Journal of Water Economics 53 (2026): 33–57

# Contributions of aquatic environments to household health expenditure: Empirical evidence from Japan

Kei Kabaya
Faculty of Economics, Chuo University

A growing number of academic works have presented empirical evidence on the positive impacts of nature exposure on human health, however, the associations between surrounding nature (especially aquatic environments) and health expenditure have been rarely examined up until now. The major purpose of this paper is to fill in this research gap using health expenditure data at the household level and fine-scale land cover information that was reclassified into four categories (i.e. deciduous forests, evergreen forests, inland freshwater and coastal saltwater). Employing a classical health economics approach, namely, a two-part model with a generalized linear model, this study found that freshwater coverage around residence would have significantly inverse effects on household health spending. As no such significant associations were found for other land cover types, this highlights the importance of surrounding freshwater environments in determining household health expenditure. These findings were mostly robust to the alternative model specifications and different buffer sizes around residence.

**JEL:** I18, Q51, Q57

**Keywords:** Health expenditure; Household; Aquatic environment; Freshwater; Forest

---

**Corresponding author:** Kei Kabaya (kkabaya901@g.chuo-u.ac.jp).

**Disclosure statement:** This work was supported by the Grants-in-Aid for Scientific Research in Japan (22K13397). The data and code used for this analysis are available from the author upon request.

---

Manuscript received: 10 Jun 2025; final version accepted: 10 Dec 2025; available online: 18 Dec 2025.

## 1 Introduction

The growing number of academic works have presented empirical evidence on the positive impacts of nature exposure on human health. It has been found that higher greenness around residence and more green space in urban areas (often measured by land cover data) would enhance physical activity, improve psychological health and reduce mortality rates (Aerts et al., 2018; Gascon et al., 2015, 2016; James et al., 2015; Kondo et al., 2018; Lachowycz and Jones, 2011; Lee and Maheswaran, 2011; van den Berg et al., 2015). Also, a number of scholars have revealed positive health effects of more direct and intentional exposure to forests (e.g. forest bathing), such as improved heart functions and immune systems, reduced stress, depression and mental fatigue, and enhanced cognitive performance, among others (Lee et al., 2017; Rajoo et al., 2020; Wen et al., 2019).

Positive associations between aquatic environments and human health have been also unveiled recently (White et al., 2020). Several studies have found that those who lived closer to the coasts were likely to report better health conditions (Hooyberg et al., 2020; Wheeler et al., 2012; White et al., 2013a) and more frequent physical activities (Bauman et al., 1999). A higher proportion of aquatic areas (including rivers and lakes) around residence was also found to be positively associated with mental health (de Vries et al., 2016; McDougall et al., 2021; Pearson et al., 2019). Some scholars addressed the importance of visibility (Dempsey et al., 2018; Garrett et al., 2019; Nutsford et al., 2016) and intentional exposure (Gascon et al., 2017; Völker et al., 2018; White et al., 2013b) in the relationship between aquatic environments and human health.

Despite such research advancement over the past few decades, associations between nature and health expenditure have been overlooked up until now. Examining such a relationship is vital for considering potential economic contributions of nature not only to household but also public finance with regard to reducing healthcare costs. Still, there are only a limited number of studies that investigated this issue. These studies basically found inverse associations between forest coverage and health expenditure (Becker et al., 2019; Becker and Browning, 2021; Kabaya, 2021). One major caveat of these works is, however, using collective data to derive individual-level healthcare costs. This approach is inconsistent with the above medical research that basically utilize personal records in the analyses.

The objective of this paper is to explore more robust evidence of nature's contribution to health expenditure using household level data in Japan. To the best of my knowledge, there is only a single study that examined the associations between surrounding natural environments and health spending at the individual level (Van Den Eeden et al., 2022). Building on this foundation, the present study focuses on health expenditure of households rather than individuals to evaluate more holistic health impacts of surrounding nature. Furthermore, this study extends the scope of natural environments to include aquatic environments, i.e. inland freshwater (e.g. rivers and lakes) and coastal saltwater (e.g. ocean and seas), in addition to forest environments. Since there is no single study that investigated the associations between water bodies and health expenditure at the micro level, this study will fill in a significant research gap in this field.

The rest of the paper is structured as follows. Section 2 reviews the literature and further clarifies the novelty of this study. Following the explanation of data collection and econometric specification in Section 3, descriptive statistics and estimation results are presented in Section 4. The next section discusses interpretations and implications of the major findings and addresses some caveats of this study. Then, Section 6 concludes.

## 2 Literature Review

As noted earlier, only a limited number of articles investigated the associations between surrounding natural environments and health expenditure. One of the earlier applications is Becker et al. (2019). They examined five specific types of green land cover and found that forest and shrub covers were significantly and inversely associated with per capita healthcare spending (whereas no such relations were found for grassland, farmland and urban vegetation). What should be noted is that they utilised collective data at the county level to derive per capita health expenditure rather than collect data from each individual.

This also holds for Becker and Browning (2021). They expanded the previous analysis by using Normalized Difference Vegetation Index (NDVI) to measure overall greenness, alongside a measure for blue space (open water), at the county level. They reported a significant and inverse association between per-capita health spending and total greenness while found no significant association with blue space. Note, however, that their study did not differentiate among types of water bodies such as rivers, lakes, and coastal areas.

Kabaya (2021) also used collective data at the prefectural level to derive per capita health spending. This study focused on forest coverage as well as forest-related indicators, finding that a mixed forest coverage and urban-forest proximity had significantly inverse long-run effects on per capita health expenditure. Meanwhile, overall, evergreen, and deciduous forest coverage, as well as regional vegetation diversity, were found to be insignificant in determining healthcare costs.

Unlike these articles, Van Den Eeden et al. (2022) utilized individual-level health expenditure. They examined the associations with residential green cover (measured by NDVI and tree canopy) and reported a significantly inverse association between higher levels of green cover and lower direct healthcare costs. Their sensitivity analyses that included adjustments for health factors showed an attenuation of the NDVI-cost association, implying that green cover would reduce healthcare costs by improving health status of individuals.

Apart from these more relevant studies, the determinants of medical expenditures at the individual or household level have long been studied in the health economics literature. Some found that the level of health expenditure increased with age (Angulo et al., 2011; Chaze, 2005; Rous and Hotchkiss, 2003), whereas the other underlined the importance of proximity to death (Werblow et al., 2007). Wong et al. (2011) further investigated this controversial issue and reported that proximity to death was a more significant predictor of high health expenditure especially for lethal diseases while age would be a more crucial factor for chronic diseases. Individual health status per se was also found to be a significant factor in determining the level of healthcare costs (Deb et al., 2006; Grassetti and Rizzi, 2019).

Beyond these factors, other demographic characteristics may likewise be important. For example, women are likely to generate a higher level of health expenditure than men (Angulo et al., 2011; Grassetti and Rizzi, 2019), and wealthier families opt to spend more for healthcare (López-Nicolás, 1998; Matsaganis et al., 2009; Okunade et al., 2010; Rous and Hotchkiss, 2003; Rout, 2008). The effects of surrounding environments, especially urban-rural contrasts, have been also examined in the literature, however, the findings were rather mixed. Rout (2008) and Rous and Hotchkiss (2003) found that the urbanization level was inversely associated with the health expenditure whereas Chaze (2005) found the opposite result.

Taken all together, these previous works suggest that various factors, including green covers, influence the level of individual or household health spending. Nevertheless, only a limited number of studies has investigated the effects of aquatic environments. More specifically, no single study has examined the impacts of freshwater and coastal saltwater on household health expenditure up until now. Thus, this study will fill in a significant research gap in this field.

# 3 Methodology

## 3.1 Data Acquisition

Japan institutionalizes universal health insurance systems that cover all national citizens. The level of health benefits varies from 70% to 90% depending on one's age primarily and income subsidiarily[1]. Since health insurance costs are often deducted from one's revenues beforehand, insurance is almost always applied at checkouts in hospitals and pharmacies. This helps Japanese people to perceive the cost adjusted after health benefits (i.e. out-of-pocket) as their medical expenditures. In line with such a mindset, the current study put a focus on the out-of-pocket health expenditure. Although there is a national healthcare database covering almost all individual medical expenses, it registers no personal residential information unfortunately. Alternatively, this study collects household health expenditure through a self-completed questionnaire survey.

An internet survey of general Japanese citizens (aged between 18 and 69) was conducted between July and October in 2022 to collect household health expenditure. Respondents were recruited through an existing online panel of a domestic research company and were requested to present total expenses for healthcare and prescriptions (except dental care) for all household members in the fiscal year 2021 (i.e. from April 2021 to March 2022). With a view to obtaining accurate values, I asked them to refer to the payment notification that has recorded all healthcare-related actual spending for a certain period of time. This notification must have been provided by their own health insurers as long as they have taken any medical treatments. Put differently, those who have not received this notification must have taken no healthcare services for an entire year, and were asked to report zero expenditure in the questionnaire.

Household characteristics were also collected during the survey. With reference to several health economic studies that examined the determinants of health expenditure at the household level (Chaze, 2005; López-Nicolás, 1998; Matsaganis et al., 2009; Okunade et al., 2010; Rous and Hotchkiss, 2003), the following information was obtained, namely,

[1]Minors are further supported financially by local governments. The eligibility and the level of assistance depend on each municipality.

family size, each family member's age, sex and chronic conditions as well as household income.

Measurement of surrounding natural environments began with identifying respondents' locations. Because it was impossible to directly ask their home addresses for privacy, I asked seven-digit postal codes alternatively during the survey. Then, representative points in the respective postal zones were identified using an address matching technique, as proxies of their home locations[2]. The surrounding natural environments were evaluated using a land cover map, obtained from the Earth Observation Research Center in Japan. This map represents average surface conditions in 10 meters grids between 2018 and 2020 and categorizes land covers into 12 distinct types. The present analysis focused on the following five categories: deciduous broad-leaf forests, deciduous needle-leaf forests, evergreen broad-leaf forests, evergreen needle-leaf forests and aquatic surfaces, among others. Considering the commonality and discrepancy, four types of these forests were merged into either 'deciduous forests' or 'evergreen forests' while water surfaces were separated into 'inland freshwater' and 'coastal saltwater' using coastlines as boundaries. Additionally, the coverage of urban area was also computed to control for the urbanization level, which may influence household healthcare behaviors (Chaze, 2005; Sanwald and Theurl, 2017). See Figure 1 for the sample land cover maps.

An appropriate size of surrounding area will be an essentially empirical question. The relevant public health research applied different buffer sizes that varied from several hundred meters to a few kilometers (Gascon et al., 2015). This study tests four alternative radiuses incrementing from 0.5 km to 2.0 km at 0.5 km steps and examines the estimation results comparably.

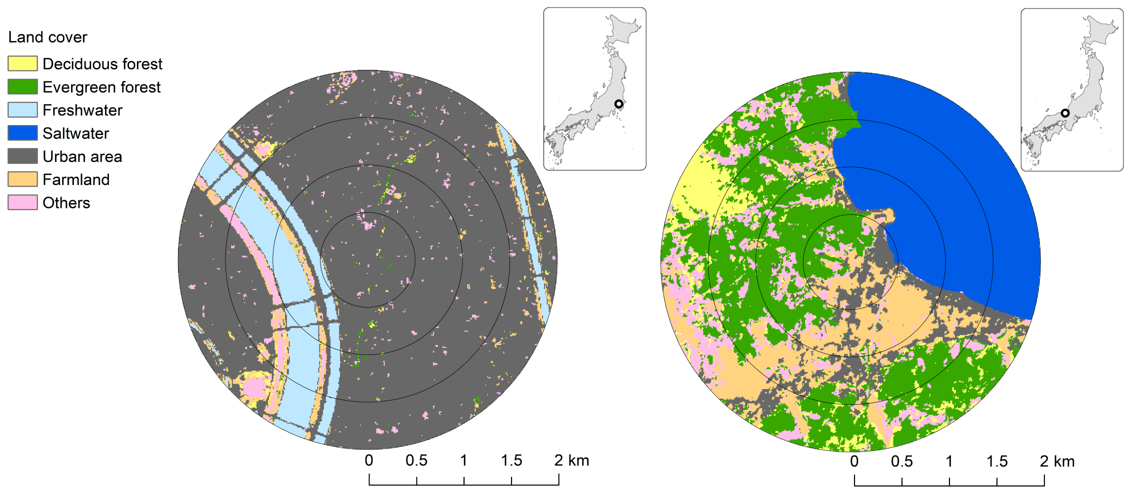


**Figure 1:** Sample land cover

[2]Applying such a methodology resulted in locating two households sharing the same zip code at the same location and defining the same surrounding environments consequently. The unforeseeable effects of such a treatment cannot be examined in this study and thus will be reserved for future work.

### 3.2 Econometric Specification

Theoretical framework of the analysis on health expenditure is described in Appendix A. Practically, heavy tail and zero spike frequently observed in the distribution of micro-level health expenditure needs to be fully taken into account. A number of approaches have been proposed to overcome such challenges (see Jones, 2000; Mihaylova et al., 2011) and some studies compared different econometric specifications to explain health expenditure (Basu et al., 2006; Buntin and Zaslavsky, 2004; Deb and Trivedi, 2002; Jones et al., 2016; Madden, 2008; Manning and Mullahy, 2001; Matsaganis et al., 2009). One of the major approaches to deal with heavy tail and zero spike is to employ log models and censored approaches. The former models include an ordinary least square (OLS) model on a logged dependent variable and a generalized linear model (GLM) with a log-link function. The latter approaches can be exemplified by a two-part model (2PM) and a sample selection model (SSM).

The present study follows such traditional approaches and opts to employ the 2PM with the GLM specification (2PM-GLM) with the log-link function. The major reasons for selecting the 2PM rather than the SSM are; 1) this study attempts to model actual health spending with genuine zeros rather than potential outcomes that may be unobserved; 2) some assumptions such as exclusion restrictions and error correlations required to use the SSM are not necessarily presumed a priori in this study. As for the selection of the GLM with the log-link function rather than the log-transformed OLS, this study follows the guidance provided by Manning and Mullahy (2001), which suggests using the former specification when a coefficient of kurtosis of the log-scale residual from the OLS is smaller than three. Preliminary analysis revealed that coefficients of kurtosis were slightly smaller than this threshold, regardless of the buffer sizes employed[3]. This is the foundation of applying the GLM with the log-link function for this study[4].

The first part of the 2PM considers the probability that a household needs medical care during a certain period (i.e., one year in this analysis). Let $y_i$ be the health expenditure of a household $i$, and define $h_i = 1$ if a household takes medical treatments and $h_i = 0$ otherwise. The probability to observe $y_i > 0$ can be:

$$\Pr(y_i > 0) = \Pr(h_i = 1) = \Pr(h_i^* > 0),$$

where $h_i^*$ is a latent variable that represents utility difference between two alternatives available to a household (i.e., whether to receive healthcare services or not). Assume that $h_i^*$ is additively separable and can be expressed as:

$$h_i^* = x_i'\beta + \varepsilon_i,$$

where $x$ is a vector of explanatory variables, $\beta$ is the corresponding vector of parameters and $\varepsilon_i$ is a normally distributed error term. This specification derives the following probit

[3] Manning and Mullahy (2001) also recommended to divide the coefficient of kurtosis by the square root of the estimated variance function so as to rule out heteroskedasticity. Following such a guideline and using the variance function proposed by Western and Bloome (2009), I found that the coefficients of kurtosis were slightly smaller than three for all buffer sizes.

[4] Still, the alternative approaches mentioned above (i.e. SSM and the log-transformed OLS) as well as one-part models (1PM) are also tested to check the robustness of the results.

function:

$$\Pr(y_i > 0) = \Pr(h_i^* > 0) = \Pr(\varepsilon_i > -x_i'\beta) = 1 - \Phi(-x_i'\beta) = \Phi(x_i'\beta),$$

where $\Phi(\cdot)$ represents the standard normal cumulative distribution function.

The second part of the 2PM explains the levels of health expenditure among those who have taken medical treatments. The positives, $E(y_i \mid y_i > 0)$, can be modeled using the GLM specification as:

$$E(y_i \mid y_i > 0) = g^{-1}(x_i'\gamma),$$

where $g(\cdot)$ denotes a monotonic and differentiable link function and $\gamma$ is the vector of estimable parameters. Applying the log-link function, as is typically the case in health expenditure applications, transforms this equation into:

$$\ln\left(E(y_i \mid y_i > 0)\right) = x_i'\gamma.$$

The overall mean can be written as the product of expectations from the first and the second parts of the model as follows:

$$E(y_i) = \Pr(y_i > 0) \times E(y_i \mid y_i > 0).$$

The likelihood contribution for each observation can be written as:

$$L = \left[\Phi(x_i'\beta)g^{-1}(x_i'\gamma)\right]^{I(y_i>0)} \times \left[1 - \Phi(x_i'\beta)\right]^{I(y_i=0)},$$

where $I(\cdot)$ is an indicator function that takes 1 if the condition is met and 0 otherwise. Then, the log-likelihood contribution is

$$\ln L = I(y_i > 0)\left[\ln\{\Phi(x_i'\beta)\} + \ln\{g^{-1}(x_i'\gamma)\}\right] + I(y_i = 0)\ln[1 - \Phi(x_i'\beta)].$$

Since $\beta$ and $\gamma$ are additively separable in the log-likelihood function for each observation, the first and the second part of the model can be estimated separately.

Finally, marginal effects of respective variables are computed using the following equation:

$$\begin{aligned}\frac{\partial \hat{y}_i}{\partial x_j} = \frac{\partial \Pr(y_i > 0)\, \hat{y}_i \mid (y_i > 0)}{\partial x_j} &= \hat{\beta}_j\, \phi(x_i'\hat{\beta})\left(\hat{y}_i \mid y_i > 0\right) + \Phi(x_i'\hat{\beta})\, \hat{\gamma}_j \left(\hat{y}_i \mid y_i > 0\right) \\ &= \left[\frac{\hat{\beta}_j\, \phi(x_i'\hat{\beta})}{\Phi(x_i'\hat{\beta})} + \hat{\gamma}_j\right]\hat{y}_i. \end{aligned} \tag{1}$$

# 4 Results

## 4.1 Descriptive Statistics

Descriptive statistics of 659 valid respondents (households) collected is summarized in Table 1. The mean household health expenditure was approximately JPY 29400 (equivalent to USD 248 as of March 2022). This value was relatively close to the national statistic (i.e. around JPY 34000) recorded in the Annual Report on the Family Income and Expenditure Survey in 2021. Though not indicated on the table, around 30% of the

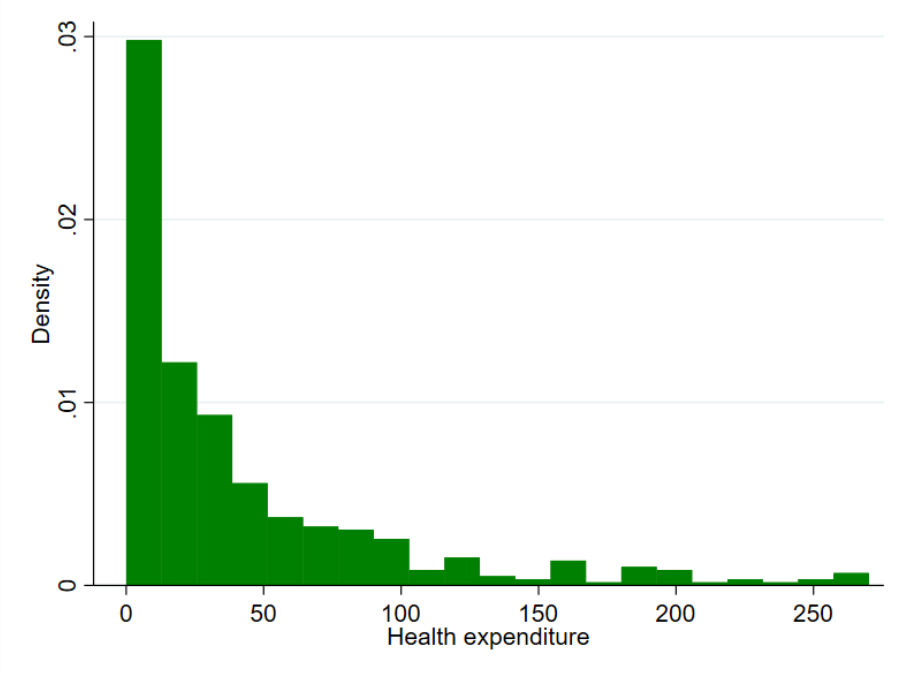


**Figure 2:** Distribution of household health expenditure

sampled households spent zero for healthcare. Figure 2 illustrates the distribution of household health expenditure excluding these zeros. As is evident, it was highly skewed to the right and seemed to follow the gamma distribution.

The mean household size (i.e. *circa* 2.4) and the average share of male families (i.e. *circa* 0.5) were also similar to the national statistic based on the National Census in 2020. In combination with the above household health expenditure, it can be said that my sample represented Japan's households well. Note, however, that the mean share of the members aged over 65 (i.e. *circa* 0.1) was below the national average (i.e. *circa* 0.3). This could be because the survey was conducted for the citizens below 70 years old in consideration of the online panel registration of the research company, despite that those who were older than this age accounted for more than 20% of the population in Japan. Thus, some cautions are required to interpret the results from the current study with this regard.

The mean coverages of natural environments by buffer size are also indicated in the lower part of Table 1. What should be noted is that the mean coverages of all natural land cover categories became larger in proportion to the expansion of buffer size. The total natural coverage on average increased from around 10% (i.e. 0.5 km buffer size) to 15% (i.e. 2.0 km buffer size) accordingly. Conversely, the proportion of urban area shrank as the radius became larger. Still, it dominated surrounding surface areas around residence (i.e. 63% in 2.0 km buffer size), implying that most of the sampled household lived in relatively developed areas.

**Table 1:** Descriptive statistics

| Variable | Description | Mean | S.D. | Min | Max |
|---|---|---|---|---|---|
| y | Health expenditure (thousand JPY) | 29.400 | 48.718 | 0.000 | 270.154 |
| Family | Family size | 2.405 | 1.258 | 1.000 | 7.000 |
| Elder | % of families aged over 65 | 11.680 | 26.855 | 0.000 | 100.000 |
| Male | % of male families | 52.268 | 33.147 | 0.000 | 100.000 |
| Chronic | % of families with health problems | 31.927 | 40.806 | 0.000 | 100.000 |
| Income | Household income (million JPY) | 5.747 | 4.160 | 0.500 | 27.500 |
| Coverage05 | Coverage within 0.5 km radius (%) | | | | |
| Deciduous | Deciduous forests | 3.330 | 7.757 | 0.000 | 88.978 |
| Evergreen | Evergreen forests | 4.642 | 10.800 | 0.000 | 82.172 |
| Freshwater | Inland freshwater | 1.352 | 4.163 | 0.000 | 36.335 |
| Saltwater | Coastal saltwater | 0.670 | 5.238 | 0.000 | 100.000 |
| Urban | Urban area | 71.405 | 26.844 | 0.000 | 99.796 |
| Coverage10 | Coverage within 1.0 km radius (%) | | | | |
| Deciduous | Deciduous forests | 3.821 | 7.612 | 0.000 | 79.340 |
| Evergreen | Evergreen forests | 5.091 | 10.177 | 0.000 | 69.293 |
| Freshwater | Inland freshwater | 1.745 | 4.239 | 0.000 | 39.839 |
| Saltwater | Coastal saltwater | 1.504 | 6.863 | 0.000 | 80.108 |
| Urban | Urban area | 67.789 | 26.032 | 0.010 | 98.202 |
| Coverage15 | Coverage within 1.5 km radius (%) | | | | |
| Deciduous | Deciduous forests | 4.110 | 7.611 | 0.000 | 76.057 |
| Evergreen | Evergreen forests | 5.542 | 10.068 | 0.011 | 75.028 |
| Freshwater | Inland freshwater | 1.851 | 3.927 | 0.000 | 32.543 |
| Saltwater | Coastal saltwater | 2.200 | 7.784 | 0.000 | 57.948 |
| Urban | Urban area | 65.218 | 25.692 | 0.014 | 97.526 |
| Coverage20 | Coverage within 2.0 km radius (%) | | | | |
| Deciduous | Deciduous forests | 4.364 | 7.707 | 0.000 | 74.081 |
| Evergreen | Evergreen forests | 5.975 | 10.043 | 0.031 | 75.851 |
| Freshwater | Inland freshwater | 1.946 | 3.660 | 0.000 | 28.618 |
| Saltwater | Coastal saltwater | 2.753 | 8.445 | 0.000 | 55.026 |
| Urban | Urban area | 63.270 | 25.601 | 0.008 | 96.652 |

*Note:* The number of observations is 659.

## 4.2 Estimation Results

Prefectural dummy variables were additionally included in the estimation models to control regional heterogeneity[5]. Table 2 presents the estimation results of the 2PM-GLM with the log-link function. The major results were qualitatively similar across four different buffer sizes. Yet, two information criteria, namely Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC), suggested that the models with 1.0 km radius (i.e. 2PM-GLM II) and 1.5 km radius (i.e. 2PM-GLM III) were slightly superior to alternative specifications. In what follows, thus, the estimation results of these two models were particularly described.

The variables *Family*, *Male* and *Chronic* were statistically significant in the first part

[5]Estimation results of these variables are omitted from the tables to preserve the space. Still, they are available from the author upon request.

**Table 2:** Estimation results of household health expenditure

| | 2PM-GLM I $r$ = 0.5 km | | 2PM-GLM II $r$ = 1.0 km | | 2PM-GLM III $r$ = 1.5 km | | 2PM-GLM IV $r$ = 2.0 km | |
|---|---|---|---|---|---|---|---|---|
| Variable | Coef | S.E. | Coef | S.E. | Coef | S.E. | Coef | S.E. |
| First-part | | | | | | | | |
| Family | 0.173*** | 0.042 | 0.175*** | 0.042 | 0.177*** | 0.042 | 0.178*** | 0.043 |
| Elder | 0.003 | 0.003 | 0.002 | 0.003 | 0.002 | 0.003 | 0.002 | 0.003 |
| Male | −0.004* | 0.002 | −0.004* | 0.002 | −0.004* | 0.002 | −0.004* | 0.002 |
| Chronic | 0.023*** | 0.002 | 0.023*** | 0.002 | 0.023*** | 0.002 | 0.023*** | 0.002 |
| Income | 0.026 | 0.021 | 0.026 | 0.020 | 0.026 | 0.020 | 0.026 | 0.020 |
| Deciduous | −0.012* | 0.007 | −0.009 | 0.010 | −0.007 | 0.011 | −0.009 | 0.011 |
| Evergreen | 0.004 | 0.011 | 0.008 | 0.012 | 0.002 | 0.011 | 0.000 | 0.012 |
| Freshwater | 0.003 | 0.015 | 0.006 | 0.015 | 0.006 | 0.017 | 0.004 | 0.018 |
| Saltwater | 0.002 | 0.016 | 0.001 | 0.013 | −0.001 | 0.012 | −0.004 | 0.011 |
| Urban | −0.001 | 0.005 | 0.001 | 0.005 | 0.000 | 0.005 | 0.000 | 0.005 |
| Constant | −0.417 | 0.449 | −0.569 | 0.513 | −0.561 | 0.503 | −0.455 | 0.487 |
| Second-part | | | | | | | | |
| Family | 0.261*** | 0.041 | 0.261*** | 0.042 | 0.255*** | 0.044 | 0.252*** | 0.045 |
| Elder | 0.007*** | 0.002 | 0.007*** | 0.002 | 0.008*** | 0.002 | 0.007*** | 0.002 |
| Male | −0.002 | 0.002 | −0.002 | 0.002 | −0.002 | 0.002 | −0.002 | 0.002 |
| Chronic | 0.013*** | 0.002 | 0.013*** | 0.002 | 0.012*** | 0.002 | 0.013*** | 0.002 |
| Income | 0.023** | 0.011 | 0.022* | 0.012 | 0.022* | 0.012 | 0.022* | 0.012 |
| Deciduous | 0.014 | 0.010 | 0.017* | 0.009 | 0.019* | 0.010 | 0.016 | 0.010 |
| Evergreen | 0.001 | 0.008 | −0.009 | 0.008 | −0.011 | 0.008 | −0.007 | 0.008 |
| Freshwater | −0.021* | 0.011 | −0.037*** | 0.009 | −0.040*** | 0.014 | −0.035** | 0.015 |
| Saltwater | 0.009 | 0.008 | 0.003 | 0.007 | 0.001 | 0.007 | 0.000 | 0.007 |
| Urban | 0.003 | 0.003 | 0.000 | 0.005 | −0.001 | 0.005 | −0.001 | 0.005 |
| Constant | 1.815*** | 0.358 | 1.998*** | 0.428 | 2.026*** | 0.433 | 2.058*** | 0.403 |
| Log-likelihood | -2358.6 | | -2355.7 | | -2355.7 | | -2357.3 | |
| AIC | 4759.3 | | 4753.4 | | 4753.5 | | 4756.7 | |
| BIC | 4852.1 | | 4846.3 | | 4846.3 | | 4849.5 | |

*Note:* The number of observations is 614 (several prefecture dummies perfectly predicted positive outcomes in the first-stage probit regression and were therefore omitted). ***, **, and * denote statistical significance at 1%, 5%, and 10% levels, respectively. Prefectural dummies are skipped to preserve space.

of the model. The estimated sign indicated that larger families are more likely to take medical treatments. Likewise, the family that has more members with health problems has higher probability to see a doctor. In contrast, the sign of *Male* was negative, implying that families with a higher proportion of females are more likely to take medical treatments. These results are generally consistent with previous findings (Chaze, 2005; Okunade et al., 2010). Meanwhile, this study found no significant associations between the probability of taking medical treatments and the share of elderly people in a family as well as household income. Also, the surrounding environments were found to be insignificant.

Turning to the second-part, which modeled the level of household health expenditure among those who have taken medical treatments. The variables *Family* and *Chronic* were statistically significant as with the first part of the model. The larger family in size and/or the family having more members with health problems not only have higher probability to take medical treatments but also pay more out-of-pocket health expenditures. Additionally, *Elder* and *Income* were also significantly positive, implying that households with more elderly people and/or with more income were likely to spend more for healthcare, as with Matsaganis et al. (2009). Whereas, the share of male families did not exhibit significant association with the level of health expenditure, unlike the first-part of the model. These results can be interpreted that the family with a higher female proportion is more probable to take medical treatments but would not necessarily spend more for them.

The focus is now shifted to the surrounding natural environments at the bottom of Table 2. The variables *Deciduous* and *Freshwater* showed statistically significant associations with the level of household health expenditure. Yet, their empirical implications are not the same. The sign of *Deciduous* was positive, indicating that those who are surrounded by more deciduous forests would spend more for medical care. What should be noted is that this variable was significant only at the 10% level and was no longer significant in the 2PM–GLM I and 2PM–GLM IV specifications. Thus, this result will be inconclusive and should be understood with caution. Meanwhile, the sign of *Freshwater* was negative, suggesting that those who reside in the area with more freshwater environments are prone to spend less for medical care. This variable was significant at the 1% level and statistically significant in all models regardless of the buffer size employed.

As can be observed, some variables were significant in one part of the two-part model but not in the other. This can be explained by the distinct mechanisms each part captures: the first part reflects the decision to seek healthcare, whereas the second part reflects the level of health spending once it is sought. In this context, it is reasonable that the land cover variables were insignificant in the first part, as natural environments are unlikely to trigger immediate care-seeking compared with more direct factors such as family size or chronic conditions. In contrast, *Freshwater* showed the significantly negative association in the second part, suggesting that surrounding freshwater environments may reduce the severity of illness or the required intensity of treatment among those who use medical services. The possible reason why only freshwater environments were significant will be discussed later.

The marginal effects are computed using Eq. (1) and based on the estimation results (see Table 3). It suggests that one more headcount in a family would increase household out-of-pocket healthcare expenditure by around JPY9000 (i.e. USD76). Likewise, increase in 1% point of the share of elderly person in a family and a member with chronic illness would require additional JPY240 and JPY540, respectively. Conversely, 1% point increase in inland freshwater in surrounding environments would decrease household health spending by approximately JPY1100 (i.e. USD9) on average (see 2PM-GLM II and 2PM-GLM III). What should be mentioned is that the marginal effects of deciduous forests were all insignificant. From these consequences, I would conclude that deciduous forests were insignificant in determining household health expenditure at least in this study.

Now, check the robustness of the estimation results with the alternative specifications. In this study, four different models, namely, the one-part model (1PM) with log-transformed OLS (1PM–OLS), the 1PM with the GLM specification and log-link function (1PM–GLM), the two-part model (2PM) with log-transformed OLS (2PM–OLS), and the sample selection

**Table 3:** Marginal effects on household health expenditure

| | 2PM-GLM I $r$ = 0.5 km | | 2PM-GLM II $r$ = 1.0 km | | 2PM-GLM III $r$ = 1.5 km | | 2PM-GLM IV $r$ = 2.0 km | |
|---|---|---|---|---|---|---|---|---|
| Variable | Coef | S.E. | Coef | S.E. | Coef | S.E. | Coef | S.E. |
| Family | 9.063*** | 1.348 | 9.088*** | 1.368 | 8.905*** | 1.381 | 8.837*** | 1.429 |
| Elder | 0.227*** | 0.061 | 0.238*** | 0.063 | 0.245*** | 0.069 | 0.233*** | 0.065 |
| Male | −0.095 | 0.059 | −0.098 | 0.062 | −0.095 | 0.062 | −0.092 | 0.062 |
| Chronic | 0.555*** | 0.051 | 0.542*** | 0.050 | 0.539*** | 0.050 | 0.543*** | 0.051 |
| Income | 0.879** | 0.370 | 0.837** | 0.378 | 0.847** | 0.383 | 0.859** | 0.389 |
| Deciduous | 0.340 | 0.289 | 0.443 | 0.288 | 0.512 | 0.317 | 0.404 | 0.300 |
| Evergreen | 0.073 | 0.240 | −0.207 | 0.262 | −0.309 | 0.257 | −0.202 | 0.256 |
| Freshwater | −0.602* | 0.343 | −1.072*** | 0.302 | −1.152*** | 0.443 | −1.011** | 0.456 |
| Saltwater | 0.280 | 0.262 | 0.094 | 0.228 | 0.020 | 0.228 | −0.041 | 0.230 |
| Urban | 0.074 | 0.109 | 0.004 | 0.145 | −0.016 | 0.158 | −0.025 | 0.154 |

*Note:* The standard errors were computed using the delta method. ***, **, and * denote statistical significance at 1%, 5%, and 10% levels, respectively.

model with log-transformed OLS (SSM–OLS), were comparably tested. Table 4 presents the estimation results with a specific focus on freshwater coverage to preserve the space (see Tables B1– B4 in the Appendix for all other variables). The implications drawn from these model outputs were broadly consistent with the main findings, especially for the 1 km and 1.5 km buffer sizes, which showed better statistical performance in every specification. Judging from these results, I conclude that the major finding of this study holds even when different analytical approaches would be employed.

**Table 4:** Estimation results of *Freshwater* with alternative specifications

| | $r$ = 0.5 km | | $r$ = 1.0 km | | $r$ = 1.5 km | | $r$ = 2.0 km | |
|---|---|---|---|---|---|---|---|---|
| Model | Coef | S.E. | Coef | S.E. | Coef | S.E. | Coef | S.E. |
| 1PM-OLS | −0.014 | 0.015 | −0.028** | 0.013 | −0.030* | 0.016 | −0.028 | 0.019 |
| 1PM-GLM | −0.021* | 0.011 | −0.037*** | 0.009 | −0.040*** | 0.014 | −0.035** | 0.015 |
| 2PM-OLS | | | | | | | | |
| First-part | 0.003 | 0.015 | 0.006 | 0.015 | 0.006 | 0.017 | 0.004 | 0.018 |
| Second-part | −0.014 | 0.015 | −0.028** | 0.013 | −0.030* | 0.016 | −0.028 | 0.019 |
| SSM-OLS | | | | | | | | |
| First-part | 0.003 | 0.015 | 0.006 | 0.015 | 0.006 | 0.017 | 0.004 | 0.018 |
| Second-part | −0.014 | 0.014 | −0.028** | 0.012 | −0.031** | 0.015 | −0.028 | 0.017 |

*Note:* ***, **, and * denote statistical significance at 1%, 5%, and 10% levels, respectively.

Furthermore, I examined potential multicollinearity among the explanatory variables using the 1PM–OLS model and variance inflation factors (VIFs). Consequently, the mean VIFs for the respective buffer zones were found to be quite small (see Table B1), and no large VIFs (i.e., greater than five) were detected for any major individual variables. Thus, I conclude that multicollinearity is not a serious concern in this study.

## 5 Discussion

This study found that the freshwater coverage around residence would have a significantly negative association with household health expenditure. A possible underlying mechanism is that nearby rivers and lakes would improve health of the residents and thereby reduce opportunities to take medical treatments and prescriptions. Such positive effects of freshwater to human health have been reported in the literature. For example, McDougall et al. (2021) and Pearson et al. (2019) found that higher neighborhood freshwater coverage was significantly associated with lower antidepressant medication prevalence and lower anxiety disorder hospitalization rate, respectively. De Bell et al. (2017) argued that such psychological benefits may be the most important health impact obtained from visiting freshwater environments. It was also found that the level of mental health was associated with the frequency of visiting freshwater zones (Völker et al., 2018).

Meanwhile, the associations between household health expenditures and forest coverage were found to be statistically insignificant in this study. This is inconsistent with Van Den Eeden et al. (2022), who found the significantly inverse associations between residential green cover and individual health expenditure. Yet, my results corroborate some findings from other public health articles that compared blue space (both freshwater and coastal saltwater) and green space. For instance, Nutsford et al. (2016) showed significant associations of psychological distress with blue space visibility but not with green space. McDougall et al. (2021) reported that neighboring blue space availability would have greater impacts on mental health than nearby green space. Likewise, de Vries et al. (2016) found that the positive effects on mental health were stronger for blue space than green space. The findings from this study can be interpreted in conjunction with these articles and suggest that aquatic environments would have a more important role in determining household health expenditure.

Still, coastal saltwater was also found to be insignificant in this study. This consequence seems to be inconsistent with the previous works that revealed positive health impacts of surrounding coastal environments (Dempsey et al., 2018; Hooyberg et al., 2020; Wheeler et al., 2012; White et al., 2013a). One possible explanation will be different pathways of health impacts between coastal and freshwater environments. The positive associations between coasts and health would be mediated by physical activities (e.g. walking), whereas the positive health effects of freshwater would not (Pasanen et al., 2019). Indeed, increase in physical activities has been frequently argued as a potential mechanism of coastal health effects (Bauman et al., 1999; Wheeler et al., 2012). Insignificance of coastal saltwater can be thus attributed to the possibility that the respondents in my dataset may not have used coastal areas for their physical activities as frequently as reported in these studies. Conversely, mental health effects are often highlighted in freshwater health studies, as discussed above. These effects can arise from more indirect or incidental exposure (e.g., viewing or passing by), which is basically rooted in residential proximity, and were therefore considered to have been detected in this study.

These consequences are far from conclusive, however, considering the small sample size and the specific sample characteristics (e.g. most of the sampled household resided in developed areas) of the current study. Essentially, further works with other datasets will be needed to corroborate such findings. Also, employing other estimation methods, e.g. generalized gamma models (Hill and Miller, 2010; Manning et al., 2005) and extended estimating equations models (Basu et al., 2006; Hill and Miller, 2010)[6], may be required to reach a more robust conclusion.

Policy implications can be drawn from the major finding of this study. Japan currently faces rapid aging of society and the national medical expenditure (including both benefits provided by health insurance systems and patients' out-of-pocket payments) amounted to 46.7 trillion JPY (i.e. 394 billion USD) in 2022. National and local governments allocate certain proportions of tax revenue to support universal health insurance system, which exerts significant pressure on the financial resources of the governments. Under such circumstances, we should be careful not to increase health expenditures any further. The major result of this study suggests that the health expenditure of those who live in the area with broader freshwater coverage has been maintained at a lower level. This, in turn, implies that we should conserve freshwater ecosystems not only for environmental concerns but also for our health and public health administrations.

One major caveat, however, is that the association revealed from this study was correlational rather than causal. It is not completely evident whether the increase in freshwater coverage in a specific area would decrease health expenditure of neighboring households. Put differently, it cannot be strongly proposed that we should restore freshwater ecosystems to reduce national medical spending. To make such a policy recommendation, we need to further clarify the causal relationship between freshwater coverage and household health expenditure with more rigorous methodological approaches (e.g. a randomized controlled trial). Although applying such a methodology is conceptually and technically challenging for this kind of study[7], some analytical methods based on natural experiments might be feasible if we could find a specific area by chance where inland freshwater has emerged recently. In doing so, one should distinguish artificial water bodies (e.g. irrigation channels and multipurpose dams) from natural freshwater ecosystems. This is because different types of inland water would influence human health in a contrastive way (McDougall et al., 2022), and environmental implications derived from such analyses will be completely opposite.

## 6 Conclusion

This study found that the freshwater coverage around residence would have a significantly negative impact on household health expenditure. Since the associations between freshwater environments and human health have been overlooked up until now (McDougall et al., 2020) and their contribution to household health expenditure has never been documented to date, this finding will be an important contribution to the literature. Meanwhile, this study found no significant associations between household health expenditure and the

[6] I estimated an extended estimating equations model preliminarily, and found that it did not converge within the reasonable number of iterations (i.e. 1000). Although I decided not to use this model in this study based on this attempt, such a modeling approach may be suitable for other datasets.

[7] For example, given that one hopes to apply a randomized controlled trial to examine causal relations between surrounding nature and health expenditure, relocating some households forcefully for such a purpose cannot be allowed from moral and human rights perspectives. Even though they agree on the relocation, the researcher may in turn suffer from massive financial costs required for tenancy and so forth.

coverage of forests as well as coastal areas. This suggests that the freshwater coverage would have a more important role in determining household health expenditure than other land cover types. As noted earlier, however, this consequence is far from conclusive and further works with other datasets will be required to corroborate such findings. Also, more rigorous methodological approaches need to be employed to draw more specific policy implications.

While this study focused on the specific types of natural environments that would be impactful on total household health expenditure, one may be more interested in the category of medical treatments (e.g. internal medicine, cardiac surgery and mental counselling) that could be affected by nearby nature. With regard to freshwater environments, most of the previous public health studies focused on mental health rather than physical conditions (De Bell et al., 2017; McDougall et al., 2021; Pearson et al., 2019; Völker et al., 2018). Although it was difficult in my case, it will be possible to break down health spending into several healthcare categories and examine the impacts of nature separately, if a holistic database has been already developed.

The effects of more intentional nature exposure will be also worth investigating. It has been well known that nature exposure would have positive health impacts as noted earlier (Gascon et al., 2017; Lee et al., 2017; Rajoo et al., 2020; Völker et al., 2018; Wen et al., 2019; White et al., 2013b), however, its consequences on health expenditure have been unexamined up until now. A primary step that could be taken for revealing such an association would be to explore correlational relationships between visitation frequency and individual health spending. Subsequently, one can take an experimental approach such as a randomized controlled trial to obtain more robust evidence of nature's contributions to our health expenditure. Fortunately, taking different groups to alternative natural environments and observing subsequent health expenditure would raise fewer conceptual and technical issues than relocating some households.

As noted, reducing healthcare costs is an important agenda in public health administrations in Japan where rapid aging of society poses a serious problem of incrementing healthcare costs on the national accounting. In the meantime, ecosystem restorations with specific focus on freshwater and wetlands have been promoted for decades in this country mainly for environmental concerns. This study will be an important first step to link these two policy agendas and re-evaluate ecosystem restorations from human health perspectives. Once more causal association is revealed in more rigorous ways, we could argue the necessity of freshwater ecosystem restoration more strongly that could have triple dividends to local biodiversity, our health expenditure and national accounting.

## Appendix A: Theoretical Framework

Theoretical framework of the analysis stands on the traditional economic theory that assumes each person would maximize his/her utility given a certain health condition and some investments in the stock of health (Grossman, 1972). Let each household have the following utility function:

$$U = U(H, Z) \tag{A1}$$

where $H$ is health and $Z$ is the aggregate of all commodities besides health. Following Angulo et al. (2011), it is assumed that a household can be affected by an illness or injury (represented by a random variable $h$) with a probability of $\Pi(h \mid A, E)$, conditioned by household attributes A and surrounding environments $E$, which can be considered exogenous. Once a family member gets sick or injured, a household makes a decision whether to take medical treatments $M$ to restore its healthy state. On these theoretical considerations, $H$ is now described as:

$$H = H(M \mid h, A, E)$$

Then, Eq. (A1) can be rewritten as:

$$U = U(Z, H \mid A, E, M) \tag{A2}$$

Each household has a budgetary restriction given its income $R$ as:

$$P_z' Z + P_M M = R \tag{A3}$$

where $P_z$ is the vector of prices of the goods and the services except healthcare and $P_M$ is the price of medical treatments $M$. In essence, each household is assumed to solve the utility maximization problem of Eq. (A2) under the budgetary condition of Eq. (A3).

The reduced form of household healthcare demand can be expressed by the following function:

$$M = f(P_z, P_M, h, A, E, R)$$

In general, a medical doctor decides which treatments and prescriptions should be applied to a patient, and thus $P_M$ is basically unknown to a household until healthcare services are actually provided. Also, the variability of $P_z$ is considered irrelevant under the setting of cross-sectional data. Therefore, these variables are omitted from the final specification of the model.

# Appendix B: Alternative Specifications

Estimation results of alternative model specifications are provided in this Appendix.

**Table B1:** Estimation results of 1PM-OLS

| | 1PM-OLS I $r$ = 0.5 km | | 1PM-OLS II $r$ = 1.0 km | | 1PM-OLS III $r$ = 1.5 km | | 1PM-OLS IV $r$ = 2.0 km | |
|---|---|---|---|---|---|---|---|---|
| Variable | Coef | S.E. | Coef | S.E. | Coef | S.E. | Coef | S.E. |
| Family | 0.238*** | 0.050 | 0.245*** | 0.049 | 0.242*** | 0.050 | 0.241*** | 0.052 |
| Elder | 0.009*** | 0.002 | 0.009*** | 0.002 | 0.009*** | 0.002 | 0.009*** | 0.002 |
| Male | −0.003 | 0.002 | −0.003 | 0.002 | −0.003 | 0.002 | −0.003 | 0.002 |
| Chronic | 0.014*** | 0.002 | 0.014*** | 0.002 | 0.014*** | 0.002 | 0.014*** | 0.002 |
| Income | 0.021** | 0.010 | 0.017* | 0.010 | 0.017* | 0.010 | 0.018* | 0.010 |
| Deciduous | 0.013 | 0.011 | 0.014 | 0.012 | 0.016 | 0.012 | 0.012 | 0.012 |
| Evergreen | 0.004 | 0.008 | −0.002 | 0.008 | −0.003 | 0.009 | 0.000 | 0.009 |
| Freshwater | −0.014 | 0.015 | −0.028** | 0.013 | −0.030* | 0.016 | −0.028 | 0.019 |
| Saltwater | 0.015 | 0.009 | 0.009 | 0.007 | 0.006 | 0.007 | 0.003 | 0.008 |
| Urban | 0.005 | 0.004 | 0.003 | 0.005 | 0.002 | 0.006 | 0.001 | 0.006 |
| Constant | 1.162*** | 0.428 | 1.302** | 0.511 | 1.341** | 0.515 | 1.460*** | 0.520 |
| $R^2$ | 0.299 | | 0.303 | | 0.302 | | 0.298 | |
| Mean VIF | 1.69 | | 1.72 | | 1.74 | | 1.78 | |

*Note:* The number of observations is 459 (zero expenditures were omitted due to log-transformation). ***, **, and * denote statistical significance at 1%, 5%, and 10% levels, respectively. Prefectural dummies are skipped to preserve the space.

**Table B2:** Estimation results of 1PM-GLM

| | 1PM-GLM I $r$ = 0.5 km | | 1PM-GLM II $r$ = 1.0 km | | 1PM-GLM III $r$ = 1.5 km | | 1PM-GLM IV $r$ = 2.0 km | |
|---|---|---|---|---|---|---|---|---|
| Variable | Coef | S.E. | Coef | S.E. | Coef | S.E. | Coef | S.E. |
| Family | 0.261*** | 0.041 | 0.261*** | 0.042 | 0.255*** | 0.044 | 0.252*** | 0.045 |
| Elder | 0.007*** | 0.002 | 0.007*** | 0.002 | 0.008*** | 0.002 | 0.007*** | 0.002 |
| Male | −0.002 | 0.002 | −0.002 | 0.002 | −0.002 | 0.002 | −0.002 | 0.002 |
| Chronic | 0.013*** | 0.002 | 0.013*** | 0.002 | 0.012*** | 0.002 | 0.013*** | 0.002 |
| Income | 0.023** | 0.011 | 0.022* | 0.012 | 0.022* | 0.012 | 0.022* | 0.012 |
| Deciduous | 0.014 | 0.010 | 0.017* | 0.009 | 0.019* | 0.010 | 0.016 | 0.010 |
| Evergreen | 0.001 | 0.008 | −0.009 | 0.008 | −0.011 | 0.008 | −0.007 | 0.008 |
| Freshwater | −0.021* | 0.011 | −0.037*** | 0.009 | −0.040*** | 0.014 | −0.035** | 0.015 |
| Saltwater | 0.009 | 0.008 | 0.003 | 0.007 | 0.001 | 0.007 | 0.000 | 0.007 |
| Urban | 0.003 | 0.003 | 0.000 | 0.005 | −0.001 | 0.005 | −0.001 | 0.005 |
| Constant | 1.815*** | 0.358 | 1.998*** | 0.428 | 2.026*** | 0.433 | 2.058*** | 0.403 |
| Log-likelihood | -2072.0 | | -2068.9 | | -2068.8 | | -2070.6 | |
| AIC | 4166.0 | | 4159.9 | | 4159.6 | | 4163.2 | |
| BIC | 4211.4 | | 4205.3 | | 4205.0 | | 4208.6 | |

*Note:* The number of observations is 459 (zero expenditures were omitted due to log-transformation). ***, **, and * denote statistical significance at 1%, 5%, and 10% levels, respectively. Prefectural dummies are skipped to preserve the space.

**Table B3:** Estimation results of 2PM-OLS

| | 2PM-OLS I $r$ = 0.5 km | | 2PM-OLS II $r$ = 1.0 km | | 2PM-OLS III $r$ = 1.5 km | | 2PM-OLS IV $r$ = 2.0 km | |
|---|---|---|---|---|---|---|---|---|
| Variable | Coef | S.E. | Coef | S.E. | Coef | S.E. | Coef | S.E. |
| First-part | | | | | | | | |
| Family | 0.173*** | 0.042 | 0.175*** | 0.042 | 0.177*** | 0.042 | 0.178*** | 0.043 |
| Elder | 0.003 | 0.003 | 0.003 | 0.003 | 0.003 | 0.003 | 0.003 | 0.003 |
| Male | −0.004* | 0.002 | −0.004* | 0.002 | −0.004* | 0.002 | −0.004* | 0.002 |
| Chronic | 0.023*** | 0.002 | 0.023*** | 0.002 | 0.023*** | 0.002 | 0.023*** | 0.002 |
| Income | 0.026 | 0.021 | 0.026 | 0.020 | 0.026 | 0.020 | 0.026 | 0.020 |
| Deciduous | −0.012* | 0.007 | −0.009 | 0.010 | −0.007 | 0.011 | −0.009 | 0.011 |
| Evergreen | 0.004 | 0.011 | 0.008 | 0.012 | 0.002 | 0.011 | 0.000 | 0.012 |
| Freshwater | 0.003 | 0.015 | 0.006 | 0.015 | 0.006 | 0.017 | 0.004 | 0.018 |
| Saltwater | 0.002 | 0.016 | 0.001 | 0.013 | −0.001 | 0.012 | −0.004 | 0.011 |
| Urban | −0.001 | 0.005 | 0.001 | 0.005 | 0.000 | 0.005 | 0.000 | 0.005 |
| Constant | −0.417 | 0.449 | −0.569 | 0.513 | −0.561 | 0.503 | −0.455 | 0.487 |
| Second-part | | | | | | | | |
| Family | 0.238*** | 0.050 | 0.245*** | 0.049 | 0.242*** | 0.050 | 0.241*** | 0.052 |
| Elder | 0.009*** | 0.002 | 0.009*** | 0.002 | 0.009*** | 0.002 | 0.009*** | 0.002 |
| Male | −0.003 | 0.002 | −0.003 | 0.002 | −0.003 | 0.002 | −0.002 | 0.002 |
| Chronic | 0.014*** | 0.002 | 0.014*** | 0.002 | 0.014*** | 0.002 | 0.014*** | 0.002 |
| Income | 0.021** | 0.010 | 0.017* | 0.010 | 0.017* | 0.010 | 0.018* | 0.010 |
| Deciduous | 0.013 | 0.011 | 0.014 | 0.012 | 0.016 | 0.012 | 0.012 | 0.012 |
| Evergreen | 0.004 | 0.008 | −0.002 | 0.008 | −0.003 | 0.009 | 0.000 | 0.009 |
| Freshwater | −0.014 | 0.015 | −0.028** | 0.013 | −0.030* | 0.016 | −0.028 | 0.019 |
| Saltwater | 0.015 | 0.009 | 0.009 | 0.007 | 0.006 | 0.007 | 0.003 | 0.008 |
| Urban | 0.005 | 0.004 | 0.003 | 0.005 | 0.002 | 0.006 | 0.001 | 0.006 |
| Constant | 1.162*** | 0.428 | 1.302** | 0.511 | 1.341*** | 0.515 | 1.460*** | 0.520 |
| Log-likelihood | -1007.0 | | -1005.9 | | -1006.3 | | -1007.4 | |
| AIC | 2054.0 | | 2051.8 | | 2052.6 | | 2054.9 | |
| BIC | 2142.4 | | 2140.2 | | 2141.0 | | 2143.3 | |

*Note:* The number of observations is 614 (several prefecture dummies perfectly predicted positive outcomes in the first-stage probit regression and were therefore omitted). ***, **, and * denote statistical significance at 1%, 5%, and 10% levels, respectively. Prefectural dummies are skipped to preserve the space.

**Table B4:** Estimation results of SSM-OLS

| | SSM-OLS I $r$ = 0.5 km | | SSM-OLS II $r$ = 1.0 km | | SSM-OLS III $r$ = 1.5 km | | SSM-OLS IV $r$ = 2.0 km | |
|---|---|---|---|---|---|---|---|---|
| Variable | Coef | S.E. | Coef | S.E. | Coef | S.E. | Coef | S.E. |
| First-part | | | | | | | | |
| Family | 0.173*** | 0.043 | 0.175*** | 0.042 | 0.177*** | 0.043 | 0.178*** | 0.043 |
| Elder | 0.003 | 0.003 | 0.003 | 0.003 | 0.003 | 0.003 | 0.003 | 0.003 |
| Male | −0.004* | 0.002 | −0.004* | 0.002 | −0.004* | 0.002 | −0.004* | 0.002 |
| Chronic | 0.023*** | 0.002 | 0.023*** | 0.002 | 0.023*** | 0.002 | 0.023*** | 0.002 |
| Income | 0.026 | 0.021 | 0.026 | 0.020 | 0.026 | 0.020 | 0.026 | 0.020 |
| Deciduous | −0.012* | 0.007 | −0.009 | 0.010 | −0.007 | 0.010 | −0.009 | 0.011 |
| Evergreen | 0.004 | 0.012 | 0.008 | 0.012 | 0.002 | 0.011 | 0.000 | 0.012 |
| Freshwater | 0.003 | 0.015 | 0.006 | 0.015 | 0.006 | 0.017 | 0.004 | 0.018 |
| Saltwater | 0.002 | 0.016 | 0.001 | 0.013 | −0.001 | 0.012 | −0.004 | 0.011 |
| Urban | −0.001 | 0.005 | 0.001 | 0.005 | 0.000 | 0.005 | 0.000 | 0.005 |
| Constant | −0.417 | 0.445 | −0.571 | 0.507 | −0.563 | 0.495 | −0.454 | 0.479 |
| Second-part | | | | | | | | |
| Family | 0.238*** | 0.054 | 0.243*** | 0.052 | 0.241*** | 0.052 | 0.242*** | 0.053 |
| Elder | 0.009*** | 0.002 | 0.009*** | 0.002 | 0.009*** | 0.002 | 0.009*** | 0.002 |
| Male | −0.003 | 0.002 | −0.003 | 0.002 | −0.003 | 0.002 | −0.002 | 0.002 |
| Chronic | 0.014*** | 0.002 | 0.014*** | 0.002 | 0.014*** | 0.002 | 0.014*** | 0.002 |
| Income | 0.021** | 0.010 | 0.017* | 0.009 | 0.017* | 0.009 | 0.018* | 0.010 |
| Deciduous | 0.013 | 0.011 | 0.015 | 0.011 | 0.016 | 0.012 | 0.012 | 0.012 |
| Evergreen | 0.004 | 0.007 | −0.003 | 0.008 | −0.003 | 0.008 | 0.000 | 0.009 |
| Freshwater | −0.014 | 0.014 | −0.028** | 0.012 | −0.031** | 0.015 | −0.028 | 0.017 |
| Saltwater | 0.015* | 0.008 | 0.009 | 0.007 | 0.006 | 0.007 | 0.003 | 0.007 |
| Urban | 0.005 | 0.004 | 0.003 | 0.005 | 0.002 | 0.005 | 0.001 | 0.006 |
| Constant | 1.162*** | 0.440 | 1.326*** | 0.509 | 1.354*** | 0.517 | 1.449*** | 0.516 |
| Log-likelihood | -1007.0 | | -1005.9 | | -1006.3 | | -1007.4 | |
| AIC | 2100.0 | | 2081.8 | | 2082.6 | | 2088.9 | |
| BIC | 2293.1 | | 2238.9 | | 2239.8 | | 2255.0 | |

*Note:* The number of observations is 659. ***, **, and * denote statistical significance at 1%, 5%, and 10% levels, respectively. Prefectural dummies are skipped to preserve the space.

## References

Aerts, R., Honnay, O., and Van Nieuwenhuyse, A. (2018). Biodiversity and human health: Mechanisms and evidence of the positive health effects of diversity in nature and green spaces. *British Medical Bulletin*, 127(1):5–22.

Angulo, A. M., Barberán, R., Egea, P., and Mur, J. (2011). An analysis of health expenditure on a microdata population basis. *Economic Modelling*, 28(1-2):169–180.

Basu, A., Arondekar, B. V., and Rathouz, P. J. (2006). Scale of interest versus scale of estimation: Comparing alternative estimators for the incremental costs of a comorbidity. *Health Economics*, 15(10):1091–1107.

Bauman, A., Smith, B., Stoker, L., Bellew, B., and Booth, M. (1999). Geographical influences upon physical activity participation: Evidence of a coastal effect. *Australian and New Zealand Journal of Public Health*, 23(3):322–324.

Becker, D. A. and Browning, M. H. (2021). Total area greenness is associated with lower per-capita medicare spending, but blue spaces are not. *City and Environment Interactions*, 11:100063.

Becker, D. A., Browning, M. H., Kuo, M., and Van Den Eeden, S. K. (2019). Is green land cover associated with less health care spending? Promising findings from county-level Medicare spending in the continental United States. *Urban Forestry and Urban Greening*, 41:39–47.

Buntin, M. B. and Zaslavsky, A. M. (2004). Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures. *Journal of Health Economics*, 23(3):525–542.

Chaze, J. P. (2005). Assessing household health expenditure with box-cox censoring models. *Health Economics*, 14(9):893–907.

De Bell, S., Graham, H., Jarvis, S., and White, P. (2017). The importance of nature in mediating social and psychological benefits associated with visits to freshwater blue space. *Landscape and Urban Planning*, 167:118–127.

de Vries, S., Ten Have, M., van Dorsselaer, S., van Wezep, M., Hermans, T., and de Graaf, R. (2016). Local availability of green and blue space and prevalence of common mental disorders in the Netherlands. *BJPsych Open*, 2(6):366–372.

Deb, P., Munkin, M. K., and Trivedi, P. K. (2006). Bayesian analysis of the two-part model with endogeneity: Application to health care expenditure. *Journal of Applied Econometrics*, 21(7):1081–1099.

Deb, P. and Trivedi, P. K. (2002). The structure of demand for health care: Latent class versus two-part models. *Journal of Health Economics*, 21(4):601–625.

Dempsey, S., Devine, M. T., Gillespie, T., Lyons, S., and Nolan, A. (2018). Coastal blue space and depression in older adults. *Health and Place*, 54:110–117.

Garrett, J. K., White, M. P., Huang, J., Ng, S., Hui, Z., Leung, C., Tse, L. A., Fung, F., Elliott, L. R., Depledge, M. H., and Wong, M. C. S. (2019). Urban blue space and health and wellbeing in Hong Kong: Results from a survey of older adults. *Health and Place*, 55:100–110.

Gascon, M., Triguero-Mas, M., Martínez, D., Dadvand, P., Forns, J., Plasència, A., and Nieuwenhuijsen, M. J. (2015). Mental health benefits of long-term exposure to residential green and blue spaces: A systematic review. *International Journal of Environmental Research and Public Health*, 12(4):4354–4379.

Gascon, M., Triguero-Mas, M., Martínez, D., Dadvand, P., Rojas-Rueda, D., Plasència, A., and Nieuwenhuijsen, M. J. (2016). Residential green spaces and mortality: A systematic review. *Environment International*, 86:60–67.

Gascon, M., Zijlema, W., Vert, C., White, M. P., and Nieuwenhuijsen, M. J. (2017). Outdoor blue spaces, human health and well-being: A systematic review of quantitative studies. *International Journal of Hygiene and Environmental Health*, 220(8):1207–1221.

Grassetti, L. and Rizzi, L. (2019). The determinants of individual health care expenditures in the Italian region of Friuli Venezia Giulia: Evidence from a hierarchical spatial model estimation. *Empirical Economics*, 56(3):987–1009.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80:223–255.

Hill, S. C. and Miller, G. E. (2010). Health expenditure estimation and functional form: Applications of the generalized gamma and extended estimating equations models. *Health Economics*, 19(5):608–627.

Hooyberg, A., Roose, H., Grellier, J., Elliott, L. R., Lonneville, B., White, M. P., Michels, N., De Henauw, S., Vandegehuchte, M., and Everaert, G. (2020). General health and residential proximity to the coast in Belgium: Results from a cross-sectional health survey. *Environmental Research*, 184:109225.

James, P., Banay, R. F., Hart, J. E., and Laden, F. (2015). A review of the health benefits of greenness. *Current Epidemiology Reports*, 2(2):131–142.

Jones, A. M. (2000). Health econometrics. In Culyer, A. J. and Newhouse, J. P., editors, *Handbook of Health Economics*, volume 1, part A, pages 265–344. Elsevier.

Jones, A. M., Lomas, J., Moore, P. T., and Rice, N. (2016). A quasi-Monte-Carlo comparison of parametric and semiparametric regression methods for heavy-tailed and non-normal data: An application to healthcare costs. *Journal of the Royal Statistical Society Series A*, 179(4):951–974.

Kabaya, K. (2021). Empirical analysis of associations between health expenditure and forest environments: A case of Japan. *Ecological Economics*, 181:106927.

Kondo, M. C., Fluehr, J. M., McKeon, T., and Branas, C. C. (2018). Urban green space and its impact on human health. *International Journal of Environmental Research and Public Health*, 15(3):445.

Lachowycz, K. and Jones, A. P. (2011). Greenspace and obesity: A systematic review of the evidence. *Obesity Reviews*, 12(5):e183–e189.

Lee, A. C. and Maheswaran, R. (2011). The health benefits of urban green spaces: A review of the evidence. *Journal of Public Health*, 33(2):212–222.

Lee, I., Choi, H., Bang, K. S., Kim, S., Song, M., and Lee, B. (2017). Effects of forest therapy on depressive symptoms among adults: A systematic review. *International Journal of Environmental Research and Public Health*, 14(3):321.

López-Nicolás, A. (1998). Unobserved heterogeneity and censoring in the demand for health care. *Health Economics*, 7(5):429–437.

Madden, D. (2008). Sample selection versus two-part models revisited: The case of female smoking and drinking. *Journal of Health Economics*, 27(2):300–307.

Manning, W. G., Basu, A., and Mullahy, J. (2005). Generalized modeling approaches to risk adjustment of skewed outcomes data. *Journal of Health Economics*, 24(3):465–488.

Manning, W. G. and Mullahy, J. (2001). Estimating log models: To transform or not to transform? *Journal of Health Economics*, 20(4):461–494.

Matsaganis, M., Mitrakos, T., and Tsakloglou, P. (2009). Modelling health expenditure at the household level in Greece. *The European Journal of Health Economics*, 10:329–336.

McDougall, C. W., Hanley, N., Quilliam, R. S., Bartie, P. J., Robertson, T., Griffiths, M., and Oliver, D. M. (2021). Neighbourhood blue space and mental health: A nationwide ecological study of antidepressant medication prescribed to older adults. *Landscape and Urban Planning*, 214:104132.

McDougall, C. W., Hanley, N., Quilliam, R. S., and Oliver, D. M. (2022). Blue space exposure, health and well-being: Does freshwater type matter? *Landscape and Urban Planning*, 224:104446.

McDougall, C. W., Quilliam, R. S., Hanley, N., and Oliver, D. M. (2020). Freshwater blue space and population health: An emerging research agenda. *Science of The Total Environment*, 737:140196.

Mihaylova, B., Briggs, A., O'Hagan, A., and Thompson, S. G. (2011). Review of statistical methods for analysing healthcare resources and costs. *Health Economics*, 20(8):897–916.

Nutsford, D., Pearson, A. L., Kingham, S., and Reitsma, F. (2016). Residential exposure to visible blue space (but not green space) associated with lower psychological distress in a capital city. *Health and Place*, 39:70–78.

Okunade, A. A., Suraratdecha, C., and Benson, D. A. (2010). Determinants of Thailand household health expenditure: The relevance of permanent resources and other correlates. *Health Economics*, 19(3):365–376.

Pasanen, T. P., White, M. P., Wheeler, B. W., Garrett, J. K., and Elliott, L. R. (2019). Neighbourhood blue space, health and wellbeing: The mediating role of different types of physical activity. *Environment International*, 131:105016.

Pearson, A. L., Shortridge, A., Delamater, P. L., Horton, T. H., Dahlin, K., Rzotkiewicz, A., and Marchiori, M. J. (2019). Effects of freshwater blue spaces may be beneficial for mental health: A first, ecological study in the North American Great Lakes region. *PLoS ONE*, 14(8):e0221977.

Rajoo, K. S., Karam, D. S., and Abdullah, M. Z. (2020). The physiological and psychosocial effects of forest therapy: A systematic review. *Urban Forestry and Urban Greening*, 54:126744.

Rous, J. J. and Hotchkiss, D. R. (2003). Estimation of the determinants of household health expenditures in Nepal with controls for endogenous illness and provider choice. *Health Economics*, 12(6):431–451.

Rout, H. S. (2008). Socio-economic factors and household health expenditure: The case of Orissa. *Journal of Health Management*, 10(1):101–118.

Sanwald, A. and Theurl, E. (2017). Out-of-pocket expenditures for pharmaceuticals: Lessons from the Austrian household budget survey. *The European Journal of Health Economics*, 18:435–447.

van den Berg, M., Wendel-Vos, W., Van Poppel, M., Kemper, H., Van Mechelen, W., and Maas, J. (2015). Health benefits of green spaces in the living environment: A systematic review of epidemiological studies. *Urban Forestry and Urban Greening*, 14(4):806–816.

Van Den Eeden, S. K., Browning, M. H., Becker, D. A., Shan, J., Alexeeff, S. E., Ray, G. T., Quesenberry, C. P., and Kuo, M. (2022). Association between residential green cover and direct healthcare costs in Northern California: An individual level analysis of 5 million persons. *Environment International*, 163:107174.

Völker, S., Heiler, A., Pollmann, T., Claßen, T., Hornberg, C., and Kistemann, T. (2018). Do perceived walking distance to and use of urban blue spaces affect self-reported physical and mental health? *Urban Forestry and Urban Greening*, 29:1–9.

Wen, Y., Yan, Q., Pan, Y., Gu, X., and Liu, Y. (2019). Medical empirical research on forest bathing (Shinrin-yoku): A systematic review. *Environmental Health and Preventive Medicine*, 24(1):1–21.

Werblow, A., Felder, S., and Zweifel, P. (2007). Population ageing and health care expenditure: A school of 'red herrings'? *Health Economics*, 16(10):1109–1126.

Western, B. and Bloome, D. (2009). Variance function regressions for studying inequality. *Sociological Methodology*, 39(1):293–326.

Wheeler, B. W., White, M., Stahl-Timmins, W., and Depledge, M. H. (2012). Does living by the coast improve health and wellbeing? *Health and Place*, 18:1198–1201.

White, M. P., Alcock, I., Wheeler, B. W., and Depledge, M. H. (2013a). Coastal proximity, health and well-being: Results from a longitudinal panel survey. *Health and Place*, 23:97–103.

White, M. P., Elliott, L. R., Gascon, M., Roberts, B., and Fleming, L. E. (2020). Blue space, health and well-being: A narrative overview and synthesis of potential benefits. *Environmental Research*, 191:110169.

White, M. P., Pahl, S., Ashbullby, K., Herbert, S., and Depledge, M. H. (2013b). Feelings of restoration from recent nature visits. *Journal of Environmental Psychology*, 35:40–51.

Wong, A., van Baal, P. H., Boshuizen, H. C., and Polder, J. J. (2011). Exploring the influence of proximity to death on disease-specific hospital expenditures: A carpaccio of red herrings. *Health Economics*, 20(4):379–400.