'revisit': an R Package for Taming the Reproducibility Problem

Norm Matloff
Dept. of Computer Science
University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Stanford R Group, 7 November, 2017

These slides will be available at
`http://heather.cs.ucdavis.edu/StanfordR.pdf`

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

If You Are Curious

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

If You Are Curious

Why is a CS professor interested in this?

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

If You Are Curious

Why is a CS professor interested in this?

- PhD in pure math, abstract probability theory.
- Joined UCD, working on statistical methodology.
- Was one of the founders of the UCD Stat Dept.
- Moved to CS Dept. long ago, but "Once a statistician, always a statistician."

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Reinhart and Rogoff Case

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Reinhart and Rogoff Case

Reinhart and Rogoff (2010) found that high government debt $\Longrightarrow$ weaker/negative economic growth.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Reinhart and Rogoff Case

Reinhart and Rogoff (2010) found that high government debt $\Longrightarrow$ weaker/negative economic growth.

The finding was strongly influential.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Reinhart and Rogoff Case

Reinhart and Rogoff (2010) found that high government debt $\Longrightarrow$ weaker/negative economic growth.
The finding was strongly influential.

- Much cited by deficit hawks in Congress.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Reinhart and Rogoff Case

Reinhart and Rogoff (2010) found that high government debt $\Longrightarrow$ weaker/negative economic growth.

The finding was strongly influential.

- Much cited by deficit hawks in Congress.
- Often reported by the *Washington Post* as "consensus among economists."

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Reinhart and Rogoff Case

Reinhart and Rogoff (2010) found that high government debt $\Longrightarrow$ weaker/negative economic growth.
The finding was strongly influential.

- Much cited by deficit hawks in Congress.
- Often reported by the *Washington Post* as "consensus among economists."

And yet:

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Reinhart and Rogoff Case

Reinhart and Rogoff (2010) found that high government debt $\Longrightarrow$ weaker/negative economic growth.
The finding was strongly influential.

- Much cited by deficit hawks in Congress.
- Often reported by the *Washington Post* as "consensus among economists."

And yet:

- Analysis had spreadsheet errors, data available but missing in the spreadsheet.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Reinhart and Rogoff Case

Reinhart and Rogoff (2010) found that high government debt $\Longrightarrow$ weaker/negative economic growth.
The finding was strongly influential.

- Much cited by deficit hawks in Congress.
- Often reported by the *Washington Post* as "consensus among economists."

And yet:

- Analysis had spreadsheet errors, data available but missing in the spreadsheet.
- Dubious weighting used (some say).

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Reinhart and Rogoff Case

Reinhart and Rogoff (2010) found that high government debt $\Longrightarrow$ weaker/negative economic growth.
The finding was strongly influential.

- Much cited by deficit hawks in Congress.
- Often reported by the *Washington Post* as "consensus among economists."

And yet:

- Analysis had spreadsheet errors, data available but missing in the spreadsheet.
- Dubious weighting used (some say).
- Should have used a longer time window (some say).

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Reinhart and Rogoff Case

Reinhart and Rogoff (2010) found that high government debt $\Longrightarrow$ weaker/negative economic growth.
The finding was strongly influential.

- Much cited by deficit hawks in Congress.
- Often reported by the *Washington Post* as "consensus among economists."

And yet:

- Analysis had spreadsheet errors, data available but missing in the spreadsheet.
- Dubious weighting used (some say).
- Should have used a longer time window (some say).
- After correction, -0.1% growth becomes +2.2%

Eichengreen Comment

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Eichengreen Comment

The brouhaha over Carmen Reinhart and Kenneth Rogoffs article "Growth in a Time of Debt" has raised troubling questions not only about the efficacy of [fiscal] austerity, but also about the reliability of economic analysis. If a flawed study could appear in a prestigious working-paper series, why should anyone trust economic research?

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Who's Right?

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Who's Right?

- Reinhart and Rogoff defended their basic findings.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Who's Right?

- Reinhart and Rogoff defended their basic findings.
- Lots of controversy back and forth.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Who's Right?

- Reinhart and Rogoff defended their basic findings.
- Lots of controversy back and forth.
- But the incident shows this:

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Who's Right?

- Reinhart and Rogoff defended their basic findings.
- Lots of controversy back and forth.
- But the incident shows this:
  - Research should be **transparent**.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Who's Right?

- Reinhart and Rogoff defended their basic findings.
- Lots of controversy back and forth.
- But the incident shows this:
  - Research should be **transparent**.
  - We need to facilitate a **healthy skepticism**, by facilitating **the asking of "What if" questions**.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Who's Right?

- Reinhart and Rogoff defended their basic findings.
- Lots of controversy back and forth.
- But the incident shows this:
  - Research should be **transparent**.
  - We need to facilitate a **healthy skepticism**, by facilitating **the asking of "What if" questions**.
  - In this case, e.g. "What if a different time frame had been used?" "What if a different weighting had been used?"

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 1

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 1

- An R-language package, usable in GUI and text forms.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 1

- An R-language package, usable in GUI and text forms. (Text form does have some advantages.)

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 1

- An R-language package, usable in GUI and text forms. (Text form does have some advantages.)
- A "statistical audit."

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 1

- An R-language package, usable in GUI and text forms. (Text form does have some advantages.)
- A "statistical audit."
- Enable other scientists to "revisit" all numerical aspects of a scientific investigation:

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 1

- An R-language package, usable in GUI and text forms. (Text form does have some advantages.)
- A "statistical audit."
- Enable other scientists to "revisit" all numerical aspects of a scientific investigation:
    - Data wrangling — dealing with outliers, missing values etc.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 1

- An R-language package, usable in GUI and text forms. (Text form does have some advantages.)
- A "statistical audit."
- Enable other scientists to "revisit" all numerical aspects of a scientific investigation:
  - Data wrangling — dealing with outliers, missing values etc.
  - Statistical analyses — assumptions, methods, variables, time frames, etc.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 1

- An R-language package, usable in GUI and text forms. (Text form does have some advantages.)
- A "statistical audit."
- Enable other scientists to "revisit" all numerical aspects of a scientific investigation:
  - Data wrangling — dealing with outliers, missing values etc.
  - Statistical analyses — assumptions, methods, variables, time frames, etc.
  - Allows users to ask the "What if?" questions, in nested manners.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 1

- An R-language package, usable in GUI and text forms. (Text form does have some advantages.)
- A "statistical audit."
- Enable other scientists to "revisit" all numerical aspects of a scientific investigation:
  - Data wrangling — dealing with outliers, missing values etc.
  - Statistical analyses — assumptions, methods, variables, time frames, etc.
  - Allows users to ask the "What if?" questions, in nested manners. E.g. asking 3 binary What Ifs forms 8 scenarios,

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 1

- An R-language package, usable in GUI and text forms. (Text form does have some advantages.)
- A "statistical audit."
- Enable other scientists to "revisit" all numerical aspects of a scientific investigation:
  - Data wrangling — dealing with outliers, missing values etc.
  - Statistical analyses — assumptions, methods, variables, time frames, etc.
  - Allows users to ask the "What if?" questions, in nested manners. E.g. asking 3 binary What Ifs forms 8 scenarios, which in **revisit** we call *branches* after GitHub.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 1

- An R-language package, usable in GUI and text forms. (Text form does have some advantages.)
- A "statistical audit."
- Enable other scientists to "revisit" all numerical aspects of a scientific investigation:
  - Data wrangling — dealing with outliers, missing values etc.
  - Statistical analyses — assumptions, methods, variables, time frames, etc.
  - Allows users to ask the "What if?" questions, in nested manners. E.g. asking 3 binary What Ifs forms 8 scenarios, which in **revisit** we call *branches* after GitHub.
- Enable the original research team itself to do the above *during the research project*.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Potti Case

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Potti Case

- Cancer, genomics research.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Potti Case

- Cancer, genomics research.
- Clinical trials patients sued.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Potti Case

- Cancer, genomics research.
- Clinical trials patients sued.
- Sloppiness, apparent fraud.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Potti Case

- Cancer, genomics research.
- Clinical trials patients sued.
- Sloppiness, apparent fraud.
- But also *poor use of statistical methods*.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

The Statistical Methodology Aspect

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

The Statistical Methodology Aspect

Nature reported on a survey of scientists about the reproducibility problem (emphasis added),

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

The Statistical Methodology Aspect

Nature reported on a survey of scientists about the reproducibility problem (emphasis added),

> *More than 60% of respondents [cited]...pressure to publish and* **selective reporting***...More than half pointed to insufficient replication in the lab, poor oversight or* **low statistical power***.*
> *Respondents were asked to rate 11 different approaches to improving reproducibility...Nearly 90% — more than 1,000 people — ticked* **"More robust experimental design," "better statistics"...**

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

ASA Statement on P-Values

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

ASA Statement on P-Values

- The problems have been known all along, e.g. Meehl, "Ronald [Fisher] has befuddled us, mesmerized us, and led us down the primrose path."

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

ASA Statement on P-Values

- The problems have been known all along, e.g. Meehl, "Ronald [Fisher] has befuddled us, mesmerized us, and led us down the primrose path."
- But as Twain said, "Everyone talks about the weather but no one does anything about it."

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

ASA Statement on P-Values

- The problems have been known all along, e.g. Meehl, "Ronald [Fisher] has befuddled us, mesmerized us, and led us down the primrose path."
- But as Twain said, "Everyone talks about the weather but no one does anything about it."
- Hence ASA's first-ever, and long overdue, policy statement, 2016.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

ASA Statement on P-Values

- The problems have been known all along, e.g. Meehl, "Ronald [Fisher] has befuddled us, mesmerized us, and led us down the primrose path."
- But as Twain said, "Everyone talks about the weather but no one does anything about it."
- Hence ASA's first-ever, and long overdue, policy statement, 2016.
- "The ASA releases this guidance on p-values to improve the conduct and interpretation of quantitative science and inform the growing emphasis on reproducibility of science research."

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Multiple Inference Methods

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Multiple Inference Methods

Lack of use of multiple/simultaneous inference procedures, a severe problem:

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Multiple Inference Methods

Lack of use of multiple/simultaneous inference procedures, a severe problem:

- Even if there are no underlying interesting effects, if one performs enough tests, some "significant" ones will be found.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Multiple Inference Methods

Lack of use of multiple/simultaneous inference procedures, a severe problem:

- Even if there are no underlying interesting effects, if one performs enough tests, some "significant" ones will be found.
- Old statistical joke: "If you beat the data long enough, they will confess."

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Multiple Inference Methods

Lack of use of multiple/simultaneous inference procedures, a severe problem:

- Even if there are no underlying interesting effects, if one performs enough tests, some "significant" ones will be found.
- Old statistical joke: "If you beat the data long enough, they will confess."
- ASA statement decries "p-hacking," "data dredging," and "publishing only significant results."

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Multiple Inference Methods

Lack of use of multiple/simultaneous inference procedures, a severe problem:

- Even if there are no underlying interesting effects, if one performs enough tests, some "significant" ones will be found.
- Old statistical joke: "If you beat the data long enough, they will confess."
- ASA statement decries "p-hacking," "data dredging," and "publishing only significant results."
- One of the problems cited in the Potti case was "overfitting," here meaning the above.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 2

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 2

Encourage statistical best practices.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 2

Encourage statistical best practices.

- The **revisit** package aims to wean users away from relying much on p-values.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 2

Encourage statistical best practices.

- The **revisit** package aims to wean users away from relying much on p-values.
- Unfortunately,scientific journals will still require p-values,

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 2

Encourage statistical best practices.

- The **revisit** package aims to wean users away from relying much on p-values.
- Unfortunately,scientific journals will still require p-values, but **revisit** users can learn to de-emphasize them, and add confidence intervals, much more informative.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 2

Encourage statistical best practices.

- The **revisit** package aims to wean users away from relying much on p-values.
- Unfortunately,scientific journals will still require p-values, but **revisit** users can learn to de-emphasize them, and add confidence intervals, much more informative.
- The package will be adding confidence interval alternatives to testing-only procedures.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 2

Encourage statistical best practices.

- The **revisit** package aims to wean users away from relying much on p-values.
- Unfortunately,scientific journals will still require p-values, but **revisit** users can learn to de-emphasize them, and add confidence intervals, much more informative.
- The package will be adding confidence interval alternatives to testing-only procedures.
- E.g. log-linear model.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 2

Encourage statistical best practices.

- The **revisit** package aims to wean users away from relying much on p-values.
- Unfortunately,scientific journals will still require p-values, but **revisit** users can learn to de-emphasize them, and add confidence intervals, much more informative.
- The package will be adding confidence interval alternatives to testing-only procedures.
- E.g. log-linear model. Presently even point estimates in R are available only on request, and even then without standard errors.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals of 'revisit': Part 2

Encourage statistical best practices.

- The **revisit** package aims to wean users away from relying much on p-values.
- Unfortunately,scientific journals will still require p-values, but **revisit** users can learn to de-emphasize them, and add confidence intervals, much more informative.
- The package will be adding confidence interval alternatives to testing-only procedures.
- E.g. log-linear model. Presently even point estimates in R are available only on request, and even then without standard errors.
- Solution: Apply the "Poisson trick" and use **glm()**.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals, Part 2, cont'd.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals, Part 2, cont'd.

Dealing with the multiple inference issue:

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals, Part 2, cont'd.

Dealing with the multiple inference issue:

- The package attempts to count how many tests/confidence intervals the user has (directly or indirectly) performed.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals, Part 2, cont'd.

Dealing with the multiple inference issue:

- The package attempts to count how many tests/confidence intervals the user has (directly or indirectly) performed. E.g., a call to **lm()** gives a p-value for each estimated coefficient.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals, Part 2, cont'd.

Dealing with the multiple inference issue:

- The package attempts to count how many tests/confidence intervals the user has (directly or indirectly) performed. E.g., a call to **lm()** gives a p-value for each estimated coefficient.
- If too many (value might be user-defined), **revisit** issues a warning, "Consider using multiple inference methods."

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals, Part 2, cont'd.

Dealing with the multiple inference issue:

- The package attempts to count how many tests/confidence intervals the user has (directly or indirectly) performed. E.g., a call to **lm()** gives a p-value for each estimated coefficient.
- If too many (value might be user-defined), **revisit** issues a warning, "Consider using multiple inference methods."
- Kind of like, "This deduction may result in an IRS audit."

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals, Part 2, cont'd.

Dealing with the multiple inference issue:

- The package attempts to count how many tests/confidence intervals the user has (directly or indirectly) performed. E.g., a call to **lm()** gives a p-value for each estimated coefficient.
- If too many (value might be user-defined), **revisit** issues a warning, "Consider using multiple inference methods."
- Kind of like, "This deduction may result in an IRS audit."

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Goals, Part 2, cont'd.

Dealing with the multiple inference issue:

- The package attempts to count how many tests/confidence intervals the user has (directly or indirectly) performed. E.g., a call to **lm()** gives a p-value for each estimated coefficient.
- If too many (value might be user-defined), **revisit** issues a warning, "Consider using multiple inference methods."
- Kind of like, "This deduction may result in an IRS audit."

- Currently only Bonferroni offered, more coming.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Wrapper Functions

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Wrapper Functions

- Users call **revisit** functions, as wrappers to standard R functions.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Wrapper Functions

- Users call **revisit** functions, as wrappers to standard R functions.
- E.g., **lm()** (not implemented yet).

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Wrapper Functions

- Users call **revisit** functions, as wrappers to standard R functions.
- E.g., **lm()** (not implemented yet). User runs our wrapper, **lm.rv()**.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Wrapper Functions

- Users call **revisit** functions, as wrappers to standard R functions.
- E.g., **lm()** (not implemented yet). User runs our wrapper, **lm.rv()**.
- Among other things, **lm.rv()** will run, say, **qr()** from the **quantreg** package, then display for the user the two sets of estimated regression coefficients. If they differ much, some outlier hunting/deletion might be warranted.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Wrapper Functions

- Users call **revisit** functions, as wrappers to standard R functions.
- E.g., **lm()** (not implemented yet). User runs our wrapper, **lm.rv()**.
- Among other things, **lm.rv()** will run, say, **qr()** from the **quantreg** package, then display for the user the two sets of estimated regression coefficients. If they differ much, some outlier hunting/deletion might be warranted.
- In addition, some plots from my **regtools** package will be run (on CRAN, coordinated with my new book).

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Structure of the revisit Package

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Structure of the revisit Package

- Available at **github/matloff**.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Structure of the revisit Package

- Available at **github/matloff**.
- GUI and text versions.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Structure of the revisit Package

- Available at **github/matloff**.
- GUI and text versions.
- Case studies.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Structure of the revisit Package

- Available at **github/matloff**.
- GUI and text versions.
- Case studies.
- GitHub-inspired branch structure. Each "What if?" scenario is saved as a separate file.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Structure of the revisit Package

- Available at **github/matloff**.
- GUI and text versions.
- Case studies.
- GitHub-inspired branch structure. Each "What if?" scenario is saved as a separate file.
- Example scenario: Delete outliers x, y and z; choose predictor variables u and v; use Bonferroni adjustments.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Structure of the revisit Package

- Available at **github/matloff**.
- GUI and text versions.
- Case studies.
- GitHub-inspired branch structure. Each "What if?" scenario is saved as a separate file.
- Example scenario: Delete outliers x, y and z; choose predictor variables u and v; use Bonferroni adjustments.
- A scientist who has explored several scenarios can package these branches and send them to others for further exploration.

Norm Matloff Dept. of Computer Science University of California at Davis

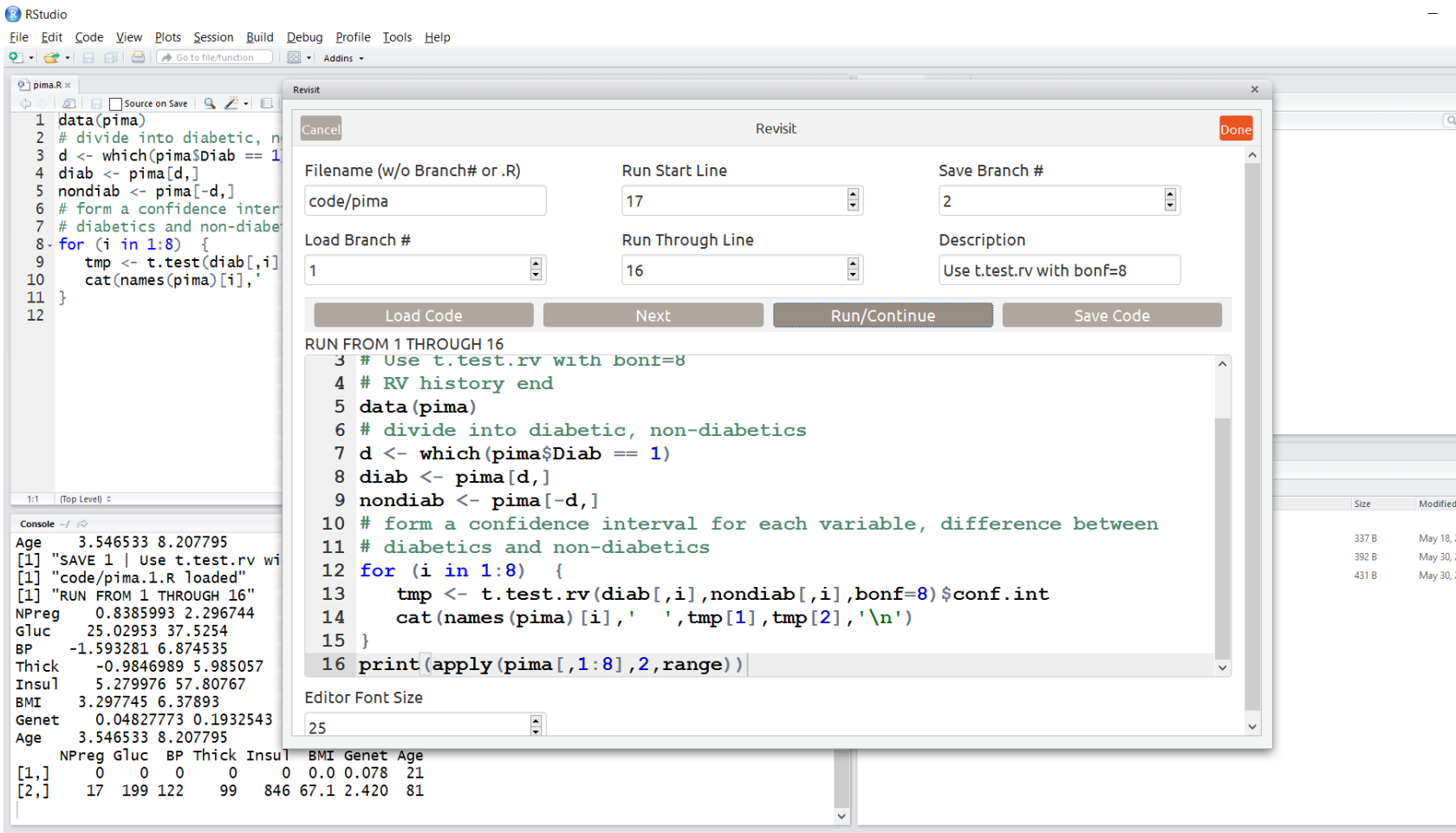
with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

GUI Example

Norm Matloff
Dept. of Computer Science
University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

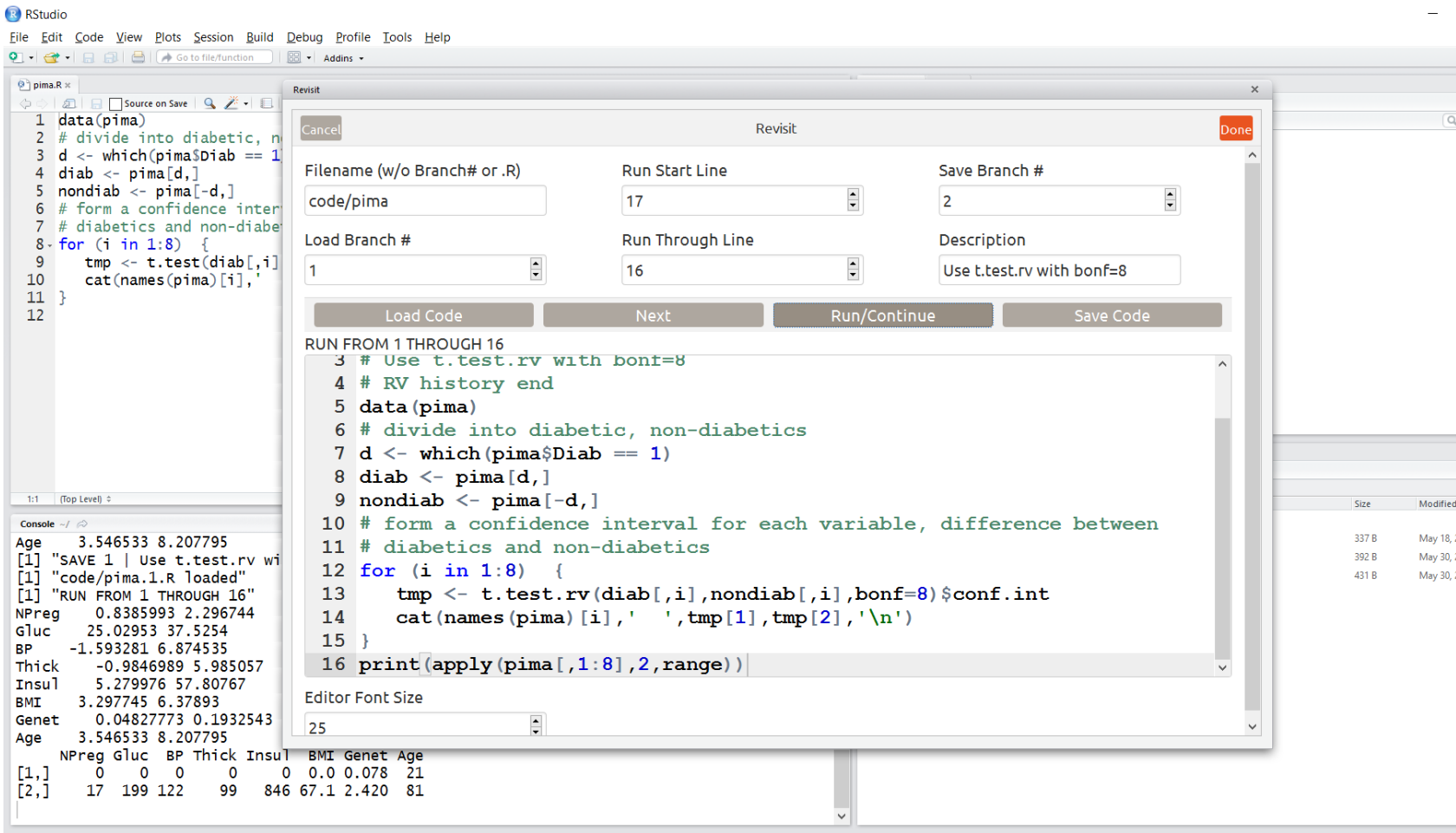
GUI Example



Norm Matloff
Dept. of Computer Science
University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

GUI Example



But it will be clearer to display text here.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Case Study: Zavodny

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Case Study: Zavodny

- Zavodny study, commissioned by an advocacy group in 2011, of impact of H-1B work visa program on U.S. workers.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Case Study: Zavodny

- Zavodny study, commissioned by an advocacy group in 2011, of impact of H-1B work visa program on U.S. workers.
- Highly controversial, much criticism of the visa by Clinton, Sanders, Trump etc. in 2016 election.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Case Study: Zavodny

- Zavodny study, commissioned by an advocacy group in 2011, of impact of H-1B work visa program on U.S. workers.
- Highly controversial, much criticism of the visa by Clinton, Sanders, Trump etc. in 2016 election.
- Zavodny found that each visa worker creates 2.62 new jobs for Americans. Peri (2014), also funded by an advocacy group, had similar findings. Gelber *et al* found the opposite, a crowding-out of U.S. workers.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Case Study: Zavodny

- Zavodny study, commissioned by an advocacy group in 2011, of impact of H-1B work visa program on U.S. workers.
- Highly controversial, much criticism of the visa by Clinton, Sanders, Trump etc. in 2016 election.
- Zavodny found that each visa worker creates 2.62 new jobs for Americans. Peri (2014), also funded by an advocacy group, had similar findings. Gelber *et al* found the opposite, a crowding-out of U.S. workers.
- Dr. Zavodny kindly shared her code and data with Reed Davis, one of the **revisit** authors.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

```
> library(revisit)
> rvinit()  # required initialization
> loadb('ols262.R')  # load the branch
> lcc()  # list the code
...
...
4 data(zav)
5 zav = zav[zav$year < 2008,] # 2008-2010 removed
...
...
```

Again, this is R code converted from Stata. Does it reproduce Zavodny's results? Yes:

```
> runb()
[1] "Slope   = 0.00446438147988468"
[1] "P-value = 0.0140870195483076"
[1] "Jobs    = 262.985782017836"
```

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

But Zavodny omitted 2008-2010. What if...?

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

But Zavodny omitted 2008-2010. What if...?

We call **revisit** function **edt()** to edit the code (visual editor in GUI),

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

But Zavodny omitted 2008-2010. What if...?
We call **revisit** function **edt()** to edit the code (visual editor in GUI), commenting out line 5. Then:

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

But Zavodny omitted 2008-2010. What if...?
We call **revisit** function **edt()** to edit the code (visual editor in GUI), commenting out line 5. Then:

```
> runb()
[1] "Slope    = 0.00180848722715659"
[1] "P-value = 0.33637275201986"
[1] "Jobs     = 124.352299406043"
```

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

But Zavodny omitted 2008-2010. What if...?
We call **revisit** function **edt()** to edit the code (visual editor in GUI), commenting out line 5. Then:

```
> runb()
[1] "Slope    = 0.00180848722715659"
[1] "P-value = 0.33637275201986"
[1] "Jobs     = 124.352299406043"
```

Now, the result is no longer significant [sic], and the point estimate has been cut in half.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

Some other What Ifs:

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

Some other What Ifs:

Adj. R^2 is 0.91, quite high.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

Some other What Ifs:

Adj. R^2 is 0.91, quite high. The author's model include dummies for state effects. What if we remove them?

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

Some other What Ifs:
Adj. R^2 is 0.91, quite high. The author's model include dummies for state effects. What if we remove them?

```
...

Coefficients:
                                  Estimate  Std. Error
(Intercept)                      4.1780416   0.0106922
lnimmshare_emp_stem_e_grad      -0.0130295   0.0036493
lnimmshare_emp_stem_n_grad       0.0005722   0.0040274
fyear2001                       -0.0098670   0.0104854
...
Multiple R-squared:  0.372,  Adj. R-squared:  0.3517
```

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

Some other What Ifs:
Adj. R^2 is 0.91, quite high. The author's model include dummies for state effects. What if we remove them?

```
...

Coefficients:
                                  Estimate  Std. Error
(Intercept)                      4.1780416   0.0106922
lnimmshare_emp_stem_e_grad      -0.0130295   0.0036493
lnimmshare_emp_stem_n_grad       0.0005722   0.0040274
fyear2001                       -0.0098670   0.0104854
...
Multiple R-squared:  0.372,  Adj. R-squared:  0.3517
```

Using only immigrant share and time effects, R^2 drops a lot.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Zavodny, cont'd.

Very complex topic, many assumptions etc.
But clearly Zavodny's "2.62 jobs created by each H-1B" figure – very widely cited in the press — cannot be taken as definitive.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Who Might Use revisit?

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Who Might Use revisit?

- Good for coordination within a research team, during a project, asking a lot pf What If questions.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Who Might Use revisit?

- Good for coordination within a research team, during a project, asking a lot pf What If questions.
- Voluntary publication of the data and **revisit** files by a research team, in the spirit of open intellectual inquiry and reproducibility.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Who Might Use revisit?

- Good for coordination within a research team, during a project, asking a lot pf What If questions.
- Voluntary publication of the data and **revisit** files by a research team, in the spirit of open intellectual inquiry and reproducibility. Would have been nice if Reinhart, Rogoff, Potti and Zavodny had done this.

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Who Might Use revisit?

- Good for coordination within a research team, during a project, asking a lot pf What If questions.
- Voluntary publication of the data and **revisit** files by a research team, in the spirit of open intellectual inquiry and reproducibility. Would have been nice if Reinhart, Rogoff, Potti and Zavodny had done this.
- Possible adoption by journals and funding agencies, as a requirement for publication/funding?

Norm Matloff Dept. of Computer Science University of California at Davis

with Laurel Beckett, Tiffany Chen, Reed Davis, Paul Thompson and Emily Watkins

Who Might Use revisit?

- Good for coordination within a research team, during a project, asking a lot pf What If questions.
- Voluntary publication of the data and **revisit** files by a research team, in the spirit of open intellectual inquiry and reproducibility. Would have been nice if Reinhart, Rogoff, Potti and Zavodny had done this.
- Possible adoption by journals and funding agencies, as a requirement for publication/funding?
- Use as a teaching tool, especially with the case studies.