Life after work: The causal effect of retirement on wellbeing

Calvin Cheng[1]

[1]Institute of Strategic and International Studies (ISIS) Malaysia.
calvin.ckw@isis.org.my, calvinchengkw@hotmail.com

Abstract

I study the effects of retirement on wellbeing using a fuzzy regression discontinuity design that exploits institutional rules surrounding access to pension funds in Malaysia. Using a Wave 1 of the Malaysian Ageing and Retirement Survey (MARS), I find that retirement has a significant and positive effect on total, mental, and social wellbeing. My estimates suggest that individuals who are induced to retire experience wellbeing improvements that exceed the magnitudes typically associated with major life events such as the birth of a first child or marriage, as reported in the existing life satisfaction literature. I also find that the retirement-wellbeing link is heterogeneous across individuals with different socio-economic and demographic characteristics. Individuals with lower education, assets, or chronically ill experience lower or negative wellbeing gains after retirement.

* * *

1 Introduction and background

For many, the transition from decades of work to a life of leisure can be daunting. Retirement can upend routines, social networks, and incomes, in turn creating significant impacts on an individual's stress and wellbeing. Yet, even as the global share of retirees continues to grow as populations age – public pension systems in many countries across the world are buckling under the weight of higher life expectancies. This, coupled with mounting inflationary pressures and the rapidly shifting nature of work, means that understanding the causal relationship between retirement and wellbeing is becoming increasingly urgent for the design of effective policies that safeguard the welfare of the growing number of individuals transitioning into retirement.

Despite this, the existing literature on the causal link between retirement and wellbeing has been mixed.

Some studies have demonstrated a positive link between retirement and wellbeing – while others have shown that it decreases wellbeing and life satisfaction (Kuhn, 2018; Picchio and van Ours, 2019). Kesavayuth et al. (2020) use the eligibility age for state pensions as an instrument for retirement status, showing that retired individuals self-report higher life satisfaction. Similarly, Yemiscigil et al. (2021) also employ an instrumental variables approach on US data, concluding that life satisfaction and purpose increased after retirement. Picchio and van Ours (2019) use a fuzzy regression discontinuity with data from the Netherlands, concluding that retiring has different effects on wellbeing – positive for partnered individuals and negative for single individuals. Conversely, Dave et al. (2006) and Sahlgren (2013) demonstrate through estimating dynamic models that retirement led to an increase in reported illnesses and a decline in mental health over time. Likewise, Atalay et al. (2019) uses two-period Australian panel data to conclude that retirement has a modest, negative effect on cognition.

One major reason for this discrepancy in the literature is that studying the causal effect of retirement and wellbeing is fraught with analytical challenges. For one, retirement decisions are typically endogenous. They are influenced by observed factors like wealth and assets, but also by unobserved factors like individual preferences and health, both of which can have profound effects on well-being. This also raises concerns about the direction of causality. Retirement influences wellbeing, but wellbeing also influences retirement timing (Kuhn, 2018). To overcome this challenge, I use a fuzzy regression discontinuity approach, exploiting a discontinuity at age 55, which is the age at which individuals are eligible to access pension funds in Malaysia. This paper contributes to the body of evidence surrounding the retirement-wellbeing link, extending the literature into a developing Asian country context using a brand new dataset from Malaysia – Wave 1 of the Malaysian Ageing and Retirement Survey (MARS).

The rest of the paper is organized as follows: the following section introduces the MARS dataset, describes variable creation, and presents some descriptive statistics. Subsequently, Section 3 outlines the empirical and identification strategy and model specifications. Section 4 presents some graphical analysis as well as the main results. Section 5 explores heterogeneity and potential channels using a propensity score matching approach and subsample analysis. Section 6 reports various robustness checks and placebo tests. The final section concludes with a brief discussion of the policy implications of the main results.

2 Data

2.1 The MARS dataset and variable construction

The dataset in use is a single wave of a nationally representative longitudinal survey of people aged 40 years and older in Malaysia, drawn from the newly published Malaysian Ageing and Retirement Survey (MARS), which was made available privately through the Social Wellbeing Research Centre (SWRC) at the University of Malaya. The MARS dataset includes information on 5,613 individuals on variety of demographic, work status, health, economic, and psycho-social indicators (Mansor et al., 2019). The key explanatory variable is retirement status. For this essay, I follow standard economic conceptions of retirement, defined as a permanent withdrawal from the labour force, i.e. a transition from working to not working (Michaels, 2023). As such, I consider an individual retired if they list their work status as "retired", have reported working before retirement, and is currently not working for pay.

For the outcome variable of interest, i.e. indicators of wellbeing, I use the MARS dataset to construct three different composite wellbeing variables. First, I construct a subjective 'negative affect' mental wellbeing indicator, which is a composite measure of self-reported stress, anxiety, depression, sadness, fulfilment, and feelings of isolation. Second, I create a 'social' wellbeing indicator, which is a composite of self-reported social interactions and activities that are thought to contribute to wellbeing – including time spent with family and social activities, exercise, volunteer work, and on hobbies. Third, I combine both the mental and social wellbeing indicators to create a 'total' wellbeing indicator. In constructing each of the composite wellbeing indicators, I follow the standard practice to order and orient each underlying variable to produce a 5-point scale, such that higher values indicate greater wellbeing (Picchio and van Ours, 2019).[1] The composite indicators are then averaged out to preserve the 5-point scale. This type of ordinal point scale is common in the life satisfaction literature, one notable example being the widely-used Satisfaction With Life Scale (Clark et al., 2008; Picchio and van Ours, 2019; Titov et al., 2022).

2.2 Descriptive statistics

Table 1 provides summary statistics of the retirement indicator, age and various other control variables along with the wellbeing outcome variables. Figs. 1 and 2 show the density of individuals along the running variable. Overall, retired individuals skew much older than non-retired individuals, but there appears to be no immediate evidence of any potential irregularities or anomalies along the age distribution. The summary statistics indicate that, as expected, there is a significant amount of self-selection into retirement. Retired

[1] For example, a variable negative for wellbeing like 'stress' would be recoded from 5 = very stressed to 5 = never stressed, before combining to create the composite wellbeing indicators

and non-retired individuals are significantly different in terms of observable (and very likely unobservable) characteristics. The mean age of the 5,613-person sample is about 57 years old, while the mean age for retired and non-retired individuals is 65 years and 55 years respectively. Women make up about 55.8 percent of the sample, but less than 28 percent of retired individuals. Retired individuals also have less education, more reported illnesses, larger family sizes, and fewer savings – and this difference is statistically significant from zero (see last column in Table 1). Consequently, a simple comparison of the wellbeing of retired and non-retired individuals would likely lead to biased estimates due to endogeneity. This is expanded further in section 3, and motivates the empirical strategy used in this paper.

Table 1: **Descriptive statistics**

	Total		Retired		Non-retired		
	Mean	Std dev.	Mean	Std dev.	Mean	Std dev.	Difference
Covariate							
Age	57.169	10.770	64.793	8.690	55.382	10.425	-9.411***
Gender	0.558	0.497	0.279	0.449	0.623	0.485	0.345***
Education	8.036	3.923	8.523	3.765	7.922	3.950	-0.602***
Marital status	0.776	0.417	0.774	0.418	0.776	0.417	0.002
Ethnicity	4.692	15.912	3.871	13.882	4.885	16.347	1.015*
Illnesses	1.068	1.201	1.480	1.407	0.971	1.126	-0.509***
Health	2.568	0.807	2.710	0.818	2.535	0.801	-0.175***
Number of children	3.721	2.448	3.859	2.418	3.684	2.456	-0.175
Malaysian born	1.204	0.881	1.169	0.805	1.212	0.897	0.043
Savings	24309.274	445432.379	23006.557	160676.635	24614.684	488760.526	1608.126
Outcome variables							
Total wellbeing	3.334	0.584	3.398	0.609	3.319	0.578	-0.078***
Social wellbeing	2.819	0.805	2.876	0.839	2.806	0.796	-0.070*
Mental wellbeing	3.935	0.638	4.006	0.634	3.918	0.638	-0.088***
Observations	5613		1066		4547		5613

3 Empirical strategy

3.1 Identification and theoretical setup

As mentioned earlier, it is well-established that retirement decisions are typically endogenous (Kuhn, 2018; Picchio and van Ours, 2019). Individuals self-select into retirement based on observed and unobserved factors like motivation, health, or labour market experiences (Picchio and van Ours, 2019). Indeed, the previous section demonstrates that retired and non-retired individuals in our sample are systematically different and very likely have different potential outcomes. Similarly, the relationship between wellbeing and retirement is

Figure 1: **Age**

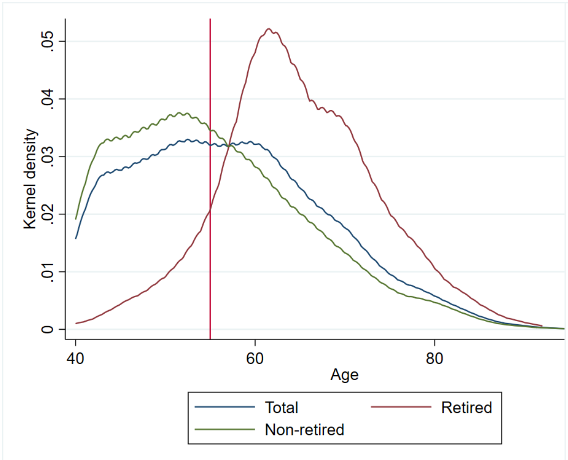
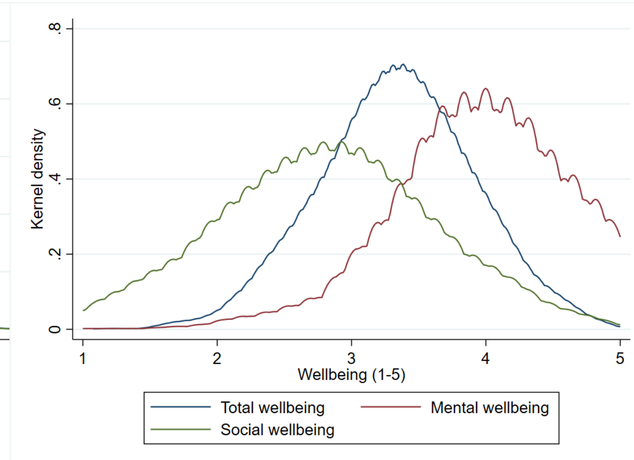


Figure 2: **Wellbeing**



Note: The vertical line in Figure 1 denotes age 55. Densities estimated via Epanechnikov kernel functions.

bidirectional: retirement affects wellbeing, but wellbeing itself also affects an individual's decision to retire. As such, a naive comparison of the wellbeing of retired and non-retired individuals would be contaminated by selection bias – driven by numerous pre-existing differences between those who choose to retire and those who did not.

Consequently, to overcome these endogeneity concerns, I exploit institutional rules surrounding eligibility for receipt of the state pension at age 55. The statutory retirement age in Malaysia is 60 years old, but individuals become eligible to access pension funds from the Employees Provident Fund (EPF) at the age of 55 (Kumpulan Simpanan Wang Pekerja, n.d.). Based on conventional labour supply theory and the life-cycle model of consumption and leisure, individuals make decisions to maximise the present discounted value of lifetime utility by choosing consumption and leisure subject to budget and time constraints (Heckman, 1974). Individuals with more financial resources are able to choose more leisure and retire earlier, while those with fewer resources will need to work longer to be able to consumption smooth across their remaining lifetime (Michaels, 2023). As such, an increase in access to pension benefits induced by the age 55 threshold provides a source of exogenous variation by decreasing the price of retirement for individuals just above the age 55 threshold relative to those just below. This decrease in the price of retirement leads to an increase in the demand for leisure (as a normal good), which in turn induces a discontinuity in retirement probability at the age 55 threshold. This motivates the use of a fuzzy regression discontinuity (RD) design to analyze the causal effect of retirement on wellbeing (Picchio and van Ours, 2019).

In this context, the relationship between the running variable and treatment is probabilistic instead of deterministic. That is, retirement probability 'jumps' at the pension eligibility age (55), but passing age 55 does not sharply induce all individuals to retire. The fuzzy RD design relies on this 'jump' in retirement probability when an individual reaches the pension eligibility age threshold to identify the treatment effect. (Imbens and Lemieux, 2008).[2] Defining D_i as the treatment (retirement) such that $D_i \in \{0, 1\}$, X_i as the running variable (age), x_0 is the threshold value of the running variable (age 55), and $\lim_{x \downarrow x_0}$ and $\lim_{x \uparrow x_0}$ as denoting the limit of the treatment probability as the running variable approaches the threshold x_0 from either side, we can then, under continuity conditions elaborated below, outline the canonical fuzzy RD estimand τ for individuals arbitrarily close to the threshold (Cattaneo and Titiunik, 2022; Imbens and Lemieux, 2008):

$$\tau = \frac{\lim_{x \downarrow x_0} \mathbb{E}[Y|X = x] - \lim_{x \uparrow x_0} \mathbb{E}[Y|X = x]}{\lim_{x \downarrow x_0} \mathbb{E}[D|X = x] - \lim_{x \uparrow x_0} \mathbb{E}[D|X = x]}$$

Here, the estimand τ is a ratio of the discontinuity in the regression of the outcome variable at the cutoff, divided by the discontinuity in the probability of treatment (Imbens and Lemieux, 2008). Put differently, since treatment take-up is not perfect (not all individuals past 55 years of age choose to retire), to estimate the average causal effect of receiving treatment, the estimand needs to rescale the 'jump' in outcome variable by the 'jump' in treatment probability induced by the threshold (Pischke, 2018, p. 15).

The validity of this fuzzy RD design hinges on three key identifying assumptions. The *first* is that there is indeed a discontinuity in the probability of receiving treatment at the threshold, i.e. that probability of retirement actually changes at the pension-eligibility age threshold x_0 (age 55) (Imbens and Lemieux, 2008):

$$\lim_{x \downarrow x_0} Pr(D_i = 1|X_i = x) \neq \lim_{x \uparrow x_0} Pr(D_i = 1|X_i = x)$$

The *second* key identifying assumption is continuity in treatment and outcomes around the threshold (Hahn et al., 2001). Formally, potential outcomes $\mathbb{E}[Y_i(Z)|X_i = x]$ should be continuous around the cutoff $X_i = x_0$ at either side of the threshold for $Z = 0, 1$ where $Z = 1\{X_i \geq x_0\}$ (Cattaneo and Titiunik, 2022). This implies that other predictors of retirement are itself do not jump at the threshold, and as such individuals close to the threshold are comparable and have the same potential outcomes (Picchio and van Ours, 2019). The *third* key identifying assumption is that there is a positive density of values of x around the threshold x_0. This precludes any "bunching" or discontinuities in the density function of the score around the threshold – which may indicate that individuals were able to precisely manipulate the running variable to end up on one side of the threshold (Cattaneo and Titiunik, 2022). Such manipulation at the threshold would invalidate the

[2] We can ignore this "fuzziness" and run a sharp regression discontinuity, which would then recover the causal effect of treatment assignment, instead of the causal effect of receiving treatment – an intention-to-treat (ITT) effect (Cattaneo and Titiunik, 2022)

RD design, because treatment assignment at the threshold would no longer be as-if random. To ensure the robustness of the design, I test each of these key identifying assumptions in Section 6.

Subsequently, to give the estimand τ as defined above a concrete causal interpretation in the fuzzy RD context, I take the approach of interpreting τ as the average causal effect of retirement for the subgroup of compliers with X_i arbitrarily close to x_0 (Cattaneo and Titiunik, 2022). This requires two main additional assumptions. The first is monotonicity: $D_i(x)$ is non-increasing in x at $x = x_0$, i.e. probability of treatment jumps at threshold, but the jump is in the same direction for all i (Hahn et al., 2001). The second is the exclusion restriction, that the only effect of passing the age threshold on the outcome (wellbeing) is through retirement – i.e. that there is no direct effect of the threshold on the outcome (Cattaneo and Titiunik, 2022).[3]

If these assumptions hold, we can interpret τ from above as a local average treatment effect (LATE) for compliers close to the threshold. That is, $\tau_{LATE} = \mathbb{E}[Y_{1i} - Y_{0i} \,|\, Complier = 1 \; and \; X_i = x_0]$, where compliers are defined as individuals who have $\lim_{x \downarrow x_0} D_i = 0; and; \lim_{x \uparrow x_0} D_i = 1$ (Imbens and Lemieux, 2008). In this context, the estimand τ would be the average causal effect of retirement, averaged over individuals near the threshold who "complied" with the pension eligibility encouragement and changed their retirement status at the threshold. This local LATE may not generalise to individuals further away from the threshold or to individuals who are insensitive to incentives to retire induced by the threshold. Nonetheless, this local LATE is still externally valid and policy-relevant for informing the design of retirement policy. For one, it captures the causal effect of retirement on wellbeing for an important subgroup: individuals who are close to the retirement age and who are the most responsive to retirement policy changes or pension eligibility thresholds at the margin. As Malaysia looks to raising the pension eligibility age threshold in the coming years (“Increase in EPF withdrawal age worth a second look?” 2023), understanding the effect of retiring for individuals at the margins would be crucial for gauging the potential impacts of such a policy change.

3.2 Model specification

As is standard for linear fuzzy RD designs, I estimate the causal effect of retirement on wellbeing using two-stage least squares (2SLS). This entails a first-stage regression of retirement status (D_i) on the threshold instrument Z_i, where $Z_i = 1\{X_i \geq x_0\}$ along with a function of the running variable. The second stage involves regressing the wellbeing indicator (Y_i) on the predicted retirement status from the first stage regressions ($\widehat{D_i}$), along with the same age polynomials. The running variable is centred around the age-55 threshold ($X_i - x_0$). Additionally, $h(X_i)$ is a flexible function in X_i, and W_i is a vector of covariates. Then,

[3] The other LATE assumptions for the validity of the instrument including first-stage/relevance and independence are captured by the other identifying assumptions above, namely the discontinuity and continuity assumptions respectively

the general model is specified as such:

$$D_i = \alpha + \delta Z_i + h(X_i - x_0) + \Phi(W_i) + \varepsilon_i$$
$$Y_i = \mu + \delta \widehat{D_i} + h(X_i - x_0) + \Phi(W_i) + u_i$$

In the preferred specification below, I estimate a parametric fuzzy RD with two age polynomials – in line with the usual age dynamics specified in the literature – along with controls for different pre-treatment characteristics. In subsequent specifications, I explore the inclusion of additional covariates, different age polynomials, allowing for separate slopes fitted above and below the threshold, and non-parametric approaches with different bandwidths.

4 Estimation

4.1 Estimation sample

For most of the main model estimates in the following sections (Section 4.2 and Section 5), I trim the sample of individuals to those aged 75 and below in the main RD specifications, effectively setting a bandwidth of 15 on either side of the threshold and decreasing the sample from 5,613 to 5,273 individuals. This is done to balance the support points and age bins around the threshold, with the aim of improving comparability between units just below and above the threshold. Additionally, the relevance of retirement decisions on wellbeing may break down at ages too far above the average life expectancy in Malaysia (76 years). For completeness, I also report estimates from the model using the full age (40-95 years old) sample in the main results below.

4.2 Standard errors and bootstrap method

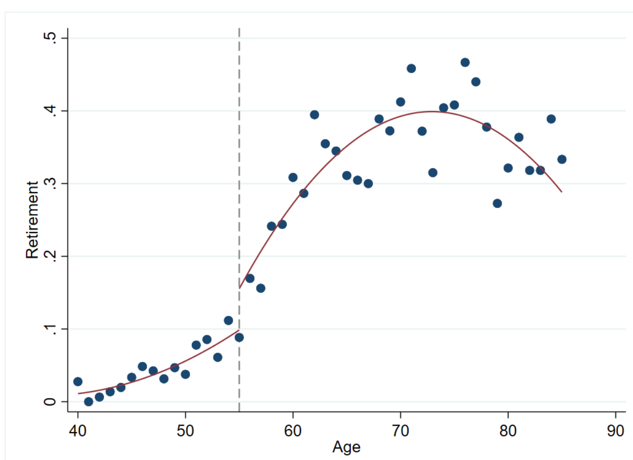
Following the convention in the regression discontinuity literature, I cluster the standard errors on age (Lee and Card, 2008). This is to account for within-group correlations, as we may expect individuals of the same age cluster to share similar characteristics. However, after trimming the data to individuals aged 75 and below, we are left with a relatively low number of clusters (36). Running conventional clustered standard errors with this number of clusters may lead to an underestimation of standard errors (Özler, 2012). As such, I use a wild bootstrap approach to better maintain the original within-cluster correlation structure in the data and provide a more accurate estimate of the sampling distribution even with a relatively low number of clusters. In the main results below, I report wild bootstrap bias-corrected p-values instead of standard errors along with the point estimates.

4.3 Graphical analysis

The discontinuity in retirement probability at the threshold

As discussed, the first key identifying assumption is that there needs to be a discontinuity in the retirement probability at the threshold age where individuals become eligible for the state pension. Within an instrumental variables framework, the threshold instrument ($Z_i = 1\{X_i \geq x_0\}$) needs to be relevant in predicting retirement. Figure 3 shows the relationship between retirement and the running variable, with a 2nd-order polynomial regression fitted separately on each side of the threshold with uniform kernel weights. The threshold instrument does indeed induce a modest discontinuity in retirement probability. Retirement probability gradually rises quite linearly starting from age 40 (the youngest age in the dataset) and there is a structural break at age 55. After 55, retirement probability exhibits a quadratic pattern, with individuals above age 75 experiencing a decline in retirement probability. At age 55, there is a modest but statistically significant jump in retirement probability of about 5.7 percentage points (Fig. 3). Though this jump is statistically significant, and the F-stat of the overall regression is 174.96, this increase in retirement probability induced by the threshold is relatively small compared to that observed in other studies. For example, Picchio and van Ours (2019) find a jump in retirement probability of about 20 to 37 percent at the pension eligibility age threshold in the Netherlands.

Figure 3: **The discontinuity in retirement probability**



Note: Each point is a local sample mean calculated for each bin. Bins are selected using equally-spaced one year intervals. Regression lines are obtained by fitting 2nd order polynomials on each side separately with uniform kernel weights. The sample is restricted to individuals aged 85 and below. The jump in retirement probability at the threshold is about 5.7 percentage points, statistically different from zero with a p-value of 0.00. The F-stat of the overall regression is 174.96, larger than the F>10 rule of thumb recommended by (Stock and Yogo, 2005). Nonetheless, the instrument only induces a small proportion of compliers.

Wellbeing measures at the threshold

Figs. 4 and 5 shows the structural relationship between age and the composite wellbeing indicators, obtained by fitting regressions with uniform kernel weights on both sides of the threshold. Figure 4 shows that, similar to retirement probability, mental wellbeing decreases linearly before the age 55 threshold, then evolves in a quadratic form – first increasing slightly until about age 70 and then subsequently declining. On the contrary, social wellbeing decreases linearly before age 55 and then continues to steadily deteriorate. For both mental and social wellbeing, there is an apparent and statistically significant discontinuity at the age 55 threshold.

4.4 Main results

Table 2 compiles the main estimates for the effect of retirement on wellbeing for each composite indicator, across six different model specifications. The different models (1-6) differ in the order of polynomial fitted, covariates included, bandwidth used, and whether or not a different slope was fitted above and below the threshold. Models 1-4 are estimated parametrically with a restricted age sample as discussed above, model 5 is a non-parametric specification with MSE-optimal bandwidths, and model 6 is estimated with the full unrestricted age sample, using the same specification as the baseline model (Model 4). All specifications were estimated using 2-stage least squares (2SLS), with bias-corrected p-values derived via wild bootstrap on standard errors clustered by age.

Figure 4: **Mental wellbeing**

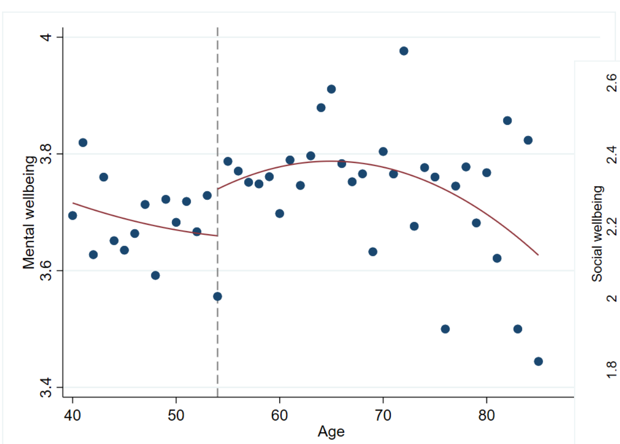
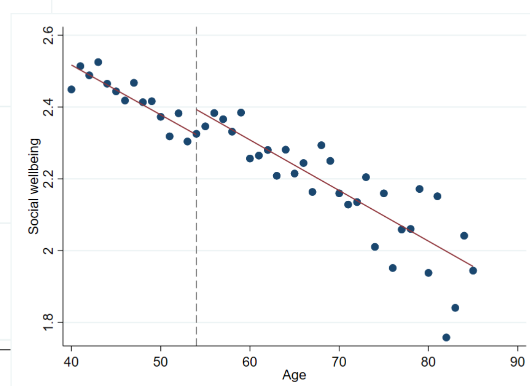


Figure 5: **Social wellbeing**



Note: Each point is a local sample mean calculated for each bin. Bins are selected using equally-spaced one year intervals. Regression lines are obtained by fitting 2nd-order (Figure 2) or 1st-order (Figure 3) polynomials on each side separately with uniform kernel weights. The sample is restricted to individuals aged 85 and below.

The results across the different specifications indicate that retirement has a positive and statistically significant effect on wellbeing (see Table 2), with the exception of the non-parametric model with MSE-optimal bandwidths (Model 5). Overall, the estimated treatment effects of retirement on total wellbeing range from about 0.6 points to 1.1 points across the different models. Broadly, across the different specifications, the estimates suggest that the effect of retirement on social wellbeing (ranging from 0.9 points to 1.6 points) is higher than the estimated effect on mental wellbeing (ranging from 0.5 points to 0.7 points). As explained in Section 2, in interpreting the magnitudes of these coefficients, a 1-point increase is equivalent to one step up on an ordinal 5-point scale like the Satisfaction With Life Scale (Clark et al., 2008).

The preferred baseline specification is Model 4 in Table 2, with a 2nd-order polynomial fitted parametri-

Table 2: **Main results**

Model	(1)			(2)			(3)		
	Total	Mental	Social	**Total**	Mental	Social	**Total**	Mental	Social
Retired	1.052***	0.671*	1.362**	0.624**	0.635**	0.601**	1.147***	0.591**	1.630***
	(0.003)	(0.054)	(0.044)	(0.013)	(0.048)	(0.048)	(0.008)	(0.030)	(0.010)
Polynomial (below)	1			2			1		
Polynomial (above)	1			2			2		
Separate#	no			no			yes		
Controls##	no			yes			yes		
Type	parametric			parametric			parametric		
Bandwidth	15,15			15,15			15,15		
Clustered SEs	yes			yes			yes		
Resampling method	Wild bootstrap			Wild bootstrap			Wild bootstrap		

	(4)			(5)			(6)		
	Total	Mental	Social	**Total**	Mental	Social	**Total**	Mental	Social
Retired	0.775***	0.490***	1.024**	-2.397	-3.992	-1.520	0.855*	0.745	0.953*
	(0.006)	(0.007)	(0.015)	(0.576)	(0.737)	(0.829)	(0.081)	(0.225)	(0.096)
Polynomial (above)	2			1			2		
Polynomial (below)	2			1			2		
Separate#	yes			yes			yes		
Controls##	yes			yes			yes		
Type	parametric			non-parametric			parametric		
Bandwidth	15,15			MSE-optimal			None		
Clustered SEs	yes			Robust standard errors			yes		
Resampling method	Wild bootstrap			None			Wild bootstrap		

Note:* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. #denotes a separate slope fitted for each side of the threshold. ##controls: gender, education, immigrant status, reported chronic illnesses. Bias-corrected p-values derived via wild bootstrap on standard errors clustered by age between brackets. 'Mental' and 'social' wellbeing are composite wellbeing indices as described in Section 2.

cally above and below the threshold, along with additional controls for gender, education, and immigrant status. The results of this baseline specification indicate retirement has a statistically significant average marginal effect of increasing total wellbeing by 0.775 points, mental wellbeing by 0.490 points and social wellbeing by 1.024 points. This is a large magnitude. To put this in perspective, life satisfaction studies have found that major life events like getting married and experiencing the birth of a first child raise life satisfaction by 0.25 points and 0.5 points on a 5-point scale respectively (Clark et al., 2008, p. 331 and Myrskylä and Margolis, 2014, p. 1853). Using these numbers as a rough comparison, these baseline estimates suggest that the total wellbeing improvements from retirement are equivalent to around 1.6 times the wellbeing gains experienced from the birth of a first child (Myrskylä and Margolis, 2014, p. 1853). Likewise the baseline estimates indicate a social wellbeing improvement of about two times that of a first child – and a mental wellbeing improvement roughly on par with a first childbirth.

The non-parametric specification in Models 5 in Table 2 was included to allow a flexible approach without imposing restrictive functional form assumptions. Non-parametric methods restrict the data used to estimate local treatment effects to a small bandwidth around the threshold, which may reduce bias from using data further away from the threshold for estimation (Imbens and Lemieux, 2008). Yet, non-parametric methods entail a bias-variance tradeoff – requiring sufficient data within the bandwidth to yield precise estimates (Gelman and Imbens, 2014). In terms of bandwidth selection, Model 5 uses optimal bandwidths selected via a single common mean squared error (MSE) optimal bandwidth selection procedure (Cattaneo et al., 2018). However, the estimates obtained in Model 5 are extremely non-significant – with p-values above 0.5 and estimated coefficients that are opposite in direction of the other specifications. This could occur due to a reduction in power and precision from a smaller effective sample size within the bandwidth – magnified by a relatively low proportion of compliers in the first stage (see Section 4.1). Nonetheless, this non-parametric specification serves as a supplement to the baseline estimates, and as such the robustness and validity of my main results do not depend solely on Model 5. The other models (1-4) are consistent in showing that retirement has a statistically significant effect on wellbeing across different specifications, and I provide additional robustness checks in the following section.

5 Heterogeneity and potential channels

My results indicate that retirement has a positive effect on social and mental wellbeing. But how exactly does retirement lead to greater wellbeing? To explore potential channels for the effect of retirement and wellbeing and examine heterogeneity, I conduct two additional analyses. First, I compare the main RD treatment effect estimates with estimates obtained from a propensity score matching approach to explore the direction of unobservable factors driving the retirement-wellbeing link. Second, I re-run the baseline specification above on different subsamples conditioned on different characteristics.

5.1 Comparing the RD estimates with matching estimators

Propensity score matching aims to match each treated unit to a non-treated unit with the most similar characteristics, balancing observable covariates between retired and non-retired individuals. Figure 6 provides a list of the observable pre-treatment covariates used for the matching procedure: age, gender, education, marital status, number of children, ethnicity, and immigrant status.

First, I fit a logistic regression to predict the probability of retirement based on these observed covariates, then use these estimated propensity scores to match each treated unit with one control unit using 1-to-1 nearest-neighbour matching with replacement[4]. To assess the effectiveness of the matching procedure, I conduct standard covariate balance checks between treatment and control groups[5]. If the propensity score matching procedure was successful, we would expect to see no significant differences in the mean of observed covariates between the treatment and control groups in the adjusted or 'matched' sample. Figure 6 displays a plot of the standardised mean differences for both the unadjusted and matching-adjusted sample, showing that the matching procedure has strongly balanced covariate values between treated and control units.[6] Subsequently, I estimate the average treatment effects of retirement produced by the matching estimator by comparing the matched treatment and control units.

The results of these matching estimates for each composite wellbeing indicator are summarised in Table 3. The matching estimates of the effect of retirement on wellbeing are in the same direction as the fuzzy RD estimates but the magnitude of the estimates is much smaller than those from the fuzzy RD design. The matching estimators indicate that retirement increases total, mental, and social well-being by about 0.116 points, 0.118 points, and 0.115 points respectively (Table 3) – compared to the baseline RD specifications which show an effect size of more than 0.7 points. Since conditional independence ($(Y_1, Y_0) \perp\!\!\!\perp D_i|X_i$) is unlikely to hold in this context due to the presence of unobserved variables that simultaneously affect both retirement and wellbeing, these matching estimates cannot be interpreted as causal. That is, even after conditioning on observed covariates, the assignment of treatment is not likely to be independent of potential outcomes. Likewise, the RD approach estimates a local LATE for compliers near the threshold, while the propensity score matching procedure estimates an average treatment effect for the treated (ATT) – and as such these estimates are not directly comparable.

[4] Using the 'Matching' package in R by Sekhon, 2011

[5] via 'Cobalt' by Greifer, 2023

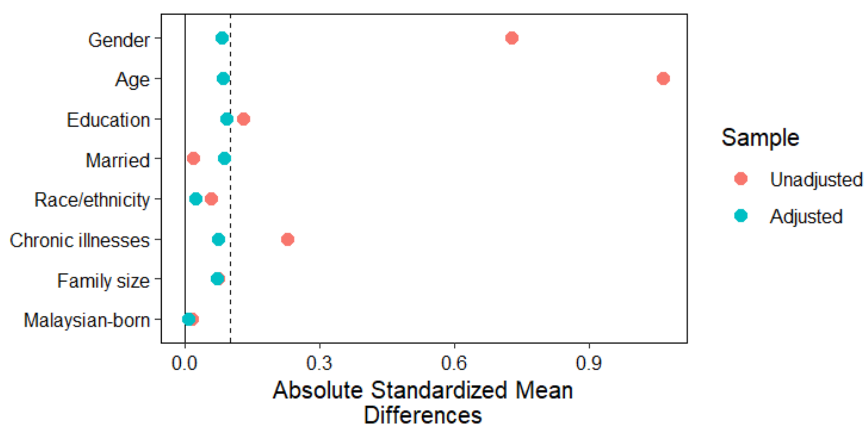
[6] In Appendix B, I provide a table of these covariate values and estimated treatment effects balanced using alternative

Nonetheless, contrasting the matching and RD estimates may still be interesting. As the matching approach only accounts for observable factors, while the RD design effectively accounts for both observable and unobservable factors at the limit of the threshold, comparing the estimated treatment effects from both approaches may help us understand the unobservable mechanisms driving the relationship between retirement and wellbeing. For one, the large difference between the RD and matching estimates suggests that unobservable factors are a significant driver of retirement decisions and wellbeing – and that a naive comparison of observable characteristics would be misleading. Similarly, this gap also suggests that the compliers who are induced to retire by the age-55 threshold (LATE) are likely very different from those who self-select into retirement (ATT) at other ages.

Additionally, the treatment effect estimates from the matching approach are lower than the RD estimates.

matching parameters.

Figure 6: **Matching estimator: covariate balance plots**



Note: Dotted line denotes threshold set at 0.1 absolute standardised mean difference. Propensity score matching, using 1-to-1 matching with replacement. Distance calculated via Mahalanobis distance.

Table 3: **Matching estimates**

	Total wellbeing	Mental wellbeing	Social wellbeing
Retired	0.116***	0.118***	0.115***
	(0.039)	(0.038)	(0.054)
Total observations	4025	4025	4025
Treated observations	875	875	875
Matched observations	875	875	875

Note:* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Abadie-Imbens (AI) standard errors displayed between brackets (Abadie and Imbens, 2004). Propensity score matching using gender, age and various other covariates (see Figure 6 below). 1-to-1 matching with replacement, distance calculated via Mahalanobis distance.

This may suggest two possibilities. The first is that these unobservable characteristics are negatively correlated with retirement decisions and positively correlated with wellbeing. For example, one potential mechanism may be ambition or 'drive'. Individuals who have more intrinsic motivation to work may both be less likely to retire and more likely to experience higher wellbeing. The second possibility is that these unobservable characteristics are positively correlated with retirement and negatively correlated with wellbeing. One potential mechanism is that individuals in high-pressure jobs with high mental and physical demands may both choose to retire earlier and experience lower wellbeing from prolonged exposure to stress. This ties back to the life-cycle labour supply model, where individuals choose consumption and leisure to maximise their lifetime utility subject to constraints. Individuals with intrinsic motivation for work may derive more utility from work and have a lower marginal utility of leisure. On the contrary, individuals with higher-stress work environments may derive lower utility from work (due to mental and physical strain) and choose to retire earlier.

5.2 Subsample analysis

I conduct additional analysis by re-running the baseline specification above on different subsamples conditioned on different variables. I explore three aspects: health, financial security and education. The results of this additional analysis are presented in Table 4.

Health and chronic illness: I construct a variable indicating if an individual reported a diagnosed chronic illness (e.g. cancer, dementia, heart disease, clinical depression), and then re-run the baseline fuzzy RD specification on subsamples of individuals who reported 1 or more chronic illnesses and individuals who did not report any diagnosed illnesses. The results in Table 4 indicate that for individuals with chronic illnesses, the estimated effect of retirement on total wellbeing is statistically insignificant with a negative coefficient. On the contrary, individuals with no reported illnesses experience a statistically significant increase in total wellbeing – a 1.6-point increase, much higher than in the overall sample. This finding suggests may suggest that health plays an important role in mediating the links between retirement and wellbeing. For individuals without chronic illnesses, retirement may offer an opportunity to make use of increased leisure time, while individuals with chronic illnesses may not be able to make full use of the extra time retirement affords. Instead, they may face physical and social limitations in retirement that offset any wellbeing gains from retiring.

Assets: I define a variable indicating if an individual reported total assets valued in excess of RM100,000 (around £18,000, which in Malaysia likely indicates that the individual owns a vehicle or housing asset), and run the baseline model on subsamples of individuals with reported asset values of more than RM100,000 and those that reported assets below RM100,000. As expected, I find that individuals with higher asset values experience an improvement in total wellbeing of about 1.0 points, though this is not statistically significant after the wild bootstrap procedure (Table 4). Meanwhile, individuals with lower asset values display a weaker positive retirement-wellbeing link, experiencing only a 0.7-point increase in total wellbeing. This supports standard labour supply theories of a life-cycle model of consumption, where individuals with more assets may be better able to maintain their level of consumption during retirement. On the contrary, individuals with lower-valued assets may have lower levels of consumption, and as such, experience a lower level of wellbeing.

Years of education: I compare individuals who reported more than 9 years of education (the median years of education in the sample) to individuals with 9 or fewer years of education. The results indicate that for higher-educated individuals, retirement has a statistically significant positive effect on total wellbeing, while lesser-educated individuals report a smaller and statistically insignificant effect. These findings may indicate that education plays a crucial role in determining the relationship between retirement and wellbeing. Higher-educated individuals may be better able to leverage the resources, information, and social networks they need during retirement. Conversely, lower-educated individuals may face larger informational and knowledge gaps related to health and wellbeing-promoting behaviours.

Table 4: Heterogeneity analyses

	Chronic illnesses		Assets		Years of education	
	>1	≤ 1	$>$RM100K	$\leq$RM100k	>9	≤ 9
Retired	-0.131	1.583**	1.003	0.741*	0.904**	0.168
Bootstrap p-value	(0.756)	(0.012)	(0.122)	(0.061)	(0.038)	(0.824)
N	1639	2350	1226	4342	3263	2305

Note:* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Bias-corrected p-values derived via wild bootstrap on standard errors clustered by age between brackets.

6 Robustness of estimates

6.1 Covariate balance tests

As discussed in the empirical strategy section (Section 3), a key identifying assumption for the fuzzy RD design is that treatment and outcomes are continuous around the threshold (Hahn et al., 2001). If this continuity assumption holds, this suggests that individuals just above and just below the threshold have the same potential outcomes (Picchio and van Ours, 2019). As such, I conduct covariate balance tests to show that there are no significant 'jumps' in observable pre-treatment characteristics at the age-55 threshold. Binned scatterplots for education, health status, number of children, illnesses, gender and savings are displayed in

Figures 7 to 12, plotted on the running variable.

Visual inspection suggests that there are no significant jumps of these characteristics at the threshold. Likewise, regressing each covariate separately on the running variable, age dynamics, and the threshold indicator (same specification as the baseline model), only a single covariate out of 10 tested had a discontinuity that was marginally statistically significantly different from zero (significant at 10% level) – quite close to about what could be expected by random chance (reported in Appendix A). Similarly, another common concern surrounding using age thresholds is that there may be other policies that use the same age eligibility rule, for instance, old-age cash benefits that are categorically targeted. However, in Malaysia, centring old-age benefit programmes, such as the Old Age Assistance (*Bantuan Warga Emas*) unconditional transfer, are targeted at age 60 – and as such there are no other policy 'jumps' at the age 55 threshold. Overall, while these analyses do not rule out potential 'jumps' in unobservables at the threshold, the results of these balance tests suggest that the continuity assumption holds. This strengthens the credibility of the fuzzy RD design and supports the assumption that the results estimated in Section 4 are not driven by other differences in observable characteristics around the threshold.

Figure 7: **Education**

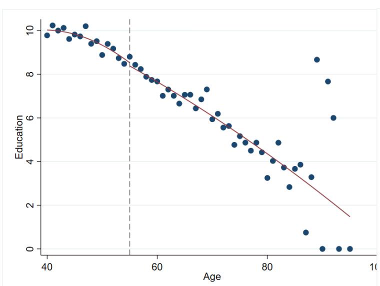


Figure 8: **Health status(1-5)**

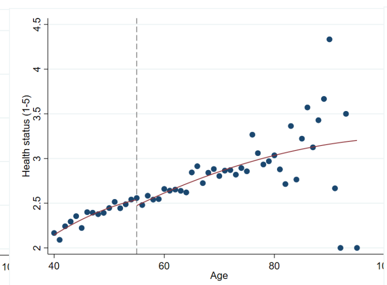


Figure 9: **Number of children**

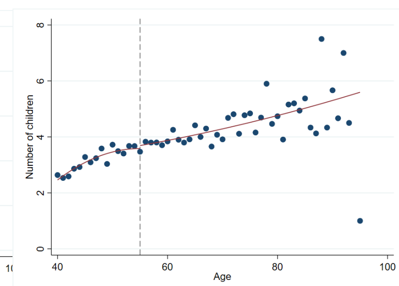


Figure 10: **Reported illnesses**

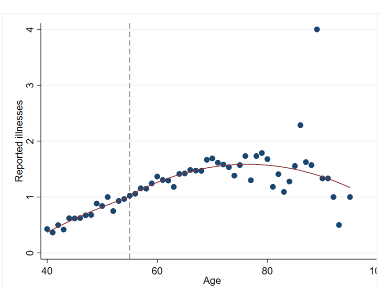


Figure 11: **Gender**

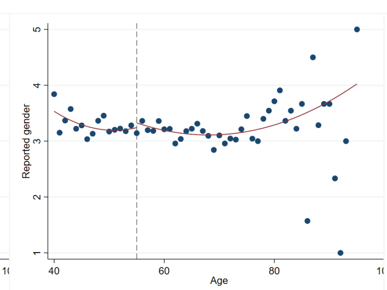
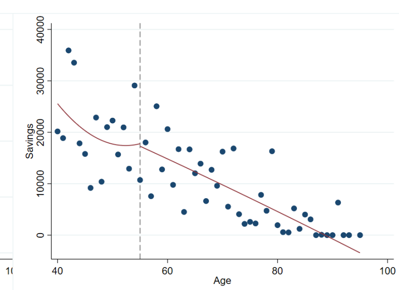


Figure 12: **Wealth**

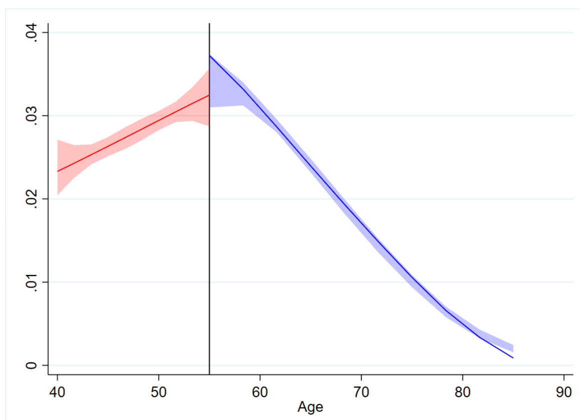


Note: 2nd-order polynomial fitted on each side of the threshold. Full list of covariate balance tests reported in Appendix A

6.2 Density tests for manipulation of the running variable

In line with Cattaneo et al. (2018) and McCrary (2006), I perform a formal test to check if there is evidence of manipulation of the running variable (age) around the threshold (55 years). As discussed in the empirical strategy section (Section 3), one key identifying assumption for the validity of the fuzzy RD design is that there is a positive density of x around the threshold – i.e. that there is no 'bunching' around the cutoff (Cattaneo and Titiunik, 2022). As such, I perform a McCrary test using the 'rddensity' package developed by Cattaneo et al. (2018). The 'rddensity' plot of the local polynomial density estimates of the running variable is shown in Figure 13. A statistical test at the threshold returns the p-value of 0.458, indicating that there is no evidence to support the notion that individuals were able to precisely manipulate their age around the threshold. Of course, this visual examination and test do not rule out two-way manipulation, but given that age is typically hard to manipulate in countries with national identity card systems (as Malaysia has), we can be reasonably confident that this key identifying assumption was not violated.

Figure 13: **Density around age threshold (manipulation testing)**



Note: The solid line is a local polynomial density estimate of the age running variable, following the 'rddensity' package as outlined in Cattaneo et al. (2018), with triangular kernel weights. The shaded areas are confidence bands around the estimate using jackknife estimates. A formal test of positive density around the threshold as in McCrary (2006) returns a p-value of 0.458.

6.3 Placebo and falsification tests

In order to further assess the robustness of the main findings, I conduct a series of falsification tests using placebo eligibility age thresholds. I created three placebo treatments for ages 54, 56, and 60, effectively 'pretending' that treatment occurred at these age thresholds in turn. If the fuzzy RD design specified is valid, we should not expect to find any significant treatment effects at these placebo age thresholds. I re-run the baseline specification from Section 4 (Model 2) with 2nd-order polynomials on either side of the threshold and covariates with each placebo treatment indicator. Table 5 summarises the results of these tests. Each of the estimated coefficients for retirement for age 54, 56, and 60, are highly statistically insignificant. These results further support that the discontinuity in retirement probability at age 55 drives the main results and bolsters the interpretation that the estimates in Section 4 are causal.

Table 5: **Placebo tests**

	(1) Age 54	(2) Age 56	(3) Age 60
Retired	0.288	-0.452	0.535
Bootstrap p-value	(0.596)	(0.251)	(0.177)
Polynomial	2	2	2
Controls	yes	yes	yes
Observations	5568	5568	5568
Joint F-stat	5568	5568	5568

Note: * p<0.10, ** p<0.05, *** p<0.01. Bias-corrected p-values derived via wild bootstrap on standard errors clustered by age between brackets.

7 Discussion and policy implications

Amid rapid shifts in demographics across the globe, in the coming decades, a growing proportion of the world's population will be retired. This, combined with the evolving nature of work, rising healthcare costs, and concerns about public pension systems in many countries mean that understanding the retirement-wellbeing link is now more important than ever before. Yet, the literature surrounding the effect of retirement and wellbeing is relatively mixed – with many studies showing that retirement leads to greater life satisfaction and health (Kesavayuth et al., 2020; Kuhn, 2018; Picchio and van Ours, 2019; Yemiscigil et al., 2021) while others have shown its potential to have negative effects (Atalay et al., 2019; Dave et al., 2006; Sahlgren, 2013).

This essay contributes to this understanding in a developing Asian country context, providing novel evidence from a new dataset from Malaysia. A central challenge in analysing the causal effects of retirement on wellbeing is that an individual's decision to retire is endogenous. To overcome these endogeneity concerns, I exploit institutional rules that allow individuals to access pension funds at age 55. This induces a discontinuity in retirement probability that allows me to specify a fuzzy RD design to estimate the effect of retirement on total, mental, and social wellbeing. The validity of the RD design hinges on numerous key identifying assumptions, which I test in Section 6.

My main findings are fourfold. First, there appears to be a positive link between retirement and wellbeing in Malaysia that is consistent across most of the different model specifications. Second, this positive impact is statistically significant and practically large in magnitude – indicating an increase in total wellbeing roughly equivalent to 1.6 times that experienced by the birth of a first child (Myrskylä and Margolis, 2014, p. 1853). The results suggest that retirement has a stronger impact on improving social wellbeing compared to mental wellbeing. Third, comparing the main RD estimates with those obtained from a propensity score matching approach suggests that unobservables are a large driver of both retirement decisions and wellbeing. Lastly, I find that health status, educational attainment, and size of financial assets significantly affect the retirement experience. Individuals with more reported chronic illnesses, lower educational attainment and smaller pool of financial assets experience smaller or negative wellbeing gains after retirement.

The main findings of this essay should be interpreted with caution. The use of a fuzzy RD design produces a local LATE, a treatment effect specifically pertaining only to compliers close to the threshold, and may not be readily generalisable to other population subgroups. Likewise, as discussed in Section 4.3, the first-stage analysis indicates that the age-55 threshold induces a relatively low percentage of compliers. Additionally, some of the estimated coefficients are sensitive to bandwidth choice, particularly in the non-parametric specification that uses a relatively narrow MSE-optimally-chosen bandwidth.

Nonetheless, these findings have important policy implications for the design of social programmes to safeguard the welfare of retirees in Malaysia and across the developing Asian region. For one, given that individuals who are induced to retire at the threshold experience a positive wellbeing impact and that preferences for leisure and work are heterogeneous, this suggests that a one-size-fits-all approach to retirement may not be optimal for improving the wellbeing of individuals close to retirement age. Rather, because individuals self-select into retirement based on expected wellbeing gains, implementing a flexible statutory retirement age would allow them to choose to retire at an age that maximises their wellbeing. Instead of a fixed mandatory retirement age, this may entail implementing a minimum pension age, after which individuals can choose to remain in full-time work, combine part-time work and collect a portion of their pension, or fully retire and collect the full amount of their pension. These flexible retirement policies and tiered pension disbursement schemes will give individuals greater choice – allowing those who want to retire to be able to do so while still maintaining incentive structures for individuals who wish to maintain some level of attachment to the labour force. Lastly, results from the heterogeneity analysis suggest that it is important for policymakers to consider how retirees with lower education, assets, and/or chronic illnesses experience lower or negative wellbeing changes after retirement. This means that social and health support programmes aimed at supporting the welfare of these vulnerable demographics after retirement are crucial.

* * *

References

Abadie, A., & Imbens, G. (2004). *Large sample properties of matching estimators for average treatment effects*. https://web.stanford.edu/class/polisci353/2004winter/reading/sme_21jan04.pdf

Atalay, K., Barrett, G. F., & Staneva, A. (2019). The effect of retirement on elderly cognitive functioning. *Journal of Health Economics*, *66*, 37–53. https://doi.org/10.1016/j.jhealeco.2019.04.006

Cattaneo, M. D., Jansson, M., & Ma, X. (2018). Manipulation testing based on density discontinuity. *The Stata Journal*, *18*(1), 234–261. https://doi.org/10.1177/1536867X1801800115

Cattaneo, M. D., & Titiunik, R. (2022). Regression discontinuity designs.

Clark, A. E., Diener, E., Georgellis, Y., & Lucas, R. E. (2008). Lags and leads in life satisfaction: A test of the baseline hypothesis. *The Economic Journal*, *118*(529), F222–F243. Retrieved April 16, 2023, from http://www.jstor.org/stable/20108833

Dave, D., Rashad, I., & Spasojevic, J. (2006). The effects of retirement on physical and mental health outcomes. https://www.nber.org/system/files/working_papers/w12123/w12123.pdf

Gelman, A., & Imbens, G. (2014). *Why high-order polynomials should not be used in regression discontinuity designs*. https://www.nber.org/system/files/working_papers/w20405/w20405.pdf

Greifer, N. (2023). *Covariate balance tables and plots: A guide to the 'cobalt' package*. https://ngreifer.github.io/cobalt/articles/cobalt.html

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, *69*(1), 201–209. https://doi.org/https://doi.org/10.1111/1468-0262.00183

Heckman, J. (1974). Life cycle consumption and labor supply: An explanation of the relationship between income and consumption over the life cycle. *The American Economic Review*, *64*(1), 188–194. Retrieved April 22, 2023, from http://www.jstor.org/stable/1814894

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice [The regression discontinuity design: Theory and applications]. *Journal of Econometrics*, *142*(2), 615–635. https://doi.org/https://doi.org/10.1016/j.jeconom.2007.05.001

Increase in epf withdrawal age worth a second look? (2023). https://www.thestar.com.my/news/nation/2023/02/17/increase-in-epf-withdrawal-age-worth-a-second-look

Kesavayuth, D., Rosenman, R. E., & Zikos, V. (2020). The dynamic effects of retirement on well-being. *Applied Economics*, *52*(54), 5921–5935. https://doi.org/10.1080/00036846.2020.1781770

Kuhn, A. (2018). The complex effects of retirement on health. https://doi.org/10.15185/izawol.430

Kumpulan Simpanan Wang Pekerja. (n.d.). *Full withdrawal: Make a smooth transition into retirement.* https://www.kwsp.gov.my/member/withdrawals/full/age-55-60#age55

Lee, D. S., & Card, D. (2008). *Regression discontinuity inference with specification error.* https://www.princeton.edu/~davidlee/wp/RDerror.pdf

Mansor, N., Awang, H., & Rashid, N. F. A. (2019). Malaysia ageing and retirement survey. In D. Gu & M. E. Dupre (Eds.), *Encyclopedia of gerontology and population aging* (pp. 1–5). Springer International Publishing. https://doi.org/10.1007/978-3-319-69892-2_344-1

McCrary, J. (2006). *Manipulation of the running variable in the regression discontinuity design: A density test.* https://eml.berkeley.edu/~jmccrary/mccrary2006_DCdensity.pdf

Michaels, G. (2023). *Topic 1: Labor supply and welfare systems.* https://moodle.lse.ac.uk/pluginfile.php/1218850/mod_resource/content/13/EC423_Michaelmas_Topic1.pdf

Myrskylä, M., & Margolis, R. (2014). Happiness: Before and After the Kids. *Demography*, *51*(5), 1843–1866. https://doi.org/10.1007/s13524-014-0321-x

Picchio, M., & van Ours, J. (2019). The mental health effects of retirement. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3496624

Pischke, J.-S. (2018). Regression discontinuity design.

Sahlgren, G. H. (2013). Work longer, live healthier: The relationship between economic activity, health and government policy. https://iea.org.uk/wp-content/uploads/2016/07/Work\%20Longer,\%20Live_Healthier.pdf

Sekhon, J. S. (2011). *Multivariate and propensity score matching software with automated balance optimization: The matching package for r.* https://www.jstatsoft.org/article/view/v042i07

Stock, J., & Yogo, M. (2005). Testing for weak instruments in linear iv regression. In D. W. Andrews (Ed.), *Identification and inference for econometric models* (pp. 80–108). Cambridge University Press. http://www.economics.harvard.edu/faculty/stock/files/TestingWeakInstr_Stock\%2BYogo.pdf

Titov, N., Dear, B. F., Bisby, M. A., Nielssen, O., Staples, L. G., Kayrouz, R., Cross, S., & Karin, E. (2022). Measures of daily activities associated with mental health (things you do questionnaire): Development of a preliminary psychometric study and replication study. *JMIR Formative Research*, *6*(7). https://doi.org/doi:10.2196/38837

Yemiscigil, A., Powdthavee, N., & Whillans, A. V. (2021). The effects of retirement on sense of purpose in life: Crisis or opportunity? *Psychological Science*, *32*, 1856–1864. https://doi.org/https://doi.org/10.1177/09567976211024248

Özler, B. (2012). *Beware of studies with a small number of clusters.* https://blogs.worldbank.org/impactevaluations/beware-of-studies-with-a-small-number-of-clusters

Appendix

A. RD covariate balance tests

Table 6: **RD: covariate balance**

Education	c101	Children	Illnesses	Gender	Savings	Assets	Married	Religion	Ethnicity
22.57*	1.653	10.33	-2.482	-5.222	1484135.6	120141631.8	-0.633	55.41	39.97
(0.070)	(0.111)	(0.316)	(0.563)	(0.713)	(0.453)	(0.604)	(0.560)	(0.105)	(0.576)

Bias-corrected p-values obtained from wild bootstrap resampling on standard errors clustered on age reported in brackets

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B. Propensity score matching covariate balance tests

Table 7: **Matching: covariate balance**

Covariate	Baseline		(2)		(3)	
	Adjusted diff.	Threshold	Adjusted diff.	Threshold	Adjusted diff.	Threshold
Retired	0.00	<0.1	0.3	>0.1	0.00	<0.1
Gender	0.03	<0.1	0.07	<0.1	0.03	<0.1
Age	0.09	<0.1	0.1	<0.1	0.10	>0.1
Education	0.09	<0.1	0.11	>0.1	0.11	>0.1
Ethnicity	0.01	<0.1	0.04	<0.1	0.10	<0.1
Religion	0.07	<0.1	0.02	<0.1	0.01	<0.1
Health	0.00	<0.1	0.01	<0.1	0.06	<0.1
Number of children	0.06	<0.1	0.05	<0.1	0.03	<0.1
Malaysian-born	0.00	<0.1	0.011	<0.1	0.00	<0.1
1-to-M matching	1		1		5	
Replacement	no		yes		yes	