**	**1.7**
Median		**3.9**	**3.1**	**2.8**	**1.9**
Number of observations = **2,317**		866	654	445	352

Debt:GDP ratio = 90%+ → 2.2% growth (!!)

Genes and Excel

Septin 2

Membrane-Associated Ring Finger (C3HC4) 1

2310009E13

	A	B
1	**Actual value**	**What Excel turns it into**
2	SEPT2	2-Sep
3	MARCH1	1-Mar
4	2310009E13	2.31E+19

20% of genetics papers between 2005–2015 (!!!)

General guidelines

Don't touch the raw data

If you do, explain what you did!

Use self-documenting, reproducible code

The whole point of this Engineering model! (Quarto!)

Use open formats

Use .csv, not .xlsx

Engineering model in real life

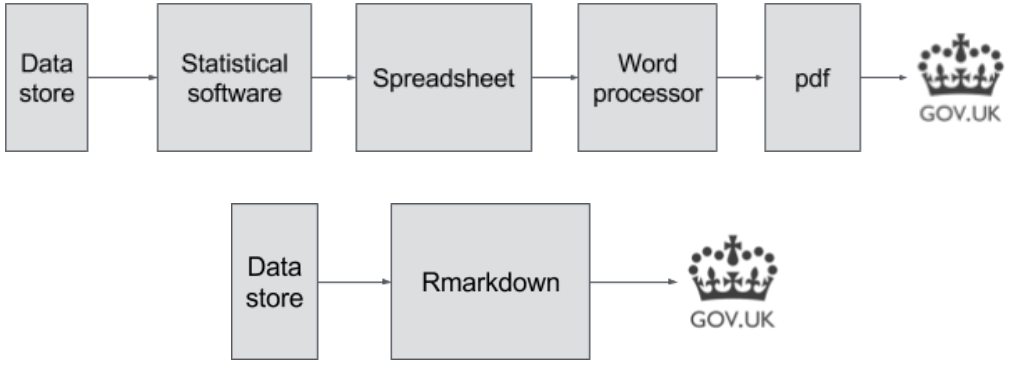
3.1.2 Data Visualization

We use ggplot2 as our main package to create ad-hoc exploratory graphics as well as polished-looking customized visualizations. When combined with tools to clean and transform data, ggplot2 allows analysts to quickly translate insights into high quality, compelling visualizations. In addition to the static graphics of `ggplot2`, we often make interactive visualizations or dashboards using R packages such as `plotly` (Sievert et al. 2017), `leaflet` (Cheng et al. 2017), `dygraphs` (Vanderkam et al. 2017), `DiagrammeR` (Sveidqvist et al. 2017), and `shiny` (Chang et al. 2017).

3.1.3 Reproducible Research

At Airbnb, all R analyses are documented in `rmarkdown`, where code and visualizations are combined within a single written report. Posts are carefully reviewed by experts in the content area and techniques used, both in terms of methodologies and code style, before publishing and sharing with the business partners. The peer review process is

Airbnb, ggplot, and rmarkdown



The UK's reproducible analysis pipeline

Universal writing

$\LaTeX$

You might know LaTeX—it's a scientific typesetting language

Great! It's a way to do the Engineering model

Run `latexmk` from the terminal or click the "compile" button in your TeX editor and you'll stitch together all your separate writing, tables, images, and citations

HOWEVER, the world runs on `.docx`

There are ways to convert from `.tex` to `.docx`, but they're a pain

Three syntaxes

If you want to be able to write in LaTeX, but also have HTML for a blog post later, and also have a Word file for a journal later, you have to learn (and write with!) all these!

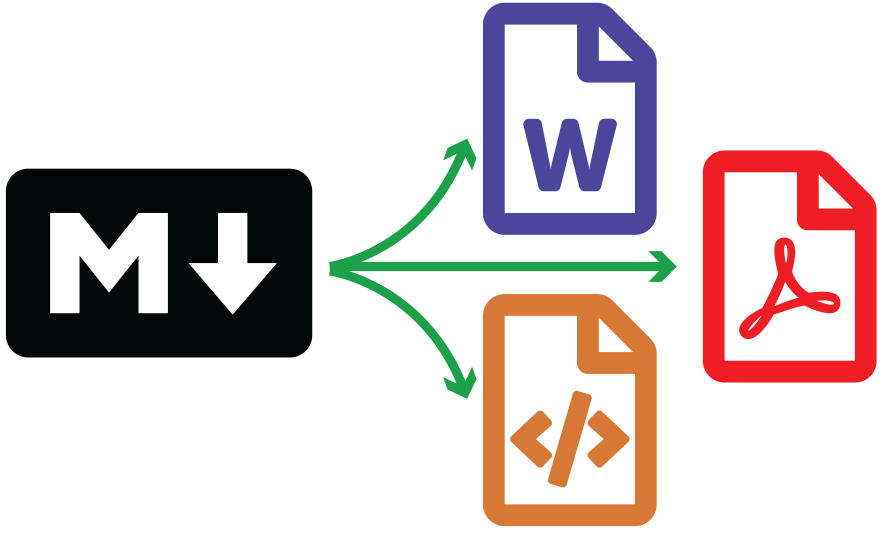
Format	LaTeX	HTML	Word
Bold	\textbf{Something}	Something	Click on stuff
Heading 2	\subsection{Something}	Something	Click on stuff
Link	\href{google.com}{Link}	Link	Click on stuff
Citation	\cite{Heiss2020}	lolz	lolz
Math	y = \beta_0 + \beta_1 x_1	lolz	Equation editor

Markdown + Quarto

Solution: Use a universal syntax

Write with one simplified syntax and convert from that to whatever output you want.

Format	Markdown
Bold	**Something**
Heading 2	## Something
Link text	[Link](google.com)
Citation	@Heiss2020
Math	y = \beta_0 + \beta_1 x_1



Quarto

The magic glue that makes this all work

```
# To HTML
quarto render manuscript.qmd --to html

# To Word
quarto render manuscript.qmd --to docx

# To PDF (through LaTeX)
quarto render manuscript.qmd --to pdf
```

(Or just click on stuff in RStudio)

Let's play with Markdown and Quarto!