Social Sciences Research Methods Centre
Replication Workshop

Lent 2014/15
Seminar Room, Department of Sociology
Mondays 5-7pm
Start: January 12, 2015
End: March 9, 2015

Bookings: http://tinyurl.com/malf62q

Instructor:	Nicole Janz (nj248@cam.ac.uk)
TA's:	Aiora Zabala (az296@cam.ac.uk)
	Sergio Martinez Cuesta (scuesta@ebi.ac.uk)
	Sun Rui (rs770@cam.ac.uk)
	Katarzyna Doniec (kd353@cam.ac.uk)
Website:	http://nicolejanz.de/teaching/replication.html
Email-List:	ucam-replication@lists.cam.ac.uk.
Dataverse:	http://dvn.iq.harvard.edu/dvn/dv/CambridgeReplication
Dropbox:	"Replication Workshop"
Blog:	https://politicalsciencereplication.wordpress.com/
Twitter:	@PolSciReplication

Overview
This workshop will introduce students to the process of reproducing published work. Replicating other scholars' work is an essential tool for becoming familiar with methods, learning to select suitable models, and getting a chance to publish early during their academic career. This replication workshop will therefore provide students with a deeper understanding of statistical modeling and professionalism in their field. With the right amount of value added, a replication study can be submitted to a journal, as has been done by several students in the past.

Duration and format
The course runs over 8 sessions with a break in between. Most of the sessions consist of a short lecture and an extensive practical part using the software R. There will be a weekly assignment in which you work on your replication before bringing the results (and problems) to the next session.

Final Project
The main focus of the course is the final project, i.e. a paper in which you have replicated and extend previously published results.

Requirements
- You can do summary statistics, correlation, Chi-square, t-tests and F-tests, multiple linear regression, data handling in R.
- You can commit to at least 6 hours self-directed work in most of the weeks.
- Your MPhil / PhD supervisor supports your participation in the course.
- You have installed R and Rstudio on your laptop.
- You commit to finishing the course and uploading your result to our dataverse.

Preparation for first session
Please bring a copy of the article(s) including the replication data set you would like to work on so we can discuss if it is suitable. A rough guideline:

1) Pick a paper where the data set is available from the Journal's webpage, the author's webpage or a data verse. If you cannot locate its data and code, don't choose it!

2) To increase the probability that your paper will eventually be published, pick an article from a top journal**.**
3) The article should have been published in 2010 or later.
4) Your paper should use methods you can manage or learn.

Schedule

Week 1 **Jan 12, 2015** **5-7pm** *Welcome!*	What is replication? The "reproducibility debate" in statistical research - why replicate? How to pick a suitable paper for replication? Where and how to download data? How to approach the original author? Bring 1-3 papers, including the data, to the first session to pick your project. **Readings:** King, Gary. 1995. Replication, Replication. PS: Political Science and Politics 28: 443–499. http://j.mp/jCyfF1 King, Gary "How to Write a Publishable Paper as a Class Project ", http://gking.harvard.edu/papers Janz, Nicole (under review). Bringing the Gold Standard Into the Class Room: Replication in University Teaching (pdf in dropbox)
Week 2: **Jan 19, 2015** **5-7pm** *Getting started*	How to plan your replication? How to work reproducibly yourself? **Guest:** **Dr Andrew Caines, Department of Theoretical and Applied Linguistics** **"How to make your own work reproducible with Rmarkdown" (30min)** **Readings:** "How to make your work reproducible" https://politicalsciencereplication.wordpress.com/2013/06/04/how-to-make-your-work-reproducible/ Christopher Gandrud, Getting Started with Reproducible Research http://christophergandrud.blogspot.co.uk/2013/07/getting-started-with-reproducible.html There may be additional readings provided by Dr Caines.
Week 3 **Jan 26, 2015** **4.30-6.30** *Practical Session with the TAs*	Practical Session with the TAs: discussing first results; problem solving and coding improvement. **Guest:** **Dr Thilo Klein, Judge Business School, Cambridge** **"Documenting your statistical workflow: How to write an R package" (30min)** **Readings:** No readings for this session; but have all your code and questions ready for the practical part of the session; this is a unique opportunity to get as much help as possible from TAs
Week 4 **Feb 2, 2015** **5-7pm** *Practical Session with the TAs*	Presenting the next set of results and your refined analysis; intensive work on the analysis; this or next week the main tables should be finished; TAs will help in class **Readings:** No readings for this session.

Week 5 **Feb 9, 2015** **5-7pm** *Presenting results; Preparing cross-check*	We will discuss how the cross-check among students works: as an assignment for next week, students will exchange the replication code and replicate each others' work; discussing guidelines for cross-checking and constructive feedback Working with the TAs on your Rcode (so that it is well commented) and a draft report which will be handed over to another student by Thursday **Readings:** Watch this video by Gary King on sharing data: http://tinyurl.com/cpzwbzb
BREAK **Feb 16, 2015** *Self-directed work*	**During this time you will:** - cross-check another student's work and write a memo with constructive feedback - read a published replication paper (we will assign these in the previous week to students) and prepare a presentation of: outline of the paper; how the paper added value - think about how you could add value to your own paper
Week 6 **Feb 23, 2015** **5-7pm** *How to add value and get your replication published*	Reports on Cross Check; Problem solving; How to bring a replication into a publishable format for a journal; how to add value to a replication; steps to advance knowledge: Robustness checks, Sensitivity analysis, Dummy variables, Interactions, Omitted variables, Adjusted and improved models **Prof. Mark Haggard, Department of Psychology** **"What can we learn from the 'Replication Crisis'? "** R part: discuss with your cross-check 'partner' what can be improved; Help in R from the TAs: prepare to add value to your replication
Week 7 **March 2, 2015** **5-7pm** *Finalizing your R code and paper draft*	Bring your Rcode and a draft for your final replication paper; get help from TAs with your 'added value component', e.g. robustness checks, sensitivity analysis, dummy variables, interactions, omitted variables, adjusted and improved models; ideas on how to publish **Guest:** **Yin Wu, Department of Psychology (alumnus of this workshop)** **"How to publish your replication paper" (30min)** **Readings:** Read a published replication paper (we will assign these in the previous week to students) and prepare a presentation of: outline of the paper; how the paper added value Mark Bell and Nicholas Miller, Guest post: How to persuade journals to accept your replication paper http://tinyurl.com/pqh7b2x "Publishing as a grad student: replication helps" http://tinyurl.com/oksyln3
Week 8 **March 9, 2015** **5-7pm** *Wrap up*	Publishing the result in our dataverse http://dvn.iq.harvard.edu/dvn/dv/CambridgeReplication **Bring your final paper draft, Rcode and data to upload in class** Have a look at this webpage:

	http://projects.iq.harvard.edu/psreplication in particular the guidelines which apply to this course: http://projects.iq.harvard.edu/psreplication/guidelines