Article:

Spearing, Joe Mead (2024) The effect of retirement eligibility on mental health in the United Kingdom:Heterogeneous effects by occupation. Health Economics. pp. 1621-1648. ISSN: 1057-9230

https://doi.org/10.1002/hec.4835

Received: 23 July 2023 | Revised: 11 March 2024 | Accepted: 10 April 2024

DOI: 10.1002/hec.4835

SHORT RESEARCH ARTICLE

Health Economics Letters

Health Economics WILEY

The effect of retirement eligibility on mental health in the United Kingdom: Heterogeneous effects by occupation

Joe Spearing

The University of Georgia, Athens, Georgia, USA

Correspondence
Joe Spearing.
Email: joe.spearing@uga.edu

Abstract

I investigate heterogeneity across occupational characteristics in the effect of retirement eligibility on mental health in the United Kingdom. I use K-means clustering to define three occupational clusters, differing across multiple dimensions. I estimate the effect of retirement eligibility using a Regression Discontinuity Design, allowing the effect to differ by cluster. The effects of retirement eligibility are beneficial, and greater in two clusters: one comprised of white-collar jobs in an office setting and another of blue-collar jobs with high physical demands and hazards. The cluster with smaller benefits mixes blue- and white-collar uncompetitive jobs with high levels of customer interaction. The results have implications for the distributional effect of raising the retirement age.

KEYWORDS
mental health, occupation, pensions, retirement

1 | INTRODUCTION

Due to the fiscal challenges of population ageing, many governments have increased the age of retirement eligibility. The distributional effect of these policy changes depends on how the health effects of retirement and retirement eligibility differ across occupations. I link the British Household Panel Survey/Understanding Society to O*NET, which contains detailed information about occupational characteristics. I use K-means clustering to define three occupation "clusters" and assess how the short-run effect of retirement eligibility on mental health differs across them using a Regression Discontinuity Design (RDD).

I find that there are two occupation clusters where workers' mental health improves more than the population average from retirement eligibility. The first is over-represented amongst skilled agricultural workers and craft and related trades, and includes mostly blue-collar occupations characterized by wearing Personal Protective Equipment (PPE), active body positioning, uncomfortable conditions, and job hazards. The second cluster with high benefits of

I am grateful to Meghan Skira, Joshua Kinsler, Matthew Knepper, Carolina Caetano, and Abigail Cormier for their time, feedback and valuable input, and participants in the University of Georgia Economics seminar series and ASHEcon for their constructive comments. Additionally, I am grateful for comments from two anonymous referees which helped substantially improve this paper. Conversations with Cormac O'Dea and Brent Davis were also valuable.

retirement eligibility is over-represented amongst senior officials, professionals, associate professionals and clerks. This cluster is characterized by sedentary body positioning, high levels of communication, comfortable conditions and few job hazards, involves working with people, and is routine, competitive, and mostly white-collar. The mental health benefits of retirement eligibility are around half the population average for one cluster, which is over-represented in elementary occupations and service workers, and mixes blue- and white-collar jobs. These occupations score the lowest for tight scheduling, highest for conflictual contact, and are the least competitive and routine.

My research relates to the literature estimating the effects of retirement eligibility and retirement on health outcomes (e.g., Geyer et al., 2020; Gorry et al., 2018; Johnston & Lee, 2009; Müller & Shaikh, 2018; Neuman, 2008; Nishimura et al., 2018; Rose, 2020), including heterogeneous effects by occupation.[1] Most studies disaggregate the effect by one occupational characteristic, such as blue- and white-collar work (Kolodziej & García-Gómez, 2019), straining and non-straining jobs (Mazzonna & Peracchi, 2017), and physically demanding and less demanding jobs (Leckcivilize and McNamee, 2021). Research considering multiple occupational characteristics mostly considers one characteristic at a time (e.g., Eibich, 2015; Heller-Sahlgren, 2017). An exception is Carrino et al. (2020), who consider heterogeneity by levels of control and levels of physical/psychological demand simultaneously.

The novelty of this paper lies in employing K-means clustering to define occupational types, using rich data on occupational characteristics while maintaining precise RDD estimates. I contribute, firstly, by providing a richer characterization of the types of jobs where gaining eligibility to retire is more beneficial. Studies often examine effect heterogeneity across only one particular aspect of work (e.g., physical demands), but they cannot rule out that results are driven by other occupational characteristics that are correlated with the characteristic under consideration. My approach does not suffer from this concern. Secondly, by using a data-driven process to define occupational groupings, I do not impose ad hoc assumptions on the occupations that are restricted to have the same mental health effects. For example, previous literature has investigated whether being eligible to retire from white-collar and blue-collar jobs has differential effects (e.g., Kolodziej & García-Gómez, 2019). My approach separates white- and blue-collar jobs which have different characteristics. There are elevated benefits from becoming eligible to retire in some blue-collar jobs but not others; and some white-collar jobs but not others.

2 | DATA

I use a restricted version of the British Household Panel Survey/Understanding Society (BHPS/US),[2] which provides date of birth in months and occupation at the 4-digit ISCO-88 level, and O*NET.[3]

BHPS/US is a representative survey of the UK population with data on retirement status, household income (which I express in real equivalized log points) and the General Health Questionnaire (GHQ12). I use the GHQ12 to construct the GHQ12 caseness (McCabe et al., 1996), an index of mental health symptoms, and its sub-indices which measure the severity of symptoms of Anxiety and Depression, Loss of Confidence, and Social Dysfunction (Graetz, 1991). Higher values of mental health variables reflect worse mental health. I also consider a dummy variable for the caseness score being higher than 8, a key screening threshold (Anjara et al., 2020). Appendix A details the construction of these indices. I use date of birth in months to construct expected age in months.[4]

The 57 O*NET "Work Context" variables report activities performed at work and the settings they are performed in. The O*NET survey interviews employees or experts in particular occupations. Respondents give a score on each variable for their occupation from an ordered list. The Department of Labor (DoL) reports a summary measure for each variable for each occupation, which I use as a measure of each occupation's score on each variable. Table 1 shows data on CEOs' levels of contact at work in the August 2020 wave of the O*NET data. Respondents choose from 5 ordered categories. The measure of CEOs' "contact at work" is calculated by assigning each response a number and calculating an "average" response:

$$(0 \times 1) + (0 \times 2) + (0.0268 \times 3) + (0.1483 \times 4) + (0.8249 \times 5) = 4.8 \tag{1}$$

For describing variation across occupational clusters, I aggregate variables using the DoL's "O*NET model",[5] which categorizes the 57 variables into 11 groups. I compute 11 indices as the average of the variables within them, choosing signs for each variable to give the index an intuitive economic meaning. I label the index to make this meaning explicit. For example, in the body positioning index, I assign a positive weight to sedentary body positioning (like sitting) and a

TABLE 1 O*NET measure of contact with people at work for CEOs, August 2020.

Category	% of respondents
No contact with others	0
Occasional contact with others	0
Contact with others about half the time	2.68
Contact with others most of the time	14.83
Constant contact with others	82.49

Note: Data are from O*NET. The O*NET study interviews either employees or experts in a given occupation and asks them to score the regularity of each variable within that occupation.

negative weight to active body positioning (like standing) so the overall index is a measure of how sedentary body positioning is in that occupation. Table 2 presents a list of the indices, their constituent variables, and the sign of each variable. I use the economic interpretation of the indices in parentheses in the remainder of the paper. Each occupation's score on each O*NET variable is its average score over the period.[6]

While I study the effects of retirement eligibility on health outcomes in the United Kingdom, the O*NET data are collected in the United States. I therefore rely on similarity of the content of occupations between the two countries. Other researchers have applied the O*NET data to the European context (e.g., Goos et al., 2014; Hardy, Keister, and Lewandowski, 2018; Lewandowski, 2020), including the UK (Jolivet & Postel-Vinay, 2020). Applying the O*NET data to the UK is appropriate for several reasons: firstly, the UK and the US are at similar stages of economic development, with similar levels of education[7] and access to technology. These determinants of working conditions are therefore roughly constant. Secondly, there is significant overlap between firms' management in the US and the UK (Edwards et al., 2010) and we might therefore expect management practices to be similar. Thirdly, when surveys similar to O*NET have been conducted in European countries, results have been closely correlated with those in the US (e.g., CEDEFOP, 2013; Handel, 2012). Hardy, Keister and Lewandowski (2018) provide further discussion of the applicability of O*NET to a European context.

I link the O*NET data to the BHPS/US data by current or last occupation at the 4-digit level using crosswalks provided by Hardy (2016).[8] My sample includes observations of people within 15 years of retirement eligibility who report a current or last occupation at least once within the survey, and which report values for all mental health variables and retirement status. I use this sample for descriptive analysis of the differences across people working in different types of occupations. Panel A of Table 3 presents summary statistics. There are 139,762 observations in the sample, although some observations have missing data on log household income. For each variable, I report the sample mean, standard deviation, and minimum and maximum values, as well as the mean in the full data set (i.e., including observations of respondents who are not within 15 years of retirement). In my main RDD estimates, I use a 5-year bandwidth. I present summary statistics for people within 5 years of retirement eligibility in Panel B of Table 3. Both samples are healthier than the data set as a whole on every measure of mental health. A greater share are retired, and male. They have roughly the same log household income.

3 | ESTIMATION

I define occupational groups using K-means clustering (Bonhomme et al., 2022; Steinley, 2006). I define a set of "centroids", each with a value for each O*NET variable, and assign each data point (work context scores for each person-wave observation) to a centroid, minimizing the total Euclidian distance between data points and the centroids they are assigned to. A "cluster" is a set of data points assigned to the same centroid.

For a given number of clusters N, I solve:

$$\min_{\{k_1,..k_I\}\in\{1,\ldots,N\}} \Sigma_{i=1}^{I} \|h_i - \tilde{h}(k_i)\| \tag{2}$$

h_i is the vector of O*NET variables for observation i, k_i is the cluster assigned to observation i, and $\tilde{h}(k_i)$ is the mean vector of values for all points in cluster k.

TABLE 2 O*NET indices and related variables.

O*NET index	Variables comprising the index
Work setting (working indoors)	Indoors, environmentally controlled (+) indoors, not environmentally controlled (+); physical proximity (−); outdoors, under cover (−);
	In an open vehicle or equipment (−); in an enclosed vehicle or equipment (−); outdoors, exposed to weather (−)
Communication (high communication)	Telephone (+); electronic mail (+); letters and memos (+); face-to-face discussions (+); contact with others (+); public speaking (+)
Role relationships (work with people)	Coordinate or lead others (+); work with work group or team (+); deal with external customers (+)
Responsibility for others (high responsibility)	Responsibility for outcomes and results (+);
	Responsible for others' health and safety (+)
Conflictual contact (high conflictual contact)	Deal with unpleasant or angry people (+); deal with physically aggressive people (+); frequency of conflict situations (+)
Environmental conditions (uncomfortable conditions)	Sounds/Noise levels are distracting or uncomfortable (+); very hot or cold temperatures (+); extremely bright or inadequate lighting (+); exposed to contaminants (+); cramped work space (+); awkward positions (+); exposed to whole body vibration (+)
Job hazards (high job hazards)	Exposed to radiation (+); exposed to disease or infections (+); exposed to high places (+);
	Exposed to hazardous equipment (+); exposed to minor burns, cuts, bites, or stings (+);
	Exposed to hazardous conditions (+)
Body positioning (sedentary body positioning)	Spend time standing (−); spend time sitting (+); spend time climbing ladders, scaffolds, or
	Poles (−); spend time walking and running (−); spend time kneeling, crouching, stooping, or
	Crawling (−); spend time keeping or regaining balance (−); spend time using hands to handle control or feel objects tools or controls (−); spend time bending or twisting the body (−); spend time making repetitive motions (−)
Work attire (wear personal protective equipment)	Wear personal protective equipment (PPE) (+); wear specialized PPE (+)
Cruciality of position (crucial position)	Freedom to make decisions (+); consequence of error (+); impact of decisions on Co-workers or company results (+); frequency of decision making (+)
Routine versus Challenging (routine)	Degree of automation (+); importance of being exact or accurate (+); importance of repeating same tasks (+); structured versus unstructured work (+)
Level of competition (competitive)	Level of competition (+)
Pace and scheduling (tight scheduling)	Time pressure (+); pace determined by speed of equipment (+); work schedules (+); duration of typical work week (+)

Note: Indices are comprised of variables grouped by the O*NET model. The index is the average of the underlying measures. I sign each variable in the index to facilitate an easier interpretation of the overall index. The sign of each variable within the index is given by the (+) and (−) signs, and the economic interpretation of these indices is given in parentheses.

I solve this problem using Lloyd's algorithm:

1. Allocate each data point to the cluster with the closest mean value.
2. Recalculate cluster means.

These steps are repeated until convergence. For my main results I use three clusters.[9] To verify my solution is the global minimum, I start the algorithm from 25 random centroids and select the smallest solution. Clustering differs

TABLE 3 Sample summary statistics.

Statistic	*N*	Mean	St. Dev.	Min	Max	Mean (full data set)
Panel A: 15 year bandwidth						
Age	139,762	62.8	7.9	45	80	59.9
GHQ12 caseness	139,762	1.49	2.8	0	12	1.74
1 (Caseness >8)	139,762	0.06	0.25	0	1	0.08
Loss of confidence	139,762	2.98	1.3	0	8	3.09
Social dysfunction	139,762	12.5	2.1	5	24	12.6
Anxiety and depression	139,762	7.15	2.5	0	16	7.42
Log household income	88,863	9.76	0.8	0	15	9.74
Retirement	139,762	0.49	0.50	0	1	0.38
Female	139,762	0.37	0.48	0	1	0.54
Panel B: 5 year bandwidth						
Age	47,235	62.9	3.4	52	70	
GHQ12 caseness	47,235	1.4	2.7	0	12	
1 (Caseness >8)	47,235	0.06	0.24	0	1	
Loss of confidence	47,235	3.0	1.2	0	8	
Social dysfunction	47,235	12.5	2.0	6	24	
Anxiety and depression	47,235	7.1	2.5	1	16	
Log household income	28,723	9.8	0.79	0	14	
Retirement	47,235	0.50	0.50	0	1	
Female	47,235	0.36	0.48	0	1	

Note: Data are from the Understanding Society/British Household Panel Survey. In Panel A, the sample includes all person-wave observations within 15 years of the retirement eligibility age, for individuals with data on their employment history. Sample statistics for each variable are reported after dropping missing values. Household income is adjusted for CPI and equivalence scales. The GHQ12 caseness score is the number of negative mental health symptoms a person currently experiences. Loss of Confidence, Social Dysfunction, and Anxiety and Depression measure the extent of negative mental health symptoms of these types. Retired is a binary variable equal to 1 if and only if a person's primary reported labor market status is retired. In Panel B I narrow the sample to observations of people within 5 years of the retirement eligibility age.

from other dimension-reduction techniques such as factor and principal component analysis by producing discrete types, which is advantageous in this context because I flexibly interact cluster membership with the baseline RDD specification, avoiding unwarranted parametric assumptions, for example, linearity of effect heterogeneity in factors.

I estimate the short-run effect of retirement eligibility using local linear regressions on the running variable, r_{it}, time to retirement eligibility in months, allowing for a discontinuity at $r_{it} = 0$:

$$y_{it} = \alpha_0 + \delta_a r_{it} + \beta \mathbb{1}\{r_{it} < 0\} + \delta_b r_{it} \times \mathbb{1}\{r_{it} < 0\} + \tau_t + \epsilon_{it} \tag{3}$$

y_{it} is the outcome of interest for person i at time t. τ_t are wave fixed effects which can improve the precision of estimates (Lee and Lemieux, 2010). ϵ_{it} is the residual. α_0, δ_a, δ_b and β are regression coefficients, where β captures the discontinuity and the causal effect of interest. While the data are a panel, any characteristics which are fixed across individuals are not discontinuous in age, and therefore I do not include person fixed effects.[10]

The basic state pension pays a maximum of GBP 142 per week to people who have made sufficient tax contributions. The retirement eligibility age is the age at which a person can claim the basic state pension (the state pension age, SPA). In my sample, it is 65 for men, and between 60 and 65 for women.[11] The SPA is often also the age of retirement eligibility for occupational and private pensions (Hammond et al., 2016). Eligibility to claim pensions therefore changes discontinuously when a person reaches the SPA. For all men in my sample, this occurs on their 65th birthday, and therefore it is always possible to identify whether they are "treated", that is, reached the SPA, from the "Age at last birthday" variable. Some women reach the SPA between birthdays. For some observations of those women it is not always possible to identify exactly which side of the cutoff their months to retirement eligibility falls. For those observations, I follow Dong's (2015) recommendation and drop them from the sample. Since healthcare is predominantly provided by the National Health Service for people of all ages, there is no confounding due to changes in health insurance eligibility. Within the age bandwidths studied, I am not aware of any other discontinuous changes which confound the relationship.

To estimate heterogeneous effects by occupation cluster, I allow the discontinuity and trend in the running variable to differ by occupation cluster membership, x_i:

$$y_{it} = \alpha_0 + \delta_a r_{it} + \beta_1 \mathbb{1}\{r_{it} < 0\} + \delta_b r_{it} \times \mathbb{1}\{r_{it} < 0\} + \Sigma_{j=2}^{J}\Big(\delta_{j0}\mathbb{1}\{x_i = j\} + \delta_{aj} r_{it} \times \mathbb{1}\{x_i = j\} + \beta_j \mathbb{1}\{r_{it} < 0\} \times \mathbb{1}\{x_i = j\} + \delta_{cj} r_{it} \times \mathbb{1}\{r_{it} < 0\} \times \mathbb{1}\{x_i = j\}\Big) + \tau_t + \epsilon_{it} \quad (4)$$

Equation (4) is formed by interacting each term of Equation (3) with dummies for occupational cluster membership. The causal effect of retirement eligibility for those in cluster one is given by β_1 while the casual effect of retirement eligibility for those in cluster $j \neq 1$ is given by $\beta_1 + \beta_j$. I use a 5-year bandwidth[12] and a triangular kernel. Appendix C and D explore robustness of results to different kernels, bandwidths, and local polynomials. I cluster standard errors at the person level. Appendix E investigates robustness to alternative standard error estimators.

Finally, while some researchers use retirement eligibility as an instrument for retirement (e.g., Heller-Sahlgren, 2017), I estimate the reduced-form effect of retirement eligibility on mental health, which has two key advantages. Firstly, it does not require the exclusion restriction assumption— that retirement eligibility only affects mental health via its effect on retirement behavior. Secondly, my approach directly assesses the effects of retirement eligibility, the key policy lever available to policymakers.

4 | RESULTS

4.1 | Occupation groups

Table 4 presents the distribution of one-digit occupations across clusters. Occupations in cluster one are disproportionately in skilled agriculture, craft and related trades, and to a lesser extent plant and machinery operator workers. Cluster two jobs are disproportionately service jobs (e.g., travel attendants) plant and machinery operator workers, and elementary occupations (e.g., street vendors). Cluster three occupations are mostly white-collar, and largely comprise senior officials, professionals, associate professionals and clerks.

Figure 1 details the characteristics of the clusters resulting from K-means clustering. I plot the percentage of observations in each cluster who score "high" (in the top 25%) less the percentage of observations who score "low" (in the bottom 25%) on each O*NET index. Cluster one occupations are mostly outside, involve the least communication, working with people and conflictual contact, are the most likely to work in uncomfortable conditions, experience job hazards, wear PPE, and have tight scheduling. Cluster two jobs have the most conflictual contact, and score the lowest for tight scheduling. On the other hand, they do not score especially high or low on average for other job characteristics. Cluster three jobs score the lowest for wearing PPE, having uncomfortable conditions, and having high levels of job hazards. They score the highest for sedentary body positioning, communication, competitiveness and routineness, working indoors, and working with people.

Table 5 shows patterns of health, gender, and mental health across clusters. Employees in cluster one have lower household income, better mental health, and mostly blue-collar jobs and are disproportionately men. Cluster two mixes blue- and white-collar jobs and workers are disproportionately women. Cluster three employees have the highest household income and are 42% male.

TABLE 4 Share of occupational employment by cluster.

One-digit ISCO-88 occupation group	Cluster 1	Cluster 2	Cluster 3
Senior officials	0.050	0.054	0.896
Professionals	0.010	0.184	0.807
Technicians and associate professionals	0.040	0.349	0.611
Clerks	0.014	0.061	0.925
Service workers	0.058	0.874	0.069
Skilled agriculture	0.925	0.014	0.061
Craft and related trades	0.798	0.121	0.081
Plant and machinery operator workers	0.513	0.440	0.047
Elementary occupations	0.413	0.529	0.058

Note: The table shows the one-digit ISCO-88 occupation shares in each occupational cluster in the sample. I cluster person-wave observations by O*NET occupation data using K-means clustering.

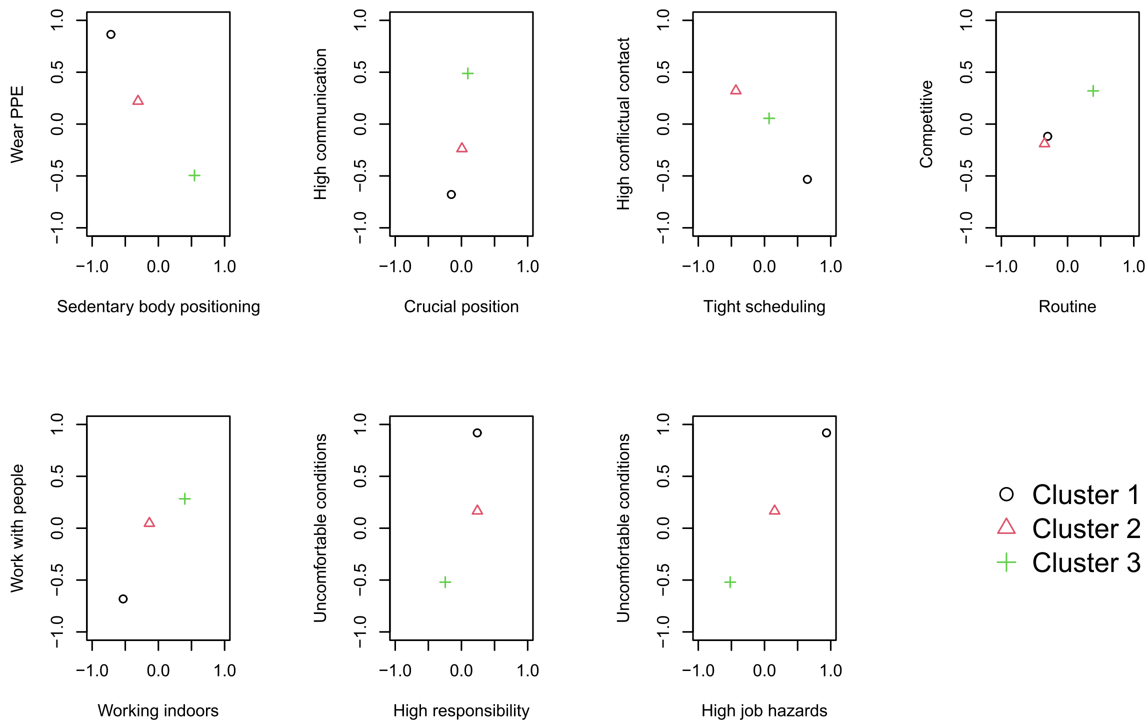

FIGURE 1 Percent of people in each cluster scoring high minus percent scoring low. I cluster person-wave observations by O*NET occupation data using K-means clustering. The measures of cluster characteristics are calculated by subtracting the percentage of people who score "low" on an O*NET variable within a cluster from the percentage of people scoring "high", where scoring "high" or "low" is determined by being in the top 25% or bottom 25% respectively of the distribution of scores. Individuals are assigned to occupation clusters based on current or last occupation. O*NET indices are calculated based on the O*NET model, which groups variables into categories. I calculate the indices as the simple average over the variables in the index, which are signed to facilitate an intuitive interpretation (see Table 2). [Colour figure can be viewed at wileyonlinelibrary.com]

4.2 | Heterogeneous effects of retirement

Table 6 reports the effect of retirement eligibility overall and by cluster. Retirement eligibility has an overall beneficial effect on the overall GHQ12 caseness, the probability of the caseness being above clinically significant levels, symptoms of Social Dysfunction, symptoms of Loss of Confidence (but only significant at the 10% level), and symptoms of Anxiety and Depression. Effects are concentrated amongst those in cluster one and cluster three occupations. In cluster one,

TABLE 5 Patterns of income and health amongst those working in different occupational clusters.

Cluster	Sample %	% male	Average monthly household income	Average caseness score	% white-collar
1	20.7	84.5	4173	1.092	11.3
2	33.1	30.6	4247	1.566	58.0
3	46.2	42.4	5703	1.549	99.0

Note: Cells are average values of each variable, conditional on working in a given occupation cluster. Variables are defined in the note to Table 3. Individuals are assigned to occupation clusters based on current occupation. While Understanding Society does not report whether a job is blue-collar or white-collar, the data do report the Goldthorpe-Hope scheme (see e.g., Erikson & Goldthorpe, 2002) an occupational categorization which can be collapsed to a blue-collar/white-collar categorization (Dhungel et al., 2021).

most of the benefit comes from symptoms of Anxiety and Depression; in cluster three, retirement eligibility affects symptoms of Anxiety and Depression and Social Dysfunction.

Men derive larger overall benefits from retiring. Disaggregating by cluster and gender, estimates are less precise and we should be more tentative about drawing conclusions. However, my results suggest occupational effect heterogeneity even within gender, especially amongst women. Therefore gender differences do not drive my main results. The overall muted effect of being eligible to retire from cluster two appears to be explained by a large beneficial effect amongst men and potentially an adverse effect amongst women (though the latter is not statistically significant). The beneficial effect of being eligible to retire from cluster one is present for men but not for women. The point estimates for cluster three are similar for both genders and are consistent with a beneficial effect which is large relative to the population average.

Differential patterns of retirement and changes in income are unlikely to drive my results: eligibility to retire from cluster two occupations has the smallest effect on mental health, but the largest effect on household income and second largest effect on the probability of retiring. My results are consistent with an adverse effect of working in cluster one and three occupations compared to retirement. The physical danger of cluster one occupations might increase stress about the risk of injury, and there are aspects of cluster three occupations associated with mental illness: they are sedentary, competitive and based indoors. Having the option of receiving the BSP may also improve mental health for those who work in these occupations, by improving the outside option for someone who is injured, fired, or otherwise loses their job. Alternatively, people who benefit most from retirement may disproportionately sort into cluster one and three occupations.

Appendix G presents robustness tests. My results are qualitatively robust to dropping wave fixed effects and observations of people who retire before their retirement eligibility age.

5 | DISCUSSION AND CONCLUSION

This paper investigates how the effect of retirement eligibility on mental health outcomes varies across occupational characteristics. I use K-means clustering to estimate the heterogeneous effect of retirement eligibility across multiple dimensions simultaneously. The benefits of retirement are concentrated in identifiable occupational groups: one group of physical, dangerous, blue-collar jobs, and one group of professional, associate professional, or clerk jobs which are white collar and have high levels of contact with people.

My results provide novel evidence about the heterogeneous effects of retirement eligibility. Previous literature has estimated heterogeneous effects using a blue-collar and white-collar categorization (e.g., Kolodziej & García-Gómez, 2019). This approach does not allow for retirement eligibility to have heterogeneous effects within the blue-collar and white-collar categorizations, an important potential omission given that there is large variation in working conditions within blue-collar and white-collar occupations. I take a more detailed and data-driven approach to constructing occupational types. My approach suggests there are both blue- and white-collar occupations (clusters one and three) where the effect of retirement eligibility is elevated, and blue- and white-collar occupations (cluster two) where the effect is muted.

Secondly, previous literature has investigated mental health effects of psychologically straining jobs and physically straining jobs either separately (e.g., Eibich, 2015; Heller-Sahlgren, 2017; Mazzonna & Peracchi, 2017), or combining them into a single measure of job demand (Carrino et al., 2020). I show that there are two non-overlapping categories of occupations which are physically straining (cluster one) and psychologically straining (cluster three), and the causal

TABLE 6 The effect of retirement eligibility by occupational cluster (using RDD).

	GHQ12 caseness	1 (caseness >8)	Loss of confidence	Social dysfunction	Anxiety and depression	Log household income	Retired
Panel A: Men and women							
All clusters	−0.115***	−0.008**	−0.039*	−0.122***	−0.116***	0.134***	0.226***
	(0.044)	(0.004)	(0.020)	(0.033)	(0.040)	(0.017)	(0.008)
Cluster 1	−0.126	−0.006	−0.049	−0.095	−0.200**	0.139***	0.299***
	(0.089)	(0.008)	(0.041)	(0.066)	(0.081)	(0.032)	(0.015)
Cluster 2	−0.048	−0.004	−0.032	−0.084	−0.029	0.190***	0.239***
	(0.078)	(0.007)	(0.036)	(0.058)	(0.071)	(0.030)	(0.013)
Cluster 3	−0.158***	−0.012**	−0.040	−0.166***	−0.136***	0.094***	0.176***
	(0.066)	(0.006)	(0.030)	(0.049)	(0.060)	(0.025)	(0.011)
Obs	47,235	47,235	47,235	47,235	47,235	28,723	47,235
Panel B: Men							
All clusters	−0.159***	−0.011***	−0.049**	−0.150***	−0.109**	0.138***	0.233***
	(0.050)	(0.004)	(0.024)	(0.037)	(0.047)	(0.019)	(0.009)
Cluster 1	−0.173*	−0.011	−0.061	−0.126*	−0.231***	0.139***	0.308***
	(0.092)	(0.008)	(0.044)	(0.068)	(0.087)	(0.035)	(0.017)
Cluster 2	−0.162*	−0.011	−0.058	−0.132*	−0.019	0.196***	0.243***
	(0.098)	(0.008)	(0.046)	(0.072)	(0.092)	(0.035)	(0.018)
Cluster 3	−0.150**	−0.012	−0.037	−0.177***	−0.082	0.095***	0.174***
	(0.076)	(0.007)	(0.036)	(0.056)	(0.072)	(0.028)	(0.014)
Obs	30,486	30,486	30,486	30,486	30,486	22,420	30,486
Panel C: Women							
All clusters	−0.013	−0.002	−0.014	−0.052	−0.135*	0.090	0.223***
	(0.086)	(0.008)	(0.037)	(0.065)	(0.073)	(0.043)	(0.013)
Cluster 1	0.020	0.008	−0.012	0.030	−0.164	0.194	0.277***
	(0.230)	(0.021)	(0.100)	(0.175)	(0.196)	(0.124)	(0.034)
Cluster 2	0.115	0.005	−0.0002	−0.013	−0.041	0.166**	0.244***
	(0.131)	(0.012)	(0.057)	(0.100)	(0.112)	(0.067)	(0.019)
Cluster 3	−0.150	−0.011	−0.032	−0.120	−0.224**	0.044	0.188***
	(0.128)	(0.012)	(0.056)	(0.097)	(0.109)	(0.059)	(0.019)
Obs	16,749	16,749	16,749	16,749	16,749	6303	16,749

Note: Results are the short-run causal effect of being eligible to retire for those in each occupational cluster. I use a local linear regression around the cutoff of 0 months to retirement eligibility. The running variable is expected months to retirement eligibility. A bandwidth of 5 years is used, and observations are weighted using a triangular kernel. For heterogeneous effects by cluster, I also interact with occupational cluster membership as in Equation (4), and use the sum of coefficients to infer the causal effects presented in the Table, that is, the causal effect of retirement eligibility for cluster one is β_1 in Equation (4), and $\beta_1 + \beta_j$ for cluster j. I control for wave fixed effects. Standard errors are clustered at the individual level. I calculate the standard error of the causal effect by using the heteroskedasticity-robust variance-covariance matrix clustered at the individual level. Where I allow the effect to differ by gender, I estimate the effect separately in sample of men and women. Variables are defined in the note to Table 3.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

TABLE 7 Optimal bandwidths for each outcome.

	MSE optimal bandwidth (years)
GHQ12 caseness	4.620
1(Caseness >8)	4.891
Loss of confidence	4.783
Social dysfunction	4.363
Anxiety and depression	4.741
Household income	4.002
Retired	4.163

Note: The sample is all person-wave observations within 15 years of the age of retirement eligibility that provide a value for each mental health variable and retirement. Optimal bandwidths are calculated using the procedures developed by Calonico et al. (2014), which minimize the mean squared error of the RDD estimator. Variables are defined in the note to Table 3.

effect of retirement eligibility varies between them. My results therefore indicate that the source of job strain is important and support differentiating between psychological and physical sources of strain.

Regarding policy, many developed economies are raising the age of retirement eligibility. The desirability of this policy may depend on who shoulders the burden. My results suggest that the largest mental health costs fall on those in the highest prestige, highest paying jobs (cluster three), followed by those in cluster one occupations, who have the lowest household income but better pre-retirement mental health.

My research has two limitations: firstly, I only identify the short-run effect of retirement eligibility;[13] secondly, I cannot distinguish between the hypothesis that the causal effect of being eligible to retire differs by occupation, and the hypothesis that people with different benefits of retirement sort into different occupations.

CONFLICT OF INTEREST STATEMENT

I have no conflict of interest to disclose.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the UK data service. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from https://ukdataservice.ac.uk/ with the permission of the UK data service.

ORCID

Joe Spearing https://orcid.org/0000-0002-1437-1780

ENDNOTES

[1] There is also a literature which investigates heterogeneity across dimensions other than occupation, such as gender (Atalay & Barrett, 2014), marital status (Della Giusta & Longhi, 2021; Picchio and Ours, 2020), education (Della Giusta & Longhi, 2021), personality traits (Kesavayuth et al., 2016), quality of social life (Kettlewell & Lam, 2022), and the business cycle (Martinez Jimenez, Hollingsworth, and Zucchelli, 2021).

[2] University of Essex, Institute for Social and Economic Research. (2022). Understanding Society: Waves 1–11, 2009–2020 and Harmonised BHPS: Waves 1–18, 1991–2009. [data collection]. 15th Edition. UK Data Service. SN: 6614, http://doi.org/10.5255/UKDA-SN-6614-16.

[3] O*NET 26.3 Database, by the US Department of Labor, Employment and Training Administration; https://creativecommons.org/licenses/by/4.0/.

[4] The survey does not report exact day of birth. I assume that day of birth is uniformly distributed. When a person is interviewed later in the month, their expected age is older than when they are interviewed earlier in the month. For respondents who are ever interviewed in their birth month, age at last birthday allows me to narrow down the day of the month that they were born on, so I tighten the bounds on their possible age accordingly.

[5] https://www.onetcenter.org/content.html?msclkid=203f5eb3bc3611ec8e6a560536451be2#cm4.

[6] The rationale for averaging over the whole period is that results should not be affected by how characteristics of an occupation have changed since a person stopped working, that is, a person's occupational cluster should not change because of changes to working conditions which have occurred since they retired. To determine the stability of occupational characteristics across the period I observe

them, I estimate the correlation between an occupation's score on each O*NET work context variable in its first wave versus the most recent wave. For all variables the correlation is greater than 0.6, and for most it is above 0.7.

7 Source: World Bank https://genderdata.worldbank.org/indicators/se-sch-life/?gender=total.

8 The crosswalk translates occupation codes across classification systems. It tells the researcher which code in, say, the ISCO-88 classification corresponds to each code in the SOC classification. Most papers (e.g., Bertoni Maggi and Weber, 2018; Carrino Glaser and Avendano, 2020; Kolodziej & García-Gómez, 2019; Leckcivilize and McNamee, 2021; Mazzonna & Peracchi, 2017) which categorize people into different kinds of jobs in an RDD base the categorization on a person's current job or last job if not currently working. If a person has a history of working within the sample period, I classify their last occupation as the occupation they were last observed working in. Additionally, individuals who never work within the sample period report the last occupation they worked in if applicable.

9 I choose three clusters in order to trade off the benefits of capturing a greater share of the variation in occupational characteristics against the costs of losing precision when the number of clusters increases. I discuss this tradeoff, and results using four clusters in Appendix F.

10 Including person fixed effects could decrease precision significantly in this case because some individuals only appear in the bandwidth a small number of times (Lee and Lemieux, 2010). For this reason, many researchers using a RDD do not use person fixed effects even when using panel data (e.g., Rose, 2020; Watson, 2020).

11 Appendix B provides details about the Pensions Acts which increased the age of retirement eligibility for women and how I incorporate the changing pension age into my analysis.

12 I present optimal bandwidths according to Calonico et al. (2014) in Table 7. The choice of bandwidth is based on trading off the decrease in variance from a larger bandwidth against the reduction in bias from reducing the bandwidth. Most optimal bandwidths are close to 5 years.

13 An important rejoinder is that some studies have shown evidence of a "halo effect", a short-run beneficial effect which is not sustained (e.g., Heller-Sahlgren, 2017). To the extent that this is true, results RDD may not be a good indicator of the longer-run effect of retirement eligibility.

14 https://www.gov.uk/government/publications/state-pension-age-timetable/state-pension-age-timetable.

15 Dong (2015) shows that with a second-order polynomial estimation of RDD and measurement error in the running variable, the true causal effect is given by:

$$\tau = \beta_1 - \mu_1\delta_b + \left(2\mu_1^2 - \mu_2\right)\delta_d \tag{5}$$

where β_1, and δ_b have the same interpretation as in Equation (3), δ_d is the coefficient on the squared term of the polynomial, and μ_i is the ith moment of the measurement error in the running variable. This equation implies that if the running variable is observed with mean-zero error and a squared term is not included, measurement error in the running variable does not bias the RDD estimate.

REFERENCES

Anjara, S. G., Bonetto, C., Van Bortel, T., & Brayne, C. (2020). Using the GHQ-12 to screen for mental health problems among primary care patients: Psychometrics and practical considerations. *International Journal of Mental Health Systems*, *14*(1), 1–13. https://doi.org/10.1186/s13033-020-00397-0

Atalay, K., & Barrett, G. F. (2014). The causal effect of retirement on health: New evidence from Australian pension reform. *Economics Letters*, *125*(3), 392–395. https://doi.org/10.1016/j.econlet.2014.10.028

Belloni, M., Carrino, L., & Meschi, E. (2022). The impact of working conditions on mental health: Novel evidence from the UK. *Labour Economics*, 102176.

Bertoni, M., Maggi, S., & Weber, G. (2018). Work, retirement, and muscle strength loss in old age. *Health Economics*, *27*(1), 115–128. https://doi.org/10.1002/hec.3517

Bonhomme, S., Lamadon, T., & Manresa, E. (2022). Discretizing unobserved heterogeneity. *Econometrica*, *90.2*(2), 625–643. https://doi.org/10.3982/ecta15238

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*(6), 2295–2326. https://doi.org/10.3982/ecta11757

Carrino, L., Glaser, K., & Avendano, M. (2020). Later retirement, job strain, and health: Evidence from the new State Pension age in the United Kingdom. *Health Economics*, *29*(8), 891–912. https://doi.org/10.1002/hec.4025

CEDEFOP. (2013). Quantifying skill needs in Europe. Occupational Skills profiles: Methodology and application.

Della Giusta M., & S. Longhi (2021). Stung by pension reforms: The unequal impact of changes in state pension age on UK women and their partners. *Labour Economics, Labour Economics 72*, 102049. https://doi.org/10.1016/j.labeco.2021.102049

Dhungel, B., Murakami, T., Wada, K., & Gilmour, S. (2021). Mortality risks among blue-and white-collar workers: A time series study among Japanese men aged 25-64 years from 1980 to 2015. *Journal of Occupational Health*, *63*(1), e12215. https://doi.org/10.1002/1348-9585.12215

Dong, Y. (2015). Regression discontinuity applications with rounding errors in the running variable. *Journal of Applied Econometrics*, *30*(3), 422–446. https://doi.org/10.1002/jae.2369

Edwards, T., Edwards, P., Ferner, A., Marginson, P., & Tregaskis, O. (2010). Multinational companies and the diffusion of employment practices from outside the country of origin. *Management International Review*, *50*(5), 613–634. https://doi.org/10.1007/s11575-010-0051-2

Eibich, P. (2015). Understanding the effect of retirement on health: Mechanisms and heterogeneity. *Journal of Health Economics*, *43*, 1–12. https://doi.org/10.1016/j.jhealeco.2015.05.001

Erikson, R., & Goldthorpe, J. H. (2002). Intergenerational inequality: A sociological perspective. *The Journal of Economic Perspectives*, *16*(3), 31–44. https://doi.org/10.1257/089533002760278695

Gathergood, J. (2013). An instrumental variable approach to unemployment, psychological health and social norm effects. *Health Economics*, *22*(6), 643–654.

Gelaye, B., Tadesse, M. G., Lohsoonthorn, V., Lertmeharit, S., Pensuksan, W. C., Sanchez, S. E., Lemma, S., Berhane, Y., Vélez, J. C., Barbosa, C., Anderade, A., & Williams, M. A. (2015). Psychometric properties and factor structure of the General Health Questionnaire as a screening tool for anxiety and depressive symptoms in a multi-national study of young adults. *Journal of Affective Disorders*, *187*, 197–202.

Geyer, J., Haan, P., Hammerschmid, A., & Peters, M. (2020). Labor market and distributional effects of an increase in the retirement age. *Labour Economics*, *65*, 101817. https://doi.org/10.1016/j.labeco.2020.101817

Goos, M., Manning, A., & Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *The American Economic Review*, *104*(8), 2509–2526. https://doi.org/10.1257/aer.104.8.2509

Gorry, A., Gorry, D., & Slavov, S. N. (2018). Does retirement improve health and life satisfaction? *Health Economics*, *27*(12), 2067–2086. https://doi.org/10.1002/hec.3821

Graetz, B. (1991). Multidimensional properties of the general health questionnaire. *Social Psychiatry and Psychiatric Epidemiology*, *26*(3), 132–138. https://doi.org/10.1007/bf00782952

Gureje, O., & Obikoya, B. (1990). The GHQ-12 as a screening tool in a primary care setting. *Social Psychiatry and Psychiatric Epidemiology*, *25*(5), 276–280.

Hammond, R., Baxter, S., Bramley, R., Kakkad, A., Mehta, S., & Sadler, M. (2016). Considerations on state pension age in the United Kingdom. *British Actuarial Journal*, *21*(1), 165–203. https://doi.org/10.1017/s1357321715000069

Handel, M. J. (2012). Trends in job skill demands in OECD countries.

Hardy, W. (2016). Hardy, 2006. Retrieved from https://ibs.org.pl/en/resources/occupation-classifications-crosswalks-from-onet-soc-to-isco/. Accessed: 2021-08-05.

Hardy, W., Roma, K., & Lewandowski, P. (2018). Educational upgrading, structural change and the task composition of jobs in Europe. *The Economics of Transition*, *26*(2), 201–231. https://doi.org/10.1111/ecot.12145

Heller-Sahlgren, G. (2017). Retirement blues. *Journal of Health Economics*, *54*, 66–78. https://doi.org/10.1016/j.jhealeco.2017.03.007

Johnston, D. W., & Lee, W.-S. (2009). Retiring to the good life? The short-term effects of retirement on health. *Economics Letters*, *103*(1), 8–11. https://doi.org/10.1016/j.econlet.2009.01.015

Jolivet, G., & Postel-Vinay, F. (2020). A structural analysis of mental health and labor market trajectories.

Kesavayuth, D., Rosenman, R. E., & Zikos, V. (2016). Retirement, personality, and well-being. *Economic Inquiry*, *54*(2), 733–750. https://doi.org/10.1111/ecin.12307

Kettlewell, N., & Lam, J. (2022). Retirement, social support and mental well-being: A couple-level analysis. *The European Journal of Health Economics*, *23*(3), 511–535. https://doi.org/10.1007/s10198-021-01374-1

Kolesár, M., & Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, *108*(8), 2277–2304.

Kolodziej, I. W. K., & García-Gómez, P. (2019). Saved by retirement: Beyond the mean effect on mental health. *Social Science and Medicine*, *225*, 85–97. https://doi.org/10.1016/j.socscimed.2019.02.003

Leckcivilize, A., & Paul, McN. (2021). The effects of retirement on health and health behaviour among retirees and their partners: Evidence from the English longitudinal study of ageing. *Journal of Population Ageing*, 1–32.

Lee, D. S., & Thomas, L. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*(2), 281–355. https://doi.org/10.1257/jel.48.2.281

Lewandowski, P. (2020). Occupational exposure to contagion and the spread of COVID-19 in Europe.

Martinez, J. M., Hollingsworth, B., & Zucchelli, E. (2021). Health, retirement and economic shocks.

Mazzonna, F., & Peracchi, F. (2017). Unhealthy retirement? *Journal of Human Resources*, *52*(1), 128–151. https://doi.org/10.3368/jhr.52.1.0914-6627r1

McCabe, C. J., Thomas, K. J., Brazier, J. E., & Coleman, P. (1996). Measuring the mental health status of a population: A comparison of the GHQ–12 and the SF–36 (MHI–5). *The British Journal of Psychiatry*, *169*(4), 517–521. https://doi.org/10.1192/bjp.169.4.516

Müller, T., & Shaikh, M. (2018). Your retirement and my health behavior: Evidence on retirement externalities from a fuzzy regression discontinuity design. *Journal of Health Economics*, *57*, 45–59. https://doi.org/10.1016/j.jhealeco.2017.10.005

Neuman, K. (2008). Quit your job and get healthier? The effect of retirement on health. *Journal of Labor Research*, *29*(2), 177–201. https://doi.org/10.1007/s12122-007-9036-8

Nishimura, Y., Oikawa, M., & Motegi, H. (2018). What explains the difference in the effect of retirement on health? Evidence from global aging data. *Journal of Economic Surveys*, *32*(3), 792–847. https://doi.org/10.1111/joes.12215

Picchio, M., & Jan, C. O. (2020). Mental health effects of retirement. *De Economist*, *3*, 419–452.

Rose, L. (2020). Retirement and health: Evidence from england. *Journal of Health Economics*, *73*, 102352. https://doi.org/10.1016/j.jhealeco.2020.102352

Steinley, D. (2006). K-Means clustering: A half-century synthesis. *British Journal of Mathematical and Statistical Psychology*, *59*(1), 1–34. https://doi.org/10.1348/000711005x48266

Watson, T. (2020). Retirement blues and boons: Health and spillover effects of retirement.

How to cite this article: Spearing, J. (2024). The effect of retirement eligibility on mental health in the United Kingdom: Heterogeneous effects by occupation. *Health Economics*, *33*(8), 1621–1648. https://doi.org/10.1002/hec.4835

APPENDIX A

GHQ12 QUESTIONS AND SUB-INDICES

The GHQ12 is a set of 12 questions about mental health symptoms, each of which has four possible answers reflecting different levels of severity. If the question is posed so that affirmative answers indicate good mental health, the answers are one of the following: "Better than usual", "Same as usual", "Less than usual", "Much less than usual". If the question is posed so that affirmative answers indicate poor mental health, the answers are one of the following: "Not at all", "No more than usual", "Rather more than usual", "Much more than usual". To construct the caseness, I first convert each question into a binary variable: a person scores 0 if their symptoms are either the best or second best possible, and 1 if their symptoms are either the worst or second worst possible. I aggregate these variables into the GHQ12 caseness score. The sub-indices are Loss of Confidence, Anxiety and Depression and Social Dysfunction. In calculating the sub-indices, one treats the individual variables as cardinal measures of mental health symptoms, scored from 1 to 4, and sum over all the symptoms of a particular type

The questions used to construct the GHQ12 variables are listed below. The sub-index a question belongs to is in parentheses after each question.

1. Have you recently been able to concentrate on whatever you're doing? (Social Dysfunction)
2. Have you recently lost much sleep over worry? (Anxiety and Depression)
3. Have you recently felt that you were playing a useful part in things? (Social Dysfunction)
4. Have you recently felt capable of making decisions about things? (Social Dysfunction)
5. Have you recently felt constantly under strain? (Anxiety and Depression)
6. Have you recently felt you couldn't overcome your difficulties? (Anxiety and Depression)
7. Have you recently been able to enjoy your normal day-to-day activities? (Social Dysfunction)
8. Have you recently been able to face up to problems? (Social Dysfunction)
9. Have you recently been feeling unhappy or depressed? (Anxiety and Depression)
10. Have you recently been losing confidence in yourself? (Loss of Confidence)
11. Have you recently been thinking of yourself as a worthless person? (Loss of Confidence)
12. Have you recently been feeling reasonably happy, all things considered? (Social Dysfunction)

The GHQ12 has been used to screen for psychological illness in a variety of contexts (e.g., Anjara et al., 2020; Gureje & Obikoya, 1990). There is also evidence that screening using sub-indices can help identify the risk of different psychiatric illnesses (Gelaye et al., 2015). Both the caseness score and sub-indices have been used as outcome measures in various economics studies (e.g., Belloni et al., 2022; Gathergood, 2013).

APPENDIX B

INSTITUTIONAL DETAILS ON WOMEN'S RETIREMENT AGES

For women born before April 1950, the SPA is 60. The Pension Acts of 1995 and 2011 increased the retirement eligibility age for women born after this date. The increase was phased in over time so that the later a woman is born, the older she is on the date she reaches SPA eligibility, until the retirement age is equalized between men and women. For women, there is therefore a mapping between date of birth and date of reaching SPA. The retirement dates for UK women by date of birth are given in Table B1.[14]

For each woman in my sample, I construct her expected months to retirement by taking a weighted average of retirement dates she could face and subtracting the current date. Dong (2015) shows that measurement error in the running variable does not lead to bias in RDD results if a local linear regression is used and the measurement error is mean zero. Additionally, following Dong's recommendation, I drop observations where it cannot be determined which side of the cutoff they fall. Note that all such observations are of women, because all of the men in my sample reach retirement age at the age of 65, and the survey reports age at last birthday.

TABLE B1 Women's retirement date by date of birth.

Date of birth	Date state pension age reached
April 6, 1950 – May 5, 1950	May 6, 2010
May 6, 1950 – June 5, 1950	July 6, 2010
June 6, 1950 – July 5, 1950	September 6, 2010
July 6, 1950 – August 5, 1950	November 6, 2010
August 6, 1950 – September 5, 1950	January 6, 2011
September 6, 1950 – October 5, 1950	March 6, 2011
October 6, 1950 – November 5, 1950	May 6, 2011
November 6, 1950 – December 5, 1950	July 6, 2011
December 6, 1950 – January 5, 1951	September 6, 2011
January 6, 1951 – February 5, 1951	November 6, 2011
February 6, 1951 – March 5, 1951	January 6, 2012
March 6, 1951 – April 5, 1951	March 6, 2012
April 6, 1951 – May 5, 1951	May 6, 2012
May 6, 1951 – June 5, 1951	July 6, 2012
June 6, 1951 – July 5, 1951	September 6, 2012
July 6, 1951 – August 5, 1951	November 6, 2012
August 6, 1951 – September 5, 1951	January 6, 2013
September 6, 1951 – October 5, 1951	March 6, 2013
October 6, 1951 – November 5, 1951	May 6, 2013
November 6, 1951 – December 5, 1951	July 6, 2013
December 6, 1951 – January 5, 1952	September 6, 2013
January 6, 1952 – February 5, 1952	November 6, 2013
February 6, 1952 – March 5, 1952	January 6, 2014
March 6, 1952 – April 5, 1952	March 6, 2014
April 6, 1952 – May 5, 1952	May 6, 2014
May 6, 1952 – June 5, 1952	July 6, 2014
June 6, 1952 – July 5, 1952	September 6, 2014
July 6, 1952 – August 5, 1952	November 6, 2014
August 6, 1952 – September 5, 1952	January 6, 2015
September 6, 1952 – October 5, 1952	March 6, 2015

TABLE B1 (Continued)

Date of birth	Date state pension age reached
October 6, 1952 – November 5, 1952	May 6, 2015
November 6, 1952 – December 5, 1952	July 6, 2015
December 6, 1952 – January 5, 1953	September 6, 2015
January 6, 1953 – February 5, 1953	November 6, 2015
February 6, 1953 – March 5, 1953	January 6, 2016
March 6, 1953 – April 5, 1953	March 6, 2016
April 6, 1953 – May 5, 1953	July 6, 2016
May 6, 1953 – June 5, 1953	November 6, 2016
June 6, 1953 – July 5, 1953	March 6, 2017
July 6, 1953 – August 5, 1953	July 6, 2017
August 6, 1953 – September 5, 1953	November 6, 2017
September 6, 1953 – October 5, 1953	March 6, 2018
October 6, 1953 – November 5, 1953	July 6, 2018
November 6, 1953 – December 5, 1953	November 6, 2018

Note: Data show the date of state pension eligibility by date of birth for women in the UK. Source: Department for Work and Pensions. Women born before April 6, 1950 reach retirement eligibility aged 60. Women born after December 5, 1953 reach retirement eligibility aged 65.

APPENDIX C

SENSITIVITY OF ALL-SAMPLE ESTIMATES TO CHANGES IN SPECIFICATION

Figures C1–C4 display how the causal estimates of the effect of being eligible to retire on mental health outcomes differ across specifications. For example, in Figure C1, the top panel shows the estimated effect of gaining eligibility to retire on the GHQ12 caseness. Each point is an estimated causal effect of retirement eligibility on the GHQ12 caseness (and confidence interval) for a given set of estimation choices. The three panels below show the bandwidth used in estimation, whether a quadratic term in the running variable is included, and which type of kernel used.

Across almost all specifications, being eligible to retire has a beneficial effect on mental health as measured by GHQ12 caseness, symptoms of Anxiety and Depression, and Social Dysfunction. The estimates are remarkably stable, and the only specifications in which they are not statistically significant are when a small bandwidth and quadratic specification are used.

Using a quadratic specification with measurement error in the running variable biases the estimated effect.[15] However, in most of my estimates, the size of the bias is small. It depends on the variance of the measurement error, and the magnitude of the estimated curvature of the trend in potential outcomes for those who are treated. Given my assumptions about the error in months to retirement being uniformly distributed, this variance is always less than $\frac{1}{12}$. The estimated curvature tends to be small. For example, when a 5-year bandwidth and triangular kernal are used to estimate the effect on the overall caseness, the coefficient on the squared term of those above the cutoff is 0.000004631, giving a total bias smaller than $0.000004631 \times \frac{1}{12} = 0.00000386$.

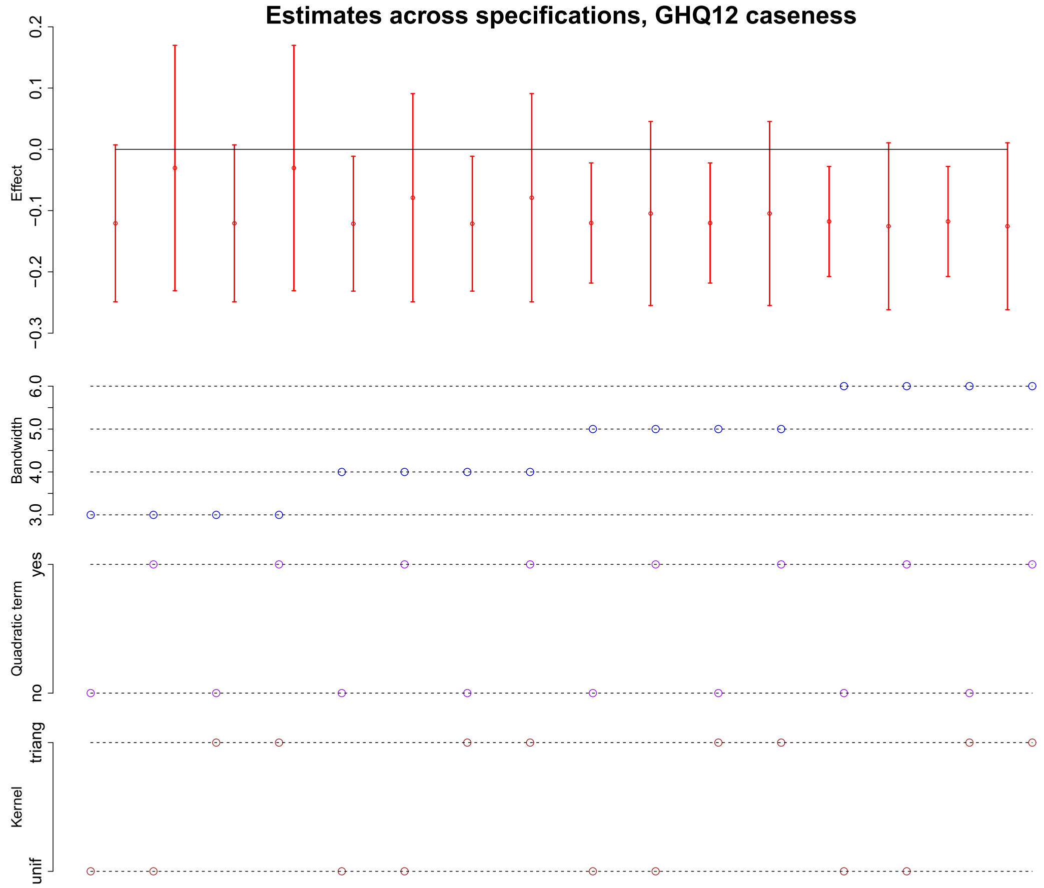

FIGURE C1 Effect of retirement eligibility on GHQ12 caseness across specifications. Results are the average short-run causal effect of being eligible to retire on the GHQ12 caseness. I estimate the causal effect using the discontinuity in the GHQ12 caseness as a person become eligible for retirement. I use a bandwidth of between 3 and 6 years. I use a triangular or uniform kernel. I use a local linear or local quadratic estimator. I calculate the standard error of the causal effect by using the heteroskedasticity-robust variance-covariance matrix clustered by individual. Variables are defined in the note to Table 3. [Colour figure can be viewed at wileyonlinelibrary.com]

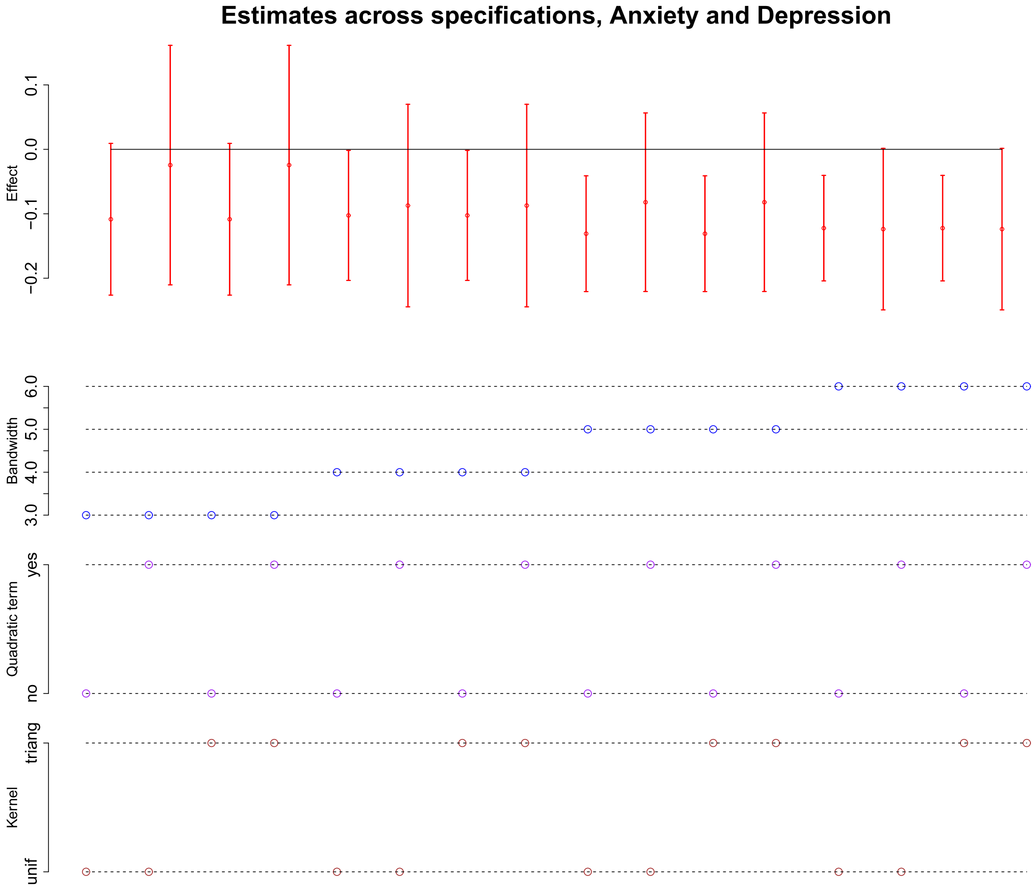

FIGURE C2 Effect of retirement eligibility on Anxiety and Depression across specifications. Results are the average short-run causal effect of being eligible on symptoms of Anxiety and Depression to retire in the population. The range of estimation specifications is described in the note to Figure C1. Variables are defined in the note to Table 3. [Colour figure can be viewed at wileyonlinelibrary.com]

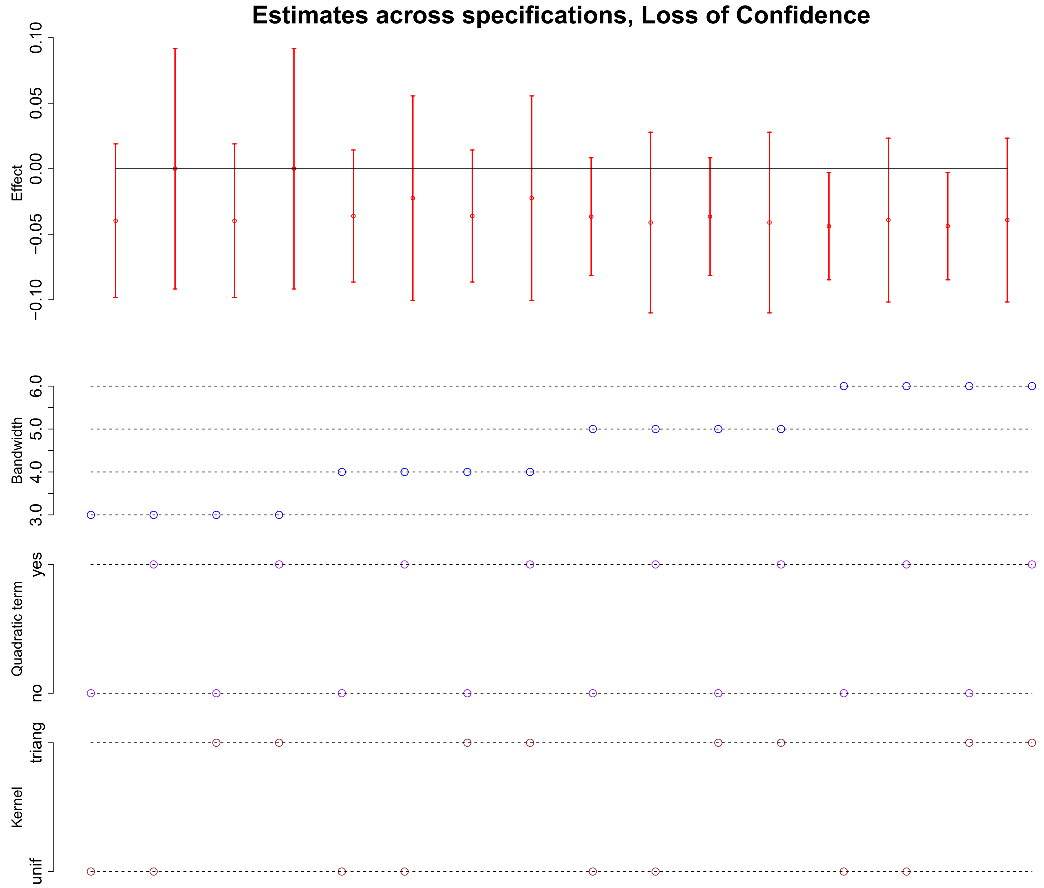

FIGURE C3 Effect of retirement eligibility on Loss of Confidence across specifications. Results are the average short-run causal effect of being eligible to retire on symptoms of Loss of Confidence. The range of estimation specifications is described in the note to Figure C1. Variables are defined in the note to Table 3. [Colour figure can be viewed at wileyonlinelibrary.com]

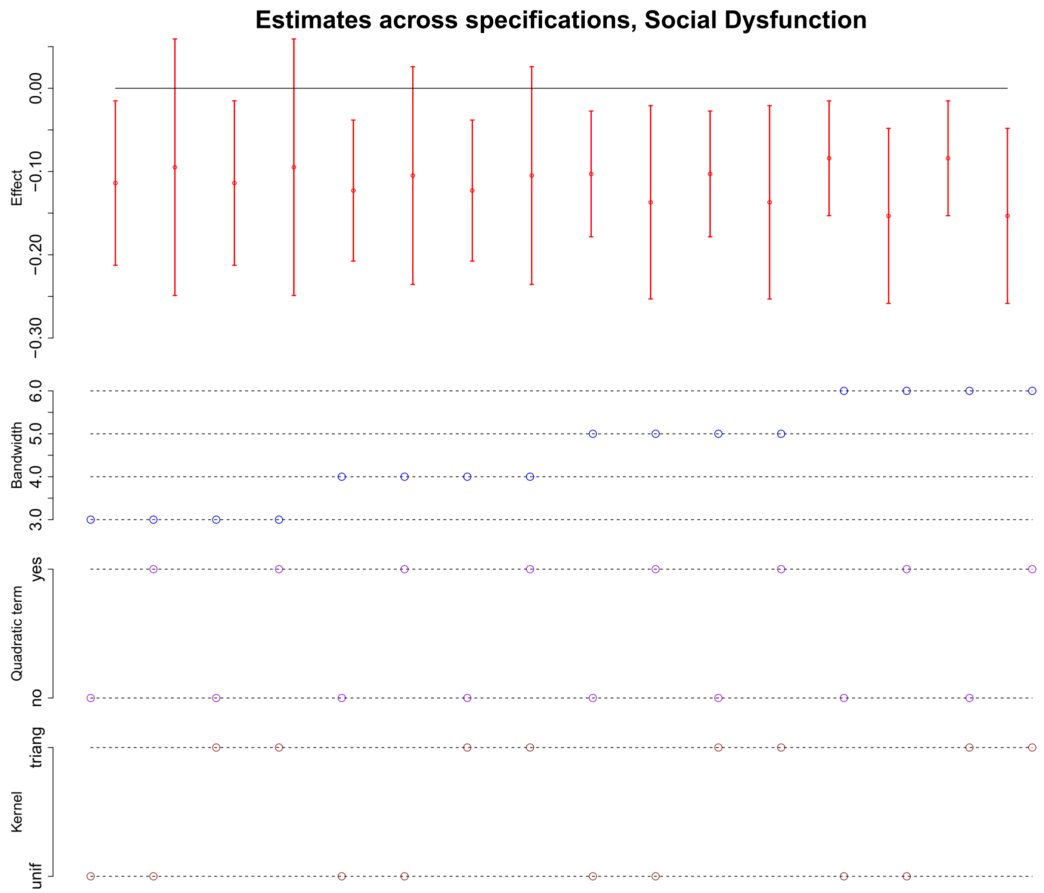

FIGURE C4 Effect of retirement eligibility on Social Dysfunction across specifications. Results are the average short-run causal effect of being eligible to retire on symptoms of Social Dysfunction. The range of estimation specifications is described in the note to Figure C1. Variables are defined in the note to Table 3. [Colour figure can be viewed at wileyonlinelibrary.com]

APPENDIX D

SENSITIVITY OF ESTIMATES BY OCCUPATION TYPE TO CHANGES IN SPECIFICATION

This section discusses and Figures D1–D4 show how the effect of retirement eligibility on mental health by occupational cluster depends on estimation choices. For example, in Figure D1, the top three panels show the estimated effect of gaining eligibility to retire from each of the three occupational clusters on the GHQ12 caseness. Each point is therefore an estimated causal effect of retirement eligibility on the GHQ12 caseness (and confidence interval) for a given set of estimation choices. The three panels below show the bandwidth used in estimation, whether a quadratic term in the running variable is included, and which type of kernel used.

The effect of retirement eligibility on the GHQ12 caseness index is greatest in cluster 3. Figure D1 shows how this effect varies by estimation choices. Being eligible to retire has a beneficial effect on the GHQ12 caseness measure amongst those in cluster 3 in every specification, and a beneficial effect amongst those in cluster 1 in most specifications (the exceptions are at small bandwidths). However, some estimates do not attain statistical significance. Including a quadratic term in the specification pushes estimates towards zero.

Figure D2 shows the range of estimates by specification for the causal effect on symptoms of Anxiety and Depression. Again, the beneficial effect of retirement eligibility is evident in all specifications for those in cluster 3, and most specifications for cluster 1. Again, the inclusion of a quadratic term is most important in changing the results.

In Figure D3, I show the pattern of estimates for symptoms of Loss of Confidence. There are no results which are statistically significant. Results are also close to zero in most specifications.

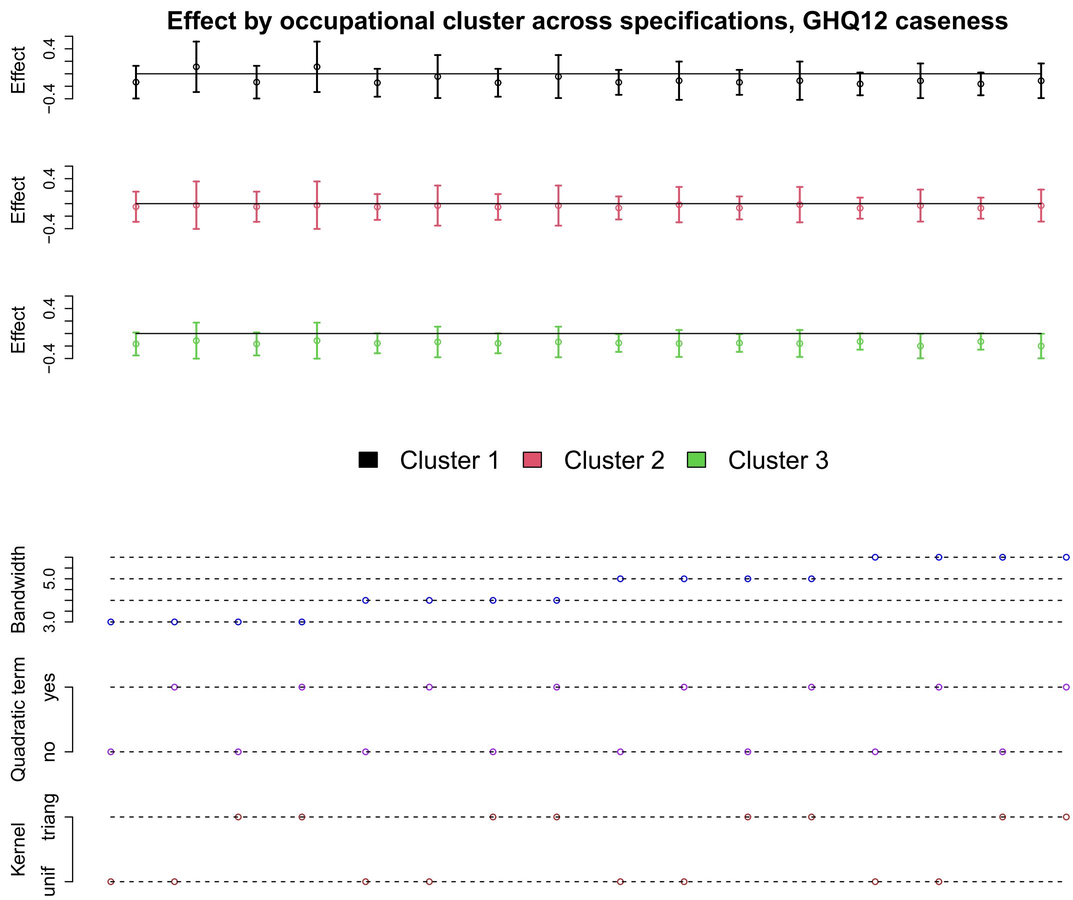

FIGURE D1 Effect of retirement eligibility on GHQ12 caseness across occupational clusters across different specifications. Results are obtained using a local linear or quadratic regression around the cutoff of 0 months to retirement, interacting each term in the regression with occupational cluster membership. The causal effects are the discontinuity interacted with the discontinutiy interacted with cluster membership (except for cluster one, which is the omitted value). The running variable is expected months to retirement. A bandwidth of 3–6 years is used; observations are weighted using a triangular or uniform kernel. Standard errors are heteroskedasticity robust and clustered at the individual level. Variables are defined in the note to Table 3. [Colour figure can be viewed at wileyonlinelibrary.com]

In my preferred specification (Equations (3) and (4) with a local linear regression, 5-year bandwidth, and triangular kernel), there is a beneficial effect of being eligible to retire on symptoms of Social Dysfunction for those in cluster three. This result is relatively consistent across specifications (see Figure D4). In most specifications, the estimated effect is statistically significant.

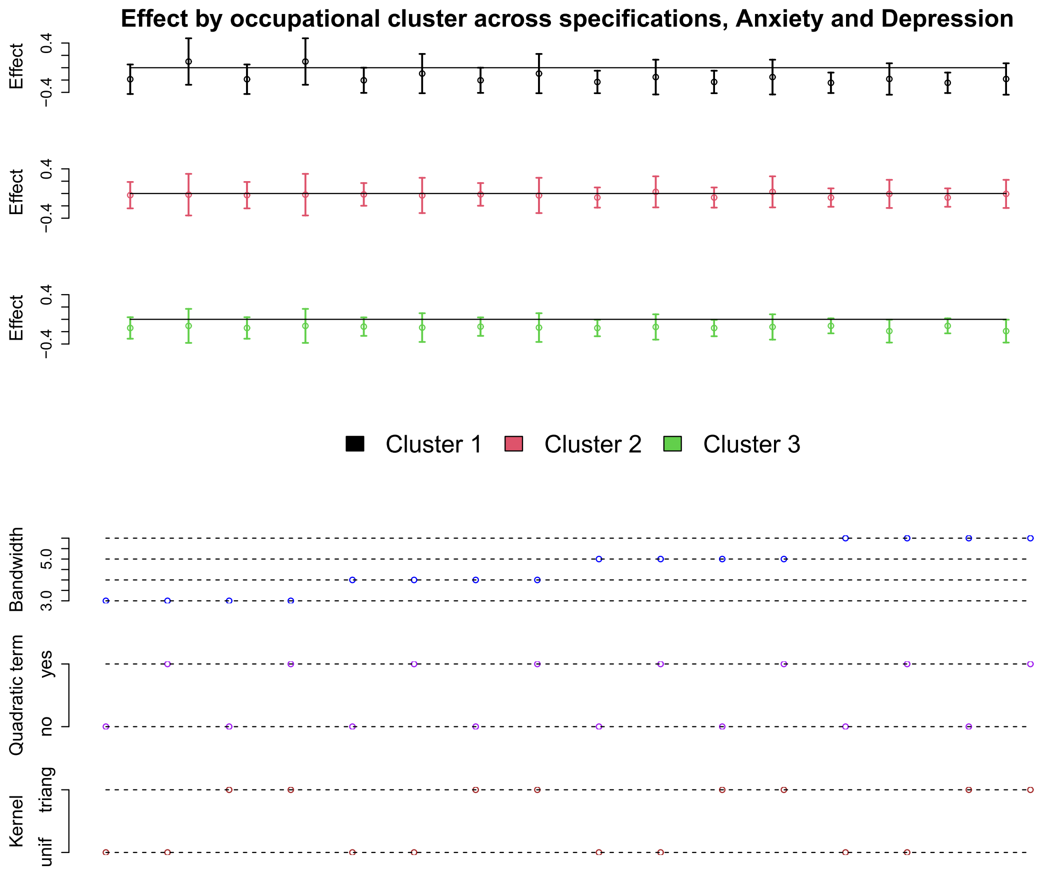

FIGURE D2 Effect of retirement eligibility on Anxiety and Depression across occupational clusters across different specifications. The range of specifications for estimation is described in the note to Figure D1. Standard errors are heteroskedasticity robust and clustered at the individual level. Variables are defined in the note to Table 3. [Colour figure can be viewed at wileyonlinelibrary.com]

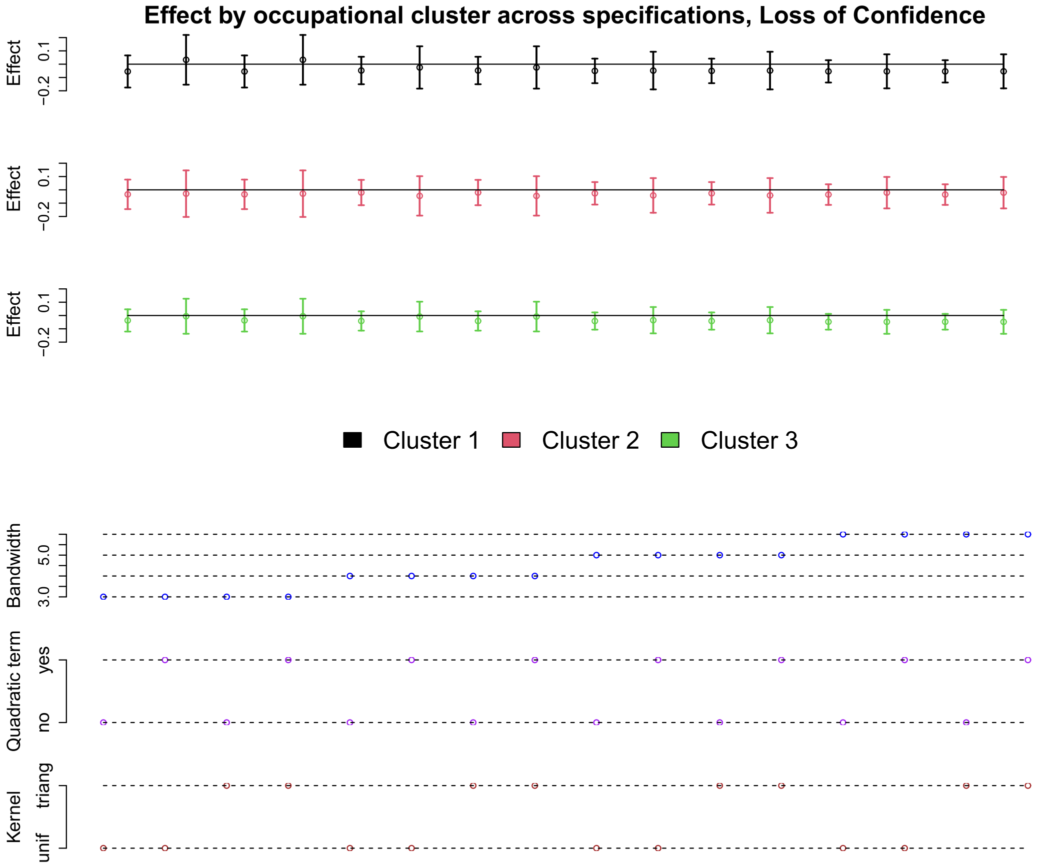

FIGURE D3 Effect of retirement eligibility on Loss of Confidence across occupational clusters across different specifications. The range of specifications for estimation is described in the note to Figure D1. Standard errors are heteroskedasticity robust and clustered at the individual level. Variables are defined in the note to Table 3. [Colour figure can be viewed at wileyonlinelibrary.com]

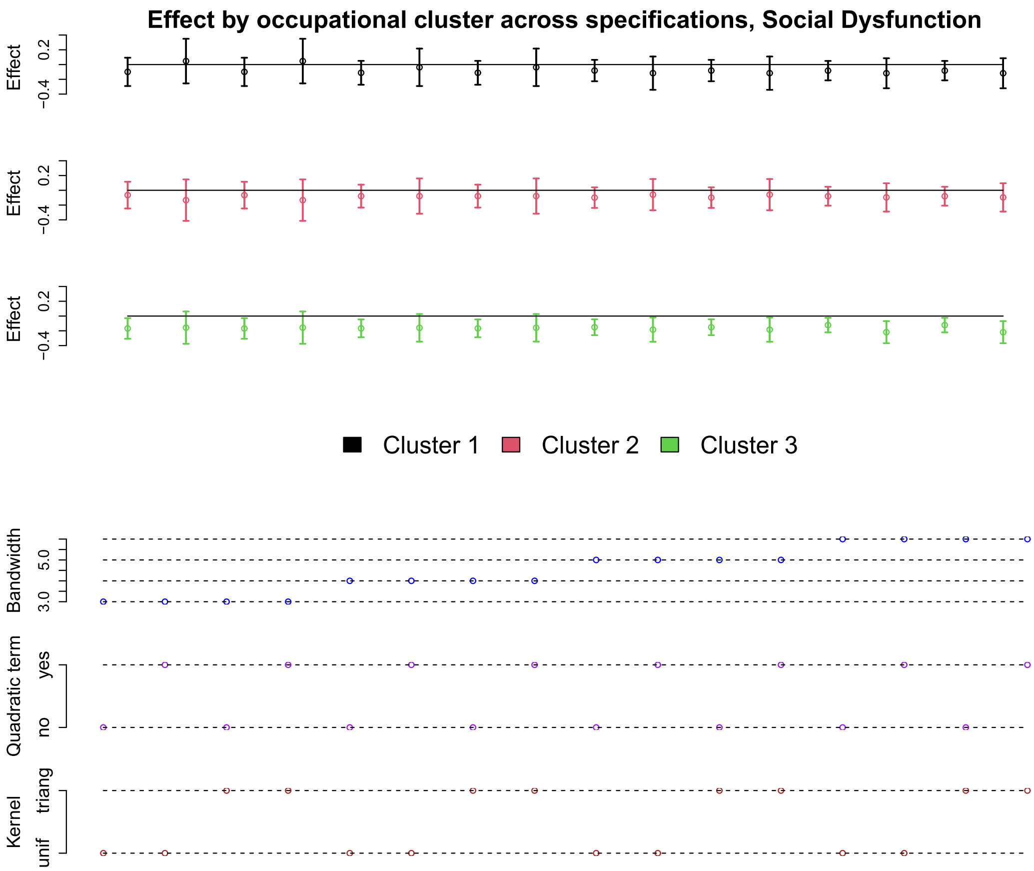

FIGURE D4 Effect of retirement eligibility on Social Dysfunction across occupational clusters across different specifications. The range of specifications for estimation is described in the note to Figure D1. Standard errors are heteroskedasticity robust and clustered at the individual level. Variables are defined in the note to Table 3. [Colour figure can be viewed at wileyonlinelibrary.com]

APPENDIX E

STANDARD ERROR ESTIMATES

Table E1 shows how the standard error changes with different estimators for the specification which investigates the average effect of retirement eligibility. For the most part, standard errors are similar when calculated in different ways. The exception is the "robust" standard error (Calonico et al., 2014). This method adjusts for the bias associated with larger bandwidths. I cannot calculate CCT standard errors for local linear estimators which include interaction terms. I therefore use OLS standard errors clustered at the individual level, and examine how results change as the bandwidth decreases (see Appendix C and D).

TABLE E1 Standard errors using different calculations, for each dependent variable.

	GHQ12 caseness	1 (caseness >8)	Loss of confidence	Social dysfunction	Anxiety and depression	Log household income	Retired
OLS	0.044	0.004	0.020	0.033	0.040	0.017	0.008
OLS clustered	0.044	0.004	0.020	0.033	0.040	0.017	0.008
cct conventional	0.050	0.004	0.023	0.038	0.046	0.020	0.009
cct bias-correct	0.050	0.004	0.023	0.038	0.046	0.020	0.009
cct robust	0.077	0.007	0.035	0.058	0.072	0.031	0.014
Honest KR	0.050	0.004	0.023	0.038	0.046	0.020	0.009

Note: Standard errors are calculated for an RDD estimator where the running variable is expected months to retirement eligibility. CCT standard errors, CCT bias-correct standard errors and CCT robust standard errors are as detailed by Calonico et al. (2014). Honest standard errors are those described by Kolesár and Rothe (2018). The estimation is described in the notes to Table 6. Variables are defined in the notes to Table 3. Note that I do not cluster standard errors by the running variable, age in months. Standard errors clustered by the running variable have poor coverage properties, that is, if the regression is correctly specified, the standard errors will tend to be too small (Kolesár & Rothe, 2018).

APPENDIX F

ANALYSIS USING FOUR OCCUPATIONAL CLUSTERS

In this section, I discuss the tradeoffs involved in choosing the number of clusters, and show that the results when the number of clusters increases to four are consistent with my main results.

Using a larger number of clusters yields a finer, more precise definition of an occupational type, but also reduces the statistical power and ability to detect heterogeneous effects. One way of quantifying the effect of increasing clusters is to observe how the share of variation explained by the cluster structure changes as the number of clusters increases. Figure F1 shows how the within-cluster sum of squares changes as the number of clusters increases. A smaller within-cluster sum of squares indicates that the cluster structure explains a larger share of the variation between occupations. We can see that the returns to increasing the number of clusters decreases significantly between three and four clusters, that is, the reduction in within-cluster sum of squares when we move from three to four clusters is small relative to the reduction in within-cluster sum of squares when we move from two to three clusters. This indicates that using three or four clusters is likely to deliver most of the benefits of clustering in explaining the structure of occupational characteristics while preserving precise estimates.

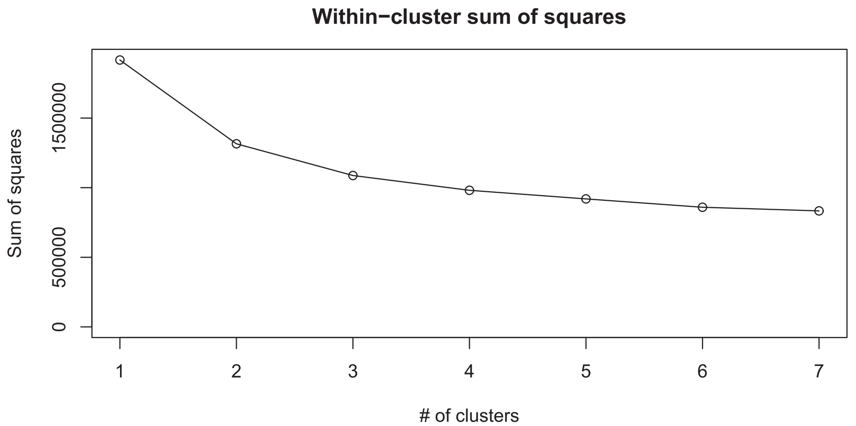

FIGURE F1 Within-cluster sum of squares at different number of clusters. I cluster person-wave observations by O*NET occupation data using K-means clustering. Within-cluster sum of squares is the sum of the Euclidian distance between data points and their centroid. I normalize by dividing by the total Euclidian distance between all data points.

Because the within-cluster sum of squares is similar for three and four clusters, I use three clusters within the main text to attain the most precise estimates. However, I also reproduce my main results using four clusters. The occupational mix of clusters produced, and the O*NET characteristics, are shown in Table F1 and Figure F2 respectively. Cluster one is over-represented amongst service workers, plant and machinery operator workers, and elementary occupations. These jobs score the lowest for cruciality of positioning and tight scheduling, are the least routine, and the least competitive. Cluster two occupations are over-represented amongst technicians and associate professionals, and to a lesser extent amongst senior officials and professionals. These jobs score highly for cruciality of position, conflictual

TABLE F1 Share of occupational employment by cluster.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Senior officials	0.022	0.339	0.612	0.027
Professionals	0.014	0.380	0.598	0.008
Technicians and associate professionals	0.091	0.513	0.366	0.030
Clerks	0.046	0.039	0.906	0.009
Service workers	0.639	0.302	0.042	0.017
Skilled agriculture	0.027	0.032	0.029	0.912
Craft and related trades	0.072	0.083	0.059	0.786
Plant and machinery operator workers	0.530	0.050	0.029	0.391
Elementary occupations	0.534	0.137	0.045	0.284

Note: The table shows the one-digit ISCO-88 occupation shares in each occupational cluster in the sample. I cluster person-wave observations by O*NET occupation data using K-means clustering.

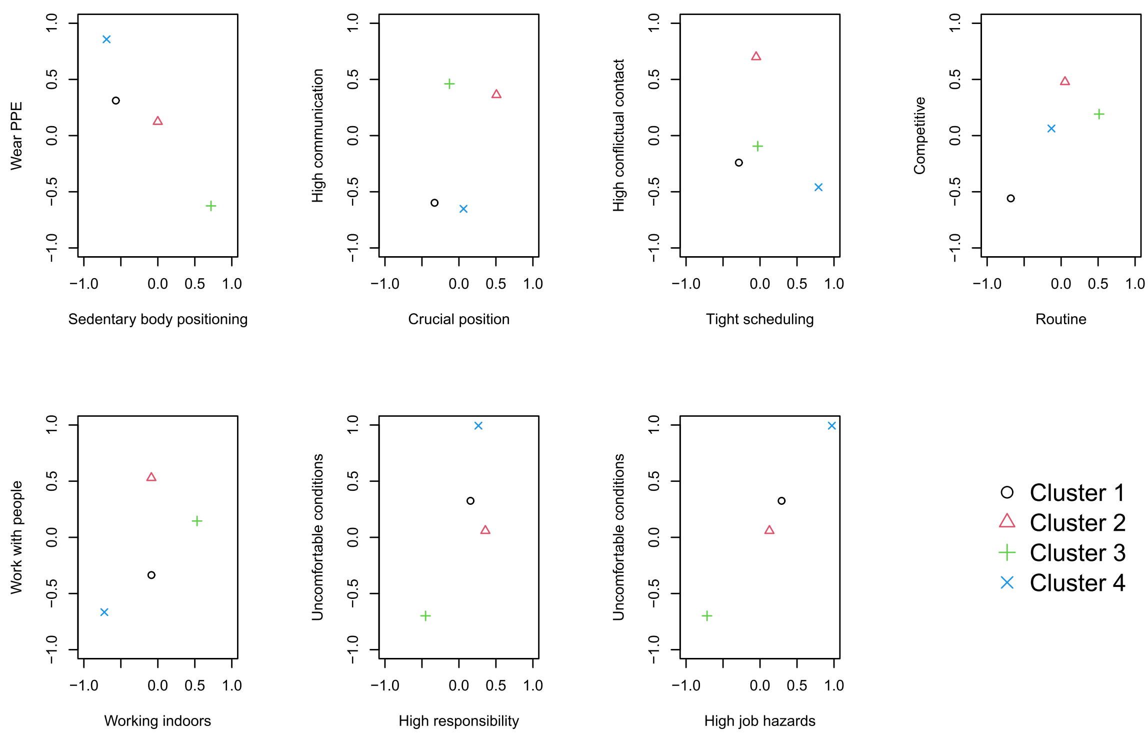

FIGURE F2 Percent of people in each cluster scoring high minus percent scoring low, four clusters. I cluster person-wave observations by O*NET occupation data using K-means clustering. The measures of cluster characteristics are calculated by subtracting the percentage of people who score "low" on an O*NET variable within a cluster from the percentage of people scoring "high", where scoring "high" or "low" is determined by being in the top 25% or bottom 25% respectively of the distribution of scores. Individuals are assigned to occupation clusters based on current or last occupation. O*NET indices are calculated based on the O*NET model, which groups variables into categories. I calculate the indices as the simple average over the variables in the index, which are signed to facilitate an intuitive interpretation (see Table 2). [Colour figure can be viewed at wileyonlinelibrary.com] contact, competitiveness, and working with people. Cluster three is over-represented amongst senior officials, professionals, and clerks. These jobs have the most sedentary body positioning, score highly for communication, are the most routine, and are the most indoor, low-hazard, comfortable jobs. Cluster four is most over-represented amongst skilled agricultural jobs, and in craft and related trades. These jobs score highest for wearing PPE, lowest for communication, highest for tight scheduling, lowest for working with people, lowest for working indoors, and highest for uncomfortable conditions and high job hazards.

There is not a direct mapping from the three cluster summary to the four cluster summary. Loosely, in terms of occupational employment and O*NET characteristics, cluster one (with three clusters) is similar in characteristics to cluster four (with four clusters); cluster three (three clusters) is similar to clusters two and three (with four clusters); and cluster two (with three clusters) is most similar to cluster one (four clusters).

Table F2 shows my main results where four clusters are used. With regard to the overall GHQ12 caseness, there is a higher benefit of retirement amongst people who work in cluster four occupations, and a higher (though statistically insignificant) benefit of retirement for people in clusters two and three. The benefit of retirement eligibility for those in cluster four comes predominantly from symptoms of Anxiety and Depression, while the benefit in cluster two comes from symptoms of Loss of Confidence, Social Dysfunction, and Anxiety and Depression. The benefit of retirement eligibility for those in cluster two is concentrated amongst men, who gain significantly more. For women in cluster four, their benefit of retirement eligibility is much larger than that for men, although for the GHQ12 caseness, this effect does not attain statistical significance.

Although we should not expect a direct mapping from cluster three results to cluster four results, the broad pattern is consistent. Cluster four (four clusters) workers have a higher than average benefit of retirement eligibility, concentrated in symptoms of Anxiety and Depression. This result is consistent with those in cluster one (three clusters) occupations having a higher than average benefit of retirement eligibility. Similarly, cluster two (four clusters) workers have somewhat elevated benefits from retirement eligibility, spread more evenly across symptoms of Anxiety and Depression, Loss of Confidence and Social Dysfunction. This result is consistent with those in cluster three (three clusters) having higher benefits from retirement eligibility spread across symptoms of Social Dysfunction and Anxiety and Depression. The inconsistencies are that, based on the three cluster results, we would think that either cluster two or three would have the highest overall benefit from retirement eligibility in the four-cluster analysis. Actually cluster four does. We would also not expect cluster two employees to gain in terms of symptoms of Loss of Confidence.

TABLE F2 The effect of retirement by occupational cluster (four clusters).

	GHQ12 caseness	1 (caseness >8)	Loss of confidence	Social dysfunction	Anxiety and depression	Log household income	Retired
Panel A: All							
Cluster 1 RD	−0.066	−0.003	−0.004	−0.032	−0.045	0.163***	0.268***
	(0.099)	(0.008)	(0.045)	(0.075)	(0.086)	(0.036)	(0.015)
Cluster 2 RD	−0.094	−0.003	−0.082**	−0.121*	−0.142*	0.137***	0.205***
	(0.091)	(0.007)	(0.041)	(0.071)	(0.083)	(0.037)	(0.015)
Cluster 3 RD	−0.108	−0.009	−0.034	−0.091	−0.067	0.105***	0.159***
	(0.078)	(0.006)	(0.035)	(0.060)	(0.072)	(0.030)	(0.013)
Cluster 4 RD	−0.238**	−0.008	−0.067	−0.099	−0.311***	0.142***	0.323***
	(0.107)	(0.008)	(0.049)	(0.079)	(0.098)	(0.036)	(0.017)
Obs	47,326	47,326	47,551	47,547	47,514	32,641	52,547
Panel B: Men							
Cluster 1 RD	−0.101	−0.001	0.015	−0.002	−0.066	0.132***	0.303***
	(0.124)	(0.010)	(0.058)	(0.093)	(0.111)	(0.037)	(0.019)
Cluster 2 RD	−0.143	−0.007	−0.075	−0.164**	−0.095	0.165***	0.212***
	(0.108)	(0.008)	(0.051)	(0.082)	(0.102)	(0.040)	(0.019)

TABLE F2 (Continued)

	GHQ12 caseness	1 (caseness >8)	Loss of confidence	Social dysfunction	Anxiety and depression	Log household income	Retired
Cluster 3 RD	−0.087	−0.007	−0.052	−0.050	−0.069	0.087***	0.185***
	(0.088)	(0.007)	(0.042)	(0.069)	(0.085)	(0.032)	(0.016)
Cluster 4 RD	−0.223**	−0.010	−0.062	−0.133*	−0.310***	0.137***	0.319***
	(0.106)	(0.008)	(0.050)	(0.078)	(0.101)	(0.037)	(0.018)
Obs	30,571	30,571	30,687	30,678	30,672	25,576	34,323
Panel C: Women							
Cluster 1 RD	−0.005	−0.005	−0.028	−0.055	−0.014	0.219*	0.246***
	(0.160)	(0.014)	(0.071)	(0.124)	(0.135)	(0.130)	(0.021)
Cluster 2 RD	−0.002	0.003	−0.107	−0.018	−0.264*	−0.112	0.234***
	(0.161)	(0.014)	(0.070)	(0.131)	(0.139)	(0.091)	(0.023)
Cluster 3 RD	−0.170	−0.018	−0.004	−0.183	−0.099	0.155	0.146***
	(0.153)	(0.013)	(0.065)	(0.119)	(0.132)	(0.105)	(0.022)
Cluster 4 RD	−0.655	−0.031	−0.181	−0.028	−0.653*	0.281**	0.428***
	(0.479)	(0.043)	(0.189)	(0.358)	(0.362)	(0.142)	(0.052)
Obs	16,755	16,755	16,864	16,869	16,842	7065	18,014

Note: Results are the short-run causal effect of being eligible to retire for those in each occupational cluster. I use a local linear regression about the cutoff of 0 months to retirement, interacted with occupational cluster membership as in Equation (4). I control for wave fixed effects. Cluster one is the omitted cluster. The causal effect for cluster one individuals is the discontinuity at the age of retirement eligibility. The causal effects for clusters two to four are the causal effect plus the coefficients on the terms which interact cluster membership and the discontinuity. I calculate the standard error of the causal effect by using the heteroskedasticity-robust variance-covariance matrix clustered at the individual level. I use a bandwidth of 5 years. I use a triangular kernel. Variables are defined in the note to Table 3.

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

APPENDIX G

ROBUSTNESS TESTS

I conduct a number of robustness tests. Firstly, I drop wave fixed effects. Results are shown in Table G1. The patterns are very similar to those of the baseline results. In the full sample, there is a beneficial effect of being eligible to retire on all measures of mental health and this benefit is higher for people in occupations in clusters one and three. In cluster one, the benefit is especially pronounced for symptoms of Anxiety and Depression. In cluster three, benefits are more evenly drawn from symptoms of Anxiety and Depression and Social Dysfunction.

Secondly, I investigate how my results change when I estimate the causal effect amongst only those who work up to retirement. I drop all individuals where (1) they are observed not working at any point within the bandwidth before the age of retirement eligibility or (2) it cannot be determined whether they were working at all points within the bandwidth before their age of retirement eligibility. This exercise involves dropping all individuals who never reach the age of retirement eligibility, and all individuals who enter the sample with fewer than 5 years until retirement eligibility. Correspondingly, standard errors are larger.

Results are shown in Table G2. Across all measures of mental health, there is a beneficial effect of retirement eligibility in the full sample. The effect is significant for overall caseness, symptoms of Social Dysfunction, and Anxiety and Depression. On the other hand, there is not as pronounced a benefit of retirement for those in cluster one occupations in this specification, though they still experience a reduction in symptoms of Anxiety and Depression, which is significant at the 10% level. There is a larger benefit of retirement on symptoms of mental illness for those in cluster three, with the benefit coming from symptoms of Social Dysfunction and Anxiety and Depression. These results are consistent with the benefits of retirement eligibility accruing mostly to those who are still working at the point they become eligible to retire in cluster three, but partially to those who are not working at the age of retirement eligibility in

TABLE G1 The effect of retirement eligibility by occupational cluster, excluding wave fixed effects.

	GHQ12 caseness	1 (caseness >8)	Loss of confidence	Social dysfunction	Anxiety and depression	Log household income	Retired
Panel A: Average effect of retirement eligibility							
All clusters	−0.119***	−0.008**	−0.038*	−0.126***	−0.116***	0.127***	0.224***
	(0.044)	(0.004)	(0.020)	(0.033)	(0.041)	(0.017)	(0.008)
Panel B: Effect of retirement eligibility across occupational clusters							
Cluster 1	−0.135	−0.007	−0.052	−0.103	−0.205**	0.134***	0.298***
	(0.089)	(0.008)	(0.041)	(0.066)	(0.082)	(0.032)	(0.015)
Cluster 2	−0.051	−0.005	−0.030	−0.089	−0.023	0.178***	0.234***
	(0.079)	(0.007)	(0.036)	(0.058)	(0.072)	(0.030)	(0.013)
Cluster 3	−0.161**	−0.012**	−0.039	−0.167***	−0.140**	0.089***	0.175***
	(0.067)	(0.006)	(0.030)	(0.049)	(0.061)	(0.025)	(0.011)
Obs	47,235	47,235	47,235	47,235	47,235	28,723	47,235

Note: Results are the short-run causal effect of being eligible to retire for those in each occupational cluster. I use a local linear estimator as described in the note of Table 6, although in this specification, I omit wave fixed effects. Variables are defined in the note to Table 3.

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

TABLE G2 The effect of retirement eligibility by occupational cluster amongst people who work until the SPA.

	GHQ12 caseness	1 (caseness >8)	Loss of confidence	Social dysfunction	Anxiety and depression	Log household income	Retired
Panel A: Average effect of retirement eligibility							
All clusters	−0.120**	−0.007	−0.043	−0.135***	−0.183***	−0.025	0.430***
	(0.059)	(0.005)	(0.027)	(0.044)	(0.054)	(0.020)	(0.009)
Panel B: Effect of retirement eligibility across occupational clusters							
Cluster 1	−0.080	0.001	−0.036	−0.056	−0.180*	0.027	0.503***
	(0.116)	(0.010)	(0.053)	(0.087)	(0.105)	(0.035)	(0.017)
Cluster 2	−0.039	−0.003	−0.070	−0.058	−0.134	0.036	0.434***
	(0.104)	(0.009)	(0.047)	(0.078)	(0.095)	(0.037)	(0.015)
Cluster 3	−0.212***	−0.014	−0.030	−0.241***	−0.235***	−0.110***	0.383***
	(0.092)	(0.008)	(0.042)	(0.069)	(0.083)	(0.030)	(0.013)
Obs	27,568	27,568	27,568	27,568	27,568	15,973	27,568

Note: Results are the short-run causal effect of being eligible to retire for those in each occupational cluster. The sample is all those people who have a full work history within the bandwidth up to the age of retirement, and are always observed to be in work. I use a local linear estimator as described in the note of Table 6. Variables are defined in the note to Table 3.

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

cluster one. If those who benefit from being eligible to retire are working, this suggests the benefit runs through the decision to retire, that is, the benefit of retiring from cluster three jobs is high. On the other hand, it is possible that many people in cluster one who benefit from being eligible to retire have stopped working by the time they reach retirement age, and benefit predominantly from the added economic security that BSP benefits give them.