The University Of Sheffield.

Department Of Economics

Difference-in-differences with an ordinal dependent variable: Assessing the impact of the London bombings on the safety perceptions of Muslims

Sheffield Economic Research Paper Series

SERPS no. 2021005

ISSN 1749-8368

29 September 2021

Difference-in-differences with an ordinal dependent variable: Assessing the impact of the London bombings on the safety perceptions of Muslims*

Arne Risa Hole† — University of Sheffield

Anita Ratcliffe‡ — University of Sheffield

29 September 2021

Abstract

In this paper, we propose a methodology for estimating treatment effects when using difference-in-differences with an ordinal dependent variable. Specifically, we derive an expression for the Average Treatment Effect on the Treated in terms of changes in response probabilities. An advantage of taking this approach is the ability to assess any distributional effects of exposure to treatment. We use the proposed estimator to evaluate the impact of the London bombings on the safety perceptions of Muslims, with our results highlighting a shift from moderately low to very high safety concerns among younger Muslims in the aftermath of the bombings.

Keywords: Difference-in-differences, ordinal dependent variables, terrorism

JEL codes: C25, I10, I31, J15

*The methodology presented in this paper first appeared in our working paper entitled 'The impact of the London bombings on the wellbeing of young Muslims' (Hole and Ratcliffe, 2015). We thank Karl Taylor, Bert Van Landeghem, participants at the Geilo Health Economics Workshop (2019) and the iHEA World Congress (Basel 2019) as well as seminar participants at the universities of Manchester (2015), Lancaster (2016), Newcastle (2016), Universitat Jaume I (2016), Erasmus University Rotterdam (2019), Oxford (2019) and Sheffield (2019) for very helpful comments. We take full responsibility for any errors. This research uses the 2005 Citizenship Survey. Home Office. Communities Group and National Centre for Social Research, Home Office Citizenship Survey, 2005 [computer file]. Colchester, Essex: UK Data Archive [distributor], June 2006. SN: 5367, `http://dx.doi.org/10.5255/UKDA-SN-5367-1`. Special thanks to Robert Rutherfoord and Natcen for help with accessing the interview date in The Citizenship Survey.

†University of Sheffield, Department of Economics, 9 Mappin Street, Sheffield, S1 4DT. E-mail: `a.r.hole@sheffield.ac.uk` (corresponding author)

‡University of Sheffield, Department of Economics, 9 Mappin Street, Sheffield, S1 4DT. E-mail: `a.ratcliffe@sheffield.ac.uk`

1 Introduction

The method of difference-in-differences is widely-applied to evaluate the effect of a policy change, intervention or other significant event on an outcome of interest. It identifies the effect of exposure to treatment by making the assumption of common trends in the mean of the outcome variable across treated and untreated groups. However, it is often implausible to make this assumption in models respecting the statistical properties of limited dependent variables. This issue, and its potential solution, has been illustrated in the context of a binary dependent variable (see for example Blundell et al., 2004; Blundell and Costa Dias, 2009; Lechner, 2011; Puhani, 2012). As noted in Lechner (2011) and Puhani (2012), the same approach can be applied to other limited dependent variables, though this possibility has received scant attention in the literature to date.

In this paper, we apply the method of difference-in-differences to an ordinal dependent variable (also widely termed ordered categorical dependent variable). Specifically, we derive an expression for the Average Treatment Effect on the Treated (ATET) in terms of changes in response probabilities. We then illustrate its application in an empirical setting to evaluate the impact of the London bombings (i.e. terror attacks on the 7th and 21st July 2005) on the safety perceptions of Muslims. We focus on an ordinal variable for several reasons. From a data-orientated perspective, ordered data are pervasive in social science disciplines and possess two key attributes that are particularly attractive in the assessment of a policy change, intervention or other notable event. The first, perhaps under-valued, attribute is that ordinal variables provide an opportunity to assess the distributional effects of exposure to treatment, allowing a more detailed picture of the treatment impact to emerge. This is particularly relevant when evaluating the impact of a policy change or medical intervention, where it might be of interest to know whether exposure to treatment has greatest effect on those initially worse off. The scope for ordered data to shed light on distributional effects is emphasised for non-parametric methods by Boes (2013) but this insight readily applies to parametric methods. The second, and widely appreciated, attribute is that data on attitudes and emotions provide supplementary evidence on outcomes that may be overlooked when prioritising objective indicators (Veenhoven, 2002). For example, our empirical application reveals heightened safety concerns among Muslims following the London bombings, that is in line with existing evidence on hate crimes (Hanes and Machin, 2014), while also indicating more widespread effects of the bombings. These two attributes combined make ordered data a powerful tool in the analysis of treatment effects that requires a better toolkit to fully unlock its potential. Moreover, as some ordered data - in particular subjective wellbeing - are seen as a priority of government policy (Frijters et al., 2020), the need for a better toolkit is only likely to increase in future.

From a methodological perspective, two well-used methods for assessing the determinants of ordinal variables (i.e. interpreting estimated coefficients from a linear regression model or from the latent equation of an ordered probit/logit model) have recently attracted considerable criticism (see for example Schröder and Yitzhaki, 2017; Bond and Lang, 2019). And while there is no consensus in the literature as to how to implement the method of difference-in-differences with an ordinal variable, the use of linear regression is widespread. Another commonly used approach converts the ordered outcome to a binary outcome and potentially discards interesting and useful variation in the outcome. The approach that we advocate, which analyses changes in response probabilities, circumvents criticisms raised in Schröder and Yitzhaki (2017) and Bond and Lang (2019) and retains all variation in the outcome. We would therefore argue that our approach provides a timely solution to existing methodological concerns while emphasising a key attribute of ordinal variables (i.e. the ability to consider distributional effects). In addition, our approach is easily implemented within the existing apparatus of Stata.

2 Related literature

Our paper primarily relates to an existing methodological literature using the method of difference-in-differences to identify the effect of a policy change, intervention or other significant event with a limited dependent variable. This literature considers these issues in the context of a binary dependent variable (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 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 variable 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 themes from a different vantage point. In response to Ai and Norton (2003), who liken the method of difference-in-differences to a model with an interaction term with the treatment effect given by the cross-difference (i.e. derivative) of the observed outcome, he clarifies that the treatment effect is the difference between two cross-differences (i.e. of the observed minus the potential outcome). In keeping with Ai and Norton (2003), Puhani (2012) uses a binary dependent variable to illustrate this result while noting it applies to any non-linear model with a strictly monotonic transformation function of a linear index (i.e. probit, logit or Tobit). The same possibility is raised in Lechner (2011). However, this point is not developed further, and as these contributions propose different estimators for the treatment effect with a binary outcome, different estimators are potentially available for an ordinal outcome. We build on this existing literature to derive an expression for the Average Treatment Effect on the Treated (ATET) for an ordinal variable in terms of changes in response probabilities. Our preferred estimator is a natural extension of the estimator suggested by Puhani (2012) for a binary variable as simulations indicate this estimator may be more efficient.

We also contribute to an empirical literature using the method of difference-in-differences with an ordinal dependent variable, which adopts a diverse range of approaches (see for example Gruber and Mullainathan, 2005; Gregg et al., 2009; Brodeur and Connolly, 2013; Leicester and Levell, 2016; Clark et al., 2020; Hole and Ratcliffe, 2020). The most dominant approach, attractive due to its simplicity, treats the ordered outcome as a continuous variable by attaching a numeric value to each response category, starting at 1 for the lowest category and increasing by 1 for each subsequent response category, and analyses its mean using linear regression methods. However, for some ordered outcomes (i.e. highest qualification attained), the mean lacks meaning. For other ordered outcomes (i.e. subjective wellbeing), the mean can be imbued with meaning when accepting strong assumptions governing the distance between response categories. However, ordinal variables provide information on the ranking of, and not distance between, response categories so that other equally valid approaches exist to assign numeric values to each response category. Schröder and Yitzhaki (2017) consider various alternative rank-preserving monotonic increasing transformations and show that the sign of estimated coefficients can be reversed using these alternative transformations. Bond and Lang (2019) make a similar point for the mean of the underlying latent variable in ordered probit/logit models, with estimated coefficients from these models sensitive to alternative, but equally valid, distributions of the error term. Both these papers suggest the estimated impact of treatment may hinge on equally valid choices made from a range of options. More recently, Kaiser and Vendrik (2020) cast doubt on the plausibility of reporting behaviour implied by alternative scale transformations where ordinal variables have a 'large' number of response categories and both Kaiser and Vendrik (2020) and Bloem and Oswald (2021) propose tests to verify the robustness of results obtained from linear regression methods. However, in the evaluation of treatment effects it is advantageous to consider distributional effects, and alternative approaches are still required where these tests indicate linear regression methods are inappropriate.

Another approach collapses the ordinal variable to a binary variable though various strategies are employed to achieve this. For example, Gruber and Mullainathan (2005) construct a series of binary dependent variables to indicate responses are equal to a particular response category, Leicester and Levell (2016) construct a series of 'cumulative' binary dependent variables with each binary variable indicating if responses are equal to, or greater than a given response category, and Hole and Ratcliffe (2020) construct a single binary variable to indicate a given strength of attachment/outcome. Such diversity arises given the lack of a basis for selecting an appropriate threshold when collapsing an ordered outcome to a binary one, and any choice invariably discards potentially interesting variation in the outcome variable.

Finally, we contribute to a well-established 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 and social outcomes including labour market outcomes, health and wellbeing and assimilation (see for example Åslund and Rooth, 2005; Dávila and Mora, 2005; Kaushal et al., 2007; Johnston and Lordan, 2012; Gould and Klor, 2016; Hole and Ratcliffe, 2020, and references therein). We focus on the safety perceptions of British Muslims following the London bombings and provide evidence of increased safety concerns among younger Muslims. Interestingly, our findings show a substantial decline in those feeling largely unconcerned about their safety and a substantial rise in those feeling very worried - as opposed to fairly worried - about it. In highlighting this shift towards acute levels of concern, these results illustrate the benefit of estimating the distributional effects of treatment - such detail is not available using other methods but is arguably very useful to policymakers. More generally, the findings provide complementary evidence to research documenting an increase in hate crimes in the immediate aftermath of terror attacks (Hanes and Machin, 2014; Gould and Klor, 2016) while also suggesting the impact of the London bombings extends beyond the victims of hate crime.

3 Methodology

As discussed earlier, several issues arise when trying to analyse ordered data in a difference-in-differences framework. In this research, we suggest that a solution is to construct the treatment effect in terms of the probability that a given response category is observed (i.e. the response probability) while also assuming common trends at the level of the latent variable.[1] This allows us to build on the expositions of Lechner (2011) and Puhani

[1]In parallel work to ours Yamauchi (2020) proposes an alternative identification strategy building on the work by Athey and Imbens (2006). Though his approach allows for a more flexible model

(2012) for a binary dependent variable, and we follow these discussions based on potential outcomes to begin with. Thus 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.[2] 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 the potential latent index underlies the potential outcome, it is the former that takes primary focus, and it is modelled as a function of group membership, time, and individual-level characteristics:

$$Y_i^{1*} = \beta_1 D_i + \delta^1 T_i + x_i'\gamma + \varepsilon_i \tag{1}$$

$$Y_i^{0*} = \beta_1 D_i + \delta^0 T_i + x_i'\gamma + \varepsilon_i \tag{2}$$

where D_i is equal to one if an individual is assigned to treatment and is zero otherwise. This feature allows the potential latent index to differ by an amount β_1 for individuals assigned to treatment, relative to those not assigned to treatment, and captures a time-invariant treatment group fixed effect. T_i is equal to one where an individual is observed in the post-treatment period and is zero otherwise. Thus an individual's potential latent index under treatment shifts by an amount δ^1 in the post-treatment period whereas it shifts by an amount δ^0 without treatment, with δ^1-δ^0 capturing the effect of treatment. Of course, we are not interested in the effect of treatment on the latent variable per se, and we show below how this translates into a treatment effect for the probability of observing a specific response category. Finally, x_i is a vector of individual characteristics and ε_i is an error term, which is assumed to be IID standard normal. The assumption

specification, our approach permits the use of standard ordered probit methods in keeping with the bulk of existing literature analysing ordinal dependent variables, and therefore provides a more natural bridge between that literature and the current endeavour to apply the method of difference-in-differences to these variables. This also means our approach can be implemented using standard software. Moreover, as detailed in Hole and Ratcliffe (2015) our approach can be linked to linear difference-in-differences estimation if researchers are prepared to assume equal distance between response categories.

[2] To simplify the exposition we have omitted time subscripts but these can easily be accommodated.

of common trends in the latent variable is embedded in equation 2, where the potential latent index without treatment follows the same trajectory for treated and control groups i.e. $E(Y_i^{0*} \mid D_i = 1, T_i = 1, x_i) - E(Y_i^{0*} \mid D_i = 1, T_i = 0, x_i) = E(Y_i^{0*} \mid D_i = 0, T_i = 1, x_i) - E(Y_i^{0*} \mid D_i = 0, T_i = 0, x_i) = \delta_0$.

Our point of departure from Lechner (2011) is that we are interested in a treatment effect with an ordered response model. We therefore turn our attention to showing that a treatment effect is identified with an ordered response model when common trends are assumed at the level of the latent variable. This first requires outlining how the potential latent index is linked to both the potential outcome and the probability of observing that the potential outcome is equal to a specific response category, which together provide a basis for constructing a treatment effect. For example, the potential latent index maps onto the potential outcome as follows:

$$Y_i^s = k \text{ if } \mu_k < Y_i^{s*} \leq \mu_{k+1}, \quad k = 1, ..., K \tag{3}$$

where s=0,1 so that Y^s denotes either of the two potential outcomes for each individual and k is one of multiple ordered response categories ranging from 1 to K. Thus we observe a potential outcome to be equal to the response category k if the associated potential latent index falls within the range defined by the two threshold parameters μ_k and μ_{k+1}. The threshold parameters are assumed to be strictly increasing in k ($\mu_k < \mu_{k+1}\ \forall k$) with $\mu_1 = -\infty$ and $\mu_{K+1} = \infty$. The probability that a potential outcome is equal to the response category k is given by:

$$\begin{aligned} P_{ik} &= E(I(Y_i^s = k) \mid D_i, T_i, x_i) \\ &= \Phi(\mu_{k+1} - E(Y_i^{s*} \mid D_i, T_i, x_i)) - \Phi(\mu_k - E(Y_i^{s*} \mid D_i, T_i, x_i)) \end{aligned} \tag{4}$$

where $I(\cdot)$ is the indicator function and $\Phi(\cdot)$ is the standard normal CDF.

As the potential latent index maps onto the probability that the potential outcome is equal to a specific response category, a natural way to think about the treatment effect is in terms of the effect of treatment on the probability of observing a specific response category. Since we never observe both potential outcomes for any individual, we cannot identify the individual-level treatment effect, and instead focus on the average treatment effect on the treated (ATET). This is simply the expected difference in the probability of observing a specific response category across the two treatment states for a randomly chosen individual in the treated group:

$$ATET_{P_k} = E(I(Y_i^1 = k) - I(Y_i^0 = k) \mid D_i = 1, T_i = 1) \tag{5}$$

Clearly the ATET requires a counterfactual response probability for individuals assigned to treatment, but this can be identified using the assumption of common trends in the latent variable contained in equation 2. For example, the common trends assumption implies:

$$\begin{aligned} &E(Y_i^{0*} \mid D_i = 1, T_i = 1, x_i) \\ &\quad = E(Y_i^{0*} \mid D_i = 1, T_i = 0, x_i) + E(Y_i^{0*} \mid D_i = 0, T_i = 1, x_i) - E(Y_i^{0*} \mid D_i = 0, T_i = 0, x_i) \end{aligned} \tag{6}$$

which shows that the expected potential latent index without treatment for treated individuals - required to model the counterfactual response probability as per equation 4 - can be expressed in terms of several expectations of the potential latent index. Notice that all information required for the right hand side of equation 6 is available: the potential latent index without treatment is realised pre/post treatment for untreated individuals while for treated individuals the potential latent wellbeing with treatment is realised in the pre-treatment period.[3] The latter is equivalent to the potential latent wellbeing without treatment, given there is no effect of treatment in this period. The required indices are easily obtained using equation 2 but it is more convenient to use the following realisation

[3]Note that we distinguish between *realised* and *observed*. While the latent index is unobserved by definition, it is realised when it maps onto a realised (and observed) outcome.

rule to convert the potential latent wellbeing into realised latent wellbeing:

$$\begin{aligned}
Y_i^* &= D_i Y_i^{1*} + (1 - D_i) Y_i^{0*} \\
&= D_i(\beta_1 D_i + \delta^1 T_i + x_i'\gamma + \varepsilon_i) + (1 - D_i)(\beta_1 D_i + \delta^0 T_i + x_i'\gamma + \varepsilon_i) \\
&= \beta_1 D_i + \delta^0 T_i + (\delta^1 - \delta^0) D_i T_i + x_i'\gamma + \varepsilon_i \\
&= \beta_1 D_i + \beta_2 T_i + \beta_3 D_i T_i + x_i'\gamma + \varepsilon_i
\end{aligned} \tag{7}$$

where $\beta_2 = \delta^0$ and $\beta_3 = \delta^1 - \delta^0$. The expected potential latent index without treatment for treated individuals in the post-treatment period is therefore:

$$\begin{aligned}
&E(Y_i^{0*} \mid D_i = 1, T_i = 1, x_i) \\
&\quad = E(Y_i^* \mid D_i = 1, T_i = 0, x_i) + E(Y_i^* \mid D_i = 0, T_i = 1, x_i) - E(Y_i^* \mid D_i = 0, T_i = 0, x_i) \\
&\quad = \beta_1 + \beta_2 + x_i'\gamma
\end{aligned} \tag{8}$$

This suggests that the counterfactual response probability in the post-treatment period is given by:

$$E(I(Y_i^0 = k) \mid D_i = 1, T_i = 1, x_i) = \Phi(\mu_{k+1} - \beta_1 - \beta_2 - x_i'\gamma) - \Phi(\mu_k - \beta_1 - \beta_2 - x_i'\gamma) \tag{9}$$

Thus an estimate of the ATET is given by:

$$\begin{aligned}
\widehat{ATET}_{P_k} = \frac{1}{N^1} \sum_{i=1}^{N} D_i T_i \Big\{ &\Big[\Phi(\hat{\mu}_{k+1} - \hat{\beta}_1 - \hat{\beta}_2 - \hat{\beta}_3 - x_i'\hat{\gamma}) - \Phi(\hat{\mu}_k - \hat{\beta}_1 - \hat{\beta}_2 - \hat{\beta}_3 - x_i'\hat{\gamma})\Big] \\
&- \Big[\Phi(\hat{\mu}_{k+1} - \hat{\beta}_1 - \hat{\beta}_2 - x_i'\hat{\gamma}) - \Phi(\hat{\mu}_k - \hat{\beta}_1 - \hat{\beta}_2 - x_i'\hat{\gamma})\Big] \Big\}
\end{aligned} \tag{10}$$

where $N^1 = \sum_{i=1}^{N} D_i T_i$.[4]

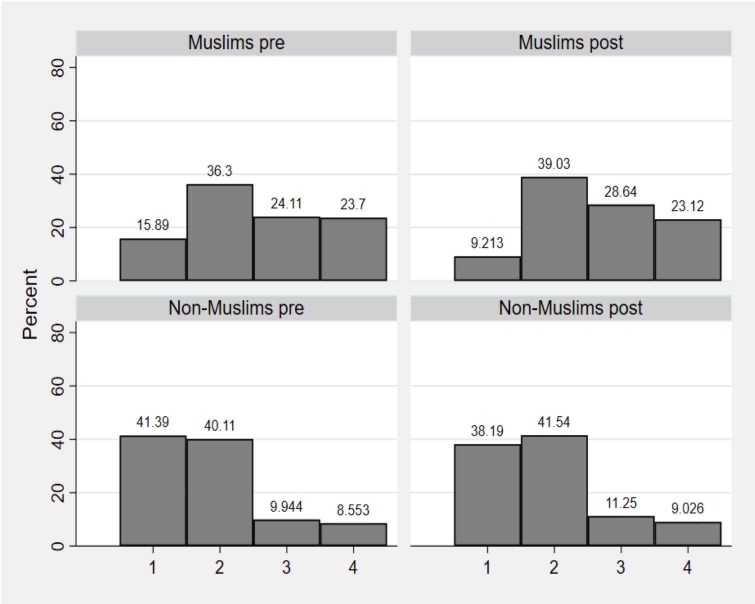
[4] An alternative estimator of $ATET_{P_k}$ is given by substituting the expression in the first square bracket of equation 10 by $I(Y_i = k)$, which is an extension of the estimator proposed by Lechner (2011). Since the average of the predicted probabilities is not, in general, identical to the observed frequency of response category k, the two estimators will not coincide. Appendix C presents a small-scale simulation experiment where we find that while both estimators are virtually unbiased, the estimator presented in equation 10 is somewhat more efficient.

4 Application

We illustrate how to implement this estimator with an empirical application. We consider the impact of the London bombings on the safety perceptions of Muslims. Previous research suggests that hate crimes targeting Muslims increased (Hanes and Machin, 2014) and that there may have been a more general rise in prejudice towards Muslims (Ratcliffe and von Hinke Kessler Scholder, 2015; Hole and Ratcliffe, 2020) following these terror attacks. We use The Citizenship Survey, a large and representative cross-section survey taking place in 2001, 2003, 2005 and then annually from 2007 until 2011. It interviews individuals aged 16+ living in England and Wales and contains a range of information on charitable and civic activity, community cohesion and race relations. We use the 2005 sweep, which contains a 'Fear of Crime' module[5] from which we take our dependent variable '*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*'.[6] Approximately 14 000 individuals (comprising a core sample of approximately 10 000 and an ethnic minority boost sample of approximately 4 000) are interviewed in the 2005 sweep, with fieldwork taking place between 8th March and 30th September.[7] Crucially, we have special permission to use the interview date to construct a clean pre and post treatment period, with post treatment defined as July 7th onwards. We exclude individuals aged 70+ (approximately 1 900) as education qualifications are not collected for this group. Although we could further restrict our sample to ethnic minorities, we follow Hole and Ratcliffe (2015) and simply compare the outcome of Muslims to non-Muslims, but our results are not sensitive to this choice. We control for a wide range of demographic, socioeconomic and attitudinal variables and after excluding individuals with missing information our final sample comprises just under 11 200 individuals. Figure 1 plots the distribution of the dependent variable by Muslims and Non-Muslims, pre and post treatment. Equivalent plots by age groups and summary statistics for the control variables are available in Appendix A.

[5]The 'Fear of Crime' module does not appear in the 2001 or 2003 sweeps of The Citizenship Survey.
[6]Less than 1% of respondents refuse to answer the question or respond with '*don't know*'.
[7]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, pre and post treatment



Notes: 1 corresponds to response category *'not at all worried'*, 2 to *'not very worried'*, 3 to *'fairly worried'*, 4 to *'very worried'*.

Table 1 presents results for the impact of the London bombings on the safety perceptions of Muslims. To estimate the ATET for each response category we fit an ordered probit model defined in equations 1-4 and 7 on the full sample and use a routine post-estimation command to construct the expression in equation 10 (see Appendix B for Stata code). We first focus on the impact on all Muslims, and subsequently on younger (age ≤ 35) versus older (age > 35) Muslims. Nandi and Luthra (2016) show that ethnic minorities are more likely to anticipate and experience harassment in public places. As it is harder for young people to avoid these spaces, due to the need to travel for education, work and childcare, it is likely that young people face greater exposure to crime and may fear it more.

Results in column 1 provide suggestive evidence of an adverse effect of the bombings on the safety perceptions of Muslims, with the probability of reporting response categories *'not at all worried'* and *'not very worried'* decreasing by about 1.8 percentage points relative to non-Muslims. Conversely, the probability of reporting *'fairly worried'* and *'very worried'* increase by 0.6 and 3.1 percentage points, respectively. However, for the

Table 1: The impact of the London bombings on the safety perceptions of Muslims

	All (1)	Age $\leq$ 35 (2)	Age > 35 (3)
$ATET_{P_1}$	-0.018	-0.033**	0.014
	(0.011)	(0.013)	(0.021)
$ATET_{P_2}$	-0.019	-0.043**	0.011
	(0.013)	(0.020)	(0.015)
$ATET_{P_3}$	0.006*	0.009***	-0.005
	(0.003)	(0.003)	(0.008)
$ATET_{P_4}$	0.031	0.067**	-0.020
	(0.021)	(0.030)	(0.028)
N	11186	4164	7022

Notes: $ATET_{P_1}$ corresponds to response category *'not at all worried'*, $ATET_{P_2}$ to *'not very worried'*, $ATET_{P_3}$ to *'fairly worried'*, $ATET_{P_4}$ to *'very worried'*. See Section 3 for details of the estimation strategy. Standard errors are calculated using the linearisation method. Significance levels are shown as *p<0.1, **p<0.05, ***p<0.01.

most part these results are statistically insignificant at conventional levels.

Columns 2 and 3 show that the impact of the bombings on safety perceptions is largely confined to younger Muslims. For this age group, and after controlling for a wide range of variables, the probability of responding *'not at all worried'* and *'not very worried'* falls by 3.3 and 4.3 percentage points respectively for Muslims relative to non-Muslims after the bombings. At the same time, the probability of responding *'very worried'* increases by 6.7 percentage points for Muslims relative to non-Muslims and while the probability of responding *'fairly worried'* also increases, it does so less markedly at 0.9 percentage points. All of these effects are statistically significant at the 5 percent level. For older Muslims, the pattern of changes in each response category is consistent with a general improvement in safety perceptions, but the effects are small and not significant at conventional levels.

These results illustrate the benefit to estimating the distributional effects of exposure to treatment. Specifically, they show that the impact of treatment is greatest for the probability of feeling *'very worried'*, with the difference in the estimated treatment effect for response categories *'very worried'* and *'fairly worried'* statistically significant (see Table 2). Conversely, there is little evidence of a statistically significant difference in the estimated treatment effect between response categories *'not at all worried'* and *'not very worried'*. These results therefore shine light on a shift from low level to utmost safety concern among younger Muslims after the bombings. Such findings would be overlooked using the usual methods to estimate treatment effects with ordered data and yet, from a policy perspective, this pattern of results may pose more of a concern than, for example, equal effects of treatment across all response categories. Moreover, Nandi and Luthra (2016) show that fear of harassment is more pervasive than experience of harassment among British Asians, suggesting the impact of the bombings may be more widespread than the evidence on hate crimes suggests, particularly since fear of harassment is associated with poorer mental health (Nandi et al., 2016).

Table 2: Differences in estimated treatment effects

	All (1)	Age $\leq$ 35 (2)	Age > 35 (3)
$ATET_{P_1} - ATET_{P_2}$	0.001	0.010	0.004
	(0.002)	(0.008)	(0.007)
$ATET_{P_1} - ATET_{P_3}$	-0.024	-0.042***	0.020
	(0.015)	(0.016)	(0.029)
$ATET_{P_1} - ATET_{P_4}$	-0.049	-0.100**	0.034
	(0.032)	(0.044)	(0.049)
$ATET_{P_2} - ATET_{P_3}$	-0.025	-0.053**	0.016
	(0.017)	(0.023)	(0.023)
$ATET_{P_2} - ATET_{P_4}$	-0.050	-0.110**	0.031
	(0.034)	(0.051)	(0.042)
$ATET_{P_3} - ATET_{P_4}$	-0.025	-0.057**	0.015
	(0.018)	(0.028)	(0.020)

Notes: $ATET_{P_1}$ corresponds to response category *'not at all worried'*, $ATET_{P_2}$ to *'not very worried'*, $ATET_{P_3}$ to *'fairly worried'*, $ATET_{P_4}$ to *'very worried'*. Significance levels are shown as *p<0.1, **p<0.05, ***p<0.01.

5 Conclusions

In this paper, we describe an expression for the Average Treatment Effect on the Treated (ATET) in terms of changes in response probabilities for the method of difference-in-differences with an ordinal dependent variable, extending existing work in this area for a binary dependent variable (Blundell and Costa Dias, 2009; Lechner, 2011; Puhani, 2012). We argue, and demonstrate in our empirical application, that a key advantage of this approach is the ability to capture distributional effects of exposure to treatment. Importantly, the proposed approach circumvents recent concerns raised in the analysis of ordinal variables when using linear and non-linear methods (see for example Schröder and Yitzhaki, 2017; Bond and Lang, 2019). Finally, our empirical application extends existing research examining the impact of extremist Islamic terrorism on the outcomes of Muslims living in non-Muslim majority countries. We show that concerns for personal safety increased sharply among (younger) Muslims relative to non-Muslims in the aftermath of the London bombings, which is in line with existing research on hate crimes (Hanes and Machin, 2014; Gould and Klor, 2016), while also indicating more widespread effects of the bombings.

References

Ai, C. and E. C. Norton (2003). Interaction terms in logit and probit models. *Economics Letters 80*(1), 123 – 129.

Athey, S. and G. W. Imbens (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica 74*(2), 431–497.

Bloem, J. R. and A. J. Oswald (2021). The analysis of human feelings: A practical suggestion for a robustness test. *Review of Income and Wealth (forthcoming).*

Blundell, R. and M. Costa Dias (2009). Alternative approaches to evaluation in empirical microeconomics. *Journal of Human Resources 44*(3), 565–640.

Blundell, R., M. Costa Dias, C. Meghir, and J. van Reenen (2004). Evaluating the employment impact of a mandatory job search program. *Journal of the European Economic Association 2*(4), 569–606.

Boes, S. (2013). Nonparametric analysis of treatment effects in ordered response models. *Empirical Economics 44*(1), 81–109.

Bond, T. N. and K. Lang (2019). The sad truth about happiness scales. *Journal of Political Economy 127*(4), 1629–1640.

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

Clark, A. E., O. Doyle, and E. Stancanelli (2020). The Impact of Terrorism on Individual Well-Being: Evidence from the Boston Marathon Bombing. *The Economic Journal 130*(631), 2065–2104.

Dávila, A. and M. Mora (2005). Changes in the earnings of Arab men in the US between 2000 and 2002. *Journal of Population Economics 18*(4), 587–601.

Frijters, P., A. E. Clark, C. Krekel, and R. Layard (2020). A happy choice: wellbeing as the goal of government. *Behavioural Public Policy 4*(2), 126–165.

Gould, E. D. and E. F. Klor (2016). The long-run effect of 9/11: Terrorism, backlash, and the assimilation of Muslim immigrants in the West. *The Economic Journal 126*(597), 2064–2114.

Gregg, P., S. Harkness, and S. Smith (2009). Welfare reform and lone parents in the UK. *The Economic Journal 119*(535), F38–F65.

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

Hanes, E. and S. Machin (2014). Hate crime in the wake of terror attacks: Evidence from 7/7 and 9/11. *Journal of Contemporary Criminal Justice 30*(3), 247–267.

Hole, A. R. and A. Ratcliffe (2015). The impact of the London bombings on the wellbeing of young Muslims. Sheffield Economic Research Paper Series 2015002, University of Sheffield.

Hole, A. R. and A. Ratcliffe (2020). The impact of the London bombings on the well-being of adolescent Muslims. *The Scandinavian Journal of Economics 122*(4), 1606–1639.

Johnston, D. W. and G. Lordan (2012). Discrimination makes me sick! An examination of the discrimination–health relationship. *Journal of Health Economics 31*(1), 99 – 111.

Kaiser, C. and M. C. M. Vendrik (2020). How threatening are transformations of reported happiness to subjective wellbeing research? INET Oxford Working Paper 2020-19, Institute for New Economic Thinking.

Kaushal, N., R. Kaestner, and C. Reimers (2007). Labor market effects of September 11th on Arab and Muslim residents of the United States. *Journal of Human Resources 42*(2), 275–308.

Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends(R) in Econometrics 4*(3), 165–224.

Leicester, A. and P. Levell (2016). Anti-smoking policies and smoker well-being: Evidence from Britain. *Fiscal Studies 37*(2), 224–257.

Nandi, A. and R. Luthra (2016). Who experiences ethnic and racial harassment? 'Written evidence submitted by the Institute for Social and Economic Research, University of Essex' to Home Affairs Committee Inquiry on Hate crime: abuse, hate and extremism online. Document HCR0090.

Nandi, A., R. Luthra, and M. Benzeval (2016). Ethnic and racial harassment and mental health: identifying sources of resilience. ISER Working Paper Series 2016-14.

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.

Åslund, O. and D. O. Rooth (2005). Shifts in attitudes and labor market discrimination: Swedish experiences after 9-11. *Journal of Population Economics 18*(4), 603–629.

Ratcliffe, A. and S. von Hinke Kessler Scholder (2015). The London bombings and racial prejudice: Evidence from housing and labour markets. *Economic Inquiry 53*, 276–293.

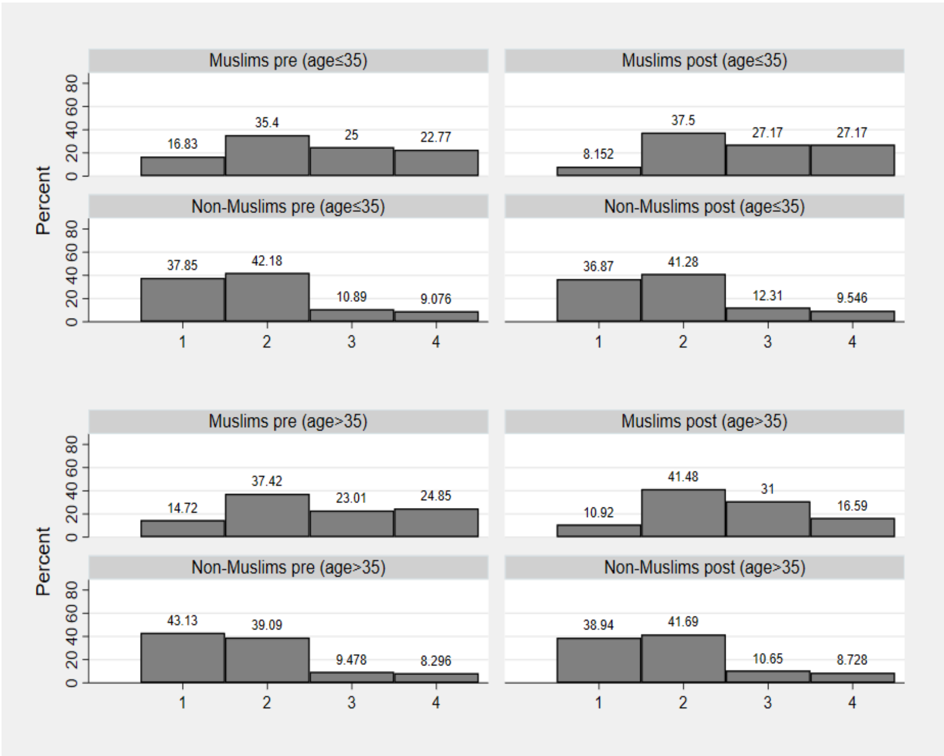
Schröder, C. and S. Yitzhaki (2017). Revisiting the evidence for cardinal treatment of ordinal variables. *European Economic Review 92*, 337 – 358.

Veenhoven, R. (2002). Why social policy needs subjective indicators. *Social Indicators Research 58*, 33 – 46.

Yamauchi, S. (2020). Difference-in-differences for ordinal outcomes: Application to the effect of mass shootings on attitudes toward gun control. arXiv:2009.13404 [stat.AP].

A Descriptive statistics

Figure A1: Distribution of the dependent variable by Muslims and Non-Muslims, pre and post treatment



Notes: 1 corresponds to response category *'not at all worried'*, 2 to *'not very worried'*, 3 to *'fairly worried'*, 4 to *'very worried'*.

Table A1: Differences in sample means of control variables by treatment status

	Muslims		non-Muslims	
	Pre	Post	Pre	Post
Ethnicity				
Asian	0.785	0.786	0.107	0.124***
Black	0.067	0.095*	0.109	0.147***
Mixed/other ethnicity	0.126	0.102	0.069	0.082**
Sociodemographics				
Female	0.490	0.529	0.559	0.552
Age	35.527	34.606	42.946	41.981***
Age squared	1404.311	1337.189	2039.957	1960.788***
Has partner	0.629	0.570**	0.592	0.535***
Widowed/divorced/separated	0.144	0.178*	0.160	0.175**
Has children	0.562	0.573	0.328	0.335
Born Abroad	0.726	0.730	0.228	0.258***
Religion important	0.801	0.774	0.340	0.348
Economic resources				
Homeowner	0.515	0.504	0.697	0.655***
Degree	0.207	0.179	0.238	0.255*
In education	0.099	0.074	0.032	0.032
No qualifications	0.353	0.355	0.222	0.201**
In work	0.468	0.442	0.662	0.679*
Manager/professional	0.184	0.164	0.353	0.361
Ln(working hours + 1)	1.613	1.534	2.311	2.381**
Personal income at least £20 000	0.155	0.134	0.305	0.320
Missing personal income	0.110	0.112	0.068	0.074
Partner's income at least £20 000	0.071	0.047*	0.191	0.183
Missing partner income	0.486	0.529	0.480	0.534***
Perceptions of discimination				
Faced discrimination in labour market	0.073	0.077	0.039	0.046*
Faced discrimination in accessing public services	0.132	0.109	0.032	0.034
Area of residence				
Lived in area 0-2 years	0.253	0.251	0.186	0.198
Mostly of same ethnicity in local area	0.277	0.290	0.653	0.591***
High Deprivation area (top decile)	0.341	0.390*	0.103	0.119**
High population density area (top decile)	0.623	0.571*	0.306	0.306
Region of residence				
North East	0.007	0.007	0.049	0.028***
North West	0.088	0.116*	0.111	0.087***
Yorkshire and the Humber	0.211	0.126***	0.094	0.078***
East Midlands	0.045	0.039	0.086	0.082
West Midlands	0.123	0.154	0.096	0.119***
East of England	0.041	0.054	0.088	0.096
South East	0.041	0.075***	0.116	0.140***
South West	0.015	0.007	0.090	0.066***
Wales	0.018	0.002***	0.045	0.032***
N	730	597	5682	4177

Notes: Difference in sample means between individuals interviewed 8th March-6th July 2005 (Pre) and individuals interviewed 7th July-30th September 2005 (Post). Significance levels are shown as *p<0.1, **p<0.05, ***p<0.01. No corrections for multiple comparisons applied.

B Stata code

The methodology proposed in this paper can be straightforwardly implemented in Stata by running the following commands:

```
oprobit Y D T DT x1 x2, vce(robust)
margins, dydx(DT) vce(unconditional) subpop(DT) post
```

Here Y is the dependent variable, D is equal to one if an individual is assigned to treatment and zero otherwise and T is equal to one if an individual is observed in the post-treatment period and zero otherwise. DT is the interaction between D and T (i.e. $DT = D \times T$) and $x1$ and $x2$ are control variables.

The margins command calculates the ATET for each response category using equation 10. The `subpop()` option of margins ensures that the expression in curly brackets is averaged over the correct subsample, i.e. the one defined by $DT = 1$. The `vce(unconditional)` option requests that standard errors should be calculated using the linearisation method, which takes the sampling variability of the control variables into account (in addition to the variability of the coefficient estimates) and in our experience produces similar results to using bootstrapping. The estimation procedure can be adapted to adjust the standard errors for clustering by simply amending the `vce()` option of the oprobit command.

The `post` option of `margins` makes it straightforward to subsequently calculate differences between the estimated ATETs, with corresponding standard errors, using the `lincom` or `nlcom` commands. As an example, the following code calculates the difference between the estimated ATETs for response categories 1 and 2:

```
nlcom [DT]1._predict - [DT]2._predict
```

C Simulation results

As discussed in Section 3 (footnote 4) an alternative estimator of $ATET_{P_k}$ is given by substituting the expression in the first square bracket of equation 10 by $I(Y_i = k)$, which is an extension of the estimator proposed by Lechner (2011) in the context of binary choice models. Since 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.

To investigate the properties of the two alternative estimators we carried out a small-scale simulation experiment. In the simulated data 13% of the respondents are specified to belong to the treatment group ($D_i = 1$) and 50% are observed post treatment ($T_i = 1$). The true latent index function is specified as:

$$Y_i^* = 0.5D_i + 0.015T_i + 0.2D_iT_i + 0.4x_i + u_i$$

where x_i is a binary control variable with 80% probability of being equal to one if the individual is a member of the treatment group and 35% otherwise. u_i is specified to be distributed standard normal, and the true values of the threshold parameters (-0.23, 0.91 and 1.49) are set to approximately reproduce the distribution of the dependent variable in the empirical application.

Given this simple setup we can calculate the true values of $ATET_{P_k}$ for each response category, and hence evaluate which estimator performs best in terms of bias and root mean square error (RMSE). As can be seen in Table C1 both estimators are virtually unbiased, as the mean values are very close to the true values for each response category. However, our preferred estimator is somewhat more efficient than the alternative estimator, as can be seen by the lower values of RMSE across all response categories.

Table C1: Simulation results ($N = 10000$, 10000 replications)

		Preferred		Alternative	
	True value	Mean	RMSE	Mean	RMSE
$ATET_{P_1}$	-0.0406	-0.0407	0.0134	-0.0408	0.0160
$ATET_{P_2}$	-0.0380	-0.0379	0.0121	-0.0377	0.0199
$ATET_{P_3}$	0.0108	0.0108	0.0039	0.0108	0.0157
$ATET_{P_4}$	0.0678	0.0678	0.0217	0.0677	0.0239

Notes: $ATET_{P_1}$ corresponds to response category *'not at all worried'*, $ATET_{P_2}$ to *'not very worried'*, $ATET_{P_3}$ to *'fairly worried'*, $ATET_{P_4}$ to *'very worried'*. The 'Preferred' column reports the results from using equation 10 to estimate the $ATET$s. The 'Alternative' column reports the results from using the alternative estimator, where the expression in the first square bracket of equation 10 is replaced by $I(Y_i = k)$.