RIETI Discussion Paper Series 25-E-002

Heterogeneous Effects of Retirement on Health: Evidence from Japan

XIE, Mingjia
Liaoning University

YIN, Ting
RIETI

USUI, Emiko
Hitotsubashi University

ZHANG, Yi
Central University of Finance and Economics

RIETI
Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 25-E-002
January 2025

Heterogeneous Effects of Retirement on Health: Evidence from Japan[*]

Mingjia Xie[†] Ting Yin[‡§] Emiko Usui[**] Yi Zhang[††]

Abstract

We examine the heterogeneous effects of retirement on retirees' health in Japan using the Marginal Treatment Effect framework. Using data from Japanese Study of Aging and Retirement (JSTAR), we find that the effect of retirement on health is trivial and statistically insignificant overall; however, there are heterogeneous effects with respect to the likelihood of being retired. Individuals who are less likely to retire are more prone to be negatively affected by retirement whereas those who are more likely to retire are more inclined to be positively affected. This finding suggests that policies restraining the likelihood of being retired, e.g., increasing the mandatory retirement age, would cause a negative health impact in the population.

Keywords: Retirement; Health; Marginal Treatment Effects; Japan
JEL classification: I12, J26

[*]This study is conducted as a part of the Project "Economic Analysis on the problem of an aging population and a declining birthrate in China and Japan in the COVID-19 pandemic" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). The draft of this paper was presented at the RIETI DP seminar for this paper.
We wish to thank the JSTAR (Japanese Study of Aging and Retirement) team for providing data. JSTAR was conducted by RIETI, Hitotsubashi University, and the University of Tokyo. We wish to thank the support from the Joint Usage and Research Center, Institute of Economic Research, Hitotsubashi University (grant ID: IERPK2312, IERPK2429), and the Program for Innovation Research at Central University of Finance and Economics. We wish to thank Fukao Kyoji, Tomiura Eiichi, Tsuru Kotaro, Sekizawa Yoichi, and the participants at the RIETI Seminar (Oct. 2024), for their helpful comments and suggestions. We also thank the RIETI staff for their kind help and cooperation. Any errors are our own. There are no conflicts of interest.

[†] Liaoning University, Email: xiemingjia@lnu.edu.cn
[‡] Center for Intergenerational Studies, Institute of Economic Research, Hitotsubashi University. Email: yin-ting@ier.hit-u.ac.jp
[§] Research Institute of Economy, Trade and Industry, IAA. Email: yin-ting@rieti.go.jp
[**] Center for Intergenerational Studies, Institute of Economic Research, Hitotsubashi University. Email: usui@ier.hit-u.ac.jp
[††] China Center for Human Capital and Labor Market Research, Central University of Finance and Economics, P.R. China. Email: yi.zhang@cufe.edu.cn

1. Introduction

Retirement is an important transition in life, with substantial consequences on income, consumption, and physical and mental health. For some individuals, retirement eliminates work-related stress and increases leisure time enjoyment with positive effects on their well- being and mental health. For others, retirement is associated with lower income, a loss of daily routines and life purpose, and fewer social contacts. These individuals may perceive retirement as a burden that negatively affects their well-being and mental health. Thus, while there are good reasons to expect an impact of retirement on individual well-being, the direction of this effect is ex-ante unclear and depends on whether positive or negative aspects of retirement dominate (van Ours, 2022).

A large number of empirical studies have found mixed evidence on the magnitude and di- rection of the retirement effect. Some studies have estimated positive effects of retirement on mental health, well-being, and related outcomes (e.g., Charles, 2004; Johnston and Lee, 2009; Eibich, 2015; Kolodziej and Garc´ıa-G´omez, 2019). Other studies found negative effects (e.g. Dave et al., 2008; Rohwedder and Willis, 2010; Bonsang et al., 2012; De Grip et al., 2012; Heller-Sahlgren, 2017; Mazzonna and Peracchi, 2017; Atalay et al., 2019) or no effects (Coe and Zamarro, 2011; Behncke, 2012; Belloni et al., 2016; Fe and Hollingsworth, 2016).[1] To some extent, these inconclusive findings can be attributed to differences in countries, institutional contexts, and chosen identification strategies. However, to better comprehend the impact of health effects of retirement, one needs to investigate the heterogeneity in a systematic framework. While a few studies have directly attempted to address the heterogeneity, the empirical approaches used could not systematically identify and estimate the effect of retirement on mental health.

In this study, we investigate the relationship between the effects of retirement on health and the likelihood of being retired in Japan. There are at least two reasons that we explore the heterogeneity in this dimension. First, it can be of interest for policy makers. The goal of our study is to better understand the health effects for individuals who are likely to retire and how the effect differs from others'. As there are more policies aiming at motivating people to keep working, individuals who are more likely to be affected by such policies may be of interest for policy makers. Understanding health impacts of such polices on these individuals provides more insights for policy makers. Second, the dimension of heterogeneity we investigate can be handled by Marginal Treatment Effect (MTE) framework that systematically describes the distribution of heterogeneous effects rather than focusing on particularly chosen dimensions. The MTE framework is introduced by Bj¨orklund and Moffitt (1987) and generalized by (Heckman and Vytlacil, 2005, 2001, 1999), which relates the treatment effect (effect on health) to the observed and unobserved characteristics that affect the likelihood of begin retired.

We find substantial heterogeneity in the effect of retirement on health with respect to both observed and

[1] For an excellent literature overview on mental health and retirement, see Picchio and van Ours (2019) and van Ours (2022).

unobserved characteristics determining retirement. For observed characteristics, take gender as an example. Women are more likely to retire and they suffer less from being retired than men, which points to a selection on gains: individuals who are more likely to retire actually suffer less or even benefit from retirement. The selection on unobserved characteristics reinforces the finding of selection on gains. Individuals with unobserved characteristics that hinder their retirement ("high unobserved cost individuals") suffer most from retirement, whereas individuals who are more likely to retire ("low unobserved cost individuals") suffer the least.

This paper contributes to the growing literature that estimates marginal treatment effects in the context of retirement. The finding of this study shows that individuals select themselves into retirement based on the effects of retirement on health. Such selection on gains have been found by the current literature (Carneiro et al., 2011; Heckman et al., 2006; Nybom, 2017; Heckman et al., 2018, e.g.,). To the best of our knowledge, this paper provides the first evidence of the selection on gains in evaluating the effect of retirement on health.[2] This paper also provides new evidence of heterogeneous effects of retirement in Japan by relating the effect to the likelihood of being retired, which is different to the existing literature mainly focusing on heterogeneity with respect to some observables.

The remainder of the paper is organized as follows. Section 2 summarizes the identification strategy of marginal treatment effect. Section 3 describes the data. Section 4 describes the estimation results. Section 5 concludes.

2. Estimating Marginal Treatment Effects

We will introduce the model setup used in the framework of Marginal Treatment Effect and the definition of treatment effect of interest. Meanwhile, we will discuss the required assumptions for identification and briefly show the estimation procedures.

2.1 Baseline Model Setup

Let Y_1 be the potential outcome in treated state ($D=1$) and Y_0 be the potential outcome in untreated state ($D=0$). The observed outcome (Y) is the realization of one potential outcome:

$$Y = (1-D)Y_0 + DY_1 \tag{1}$$

The potential outcomes are specified as:

$$Y_j = \mu_j(X) + U_j, j \in \{0,1\} \tag{2}$$

where μ_j is a state-specific function of the observable X, and U_j is the unobservable which is normalized

[2] Though uncommon in literature, the reverse selection is found in the migration literature, which find more skilled workers are easier to migrate (Chiquiar and Hanson, 2005; Rooth and Saarela, 2007; McKenzie and Rapoport, 2010).

to $E[U_j|X] = 0$. Equation 2 indicates that the heterogeneity in the treatment effect $Y_1 - Y_0 = \mu_1(X) - \mu_0(X) + U_1 - U_0$ results from both the observed characteristics X and the unobserved characteristics. This specification defines a more flexible heterogeneity than the commonly used specification in which the treatment D is separately additive to all X (homogeneous treatment effect) and the specification in which the interaction terms between D and X are allowed (heterogeneous treatment effect with respect to only the observable). For selection to treatment (defined in this study as being retired), the following latent index model is used:

$$I_D = \mu_D(Z) - U_D \tag{3}$$

$$D = 1\{I_D > 0\} \tag{4}$$

Where μ_D is a function of $Z \equiv \{X, Z_0\}$, and Z_0 is the instrument(s) for D. μ_D represents the gross benefit of receiving treatment, and U_D represents the cost of treatment. In this study, U_D captures not only some unobserved individual characteristics but also some unobserved family background factors that affect retirement decisions. The latter could be even more important because the decision on retirement is heavily affected by various unobserved factors in the family.

In the MTE literature, the distribution of U_D is often normalized to uniform distribution on a unit interval. As a consequence, function $\mu_D(Z)$ can be interpreted as the propensity score (the probability of receiving treatment conditional on the unobservable Z), $P(Z) \equiv Pr\,(D = 1 \,\big|\, Z) = Pr\,(\mu_D(Z) > U_D) = \mu_D(Z)$, where the last equality holds when $U_D \sim U(0,1)$. Henceforth, the selection equation for treatment is re-defined as

$$D = 1\{P(Z) > U_D\} \tag{5}$$

MTE as a function of X and U_D accesses the heterogeneous treatment effect as follows:

$$\begin{aligned} MTE(x,u) &= E(Y_1 - Y_0|X = x, U_D = u) \\ &= \mu_1(x) - \mu_0(x) + E(U_1 - U_0|X = x, U_D = u) \end{aligned} \tag{6}$$

MTE is the average treatment effect for the individual with observed characteristics $X = x$ and unobserved cost to treatment $U_D = u$ (or the u_{th} quantile of U_D).[3] MTE also allows for the heterogeneity in both the observable X and the unobservable cost to receive treatment u. In this study, the MTE summarizes the heterogeneous effects of retirement with respect to the observable (e.g., gender) and the unobservable cost to retire (e.g., preference towards working). Consequently, we can directly examine the heterogeneous effects with respect to the likelihood of being retired that is described by X and U_D, which is the treatment effect of interest in this study. Moreover, compared to the average treatment effect for the whole population, MTE focuses on a more granular subpopulation and thus can be used to construct some other treatment effects of

[3] MTE is defined on Marginal individuals in receiving treatment because individuals with $U_D = u$ are also ones with $\{P(Z) = u\} \cap \{I_D = 0\}$ (indifferent in receiving treatment with propensity score u).

interest. For example, with a binary instrument Z_0 that shifts the propensity score from $P_0(x) \equiv Pr\,(D = 1|X = x, Z_0 = 0)$ to $P_1(x) \equiv Pr(D = 1|X = x, Z_0 = 1)$, Local Average Treatment Effect (LATE) based on Wald estimator is the average of MTEs for a subgroup of individuals:

$$\begin{aligned} LATE(x) &= \frac{E(Y|Z_0 = 1, X = x) - E(Y|Z_0 = 0, X = x)}{E(D|Z_0 = 1, X = x) - E(D|Z_0 = 0, X = x)} \\ &= \frac{1}{p_1(x) - p_0(x)} \int_{p_0(x)}^{p_1(x)} MTE(x,p)dp \end{aligned} \tag{7}$$

2.2 Identification

One way of identifying MTE is using the method of local IV developed by Heckman (1999; 2001; 2005). This method identifies MTE as the derivative of the conditional expectation of Y with respect to the propensity score. More precisely, we have

$$\begin{aligned} E(Y|X = x, P(Z) = p) &= \mu_0(x) + p\big(\mu_1(x) - \mu_0(x)\big) \\ &\quad + pE(U_1 - U_0|X = x, U_D \le p) \\ &= \mu_0(x) + p\big(\mu_1(x) - \mu_0(x)\big) + K(x,p) \end{aligned} \tag{8}$$

where $K(x,p) \equiv pE(U_1 - U_0\,|X = x, U_D \le p)$. $K(x,p)$ is a function of X and p that captures heterogeneity along the unobserved cost to treatment U_D. Taking the derivative of Equation 8 with respect to p and evaluating it at u, we get MTE

$$\begin{aligned} MTE(X = x, U_D = u) &= \frac{\partial E(Y|X = x, P(Z) = p)}{\partial p}\Big|_{p=u} \\ &= \mu_1(x) - \mu_0(x) + k(x,u) \end{aligned} \tag{9}$$

where $k(x,u) = E(U_1 - U_0\,|X = x, U_D = u)$. Intuitively, conditioning on $X = x$, when an infinitesimal shift occurs in the propensity score at p (changing the treatment status from untreated state to treated state), the corresponding change in Y is the treatment effect for individuals who have $X = x$ and have p as the propensity score (or unobserved cost), which is exactly MTE. Equation 9 also indicates that, without further assumptions, we need additional variation conditional on X to identify $\mu_1(x) - \mu_0(x)$ and $k(x,u)$ separately to identify MTE. This additional variation comes from the excluded instrument Z_0, and $MTE(x,p)$ is identified under the following assumption on the instrument.

Assumption 1

(U_0, U_1, U_D) is independent of Z_0, conditional on X

The conditional independence assumption requires that the instrument is independent of the unobservable in the outcome equations and the selection equation. The conditional independence between Z and (U_0, U_1, U_D) implies and is also implied by the standard IV assumptions of conditional independence and monotonicity (Vytlacil, 2002).

Besides the assumptions that are required in the literature using Instrumental Variable (IV), there are often more assumptions in estimating MTE. The local IV estimator motivated by Equation 9 indicates that the support of the propensity score P conditional on X determines the support of the unobserved cost U_D in MTE. Therefore, substantial variation in P conditional on X (which solely comes from the excluded instrument Z_0) is needed to identify $MTE(x,u)$ on a wide range of $U_D \in [0,1]$. For this reason, additional assumptions are usually required, e.g., at least one of the instruments is continuous, which makes it possible to have full support in MTE. However, it can be challenging to find proper continuous instrument(s) with sufficient variation conditional on observed covariates in many em piri- cal studies, including this study. In the case of discrete instrumental variables, alternative approaches include restricting the specifications in the model and specifying a less flexible relation among random variables.[4] Following Brinch et al. (2017), we impose the second assumption as follows:

Assumption 2

$$\mathrm{E}\left(Y_j \middle| U_D, X = x\right) = \mu_j(x) + E\left(U_j \middle| U_D\right), \qquad j \in \{0,1\}$$

Assumption 2 specifies a more restrictive version of Equation 2 because it implies that the observable and the unobservable contribute to the potential outcome in a substitute manner. consequently, MTE in Equation 6 can be written as

$$MTE(x,u) = \mu_1(x) - \mu_0(x) + E(U_1 - U_0 | U_D = u) \tag{10}$$

Equation 10 implies that $MTE(x,u)$ can be identified over the support of u, which is determined by the support of the estimated propensity score P, unconditional on X. Therefore, Assumption 2 makes the discrete instrumental variable feasible in identifying MTE.

After imposing Assumption 2, the treatment effect is still allowed to vary by X and U_D but not by the interaction between the two, and it is weaker than the additive separability assumption between D and X, which is commonly used in empirical analysis such as a linear specification $Y = \alpha D + \beta X + U$. Furthermore, Assumption 2 is implied by (but does not imply) the full independence assumption about random variables, i.e., $(Z, X \perp U_0, U_1, U_D)$ which is assumed in some applied works estimating MTE. Assumption 2 holds when there is no endogenous variable in X in the outcome (health) equation, which is also required in many applied works like the standard IV estimation approach.

Under Assumption 1 and 2, we have:

$$E(Y|X = x, P(Z) = p) = \mu_0(x) + p\left(\mu_1(x) - \mu_0(x)\right) + K(p) \tag{11}$$

and thus

[4] See a more detailed discussion in Brinch et al. (2017).

$$MTE(x,p) = \frac{\partial E(Y|P(Z) = p, X = x)}{\partial p} = \mu_1(x) - \mu_0(x) + k(p) \quad (12)$$

where $K(p) = pE(U_1 - U_0 \,|U_D \leq p)$ and $k(p) = E(U_1 - U_0 \,|U_D = p)$.

2.3 Estimation procedures

For ease of interpretation, we illustrate the idea of estimation procedure with a parametric approach. However, to make our estimates as flexible as possible, we adopt a semi-parametric approach in our estimates. Equation 12 suggests the following estimation procedures: We start by estimating the propensity score $\hat{P}(Z)$ based on Equation 4 using a probability model such as probit or logit model. We then make assumptions about the functional form of the unknown function μ_1, μ_0 and $K(p)$. With these assumed functional forms, we estimate $\widehat{\mu_0}$, $\widehat{\mu_1 - \mu_0}$, and $\widehat{K(p)}$ separately based on the equation $E(Y|X = x, P = p)$ in Equation 12. Last, we calculate MTE by taking the derivative with respect to p.

In the main specification, the propensity score P is estimated from the logistic regression. Both μ_0 and μ_1 are specified to be linear: $\mu_0(x) = \beta_0 x$ and $\mu_1(x) = \beta_1 x$. Thereby, the conditional expectation of Y is written as:

$$E(Y|X = x, P(Z) = p) = x\beta_0 + x(\beta_1 - \beta_0)p + K(p) \quad (13)$$

Furthermore, $K(p)$ is specified as a polynomial function of p with order 2 in the main specification. Note that MTE is then a linear formula in p as follows:

$$MTE(x,u) = x(\beta_1 - \beta_0) + \gamma u \quad (14)$$

$(\beta_1 - \beta_0)$ captures the heterogeneous treatment effects with respect to the observable char- acteristics X, while γ corresponds to the heterogeneous treatment effects with respect to the unobserved cost to treatment. A negative γ indicates that the treatment effect is larger for those who are more likely to be selected to treatment because of lower unobserved cost to treatment, which is in line with the prediction of the Roy Model, namely selection on gains. On the contrary, a positive γ indicates the reverse selection on gains, i.e., individuals who are less likely to receive the treatment due to the higher unobserved cost are with larger treatment effects.

To make our estimates as flexible as possible, we adopt a semi-parametric estimation approach. We first obtain the estimated $\hat{p}$ from a logistic regression. We then use local polynomial (second order) regressions of Y, X, and $X \times \hat{p}$ on $\hat{p}$ to get residuals e_Y, e_X, and $e_{X\times P}$. With these residuals, we estimate the following equation using regression and

$$e_Y = e_X\beta_0 + e_{X\times p}(\beta_1 - \beta_0) + \epsilon \quad (15)$$

Construct residual $\tilde{Y} = Y - X\widehat{\beta_0} - X(\widehat{\beta_1 - \beta_0})\hat{p}$ where $\widehat{\beta_0}$ and $(\widehat{\beta_1 - \beta_0})$ are estimated coefficients from above. Furthermore, we use the local polynomial (second order) regression of $\tilde{Y}$ on $\hat{p}$, saving level $\widehat{K(p)}$ and slope $\widehat{k(p)} = \widehat{K'(p)}$. Finally, we have $MT\widehat{E(x,u)} = x(\widehat{\beta_1 - \beta_0}) + \widehat{k(p)}$. In the nonparametric regression above, the bandwidths are chosen by rule-of-thumb using polynomial of order 4, and Gaussian kernels are used.

3. Data and variables

The data comes from Japanese Study of Aging and Retirement (JSTAR), which is a biannual panel survey starting in 2007. The survey collects the information of respondents who are aged 50 or above about their basic demographics, employment status, and health outcomes. A national representative sample of households from five cities participated in the first wave of the survey in 2007, and the number of participants increased to more households from 10 cities in 2013. This study focuses on individual-year observations with valid information on all variables used in the analysis.

3.1 Outcome: health status

The outcome variables measure the health status of the respondents. We use two measures in this study: self-rated health and health conditions. Self-rated health is reported by respondents about the feeling of their health. This is a categorical indicator with 5 levels: Good, Fairly good, Average, Not very good, and Not good. The second measure asks whether there is any health condition that interferes the respondents' casual life. This is also a categorical indicator with 4 levels: Has significantly interfered, Has interfered, Has not interfered, Has not interfered at all. To make the outcome variables consistent and interpretable, instead of directly using the two rough measures, we construct two binary variables that equals one if the respondent is in "good" health and zero otherwise. Specifically, as for self-rated health, the new measure is one if the reported health is at least "Fairly good", as for health conditions, the new measure is one if the health condition "Has not interfered" or "Has not interfered at all" their casual life.

3.2 Retirement status and its instrument

To measure a respondent's retirement status, we utilize a question in JSTAR asking respondents which situation best describes their current work status: Currently working, Temporarily not working, Not working, or other. We compute a binary measure which takes the value 1 if a respondent considers his/herself as not working, and 0 otherwise.[5]

A key problem in the retirement literature is that retirement decisions are endogenous (see e.g. Mazzonna and Peracchi, 2012; Insler, 2014; Mazzonna and Peracchi, 2017). Endogeneity can arise from reverse causality or from unobserved confounders, such as cognitive functioning or health limitations. A common way of dealing

[5] There are several definitions of retirement status. Insler (2014) discusses two common definitions of being retired: self-reported retirement status, or not being in paid labor. Both have been used in the literature.

with this is to use statutory retirement ages in each country's social security scheme as instrumental variable (e.g. Rohwedder and Willis, 2010; Coe and Zamarro, 2011; Mazzonna and Peracchi, 2012, 2017). In our application, we construct an instrument measuring the years to the statutory retirement age in Japan (60 years old). It is defined as the years to the cutoff (60 years old) when the respondent's age is over 60 years old and zero when the respondent's age is 60 or younger. We winsorize the value of the instrument at 60 rather than solely using the years to retirement age 60 when age is 60 or younger because we need to control for age and year fixed effect that can be collinear with our instrument.

Our instrument exploits the exogeneous variation from the country-level retirement system. As discussed in Gruber and Wise (2009), retirement behavior responds very strongly to incentives set by social security pension systems. Since such policies are defined on the national level and are outside of individual control, they provide credible exogenous variation to individual retirement decisions. It is therefore unlikely that mental health shows discontinuities around retirement eligibility ages that can be attributed to reasons other than retirement.[6]

3.3 Control variables

In the baseline regression, we control for age, age quadratic, survey year fixed effects, educational level, and marital status. Education level is a categorical variable: 1 "Elementary/middle school" 2 "High school (including old-system middle school, girl's school, trade school, normal school" 3 "Junior college (including technical high school, etc.)" 4 "Vocational school" 5 "University (including old-system high school, old-system technical college)" 6 "Graduate school (Master's)" 7 "Graduate school (Ph.D.)" 8 "Other".

3.4 Summary

Table 1 shows the summary statistics of all variables. Slightly less than half of the respondents report that they have at least Average health status, and more than 80% of the respondents report that their health conditions do not interfere their causal life. Around 48% of the respondents are retired, and the average years to the statutory retirement age (60) is 6.4 years. For all respondents with slightly more women, the average age is 65, and most of them have got married. Table 1 also shows the summary statistics by retirement status. There are mainly two noticeable differences. First, the retired respondents have relatively worse health status than the working population. Second, the ratio of women and the high-educated in the retired population is higher than the working population.

4. Results

[6] One concern with the instrument could be that health insurance benefits are correlated with retirement schemes. Since Japanese health insurance benefits are not contingent on age, this is not an issue here.

4.1 First stage estimation

Table 2 shows the first stage estimation results, i.e., how the instrument (years to statutory retirement age) affects retirement status. Column (1) indicates that, using a linear regression model, with one more year above the statutory retirement age, the likelihood to be retired significantly increased by 0.043. There is also little concern on the weak instrument as the Kleibergen-Paap rk LM statistic is 61.51, well above the rule-of-thumb cutoff 10. Since the first stage of the estimation procedure of MTE is based on a logistic regression, we also report the estimated parameters of the first-stage logit model in column (2). To ease the interpretation, we report the marginal effects of the instruments while fixing all other covariates at sample means. The finding is consistent with the results from the linear model that, one more year above the statutory retirement age, the likelihood of being retired increased by 3.3% conditional on all other covariates at their sample means.

The first-stage estimation generates sizable common support for the propensity score $P(Z)$ as shown in Figure 1. The estimated propensity score in the common support, namely the overlapped set of $P(Z)$ between treated and untreated, ranges from 0.13 to 0.87.[7] Without additional parametric assumptions on curvature, MTE can only be identified up to the range of common support of $P(Z)$. For the range outside the common support or the common support with a few observations, the identification of MTE is totally or heavily determined by the parametric assumption. Therefore, to further ease the concern that the identification heavily rests on the arbitrary parametric specification, we trim the points of support with the 1% lowest densities and construct the common support as the points of overlapping support between the treated and untreated. As a result, the common support after trimming ranges from 0.13 to 0.85.[8]

4.2Treatment Effect Heterogeneity in observed and unobserved characteristics

As explained in Section 2, the MTE investigates the heterogeneous effects in both the observed and the unobserved dimensions. In the semi-parametric estimation implemented in the analysis, the estimated heterogeneous effects $k(p)$ defined in Equation 12 is estimated non-parametrically, thus we use figures to present the estimation result as shown in Figure 2 and Figure 3. On the X-axis, it is the unobserved resistance to treatment U_D; on the Y-axis, it is the health effect of retirement. As shown in both figures, the effect is statistically negatively associated with the unobserved resistance (The p-value of the test for the negative slope is smaller than 0.01): with relatively low resistance, the treatment can be positive; whereas with relatively large resistance, the treatment is negative. Therefore, the health effect of retirement for individuals who are

---

[7] Although sizable support is found, the dispersion of the propensity score within the treated or the untreated indicates some constraints on who gets into retirement status. In other words, the included observed variables do not capture all determinants of retirement; Meanwhile, the unobserved characteristics captured by UD explain the remaining. UD is the unobserved cost or constraint of begin retired by definition. For example, the family background, which is not included in the covariates but heavily affects the retirement decision, can explain the dispersion of the propensity score. The dispersion also highlights the importance of the unobservable affecting retirement status, which is addressed by the MTE.

[8] A relatively small fraction of observations (166 out of 16,928) are dropped due to the trimming. After removing these these observations, we fit the baseline propensity score model on the trimmed sample again. A similar trimming strategy is used by Nybom (2017).

very likely to retire for various reasons that are unobserved to this study is relatively trivial and even positive; whereas individuals who are very unlikely to be retired due to these unobserved reasons suffer from retirement in terms of their health. For a more concrete understanding of the result, we use an example to illustrate the findings. Suppose that preference towards work is a key component when making retirement decision, and such preference is not observed by us (not included in control variables). Our finding suggests that individuals who have little interest in working (likely to be retired) suffer less negative health impact of retirement than individuals who have strong preference towards working. In summary, when individuals are less likely to retire due to various unobserved reasons, retirement can be more detrimental to their health.

The estimated MTE also indicates the heterogeneous effects of retirement on health due to observed factors, as summarized in Table 3 and Table 4. We can find some heterogeneity with respect to some observed dimensions considered in this study. For example, the effect of retirement on health for women is higher than men, regardless of the outcome measure. Given most of the estimates are insignificant, we therefore do not find evidence of heterogeneous treatments in one single dimension that is included in the control variable in this study. However, it does not necessarily mean that there is no heterogeneity in all observed dimensions. To answer this question, we investigate the relationship between the effect and the likelihood of being retired explained by the observable. Specifically, following Zhou and Xie (2019), we summarize the likelihood explained by the observable by propensity score $P(D = 1|X)$, which is the prediction of the retirement status based on the observed characteristics.[9] Then, the correlation between the propensity score and the effect contributed by the observable, i.e., $(\beta_1 - \beta_0)X$, shows the relation of interest. For example, a positive correlation between $(\beta_1 - \beta_0)X$ and the propensity score indicates a positive correlation between the MTE and the propensity score, which means that the individual who is less likely to retire explained by the observable is with smaller health effects from retirement. Table 5 confirms such positive correlations for both outcomes. In other words, we also find that individuals who are less likely to retire benefit less or even negatively from retirement, which is consistent with the finding from the analysis based on the unobserved characteristics.

The estimate based on Table 3 and Table 4 is comparable to the results based on the classical methods such as Two-Stage-Least-Square (2SLS). MTE can be used to construct LATE, which can also be estimated by a standard IV estimation approach. Table 3 and Table 4 summarize the LATE estimated by two different approaches. The point estimate of the LATE estimator constructed based on MTE is comparable to the one obtained by 2SLS. Therefore, for compliers whose retirement status is in line with whether the age is above the statutory retirement age, the effect of retirement on health is trivial and insignificant.

[9] The key idea of the refined MTE introduced by Zhou and Xie (2019) is that the latent index structure in the choice-making equation implies that all the treatment effect heterogeneity occurs along only two dimensions: (1) the propensity score $P(D = 1|X)$ and (2) the unobserved cost to treatment U_D. They also prove that we can replace the multi-dimension observed characteristics with the propensity score without loss of generosity. See more in their paper.

4.3 Robustness check

One feature of the pension system in Japan is the gap between the statutory retirement age (60) and the age after which retirees can fully receive their pension (65). Workers are officially entitled as retirees after the statutory retirement age (60), and they start to receive some proportion of the pension depending on their different conditions before being entitled to the full reception of their pensions at age (65). Therefore, the retirement decisions may response to this financial incentive so that, conditional on being above the statutory retirement age, people choose to retire even later to reduce the gap between the age of retirement and age of pension reception.10 To account for the additional incentive due to the age entitled for pension, we introduce a second instrument that equals to one if age is above 65 and zero otherwise. Table A.1 shows the first stage estimation results with two instruments. The probability of retirement is higher after being above 65 conditional of being eligible to retirement (above 60 years old), which is in line with our expectation, even though it is not statistically significant. With these two instrument, we estimate the marginal treatment effects of interest as shown by Figure A.1 and A.2. The two figures are consistent with our baseline results that individuals select themselves into retirement based on the health effect, i.e., individuals who are more likely to retire are those for whom the retirement is less detrimental or even beneficial to them.

5. Conclusion and policy implementation

This paper assesses the heterogeneity in the effects of retirement on health by estimating marginal treatment effects. The effects are heterogeneous with respect to the observed characteristics. When summarizing the likelihood of being retired by the propensity score explained by the observable, we find individuals who have a lower probability of begin retired suffer more from retirement in terms of health effects. The heterogeneity with respect to the unobserved characteristics reinforces the finding: individuals with higher unobserved cost to retire, thus less likely to be retired, suffer less or even benefit from retirement in terms of health effects. Overall, we find a positive relation between the health effect and the likelihood of being retired, namely selection on gains.

Overall, the results suggest heterogeneous treatment effects of retirement on health in Japan. The selection on gains also suggests that the health effects are consistent with the propensity of being retired. In the perspective of health impact, those who are more likely to retire should retire; whereas those who are less likely to retire should not retire. Therefore, this paper posts concerns over polices that disproportionately change the motivations in retirement decisions. For instance, when a policy is aiming at motivating people especially those who can retire easily to keep working, the probability of being retired becomes lower in the population. In such cases, the overall health effect may be negative because the individuals who are less likely to retire are with more negative impacts. Alternatively, we use an example to illustrate the policy implication further. Suppose that there is a change in retirement policy that increases the retirement age from 60 to 65. For those who are aged between 60 and 65, they are affected most as they are not eligible to normal retirement. Their probability of retirement in our sample is 0.36, which is around the 40% percentile in the population when ordering from the lowest to the highest. Suppose they are also the sub-population with 40%th likelihood to retire, the health effect of such policy change would be rather limited as the health effect of retirement for these sub-population is close to zero as shown in our baseline estimation results (Figure 2 and 3). However, we should emphasize that it is not an accurate calculation and will require a more nuanced analysis in the future. Further studies on who is less/more likely to retire may provide more detailed insights on evaluating the effect of such a policy.

Table 1 Summary Statistics

Variable	All	Retired	
		No	Yes
Outcome variable			
Good self-rated health	0.467	0.532	0.399
	(0.499)	(0.499)	(0.490)
No health condition	0.821	0.879	0.760
	(0.384)	(0.326)	(0.427)
Treatment variable			
Retired	0.480	-	-
	(0.500)	-	-
Instrumental variable			
Years to retirement age 60	6.420	3.773	9.134
	(5.999)	(4.862)	(5.738)
Covariates			
Age	65.349	62.041	68.743
	(7.393)	(6.616)	(6.484)
Male	0.485	0.586	0.379
	(0.500)	(0.492)	(0.485)
Married	0.794	0.816	0.771
	(0.404)	(0.387)	(0.419)
At least junior college degree	0.304	0.352	0.23
	(0.46)	(0.477)	(0.421)
Number of observations	16,928	8,800	8,128

Note: Sample average is in number, and the standard deviation is in parenthesis.

Table 2 First stage estimation: 2SLS and Marginal effects at means from the logistic regression

Independent variables	Retired	
Years to retirement age 60	0.043***	0.033***
	(0.005)	(0.008)
Kp Wald F statistics	61.51	-
χ^2 for test of the excluded instruments	-	17.62
Observations	16,694	16,694

Robust standard errors in parentheses are clustered at individual level. All regressions include covariates: age, age squared, gender, marital status, education level, survey year fixed effects.

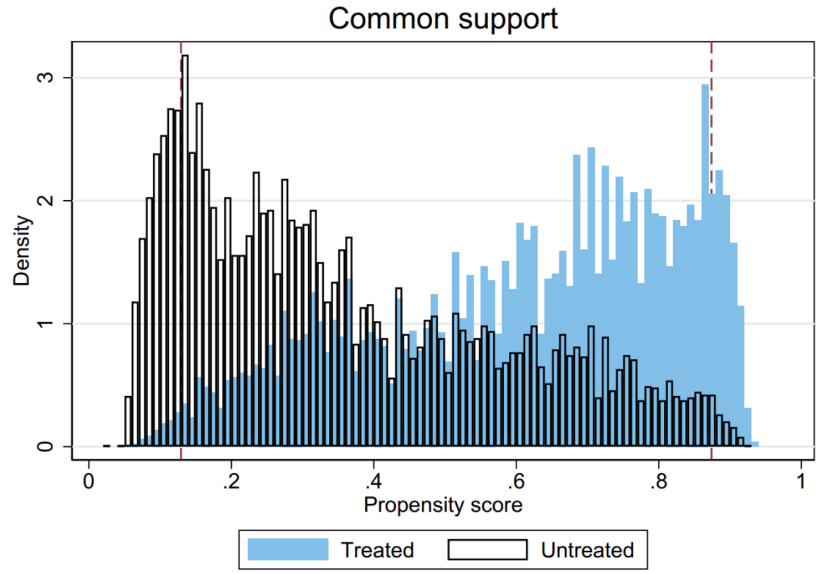

Figure 1 Estimated propensity score

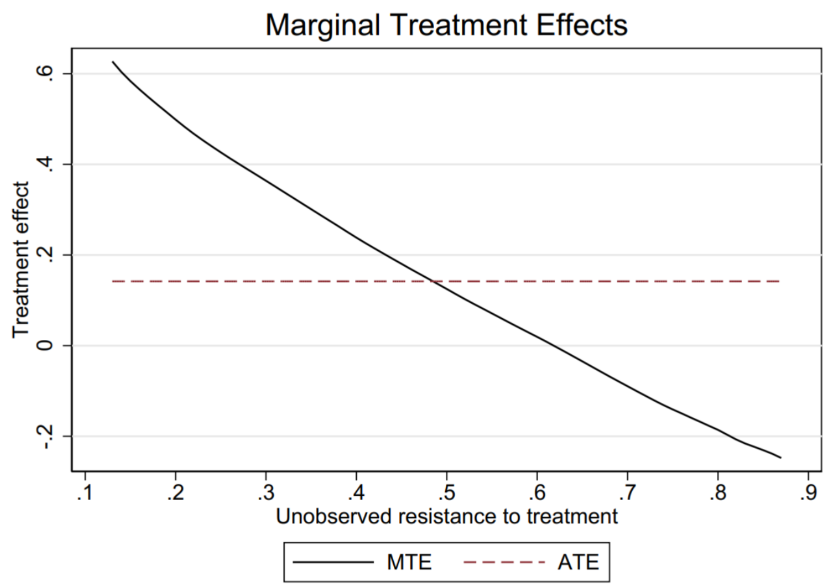

Figure 2 Estimation results of MTE for self-rated health

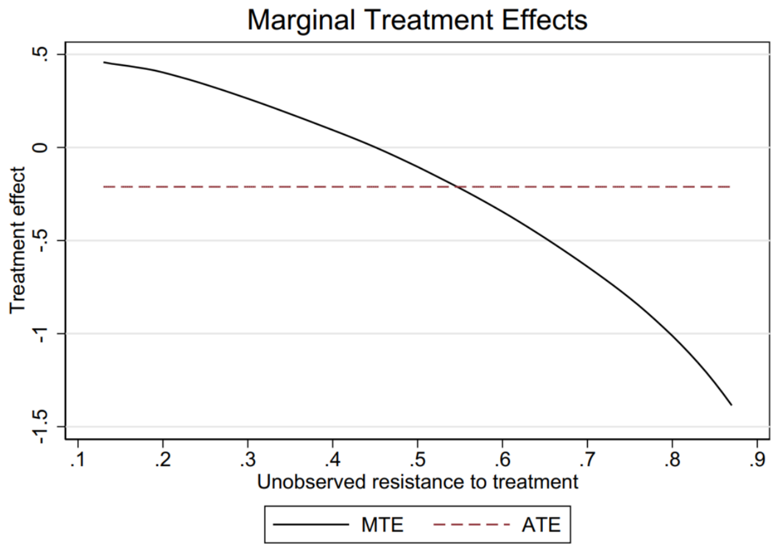

Figure 3 Estimation results of MTE for health conditions

Table 3 Estimation results for self-rated health

$\boldsymbol{MTE = (\beta_1 - \beta_0)x + k(u)}$	Coef.	Std. Err.	P-value
$\boldsymbol{\beta_1 - \beta_0}$			
Age	-0.025	0.052	0.627
Age squared	0.0002	0.0005	0.596
Male	-0.073	0.759	0.135
Married	-0.015	-0.1	0.923
Education (Categorical value)			
- Level 2	.080	0.048	0.094
- Level 3	.007	0.08	0.929
- Level 4	.051	0.055	0.357
- Level 5	.032	0.052	0.535
- Level 6	.059	0.159	0.707
- Level 7	-.135	0.325	0.677
- Level 8	.218	0.222	0.326
- Level 9	-.165	2.926	0.955
- Level 10	-.143	0.56	0.798
$\boldsymbol{k(u)}$ (See Figure 2)			
LATE (based on MTE)	-0.378	0.351	0.282
LATE (based on 2SLS)	-0.634	0.133	0.636
Test of observable heterogeneity			0.000

The estimation includes age, age square, gender, marital status, education level, and survey year fixed effects. Bootstrap standard error is reported in parenthesis. The null hypothesis of the test of observable heterogeneity is that $\beta_1 - \beta_0 = 0$ are jointly true for all observable.

Table 4 Estimation results for health conditions

$\boldsymbol{MTE = (\beta_1 - \beta_0)x + k(u)}$	Coef.	Std. Err.	P-value
$\boldsymbol{\beta_1 - \beta_0}$			
Age	-0.025	-0.028	0.034
Age squared	0.0002	0.0004	0.0003
Male	-0.073	-0.255	0.171
Married	-0.015	-0.015	-0.1
Education (Categorical value)			
- Level 2	.080	0.069	0.065
- Level 3	.007	-0.036	0.11
- Level 4	.051	0.101	0.063
- Level 5	.032	-0.045	0.064
- Level 6	.059	0.909	0.181
- Level 7	-.135	-1.11	0.476
- Level 8	.218	0.175	0.325
- Level 9	-.165	0.185	1.24
- Level 10	-.143	0.255	0.606
$\boldsymbol{k(u)}$ (See Figure 3)			
LATE (based on MTE)	-0.023	0.465	0.96
LATE (based on 2SLS)	-0.007	0.103	0.945
Test of observable heterogeneity			0.000

The estimation includes age, age square, gender, marital status, education level, and survey year fixed effects. Bootstrap standard error is reported in parenthesis. The null hypothesis of the test of observable heterogeneity is that $\beta_1 - \beta_0 = 0$ are jointly true for all observable.

Table 5 The correlation between $(\beta_1 - \beta_0)X$ and the propensity score from OLS

$(\beta_1 - \beta_0)X$	Coef	Std. Err.	P-value
Outcome: Self-rated health			
Propensity score	0.376	0.004	0
Constant	-0.352	0.002	0
Outcome: Health conditions			
Propensity score	1.075	0.002	0
Constant	-0.347	0.001	0

The estimation include age, age square, gender, marital status, education level, and survey year fixed effects. Standard error is clustered at individual level.

Reference

ATALAY, K., G. F. BARRETT, And A. STANEVA (2019): "The effect of retirement on elderly cognitive functioning," Journal of Health Economics, 66, 37–53.

BEHNCKE, S. (2012): "Does retirement trigger ill health?" Health Economics, 21, 282–300.

BELLONI, M., E. MESCHI, AND G. PASINI (2016): "The effect on mental health of retiring during the economic crisis," Health Economics, 25, 126–140.

BJÖRKLUND, A. AND R. MOFFITT (1987): "The estimation of wage gains and welfare gains in self-selection models," The Review of Economics and Statistics, 42–49.

BONSANG, E., S. ADAM, AND S. PERELMAN (2012): "Does retirement affect cognitive functioning?" Journal of Health Economics, 31, 490–501.

BRINCH, C. N., M. MOGSTAD, AND M. WISWALL (2017): "Beyond LATE with a discrete instrument," Journal of Political Economy, 125, 985–1039.

CARNEIRO, P., J. J. HECKMAN, AND E. J. VYTLACIL (2011): "Estimating marginal returns to education," American Economic Review, 101, 2754–81.

CARNEIRO, P., M. LOKSHIN, AND N. UMAPATHI (2017): "Average and marginal returns to upper secondary schooling in Indonesia," Journal of Applied Econometrics, 32, 16–36.

CHARLES, K. K. (2004): "Is retirement depressing?: Labor force inactivity and psychological well-being in later life," Research in Labor Economics, 23, 269–299.

CHIQUIAR, D. AND G. H. HANSON (2005): "International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States," Journal of political Economy, 113, 239–281.

COE, N. B. AND G. ZAMARRO (2011): "Retirement effects on health in Europe," Journal of Health Economics, 30, 77–86.

DAVE, D., I. RASHAD, AND J. SPASOJEVIC (2008): "The effects of retirement on physical and mental health outcomes," Southern Economic Journal, 75, 497–523.

DE GRIP, A., M. LINDEBOOM, AND R. MONTIZAAN (2012): "Shattered dreams: The effects of changing the pension system late in the game," Economic Journal, 122, 1–25.

EIBICH, P. (2015): "Understanding the effect of retirement on health: Mechanisms and heterogeneity," Journal of Health Economics, 43, 1–12.

FE, E. AND B. HOLLINGSWORTH (2016): "Short- and long-run estimates of the local effects of retirement on health," Journal of the Royal Statistical Society - Series A, 179, 1051– 1067.

GRUBER, J. AND D. A. WISE (2009): Social security programs and retirement around the world: Micro-estimation, University of Chicago Press.

HECKMAN, J. J., J. E. HUMPHRIES, AND G. VERAMENDi (2018): "Returns to education: The causal effects of education on earnings, health, and smoking," Journal of Political Economy, 126, S197–S246.

HECKMAN, J. J., S. URZUA, AND E. VYTLACIl (2006): "Understanding instrumental variables in models with essential heterogeneity," The Review of Economics and Statistics, 88, 389–432.

HECKMAN, J. J. AND E. VYTLACIL (2001): "Policy-relevant treatment effects," American Economic Review, 91, 107–111. (2005): "Structural equations, treatment effects, and econometric policy evaluation 1," Econometrica, 73, 669–738.

HECKMAN, J. J. AND E. J. VYTLACIL (1999): "Local instrumental variables and la- tent variable models for identifying and bounding treatment effects," Proceedings of the national Academy of Sciences, 96, 4730–4734.

HELLER-SAHLGREN, G. (2017): "Retirement blues," Journal of Health Economics, 54, 66– 78.

INSLER, M. (2014): "The health consequences of retirement," Journal of Human Resources, 49, 195–233.

JOHNSTON, D. W. AND W. LEE (2009): “Retiring to the good life? The short-term effects of retirement on health,” Economics Letters, 103, 8–11.

KOLODZIEJ, I. W. AND P. GARC´IA-GO´MEZ (2019): “Saved by retirement: Beyond the mean effect on mental health,” Social Science & Medicine, 27, 85–97.

MAZZONNA, F. AND F. PERACCHI (2012): “Ageing, cognitive abilities and retirement,” European Economic Review, 56, 691–710 (2017): “Unhealthy retirement?” Journal of Human Resources, 52, 128–151.

MCKENZIE, D. AND H. RAPOPORT (2010): “Self-selection patterns in Mexico-US migra- tion: the role of migration networks,” the Review of Economics and Statistics, 92, 811–821.

NYBOM, M. (2017): “The distribution of lifetime earnings returns to college,” Journal of Labor Economics, 35, 903–952.

PICCHIO, M. AND J. C. VAN OURS (2019): “Mental health effects of retirement,” De Economist, 168, 419–452.

ROHWEDDER, S. AND R. J. WILLIS (2010): “Mental retirement,” Journal of Economic Perspectives, 24, 119–138.

ROOTH, D.-O. AND J. SAARELA (2007): “Selection in migration and return migration: Evidence from micro data,” Economics letters, 94, 90–95.

Van Ours, J. C. (2022): “How Retirement Affects Mental Health, Cognitive Skills and Mortality; An Overview of Recent Empirical Evidence,” De Economist, 1–26.

VYTLACIL, E. (2002): “Independence, monotonicity, and latent index models: An equivalence result,” Econometrica, 70, 331–341.

ZHOU, X. AND Y. XIE (2019): “Marginal treatment effects from a propensity score perspective,” Journal of Political Economy, 127, 3070–3084.

Online Appendix

A Additional Tables and Figures

Table A.1: First stage estimation with two instruments: 2SLS and Marginal effects at means from the logistic regression

	(1)	(2)
Independent variables	Retired	
Years to statutory retirement age (60)	0.036***	0.028***
	(0.007)	(0.0089)
Above pension age (65)	0.026	0.020
	(0.018)	(0.020)
Kp Wald F statistics	30.75	-
$\chi 2$ for test of the excluded instruments	-	17.98
Observations	16,694	16,694

Robust standard errors in parentheses are clustered at individual level. All regressions include covariates: age, age squre, gender, marital status, education level, survey year fixed effects.

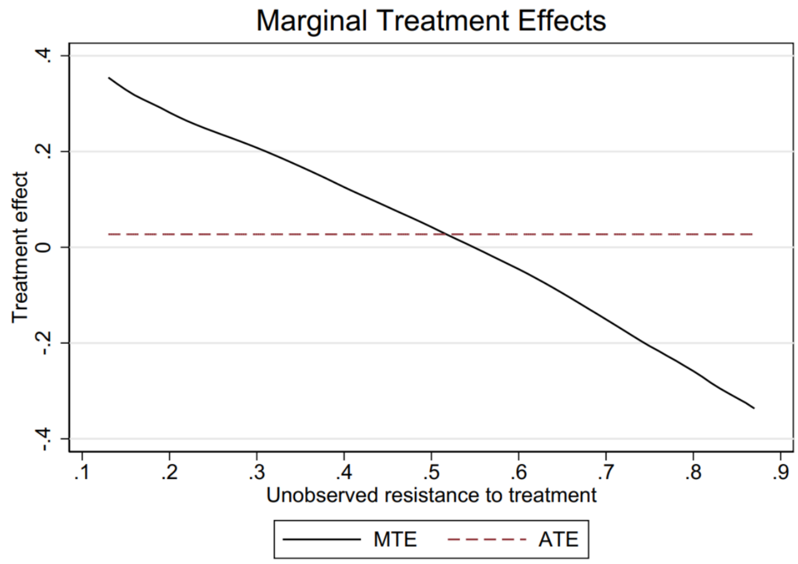

Figure A.1: Estimation results of MTE for self-rated health

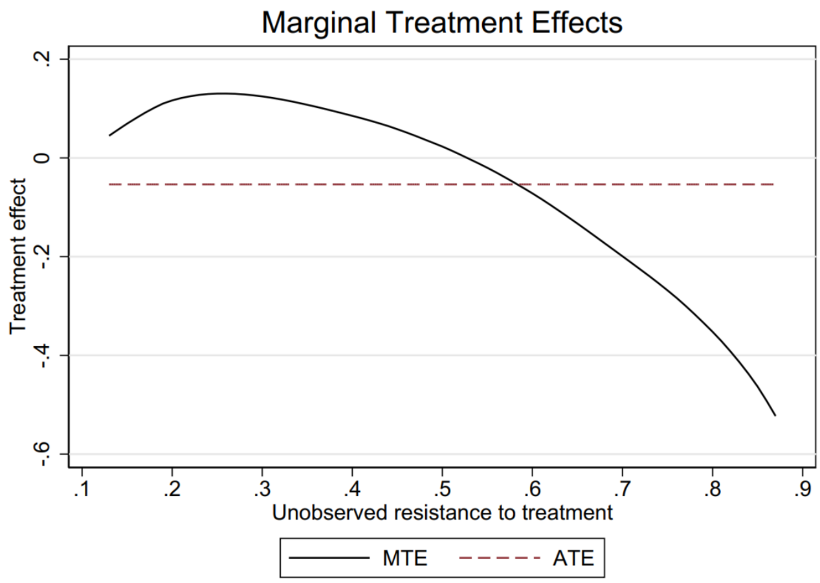

Figure A.2: Estimation results of MTE for health conditions