(Very preliminary draft)

Electoral participation over the life cycle: Evidence from British longitudinal survey data

Kazuya Masuda [*] Tomohito Okabe [†] Christopher Spencer [‡]

September 20, 2019

Abstract

In the national elections of many developed countries, voter turnout among older citizens is higher than that among younger citizens. This may imply that people are more likely to vote as they get older. However, it remains unclear how and what socio-economic channels shape this pattern. We address this question by arguing that life events (e.g. marriage, living stability) cause people to mature over time and foster political engagement. Using British long-term panel survey, we show that life events are important sources for the formation of electoral participation.

Keywords: political lifecycle theory; voter turnout; panel data analysis

1 Introduction

In the national elections of many developed countries, aggregate voter turnout is much lower among younger citizens than among older citizens. Figure 1 and 2 show changes in aggregate voter turnout in the United Kingdom. The latter figure consists of six panels, each of which shows voter turnout for citizens of different ages. The youngest cohort (panel (a)) is the least likely to vote, while the percentage of voters increases as citizens age, with the highest percentage of voters being 65 or above (panel (f)). This inequality in aggregate voter turnout can impact policy development through democratic elections in which conflicts of interests lie between generations (e.g. social security). Theoretically, we can describe such a policy development using a Downsian voting model (Downs, 1957). Consider a simple environment where the population of young voters is smaller than that of old voters in a median-voter setup. Since the age of the median voter is high, the electoral competition results in an increase in pension benefits that is less favoured by young voters. Given the lower turnout of young voters, the policy development

[*]Hitotsubashi University

[†]Hitotsubashi University, Corresponding Author, Email: t-okabe@ier.hit-u.ac.jp

[‡]Loughborough University

The authors are grateful to Heinrich Ursprung and Guido Schwerdt for fruitful suggestions, and participants at the University of Konstanz public economics seminar for helpful comments.

turns out more skewed towards old voters. This issue in an aging economy has been investigated in the relevant politico-economic literature. See, for example, Cooley and Soares (1999); Gonzalez-Eiras and Niepelt (2008); Piggott and Woodland (Ch.7 2017).

Why, however, does the electorate become more engaged in voting as people age? In fact, the answer to this question remains unclear. We seek to clarify this phenomenon and to determine the factors influencing lifecycle voting patterns. Our study provides important implications for electoral administrators who consider effective initiative programs to raise voter turnout, as well as for political parties who struggle to get the attention of the electorate. For the purpose of this study, we rely on the political lifecycle theory. In line with Parry (1992), this theory argues that life events (e.g. marriage, home ownership) stimulate voters' interest in electoral campaigns, resulting in greater involvement in democracy over time. We can imagine that as people experience life events, they will have more stake in socio-economic policies. They then get more interested in electoral campaigns where politicians announce various policy platforms, thus motivating their electoral participation. This idea for motivating electoral participation stems from consumption voting proposed by Riker and Ordeshook (1968). They assume that the voter obtains a consumption benefit from the act of voting. We construct a simple algebraic model of this behavioural theory, and show how individual electoral participation is characterized by life events.

In this study, we use long-term British panel data, *Harmonized BHPS-UKHLS data* from 1992 to 2015 to investigate the causality between life events and individual electoral participation. This panel data enables us to control for unobserved factors (i.e. individual fixed effects) that affect the dependent variable. Reverse causality between life events and electoral participation is limited; however, without controlling for the unobserved factors, potential 'spurious correlation' remains a problem. Controlling for unobserved fixed effects, as we do in this study, rules out such spurious correlation and omission bias caused by time-invariant confounders. Our linear and nonlinear probability models demonstrate that most life events have a significantly positive impact on electoral participation. More specifically, we find that demographic variables (i.e. marriage, having children and living stability) and an asset ownership variable (i.e. home ownership) raise the likelihood of voting whereas labor supply variables (i.e. employment and retirement) do not. The former finding implies that experiences of those life event expand the political interests of electorates, thus encouraging them to cast votes. The latter finding implies that working experiences do not matter for motivating electoral participations, holding all else constant. The average partial effect of all significant life events increases the likelihood of voting by approximately 25 %, given a benchmark of linear regression with no life events. To further investigate our results, we conduct three robustness checks. First, we employ the panel matching method proposed by Imai et al. (2019). This method allows to control for sampling bias which is potentially caused by missing observations of vote responses. Second, we re-estimate with a dynamic panel model by adding one lagged variable of the dependent variable. The estimate for the lagged variable accommodates the habituation hypothesis where the voting experience in the last period fosters the political interest of the electorate, thereby motivating electoral participation. Lastly, we again conduct baseline estimations with another set of panel data, *National Development Child Study* (NCDS). Overall, the robustness check exercises support our basic results; however, we need to conduct further research to improve our exercises.

We review the empirical literature relevant to our study. As the political lifecycle theory was originally developed only recently, its validity has not yet been fully examined in existing studies. For instance, Smets (2016) investigates the validity of the lifecycle theory, focusing on young adults in the United Kingdom from 1964 to 2010. She empirically confirms that life events have positive and significant effects on young citizens' turnout. She also shows that modern young adults experience fewer life events than their parents and grandparents at the same age, thus causing a decline in the voter turnout of young generations compared to past decades. Likewise, Highton and Wolfinger (2001) examine a similar issue with American survey data. Both these studies differ from ours in that they employ repeated cross-sectional data, whereas our study relies on panel data. The cross-sectional analysis in these previous works failed to control for unobserved heterogeneities that affect individual voting patterns, meaning that there may be missed sources of bias. This shortcoming is remedied in our study. While other studies examine the relationship between voter turnout and life events, most focus on a single event. Jankowski and Strate (1995) study the effect of getting a job on voting. Stoker and Jennings (1995) examine the effect of marriage, Flanagan and Sherrod (1998) explore the effect of having children, and Highton (2000) investigates the effect of residential stability. Meanwhile, Jankowski and Strate (1995) focus on the effect of home on voting. Smets and Neundorf (2014) investigate differences in cohort voting patterns using British survey data from 1972 to 2010. They show that life events, including marriage and employment, are positively associated with individual voter turnout. They demonstrate that electorates have a higher propensity to vote despite age or survey period. Other papers also investigate such dynamic propensity, labelling it as habituation (Green and Shachar, 2000; Plutzer, 2002; Aldrich et al., 2011). To the best of our knowledge, Denny and Doyle (2009) is the paper that is the closest to our study in terms of estimation methodology. Using the panel data set, they controled for individual fixed-effects, thus removing the omitted bias which generally exist in cross-sectional estimation. Although their main research question is regarding the formation of electoral habituation, they also show the effectiveness of marriage and having children on voting, using NCDS data with three waves from 1981 to 2000. Although they provide important evidence for the lifecycle theory, their evidence is valid only for one particular cohort born in 1958. Moreover, they did not examine some life events. In contrast, our study covers the observations of various generations and additional lifecycle variables, including employment, living stability, and home ownership. We also use NCDS data for our robustness check. For other relevant papers on voter turnout, habit, and life events, see Smets (2016), which provides a comprehensive literature review.

Theoretically, we rely on consumption voting framework, but other approaches to voter turnout have been proposed by existing studies. For instance, Feddersen and Pesendorfer (1996) provides a game theory framework to account for the abstentions of voting. They assume that externality among voters controls for voting decisions, but this externality would not exist in a large election where many voters cast their votes. Bendor et al. (2003) propose a simulation model to account for the endogenous formation of turnout in a large election. They certainly show the endogenous turnout, but their model is mechanical in the sense that the likelihood of individual voting is stochastically given and mechanically updated based on the history of voting. The simulated model does not provide much information on the factors that foster political participation. Thus, our study contributes to the future development of a formal theory of endogenous voter turnout by providing new empirical evidence.

The remainder of this paper is organized as follows: Section 2 presents the theoretical foundation and econometric models. Section 3 outlines the results of our estimation. Finally, Section 4 summarizes our findings and discusses remaining issues to be addressed in further investigations.

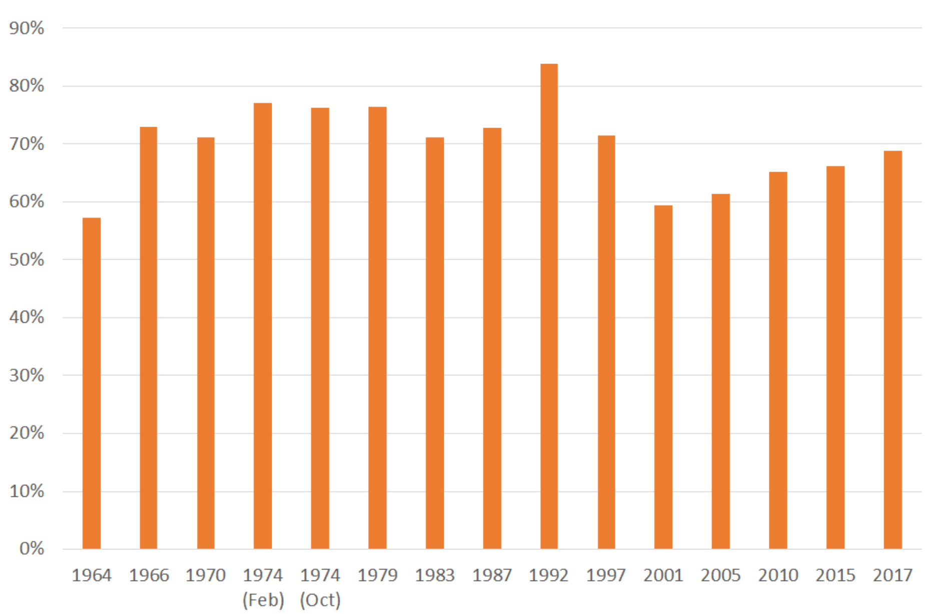
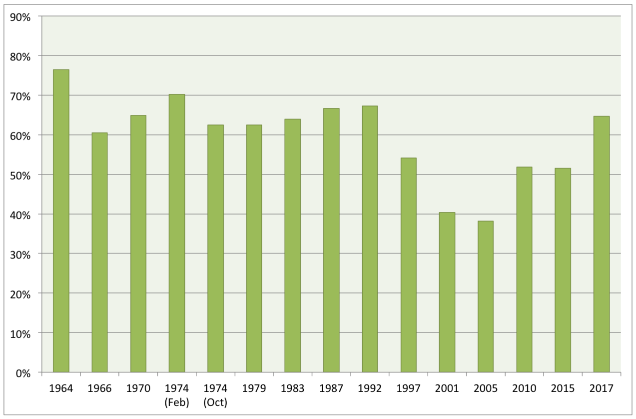
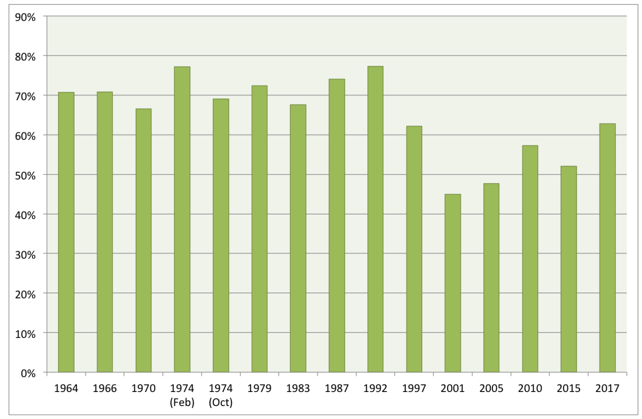


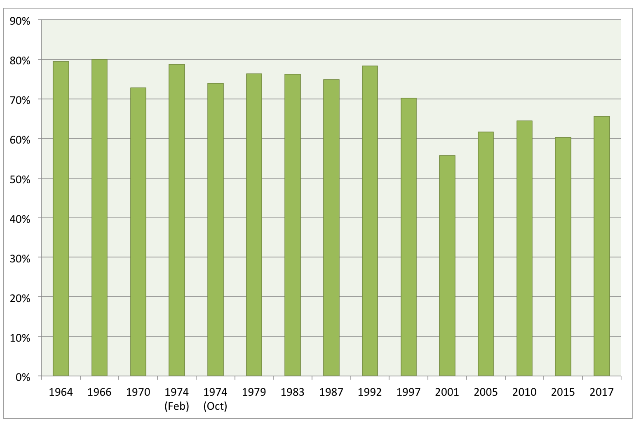
Figure 1: Turnout in general elections by UK Election Database



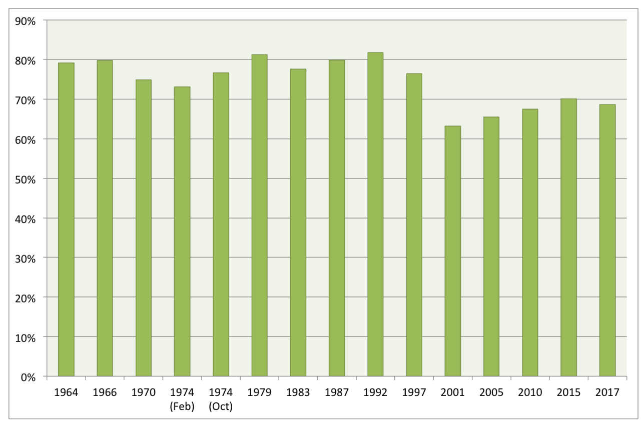
(a) 18 - 24



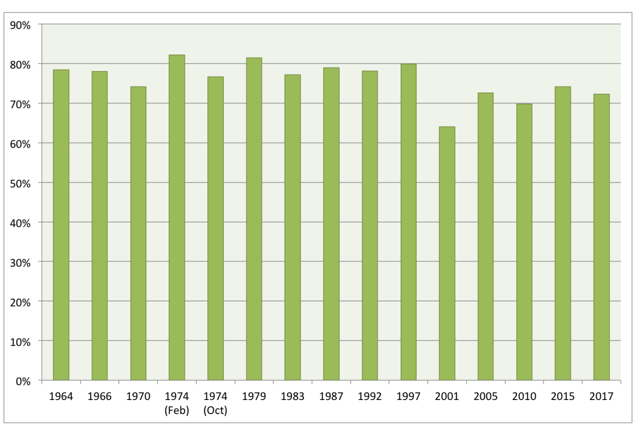
(b) 25 - 34



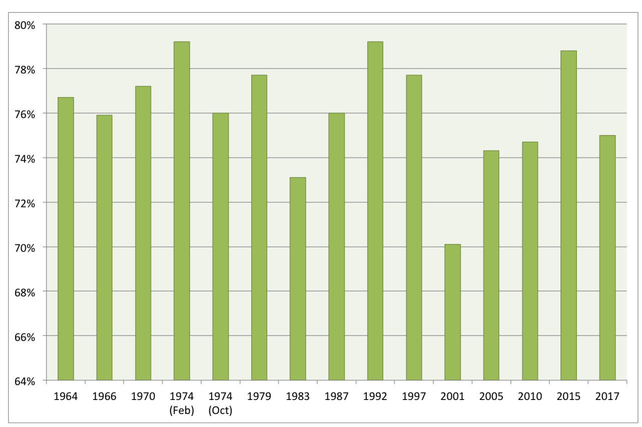
(c) 35 - 44



(d) 45 - 54



(e) 55 - 64



(f) 65+

Figure 2: Estimated turnout by age

Note: Data for 1964 and 1966 refer to people aged 21-24, as the voting age was 21 not 18.

2 Methods

2.1 Consumption voting framework

We first show how difficult it is to account for voter turnout patterns based on a conventional rational theory. What follows is along the theoretical arguments by Feddersen (2004). Suppose there are two political parties with policy platforms G and H. An individual i prefers voting for candidate $J \in \{G, H\}$ rather than abstaining if and only if

$$p_J B - C > 0 \tag{1}$$

where $p_J \in [0, 1]$ is the probability that a vote by individual i is pivotal for candidate J. $B \geq 0$ is the utility difference to individual i when the favoured candidate wins, given a benchmark utility level in the event that the other candidate wins. C is the cost of voting that includes pecuniary and psychological burdens. Obviously, in an election where many voters cast votes, the probability p_J is very small, and hence, the gross benefit of the voter $p_J B$ becomes close to zero. This implies that inequality (1) should not hold in such an electoral environment, thus failing to give a rational explanation of voting at the individual level.

To avoid the voting paradox in the rational theory, Riker and Ordeshook (1968) assume an additional benefit to the voter. They assume that the voter obtains a consumption benefit $D > 0$ from the act of voting. Now, inequality (1) is modified as

$$p_J B - C + D > 0$$

In this model set up, voters cast votes with a sufficiently large benefit $D > C$ even if $p_J B$ is zero. This is the idea of consumption voting, which is employed in further extended models, for example, Matsusaka (1995) and Becker and Mulligan (2017).

2.2 Behavioural theory

The political lifecycle theory explains generational patterns of voter turnout. For example, Parry (1992, p.154) finds that young people experience 'start-up' problems with voting due to their greater mobility, shorter spells of residence in any particular place, failure to develop an established pattern of registration and voting, and separation from strong family and/or community ties. In other words, citizens at early life stages are inclined to abstain from voting due to political apathy; however, as time goes on, life events (e.g. starting a job or forming a family) foster their attachment to civic life and facilitate democratic engagement.

Now, we argue that the political lifecycle theory can be reconstructed through the lens of the consumption voting framework described above. Specifically, we modify the framework by Becker and Mulligan (2017), and show how life events relate to electoral participation. Suppose there exists a continuum of voters with unit mass and two political parties with policy platforms G, H. These parties commit to implementing their policies if they win a given election. The utility of an individual voter i is given by:

$$U(G, H, V\{G \text{ or } H \text{ or } Z\},\ x^i, A^i_G,\ A^i_H), \quad A^i_G = e^i A_G \quad A^i_y = e^i A_H$$

where V refers to a voting process whereby individual i votes for a candidate who will implement G or H or chooses to abstain, Z. $x^i \in \{0, 1\}$ is individual i's experience of a life event. The event variable is assumed to be exogenous and $= 1$ if an individual has experienced the event and $= 0$ otherwise. A^i_G and A^i_H refer to each candidate's effective degree of political advertising (e.g. media campaigns) on individual i. A_G and A_H are the gross degrees of political advertising. The parameter $e^i \in [0, 1]$ measures the input effectiveness. This indicates that, given campaign efforts, the information input from advertising may differ between individuals.[1] If the parameter $= 0$, then regardless of the policies discussed, the individual is apathetic to electoral campaigns. Moreover, we assume that this sensitivity parameter is dependent on life events:

$$e^i = \tilde{e}(x^i), \quad \Delta\tilde{e} := \tilde{e}(x^i = 1) - \tilde{e}(x^i = 0) > 0$$

In other words, a series of life events increases one's sensitivity to campaign information.

Suppose that the outcome of an election is independent of an individual's voting decision, such that he/she receives direct consumption benefits from voting (i.e. consumption voting). The individual, therefore, seeks to maximize his or her utility by making a voting decision:

$$\max_{\{G,H,Z\}} U(V\{G \text{ or } H \text{ or } Z\}; x^i, G, H, A_G, A_H) \tag{2}$$

$$\Leftrightarrow \max_{\{G,H,Z\}} \{U_v(\text{voting for } G),\ U_v(\text{voting for } H),\ U_v(\text{choosing } Z)\} \tag{3}$$

where $U_v()$ is the marginal utility from a voting decision. According to the lifecycle theory, a life event is positively associated with the benefits of voting (i.e. $U_v(\text{voting for } G), U_v(\text{voting for } H)$) and ambiguously changes reservation utility (i.e. $U_v(\text{voting for } Z)$):

$$\frac{dU_v(\text{voting for } G)}{de^i}\Delta\tilde{e} > 0, \quad \frac{dU_v(\text{voting for } H)}{de^i}\Delta\tilde{e} > 0$$

$$\Delta U_v(\text{voting for } Z) := U_v(\text{voting for } Z, x^i = 1) - U_v(\text{voting for } Z, x^i = 0) \neq 0$$

The decision outcome is thus dependent on the magnitude of each option's marginal utility. While this model framework does not forecast the final outcome of voting decisions, it highlights how life events stimulate voters' interest in political campaigns, thus encouraging them to vote.

2.3 Panel data analysis

We empirically test the lifecycle theory which argues that electoral participation is fostered via life events. In doing so, we employ two types of binary response models: the linear probability model (LPM) and a probit model (Wooldridge (2010, Ch.15)). Although the LPM is known to provide a good approximation of covariates' partial effects, the support of the error term falls outside the unitary interval of probability. Therefore, to enhance the robustness of our estimation, we also utilize the probit model,

[1] Becker and Mulligan (2017) do not introduce the e^i parameter, because they implicitly assume an identical level of information sensitivity across individuals.

which has a functional form logically consistent with the nature of the binary variable y_{it}. In our model, subscripts $i \in \{1, 2, \cdots, N\}$ and $t \in \{0, 1, \cdots, T-1\}$ stand for individual and time indexes, respectively. In addition, $y_{it} \in \{0, 1\}$ is the dependent variable indicating whether or not an individual votes, and $\mathbf{x}_{it}$ and $\mathbf{x}_i := (\mathbf{x}_{i1}, \cdots, \mathbf{x}_{iT})$ are vectors of covariates of interest. The unobservable propensity of casting a vote is given by:

$$y_{it}^* = \mathbf{x}_{it}\boldsymbol{\beta} + \mathbf{w}_{it}\boldsymbol{\gamma} + \tau_t + c_i + \varepsilon_{it} \tag{4}$$

where $\boldsymbol{\beta}$ is the vector of parameters for the covariates of interest. $\mathbf{w}_{it}$ is a vector of control variables and $\boldsymbol{\beta}$ is its parameters. τ_t is a time-fixed effect. c_i is an unobservable individual effect. ε_{it} is a random disturbance. The superscript * is used to indicate a latent variable.

2.3.1 Linear probability model

The LPM is established by applying the zero mean assumption to the random disturbance (i.e. $\mathbf{E}\left[\varepsilon_{it} \mid \mathbf{x}_{it}, \mathbf{w}_{it}, \tau_t, c_i\right] = 0$). The response probability can be expressed by:

$$\begin{aligned} \mathbf{P}\left(y_{it} = 1 \mid \mathbf{x}_{it}, \mathbf{w}_{it}, \tau_t, c_i\right) &= \mathbf{E}\left[y_{it}^* \mid \mathbf{x}_{it}, \mathbf{w}_{it}, \tau_t, c_i\right] && (5) \\ &= \mathbf{x}_{it}\boldsymbol{\beta} + \mathbf{w}_{it}\boldsymbol{\gamma} + \tau_t + c_i && (6) \end{aligned}$$

We can estimate the parameters of the LPM using the pooled ordinary least square method. The values of the estimated parameters are then interpreted as the approximated partial effects of covariates. As for the LPM, the unobserved effect c_i is labelled as the fixed effect term. Letting $\mathbf{D}(\cdot|\,)$ denote a conditional distribution, the LPM does not require a distribution assumption on $\mathbf{D}(c_i \mid \mathbf{x}_i)$, which allows an arbitrary correlation of c_i with $\mathbf{x}_{it}$.

2.3.2 Probit model

We next establish the probit model by assuming that the random disturbance follows a standard normal distribution (i.e. $\varepsilon_{it} \sim N(0, 1)$) and that the relation between y_{it} and y_{it}^* is given by:

$$y_{it} = \begin{cases} 1 & \text{if } y_{it}^* > 0 \\ 0 & \text{if } y_{it}^* \leq 0 \end{cases} \tag{7}$$

The response probability for the probit model is then given by:

$$\begin{aligned} \mathbf{P}\left(y_{it} = 1 \mid \mathbf{x}_{it}, \mathbf{w}_{it}, \tau_t, c_i\right) &= \mathbf{P}\left(y_{it}^* > 0 \mid \mathbf{x}_{it}, \mathbf{w}_{it}, \tau_t, c_i\right) \\ &= \Phi\left(\mathbf{x}_{it}\boldsymbol{\beta} + \mathbf{w}_{it}\boldsymbol{\gamma} + \tau_t + c_i\right) && (8) \end{aligned}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. Unlike in the LPM, we cannot obtain estimates without assuming the distribution of c_i given $\mathbf{x}_i$.[2] First, we impose a strict exogeneity condition on the unobserved effect:

$$\mathbf{D}(y_{it} \mid \mathbf{x}_i, \mathbf{w}_i, \tau_t, c_i) = \mathbf{D}(y_{it} \mid \mathbf{x}_{it}, \mathbf{w}_{it}, \tau_t, c_i) \quad \text{for} \quad t = 0, \cdots, T-1 \tag{9}$$

[2]This is known as an incidental parameter problem. See Wooldridge (2010, Ch.15).

Allowing correlation between c_i and $\mathbf{x}_i$, we impose an assumption about the conditional distribution of c_i, as proposed by Chamberlain (1980):

$$c_i = \psi + \overline{\mathbf{x}_i}\boldsymbol{\xi} + \overline{\mathbf{w}_i}\boldsymbol{\eta} + a_i, \quad a_i|\mathbf{x}_i, \mathbf{w}_i \sim \text{Normal}(0, \sigma_a^2) \tag{10}$$

where ψ is a constant. $\overline{\mathbf{x}}_i$ and $\overline{\mathbf{w}}_i$ are the averages of $\mathbf{x}_{it}$ and $\mathbf{w}_{it}$ for $t = 0, \cdots, T-1$, respectively. Finally, σ_a^2 is the variance of a_i. Imposing assumptions (9) and (10) on equation (8) yields:

$$\mathbf{P}\left(y_{it} = 1 \mid \mathbf{x}_{it}, \mathbf{w}_{it}, \tau_t, c_i\right) = \Phi\left(\psi_a + \mathbf{x}_{it}\boldsymbol{\beta}_a + \overline{\mathbf{x}_i}\boldsymbol{\xi}_a + \mathbf{w}_{it}\boldsymbol{\gamma_a} + \overline{\mathbf{w}}_i\boldsymbol{\eta}_a + \tau_{at}\right) \tag{11}$$

where a subscript means that a parameter vector is multiplied by $(1 + \sigma_a^2)^{1/2}$. This is known as Chamberlain's correlated random effects probit model and can be estimated using the maximum likelihood method, specifically by maximizing the partial (pooled) log-likelihood:

$$\begin{aligned} \log \mathcal{L} = \sum_{i=1}^{N}\sum_{t=0}^{T-1} &\left\{y_{it} \log \Phi\left(\psi_a + \mathbf{x}_{it}\boldsymbol{\beta}_a + \overline{\mathbf{x}_i}\boldsymbol{\xi}_a + \mathbf{w}_{it}\boldsymbol{\gamma_a} + \overline{\mathbf{w}}_i\boldsymbol{\eta}_a + \tau_{at}\right)\right. \\ &\left. + \left(1 - y_{it}\right) \log\left[1 - \Phi\left(\psi_a + \mathbf{x}_{it}\boldsymbol{\beta}_a + \overline{\mathbf{x}_i}\boldsymbol{\xi}_a + \mathbf{w}_{it}\boldsymbol{\gamma_a} + \overline{\mathbf{w}}_i\boldsymbol{\eta}_a + \tau_{at}\right)\right]\right\} \end{aligned}$$

Otherwise, the generalized estimating equation (GEE) approach is also applicable. The GEE approach potentially improves the efficiency of estimates.

2.3.3 Dynamic panel model

Dynamic nonlinear unobserved models allow us to study the effect of a lagged dependent variable.[3] Our dynamic probit model is the variant proposed by Wooldridge (2005). The response probability for the dynamic probit model is given by:

$$\mathbf{P}\left(y_{it} = 1 \mid y_{it-1}, \cdots, y_0, \mathbf{x}_{it}, \mathbf{w}_{it}, \tau_t, c_i\right) = \Phi\left(\rho y_{it-1} + \mathbf{x}_{it}\boldsymbol{\beta} + \mathbf{w}_{it}\boldsymbol{\gamma} + \tau_t + c_i\right) \tag{12}$$

where ρ is a state dependence parameter that captures the persistence of the lagged variable. *A la* Chamberlain (1980), we assume a distribution for c_i. Specifically, c_i is given as a function of y_{i0} and other covariates:

$$c_i = \psi + \xi_0 y_{i0} + \overline{\mathbf{x}_i}\boldsymbol{\xi} + \overline{\mathbf{w}_i}\boldsymbol{\eta} + a_i, \quad a_i \sim \text{Normal}\ \left(0, \sigma_a^2\right) \tag{13}$$

Given equations (12) and (13), the likelihood function is derived as:

$$\begin{aligned} \log \mathcal{L} = \log \Bigg\{ \int_{-\infty}^{\infty} \prod_{t=0}^{T-1} &\left[\Phi\left(\rho y_{it-1} + \mathbf{x}_{it}\boldsymbol{\beta} + \mathbf{w}_{it}\boldsymbol{\gamma} + \tau_t + \psi + \xi_0 y_{i0} + \overline{\mathbf{x}_i}\boldsymbol{\xi} + \overline{\mathbf{w}_i}\boldsymbol{\eta} + a\right)\right]^{y_{it}} \\ &\left[1 - \Phi\left(\rho y_{it-1} + \mathbf{x}_{it}\boldsymbol{\beta} + \mathbf{w}_{it}\boldsymbol{\gamma} + \tau_t + \psi + \xi_0 y_{i0} + \overline{\mathbf{x}_i}\boldsymbol{\xi} + \overline{\mathbf{w}_i}\boldsymbol{\eta} + a\right)\right]^{1-y_{it}} \left(\frac{1}{\sigma_a}\right) \phi\left(\frac{a}{\sigma_a}\right) da \Bigg\} \end{aligned} \tag{14}$$

where $\psi(\cdot)$ is the density function of the standard normal distribution. The estimates are obtained by maximizing the above log-likelihood.

[3]We do not employ a dynamic LPM because various studies suggest inconsistencies in its estimation. See Pua (2015).

3 Results

3.1 Data

We next conduct estimations using data from *Harmonized BHPS-UKHLS data* which is a combination of *British Household Panel Survey (BHPS)* and its continuation survey *UK Household Longitudinal Study (UKHLS)*. The data sets are British long-term panel surveys which share similarities in terms of survey design and questions. The survey documents respondents' voting records by asking the following question: *'Did you vote in the last general election in year XXXX / month XXXX?'* The survey also documents various individual characteristics, including employment status, marital status, living stability (i.e. whether the respondent has lived at the same address for more than one year), the number of children in the household, and property ownership (i.e. whether the respondent and/or his/her partner own a home). From the entire data sets, we extract and employ six survey waves which contain a national election between two subsequent surveys: 1992 (wave 1), 1997 (wave 2), 2001 (wave 3), and 2006 (wave 4), and 2010 (wave 5) and 2015 (wave 6).[4] We exclude respondents who are ineligible to vote (i.e. under 18 years).

Table 1 provides summary statistics of the variables in our data. The variables are all dummy variables. Notice that the order of magnitude in standard deviations is overall the same as that between deviations. This suggests that the variables largely vary not only across respondents but also over time.

Table 2 shows the cross-sectional descriptive statistics by wave year. For the dummy variables, the mean values are equivalent to the fraction of respondents applicable to each attribute. Observe that voter turnout fluctuates over time. This cyclical trend is consistent with the recent turnout of the whole electorate shown in figure 1. In contrast, the mean of the other variables is inclined to be stable over time.

Table 3 re-summarizes the cross-sectional descriptive statistics by age. The respondents are grouped by every 10 years. Notice that the mean values of life events, except for retirement, increase with age. The mean of the retirement variable decreases with age. This implies that the electorate, on average, experiences more life events over time.

[4] The original samples cover 5,000 households in waves 1 - 4, and 26,000 households in waves 5-6 because of the survey renewal to UKHLS in 2009.

Table 1: Summary statistics

Variable		Mean	Std. Dev.	Min	Max	Observations
Vote	overall	0.752	0.432	0	1	N_{obs} 80,954
(1 if voted)	between		0.409			N_{ind} 40,178
	within		0.222			
Employed	overall	0.569	0.495	0	1	N_{obs} 146,499
(1 if employed)	between		0.467			N_{ind} 77,949
	within		0.217			
Retired	overall	0.225	0.418	0	1	N_{obs} 146,499
(1 if retired)	between		0.389			N_{ind} 77,949
	within		0.165			
Married	overall	0.545	0.498	0	1	N_{obs} 140,050
(1 if married)	between		0.481			N_{ind} 73,847
	within		0.177			
Living stability	overall	0.912	0.283	0	1	N_{obs} 138,082
(1 if lives for more than 3 years)	between		0.254			N_{ind} 73,069
	within		0.179			
Children	overall	0.289	0.453	0	1	N_{obs} 146,550
(1 if has a natural child)	between		0.421			N_{ind} 77,979
	within		0.206			
	overall	0.572	0.495	0	1	N_{obs} 145,650
Own house	between		0.481			N_{ind} 77,678
(1 if owns a house)	within		0.178			
Unhealthy	overall	0.171	0.377	0	1	N_{obs} 142,111
(1 if is unhealthy)	between		0.355			N_{ind} 76,031
	within		0.194			
Strong party ID	overall	0.269	0.444	0	1	N_{obs} 137,938
(1 if has a strong party identification)	between		0.398			N_{ind} 73,075
	within		0.240			
Age	overall	47.6	18.0	18	102	N_{obs} 146,541
	between		18.5			N_{ind} 77,975
	within		3.9			

[1] The between standard deviation measures variations across individuals, i.e., the standard deviation of the time-averaged individual mean.
[2] The within standard deviation measures variations across the periods of time, i.e, the standard deviation from the time-averaged individual mean.
[3] N_{obs} is the number of entire observations. N_{ind} is the number of individuals.

Table 2: Mean and standard deviation of aggregate samples

Mean (Std. Dev.) / Year (wave)	1992 (wave 1)	1997 (wave 2)	2001 (wave 3)	2006 (wave 4)	2010 (wave 5)	2015 (wave 6)
Vote	0.848	0.806	0.711	0.684	0.747	0.774
(1 if voted)	(0.359)	(0.396)	(0.453)	(0.465)	(0.435)	(0.418)
Employed	0.574	0.579	0.575	0.590	0.556	0.571
(1 if employed)	(0.494)	(0.494)	(0.494)	(0.492)	(0.497)	(0.495)
Retired	0.178	0.203	0.220	0.222	0.226	0.244
(1 if retired)	(0.382)	(0.403)	(0.414)	(0.416)	(0.418)	(0.429)
Married	0.589	0.551	0.551	0.540	0.540	0.536
(1 if married)	(0.492)	(0.497)	(0.497)	(0.498)	(0.498)	(0.499)
Living stability	0.880	0.858	0.885	0.898	0.938	0.919
(1 if lives for more than 3 years)	(0.325)	(0.349)	(0.318)	(0.302)	(0.241)	(0.273)
Children	0.296	0.292	0.305	0.287	0.296	0.273
(1 if has a child in household)	(0.457)	(0.455)	(0.460)	(0.452)	(0.456)	(0.446)
Own house	0.581	0.532	0.579	0.625	0.572	0.559
(1 if owns a house)	(0.494)	(0.499)	(0.494)	(0.484)	(0.495)	(0.496)
Unhealthy	0.087	0.104	0.110	0.097	0.222	0.202
(1 if is unhealthy)	(0.282)	(0.305)	(0.313)	(0.296)	(0.415)	(0.401)
Strong party ID	0.346	0.315	0.258	0.203	0.273	0.264
(1 if has a strong party identification)	(0.476)	(0.464)	(0.437)	(0.402)	(0.446)	(0.441)
Age	44.952	45.747	46.503	47.431	47.808	49.096
	(17.995)	(18.266)	(17.999)	(18.163)	(17.814)	(18.092)

[1] Standard deviations are shown in parenthesis.

[2] In 2013, many observations for living stability are missing due to errors in survey interviews, thus increasing the magnitude of mean and decreasing the standard deviation.

Table 3: Mean and standard deviation of aggregate samples by age

Mean (Std. Dev.) / Age	20-29	30-39	40-49	50-59	60-	All
Vote	0.571	0.695	0.766	0.817	0.864	0.738
(1 if voted)	(0.495)	(0.460)	(0.423)	(0.386)	(0.343)	(0.440)
Employed	0.664	0.773	0.804	0.722	0.286	0.668
(1 if employed)	(0.472)	(0.419)	(0.397)	(0.448)	(0.452)	(0.471)
Retired	0.000	0.000	0.003	0.058	0.619	0.115
(1 if retired)	(0.009)	(0.016)	(0.052)	(0.235)	(0.486)	(0.319)
Married	0.198	0.597	0.663	0.685	0.688	0.571
(1 if married)	(0.398)	(0.491)	(0.473)	(0.465)	(0.463)	(0.495)
Living stability	0.764	0.885	0.942	0.960	0.968	0.905
(1 if lives for more than 3 years)	(0.425)	(0.319)	(0.233)	(0.197)	(0.177)	(0.294)
Children	0.259	0.671	0.550	0.127	0.010	0.348
(1 if has a child in household)	(0.438)	(0.470)	(0.497)	(0.333)	(0.100)	(0.476)
Own house	0.203	0.556	0.674	0.712	0.732	0.577
(1 if owns a house)	(0.402)	(0.497)	(0.469)	(0.453)	(0.443)	(0.494)
Unhealthy	0.082	0.101	0.151	0.206	0.239	0.153
(1 if is unhealthy)	(0.275)	(0.301)	(0.358)	(0.405)	(0.426)	(0.360)
Strong party ID	0.195	0.211	0.242	0.281	0.346	0.252
(1 if has a strong party identification)	(0.396)	(0.408)	(0.428)	(0.449)	(0.476)	(0.434)

[1] Standard deviations are shown in parenthesis.

[2] Each column shows the age of respondents. 20 is twenties, 30 is thirties and so on.

3.2 Panel estimates

First, we examine the impacts of all life events. In doing so, we construct an indicator which simply sums up the dummy variables of life events (i.e. employed, married, living stability, children, own house). We label this indicator as the LC score. The LC score then takes the value of 0 to 6. Table 4 shows the estimates of the LC score impact on electoral participation using both linear and nonlinear binary response models. We use three specifications to examine the robustness of our findings. Time dummy variables underlie all specifications, and specification 2 and 3 include control variables (dummies for unhealthy status and strong party identification, and age). In all specifications, the estimated parameter of the LC score is significant. The results qualitatively imply that experiencing life events increases the likelihood of casting a vote. We then illustrate the qualitative impacts focusing on the third specification, that is, the linear fixed-effects model. The estimate of the third specification implies that approximately 3 percentage points increases the likelihood of voting with an extra unit increase in the LC score. However, this is based on an assumption that the marginal impact of any life event is identical, whatever the event may be, such that an extra unit increase in the LC score constantly increases the likelihood of voting. Therefore, we interpret this result as a qualitative implication that experiencing life events motivate the electorate to vote.

Next, we study the influence of individual life events on voting. Table 5 shows our estimates of the factors affecting electoral participation. Specifications 2 to 6 include control variables for health status, strong party identification and age. In the first specification, four variables—marriage, living stability, children, and home ownership—are all significant and positively associated with electoral participation. This finding is robust even after controlling for unobserved individual effects with linear (third specification) and nonlinear parameterizations (sixth specification). Thus, the above life events increase the likelihood of voting. The employment variables, employed and retired, are weakly significant and positively associated with electoral participation in the first specification. However, this is no longer statistically significant after controlling for unobserved individual effects in the third and sixth specifications. We argue that employment status would seemingly affect voting in cross-sectional analysis, but this would be caused by omission bias of unobserved individual effects.

We conduct a robustness check to remove a potential sampling bias. As indicated in Table 1, some individuals refused to give a response with respect to their vote. This may generate unknown sampling bias. For our purpose, we employ a matching method proposed by Imai et al. (2019). This matching method is designed for panel data, which allows to find control observations with an identical history up to the pre-specified number of lags. We can assess the quality of matching by checking covariate balance. Table 6 shows the results. The estimates of married, living stability, and home ownership are all statistically significant, which is consistent with the results with linear and nonlinear regressions. Having children is not statistically significant presumably because of sampling bias. See Appendix A for more details of the panel matching method.

Table 4: Effects of life event scores on electoral participation

	Spec 1 Linear No control variables	Spec 2 Linear	Spec 3 Linear
	OLS	OLS	Fixed effects
LC score	0.0457***	0.0453***	0.0292***
	(0.00129)	(0.00121)	(0.00197)
Constant	0.674***	0.465***	0.888***
	(0.00637)	(0.00678)	(0.0921)
N	80,535	79,606	79,606

[1] Standard errors are shown in parentheses.
[2] The p-values are shown with the following symbols: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$
[3] All models are associated with time dummies. Spec 2 and 3 include control variables (dummies for unhealthy status and strong party ID, and age) which estimates are suppressed.

Table 5: Effects of lifec events on turnout

	Spec 1 Linear No control variables	Spec 2 Linear	Spec 3 Linear	Spec 4 Probit	Spec 5 CRE Probit	Spec 6 CRE Probit
	OLS	OLS	Fixed effects	Pooled MLE	MLE	GEE
Employed (1 if employed)	0.0299*** (0.00479)	0.0297*** (0.00463)	-0.0108 (0.00633)	0.0755*** (0.0152)	-0.0291 (0.0323)	-0.0191 (0.0211)
Retired (1 if retired)	0.126*** (0.00539)	0.00181 (0.00615)	-0.0117 (0.00737)	0.0380 (0.0253)	-0.0879 (0.0496)	-0.0565 (0.0313)
Married (1 if married)	0.0975*** (0.00389)	0.0701*** (0.00379)	0.0377*** (0.00672)	0.247*** (0.0143)	0.195*** (0.0357)	0.126*** (0.0234)
Living stability (1 if lives for more than 3 years)	0.122*** (0.00585)	0.0907*** (0.00574)	0.0904*** (0.00703)	0.278*** (0.0171)	0.448*** (0.0329)	0.297*** (0.0216)
Children (1 if has a child in household)	-0.0677*** (0.00416)	-0.0291*** (0.00415)	0.0161** (0.00562)	-0.119*** (0.0139)	0.0943** (0.0302)	0.0580** (0.0193)
Own house (1 if owns a house)	0.112*** (0.00406)	0.0862*** (0.00396)	0.0249*** (0.00659)	0.296*** (0.0142)	0.145*** (0.0344)	0.0968*** (0.0224)
Constant	0.598*** (0.00724)	0.494*** (0.00740)	0.887*** (0.0923)	-0.0268 (0.0251)	-0.641*** (0.0822)	-0.412*** (0.0537)
N	80,535	79,606	79,606	79,606	79,145	79,145

[1] Standard errors are shown in parentheses.
[2] The p-values are shown with the following symbols: $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$
[3] All models are associated with time dummies. Spec 2 - 6 include additional control variables (dummies for unhealthy status and strong party ID, and age). Those estimates are suppressed.

Table 6: Panel matching estimates

	Panel Matching	
Employed	0.000	
	(0.018)	
Retired	0.011	
	(0.010)	
Married	0.045	***
	(0.016)	
Living stability	0.175	***
	(0.031)	
Children	-0.011	
	(0.014)	
Own house	0.041	***
	(0.013)	

[1] Standard errors are shown in parentheses.
[2] The p-values are shown with the following symbols: $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

We then conduct another robustness check, given a habituation hypothesis. The habituation hypothesis is an argument that an individual who went to vote in the last period is more likely to go to vote in the current period because a habituation of electoral participation is fostered by repetitive actions.[5] Table 7 presents cross-tabulations between electoral participation in two subsequent periods. The diagonal elements, listed in bold, highlight the persistence of voting decisions. We then perform dynamic panel estimations. Table 8 shows the results. The seventh specification includes unobserved individual effects, whereas the eighth specification does not. In the results of the seventh specification, the lagged variable for votes is significant with a 0.55 coefficient, which implies past voting has a strong persistent effect. Since this is a dummy variable, its effects on current voting behaviour are valid only when a respondent has voted in the last survey period. In the eighth specification, we drop the unobserved effect term. Here, the coefficient of the lagged variable is 1.232, which intensifies the impact of voting habituation. This implies that much of the persistence of voting is due to unobserved individual effects. The seventh specification also shows that marriage and living stability are significant and positively associated with electoral participation. This is consistent with previous results shown in Table 5. The estimates of the employment variables are weakly significant and negatively associated with electoral participation which is not a robust result. The estimates of children and home ownership turned out to

[5] See Denny and Doyle (2009) for the relevant literature.

be insignificant. This may be because of the replacement of samples. The responses of individuals who participated in the survey only for any two subsequent waves may overestimate the impacts of the lagged variable. We need to conduct further investigation in regard to this issue.

Table 7: Transitions in voter participation

		Vote in wave 2 (1997)		
		0 (No)	1 (Yes)	Total
Vote in wave 1 (1992)	0	**456**	424	880
	(No)	**52%**	48%	100%
	1	583	**5,214**	**5,797**
	(Yes)	10%	**90%**	**100%**
	Total	1,039	5,638	6,677
		16%	84%	100%

		Vote in wave 5 (2010)		
		0 (No)	1 (Yes)	Total
Vote in wave 4 (2006)	0	**831**	689	1,520
	(No)	**55%**	45%	100%
	1	381	**3,294**	3,675
	(Yes)	10%	**90%**	100%
	Total	1,212	3,983	5,195
		23%	77%	100%

		Vote in wave 3 (2001)		
		0 (No)	1 (Yes)	Total
Vote in wave 2 (1997)	0	**898**	415	1,313
	(No)	**68%**	32%	100%
	1	1,049	**5,051**	6,100
	(Yes)	17%	**83%**	100%
	Total	1,947	5,466	7,413
		26%	74%	100%

		Vote in wave 6 (2015)		
		0 (No)	1 (Yes)	Total
Vote in wave 5 (2010)	0	**1,065**	797	1,862
	(No)	**57%**	43%	100%
	1	612	**6,386**	6,998
	(Yes)	9%	**91%**	100%
	Total	1,677	7,183	8,860
		19%	81%	100%

		Vote in wave 4 (2006)		
		0 (No)	1 (Yes)	Total
Vote in wave 3 (2001)	0	**1,898**	974	2,872
	(No)	**66%**	34%	100%
	1	1,141	**6,917**	8,058
	(Yes)	14%	**86%**	100%
	Total	3,039	7,891	10,930
		28%	72%	100%

		Vote in t		
		0 (No)	1 (Yes)	Total
Vote in $t-1$	0	**5,148**	3,299	8,447
	(No)	**61%**	39%	100%
	1	3,766	**26,862**	30,628
	(Yes)	12%	**88%**	100%
	Total	8,914	30,161	39,075
		23%	77%	100%

[1] Integers present the number of observations.
[2] Percentages present the proportions of observations.
[3] The last bottom-right subtable shows the transitions of the entire samples.

Table 8: Effects of life events and the lagged vote y_{it-1}

	Spec 7 Dynamic Wooldridge Probit	Spec 8 Dynamic Probit without c_i
	MLE	Pooled MLE
Employed	-0.131*	0.0788***
(1 if employed)	(0.0602)	(0.0237)
Retired	-0.0660	0.0805*
(1 if retired)	(0.0783)	(0.0342)
Married	0.347***	0.192***
(1 if married)	(0.0628)	(0.0183)
Living stability	0.548***	0.294***
(1 if lives for more than 3 years)	(0.0586)	(0.0290)
Children	0.0382	-0.0984***
(1 if has a child in household)	(0.0503)	(0.0200)
Own house	0.0569	0.179***
(1 if owns a house)	(0.0605)	(0.0188)
y_{t-1}	0.550***	1.232***
(1 if voted in the last period)	(0.0479)	(0.0201)
Constant	-1.562***	-1.011***
	(0.147)	(0.0498)
N	20,033	38,488

[1] Standard errors are shown in parentheses.
[2] The p-values are shown with the following symbols: $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$
[3] All models are associated with time dummies, control variables (dummies for unhealthy status and strong party ID, and age) which estimates are suppressed.

Table 9 presents the computed average partial effects (APE) for three specifications. The last column presents the conditional probability of casting a vote with $\mathbf{x} = \mathbf{0}$ (except for the constant term). We use this conditional probability as a benchmark of a case where an individual has no experience of any of the life events measured. While the significant APE values range from 0.01 to 0.09, which seems small, the relative impacts are remarkable. Observe that the percentages shown below the APE values range from 2 to 13 percentage points. If an individual experiences all significant life events (i.e. marriage, living stability, children, home ownership), then the total effect is 25 % for the third specification and 18 % for the sixth specification. In the seventh specification, the total effect of marriage and living stability is 46 %, comparable to the effect of the lagged variable (42 %). In sum, our results show that life events significantly increase the likelihood of voting and that voting in the previous election also positively affects the likelihood of voting in the current election.

We make additional robustness check exercises with another set of longitudinal data, NCDS (*National Children Development Study*). NCDS is a British long-term panel survey tracking a cohort of people born in 1958. The survey waves cover periods during which respondents are eligible to vote: 1981 (cohort = 23 years old), 1991 (33 years), 2000 (42 years), 2008 (50 years), and 2013 (55 years). Table 10 shows the main results. The results are overall similar to the results with UKLHS which support the positive impacts of life events on voting. Having children is statistically significant even with the dynamic panel specification. See Appendix B for the descriptive statistics of NCDS data.

Table 9: Average partial effects

	Spec 3 Linear		Spec 6 CRE Probit		Spec 7 Dynamic Wooldridge Probit	
	Fixed effects		GEE		MLE	
Employed	-0.0108		-0.00544		-0.0231	**
(1 if employed)	-2%		-1%		-7%	
Retired	-0.0117		-0.01604	*	-0.0127	
(1 if retired)	-2%		-2%		-4%	
Married	0.0377	***	0.033207	***	0.0611	***
(1 if married)	6%		4%		18%	
Living stability	0.0904	***	0.07582	***	0.093	***
(1 if lives for more than 3 years)	13%		9%		28%	
Children	0.0161	**	0.01626	***	0.00861	
(1 if has a child in household)	2%		2%		3%	
Own house	0.0249	***	0.02446	***	0.0123	
(1 if owns a house)	4%		3%		4%	
y_{t-1}					0.1426	***
(1 if voted in the last period)					42%	
$P(y = 1 \mid \hat{\boldsymbol{\beta}})$	0.681		0.800		0.336	

[1] $P(y = 1 \mid \hat{\boldsymbol{\beta}})$ is the probability of casting a vote computed with the estimates, conditional on $\mathbf{x} = 0$ (except for the constant coefficient).

[2] The percentages are the ratio of the APE to $P(y = 1 \mid \hat{\boldsymbol{\beta}})$.

Table 10: Panel estimates with NCDS data

	Spec 2 Linear	Spec 3 Linear	Spec 6 CRE Probit	Spec 7 Dynamic Wooldridge Probit
	OLS	Fixed effects	GEE	MLE
Employed	0.0276***	0.00794	0.0186	-0.00847
(1 if employed)	(0.00648)	(0.00723)	(0.0212)	(0.0409)
Married	0.0220**	0.0396***	0.111***	0.299***
(1 if married)	(0.00674)	(0.00779)	(0.0227)	(0.0428)
Living stability	0.0552***	0.0488***	0.147***	0.0936**
(1 if lives for more than 3 years)	(0.00558)	(0.00575)	(0.0174)	(0.0289)
Children	-0.0147**	0.0176**	0.0482*	0.265***
(1 if has a child in household)	(0.00568)	(0.00644)	(0.0188)	(0.0342)
Own house	0.0899***	0.0330***	0.101***	0.0475
(1 if owns a house)	(0.00663)	(0.00808)	(0.0230)	(0.0456)
y_{t-1}				0.543***
(1 if voted in the last period)				(0.0367)
Constant	0.483***	0.621***	-0.577***	-1.176***
	(0.0154)	(0.0166)	(0.0734)	(0.147)
N	40,724	40,724	40,724	25,369

[1] Standard errors are shown in parentheses.
[2] The p-values are shown with the following symbols: $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$
[3] All models are associated with time dummies, control variables (a dummy for unhealthy status, income and age) which estimates are suppressed.

4 Concluding remarks

Our empirical analysis demonstrates that life events significantly and positively affect the electoral participation of adult British people. It also suggests that the impacts of those events are associated with voting participation in the last wave. In future research, we aim to improve our robustness checks. First, we will include additional control variables. Residential addresses of respondents would be crucial in our analysis. Since voter turnout may differ across regions for various reasons, we need to control for the geographical factor in our regression analysis. We will review the original variables in two different data sets (i.e. UKLHS and NCDS) and seek to get complete identical control variables. Second, we will carefully check the data attribution of UKLHS and examine how often responses are inadequate, especially for the variable of vote, over the entire waves. Our results with the dynamic panel model do not support the significant impacts of having children and home ownership on electoral participation. However, this may be because many respondents participated in only two waves. If this is the case, the voting experience in the last wave may be sufficient to account for the vote in the present wave, thus underestimating the impacts of life events. Narrowing down samples to just respondents who certainly participated in the survey for multiple waves, we will perform a re-estimation. Lastly, we will examine the heterogeneous effects of life events in terms of gender and generations. It would be possible that impacts of life events on electoral participations vary between men and women or between cohorts born in different ages. We seek to obtain further findings in this regard by conducting panel estimations with sub-samples.

References

Aldrich, John H., Jacob M. Montgomery, and Wendy Wood, "Turnout as a Habit," *Political Behavior*, 2011, *33* (4), 535–563. 3

Becker, Gary S. and Casey B. Mulligan, "Is Voting Rational or Instrumental?," in "Explorations in Public Sector Economics," Springer, Cham, 2017, pp. 1–11. 6, 7

Bendor, Jonathan, Daniel Diermeier, and Michael Ting, "A Behavioral Model of Turnout," *American Political Science Review*, 2003, *97* (2), 261–280. 3

Chamberlain, Gary, "Analysis of Covariance with Qualitative Data," *The Review of Economic Studies*, 1980, *47* (1), 225–238. 9

Cooley, Thomas F. and Jorge Soares, "A Positive Theory of Social Security Based on Reputation," *Journal of Political Economy*, 1999, *107*, 135–160. 2

Denny, Kevin and Orla Doyle, "Does Voting History Matter? Analysing Persistence in Turnout," *American Journal of Political Science*, 2009, *53* (1), 17–35. 3, 17

Downs, Anthony, *An Economic Theory of Democracy*, Harper and Row, 1957. 1

Feddersen, Timothy J., “Rational Choice Theory and the Paradox of Not Voting,” *Journal of Economic Perspectives*, 2004, *18* (1), 99–112. 6

_ **and Wolfgang Pesendorfer**, “The Swing Voter’s Curse,” *The American Economic Review*, 1996, *86* (3), 408–424. 3

Flanagan, C. A. and L. R. Sherrod, “Youth political development: An introduction,” *Journal of Social Issues*, 1998, *54* (3), 447–456. 3

Gonzalez-Eiras, Martin and Dirk Niepelt, “The future of social security,” *Journal of Monetary Economics*, 2008, *55* (2), 197–218. 2

Green, Donald P. and Ron Shachar, “Habit Formation and Political Behaviour: Evidence of Consuetude in Voter Turnout,” *British Journal of Political Science*, 2000, *30* (4), 561–573. 3

Highton, B., “Residential mobility, community mobility, and electoral participation,” *Political Behavior*, 2000, *22* (2), 109–120. 3

_ **and W. E. Wolfinger**, “The first seven years of the political life cycle,” *American Journal of Political Science*, 2001, *45* (1), 202–209. 3

Imai, Kosuke, In Song Kim, and Erik Wang, “Matching Methods for Causal Inference with Time-Series Cross-Sectional Data,” *mimeo*, 2019. 2, 14, 27

Jankowski, Thomas B. and John M. Strate, “Modes of participation over the adult life span,” *Political Behavior*, 1995, *17* (1), 89–106. 3

Matsusaka, John G., “Explaining voter turnout patterns: An information theory,” *Public Choice*, 1995, *84*, 91–117. 6

Parry, Geraint, *Political Participation and Democracy in Britain*, Cambridge England ; New York: Cambridge University Press, 1992. 2, 6

Piggott, John and Alan Woodland, *Handbook of the Economics of Population Aging, Volume 1B*, North Holland, 2017. 2

Plutzer, Eric, “Becoming a Habitual Voter: Inertia, Resources, and Growth in Young Adulthood,” *The American Political Science Review*, 2002, *96* (1), 41–56. 3

Pua, Andrew Adrian Yu, “On IV estimation of a dynamic linear probability model with fixed effects,” *UvA-Econometrics Working Papers*, 2015, *No 15-01.* 9

Riker, William H. and Peter C. Ordeshook, “A Theory of the Calculus of Voting,” *The American Political Science Review*, 1968, *62* (1), 25–42. 2, 6

Smets, Kaat, “Revisiting the political life-cycle model: later maturation and turnout decline among young adults,” *European Political Science Review*, 2016, *8* (2), 225–249. 3

— **and Anja Neundorf**, "The hierarchies of age-period-cohort research: Political context and the development of generational turnout patterns," *Electoral Studies*, 2014, *33*, 41–51. 3

Stoker, L. and Mk Jennings, "Life-Cycle Transitions and Political-Participation - the Case of Marriage," *American Political Science Review*, 1995, *89* (2), 421–433. 3

Wooldridge, Jeffrey M., "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity," *Journal of Applied Econometrics*, 2005, *20* (1), 39–54. 9

— , *Econometric Analysis of Cross Section and Panel Data, 2nd Edition*, MIT Press, 2010. 7, 8

Appendix A Panel Matching Method

We use the matching method for five life event variables: employment, marriage, living stability, having children, and home ownership. In doing so, we first find treatment and control observations with an identical history of treatment up to the one lagged period. Given the matched samples, we then conduct refinement by finding observations which have similar features in terms of the observed confounders. As for the confounders, we use all life event variables, save for the treatment variable of interest, its lagged variables, and the variables of health status, party identification, and age. Finally, we confirm the quality of matching by checking covariate balance. Figure 3 to 7 show how the covariate balance changes before and after the observation refinement. In each figure, a single X marking corresponds to a confounder at the present or one lagged period. Observe that the values of the covariate balance shrink after the refinement. This indicates that we successfully narrow down samples with similar features after the refinement. See Imai et al. (2019) for the technical operations of matching and refinement.

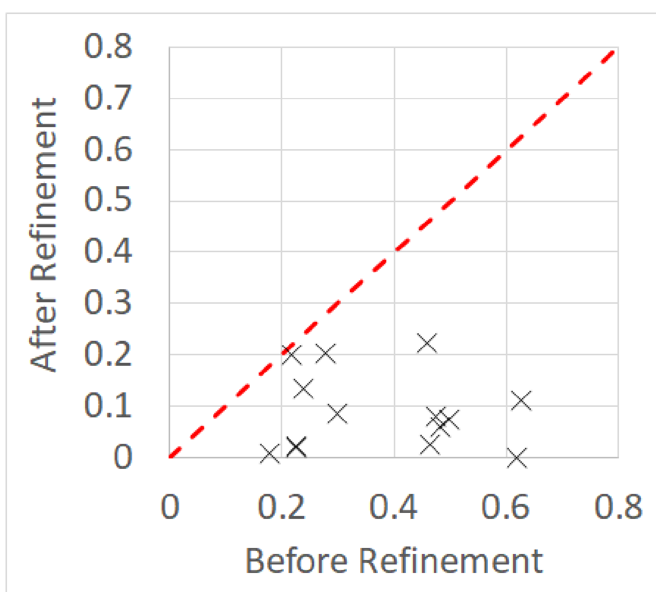


Figure 3: Covariate balance for the variable of employed

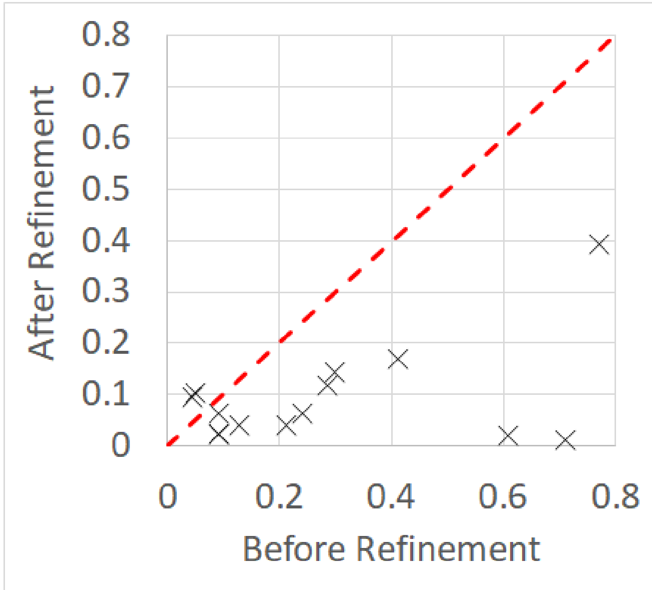


Figure 4: Covariate balance for the variable of married

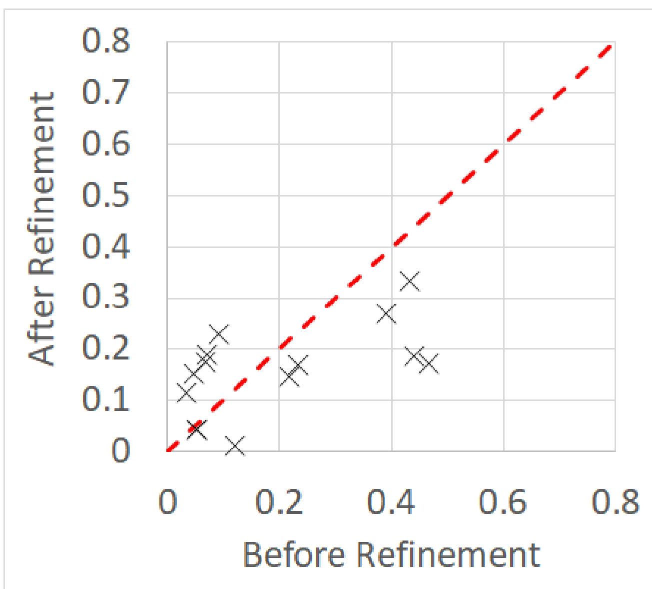


Figure 5: Covariate balance for the variable of living stability

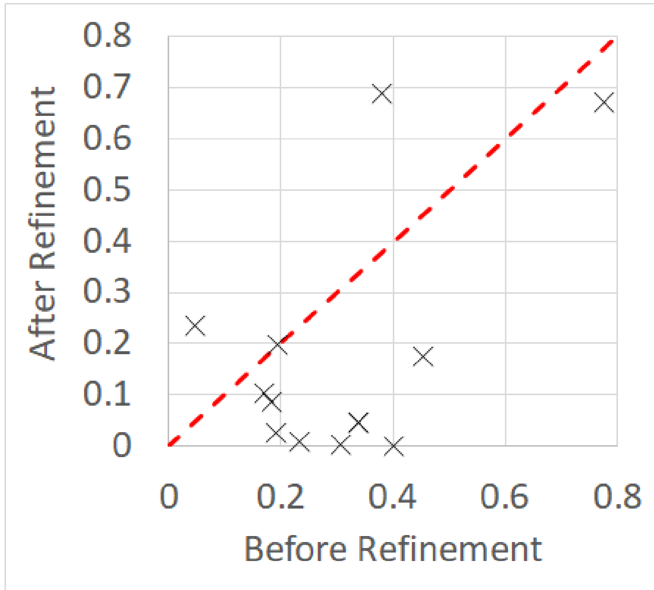


Figure 6: Covariate balance for the variable of children

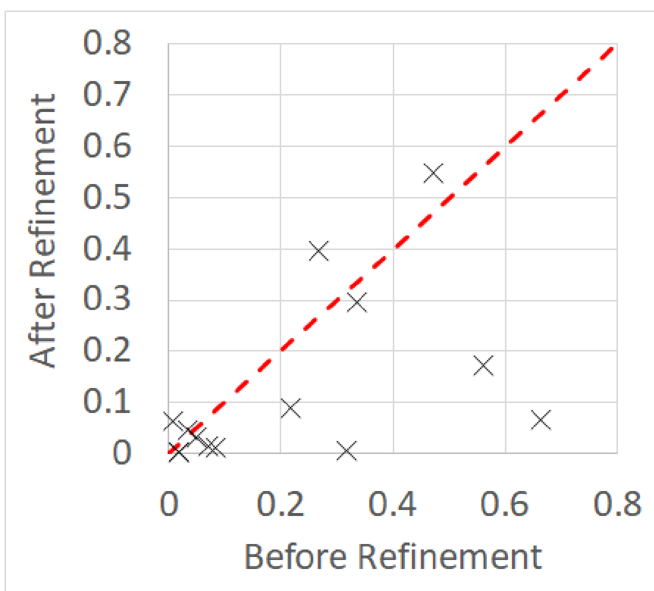


Figure 7: Covariate balance for the variable of own house

Appendix B Descriptive statistics of NDCS data

Table 11: Summary statistics of NCDS data

Variable		Mean	Std. Dev.	Min	Max	Observations
Vote	overall	0.733	0.443	0	1	N = 53579
(1 if voted)	between		0.355			n = 14778
	within		0.306			
Employed	overall	0.803	0.398	0	1	N = 54017
(1 if employed)	between		0.318			n = 14839
	within		0.285			
Married	overall	0.684	0.465	0	1	N = 54033
(1 if married)	between		0.366			n = 14838
	within		0.325			
Living stability	overall	0.633	0.482	0	1	N = 48884
(1 if lives for more than 3 years)	between		0.315			n = 14368
	within		0.407			
Natural children	overall	0.521	0.500	0	1	N = 54289
(1 if has a natural child)	between		0.363			n = 14853
	within		0.371			
Own house	overall	0.665	0.472	0	1	N = 50252
(1 if owns a house)	between		0.373			n = 14795
	within		0.333			
Unhealthy	overall	0.045	0.208	0	1	N = 53943
(1 if has above college degree)	between		0.152			n = 14828
	within		0.155			
Log (family income)	overall	9.407	1.821	0	15.799	N = 49189
	between		1.433			n = 14784
	within		1.447			

[1] The logarithmic value of family income is adjusted so that it equals zero if the family income is less than one pound.

[2] The between standard deviation measures variations across individuals, i.e., the standard deviation of the time-averaged individual mean.

[3] The within standard deviation measures variations across the periods of time, i.e, the standard deviation from the time-averaged individual mean.

Table 12: NCDS samples by year

Mean (Std. Dev.) / Year (age)	1981(23)	1991(33)	2000(42)	2008(50)	2013(55)
Vote	0.664	0.769	0.770	0.731	0.737
(1 if voted)	(0.472)	(0.422)	(0.421)	(0.444)	(0.440)
Employed	0.737	0.790	0.845	0.846	0.812
(1 if employed)	(0.440)	(0.407)	(0.361)	(0.361)	(0.390)
Married	0.446	0.797	0.801	0.691	0.717
(1 if married)	(0.497)	(0.402)	(0.400)	(0.462)	(0.451)
Living stability	0.159	0.435	0.763	0.886	0.993
(1 if lives for more than 3 years)	(0.366)	(0.496)	(0.425)	(0.317)	(0.082)
Natural children	0.253	0.667	0.718	0.561	0.420
(1 if has a natural child in household)	(0.435)	(0.471)	(0.450)	(0.496)	(0.494)
Own house	0.298	0.729	0.806	0.829	0.764
(1 if owns a house)	(0.457)	(0.444)	(0.395)	(0.377)	(0.425)
Unhealthy	0.009	0.073	0.036	0.057	0.060
(1 if is unhealthy)	(0.094)	(0.260)	(0.185)	(0.232)	(0.238)
Log [familiy income]	7.542	9.292	10.100	10.459	10.746
	(1.781)	(1.238)	(1.250)	(1.001)	(1.410)

[1] Standard deviations are shown in parenthesis.