Transparent and Reproducible Social Science Research

University of Luxembourg
October 2018

Vernon Gayle
University of Edinburgh

Roxanne Connelly
University of York

RUSSELL GROUP

A Thought Experiment

You are examining a PhD thesis or refereeing a paper

You turn the page and read the following output…

A Thought Experiment

A logistic regression model of some data from

The Youth Cohort Study of England and Wales (1997)

A Thought Experiment

The outcome is passing 5+ GCSEs at grades A*- C

The explanatory variables are

Gender, Ethnicity and Parental Social Class

n = 19,216

Logistic Regression 5+ GCSEs (A*-C) YCS Cohort 9

	b	se	t	p
5+ GCSEs (A*-C)				
Girls	-0.039	0.392	10.066	0.000
Boys	0.000	.	.	.
Chinese	1.719	0.287	5.988	0.000
Indian	1.028	0.210	4.899	0.000
White	0.997	0.160	6.241	0.000
Bangladeshi	1.351	0.317	4.267	0.000
Pakistani	0.677	0.209	3.241	0.001
Black	0.000	.	.	.
Prof/Managerial	1.945	0.078	25.093	0.000
Non-Manual(other)	1.335	0.085	15.687	0.000
Skilled Manual	0.779	0.079	9.908	0.000
Semi-skilled	0.310	0.086	3.611	0.000
Unskilled	0.000	.	.	.
Constant	-0.958	0.175	-5.475	0.000
Pseudo R^2	0.80			
n	19216			

Logistic Regression 5+ GCSEs (A*-C) YCS Cohort 9

	b	se	t	p
5+ GCSEs (A*-C)				
Girls	-0.039	0.392	10.066	0.000
Boys	0.000	.	.	.
Chinese	1.719	0.287	5.988	0.000
Indian	1.028	0.210	4.899	0.000
White	0.997	0.160	6.241	0.000
Bangladeshi	1.351	0.317	4.267	0.000
Pakistani	0.677	0.209	3.241	0.001
Black	0.000	.	.	.
Prof/Managerial	1.945	0.078	25.093	0.000
Non-Manual(other)	1.335	0.085	15.687	0.000
Skilled Manual	0.779	0.079	9.908	0.000
Semi-skilled	0.310	0.086	3.611	0.000
Unskilled	0.000	.	.	.
Constant	-0.958	0.175	-5.475	0.000
Pseudo R^2	0.80			
n	19216			

Logistic Regression 5+ GCSEs (A*-C) YCS Cohort 9

	b	se	t	p
5+ GCSEs (A*-C)				
Girls	-0.039	0.392	10.066	0.000
Boys	0.000	.	.	.
Chinese	1.719	0.287	5.988	0.000
Indian	1.028	0.210	4.899	0.000
White	0.997	0.160	6.241	0.000
Bangladeshi	1.351	0.317	4.267	0.000
Pakistani	0.677	0.209	3.241	0.001
Black	0.000	.	.	.
Prof/Managerial	1.945	0.078	25.093	0.000
Non-Manual(other)	1.335	0.085	15.687	0.000
Skilled Manual	0.779	0.079	9.908	0.000
Semi-skilled	0.310	0.086	3.611	0.000
Unskilled	0.000	.	.	.
Constant	-0.958	0.175	-5.475	0.000
Pseudo R^2	0.08			
n	19216			

Logistic Regression 5+ GCSEs (A*-C) YCS Cohort 9

	b	se	t	p
5+ GCSEs (A*-C)				
Girls	-0.039	0.392	10.066	0.000
Boys	0.000	.	.	.
Chinese	1.719	0.287	5.988	0.000
Indian	1.028	0.210	4.899	0.000
White	0.997	0.160	6.241	0.000
Bangladeshi	1.351	0.317	4.267	0.000
Pakistani	0.677	0.209	3.241	0.001
Black	0.000	.	.	.
Prof/Managerial	1.945	0.078	25.093	0.000
Non-Manual(other)	1.335	0.085	15.687	0.000
Skilled Manual	0.779	0.079	9.908	0.000
Semi-skilled	0.310	0.086	3.611	0.000
Unskilled	0.000	.	.	.
Constant	-0.958	0.175	-5.475	0.000
Pseudo R^2	0.08			
n	19216			

Logistic Regression 5+ GCSEs (A*-C) YCS Cohort 9

	b	se	t	p
5+ GCSEs (A*-C)				
Girls	-0.392	0.039	10.066	0.000
Boys	0.000	.	.	.
Chinese	1.719	0.287	5.988	0.000
Indian	1.028	0.210	4.899	0.000
White	0.997	0.160	6.241	0.000
Bangladeshi	1.351	0.317	4.267	0.000
Pakistani	0.677	0.209	3.241	0.001
Black	0.000	.	.	.
Prof/Managerial	1.945	0.078	25.093	0.000
Non-Manual(other)	1.335	0.085	15.687	0.000
Skilled Manual	0.779	0.079	9.908	0.000
Semi-skilled	0.310	0.086	3.611	0.000
Unskilled	0.000	.	.	.
Constant	-0.958	0.175	-5.475	0.000
Pseudo R^2	0.08			
n	19216			

Logistic Regression 5+ GCSEs (A*-C) YCS Cohort 9

	b	se	t	p
5+ GCSEs (A*-C)				
Girls	-0.392	0.039	10.066	0.000
Boys	0.000	.	.	.
Chinese	1.719	0.287	5.988	0.000
Indian	1.028	0.210	4.899	0.000
White	0.997	0.160	6.241	0.000
Bangladeshi	1.351	0.317	4.267	0.000
Pakistani	0.677	0.209	3.241	0.001
Black	0.000	.	.	.
Prof/Managerial	1.945	0.078	25.093	0.000
Non-Manual(other)	1.335	0.085	15.687	0.000
Skilled Manual	0.779	0.079	9.908	0.000
Semi-skilled	0.310	0.086	3.611	0.000
Unskilled	0.000	.	.	.
Constant	-0.958	0.175	-5.475	0.000
Pseudo R^2	0.08			
n	19216			

Logistic Regression 5+ GCSEs (A*-C) YCS Cohort 9

	b	se	t	p
5+ GCSEs (A*-C)				
Girls	0.392	0.039	10.066	0.000
Boys	0.000	.	.	.
Chinese	1.719	0.287	5.988	0.000
Indian	1.028	0.210	4.899	0.000
White	0.997	0.160	6.241	0.000
Bangladeshi	1.351	0.317	4.267	0.000
Pakistani	0.677	0.209	3.241	0.001
Black	0.000	.	.	.
Prof/Managerial	1.945	0.078	25.093	0.000
Non-Manual(other)	1.335	0.085	15.687	0.000
Skilled Manual	0.779	0.079	9.908	0.000
Semi-skilled	0.310	0.086	3.611	0.000
Unskilled	0.000	.	.	.
Constant	-0.958	0.175	-5.475	0.000
Pseudo R^2	0.08			
n	19216			

Logistic Regression 5+ GCSEs (A*-C) YCS Cohort 9

	b	se	t	p
5+ GCSEs (A*-C)				
Girls	0.392	0.039	10.066	0.000
Boys	0.000	.	.	.
Chinese	1.719	0.287	5.988	0.000
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
White	0.997	0.160	6.241	0.000
Bangladeshi	1.351	0.317	4.267	0.000
[illegible]	[illegible]	[illegible]	3.241	0.001
Black	0.000	.	.	.
Prof/Managerial	1.945	0.078	25.093	0.000
Non-Manual(other)	1.335	0.085	15.687	0.000
Skilled Manual	0.779	0.079	9.908	0.000
Semi-skilled	0.310	0.086	3.611	0.000
Unskilled	0.000	.	.	.
Constant	-0.958	0.175	-5.475	0.000
Pseudo R^2	0.08			
n	19216			

The final set of results looked highly plausible.

Alarmingly the results were produced on data that are completely fabricated!

Logistic Regression 5+ GCSEs (A*-C) YCS Cohort 9

	b	se	t	p
5+ GCSEs (A*- C)				
Girls	0.405	0.039	10.305	0.000
Boys	0.000	.	.	.
Chinese	2.002	0.377	5.306	0.000
Indian	1.066	0.208	5.117	0.000
White	0.643	0.171	3.757	0.000
Bangladeshi	0.766	0.345	2.222	0.026
Pakistani	0.531	0.245	2.169	0.030
Black	0.000	.	.	.
Professional/Non-Man	2.192	0.109	20.179	0.000
Other Non-Manual	1.773	0.108	16.423	0.000
Skilled Manual	0.932	0.104	8.954	0.000
Semi-Skilled Manual	0.576	0.113	5.112	0.000
Unskilled	0.000	.	.	.
Constant	-2.208	0.198	-11.152	0.000
n	12789			

Produced in Stata using svy; Connolly (2006) used SPSS with data weighted incorrectly!

Logistic Regression 5+ GCSEs (A*-C) YCS Cohort 9

	b	se	t	p
5+ GCSEs (A*- C)				
Girls	0.405	0.039	10.305	0.000
Boys	0.000	.	.	.
Chinese	2.002	0.377	5.306	0.000
Indian	1.066	0.208	5.117	0.000
White	0.643	0.171	3.757	0.000
Bangladeshi	0.766	0.345	2.222	0.026
Pakistani	0.531	0.245	2.169	0.030
Black	0.000	.	.	.
Professional/Non-Manual	2.192	0.109	20.179	0.000
Other Non-Manual	1.773	0.108	16.423	0.000
Skilled Manual	0.932	0.104	8.954	0.000
Semi-Skilled Manual	0.576	0.113	5.112	0.000
Unskilled	0.000	.	.	.
Constant	-2.208	0.198	-11.152	0.000
n	12789			

Produced in Stata using svy; Connolly (2006) used SPSS with data weighted incorrectly!

Appendix: details of binary logistic regression models derived from the three cohorts (see Table 3)

Table 5. Binary logistic regression on whether school leavers in England and Wales in 1997 gained five or more GCSE grades A*–C or not[1]

	B	S.E.	Wald	df	Sig.	Exp(B)
Gender[2]						
Girls	0.405	0.038	114.340	1	<.001	1.499
Ethnicity[3]						
Chinese	2.002	0.341	34.436	1	<.001	7.406
Indian	1.066	0.193	30.389	1	<.001	2.903
White	0.643	0.159	16.372	1	<.001	1.902
Bangladeshi	0.766	0.332	5.330	1	.021	2.151
Pakistani	0.531	0.230	5.338	1	.021	1.701
Social Class[4]						
Professional/ Managerial	2.192	0.110	396.863	1	<.001	8.954
Other Non-Manual	1.773	0.110	261.000	1	<.001	5.886
Skilled Manual	0.932	0.107	76.255	1	<.001	2.540
Semi-Skilled Manual	0.576	0.115	24.965	1	<.001	1.779
Constant	−2.208	0.189	136.885	1	<.001	0.110

[1]Source of data: Secondary analysis of data derived from first sweep of Cohort 9 of the Youth Cohort Study of England and Wales.
[2]Reference category: boys.
[3]Reference category: Black pupils.
[4]Reference category: Unskilled Manual Occupations.

Connolly, Paul. "The effects of social class and ethnicity on gender differences in GCSE attainment: a secondary analysis of the Youth Cohort Study of England and Wales 1997–2001." *British Educational Research Journal* 32.1 (2006): 3-21.

Logistic Regression 5+ GCSEs (A*-C) YCS Cohort 9

	b	se	t	p
5+ GCSEs (A*-C)				
Girls	0.302	0.039	10.000	0.000
Boys	0.000	.	.	.
Chinese	1.719	0.287	5.988	0.000
Indian	1.028	0.210	4.899	0.000
White	0.997	0.160	6.241	0.000
Bangladeshi	1.351	0.317	4.267	0.000
Pakistani	0.677	0.209	3.241	0.001
Black	0.000	.	.	.
Prof/Managerial	1.945	0.078	25.093	0.000
Non-Manual(other)	1.335	0.085	15.687	0.000
Skilled Manual	0.779	0.079	9.908	0.000
Semi-skilled	0.310	0.086	3.611	0.000
Unskilled	0.000	.	.	.
Constant	-0.958	0.175	-5.475	0.000
Pseudo R^2	0.08			
n	19216			

Are we really examining or refereeing what we think we are?

www.vernongayle.com/talks-recent.html

Research Talks

This page contains files related to research talks that I have delivered

Gayle, V. (2016) 'Is the Paper Just a Palimpsest? An appeal for reproducible social stratification research', *Social Stratification Research Seminar*, Cambridge.

Files supporting reproducibility

Stata syntax file (research code) cambridge_20160901_vg_v4.do

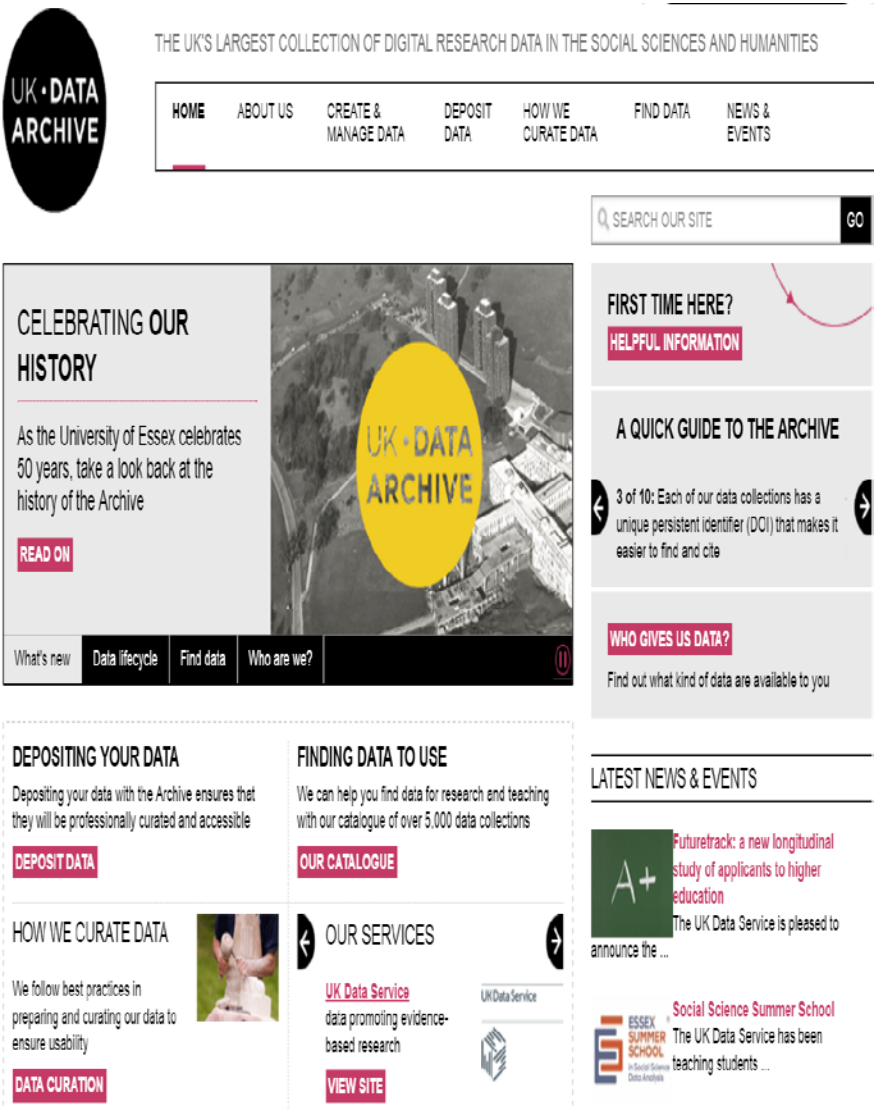
Stata 14 data file fake_data_20160828_vg_v1.dta

Stata 13 data file fake_data_20160828_stata13_vg_v1.dta

Excel versions of the data file fake_data_20160828_vg_v1.xlsx

Excel versions of the data file (csv) fake_data_20160828_vg_v1.csv

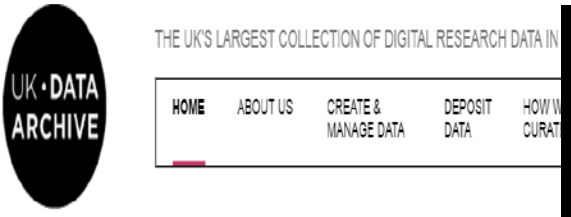
Access to Data



Access to Outputs

Access to Data

Access to Outputs



A BLACK BOX

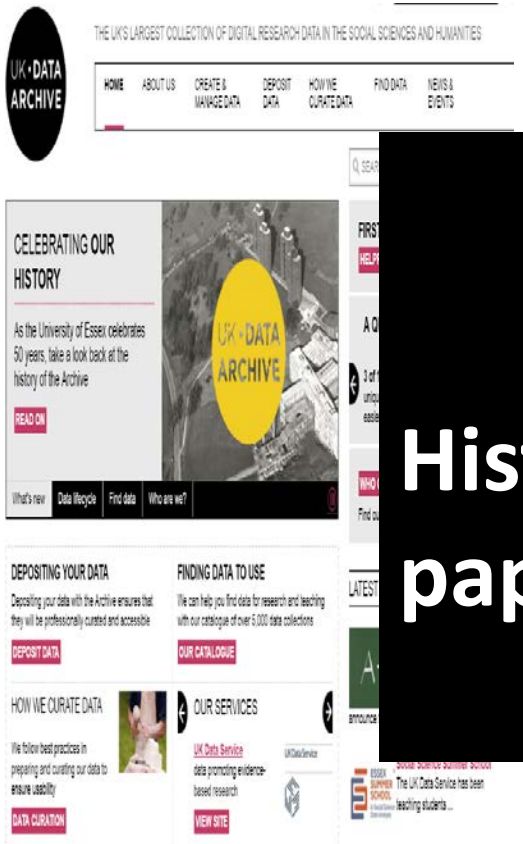
Lack of access to the research code that produced the research output

Command files
Syntax files
.do Stata files
R scripts
.sps SPSS files
Jupyter notebooks

A BLACK BOX

???

[How was NVivo used]



A BLACK BOX

Historically lack of space in paper journals

Reproducibility in Social Research

Duplication

Work can be duplicated if sufficient information is made available which ensures consistent results can be produced using the same data and the same analytical techniques

Replication

A replication study can ‘duplicate’ the original work….

But can also further test the robustness of the original work by employing new or additional data (or measures) and alternative data analysis techniques

A parchment or other writing surface on which the original text has been effaced or partially erased, and then overwritten by another; OED

palimpsest, *n.* and *adj.*

Pronunciation: Brit. /ˈpalɨm(p)sɛst/ , U.S. /ˈpæləm(p)ˌsɛst/

Young people and school General Certificate of Secondary Education attainment: looking for the 'missing middle'

Vernon Gayle[a]*, Susan Murray[b] and Roxanne Connelly[c]

[a]*School of Social and Political Science, University of Edinburgh, Edinburgh, UK;* [b]*School of Applied Social Science, University of Stirling, Stirling, UK;* [c]*Centre for Longitudinal Studies, Institute of Education, University of London, London, UK*

(*Received 6 July 2013; final version received 12 June 2014*)

In Britain, educational qualifications gained at school continue to play an important and central role in young people's educational and employment pathways. Recently there has been growing interest in documenting the lives of 'ordinary' young people. In this paper we analyse the Youth Cohort Study of England and Wales in order to better document the experiences of those with 'middle' levels of school General Certificate of Secondary Education (GCSE) attainment. The overall pattern of school GCSE attainment is one of increasing levels of performance. GCSE attainment is still highly stratified. Girls performed better than boys, and there were some marked differences in attainment for pupils from the main minority ethnic groups. Most notably, parental socio-economic positions are the most important factor. The analyses fail to persuade us that there are clear boundaries that demark a 'middle' category of school GCSE attainment. We conclude that sociologists should study 'ordinary' young people; however, school GCSE attainment is best understood as a continuum, and measures such as the number of GCSEs or point scores are preferable.

Keywords: youth transitions; sociology of youth; educational attainment; General Certificate of Secondary Education; missing middle; Youth Cohort Study of England and Wales

Introduction

The idea of a 'marginalised mainstream' or 'missing middle' is currently

```
capture drop t03cat
gen t03cat = .
replace t03cat = 0 if (t0examac2==0)
replace t03cat = 1 if (t0examac2>=1)&(t0examac2<=4)
replace t03cat = 2 if (t0examac2>=5)&(t0examac2<=14)
tab t03cat, missing

label define t03catll 0 "none" 1 "1-4" 2 "5+"
label values t03cat t03catll

* first model mlogit

svyset [pw=t1weight]

svy:mlogit t03cat ///
      cohort93 cohort95 cohort97 cohort99 cohort01   boys  ///
        black indian pakistani bangladeshi oasian other ///
                        rented ohouse ///
                                        mumonly dadonly ohh ///
                                        gradpar ///
                                        nssec11 nssec12 nssec2 nssec4 nssec5 nssec6 nssec7 ///
                       ,baseoutcome(1)

                      estimates store mlogit1


*************************************************************************
```

Young people and school General Certificate of Secondary Education attainment: looking for the 'missing middle'

Vernon Gayle[a]*, Susan Murray[b] and Roxanne Connelly[c]

[a]School of Social and Political Science, University of Edinburgh, Edinburgh, UK; [b]School of Applied Social Science, University of Stirling, Stirling, UK; [c]Centre for Longitudinal Studies, Institute of Education, University of London, London, UK

(Received 6 July 2013; final version received 12 June 2014)

In Britain, educational qualifications gained at school continue to play an important and central role in young people's educational and employment pathways. Recently there has been growing interest in documenting the lives of 'ordinary' young people. In this paper we analyse the Youth Cohort Study of England and Wales in order to better document the experiences of those with 'middle' levels of school General Certificate of Secondary Education (GCSE) attainment. The overall pattern of school GCSE attainment is one of increasing levels of performance. GCSE attainment is still highly stratified. Girls performed better than boys, and there were some marked differences in attainment for pupils from the main minority ethnic groups. Most notably, parental socio-economic positions are the most important factor. The analyses fail to persuade us that there are clear boundaries that demark a 'middle' category of school GCSE attainment. We conclude that sociologists should study 'ordinary' young people; however, school GCSE attainment is best understood as a continuum, and measures such as the number of GCSEs or point scores are preferable.

Keywords: youth transitions; sociology of youth; educational attainment; General Certificate of Secondary Education; missing middle; Youth Cohort Study of England and Wales

Introduction

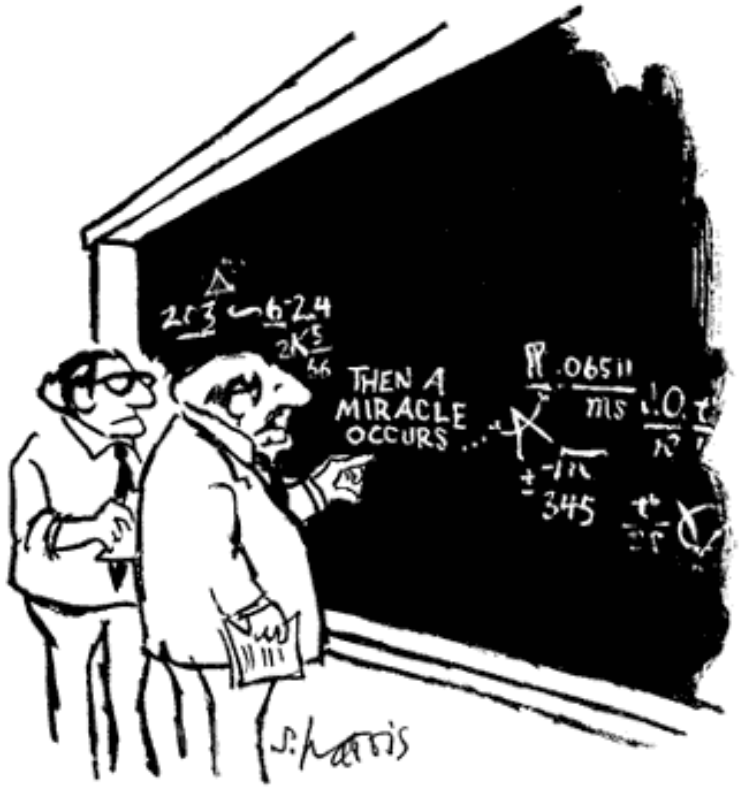
The idea of a 'marginalised mainstream' or 'missing middle' is currently

```
capture drop t03cat
gen t03cat = .
replace t03cat = 0 if (t0examac2==0)
replace t03cat = 1 if (t0examac2>=1)&(t0examac2<=4)
replace t03cat = 2 if (t0examac2>=5)&(t0examac2<=14)
tab t03cat, missing

label define t03catll 0 "none" 1 "1-4" 2 "5+"
label values t03cat t03catll

* first model mlogit

svyset [pw=t1weight]

svy:mlogit t03cat ///
    cohort93 cohort95 cohort97 cohort99 cohort01 boys ///
     black indian pakistani bangladeshi oasian other ///
                                    rented ohouse ///
                                    mumonly dadonly ohh ///
                                    gradpar ///
                                    hssec11 hssec12 hssec2 hssec4 hssec5 hssec6 hssec7 ///
,baseoutcome(1)

estimates store mlogit1
**********************************************************************
```

“Show me your working out”

“Show me your working out”

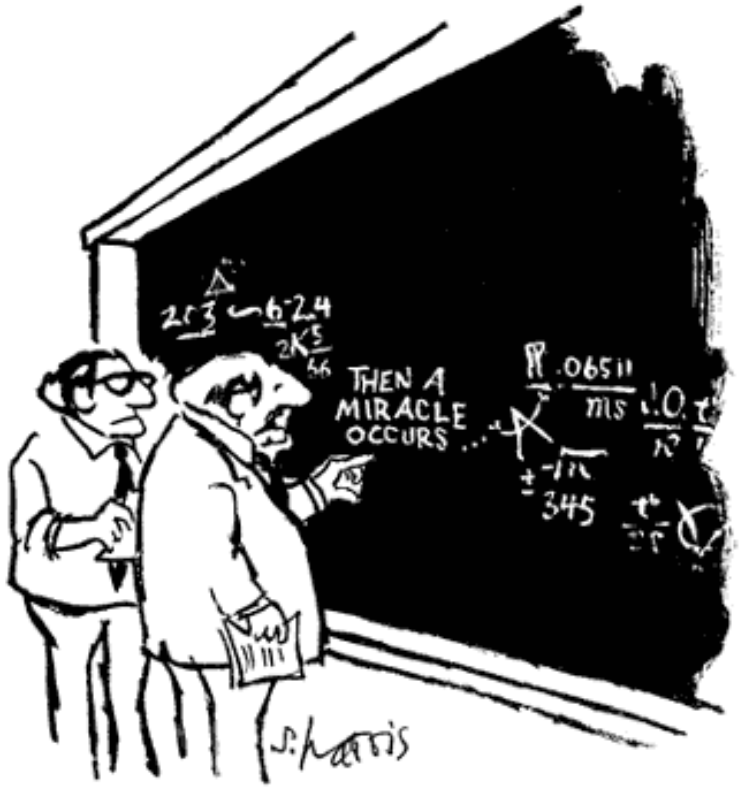
Show me → Trust me



“I think you should be more explicit here in step two.”

“Show me your working out”

The Royal Society's motto ‘*Nullius in verba*’ (take nobody’s word for it)



“I think you should be more explicit here in step two.”

Do Sociologists Release Their Data and Code?
Disappointing Results from a Field Experiment on Replication

sociologists need to be better at replication – a guest post by Cristobal Young
https://orgtheory.wordpress.com/2015/08/11/sociologists-need-to-be-better-at-replication-a-guest-post-by-cristobal-young/

Do Sociologists Release Their Data and Code?
Disappointing Results from a Field Experiment on Replication

Table 1. Response to Replication Request

Response	Frequency	Percent
Yes: Released data and code for paper	15	28%
No: Did not release	38	72%
Reasons for "No"		
IRB / legal / confidentiality issue	12	23%
No response / no follow up	10	19%
Don't have data	6	11%
Don't have time / too complicated	6	11%
Still using the data	2	4%
'See the article and figure it out'	2	4%
Total	53	100%

Figure 1: Illustrative Quotes from Student Correspondence with Authors:

Positive:

1. "Here is the data file and Stata .do file to reproduce [the] Tables…. Let me know if you have any questions."
2. "[Attached are] data and *R* code that does all regression models in the paper. Assuming that you know *R*, you could literally redo the entire paper in a few minutes."

Negative:

3. "While I applaud your efforts to replicate my research, the best guidance I can offer is that the details about the data and analysis strategies are in the paper."
4. "I don't keep or produce 'replication packages'… Data takes a significant amount of human capital and financial resources, and serves as a barrier-to-entry against other researchers… they can do it themselves."

Sharing Research Code

Sufficient information to enable other researchers to

understand

evaluate

build upon the work

Sharing Research Code

Sufficient information to enable other researchers to understand, evaluate and build upon the work

Enough information for a third party to reproduce results without needing to get additional information from the authors

Required by *Science, Am. Econ. Rev., Econometrica, Rev. Econ. Studies*

500+ journals signed up to the Transparency and Openness Promotion (TOP) Guidelines

Some Existing Examples of Code Sharing in Sociology

(e.g. social stratification research)

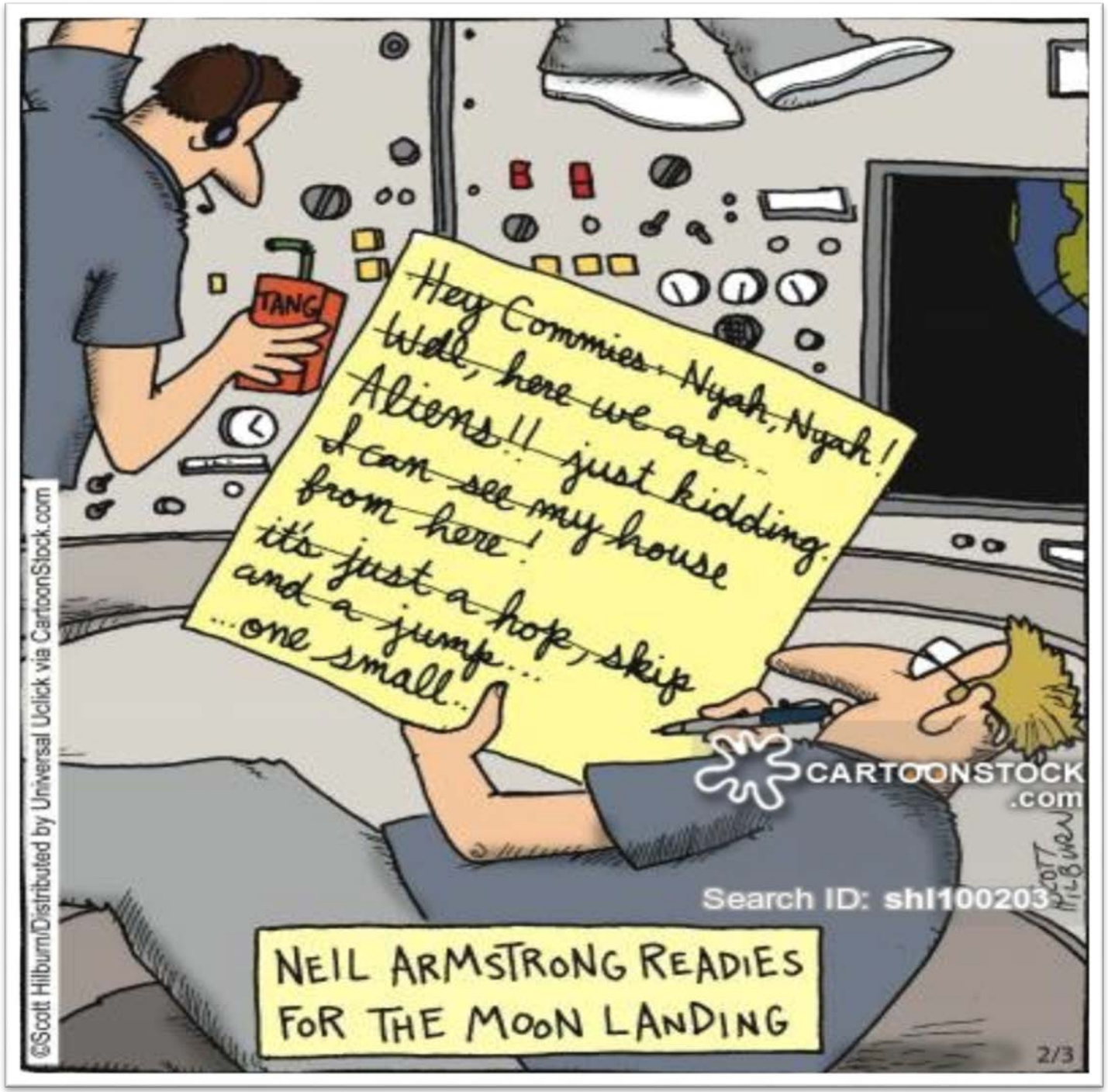
Harry Ganzeboom
http://www.harryganzeboom.nl/index.htm

Don Treiman
https://ccpr.ucla.edu/dtreiman/

Paul Lambert
http://www.camsis.stir.ac.uk/

Kim Weeden
http://www.kimweeden.com/research/occupations/

What really went on…



A small step from private reproducibility could lead to a giant leap for public reproducibility

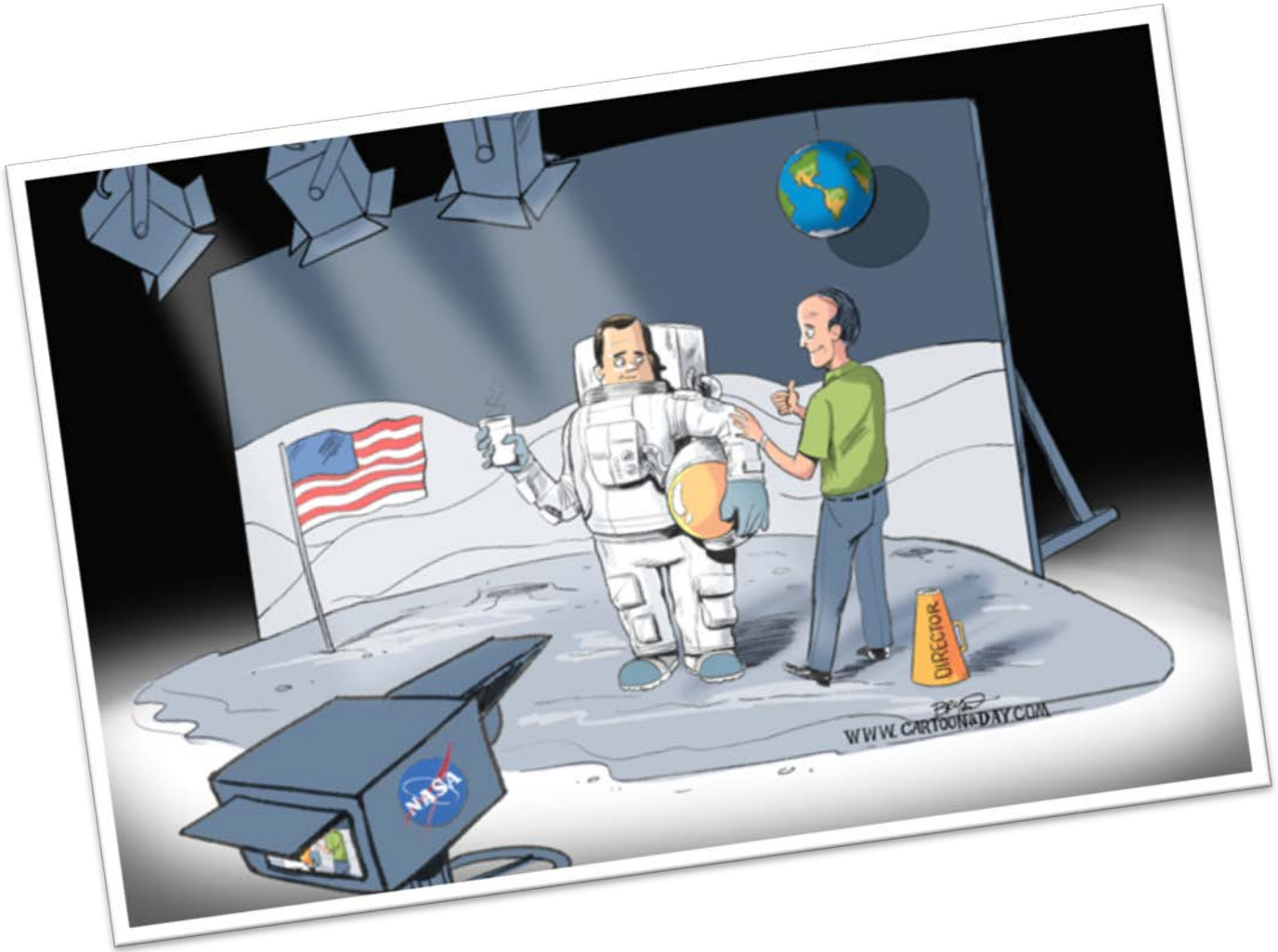


Image credit http://www.cartoonaday.com/armstrongs-moon-landing-was-a-fake/

Why bother?

- Improves transparency - don’t just trust me – I will show you
- Allows others to understand, evaluate, and build upon the work
- Checks on accuracy
- Facilitates incremental development (and comparative work)

Are we serious about what we do? (Edinburgh University’s mission is the creation, dissemination and curation of knowledge)

5 Simple Newer Rules of the Sociological Method

Analysing Large-Scale and Complex Social Science Datasets

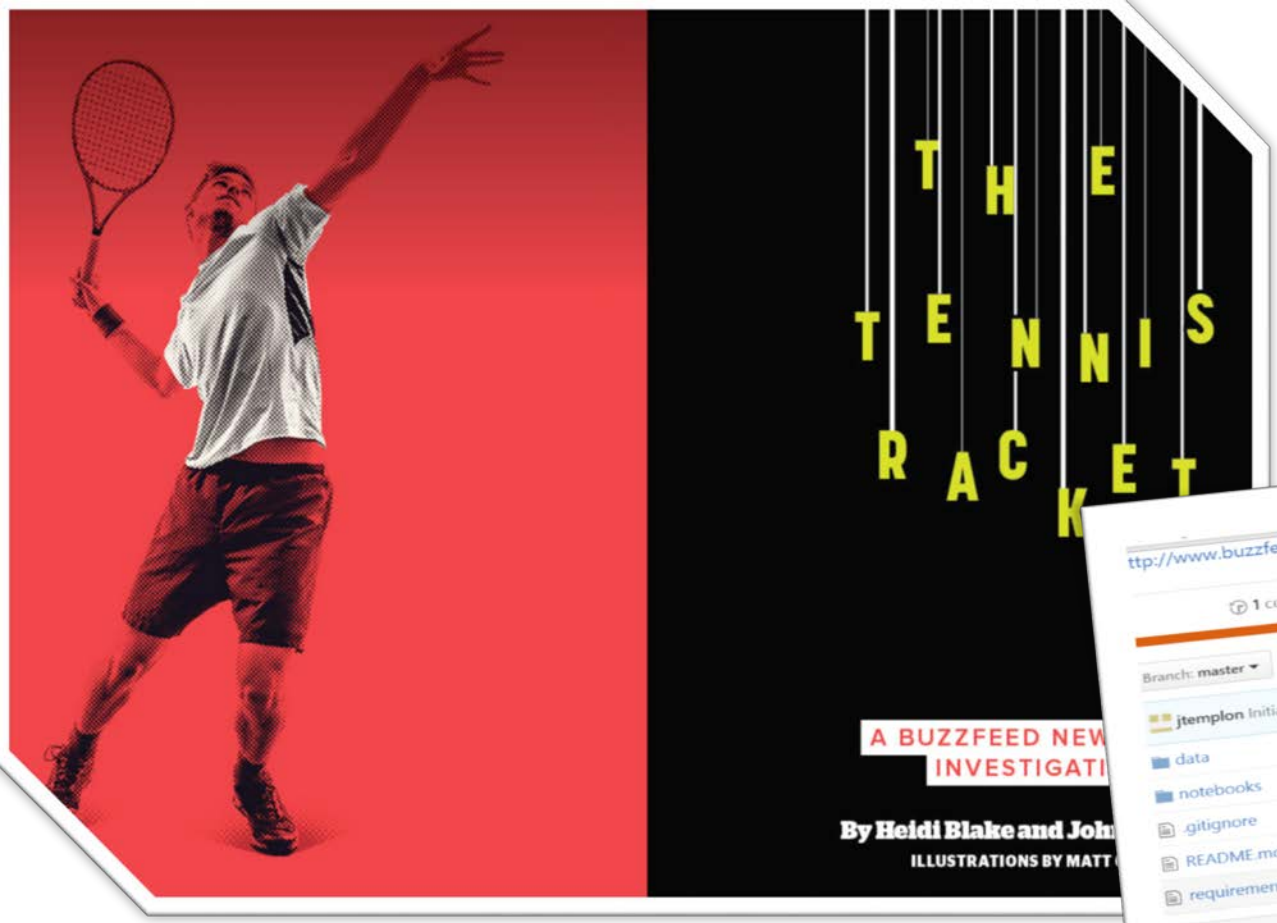
1. Tell us about your software
2. Tell us about your data
3. Show us how you got your data ready
4. Show us all the analysis you did
5. Save all of this work openly

Ottoboni, K., Boring, A. and Stark, P., 2016. Student evaluations of teaching (mostly) do not measure teaching effectiveness. *ScienceOpen Research*.

Student evaluations of teaching (mostly) do not measure teaching effectiveness

Authors: Anne Boring[1,2], Kellie Ottoboni[3], Philip B. Stark[*,3]

Publication date: 07 January 2016
Journal: ScienceOpen Research – Section: SOR-EDU
Publisher: ScienceOpen
Keywords: Assessment, Evaluation & Research methods, Labor law, Nonparametric Statistics, Disparate Impact, Gender Bias, Permutation Tests



ttp://www.buzzfeed.com/heidiblake/the-tennis-racket

1 commit | 1 branch | 0 releases | 1 contributor

Branch: master | New pull request | Create new file | Upload files | Find file | Clone or download

Latest commit 1f08da2 on Jan 17

jtemplon	Initial commit	7 months ago
data	Initial commit	7 months ago
notebooks	Initial commit	7 months ago
.gitignore	Initial commit	7 months ago
README.md	Initial commit	7 months ago
requirements.txt		

README.md

Methodology and Code: Detecting Match-Fixing Patterns In Tennis

A closer look at the data analysis behind BuzzFeed News' investigation into corruption in tennis.

General Notes

In "The Tennis Racket," a yearlong investigation into match-fixing in professional tennis, BuzzFeed News published findings from an original data analysis we performed. That analysis revealed many examples of one particularly suspicious pattern: heavy betting against a player, followed by that player's loss.

Betting patterns alone aren't proof of fixing. Players can underperform for all sorts of reasons — injury, fatigue, bad luck — and sometimes that underperformance will just happen to coincide with heavy betting against them. But it's extremely unlikely for a player to underperform repeatedly in matches on which people just happen to be betting massive sums against

Connelly, R. and Gayle, V., 2017. An investigation of social class inequalities in general cognitive ability in two British birth cohorts. *The British journal of sociology*. https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-4446.12343

RoxanneConnelly / Social-Class-Inequalities-in-General-Cognitive-Ability-in-Two-British-Birth-Cohorts

No description, website, or topics provided.

README.md

An investigation of Social Class Inequalities in General Cognitive Ability in Two British Birth Cohorts

British Journal of Sociology

Roxanne Connelly (R.Connelly@warwick.ac.uk)

Vernon Gayle (vernon.gayle@ed.ac.uk)

This repository hosts a Jupyter Notebook which accompanies the paper above. Details of how to access the required data are provided in the notebook.

New Practices

- ‘literate computing’ (Fernando Perez http://blog.fperez.org/)
  - Weaving narrative directly into live computation
  - Interleaving text with code and results
  - Complete piece achieves goals of communicating result
  - See also (Knuth 1992)

New Practices

- ‘literate computing’ (Fernando Perez http://blog.fperez.org/)
  - Weaving narrative directly into live computation
  - Interleaving text with code and results
  - Complete piece achieves goals of communicating result
  - See also (Knuth 1992)

- ‘pair programming’
  - One player on the ball, one player off the ball

New Practices

- ‘literate computing’ (Fernando Perez http://blog.fperez.org/)
  - Weaving narrative directly into live computation
  - Interleaving text with code and results
  - Complete piece achieves goals of communicating result
  - See also (Knuth 1992)

- ‘pair programming’
  - One player on the ball, one player off the ball

- ‘code peer review’
  - Each author has run the complete workflow independently using a different computer and software set-up
  - Enabled us to undertake an in-depth test of the reproducibility of the work

The Horror of a Retraction…

The authors of a March 2015 *Journal of Health and Social Behavior* (JHSB) study, "In Sickness and in Health? Physical Illness as a Risk Factor for Marital Dissolution in Later Life" (2015, 56(1):59-73), have retracted the article.

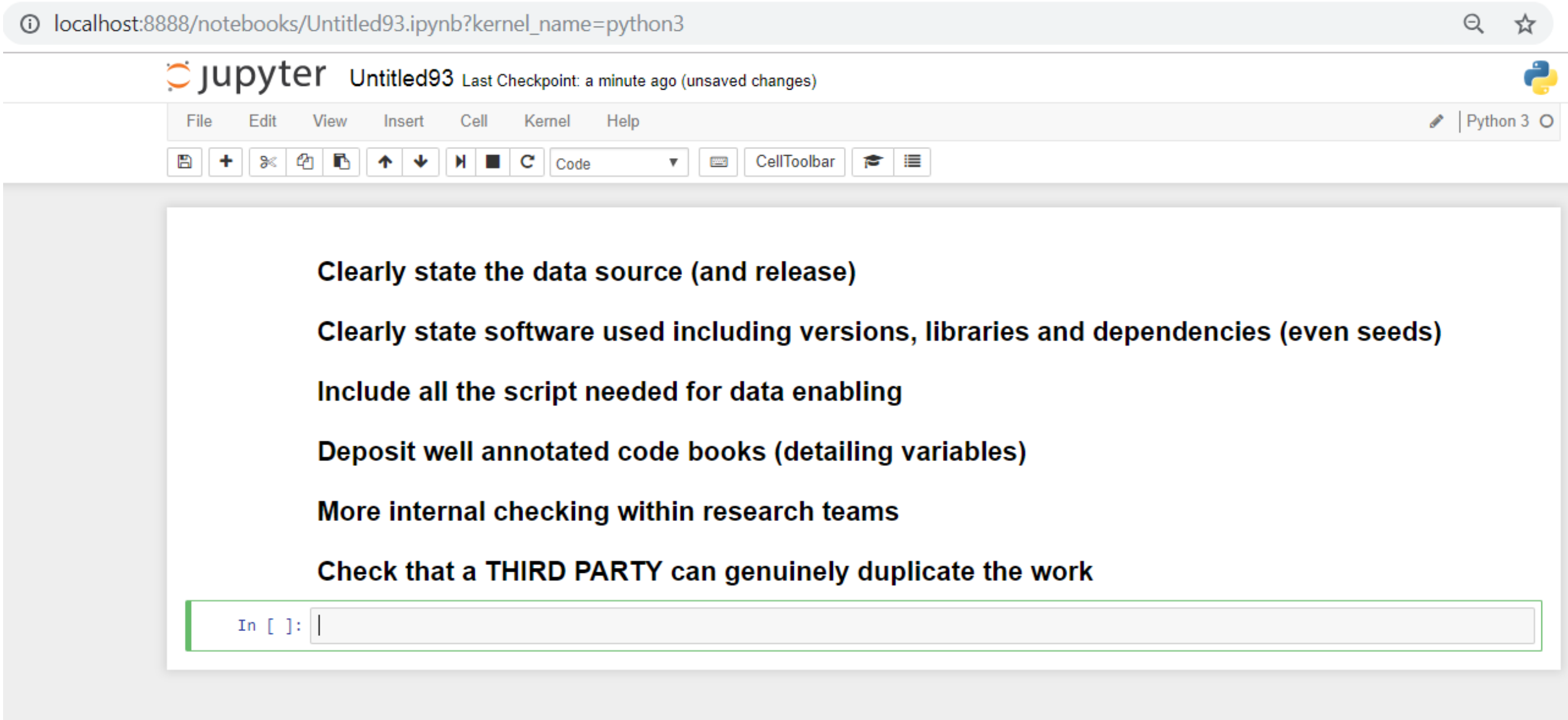
There was a major error in the coding in their dependent variable of marital status. The conclusions of that study should be considered invalid.

A corrected version of the article will appear in the September 2015 issue of JHSB.

http://www.asanet.org/press-center/press-releases/journal-health-and-social-behavior-publish-corrected-version-study

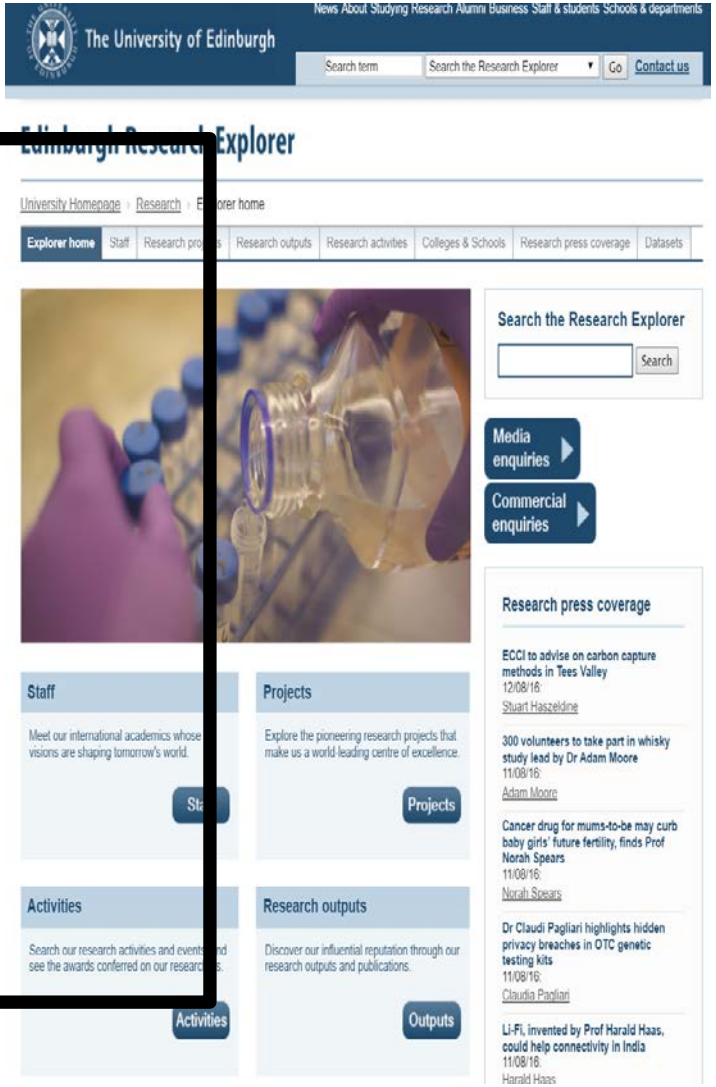
Things we should do immediately

- Deposit annotated **electronic notebooks** that a THIRD PARTY can use to completely duplicate ALL the results included in the published work


localhost:8888/notebooks/Untitled93.ipynb?kernel_name=python3
jupyter
Untitled93
Last Checkpoint: a minute ago (unsaved changes)
File
Edit
View
Insert
Cell
Kernel
Help
Python 3
Code
CellToolbar
Clearly state the data source (and release)
Clearly state software used including versions, libraries and dependencies (even seeds)
Include all the script needed for data enabling
Deposit well annotated code books (detailing variables)
More internal checking within research teams
Check that a THIRD PARTY can genuinely duplicate the work
In [ ]:

Open the BLACK BOX

Routinely sharing research code



Transparent and Reproducible Social Science Research

University of Luxembourg
October 2018

Vernon Gayle
University of Edinburgh

Roxanne Connelly
University of York

RUSSELL GROUP