Age-period-cohort modelling and covariates, with an application to obesity in England 2001-2014: Replication

Zoë Fannon

7 December 2018

1 Can the results in the paper be exactly reproduced?

Yes. It will be necessary to download data from the UK Data Service and the Office for National Statistics, and have my zipped folder "replication". Together these have everything necessary to reproduce the results.

2 What data is used?

The main data used comes from the Health Survey for England. This was accessed through the UK data service. If appropriate permission is received from UKDS, I can also provide the raw data files. These downloads have filenames in the format UKDA-1111-stata8. These should be unzipped in the directory `replication/Data_modified/HSE_downloads`.

A GDP deflator dataset is also used in the robustness checks. This is already in the folder but can also be retrieved from the Office for National Statistics. I use the version as published on 23 Dec 2016, as of 29 November 2018 it could be found at https://www.ons.gov.uk/economy/grossdomesticproductgdp/datasets/ ukquarterlynationalaccountsdatatables/current under 31 March 2017 version.

2.1 All data citations

Office of Population Censuses and Surveys. Social Survey Division, Health Survey for England, 1991-1992 : Combined Data File [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor], 1997. SN: 3238, http://dx.doi.org/10.5255/UKDA-SN-3238-1

Office of Population Censuses and Surveys. Social Survey Division, Health Survey for England, 1993 [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor], 1997. SN: 3316,

http://dx.doi.org/10.5255/UKDA-SN-3316-1

Joint Health Surveys Unit of Social and Community Planning Research and University College London, Health Survey for England, 1994 [computer file]. 4th Edition. Colchester, Essex: UK Data Archive [distributor], March 2001. SN: 3640, http://dx.doi.org/10.5255/UKDA-SN-3640-1

Joint Health Surveys Unit of Social and Community Planning Research and University College London, Health Survey for England, 1995 [computer file]. 4th Edition. Colchester, Essex: UK Data Archive [distributor], April 2010. SN: 3796, http://dx.doi.org/10.5255/UKDA-SN-3796-1

Joint Health Surveys Unit of Social and Community Planning Research and University College London, Health Survey for England, 1996 [computer file]. 4th Edition. Colchester, Essex: UK Data Archive [distributor], April 2010. SN: 3886, http://dx.doi.org/10.5255/UKDA-SN-3886-1

Joint Health Surveys Unit of Social and Community Planning Research and University College London, Health Survey for England, 1997 [computer file]. 3rd Edition. Colchester, Essex: UK Data Archive [distributor], April 2010. SN: 3979, http://dx.doi.org/10.5255/UKDA-SN-3979-1

National Centre for Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 1998 [computer file]. 5th Edition. Colchester, Essex: UK Data Archive [distributor], April 2010. SN: 4150, http://dx.doi.org/10.5255/UKDA-SN-4150-1

National Centre for Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 1999 [computer file]. 4th Edition. Colchester, Essex: UK Data Archive [distributor], April 2010. SN: 4365, http://dx.doi.org/10.5255/UKDA-SN-4365-1

National Centre for Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 2000 [computer file]. 4th Edition. Colchester, Essex: UK Data Archive [distributor], July 2011. SN: 4487, http://dx.doi.org/10.5255/UKDA-SN-4487-1

National Centre for Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 2001 [computer file]. 3rd Edition. Colchester, Essex: UK Data Archive [distributor], April 2010. SN: 4628, http://dx.doi.org/10.5255/UKDA-SN-4628-1

National Centre for Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 2002 [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor], April 2010. SN: 4912.

National Centre for Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 2003 [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor], April 2010. SN: 5098, http://dx.doi.org/10.5255/UKDA-SN-5098-1

National Centre for Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 2004 [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor], April 2010. SN: 5439, http://dx.doi.org/10.5255/UKDA-SN-5439-1

National Centre for Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 2005 [computer file]. 3rd Edition. Colchester, Essex: UK Data Archive [distributor], July 2011. SN: 5675, http://dx.doi.org/10.5255/UKDA-SN-5675-1

National Centre for Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 2006 [computer file]. 4th Edition. Colchester, Essex: UK Data Archive [distributor], July 2011. SN: 5809, http://dx.doi.org/10.5255/UKDA-SN-5809-1

National Centre for Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 2007 [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor], April 2010. SN: 6112, http://dx.doi.org/10.5255/UKDA-SN-6112-1

National Centre for Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 2008 [computer file]. 4th Edition. Colchester, Essex: UK Data Archive [distributor], August 2013. SN: 6397, http://dx.doi.org/10.5255/UKDA-SN-6397-2

National Centre for Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 2009 [computer file]. 3rd Edition. Colchester, Essex: UK Data Archive [distributor], January 2015. SN: 6732, http://dx.doi.org/10.5255/UKDA-SN-6732-2

NatCen Social Research and Royal Free and University College Medical School. Department of Epidemiology and Public Health, Health Survey for England, 2010 [computer file]. 3rd Edition. Colchester, Essex: UK Data Archive [distributor], January 2015. SN: 6986, http://dx.doi.org/10.5255/UKDA-SN-6986-3

NatCen Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 2011 [computer file]. Colchester, Essex: UK Data Archive [distributor], April 2013. SN: 7260,

http://dx.doi.org/10.5255/UKDA-SN-7260-1

NatCen Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 2012 [computer file]. Colchester, Essex: UK Data Archive [distributor], April 2014. SN: 7480, http://dx.doi.org/10.5255/UKDA-SN-7480-1

NatCen Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 2013 [computer file]. Colchester, Essex: UK Data Archive [distributor], January 2015. SN: 7649, http://dx.doi.org/10.5255/UKDA-SN-7649-1

NatCen Social Research, University College London. Department of Epidemiology and Public Health. (2016). Health Survey for England, 2014. [data collection]. UK Data Service. SN: 7919, http://dx.doi.org/10.5255/UKDA-SN-7919-1.

3   How is the data cleaned?

I use the Stata script `gettinghsedata.do` which is based on the Stata code provided in `userguide1991-2009_hse.pdf` from Higgins & Marshall (2012).

3.1   Cleaning citations

StataCorp. 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC . Version 15.1 Revision 15 Oct 2018.

Higgins, V., Marshall, A. (2012). Health Survey for England Time Series Dataset, 1991-2009. [data collection]. UK Data Service. SN: 7025, http://doi.org/10.5255/UKDA-SN-7025-1

4   How is the data analysed?

The data is analysed in R. I use the script `HSE_dataclean.R` to clean the data. This script uses the following packages: `plyr`, `reshape`, `readstata13`, `plotly`, `xtable`, `ggplot2`, `Cairo`. This script also calls the following additional scripts: `revised.functions2.R`; `TS_logit_functions.R`, and `further_functions.R`. The first of these calls the package `apc`.

I use the script `HSE_analysis.R` to analyse the data. In addition to the above-mentioned packages and scripts this script calls the package `tseries`. Robustness checks are performed using `HSE_robustcheck.R`.

4.1   Analysis citations

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. Version 3.5.0. "Joy in Playing"

Hadley Wickham (2011). The Split-Apply-Combine Strategy for Data Analysis. Journal of Statistical Software, 40(1), 1-29. URL http://www.jstatsoft.org/v40/i01/. Package: plyr. Version: 1.8.4

H. Wickham. Reshaping data with the reshape package. Journal of Statistical Software, 21(12), 2007. Package: reshape. Version: 0.8.7.

Jan Marvin Garbuszus and Sebastian Jeworutzki (2018). readstata13: Import 'Stata' Data Files. R package version 0.9.2. https://CRAN.R-project.org/package=readstata13

Carson Sievert, Chris Parmer, Toby Hocking, Scott Chamberlain, Karthik Ram, Marianne Corvellec and Pedro Despouy (2017). plotly: Create Interactive Web Graphics via 'plotly.js'. R package version 4.7.1. https://CRAN.R-project.org/package=plotly

H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2009. Package: ggplot2. Version: 2.2.1

David B. Dahl (2016). xtable: Export Tables to LaTeX or HTML. R package version 1.8-2. https://CRAN.R-project.org/package=xtable

Simon Urbanek and Jeffrey Horner (2015). Cairo: R graphics device using cairo graphics library for creating high-quality bitmap (PNG, JPEG, TIFF), vector (PDF, SVG, PostScript) and display (X11 and Win32) output. R package version 1.5-9. https://CRAN.R-project.org/package=Cairo

Bent Nielsen (2016). apc: Age-Period-Cohort Analysis. R package version 1.3. https://CRAN.R-project.org/package=apc

Adrian Trapletti and Kurt Hornik (2018). tseries: Time Series Analysis and Computational Finance. R package version 0.10-45.

5   Where do the graphs and tables in the paper come from?

The R scripts described above generate a large number of .tex and .eps files. To help the referee, I provide a tex file `check_graphs_tables.tex` which can be run following the script `HSE_analysis.R`. This will collect all of the tables and graphs that appear in the final paper in a single pdf. Note that some of the tables appearing in the final

paper are "tidied" versions of the tables generated in R, but all the results are the same subject to rounding.

The skewness and kurtosis results in table 3 of the paper will not be collected as part of this procedure. They can be found in lines 1276-1311 and 1464-1497 of HSE_analysis.R, i.e. the sections relating to the BMI and log BMI models for women.