Understanding spending and savigns decisions using financial transaction data

by

Fabian Gunzinger

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Warwick Business School

University of Warwick

September 2023

Contents

List of Tables

List of Figures

Acknowledgements

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. All work presented was carried out in collaboration with co-authors as follows:

Chapter 1 is co-authored with Neil Stewart (Warwick Business School, University of Warwick). Data was provided by Money Dashboard to Neil Stewart. Neil developed the concept of the paper. I collated, analysed the data, and wrote the manuscript with edits by Neil Stewart.

Chapter 2 is co-authored with Neil Stewart (Warwick Business School, University of Warwick). Data was provided by Money Dashboard to Neil Stewart. I developed the concept of the paper, collated and analysed the data, and wrote the manuscript with edits by Neil Stewart.

Chapter 3 is a single-author paper. Data was provided by Money Dashboard to Neil Stewart. I developed the concept of the paper, collated and analysed the data, and wrote the manuscript.

Abstract

This thesis consists of three independent research studies in the fields of statistical and behavioural science. Each study is concerned with modelling complex spatio-temporal decisions recorded in police data. Analysing decisions at a high resolution requires a comprehensive understanding of the social phenomenon and data-generating mechanism, combined with careful modelling choices. Chapter 1 is a novel model of ethnic bias at the officer-level in stop and search. Using a Bayesian hierarchical model, we model officer over-searching against two officer-specific baselines: the crime suspects that the officer encounters and the local patrolling area of the officer. We find that most police officers are biased against Black and Asian people in their search decisions, independently of which baseline we use. Furthermore, we decompose bias against ethnic minority groups into bias due to officer over-searching and over-patrolling. Chapter 2 showcases the use of a spatio-temporal Hawkes-type point process to model the reporting of domestic abuse. Extending existing Hawkes models, we test for the existence of two spillover channels in crime victim reporting. Despite well-documented spillover effects in other human behaviour, we find no evidence to support such effects in the reporting of domestic abuse. Chapter 3 introduces a new, robust statistical inference procedure for discrete outcomes. We propose using the Total Variation Distance together with Bayesian Nonparametric Learning to robustify inference. We show that this procedure pos- sesses a range of desirable theoretical properties. Furthermore, we demonstrate that our method outperforms standard inference both in terms of inference and out-of-sample performance on simulated data. Lastly, we show that robust infer- ence is important for modelling police-recorded incidence of sexual offences where fluctuations in reporting can drastically affect inference. I conclude by discussing the importance of sophisticated statistical approaches to reflect often complicated underlying social phenomenon and the equally com- plex process by which it is recorded in data.

Chapter 1

Introduction

1.1 Theoretical context

The question is important because a large number of adults in the UK and the US do not have enough savings to cover unexpected expenses like car or medical bills: in the UK, 25 percent of adults would be unable to cover an unexpected bill of £300 (Philipps et al. 2021), while in the US, about 30 percent would be unable to cover a $400 bill (Governors of the Federal Reserve System 2022). But while there is a large body of research that studies reasons for why savings are low, little is known about what could help people save more.[1]

- Financial wellbeing is important.

- People in UK and US also don't have enough to cover unexpected outlays. (See reports). Also, Sabat and Gallagher (2019)

- This has important consequences:

  - Short-term: financial well-being (see reports)

  - Long-term (viscious cycle): scarcity hypothesis - makes it harder to focus on important things (plan for retirement, focus on healthy lifestyle,

---

[1]Well-documented behavioural biases that help explain undersaving are, among others, present bias (Laibson 1997, Laibson and Marzilli-Ericson 2019), inertia (Madrian and Shea 2001), over-extrapolation (Choi et al. 2009), and limited self-control and willpower (Thaler and Shefrin 1981, Benhabib and Bisin 2005, Fudenberg and Levine 2006, Loewenstein and O'Donoghue 2004, Gul and Pesendorfer 2001). One danger of viewing low savings mainly as a result of behavioural biases is that while these biases likely do play some role and designing environments and tools to help correct them are thus part of the solution, it is at least conceivable that this is an area where the focus on behaviour-level solutions distracts from an effort to find more effective society-level solutions, a danger inherent in behavioural science research convincingly highlighted in Chater and Loewenstein (2022): if the main problem is that many people are unable to earn enough to save, then the effectiveness of helping them manage their low incomes more effectively pales in comparison with efforts to help them earn more.

support children, ...) and might lead to vicious cycle (less savings lead-
ing to increased risk of financial hardship leading to more stress leading
to less savings...)

– Buffering agains financial hardship (Roll and Despard 2020)

- Spending and savings behaviour is important component - it's not all about
lack of income.

1.2 Objectives and research questions

- I use a single dataset, provided by Money Dashboard, for all three papers.

- The data presented a number of challenges.

- Handling data of this size.

- Secure storage with easy remote access (AWS).

- Consistenc and efficient preprocessing (my process).

- Open science contribution: all work available on Github, code available for
preprocessing from beginning to end.

Motivation: My experience as a researcher in well-known academic and private-
sector institutions has made clear to me over the years that careful, reliable, and
replicable data preprocessing is undervalued in many settings.

The result are plain data errors that overturn results (Reinhard and Rogoff)
and replication issues (see Ariely controversies). Given my experience, I'd think
that all we know is the tip of the iceberg.

In the vast majority of empirical research projects there is no good reason
to not make the code public, even when the data is proprietary. And yet, for
the vast majority of papers, there is no code available to replicate findings and
check precise implementation details, which often matter but are not described
in papers.

All of this harms science - the quality of it and the trust in it.

I have dedicated a lot of time during my PhD to ensure that I can rectify at
least some of those issues.

The code for all my projects is available online and includes scripts that can
be used to easily run the entire analysis.

1.3 Chapter summaries

1.4 Contribution

1.5 Limitations and directions for further research

Chapter 2

Machine-learned behaviour traits

Chapter 3

Spending entropy predicts savings behaviour

Chapter 4

Do Money Dashboard users spend more and save less?

Chapter 5

Conclusion

This thesis consists of three independent research studies on ...

In the first part of this conclusion, I will summarise the three chapters, before discussing challenges...

5.1   Summary

The two applied studies presented in Chapters 1 and 2 are purely exploratory. A natural next step to exploring their generalisability is a pre-registered replication on new data (Van den Akker et al., 2019). We did not pre-register any studies presented here since many analysis decisions were taken conditional on the data observed. It is a well-known problem that such data-dependent decisions can affect the conclusions drawn from the research (Simmons, Nelson, and Simonsohn, 2011; Gelman and Loken, 2014). The studies are thus a useful first step in exploring the questions of ethnic bias and domestic abuse and will require further, pre-registered analyses.

5.2   Discussion

Bibliography

Benhabib, Jess and Alberto Bisin (2005). "Modeling internal commitment mechanisms and self-control: A neuroeconomics approach to consumption–saving decisions". In: *Games and economic Behavior* 52.2, pp. 460–492.

Chater, Nick and George Loewenstein (2022). "The i-frame and the s-frame: How focusing on the individual-level solutions has led behavioral public policy astray". In: *Available at SSRN 4046264*.

Choi, James J, David Laibson, Brigitte C Madrian, and Andrew Metrick (2009). "Reinforcement learning and savings behavior". In: *The Journal of finance* 64.6, pp. 2515–2534.

Fudenberg, Drew and David K Levine (2006). "A dual-self model of impulse control". In: *American economic review* 96.5, pp. 1449–1476.

Governors of the Federal Reserve System, Board of (2022). "Economic Well-Being of U.S. Households in 2021". Tech. rep.

Gul, Faruk and Wolfgang Pesendorfer (2001). "Temptation and self-control". In: *Econometrica* 69.6, pp. 1403–1435.

Laibson, David (1997). "Golden eggs and hyperbolic discounting". In: *The Quarterly Journal of Economics* 112.2, pp. 443–478.

Laibson, David and Keith Marzilli-Ericson (2019). "Intertemporal choice". In: *Handbook of Behavioral Economics* 2.

Loewenstein, George and Ted O'Donoghue (2004). "Animal spirits: Affective and deliberative processes in economic behavior". In: *Available at SSRN 539843*.

Madrian, Brigitte C and Dennis F Shea (2001). "The power of suggestion: Inertia in 401 (k) participation and savings behavior". In: *The Quarterly journal of economics* 116.4, pp. 1149–1187.

Philipps, Jo, Annick Kuipers, and Will Sandbrook (2021). "Supporting emergency savings: early learnings of the employee experience of workplace sidecar savings". Tech. rep.

Roll, Stephen and Mathieu Despard (2020). "Income loss and financial distress during COVID-19: The protective role of liquid assets". In: *Available at SSRN 3733862*.

Sabat, Jorge and Emily Gallagher (2019). "Rules of thumb in household savings decisions: Estimation using threshold regression". In: *Available at SSRN 3455696*.

Thaler, Richard H and Hersh M Shefrin (1981). "An economic theory of self-control". In: *Journal of political Economy* 89.2, pp. 392–406.