Article:

Hole, A.R. orcid.org/0000-0002-9413-8101 and Ratcliffe, A. (2025) Difference-in-differences with ordinal data: analysing the impact of terror attacks on racial harassment fears of UK Muslims. Fiscal Studies. ISSN: 0143-5671

https://doi.org/10.1111/1475-5890.70006

DOI: 10.1111/1475-5890.70006

Fiscal Studies
The Journal of Applied Public Economics
IFS Institute for Fiscal Studies

CONTRIBUTED PAPER

Difference-in-differences with ordinal data: analysing the impact of terror attacks on racial harassment fears of UK Muslims

Arne Risa Hole[1] | **Anita Ratcliffe**[2]

[1]Universitat Jaume I

[2]University of Sheffield

Correspondence
Arne Risa Hole, Department of Economics, Universitat Jaume I, Avenida de Vicent Sos Baynat s/n, 12006 Castellón de la Plana, Spain.
Email: hole@uji.es

Anita Ratcliffe suddenly passed away during the late stages of preparing this paper for publication. She was an excellent researcher who was the driving force behind this project. She is greatly missed both as a colleague and a friend.

Submitted: December 2023

Abstract

The method of difference-in-differences is central to public policy analysis, but several challenges arise in applying this method to ordinal outcomes, which are routinely collected in surveys and widely analysed in social sciences. In this paper, we propose a user-friendly estimator to implement the method of difference-in-differences with ordinal outcomes to address these challenges. This estimator quantifies the average treatment effect on the treated in terms of response probabilities and allows an assessment of the distributional impacts of treatment. We use this estimator to analyse fear of racial harassment among Muslims living in a non-Muslim majority country following extremist Islamic terror attacks. Our findings reveal a shift in feeling 'not at all worried' to 'fairly worried' about racial harassment after terror attacks, with little change in feeling 'not very worried' or 'very worried'.

KEYWORDS
Difference-in-differences, ordinal data, racial harassment, terrorism

JEL CLASSIFICATION
C25, I10, I31, J15

1 | INTRODUCTION

The method of difference-in-differences (DD) is widely used to evaluate the average treatment effect on the treated (ATET) but applying this method to ordinal outcomes – such as credit ratings, attitudes or beliefs, and subjective well-being – is challenging. The usual empirical strategy for ordinal outcomes models response probabilities (i.e. the probability of selecting a given response category). However, as these are a non-linear function of the treatment indicator, the common trends assumption required to identify the ATET cannot hold unless there is no systematic variation in response probabilities by treatment group. While this issue, and solution to assume common trends in the latent variable, is well documented for binary outcomes (see, e.g. Blundell et al., 2004; Blundell and Costa Dias, 2009; Lechner, 2011; Puhani, 2012), this solution has not yet been extended for use with ordinal outcomes. As a recent survey reveals, most of the latest developments for the method of DD are for continuous outcomes (Roth et al., 2023). Into this vacuum, empirical research using the method of DD with ordinal data adopts a diverse range of approaches, from linear regression to several different methods of converting an ordinal outcome to a binary outcome (see, e.g. Gruber and Mullainathan, 2005; Gregg, Harkness and Smith, 2009; Brodeur and Connolly, 2013; Leicester and Levell, 2016; Hole and Ratcliffe, 2020; Deal, 2022). However, as recently highlighted in Schröder and Yitzhaki (2017), a concern with applying linear regression to ordinal outcomes is the scope for sign reversals (i.e. the sign of coefficients from linear regression can flip when alternative monotonically increasing values are assigned to response categories), whereas converting an ordinal outcome to a binary outcome may entail arbitrary choices and a loss of information. At best, these limitations curtail the usefulness of vast amounts of routinely collected data in public policy evaluation; at worst, they impede the ability to make meaningful policy recommendations.

In this paper, we extend the solution proposed for the method of DD with binary outcomes to ordinal outcomes (see, e.g. Blundell et al., 2004; Blundell and Costa Dias, 2009; Lechner, 2011; Puhani, 2012) to allow estimation of the ATET in terms of response probabilities.[1] There are three advantages to focusing on response probabilities. First, marginal effects in terms of response probabilities are not susceptible to sign reversals. Secondly, by retaining the full spectrum of information in the ordinal outcome, this approach avoids making arbitrary choices in dichotomising the ordinal outcome. Thirdly, the ATET in terms of response probabilities provides a distributional assessment of the impact of treatment. For policy purposes, it may be important to know if exposure to treatment has the largest impact on those initially worse off, and more generally if there are unequal effects of treatment across response categories. We note that shifting the focus to response probabilities forgoes an understanding of the ATET in terms of the mean of the observed and continuous variable underlying the ordinal outcome (henceforth the underlying variable), which may not be desirable in, for instance, the well-being literature where the focus typically centres on the underlying intensity of feelings. Nevertheless, because individuals with a greater intensity of feelings select higher response categories, response probabilities may be valuable even in this context; see Clark (1997) for a discussion explicitly linking estimated coefficients from ordinal regression analysis of job satisfaction to response probabilities.

In our empirical application, we analyse fears of racial harassment among Muslims living in a non-Muslim majority country following extremist Islamic terror attacks. Specifically, we focus on the first extremist Islamic terror attacks in the UK, defined by an initial attack on 7 July 2005 and another (failed) attack on 21 July 2005. Our goal is to provide practical guidance for practitioners seeking to use parametric methods to implement DD with ordinal data to help unlock the huge potential of these data in public policy analysis. We therefore start by using linear regression – the most widely adopted approach to implement the method of DD with ordinal outcomes – and show that its use leads to sign reversals in our setting. We then estimate the ATET in terms of response probabilities. Our findings reveal a delayed response to the terror attacks insofar as there is little evidence of a treatment effect, however defined, in the short window between 7 and 20 July. We then show that from 21 July onward, there is a large and statistically significant increase in racial harassment fears. Our preferred specification, which allows for heteroscedasticity in the latent index, suggests a 5.3 percentage point reduction in the probability of Muslims responding 'not at all worried' and a 4.7 percentage point increase in 'fairly worried' relative to non-Muslims from 21 July, with little evidence of changes in response categories 'not very worried' and 'very worried'.

[1] Other approaches exist for the method of DD with ordinal data that do not use parametric methods as we do here (see, e.g. Athey and Imbens, 2006; Boes, 2013) but empirical applications of these approaches are extremely limited. Our approach is straightforward and easy to implement but makes stronger assumptions to achieve this goal.

The remainder of this paper is structured as follows. In Section 2, we discuss the related literature. In Section 3, we discuss the research question we seek to address and the data at our disposal. In Section 4, we discuss the methodology. In Section 5, we present our findings, and we conclude in Section 6.

2 | RELATED LITERATURE

This paper bridges various strands of literature across different fields. First, our research relates to an existing methodological literature using the method of DD to estimate an ATET with binary data (Blundell et al., 2004; Blundell and Costa Dias, 2009; Lechner, 2011; Puhani, 2012). In this case, the mean has a natural interpretation as a proportion, and the binary outcome can be modelled as a response probability. Retaining the common trends assumption for this probability is problematic because it is a non-linear function of the treatment indicator, which is not removed by taking differences. Thus, the treatment effect is only identified by assuming there are no systematic differences between treated and control groups. This issue, and its solution to assume common trends at the level of the latent index instead of the response probability, is originally explored in Blundell et al. (2004) and discussed in methodological surveys by Blundell and Costa Dias (2009) and Lechner (2011). Puhani (2012) tackles the same theme from a different vantage point. Building on Ai and Norton (2003), who derive the interaction effect in non-linear probability models as the cross-derivative of the response probability with respect to the interacted variables, Puhani shows that the ATET is the difference between two cross-derivatives (one for the observed and one for the potential outcome), which after some manipulation greatly simplifies its calculation. We build on this existing literature to formalise an expression for the ATET for an ordinal outcome in terms of response probabilities.

Our research also relates to an ongoing debate in the well-being literature where two dominant approaches for analysing ordinal data rely on coefficients from linear and ordinal regression methods. Linear regression is attractive due to its simplicity, and is operationalised by assigning numeric values to response categories (i.e. 1 to the lowest response category and increasing by 1 for each subsequent response category) and then analysing the mean of the ordered outcome. This mean can be imbued with meaning by assuming linear response scale use (i.e. each adjacent and increasing response category spans an equal distance of the underlying variable, where the latter corresponds to an intensity of feelings in this instance). However, because ordinal variables only provide information on the ranking of response categories, other approaches to assign numeric values to response categories (i.e. other labelling schemes) exist. For example, response scale use may be concave (i.e. each response category spans an increasingly smaller distance of the underlying variable) or convex (i.e. each response category spans an increasingly larger distance of the underlying variable). Schröder and Yitzhaki (2017) show that the sign of coefficients from linear regression can be reversed with alternative rank-preserving and monotonic labelling schemes that differ in the convexity/concavity of the assumed response scale use. This is problematic because, in principle, these alternative labelling schemes are equally valid. Sign reversals do not occur in every instance, but raise serious questions about research findings, and associated policy advice, based on linear regression applied to ordinal outcomes where they do occur. In an important and insightful contribution, Kaiser and Vendrik (2023) demonstrate that sign reversals occur across different labelling schemes if there are heterogeneous effects of a variable across response categories (for example, if increasing the variable X generally involves movement from lower to higher response categories but also reduces responses in the highest response category). They propose a useful test to investigate the scope for sign reversals in linear regression; see also Bloem (2022) and Bloem and Oswald (2022) for various other tests/sensitivity analysis. They further emphasise that coefficients from linear regression can be informative of marginal effects on the mean of the underlying variable when also assuming the mean of that variable within each response category is independent of the variable X. Hence, linear regression can be a powerful tool for estimating the ATET on the mean of the ordered outcome, and also the underlying variable, if certain conditions are satisfied.

Bond and Lang (2019) make a related critique regarding the use of coefficients from the latent index obtained from ordinal regression to infer marginal effects on the mean of the underlying variable. Specifically, they argue that alternative transformations of the latent index, which differ in assumed response scale use, provide equally valid characterisations of the underlying variable. However, these transformations can reverse the sign of marginal effects on the mean of the underlying variable in the presence of heteroscedasticity. Chen et al. (2022) suggest researchers using parametric methods to analyse the underlying variable instead focus on the median, which is invariant to all monotonic transformations of the latent index. From this vantage point, coefficients from the latent index remain relevant as the median coincides with the mean due to the symmetry of normal/logistic distributions.

We emphasise that marginal effects in terms of response probabilities are neither reversed by transformations to the labelling scheme nor the latent index. Rank-preserving monotonic transformations of the labelling scheme leave the latent index and threshold parameters unaltered and therefore have no bearing on response probabilities. Rank-preserving monotonic transformations of the latent index re-scale the threshold parameters to leave the mass of probabilities across response categories unchanged. As these marginal effects do not change sign regardless of the assumed response scale use, greater confidence can be placed on research findings and policy advice stemming from the analysis of response probabilities. However, to operationalise the ATET in terms of response probabilities, we do need to assume a linear latent index to satisfy the common trends assumption. This means that, in the specific context of estimating the ATET, our approach overcomes a key issue of affecting linear regression but remains susceptible to criticisms related to non-linear transformations of the latent index. Our interest lies in calculating changes in response probabilities, using the latent index as a vehicle to calculate these probabilities, as opposed to making inferences about the underlying variable though.

Finally, this research relates to an empirical literature on the impact of extremist Islamic terrorist attacks on the outcomes of Muslims living in non-Muslim majority countries. This literature considers a range of economic outcomes, such as wages and employment prospects, but also increasingly social outcomes, such as health, well-being and assimilation (see, e.g. Åslund and Rooth, 2005; Dávila and Mora, 2005; Kaushal, Kaestner and Reimers, 2007; Johnston and Lordan, 2012; Gould and Klor, 2016; Zorlu and Frijters, 2019; Hole and Ratcliffe, 2020). At the heart of these papers lies the idea that terror attacks increase prejudice towards Muslims, with spikes in hate crimes (Hanes and Machin, 2014; Gould and Klor, 2016) as well as shifts in attitudes towards immigrants (Åslund and Rooth, 2005) supporting this point of view. Few papers, however, consider changes in discrimination from the perspective of Muslims, with notable exceptions including Elsayed and de Grip (2018), Hole and Ratcliffe (2020) and Giani and Merlino (2021). We focus on fears of racial harassment, which have not been considered to date, and which have repercussions for access to, and enjoyment of, public spaces, as well as mental and physical health. We show that fears of racial harassment increase among Muslims relative to non-Muslims after terror attacks. By harnessing the ATET in terms of response probabilities and documenting an unequal impact of treatment across response categories, we also build a more detailed picture of how circumstances changed for Muslims.

3 | RESEARCH QUESTION AND DATA

Our aim is to examine the outcomes of Muslims living in a non-Muslim majority country following extremist Islamic terror attacks. We focus on the London 2005 terror attacks, marking a watershed event for race relations, as the first terror attacks on British soil by British Muslims. The first attack occurred on 7 July, with deadly bombings in multiple locations, followed by a second attack, which failed in its aim to bomb multiple locations, just two weeks later on 21 July. We exploit a 'Fear of Crime' module, first introduced in 2005 as part of the Citizenship Survey, which surveys individuals aged 16+ living in England and Wales (Home Office. Communities Group, National Centre for Social Research, 2006). This module asks respondents 'How worried are you about being subject to a physical attack because of your skin colour, ethnic origin or religion?' with response categories 'not at all worried', 'not very worried', 'fairly worried' and 'very worried'.[2] We assign a value of 1 to the lowest response category 'not at all worried', increasing by 1 for each subsequent response category, with response category 'very worried' assigned a value of 4. Approximately 14,000 individuals (comprising a core sample of approximately 10,000 and an ethnic minority boost sample of approximately 4,000) were interviewed between 8 March and 30 September 2005,[3] and we have special access to interview dates. We control for a wide range of individual circumstances, including experiences of discrimination, as well as local area characteristics. Our final sample, after excluding individuals with missing information (in particular, approximately 1,900 individuals aged 70+ are not asked about their educational qualifications), comprises just under 11,200 individuals.

We follow the strategy adopted in Hole and Ratcliffe (2020) to compare the change in racial harassment fears among Muslims (i.e. treated group) to non-Muslims (i.e. untreated group) pre- and post-treatment, where the post-treatment period is from 7 July onward. As we have access to the date of interview, we further split the post-treatment period into two windows 7–20 July (i.e. the day of the first attack until the day before the second attack) and 21+ July onwards (i.e. the day of the second attack onwards). The second attack marked the end of any hope that the first attack was an isolated incident and potentially differed in its implications for Muslims. Summary statistics for the dependent and control variables are available by treated and control group, and further disaggregated by pre- and post-treatment period, in Tables B1 and B2 in the online Appendix. These show that Muslims differ from non-Muslims across various dimensions; they are typically younger and poorer, and more likely to live in more-deprived and densely populated areas. However, the characteristics of Muslims interviewed in pre- and post-treatment periods are generally quite similar, although a larger proportion of Muslims interviewed 7–20 July are of mixed ethnic heritage, work in professional roles and live in London.

Figure 1 plots the distribution of racial harassment fears across different windows in time for Muslims and non-Muslims. Interestingly, for Muslims, there are mixed reactions in the period 7–20 July, where Muslims became less likely to report 'not at all worried' and much more likely to report 'not very worried' but also less likely to report 'very worried'. This heterogeneity in the effect of exposure to treatment across response categories previews our finding below that sign reversals in coefficients from linear regression occur in our setting. From 21 July onwards, there is a clear shift from 'not at all worried' to 'fairly worried'.

4 | METHODOLOGY

As our goal is to provide practical guidance on implementing the method of DD with ordinal data using parametric methods, we first apply linear regression given its popularity in empirical research. The ATET for the mean of the ordered outcome using linear regression is obtained via a standard DD specification,

$$Y_i = \alpha_1 D_i + \alpha_2 T_i + \alpha_3 D_i T_i + x_i' \gamma + u_i, \tag{1}$$

where Y_i is the ordered outcome for individual i, x_i is a vector of control variables and u_i is a random error term. D_i is equal to one if individual i is in the treated group and zero otherwise, while T_i is equal to one if individual i is observed in the post-treatment period and zero otherwise. The coefficient α_3 provides an estimate of the ATET. As noted earlier, this may also provide an ATET for the mean

[2] Fewer than 1 per cent of respondents refuse to answer the question or respond with 'don't know' and we drop these individuals.

[3] In practice, a handful of interviews take place in the first week of October.

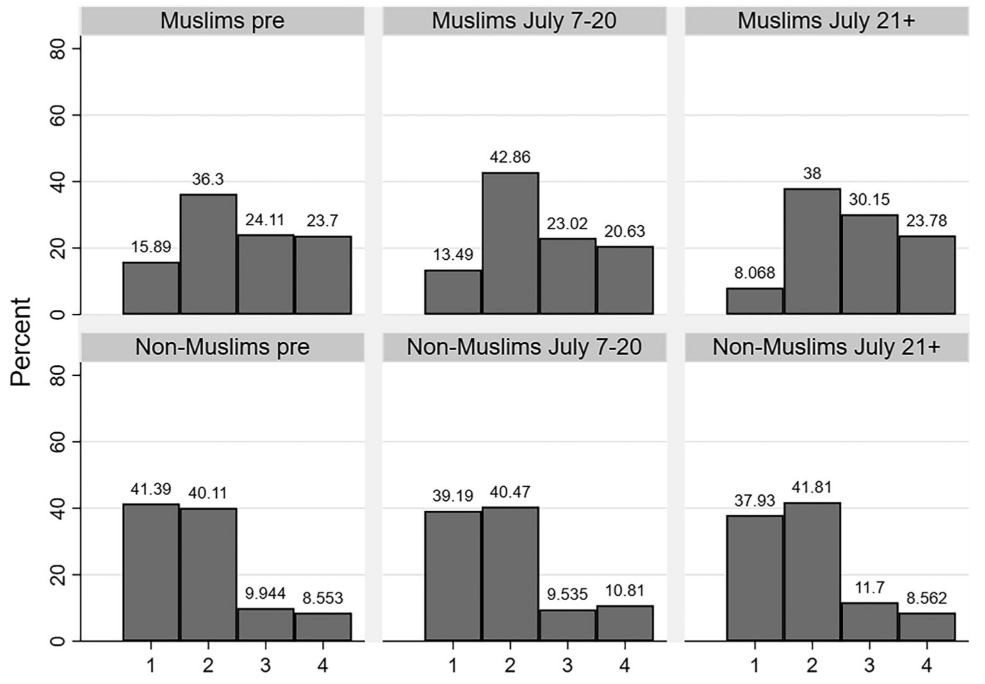

FIGURE 1 Distribution of the dependent variable by Muslims and non-Muslims, before and after terror attacks
Note: 1 corresponds to response category 'not at all worried', 2 to 'not very worried', 3 to 'fairly worried' and 4 to 'very worried'.

of the underlying variable under specific circumstances (Kaiser and Vendrik, 2023). However, a key issue with applying linear regression to ordinal data is the scope for sign reversals (Schröder and Yitzhaki, 2017).

The ATET in terms of a latent index is

$$Y_i^* = \beta_1 D_i + \beta_2 T_i + \beta_3 D_i T_i + x_i' \gamma + \varepsilon_i, \tag{2}$$

where Y_i^* is a latent variable and ε_i is a random error term, which is assumed to be independent and identically distributed standard normal. As discussed above, the coefficient β_3 can reliably provide an estimate of the ATET in terms of the median of the underlying variable (Chen et al., 2022), though we are not aware of empirical applications taking this approach. However, this coefficient should not be used to make inferences about changes in its mean as these can be reversed under some conditions (Bond and Lang, 2019).

We expand the existing toolkit for implementing the method of DD with ordinal data by extending the solution proposed for the method of DD with binary outcomes to ordinal outcomes (see, e.g. Blundell et al., 2004; Blundell and Costa Dias, 2009; Lechner, 2011; Puhani, 2012; Wooldridge, 2023). A full exposition of this approach based on potential outcomes and detailing key assumptions can be found in online Appendix A. Briefly, we assume that each individual has two potential outcomes, and assignment to treatment determines which of these potential outcomes is realised. In this context, the potential outcomes are the response categories, with Y_i^1 denoting the potential outcome with treatment and Y_i^0 the potential outcome without treatment. We assume that there is some underlying unobserved potential latent index that drives these potential outcomes. Thus, each individual has two potential latent indices, Y_i^{1*} and Y_i^{0*}, similarly linked to treatment states. As in the case of linear DD (see, e.g. Roth et al., 2023), two key assumptions are required to identify the ATET: common trends and no anticipatory effects. The difference between our case and the linear case is that we assume, as in Lechner (2011), that these assumptions apply to the latent index. Further assuming that the potential latent indices are related to the potential outcomes via an ordered probit model, we show that the estimator for the ATET in terms of response probabilities is given by

$$\widehat{ATET_{P_k}} = \frac{1}{N^1}\sum_{i=1}^{N} D_i T_i \left\{ \left[G(\hat{\mu}_{k+1} - \hat{\beta}_1 - \hat{\beta}_2 - \hat{\beta}_3 - x_i'\hat{\gamma}) - G(\hat{\mu}_k - \hat{\beta}_1 - \hat{\beta}_2 - \hat{\beta}_3 - x_i'\hat{\gamma}) \right] - \left[G(\hat{\mu}_{k+1} - \hat{\beta}_1 - \hat{\beta}_2 - x_i'\hat{\gamma}) - G(\hat{\mu}_k - \hat{\beta}_1 - \hat{\beta}_2 - x_i'\hat{\gamma}) \right] \right\}, \tag{3}$$

where k denotes the response categories ranging from 1 to K, μ_k are the threshold parameters, $G(\cdot)$ is the standard normal cumulative distribution function (CDF) and N^1 is the number of individuals in the treated group observed post-treatment ($N^1 = \sum_{i=1}^{N} D_i T_i$). As emphasised earlier, the ATET in terms of response probabilities is not susceptible to sign reversals. This estimator assumes that the error term in the latent index is homoscedastic, which may be overly restrictive. It is possible to allow for heteroscedasticity by assuming instead that ε_i is normally distributed with mean zero and variance equal to $\exp(w_i'\tau)^2$, where w_i is a vector of variables that may include D_i, T_i and $D_i T_i$. With this modification, the estimator becomes[4]

$$\widehat{ATET_{P_k}} = \frac{1}{N^1}\sum_{i=1}^{N} D_i T_i \left\{ \left[G\left(\frac{\hat{\mu}_{k+1} - \hat{\beta}_1 - \hat{\beta}_2 - \hat{\beta}_3 - x_i'\hat{\gamma}}{\exp(w_i'\hat{\tau})} \right) - G\left(\frac{\hat{\mu}_k - \hat{\beta}_1 - \hat{\beta}_2 - \hat{\beta}_3 - x_i'\hat{\gamma}}{\exp(w_i'\hat{\tau})} \right) \right] - \left[G\left(\frac{\hat{\mu}_{k+1} - \hat{\beta}_1 - \hat{\beta}_2 - x_i'\hat{\gamma}}{\exp(w_i'\hat{\tau})} \right) - G\left(\frac{\hat{\mu}_k - \hat{\beta}_1 - \hat{\beta}_2 - x_i'\hat{\gamma}}{\exp(w_i'\hat{\tau})} \right) \right] \right\}. \tag{4}$$

It should be noted, however, that if the source of the heteroscedasticity is unknown, trying to account for it may not always be beneficial (Keele and Park, 2006). The estimators proposed in equations (3) and (4) can be calculated using a routine post-estimation command following ordered probit or heteroscedastic ordered probit estimation (see online Appendix C for the Stata code).

5 | RESULTS

Table 1 presents the ATET from a linear regression (see equation (1)) for the post-treatment period (i.e. 7 July onwards) and then separately for 7–20 July and 21 July onwards. Full regression results are available in Table B3 in the online Appendix and suggest that being female, having an ethnic minority or non-native background, having experienced discrimination in other contexts and living in highly densely populated areas are associated with greater fears of racial harassment. However, economic advantage, living in an area with mostly people of the same ethnic background and living outside London are associated with lower fears. In terms of the treatment effect, there is suggestive evidence that fears of racial harassment increased for Muslims relative to non-Muslims after the terror attacks (column 1), but the effect is only statistically significant during the window after the second attack (column 3), where the change in fears represents a 5 per cent increase in its pre-treatment mean (see

[4] In parallel work to ours, Yamauchi (2020) derives an alternative version of this estimator by assuming that the shift in the distribution of the latent index over time is constant in the treatment and the control groups in the absence of treatment. This is equivalent to the assumption that underlies the changes-in-changes estimator proposed by Athey and Imbens (2006), with the important distinction that it applies to the distribution of the latent index instead of the outcome itself. We instead maintain the assumption of common trends in the mean of the latent index and assume a specific functional form for the variance of ε_i to identify the error variance under the counterfactual.

TABLE 1 The ATET using linear regression across different post-treatment windows

	7+ July (1)	7–20 July (2)	21+ July (3)
Treated × post	0.08	−0.07	0.12**
	(0.06)	(0.10)	(0.06)
N	11,186	7,398	10,200

Note: See Table B3 for full regression results. Robust standard errors. Significance levels are shown as ***$p < 0.01$, **$p < 0.05$ and *$p < 0.1$.

TABLE 2 Testing for sign reversals in the ATET

	≤ Response category 1	≤ Response category 2	≤ Response category 3
Panel A. 7+ July			
Treated × post	−0.054***	−0.035	0.006
	(0.020)	(0.028)	(0.024)
N	11,186	11,186	11,186
Panel B. 7–20 July			
Treated × post	−0.023	0.046	0.047
	(0.038)	(0.049)	(0.041)
N	7,398	7,398	7,398
Panel C. 21+ July			
Treated × post	−0.061***	−0.056*	−0.005
	(0.021)	(0.030)	(0.026)
N	10,200	10,200	10,200

Note: Response category 1 is 'not at all worried', response category 2 is 'not very worried' and response category 3 is 'fairly worried'. Each column presents results from a linear probability model where the dependent variable is equal to one if the respondent selects a response category less than or equal to the specified response category (i.e. response category 2). Includes full set of control variables. Robust standard errors. Significance levels are shown as ***$p < 0.01$, **$p < 0.05$ and *$p < 0.1$.

Table B1). In the short window between the first and second attack (column 2), the mixed responses observed in Figure 1 result in a reduction in fears overall, though this effect is not statistically significant.[5]

A limitation of using linear regression with ordinal outcomes is the scope for sign reversals in coefficients (Schröder and Yitzhaki, 2017). We use the test proposed by Kaiser and Vendrik (2023) to assess if sign reversals emerge in our setting. They suggest creating a series of dummy variables, each equal to one if an individual's response is less than or equal to a given response category. For example, in our setting, the first dummy variable is equal to one if an individual responds 'not at all worried' and zero otherwise, the second dummy variable is equal to one if an individual responds with either 'not at all worried' or 'not very worried' and zero otherwise, and so on. Coefficients from regressions on these dummy variables should not switch sign, as this would indicate that the (conditional) CDFs cross, which violates first-order stochastic dominance. As Table 2 shows, the coefficient switches sign in Panel A (after the terror attacks) and Panel B (between the first and second attack), but not in Panel C (after the second attack). This test reveals that alternative monotonically increasing values assigned to response categories can flip the sign of the ATET with linear regression. This is perhaps not surprising as Muslims reduce reporting both 'not at all worried' and 'very worried' in favour of

[5] As Muslims interviewed in the short window 7–20 July are more likely to live in London, we repeat this analysis for the (smaller) London sample. We find positive a positive ATET across all post-treatment windows but the effect is largest and only statistically significant in the window 21+ July (results available on request).

TABLE 3 ATET in terms of the response probabilities (ordered probit)

	7+ July	7–20 July	21+ July
$ATET_{P_1}$: 'not at all worried'	−0.018	0.015	−0.024**
	(0.011)	(0.026)	(0.011)
$ATET_{P_2}$: 'not very worried'	−0.019	0.010	−0.029*
	(0.013)	(0.015)	(0.015)
$ATET_{P_3}$: 'fairly worried'	0.006*	−0.005	0.007**
	(0.003)	(0.010)	(0.003)
$ATET_{P_4}$: 'very worried'	0.031	−0.020	0.046**
	(0.021)	(0.032)	(0.023)
N	11,186	7,398	10,200

Note: Robust standard errors calculated using the delta method. Significance levels are shown as $^{***}p < 0.01$, $^{**}p < 0.05$ and $^{*}p < 0.1$. See Table B4 for full regression results from the ordered probit model used to estimate the ATET in terms of response probabilities.

'not very worried' in the short window after the bombings (see Figure 1). While limiting the usefulness of the estimator in equation (1) for our purposes, this will not be the case in every setting, and it may also be that sensitivity analysis reveals sign reversals require extreme response scale use (see Bloem, 2022; Kaiser and Vendrik, 2023).

Table 3 reports the ATET in terms of response probabilities as described in equation (3). Results from ordered probit estimation of the latent index are also presented in Table B4 in the online Appendix. Table 3 also shows a delayed reaction, with little impact in the period immediately after the first attack, and a non-trivial impact in the period following the second attack. However, these response probabilities also indicate where changes are taking place across response categories. For example, there are falls of 2.4 and 2.9 percentage points in the probability of responding 'not at all worried' and 'not very worried', respectively, a 0.7 percentage point increase in the probability of responding 'fairly worried' and a 4.6 percentage point increase in the probability of responding 'very worried' after the second attack. The magnitude of the change in probabilities for 'not at all worried' or 'not very worried' is also statistically different to those for 'fairly worried' or 'very worried'. Hence, these results indicate a shift from low to substantial racial harassment fears. However, Figure 1 suggests shifts occur from 'not at all worried' to 'fairly worried' in the period following the second attack, which may imply a more flexible specification that allows for heteroscedasticity, such as that in equation (4), provides a better fit for the data.

Table 4 presents the ATET in terms of response probabilities allowing all variables to influence the mean and variance of the latent index (see equation (4)), with results from the accompanying heteroscedastic ordered probit model presented in Table B5. Two points are worth noting: (1) the likelihood ratio test rejects the hypothesis of homoscedasticity, making this our preferred specification; (2) the effect of the terror attacks on racial harassment fears arises by reducing variability in the latent index after the terror attacks rather than through an effect on its mean per se. The coefficient on treated × post for the mean of the latent index remains positive in the final column of Table B5 but is somewhat reduced in magnitude (0.10 instead of 0.14) with a p-value of 0.14. The coefficient on treated × post in the variance component is −0.2 with a p-value of 0.004. Thus, after the second attack, the responses of Muslims conditional on control variables become noticeably less noisy. So, whereas prior to the terror attacks, responses reflect a variety of unobserved idiosyncratic influences, the second terror attack seems to have brought into sharper focus the role of individual characteristics and circumstances in shaping racial harassment fears. The ATET in terms of response probabilities, presented in Table 4, reflects a combination of these effects. In line with Figure 1, they show a 5.3 percentage point reduction in the probability of responding 'not at all worried' and a 4.7 percentage point increase for 'fairly worried' after the second attack.

TABLE 4 ATET in terms of the response probabilities (heteroscedastic ordered probit)

	7+ July	**7–20 July**	**21+ July**
$ATET_{P_1}$: 'not at all worried'	-0.049^{***}	-0.024	-0.053^{***}
	(0.015)	(0.033)	(0.015)
$ATET_{P_2}$: 'not very worried'	0.023	0.067^{*}	0.007
	(0.022)	(0.040)	(0.024)
$ATET_{P_3}$: 'fairly worried'	0.041^{***}	0.017	0.047^{***}
	(0.014)	(0.021)	(0.016)
$ATET_{P_4}$: 'very worried'	-0.016	-0.061	-0.000
	(0.025)	(0.039)	(0.028)
N	11,186	7,398	10,200

Note: Robust standard errors calculated using the delta method. Significance levels are shown as $^{***}p < 0.01$, $^{**}p < 0.05$ and $^{*}p < 0.1$. See Table B5 for full regression results from the heteroscedastic ordered probit model used to estimate the ATET in terms of response probabilities.

TABLE 5 ATET in terms of the response probabilities by gender (heteroscedastic ordered probit)

	21+ July	**21+ July**	**21+ July**
	All	**Men**	**Women**
$ATET_{P_1}$: 'not at all worried'	-0.053^{***}	-0.060^{***}	-0.050^{**}
	(0.015)	(0.022)	(0.020)
$ATET_{P_2}$: 'not very worried'	0.007	0.010	0.006
	(0.024)	(0.036)	(0.034)
$ATET_{P_3}$: 'fairly worried'	0.047^{***}	0.053^{**}	0.045^{**}
	(0.016)	(0.024)	(0.022)
$ATET_{P_4}$: 'very worried'	-0.000	-0.003	-0.002
	(0.028)	(0.039)	(0.042)
N	10,200	4,599	5,601

Note: Robust standard errors calculated using the delta method. Significance levels are shown as $^{***}p < 0.01$, $^{**}p < 0.05$ and $^{*}p < 0.1$.

Estimating the ATET separately by gender does not suggest that there are any substantial differences in the impact of the bombings on fears of racial harassment between male and female Muslims. Table 5 shows a 6 percentage point reduction in the probability of responding 'not at all worried' and a 5.3 percentage point increase for 'fairly worried' after the second attack for men, while the corresponding figures for women are 5 and 4.5 percentage points, respectively. This contrasts with the findings in Hole and Ratcliffe (2020), where we found marked gender differences among Muslim adolescents in terms of the impact of the London bombings on their happiness and expectations of facing discrimination in the labour market.[6]

5.1 | Assessing common trends in the pre-treatment period

Although we have a single year of data, because we have access to the interview date, we are able to consider if common trends hold in a pre-treatment period. While this does not guarantee that common trends hold in the post-treatment period – see Kahn-Lang and Lang (2020) for an interesting discussion – it can provide reassurance. The assumption of common trends may be sensitive to functional form

[6] While we found evidence of a decline in happiness and a rise in expectations of facing discrimination in the labour market among Muslim teenage girls after the bombings, no corresponding effects were found for Muslim teenage boys.

TABLE 6 Assessing common trends in ordered outcome/latent index pre-treatment

	26+ May	26 May–8 June	9 June–6 July
Panel A. OLS			
Treated × post	−0.09	−0.10	−0.08
	(0.08)	(0.11)	(0.09)
Panel B. Ordered probit			
Treated × post	−0.10	−0.11	−0.10
	(0.09)	(0.13)	(0.10)
Panel C. Heteroscedastic ordered probit			
Mean function: Treated × post	−0.13	−0.14	−0.13
	(0.10)	(0.13)	(0.12)
Variance function: Treated × post	−0.03	−0.17	0.05
	(0.09)	(0.13)	(0.10)
N	6,412	4,367	5,559

Note: Includes full set of control variables. Robust standard errors.

(see, e.g. Meyer, 1995) and different approaches for analysing the ATET with ordinal data impose different common trends assumptions. For example, the ATET for the mean of the ordered outcome assumes common trends in this mean, and the ATET in terms of response probabilities assumes common trends in the latent index. Where practical, recourse to theory or context-specific knowledge should help guide the choice of functional form (Kahn-Lang and Lang, 2020; Roth and Sant'Anna, 2023) – but in the current context, there is no obvious frontrunner. In general, one advantage to assuming common trends for the mean of the ordered outcome, as with linear regression, is the availability of tools to assess the credibility of the common trends with continuous outcomes (see Rambachan and Roth, 2023). In a test somewhat reminiscent of Kaiser and Vendrik (2023), Roth and Sant'Anna (2023) also propose checking common trends for the CDF of the untreated potential outcomes as a means to assess whether the assumption of common trends holds for all transformations of the outcome (i.e. alternative labelling schemes in the current setting).

When several pre-treatment periods are available, an event study plot remains a useful starting point (Roth et al., 2023). An event-study plot is possible for the mean of the ordered outcome and latent index but, with a single year of data, we opt for a pseudo-treatment test. We split the pre-treatment period into a pre- and pseudo-post-treatment period with roughly an equal proportion of Muslims interviewed across both these periods. For example, 55 per cent of Muslims are interviewed prior to 7 July, with 26 per cent of Muslims interviewed prior to 26 May, which we designate the pre-treatment period, and 29 per cent of Muslims interviewed from 26 May to 6 July, which we designate the pseudo-post-treatment period. As with our previous analysis, we further split the pseudo-post-treatment period into 26 May–8 June and 9 June onward. Table 6 shows the effect of the pseudo-treatment on the mean of the ordered outcome (Panel A) and the latent index with and without heteroscedasticity (Panels B and C, respectively), providing little evidence of differential trends in either the mean of the ordered outcome or the latent index in the pre-treatment period. For completeness, Tables B6 and B7 in the online Appendix show the corresponding ATET in terms of response probabilities in the pseudo-post-treatment period, also showing little evidence of a treatment effect. Overall, therefore, the assumption of common trends holds in the pre-treatment period across different functional forms, which includes our preferred specification; this is encouraging.

TABLE 7 ATET in terms of the response probabilities (heteroscedastic ordered probit), alternative estimator

	7+ July	7–20 July	21+ July
$ATET_{P_1}$: 'not at all worried'	−0.054***	−0.004	−0.058***
	(0.019)	(0.035)	(0.019)
$ATET_{P_2}$: 'not very worried'	0.008	0.047	−0.014
	(0.024)	(0.047)	(0.025)
$ATET_{P_3}$: 'fairly worried'	0.087***	0.026	0.096***
	(0.020)	(0.039)	(0.022)
$ATET_{P_4}$: 'very worried'	−0.040	−0.068*	−0.024
	(0.025)	(0.041)	(0.025)
N	11,186	11,186	11,186

Note: Standard errors are calculated using bootstrapping (1,000 replications) in order to take the sampling variation in $I(Y_i = k)$ into account. Significance levels are shown as $^{***}p < 0.01$, $^{**}p < 0.05$ and $^{*}p < 0.1$.

5.2 | Alternative estimator

As explained in online Appendix A, an alternative estimator of $ATET_{P_k}$ is given by substituting the expression in the first square bracket of equation (3) by $I(Y_i = k)$, where $I(\cdot)$ is the indicator function.[7] Because the average of the predicted probabilities is not, in general, identical to the observed frequency of response category k, the two estimators will not coincide. Online Appendix D presents a small-scale simulation experiment, which shows that while both estimators are virtually unbiased, the estimator presented in equation (3) is somewhat more efficient (assuming that the assumptions underlying the estimator are met). However, as shown in online Appendix A, because the alternative estimator is based on weaker assumptions, it is useful to compare the results from the two approaches.

Table 7 presents the ATET in terms of response probabilities estimated using the alternative estimator allowing for heteroscedasticity. Comparing the results to those in Table 4 it can be seen that the findings are qualitatively very similar. However, the results from the alternative estimator suggests that the impact on fears of racial harassment are greater: for example, they show a 9.6 percentage point increase for 'fairly worried' after the second attack, compared to an estimated 4.7 percentage point increase according to the results in Table 4.

6 | CONCLUSION

Despite several recent developments for the method of DD with continuous outcomes – usefully summarised in Roth et al. (2023) – correspondingly little guidance exists for practitioners seeking to implement the method of DD with ordinal outcomes. Perhaps, as a result, empirical research adopts a diverse range of approaches, from linear regression to several different methods of converting an ordinal outcome to a binary outcome (see, e.g. Gruber and Mullainathan, 2005; Gregg, Harkness and Smith, 2009; Brodeur and Connolly, 2013; Leicester and Levell, 2016; Hole and Ratcliffe, 2020; Deal, 2022). Applying linear regression to ordinal outcomes can result in sign reversals of coefficients (see Schröder and Yitzhaki, 2017), with implications for research findings and policy advice based on this approach. However, converting an ordinal outcome to a binary outcome may involve arbitrary choices and a loss of information. These shortcomings limit the usefulness of a wealth of data routinely collected in surveys in public policy evaluation.

[7] Analogously, an alternative estimator allowing for heteroscedasticity is given by substituting the expression in the first square bracket of equation (4) by $I(Y_i = k)$.

In this paper, we extend the solution proposed for the method of DD with binary outcomes to ordinal outcomes (see, e.g. Blundell et al., 2004; Blundell and Costa Dias, 2009; Lechner, 2011; Puhani, 2012) to construct the ATET in terms of response probabilities. There are three advantages to focusing on response probabilities. First, marginal effects in terms of response probabilities are not susceptible to sign reversals. Secondly, this approach retains all information in the ordinal outcome. Thirdly, the ATET in terms of response probabilities provides a distributional assessment of the impact of treatment, allowing a more complete picture of the impact of treatment to emerge. We emphasise, however, that our goal is to increase the menu of available options to implement the method of DD with ordinal data rather than dictate which approach should be used.

In our empirical application, we analyse fears of racial harassment among Muslims living in a non-Muslim majority country following extremist Islamic terror attacks. Specifically, we focus on the first extremist Islamic terror attacks in the UK, defined by an initial attack on 7 July 2005 and another (failed) attack on 21 July 2005. After showing that sign reversals are possible when applying linear regression to our ordinal outcome, we switch to our proposed estimator. Our preferred specification, which allows for heteroscedasticity in the latent index, suggests a 5.3 percentage point reduction in the probability of Muslims responding 'not at all worried' and a 4.7 percentage point increase in 'fairly worried' relative to non-Muslims after the second terror attack, with little evidence of changes in response categories 'not very worried' and 'very worried'.

Our results contribute to the literature on the impact of extremist Islamic terror attacks on discrimination from the perspective of Muslims in two main ways. First, our findings are in line with those of Elsayed and de Grip (2018) and Giani and Merlino (2021) who find that extremist Islamic terror attacks worsen Muslim immigrants' attitudes toward integration and increase perceived discrimination among Muslims, respectively. Our analysis adds a new aspect to the existing evidence by focusing on fears of racial harassment, which have not been considered to date. Secondly, by using our proposed methodology we are able to build a more detailed picture of how circumstances changed for Muslims than is possible by using standard linear regression methods, by showing that the increase in fears of racial harassment are driven by a decrease in the probability of responding 'not at all worried' and an increase in the probability of reporting 'fairly worried' after the terror attacks.

ACKNOWLEDGEMENTS

Thanks to Karl Taylor, Nicolas Van de Sijpe, Bert Van Landeghem, Peter Wright, participants at the Geilo Health Economics Workshop (2019), iHEA World Congress (Basel 2019), and the universities of Erasmus Rotterdam (2019), Oxford (2019), Sheffield (2019), Lancaster (2016), Newcastle (2016), Universitat Jaume I (2016) and Manchester (2015) for helpful comments. This research uses the 2005 Citizenship Survey (Home Office. Communities Group, National Centre for Social Research, 2006). Special thanks to Robert Rutherfoord and NatCen for help with accessing the interview date in The Citizenship Survey. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

FUNDING

No funding was received for this manuscript.

CONFLICT OF INTEREST STATEMENT

The authors have nothing to disclose.

DATA AVAILABILITY STATEMENT

The version of the Citizenship Survey used in this paper is made available through the UK Data Service but this does not contain the date of interview. To obtain the date of interview, researchers must separately apply to NatCen. The interview date can then be matched to The Citizenship Survey obtained via the UK Data Service. The authors are happy to provide replication files on request.

ORCID

Arne Risa Hole https://orcid.org/0000-0002-9413-8101

REFERENCES

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, 80, 123–29.

Åslund, O., & Rooth, D. O. (2005). Shifts in attitudes and labor market discrimination: Swedish experiences after 9–11. *Journal of Population Economics*, 18, 603–29.

Athey, S., & Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica*, 74, 431–97.

Bloem, J. R. (2022). How much does the cardinal treatment of ordinal variables matter? An empirical investigation, *Political Analysis*, 30, 197–213.

Bloem, J. R., & Oswald, A. J. (2022). The analysis of human feelings: a practical suggestion for a robustness test. *Review of Income and Wealth*, 68, 689–710.

Blundell, R., & Costa Dias, M. (2009). Alternative approaches to evaluation in empirical microeconomics. *Journal of Human Resources*, 44, 565–640.

Blundell, R., Costa Dias, M., Meghir, C., & van Reenen, J. (2004). Evaluating the employment impact of a mandatory job search program. *Journal of the European Economic Association*, 2, 569–606.

Boes, S. (2013). Nonparametric analysis of treatment effects in ordered response models. *Empirical Economics*, 44, 81–109.

Bond, T. N., & Lang, K. (2019). The sad truth about happiness scales. *Journal of Political Economy*, 127, 1629–40.

Brodeur, A., & Connolly, M. (2013). Do higher child care subsidies improve parental well-being? Evidence from Quebec's family policies. *Journal of Economic Behavior & Organization*, 93, 1–16.

Chen, L-Y., Oparina, E., Powdthavee, N., & Srisuma, S. (2022). Robust ranking of happiness outcomes: a median regression perspective. *Journal of Economic Behavior & Organization*, 200, 672–86.

Clark, A. E. (1997). Job satisfaction and gender: why are women so happy at work? *Labour Economics*, 4, 341–72.

Dávila, A., & Mora, M. (2005). Changes in the earnings of Arab men in the US between 2000 and 2002. *Journal of Population Economics*, 18, 587–601.

Deal, C. (2022). Bound by Bostock: the effect of policies on attitudes. *Economics Letters*, 217, 110656.

Elsayed, A., & de Grip, A. (2018). Terrorism and the integration of Muslim immigrants. *Journal of Population Economics*, 31, 45–67.

Giani, M., & Merlino, L. P. (2021). Terrorist attacks and minority perceived discrimination. *British Journal of Sociology*, 72, 286–99.

Gould, E. D., & Klor, E. F. (2016). The long-run effect of 9/11: terrorism, backlash, and the assimilation of Muslim immigrants in the West. *Economic Journal*, 126, 2064–114.

Gregg, P., Harkness, S., & Smith, S. (2009). Welfare reform and lone parents in the UK. *Economic Journal*, 119, F38–F65.

Gruber, J. H., & Mullainathan, S. (2005). Do cigarette taxes make smokers happier? *B.E. Journal of Economic Analysis & Policy*, 5, 1–45.

Hanes, E., & Machin, S. (2014). Hate crime in the wake of terror attacks: evidence from 7/7 and 9/11. *Journal of Contemporary Criminal Justice*, 30, 247–67.

Hole, A. R., & Ratcliffe, A. (2020). The impact of the London bombings on the well-being of adolescent Muslims. *Scandinavian Journal of Economics*, 122, 1606–39.

Home Office. Communities Group, National Centre for Social Research (2006). Home Office Citizenship Survey, 2005. [data collection]. UK Data Service. SN: 5367, http://doi.org/10.5255/UKDA-SN-5367-1.

Johnston, D. W., & Lordan, G. (2012). Discrimination makes me sick! An examination of the discrimination–health relationship. *Journal of Health Economics*, 31, 99–111.

Kahn-Lang, A., & Lang, K. (2020). The promise and pitfalls of differences-in-differences: reflections on 16 and pregnant and other applications. *Journal of Business & Economic Statistics*, 38, 613–20.

Kaiser, C., & Vendrik, M. C. M. (2023). How much can we learn from happiness data? Working paper, https://doi.org/10.31235/osf.io/gzt7a.

Kaushal, N., Kaestner, R., & Reimers, C. (2007). Labor market effects of September 11th on Arab and Muslim residents of the United States. *Journal of Human Resources*, 42, 275–308.

Keele, L., & Park, D. K. (2006). Difficult choices: an evaluation of heterogenous choice models. Working paper, https://academicweb.nd.edu/~rwilliam/oglm/ljk-021706.pdf.

Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics*, 4, 165–224.

Leicester, A., & Levell, P. (2016). Anti-smoking policies and smoker well-being: evidence from Britain. *Fiscal Studies*, 37, 224–57.

Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics*, 13, 151–61.

Puhani, P. A. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear 'difference-in-differences' models. *Economics Letters*, 115, 85–87.

Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90, 2555–91.

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, 235, 2218–44.
Roth, J., & Sant'Anna, P. H. C. (2023). When is parallel trends sensitive to functional form? *Econometrica*, 91, 737–47.
Schröder, C., & Yitzhaki, S. (2017). Revisiting the evidence for cardinal treatment of ordinal variables. *European Economic Review*, 92, 337–58.
Wooldridge, J. M. (2023). Simple approaches to nonlinear difference-in-differences with panel data. *Econometrics Journal*, 26, C31–C66.
Yamauchi, S. (2020). Difference-in-differences for ordinal outcomes: application to the effect of mass shootings on attitudes toward gun control. Preprint, arXiv:2009.13404v1.
Zorlu, A., & Frijters, P. (2019). The happiness of European Muslims post-9/11. *Ethnic and Racial Studies*, 42, 23–44.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Hole, A. R., & Ratcliffe, A. (2025). Difference-in-differences with ordinal data: analysing the impact of terror attacks on racial harassment fears of UK Muslims. *Fiscal Studies*, 1–15. https://doi.org/10.1111/1475-5890.70006