CURRICULUM VITAE

Name   Paul Simon Clarke

Qualifications
- 1998   PhD in Social Statistics
  Department of Social Statistics, University of Southampton
- 1993   MSc in Statistics
  Institute of Mathematics and Statistics, University of Kent
- 1992   BSc in Statistics
  Department of Mathematics, University College of Wales, Aberystwyth

Employment History

From Sep 2013: *Professor of Social Statistics*, Institute for Social & Economic Research, University of Essex

Aug 2007 to Aug 2013: *Senior Lecturer*, Centre for Market & Public Organisation, School of Economics, Finance & Management, University of Bristol

Apr 2005 to Jul 2007: *Lecturer* then *Senior Lecturer*, Infectious Disease Epidemiology Unit, London School of Hygiene & Tropical Medicine

Jul 2003 to Mar 2005: *Research Fellow*, Department of Infectious Disease Epidemiology, Imperial College London

May 2000 to Jul 2003: *Research Statistician*, Departments of Epidemiology & Public Health and Statistical Science, UCL

Feb 1999 to May 2000: *Researcher*, Survey Methods Centre, National Centre for Social Research

Aug 1998 to Feb 1999: *Temporary Statistical Officer*, Office for National Statistics

Oct 1993 to Aug 1998: *Research Assistant*, Department of Social Statistics, University of Southampton

Successful Grant Applications

2025-28 Element compositor methods for out-of-distribution machine learning, £500k, UKRI: EPSRC (UKRI 1076), project co-lead with Samothrakis S (PL)

2024-27 The Research Centre for Micro-Social Change (MiSoC) Transition Funding, £1.8M, UKRI: ESRC (ES/S012486/1), project co-lead with Del Bono E (PL), Boda Z, Etheridge B, Grundy E, Kumari M Luthra R, Rabe B, Samothrakis S

2023-26 RESTED - REbalancing circadian rhythm of sleep and heart rate to ease dissociative symptoms, £1.1M, Wellcome Trust, consultant with Gillmeister H (PI), De Fio V, Carbonero S, Poerio G

2023-32 Understanding Society: The UK Household Longitudinal Survey Waves 17-22, £99M, UKRI: ESRC (ES/Y003071/1), co-investigator with Benzeval M (PI), Al Baghal T, etc.

2019-24 The Research Centre for Micro-Social Change (MiSoC), £7.8M, ESRC (ES/S012486/1), co-investigator with Brewer M (PI), Barban N, Bhalotra S, Delavande A, Del Bono E, Harkness S, Grundy E, Luthra R, Rabe B, Samothrakis S

2019-22 The Research Centre for Micro-Social Change (MiSoC) Transition Funding, £1.1M, ESRC (ES/S007288/1), co-investigator with Brewer M (PI), Bhalotra S, Delavande A, Del Bono E, Rabe B

2018-23 Understanding Society: The UK Household Longitudinal Survey Wave 12, £15.7M, ESRC (ES/N00812X/1), co-investigator with Benzeval M (PI), Crossley T, Burton J, Harkness S, Jackle A, Kumari M, Lynn P, Rabe B, Platt L, Payne J

2016-21 Understanding Society: The UK Household Longitudinal Survey Waves 9-11, £32.1M, ESRC (ES/N00812X/1), co-investigator with Benzeval M (PI), Burton J, Jackle A, Kumari M, Lynn P, Pudney S, Rabe B, Saggar S, Platt L, Crossley T, Buck N

2015-8 How can biomarkers and genetics improve our understanding of society and health?, £1.86M, ESRC (ES/M008592/1), co-investigator with Benzeval M (PI), Chandola T, Jones A, Kumari M, Mitchell R, Pudney S

2014-9 The Research Centre for Micro-Social Change (MiSoC): Understanding individual and family behaviours in a new era of uncertainty and change, £5.78M, ESRC Centre (ES/L009153/1) co-investigator with Brewer M (PI), Bhalotra S, Del Bono E, Devalande A, Pudney S, Voas D

2013-6 The use of interactive electronic books in the teaching and application of modern quantitative methods in the social sciences, £820k, co-investigator with Browne WJ (PI), Goldstein H, Moreau L, Steele F

2013-6 Developing methods for understanding mechanism in complex interventions, £540k, MRC Methodology Research Strategic Grant Programme (MR/K006185/1), co-investigator with Landau S (PI), Emsley R, Pickles A, White I, Wykes T

2011-4 Longitudinal effects, multilevel modelling and applications (Lemma 3), ESRC (RES-576-25-0032), £1.4M, co-investigator with Steele F (PI), Browne WJ, Charlton CJ, Goldstein H, Jones K, Leckie G

2010-2 Interrelationships between housing transitions and fertility in Britain and Australia, ESRC (RES-062-23-2265), £600k, co-investigator with Steele F (PI), Haynes M, Ermisch J, Kulu H

2010 Social science and bad behaviours: causes, consequences, solutions, ESRC Festival of Social Science (RES-622-26-0263), £1.8k, lead applicant with Andrews H, Dickson M, Scholder S, Washbrook E

2007-8 Estimating the basic reproduction number from outbreak data, MRC Capacity Building Studentship, co-investigator with Ghani AC (PI)

2004-9 Socio-economic status and pathways, US National Institutes of Health, National Institute on Aging (AG13196), $2M, co-investigator with Marmot MG (PI) et al

2003-4 Pathways between education and health: a causal modelling approach, ESRC (RES-000-22 -0290), £40k, co-investigator with Chandola T (PI), Blane D

1996-7 Nonignorable nonresponse models for categorical survey data, ESRC Research Studentship (R00429614273), £6k

In-Progress

- Baiardi A, Clarke PS, Naghi AA, Polselli A (2026) Double machine learning for static panel data with instrumental variables: new method and applications. https://arxiv.org/abs/2603.20464
- Machlanski D, Samothrakis S, Clarke P (2025) The challenges of hyperparameter tuning for accurate causal effect estimation (revised and resubmitted Sep 2025)
- Steele F, Zhang S, Clarke P (2024) Modelling correlations among grouped random effects in multilevel models with an application to the estimation of household effects on longitudinal health outcomes (slides; manuscript revised & resubmitted Feb 2026).
- Moore JM, Clarke PS (2025) Two new solutions to the zero-weights problem (preparing for resubmission)

Published in Peer-Review Journals

1. Clarke PS, Polselli A (2026). Double machine learning for static panel models with fixed effects. *Econometrics Journal* **29**(1), 69-86
2. Machlanski D, Samothrakis S, Clarke P (2024) Robustness of algorithms for causal structure learning to hyperparameter choice. In: *Third Conference on Causal Learning and Reasoning* PMLR **236**, 703-739. Carnesale Commons, UC Los Angeles: California, USA.
3. Machlanski D, Samothrakis S and Clarke P (2024) Undersmoothing Causal Estimators with Generative Trees. *IEEE Access* **12**, 38562-38574
4. Lynn P, Cabrera Alvarez P and Clarke P (2023) Sample composition and representativeness on Understanding Society. *Fiscal Studies* **44**(4), 341-359

5. Clarke PS, Bao Y (2022) Estimating mode effects from a sequential mixed-modes experiment using structural moment models. *Annals of Applied Statistics* **16**(3), 1563-1585

6. Lynch T, Hempel RJ, Whalley B, Chamba R, Clarke P, et al (2020) Refractory depression - Mechanisms & Efficacy of Radically Open Dialectical Behaviour Therapy (RefraMED): findings of randomised trial on benefits and harms. *British Journal of Psychiatry* **216**(4), 204-212

7. Steele F, Clarke PS, Kuha J (2019). Modeling within-household associations in household panel studies. *Annals of Applied Statistics* **13**(1), 367-392

8. Bao Y, Clarke PS, Smart M, Kumari M (2018) Assessing the robustness of sisVIVE in a Mendelian randomization study to estimate the causal effect of body mass index on income using multiple SNPs from Understanding Society. *Statistics in Medicine* **38**(9), 1529-42

9. Lynch TR, Hempel RJ, Whalley B, Byford S, Chamba R, Clarke P, et al. (2018) Radically open dialectical behaviour therapy for refractory depression: The RefraMED RCT. *Efficacy and Mechanism Evaluation* **5**(7), 1-112

10. Steele F, Clarke P, Leckie, G, Allen J, Johnston D (2017) Multilevel structural equation models for longitudinal data where predictors are measured more frequently than outcomes: an application to the effects of stress on the cognitive function of nurses. *Journal of the Royal Statistical Society* Series A **180**(1), 263-283

11. Hanly M, Clarke P, Steele F (2016). Sequence analysis of call record data: exploring the role of different cost settings. *Journal of the Royal Statistical Society* Series A **179**(3), 793-808.

12. Lynch T, Whalley B, Hempel R, Byford S, Clarke P, et al. (2015). Mechanisms and evaluation of radically open dialectical behaviour therapy (RO-DBT) [REFRAMED]: protocol for randomised trial. *BMJ Open* **5**, e008857

13. Clarke PS, Palmer TM, Windmeijer F (2015). Estimating structural mean models with multiple instrumental variables using the generalised method of moments. *Statistical Science* **30**(1), 96-117

14. Clarke P, Crawford C, Steele F, Vignoles A (2015). Revisiting fixed and random effects models: some considerations for policy-relevant education research. *Education Economics* **23**(3), 259-277

15. Washbrook E, Clarke PS, Steele F (2014). Investigating non-ignorable dropout in panel studies of residential mobility. *Applied Statistics* **63**(2), 239-266

16. Steele F, Clarke P, Washbrook E (2013). Modeling household decisions using longitudinal data from household panel surveys with applications to residential mobility. *Sociological Methodology* **43**, 220-271

17. Szmaragd C, Clarke P, Steele F (2013). Subject specific and population average models for binary longitudinal data: a tutorial. *Longitudinal & Lifecourse Studies* **4**(2), 147-165

18. Clarke PS, Windmeijer F (2012). Instrumental variable estimators for binary outcomes. *Journal of the American Statistical Association* **107**(500), 1638-1652

19. Griffin J, Garske T, Ghani AC, Clarke PS (2011). Estimating the basic reproduction number with incomplete data. *Biostatistics* **12**(2), 303-312

20. Chadeau-Hyam M, Clarke PS, et al. (2010). An application of Hidden Markov models to the French vCJD epidemic. *Applied Statistics* **59**(5), 839-853

21. Clarke PS, Windmeijer F (2010). Identification of causal effects on binary outcomes using structural mean models. *Biostatistics* **11**(4), 756-770

22. Clarke P, Will RG, Ghani AC (2007). Is there the potential for a self-sustaining epidemic of variant Creutzfeld-Jakob disease via blood transfusion in the UK? *Journal of the Royal Society Interface* **4**(15), 675-684

23. Pearson SC, Clarke P (2007). Changes in British men's use of family planning services from 1991 to 2000. *Journal of Family Planning & Reproductive Health Care* **33**(3), 183-187

24. Garske T, Clarke P, Ghani AC (2007). The transmissibility of highly pathogenic avian influenza in commercial poultry in industrialized countries. *PLoS ONE* **2**(4), e349

25. Clarke PS, Ghani AC (2007). A note on parameter estimation for variant Creutzfeldt-Jakob disease epidemic models. *Statistics in Medicine* **26**(3), 546-552

26. Garske T, Ward HJT, Clarke P, Will RG, Ghani AC (2006). Factors determining the potential for transmission of vCJD via surgical instruments. *Journal of the Royal Society Interface* **3**(11), 757-766

27. Pion SDS, Clarke P, et al. (2006). Co-infection with Onchocerca volvulus and Loa loa microfilariae in central Cameroon. *Parasitology* **132**(6), 843-854

28. Chandola T, Clarke P, Morris JN, Blane D (2006). Pathways between education and health: a causal modelling approach. *Journal of the Royal Statistical Society* Series A **169**(2), 337-359

29. Clarke PS (2005). Analysing change using two measures taken under different conditions. *Statistics in Medicine* **24**(22), 3401-3415

30. Clarke PS, Smith PWF (2005). On maximum likelihood estimation for log-linear models with non-ignorable non-response. *Statistics & Probability Letters* **73**(4), 441-448

31. Sacker A, Clarke P, Wiggins RD, Bartley M (2005). The social dynamics of health inequalities: a growth curve analysis of self-assessed health in the British Household Panel Study 1991-2000. *Journal of Epidemiology & Community Health* **59**(6), 495-501

32. Clarke P, Ghani AC (2005). Projections of the future course of the primary vCJD epidemic in the UK: inclusion of subclinical infection and the possibility of wider genetic susceptibility. *Journal of the Royal Society Interface* **2**(2), 19-31

33. Chandola T, Clarke P, Wiggins RD, Bartley M. (2005). Who you are and where you live: setting the context for health using multiple membership multilevel models. *Journal of Epidemiology & Community Health* **59**(2), 170-175

34. Bartley M, Sacker A, Clarke P (2004). Employment conditions, employment status and limiting illness: prospective evidence from the British Household Panel Survey 1991-200. *Journal of Epidemiology & Community Health* **58**(6), 501-506

35. Clarke PS (2004). Causal analysis of individual change using the difference score. *Epidemiology* **15**(4), 414-421

36. Clarke PS, Smith PWF (2004). Interval estimation for log-linear models with one variable subject to non-ignorable non-response. *Journal of the Royal Statistical Society* Series B **66**(2), 357-368

37. Clarke P, Sproston K, Thomas R (2003). An investigation into expectation-led interviewer effects in health surveys. *Social Science & Medicine* **56**(10), 2221-2228

38. Sacker A, Wiggins RD, Clarke P, Bartley M (2003). Making sense of symptom checklists: a latent class approach to the first 9 years of the British Household Panel Survey. *Journal of Public Health Medicine* **25**(3), 215-222

39. Singh-Manoux A, Clarke P, Marmot M (2002). Multiple measures of socioeconomic position and psychosocial health: proximal and distal effects. *International Journal of Epidemiology* **31**(6), 1192-1199

40. Clarke PS, Tate PF (2002). An application of non-ignorable non-response models for gross flows estimation in the British Labour Force Survey. *Australian & New Zealand Journal of Statistics* **44**(4), 413-425

41. Brunner EJ, Hemingway H, Walker BR, Page M, Clarke P, et al. (2002). Adrenocortical, autonomic, and inflammatory causes of the metabolic syndrome: nested case-control study. *Circulation* **106**(21), 2659-2665

42. Clarke PS (2002). On boundary solutions and identifiability in categorical regression with non-ignorable non-response. *Biometrical Journal* **44**(6), 701-717

43. Lynn P, Clarke P (2002). Separating non-contact and refusal bias: evidence from UK national surveys. *The Statistician* **51**(3), 319-333

44. Smith PWF, Skinner CJ, Clarke PS (1999). Allowing for non-ignorable non-response in the analysis of voting intention data. *Applied Statistics* **48**(4), 563-577

45. Clarke PS, Chambers RL (1998). Estimating labour force gross flows from surveys subject to household-level nonignorable nonresponse. *Survey Methodology* **24**(2), 123-129

Book Chapters

46. Clarke P, Hardy R (2007). Methods for handling missing data. In: Pickles AR, Maughan B, Wadsworth MEJ (eds), *Epidemiological Methods in Life Course Research*, 157-179. Oxford: Oxford University Press

47. Lynn P, Clarke P, Martin J, Sturgis P (2002). The effects of extended interviewer efforts on non-response bias. In: Groves RM, Dillman DA, Eltinge JL, Little RJA (eds), *Survey Nonresponse*, 135-148. New York: John Wiley

Commentaries

48. Blane D, Clarke P (2025). Contribution to Heinz W, Blane D, Clarke P, Dannefer D, Mortimer J, Robertson T, Sacker A., Schoon I, Spini D, The Legacy of John Byrnner. *Longitudinal & Life Course Studies* **16**(4), 1-15.

49. Clarke P (2012). Commentary on 'Social class returns to higher education' by Bukodi and Goldthorpe. *Longitudinal & Lifecourse Studies* **3**(2), 285-288

50. Smith PWF, Clarke PS (2005). Contribution to discussion of 'Local model uncertainty and incomplete data bias' by Copas & Eguchi. *Journal of the Royal Statistical Society* Series B **67**(4), 499-500

Other Published Work

51. Moore JM, Clarke PS (2025) Two new solutions to the zero-weights problem. Understanding Society Working Paper 2025-06. Institute for Social & Economic Research, University of Essex, UK

52. Bao Y, Clarke P, Smart M, Kumari M (2018) A comparison of robust methods for Mendelian randomization using multiple genetic variants. ISER Working Paper 2018-08. Institute for Social & Economic Research, University of Essex, UK

53. Clarke PS, Palmer TM, Windmeijer F (2011) Estimating structural mean models with multiple instrumental variables using the generalised method of moments. CMPO Working Paper 11/264. Centre for Market & Public Organisation, University of Bristol, UK

54. Clarke P, Crawford C, Steele F, Vignoles, A (2010). The choice between fixed effects and random effects models: some considerations for educational research. CMPO Working Paper 10/240. Centre for Market & Public Organisation, University of Bristol, UK

55. Clarke P, Windmeijer F (2009). Identification of causal effects on binary outcomes using structural mean models. CMPO Working Paper 09/217. Centre for Market & Public Organisation, University of Bristol, UK

56. Clarke P, Windmeijer F (2009). Instrumental variable estimators for binary outcomes. CMPO Working Paper 09/209 (see also WP 10/239). Centre for Market & Public Organisation, University of Bristol, UK

57. Ghani AC, Clarke P (2006). Variant Creutzfeldt-Jakob disease and the risk from blood transfudsion. *Behind the Medical Headlines* (http://www.behindthemedicalheadlines.com/articles/cjd.shtml)

58. Ferguson N, Ghani A, Clarke P, Donnelly C (2004). Extension of the 2003 OTM risk assesddsment. Report for Food Standards Agency (http://www.food.gov.uk/foodindustry/meat/otmreview)

59. Clarke PS, Smith PWF (2003). Results on point and interval estimation for log-linear models with non-ignorable non-response. S^{3}RI Methodology Working Paper M03/23. Southampton Statistical Sciences Research Unit, University of Southampton

60. Clarke P, Lynn P (2001). Using information on call-backs to correct for non-response bias in face-to-face interview surveys. In: *Proceedings of the Fifth International Conference on Logic and Methodology.* Opladen: Leske and Budrich

61. Lynn P, Clarke P (2000). An experiment with advance letters on the Scottish School Levers Survey. *Survey Methods Newsletter* **20**(2), 8-12. London: National Centre for Social Research

62. Clarke PS, Tate PF (1999). Methodological issues in the production and analysis of longitudinal data from the Labour Force Survey. *GSS Methodology Series No. 17*. London: Office for National Statistics

63. Clarke PS, Tate PF (1999). Weighting versus modelling in adjusting for nonignorable nonresponse in the British Labour Force Survey. In: *Proceedings of the Statistics Canada Symposium 98: Longitudinal Analysis of Complex Surveys*, 51-57. Ottawa: Statistics Canada

Professional Activities & Other Relevant Experience
- Advisory Board, Bloomsbury Research Methods Series (2022-)
- Editorial Board, SAGE Research Methods & Foundations (2018-21)
- Judging Panel, Jon Rasbash Prize (2016-18)
- Research Fellow, Centre for Evaluation & Development, University of Mannheim (since 2016)
- Associate Editor, *Applied Statistics* (2014-7)
- Advisory Board member, CMIST, University of Manchester (2013-8)
- Member, ESRC Responsive Mode Grant Assessment Panel (2013-14)
- Member, Steering Committee, UK Causal Inference Meeting (since 2012)
- Member, RSS Conference Board (2011-2014)
- Member, RSS General Applications Section (2010-11; Secretary 2012-13;chair 2014)
- Member, Society of Longitudinal and Lifecourse Studies (since 2010)
- Affiliated fellow, Centre for Multilevel Modelling, University of Bristol (since 2010)
- Attached member, ESRC International Centre for Life Course Studies (since 2010)
- Academic reviewer, ESRC Peer Review College (since 2010)
- Member, MRC Mental Health Research Network, Methodology Research Group (2009-14)
- Associate Editor, *Journal of the Royal Statistical Society* Series A (2008-11)
- Fellow, Royal Statistical Society (since 1999)
- Consultant to Office for National Statistics on Labour Force Survey non-response (1997–98)

- I have refereed for the following statistics journals: *Journal of the Royal Statistical Society Series A (Statistics in Society)*, *Series C (Applied Statistics)*, and *Series B (Statistical Methodology)*, *Biostatistics*, *Journal of the American Statistical Association* (*Theory & Methods* and *Applications & Case Studies*), *Statistics in Medicine*, *Annals of Statistics*, *Biometrika*, *International Journal of Biostatistics*, *Statistics & Computing*, *Journal of Statistical Software*, *Statistics*, *Journal of Survey Statistics & Methodology*, *Journal of Educational and Behavioral Statistics*, *Journal of the German Statistical Society*, and *Methodology*. I have also refereed for the following other journals: *BMC Medical Research Methodology*, *Longitudinal & Lifecourse Studies*, *Oxford Bulletin of Economics & Statistics*, *Sociological Methodology*, *International Journal of Social Research Methodology*, *Mathematical Biosciences*, *American Journal of Public Health*, *Journal of Epidemiology & Community Health*, *Journal of Gerontology: Social Sciences*, *PLoS One*, *Social Science & Medicine*, *Scientific Reports*, *CNS Drugs*, *British Journal of Psychiatry*, *Preventative Veterinary Medicine*, *Health Education Research*, *Tropical Medicine & International Health*, and *The Lancet* (as a statistical reviewer)