# The behavioural economics of domestic abuse

## Unique and new big-data perspective

We bring a data science approach to the study of domestic abuse data. We use data from West Midlands Police (serving X.X million people), Lloyds Banking Group (the UK’s largest retail bank with 22 million customers), and the Crime survey for England and Wales. Our approach is new and ambitious, yet eminently feasible as we already hold the necessary data. We have support from Deputy Chief constable Louisa Rolfe at West Midlands Police, who is also the National Police Chiefs’ Council lead for domestic abuse. Named postdoc Trendl works as a postdoc here at Warwick and is also an employee of Lloyds Banking Group in their behavioural and data science team. We have director level support from Lloyds Banking Group for the project. Applying data at this scale to the study of domestic abuse has never been done before.

We were dismayed by the reviews. Only reviewer XXX appreciated the need for a large-scale quantitative approach in domain dominated by qualitative research. Reviewer 158241287 raised issues about access to data, even though we already have access to all of the data required and stated so in the proposal. Our approach contrasts with the qualitative approach using case studies and interviews. Our findings will complement this qualitative literature. It is from this qualitative background that the reviewers appear to have come—no reviewer had expertise in data science or econometrics. Their criticism is focussed on us as new researchers in their domain and on the feasibility of the work. We mitigate the former risk by working with an advisory board and including coauthors from a domestic abuse background. We mitigate the latter risk by already holding or having access to all of the data needed for the project, with necessary data sharing and other legal agreements signed and in place.

## Research team experience

No one else has or is currently proposing to take a data science view of domestic abuse across multiple, diverse datasets. As 199347314 describes, “most research on domestic abuse in the UK context is qualitative in nature”. Quantitative work in the area has only incorporated a single data source (such as the CSEW, or data from a police force), and does not incorporate external data which can provide additional data on potential causes and environmental risk factors. We will be the first to demonstrate how mass administrative data can be combined from multiple sources and used to address domestic abuse. We have a track record in using such data for social science (see Mullett, T. L., Brown, G. D. A., Fincher, C. L., Kosinski, M., & Stillwell, D. (2019). Individual-Level Analyses of the Impact of Parasite Stress on Personality: Reduced Openness Only for Older Individuals. Personality and Social Psychology Bulletin. THIS IS THE FACEBOOK DATA PAPER, SHOULD WE WORRY THAT ALL PAPERS ARE NEIL’S?

Quispe-Torreblanca, E., & Stewart, N. (2019). Causal peer effects in police misconduct. Nature Human Behaviour. doi: 10.1038/s41562-019-0612-8 Available at <https://rdcu.be/bEAue>

Adams, P., Guttman-Kenney, B., Hayes, L., Hunt, S., Laibson, D., & Stewart, N. (2018). The semblance of success in nudging consumers to pay down credit card debt. Financial Conduct Authority Occasional Papers in Financial Regulation, 45. <https://www.fca.org.uk/publication/occasional-papers/occasional-paper-45.pdf>

Muggleton, Parpart, Newall, Leake, Gathergood, Stewart “Gambling-related harm in mass transaction data” Under review at *Nature*.) For example, in the paper on gambling we use Lloyds Banking Group transaction data from millions of individuals to demonstrate that gambling is associated with increased financial distress, reduced lifestyle and health outcomes, higher rates of future unemployment and disability, and higher mortality rates. This view is unique and the first ever to demonstrate such associations—again complementing the traditional qualitative case-study and small survey based approach. I THINK THE GAMBLING EXAMPLE IS A GOOD EXAMPLE OF WHAT WE WILL DELIVER FOR DOMESTIC ABUSE.)

We will establish an advisory board for the project, comprising a senior police officer from West Midlands Police, a senior member of the WMP Police and Crime Commissioner’s Office, an academic with experience in domestic abuse, Prof Vanessa Munro, a lawyer based here at Warwick with extensive experience of domestic abuse, and a member of the charity sector (NAME CHARITY) and a survivor of domestic abuse (NOMINATED BY THE CHARITY).

1. As the application of multiple, very large datasets to research in domestic abuse is novel, there is no one with the required skills in econometrics and data science who also has a substantial track record in domestic abuse research to conduct this research. Requiring such a background would make it impossible to complete the research, because no such individuals exist. We have acknowledged this from the start, hence our working closely with domestic abuse leads in the police, and why we propose including an advisory panel.

## Theory

1. It is not possible to provide a complete review of the literature in a six page proposal. Not is it reasonable to cite studies that are not directly relevant. We are fully aware of the studies XXXXX and the ESRC funded proposal XXXXX. The latter focusses upon testing the existing DASH framework, and recent outputs have shown that it has little to no explanatory power for identifying at-risk individuals. Hence our proposed research to build upon this by identifying reliable indicative patterns in a data science approach.

We are familiar with the Walby papers described by XXX on the cost of domestic abuse. We agree they are important seminal works. Our proposal builds upon the qualitative findings and analysis of the CSEW, by integrating far more data, from more sources, and using more advanced data analytic techniques.

## With Data Access Secured Already, the Project is Viable

1. We already have access to all relevant data, as we stated in the proposal. We hold the police data and CSEW data here on servers at Warwick. Postdoc Trendl is already working with Lloyds data and is an employee of Lloyds Banking Group as well as Warwick University. Approvals and data sharing agreements are already signed. We stated in the proposal that access to data is already in place and so were dismayed to see the issue raised by Reviewer 158241287.
2. Having Lloyds Banking Group as a partner in the research is a huge asset. 130837269 was concerned that we are only using data from one retail bank, but this criticism is a misunderstanding. Lloyds are the largest retail bank with 22 million customers. It is unnecessary to have data from a second retail bank and would be logistically difficult. By using weighted sampling of the 22 million, we can construct any representative sample and still achieve a sample size far in excess of any ever used to study domestic abuse.
3. We are aware of the issues when using police data and CSEW data. For example, Walby and Towers have found that the CSEW underestimates the true extent of domestic abuse as it caps the maximum number of incidents (whereas domestic abuse is often repeated). We are extremely familiar with the West Midlands Police data, and have worked with WMP on projects in stop and search, domestic abuse, risk assessment in custody, and alcohol and police resources. We are mitigating the issues with each data set by using multiple data sets .

Walby and Towers have found issues with the CSEW - in that it underestimates the true extent of domestic abuse given that it caps the maximum number of incidents (whereas domestic abuse is often repeated). We are aware of this and plan to combine CSEW data with data from other sources.

## The sensitive nature of the research topic

We take the semantic point about describing children who witness male to female intimate partner violence as victims of domestic abuse, noting that the act of witnessing such events is what makes them victims. The user reviewer has made clear their views of the language used to describe such circumstances. The reviewer is clearly reluctant for this component of the research to be completed. We believe this arises from a difference of perspectives: we aim to identify the risk factors for later criminal behaviour, not out of a desire to stereotype or assign blame, but to better identify vulnerable individuals. It is an empirical question, whether such individuals are more likely to commit crime later. If they are, we have a societal responsibility to offer support to help them avoid this, not to simply blame.

The reviewer is also reluctant about the investigation of certain socio-economic factors such as the effect of benefit payment systems, expressing concern about potential working class stereotypes. Their assertion is absolutely correct that domestic abuse exists in all sections of society. But note, the word class appears nowhere in the proposal—our proposed research is not about class. Prior research shows a significant role of financial poverty, with the CSEW showing domestic abuse rates 3.5 times higher in very low income homes compared to middle and high income. This is an uncomfortable fact, but ignoring it does not make it false.

## Specific points

# ################## Unused Text Below Here

%We have studied the currently funded ESRC project INSERT PROJECT NAME AND PI. Our proposal is quite different and compliments this work without duplication.

## Anna comments

### no previous experience in domestic abuse research/no collaboration with experts

not sure what’s the appropriate response to this

### too data-driven/no clear theoretical framework or mention or previous studies/issues about originality as some of the questions have either been examined in the context of other countries or with a qualitative approach

we are combining data from various sources to address key questions that either have never been examined or have only been examined using one approach - fresh perspectice argument, and emphasize the importance of data driven research. The mentioned {r} [ESRC grant] (<https://gtr.ukri.org/projects?ref=ES/M01178X/1>) looks at improving the DASH using “more sophisticated algorithms than a traditional logistic regression framework.” – I can’t find more details on this, but they seem to be only using police data, whereas we are taking a broader view by relying on the very detailed information in the CSEW as well. We are also aware of the US literature on domestic abuse victimisation risk factors (which informed our selection of explanatory variables in the proposed analyses), however, we believe that we still lack the proper evidence base for understanding domestic abuse victimisation in the Uk. Despite of the documented issues with the CSEW, it is an incredibly rich and useful source of information on domestic abuse victimisation, which has not been exploited. Combining insights from the CSEW and police data will complement and validate existing knowledge from qualitative invetsigations.

### issues with the CSEW

#### data quality/data access issues/timescales

#### “worrying” assumptions

### specific points:

young person living at the address – we’ll know this from police records police mis-reporting – comparison of csew and police reports reliability of data, ethnicity – children, only a year of data // different explanatory variables for question 1– random forest not explaiend– causal effects – no domestic abuse journals– no mention of cost–