
Competing Demographic Drivers of Hospital Expenditures: Coexistence of the Red Herring and the Steepening Effects

Malene Kallestrup-Lamb, Alexander O.K. Marin, Jes Søgaard

Competing Demographic Drivers of Hospital Expenditures: Coexistence of the Red Herring and the Steepening Effects

Published by:

© The ROCKWOOL Foundation Research Unit

Address:

The ROCKWOOL Foundation Research Unit

Ny Kongensgade 6

1472 Copenhagen, Denmark

Telephone +45 33 34 48 00

E-mail: kontakt@rff.dk

en.rockwoolfonden.dk/research/

June 2026

Competing Demographic Drivers of Hospital Expenditures: Coexistence of the Red Herring and the Steepening Effects*

Malene Kallestrup-Lamb

Aarhus University and PeRCent

Alexander O.K. Marin[†]

University of Southern Denmark, CPop

Jes Søgaard

University of Southern Denmark, CPop

November 21, 2025

Abstract

The fiscal sustainability of healthcare systems is increasingly strained by aging populations with two competing hypothesis dominating the literature. The Red Herring hypothesis suggests that healthcare expenditures are driven more by proximity to death than by chronological age, while the Steepening Hypothesis examines whether expenditures increase faster for older individuals over time. Nevertheless, these two frameworks have traditionally been studied independently rather than in conjunction. This paper offers a unified econometric model, allowing for rigorous assessment of their concurrent validity and interaction. Using comprehensive register-based monthly somatic hospital expenditure data for the entire Danish population from 2002 to 2017, we provide robust evidence that both hypotheses coexist. Although Red Herring effects modestly mitigate the expenditure burden of an increasingly older population, they are substantially outweighed by pronounced Steepening effects, which account for nearly 60% of hospital expenditure growth during the observation period. Through a novel decomposition method, we identify a previously unrecognized interaction between these phenomena, a *Red Herring Steepening effect*, which accelerates expenditure growth in the final years of life for older individuals. Our findings suggest that healthcare systems face considerably greater fiscal challenges from population aging than previously recognized under the Red Herring Hypothesis alone.

Keywords: Red Herring | Steepening | Hospital Expenditures | Population Aging | Expenditure Decomposition

*Authors listed alphabetically.

[†]Corresponding Author. Campusvej 55, 5230 Odense M, Denmark. Email: aokm@sam.sdu.dk

JEL code: H51, I10, I18

Acknowledgment: This paper was improved by the feedback from participants at the CPop 2022 seminar series, NHESG 2023 conference, ROCKWOOL 2023 seminar series, VIVE 2023 seminar series, EuHEA conference 2024, Nordic Demographic Symposium 2025, Ilya Kashnitsky, and Tom Engsted.

Conflict of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding: Funding for this research was provided by the ROCKWOOL Foundation through the project “Challenges of Aging Populations: Physical Functioning, Cognitive Functioning, and Health Economic Consequences.”

Declaration of Generative AI and AI-assisted Technologies in the Writing Process: During the preparation of this work, the authors used generative AI in order to improve language. After using this tool, the authors reviewed and edited the content extensively and take full responsibility for the content of the publication.

1 Introduction

Healthcare expenditures constitute a significant portion of public and private spending in high-income countries. In 2019, healthcare accounted for 16.8% of the U.S. gross domestic product (GDP), compared to 10.2% in the Euro area (WHO, 2024). A major challenge facing healthcare systems is the increasing number of older individuals (Christensen et al., 2009). As the large baby boomer generation enters old age, the demand for healthcare services grows (Ricketts, 2011), and concerns over escalating healthcare expenditures intensify.

Two major hypotheses have emerged in the literature, to address the relationship between the population age structure and healthcare expenditures. The Red Herring Hypothesis posits that time-to-death, rather than age, is the primary determinant of healthcare expenditures (Zweifel et al., 1999), a claim repeatedly supported in the context of hospital expenditures (Seshamani and Gray, 2004a; Kallestrup-Lamb et al., 2024). Empirical analyses show that accounting for time-to-death substantially reduces projected hospital expenditures relative to a Naïve scenario that excludes it (Polder et al., 2006; Bjørner and Arnberg, 2012; Geue et al., 2014). Thus, simply having an older population does not automatically mean proportionally higher expenditures, although the impact of population aging on healthcare expenditures can be more nuanced (Breyer and Lorenz, 2021). The Steepening Hypothesis argues that the age-expenditure profile becomes steeper over time (Buchner and Wasem, 2006), meaning that healthcare expenditures for older individuals rise more than those for younger individuals, typically considered in absolute terms. The literature remains divided over the existence of Steepening effects. Felder and Werblow (2008) reject the hypothesis in the context of Switzerland, whereas more recent empirical evidence supports the presence of Steepening effects for hospital expenditures in Norway (Gregersen, 2014), and for hospital expenditures plus physician care in Denmark (Kollerup et al., 2022). Unlike the Red Herring Hypothesis, the Steepening Hypothesis presents a more concerning outlook, suggesting that an increasing number of older individuals will drive significant increases in healthcare expenditures due to expenditure growth at older ages.

Despite their contrasting implications, the two hypotheses have not been formally tested within a unified statistical framework. This paper presents the first nested investigation of the Red Herring and Steepening Hypotheses using detailed register data for the entire Danish population aged 30 and above, focusing on monthly somatic hospital expenditures from 2002 to 2017. We introduce a novel parametrization that allows for rigorous hypothesis testing, filling a notable gap in the literature, which has seen limited application of formal hypothesis testing. We propose employing the Delta Method (Dorfman, 1938) to estimate marginal nominal effects and use the Wald (1943) test to assess whether the marginal expenditure effects support the two hypotheses. Additionally, we introduce a decomposition approach that is new to this literature,

demonstrating how changes in hospital expenditures can be attributed to each hypothesis.

Empirically, we present robust evidence supporting the coexistence of the effects of both hypotheses, with their relative influence varying by age and time-to-death. We perform sex-disaggregated analyses across subsamples of the population, offering a more flexible framework for rigorous analysis and hypothesis testing within a demographic subgroup. Results uncover how age, time-to-death, and Steepening effects significantly differ by sex, age, and time-to-death, confirmed through Welch tests. Specifically, while both the Red Herring and Steepening effects are present across all groups, time-to-death effects increase in age. Steepening is particularly strong for ages 50 to 80, while the age-specific growth in somatic hospital expenditures among the 80+ year old is lower. Robustness tests including a Bayesian hypothesis testing approach (Kass and Raftery, 1995) and an expanded model parametrization further validate these findings. Applying our decomposition framework, we find that the Steepening related effect is the primary driver of hospital expenditure growth, accounting for 60% of the increase, while the Red Herring effect mitigate only 19% of the expenditure growth. Additionally, we identify a novel interaction between the two hypotheses, demonstrating that the mitigating effect posited by the Red Herring Hypothesis can exhibit an age-specific growth pattern, characteristic of the Steepening Hypothesis. This interaction term accounts for 19% of the change in hospital expenditures and precisely offsets the Red Herring mitigation in the decomposition.

The remainder of this paper is structured as follows: Section 2 outlines the model, estimation approach, hypothesis tests, and decomposition method. Section 3 describes the dataset, while Section 4 presents the empirical findings. Section 6 discusses the implications of our results, and Section 7 concludes.

2 Methodology

To analyze the relationship between hospital expenditures and the demographic drivers related to population aging, we employ the widely used two-part model (Cragg, 1971; Seshamani and Gray, 2004a), a framework predominantly used in existing literature (Kallestrup-Lamb et al., 2024). This model is particularly well-suited for capturing the distinctive distribution of hospital expenditures, which is characterized by a large number of zero observations and a heavy right tail for non-zero expenditures; see Deb and Norton (2018) for a comprehensive overview of this approach. The two-part model is defined as follows, capturing the dual nature of hospital expenditures in the following probability density function:

$$\psi(H_i) = F(X_i'\delta) g(X_i'\gamma). \quad (1)$$

The first component captures the probability that an individual incurs positive hospital expenditures, $H_i > 0$, conditional on $(k \times 1)$ explanatory variables, X_i , and their associated parameters, δ . Specifically, this probability is expressed as $Pr(H_i > 0 | X_i) = F(X_i' \delta)$, where F is a cumulative density function. Modeling this probability is crucial due to the prevalence of zero-observations in individual-level hospital expenditure data. We refer to changes in the probability of incurring hospital expenditures as the *extensive margin*. The second component models the distribution of positive hospital expenditures, $H_i | X_i, H_i > 0$, as $g(X_i' \gamma)$, where g is a density function and γ a k -dimensional parameter vector. This density function is typically specified to account for the heavy right-tail characteristic of hospital expenditure. This component provides insights into the *intensive margin*, describing the extent of resource utilization once individuals enter the hospital system. When the realized value from the first part is zero, individual i incurs no hospital expenditures. Conversely, if the realized value is one, hospital expenditures are drawn randomly from the second part, $g(\cdot)$, yielding a positive amount.

To estimate the parameters of our model in equation (1) we use quasi maximum likelihood estimation, following existing literature (Atella and Conti, 2014; Moore et al., 2017).¹ The first part of the model utilizes a Probit specification, similar to that used by Seshamani and Gray (2004b), Atella and Conti (2014), and Moore et al. (2017). The second part, $g(\cdot)$, is modeled using a Poisson-distributed generalized linear model with a log-link, as suggested by (Carreras et al., 2018; von Wyl, 2019). This two-part model is further supported by Kollerup et al. (2022), who identify the Poisson-distributed generalized linear model with a log-link as the most appropriate for analyzing Danish hospital expenditure data using the modified Park's test (Manning and Mullahy, 2001). Additional details on the estimation of two-part models can be found in Belotti et al. (2015).

The estimated parameters, δ and γ , enter the model through nonlinear Probit and Poisson functions, respectively. Because of this nonlinearity, these estimates lack a straightforward interpretation and should not be treated as marginal effects on either the probability of using hospital services or the level of expenditures. However, understanding marginal effects is essential for evaluating the Red Herring and Steepening hypotheses, which focus on the direction and magnitude of partial derivatives and not on the coefficients themselves. For example, the Red Herring Hypothesis states that hospital expenditures increase substantially as time-to-death shortens, while the Steepening Hypothesis posits a positive second derivative of expenditures with respect to age and time. These effects cannot be directly assessed from parameter esti-

¹We employ quasi-maximum likelihood estimation with robust standard errors, which provides consistent parameter estimates and valid inference even under distributional misspecification, requiring only correct specification of the conditional mean function. Hence, the choice of the distribution of the second part is of lesser importance (White, 1982; Gourieroux et al., 1984). We conduct robustness estimates against the conditional mean function specification.

mates, fitted values, or conditional means, which do not convey the uncertainty needed for hypothesis testing. To address this, we apply the Delta Method (Dorfman, 1938) to estimate unbiased marginal effects for key variables such as age, time-to-death, and their interaction, complete with asymptotic properties. Although fully parameterized models are often evaluated via predicted outcomes—as in Seshamani and Gray (2004b) and Kollerup et al. (2022)—such predictions should not be mistaken for marginal effects. Our use of the Delta Method represents a methodological advance in testing the Red Herring and Steepening hypotheses by enabling formal inference. Specifically, we report the sample average of individual marginal effects, rather than marginal effects at the sample mean.²

Conceptually, the two-part model in equation (1) aligns with the approach proposed by Heckman (1976), but it diverges significantly in its handling of zero hospital expenditures, treating them as actual zeros rather than censored observations. Moreover, Seshamani and Gray (2004a) underscore the problem of multicollinearity in the analysis of hospital expenditures arising from the inclusion of the inverse Mills ratio in the second stage of the Heckman (1976) model, advocating for the two-part model as a more robust alternative for analyzing hospital expenditures.

2.1 Linear Parametrization

This section provides a detailed examination of the linear arguments, $X_i'\delta$ and $X_i'\gamma$, in equation (1). Specifically, it explores how these terms capture the relationship between demographic drivers and hospital expenditures, highlighting key variations in expenditure dynamics across time and age. In research on hospital expenditures and the demographic composition of a population, individual age plays a crucial role, as expenditures typically increase with advancing age (Howdon and Rice, 2018; Maynou et al., 2023; Laudicella et al., 2022). The assumption that age, A , is the primary determinant of an individual's, i , hospital expenditures, is commonly referred to as the *Naïve* model (van Baal and Wong, 2012). In equation (2), age serves as the primary variable of interest, while additional control variables, \check{X} , are incorporated to account for other influencing factors

$$\begin{aligned} \text{Naïve:} \quad X_i'\delta^N &= A_i\delta_A^N + \check{X}_i'\delta_{\check{X}}^N \\ X_i'\gamma^N &= A_i\gamma_A^N + \check{X}_i'\gamma_{\check{X}}^N. \end{aligned} \tag{2}$$

In their seminal work, Zweifel et al. (1999) argue that once an individual's proximity to death

²Maynou et al. (2023) also use average marginal effects, as defined by Norton and Dowd (2018), to analyze hospital expenditures. However, their model does not explicitly include time-to-death or Steepening variables; instead, it controls for timing via month-of-death dummies.

is accounted for, age has a limited effect on hospital expenditures. Consequently, while age appears to be a highly significant predictor in the Naïve model, this significance primarily reflects omitted variable bias due to the exclusion of time-to-death, TTD , as a regressor. Equation (3) considers a parametrization of the Red Herring model by introducing the time-to-death regressor and associated effects to the Naïve model in equation (2).³ Following Moore et al. (2017), we include a binary indicator, \check{X}^* , which equals one if time-to-death is censored,⁴ indicating that the individual is still alive

$$\begin{aligned} \text{Red Herring:} \quad X_i' \delta^{RH} &= TTD_i \delta_{TTD}^{RH} + A_i \delta_A^{RH} + \check{X}_i' \delta_{\check{X}}^{RH} + \check{X}_i^* \delta_{\check{X}^*}^{RH} \\ X_i' \gamma^{RH} &= TTD_i \gamma_{TTD}^{RH} + A_i \gamma_A^{RH} + \check{X}_i' \gamma_{\check{X}}^{RH} + \check{X}_i^* \gamma_{\check{X}^*}^{RH}. \end{aligned} \quad (3)$$

Buchner and Wasem (2006) proposed a supplementary hypothesis on the relationship between hospital expenditures and the population age structure, known as the Steepening Hypothesis. As defined by Felder and Werblow (2008), "*Steepening ... means that the growth rate of healthcare expenditures rises with an increasing age*" which implies that the hospital expenditures of older individuals grow at a faster rate compared to younger individuals. In this paper, we consider growth in absolute terms following most analyses in existing literature (Buchner and Wasem, 2006; Felder and Werblow, 2008; Gregersen, 2014) and focuses on individual level expenditures like Kollerup et al. (2022).⁵ Gregersen (2014) proposes a parametrization of a hospital expenditure equation suitable for testing the existence of a Steepening effect. This approach includes an interaction between age and calendar time, $A \cdot T$, where T represents calendar time. If the coefficient estimate on the age-time interaction is positive, the age profile of hospital expenditures steepens over time for a linear model. Meanwhile, for nonlinear specifications of hospital expenditures, such as equation (1), Steepening effects can occur despite a null age-time interaction, which we provide a general theoretical example of in Appendix B.1. We address how to assess Steepening effects for such models in the following section. We formulate equation (4) which incorporates the Steepening Hypothesis, by extending the Naïve model in equation (2)

³We consider one of several types of Red Herring hypotheses, as discussed by, e.g., Breyer and Lorenz (2021), which is suitable for our dataset and research question.

⁴In practice, time-to-death is set to four and the binary indicator to one, if an individual has not died through our period of observation, described later in the data section.

⁵Analyzing the Steepening Hypothesis in relative terms (i.e., by comparing percentage growth across age groups) could provide complementary insights. A relative perspective helps to standardize across different baseline expenditure levels, identify age groups where costs are accelerating fastest, inform targeted health policy responses, and integrate into macroeconomic models considering growth rates in, e.g., GDP and other macroeconomic outputs. Meanwhile, accounting practices and healthcare budgets are typically considered in absolute terms.

with time and an age-time interaction term.⁶ Crucially, the Steepening Hypothesis disregards any influence of time-to-death as argued under the Red Herring Hypothesis.

$$\begin{aligned} \text{Steepening:} \quad X_i' \delta^S &= A_i \delta_A^S + T_i \delta_T^S + A_i T_i \delta_{A \cdot T}^S + \check{X}_i' \delta_{\check{X}}^S \\ X_i' \gamma^S &= A_i \gamma_A^S + T_i \gamma_T^S + A_i T_i \gamma_{A \cdot T}^S + \check{X}_i' \gamma_{\check{X}}^S. \end{aligned} \quad (4)$$

To develop a comprehensive framework that unifies the Red Herring and Steepening Hypotheses, we introduce a novel parameterization that seamlessly integrates both perspectives. This approach enables more precise estimation of how hospital expenditures evolve with individual aging, capturing both time-to-death effects and age-related expenditure dynamics within a single, cohesive model. While [Gregersen \(2014\)](#) demonstrates that the Red Herring and Steepening Hypotheses can both be valid, a formal model that integrates the two hypotheses has yet to be developed. In this paper, we address this gap by proposing a novel parameterization that enables such testing. In equation (5), we extend the Naïve model by incorporating the time-to-death effects from the Red Herring Hypothesis, along with calendar time and the age-time interaction from the Steepening Hypothesis.

$$\begin{aligned} \text{Joint Model:} \quad X_i' \delta &= TTD_i \delta_{TTD} + A_i \delta_A + T_i \delta_T + A_i T_i \delta_{A \cdot T} + \check{X}_i' \delta_{\check{X}} + \check{X}_i^* \delta_{\check{X}^*} \\ X_i' \gamma &= TTD_i \gamma_{TTD} + A_i \gamma_A + T_i \gamma_T + A_i T_i \gamma_{A \cdot T} + \check{X}_i' \gamma_{\check{X}} + \check{X}_i^* \gamma_{\check{X}^*}. \end{aligned} \quad (5)$$

Related studies, such as [von Wyl \(2019\)](#) and [Kollerup et al. \(2022\)](#), run regressions including Steepening effects alongside indicators of being in the last year of life.⁷ [von Wyl \(2019\)](#) also includes an indicator of whether an individual dies in the next calendar year. Although these variables serve as crude proxies for time-to-death, our parametrization in equation (5) is the first expression that allows for a rigorous statistical test of the joint validity of both hypotheses.

Our linear specifications in equations (2) to (5) are parsimonious, focusing solely on age, time-to-death, time, and the age-time interaction, capturing the effects in a few key parameters which are easy to interpret. In contrast, existing literature often applies an expanded parametrization

⁶Our specification of the variables related to the Steepening Hypothesis follows the most comprehensive discussion of the coexistent Red Herring and Steepening Hypotheses by [Gregersen \(2014\)](#), where calendar age and time are considered related to Steepening while age and time-to-death are the central factors for the Red Herring Hypothesis. We, however, acknowledge that many empirical analyses of the Red Herring Hypothesis and Naïve models control for calendar time in their regressions, e.g., [Zweifel et al. \(1999\)](#) and [Seshamani and Gray \(2004a\)](#). Yet, instances of the Naïve model and Red Herring models without changes in hospital expenditures over time, i.e., zero growth rates, have also been considered by, e.g., [van Baal and Wong \(2012\)](#) and for projections of healthcare expenditures ([Polder et al., 2006](#); [Bjørner and Arnberg, 2012](#); [Geue et al., 2014](#)).

⁷The indicator accounts for the fact that people in their final year have fewer than 12 full months to consume hospital, which reduces the total expenditures in that period.

seeking to capture age, sex, and time-to-death patterns in a single regression (Atella and Conti, 2014; von Wyl, 2019; Kollerup et al., 2022) which can be challenging to analyze and interpret due to the high dimensionality of its inputs. To explore heterogeneous patterns across subgroups, our sparse parametrization enables analyses of subgroups of individuals across age, sex, and time-to-death, where a parsimonious model likely offer a better fit. Meanwhile, we also conduct a robustness check with and expanded parametrization as discussed in Appendix C.5. Regressors such as health and income have also considered (Carreras et al., 2018; Werblow et al., 2007; Kollerup et al., 2022; Gregersen, 2014), yet, as these factors are beyond the scope of the Red Herring and Steepening Hypotheses as defined in the literature (Zweifel et al., 1999; Buchner and Wasem, 2006) we exclude them from our analysis although we acknowledge their potential impact on the aging-hospital expenditure relationship. Meanwhile, before an assessment of the potentially mediating effects of these factors can be valued, it is imperative to first establish a comprehensive understanding of the Red Herring and the Steepening effects.

2.2 Parameter Hypotheses

A key feature of our parameterization in equation (5) is its ability to perform rigorous statistical testing on whether neither hypothesis is valid, only one holds, or the two hypotheses are simultaneously valid. Specifically, equation (6) defines three hypotheses on the parameters associated with age, time-to-death, and the age-time interaction. The first null hypothesis, H_0^1 , asserts that the Naïve model in equation (2) is correct, implying parameters on time-to-death, time, and the age-time interaction are set to zero. The second null hypothesis, H_0^2 , upholds the Red Herring equation, setting all parameters related to age and the Steepening Hypothesis to zero. Lastly, the third null hypothesis, H_0^3 , assumes the Steepening specification in equation (5) is correct, setting the time-to-death parameters to zero. If all three null hypotheses are rejected, we accept the alternative hypothesis that all parameters are nonzero, indicating that the joint model with both Red Herring and Steepening effects is preferred.

$$\begin{aligned}
 H_0^1: \quad & \delta_{TTD} = \gamma_{TTD} = \delta_T = \gamma_T = \delta_{A \cdot T} = \gamma_{A \cdot T} = 0, \\
 H_0^2: \quad & \delta_A = \gamma_A = \delta_T = \gamma_T = \delta_{A \cdot T} = \gamma_{A \cdot T} = 0, \\
 H_0^3: \quad & \delta_{TTD} = \gamma_{TTD} = 0.
 \end{aligned} \tag{6}$$

Since our parameter hypotheses relate to multiple parameters, standard t-tests are not applicable. Instead, we apply the Wald test, which allows us to test multi-parameter hypotheses. The Wald test has previously been used for statistical tests of the Red Herring Hypothesis (Zweifel

et al., 1999).⁸ In conducting the Wald test, it is essential to account for the joint uncertainty of the parameters, δ and γ , arising from both parts of the model. To address this, we estimate the variance of $(\delta', \gamma)'$ and apply a cluster-robust variance estimator (Liang and Zeger, 1986), with clustering at the individual level, to properly capture the dependencies in the data.

Although the test in equation (6) provide important insights into how a model can be parameterized to encompass regressors related to both Red Herring and Steepening effects, the parametrization test do not necessarily guarantee that the effects of the Red Herring and Steepening Hypotheses are nonzero. Specifically, Steepening effects can still be observed without an age-time interaction, since the second-order derivative of F , g , or the two part model, ψ , with respect to age and time can still be significant, especially for nonlinear functions (see Appendix B.1 for derivations). To assess whether the average marginal effects for the Red Herring or Steepening effects are nonzero, we as a supplement, set up equivalent tests to those in equation (6) but do so with respect to the average marginal effects. Specifically, we test

$$\begin{aligned} H_0^{me,1} &: \mathbb{E} \left[\begin{array}{ccc} \frac{\partial f(H_i, \theta)}{\partial TTD} & \frac{\partial f(H_i, \theta)}{\partial T} & \frac{\partial f(H_i, \theta)}{\partial A \cdot T} \end{array} \right] = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}, \\ H_0^{me,2} &: \mathbb{E} \left[\begin{array}{ccc} \frac{\partial f(H_i, \theta)}{\partial A} & \frac{\partial f(H_i, \theta)}{\partial T} & \frac{\partial f(H_i, \theta)}{\partial A \cdot T} \end{array} \right] = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}, \\ H_0^{me,3} &: \mathbb{E} \left[\frac{\partial f(H_i, \theta)}{\partial TTD} \right] = \begin{bmatrix} 0 \end{bmatrix}, \end{aligned} \quad (7)$$

for the full two-part model where $f = \psi$, at the extensive margin with $f = F$, and at the intensive margin as $f = g$. These hypotheses can be implemented with Wald tests on the marginal effects as calculated with the Delta Method as discussed in Appendix B.2. A rejection of all three null hypotheses, $H_0^{me,1}$ to $H_0^{me,3}$, leads us to accept the alternative hypotheses of coexistent Red Herring and Steepening effects, for a given function, f . As a final robustness check, we compute Bayesian p-values (Kass and Raftery, 1995) to test the parameter hypotheses in equation 6, see Appendix B.3 for further details.

2.3 Decomposing Hospital Expenditure Changes

To assess the relative contributions of the Red Herring and Steepening Hypotheses to rising hospital expenditures, we introduce a new decomposition approach, novel to this literature, which shares similarities with the Kitagawa (1955) and the Oaxaca (1973)-Blinder (1973) decompositions. This framework quantifies how changes in population structure and hospital spending

⁸The Wald test tends to perform similarly to the Likelihood Ratio tests and the Score test in large samples, but the Wald test is often the simplest to compute and interpret (Engle, 1984). A limitation of our testing procedure is that we cannot combine two-sided and one-sided tests. Specifically, the Steepening Hypothesis posits that the marginal effect of the age-time interaction is positive, i.e., a one-sided hypothesis.

patterns drive aggregate hospital expenditures between two periods, t_1 and t_2 .

The decomposition is based on a hospital expenditure model, similar to [van Baal and Wong \(2012\)](#), that facilitates the attribution of expenditure changes to the different model concepts discussed in equations (2) to (5). Besides the concepts, this decomposition presents a different analytical concept from the regression analysis discussed above. We define total annual somatic hospital expenditures, \mathcal{H}_T , as the sum of expenditures across individuals grouped by age, A , and time-to-death, TTD , in period, T ,

$$\mathcal{H}_T := \sum_A \sum_{TTD} N_{T,A,TTD} \bar{H}_{T,A,TTD}, \quad (8)$$

where $N_{T,A,TTD}$ is the number of individuals in each group, and $\bar{H}_{T,A,TTD}$ is their average hospital expenditure. The total change in expenditures, $\mathcal{H}_{t_1} - \mathcal{H}_{t_2}$, can be decomposed into two broad components,

$$\underbrace{\sum_A \sum_{TTD} [N_{t_1,A,TTD} - N_{t_2,A,TTD}] \bar{H}_{t_1,A,TTD}}_{\text{Population Change}} + \underbrace{\sum_A \sum_{TTD} [\bar{H}_{t_1,A,TTD} - \bar{H}_{t_2,A,TTD}] N_{t_2,A,TTD}}_{\text{Expenditure Change}}, \quad (9)$$

relating to the change in the population size and the change in mean hospital expenditures, respectively.

A more detailed four-part decomposition, derived from equation (9) in [Appendix A](#), further isolates key drivers

$$\begin{aligned} & \underbrace{\sum_A [N_{t_1,A} - N_{t_2,A}] \bar{H}_{t_1,A}}_{\text{Naïve}} + \underbrace{\sum_A \sum_{TTD} [N_{t_1,A,TTD} - N_{t_2,A,TTD}] [\bar{H}_{t_1,A,TTD} - \bar{H}_{t_1,A}]}_{\text{Red Herring Error}} + \\ & \underbrace{\sum_A [\bar{H}_{t_1,A} - \bar{H}_{t_2,A}] N_{t_2,A}}_{\text{Steepening effect}} + \underbrace{\sum_A \sum_{TTD} \left[(\bar{H}_{t_1,A,TTD} - \bar{H}_{t_1,A}) - (\bar{H}_{t_2,A,TTD} - \bar{H}_{t_2,A}) \right] N_{t_2,A,TTD}}_{\text{Red Herring-Steepening}}. \end{aligned} \quad (10)$$

The *Naïve* term captures changes in aggregated hospital expenditures attributable to population shifts, $N_{t_1,A} - N_{t_2,A}$, assuming time-invariant mean hospital expenditures, $\bar{H}_{t_1,A}$, while ignoring time-to-death effects.⁹ Despite its simplicity, the Naïve term incorporates considerable detail by accounting for age-structure in both population demographics and hospital expenditures through summation across age groups, A . The *Red Herring Error* term quantifies the

⁹Other Naïve models examining the impact of the Red Herring Hypothesis on aggregate expenditures also exclude time trends in expenditures (see [Polder et al. \(2006\)](#); [Bjørner and Arnberg \(2012\)](#); [Geue et al. \(2014\)](#)), making our approach comparable.

systematic underestimation of near-death expenditures, $\bar{H}_{t_1,A,TTD} - \bar{H}_{t_1,A}$, that occurs when age-based projections exclude time-to-death information. The underestimation term remains time-invariant at the t_1 level signifying that the Red Herring hypothesis addresses the relationship between age and time-to-death without making statements about intertemporal changes.

The third term captures age-specific temporal changes in average hospital expenditures, $\bar{H}_{t_1,A} - \bar{H}_{t_2,A}$, closely resembling the conceptualization of the Steepening Hypothesis by [Gregersen \(2014\)](#) as an interaction between age, time, and healthcare expenditures. The term capture the age-specific temporal changes making the term more comprehensive than the pure age-time interactions considered in the parametrization in Section 2.1 as broader time-trends are also captured in this term. The final term in equation (10) introduces temporal changes in the Red Herring error between periods t_1 and t_2 as $(\bar{H}_{t_1,A,TTD} - \bar{H}_{t_1,A}) - (\bar{H}_{t_2,A,TTD} - \bar{H}_{t_2,A})$. We label this component the *Red Herring-Steepening* term since it captures not only the time-to-death related expenditures addressed by the Red Herring Hypothesis but also the age-specific time-changes considers by the Steepening Hypothesis. Although this interaction of the two hypotheses is not explicitly parameterized in our main linear parametrization in equation (5), subgroup analyses of the Steepening effect by time-to-death can yield insights into its presence within the regression framework.

An interesting result arises from the decomposition: the sum of the Red Herring Error term and the Red Herring-Steepening term in equation (10) must be zero (see Appendix A.1 for a proof). This implies that the Red Herring Error term is exactly equal in magnitude, but opposite in sign, to the Red Herring-Steepening term. While the Red Herring Error term is often interpreted as mitigating part of the expenditure increases associated with demographic changes, its influence is fully counterbalanced by growth rates of end-of-life expenditures, captured by the Red Herring-Steepening term. As a result, although growth in expenditures during the final years of life has received comparatively limited attention in the literature, it plays a role equally important to the Red Herring Hypothesis itself in shaping long-run trends in aggregate hospital expenditure.

3 Data

This paper leverages comprehensive Danish register data from [Statistics Denmark \(2023\)](#), tracked at the individual level, covering the period from 2002 to 2022. We include all individuals and, in line with existing literature, specifically focus on those aged 30 and older ([Carreras et al., 2018](#); [Felder et al., 2010](#); [Werblow et al., 2007](#); [Buchner and Wasem, 2006](#)). This selection includes the age groups that incur the highest hospital expenditures, as well as the ages at which hospital expenditures begin to rise. Our primary focus is on expenditures related to somatic hospital care. All hospital visits are documented in medical records and accessible to researchers, including

the government-insured hospital expenditures at both private and public hospitals. Government insurance covers nearly all hospital procedures in Denmark, with out-of-pocket expenses accounting for 4.6% of total hospital expenditures, while private healthcare insurance covers 0.9% of hospital expenditures ([Insurance and Pension \(FP\), 2023](#)). Somatic hospital expenditures represent a well-defined and significant share of healthcare expenditures in Denmark amounting to 47% of total publicly funded healthcare expenditures in Denmark in 2017 ([Statistics Denmark, 2025](#)).

Each interaction with a hospital is associated with a pseudonymized person identifier, allowing hospital services to be attributed to individual patients. Patients are admitted based on a referral diagnosis and receive either inpatient or outpatient care. The Danish Health Data Authority calculates Diagnosis-Related Grouping (DRG) tariffs for hospital services at a highly detailed procedure-specific level, as described in [Statens Serum Institut \(2014\)](#). These tariffs account for variable costs associated performed procedures such as medical staff time, medical equipment, and most specialized pharmaceuticals, while attributing fixed costs to individual patients is generally considered impractical. If a patient remains hospitalized beyond the standard treatment duration for a specific procedure, additional costs for the extended stay are added. DRG-based hospital expenditures are widely used in existing literature on Danish healthcare data ([Christensen et al., 2016](#); [Kollerup et al., 2022](#)). For the hospital expenditure data, we consider the observation period from 2002 to 2017 where we recorded 109,381,459 individual hospital interactions. Due to substantial changes in the hospital registers in 2018 and the absence of an established method to ensure cross-year comparability, we restrict our analysis of hospital services to the period prior to 2018.

For our analysis, we aggregate hospital expenditures into monthly amounts, capturing all expenditures incurred from the first to the last day of each month.¹⁰ This allocation is possible due to detailed admission and discharge dates in the DRG registers. The use of monthly data strengthens our study by capturing the progressive increase in hospital expenditures as individuals near the end of life ([Christensen et al., 2016](#)). Additionally, this approach helps prevent the underestimation of hospital expenditures for individuals with less than a year to live, a limitation often associated with annual data. To ensure the accuracy of our hospital expenditure data, we confirm that the aggregate annual expenditures in our dataset matches the raw DRG register data and the age-specific averages of hospital expenditures in [Christensen et al. \(2016\)](#) and [Kollerup](#)

¹⁰If a hospital stay extends across multiple months, expenditures are distributed proportionally based on the number of days spent per month. Hospital expenditures are inflated to 2020 levels in DKK using the consumer price index from [Statistics Denmark \(2024a\)](#) and converted to 2020 USD using the 2020 exchange rate of 1/6.534 from [Danmarks Nationalbank \(2024\)](#).

et al. (2022).¹¹

From the population register (BEF) (Statistics Denmark, 2024b), we obtain demographic information, including sex and age. The time-to-death of deceased individuals is determined using the date of birth from the population register (BEF) and the date of death from the deaths register (DODAASG) (Danish Health Data Authority, 2024), which covers data through the end of 2022. This approach allows us to precisely determine the time-to-death for all individuals within the four years preceding their passing, offering a detailed and comprehensive perspective on hospital utilization during the final stages of life. Seshamani and Gray (2004b) find that time-to-death effects on hospital expenditures become negligible beyond four years, and our dataset spans a sufficient duration to capture these effects.¹² Both age and time-to-death are measured in years, with decimal values representing additional days. We measure calendar time in years at a monthly resolution as $(year - 2002) + (month - 1)/12$, setting January 1, 2002, as the base date. After combining the registers, we are left with 638,184,716 monthly observations, corresponding to 3,323,879 person-years annually.¹³

3.1 Descriptive Statistics

To gain an initial understanding of how hospital expenditures correlate with key variables, we present mean hospital expenditures in Figure 1, stratified by age, sex, time-to-death, and time period.

Mean hospital expenditures by age inform about the Naïve model of equation (2). In Figure 1a, mean hospital expenditures are shown by age, with a solid purple line representing females in 2002 and a solid yellow line representing males in 2002. The dashed lines show their 2017 equivalents. We observe that average hospital expenditures generally increase with age, peaking around age 90, after which they level off. An exception occurs for females in their 30s and early 40s, where expenditures are higher due to reproductive health needs. This pattern is mirrored on the extensive margin in 1c, which illustrates the outcome of the first part of equation (1), depicting the percentage of observations with any hospital expenditure by age. Notably, these probabilities peak slightly earlier, around age 80. In contrast, the intensive margin—the

¹¹Specifically for Kollerup et al. (2022), that is, after accounting for the inclusion of additional expenditure types in their analysis. We do not report these comparisons in this paper for brevity.

¹²For individuals alive beyond 2021, we cannot observe their exact time-to-death if it exceeds four years as their death is right-censored.

¹³The project, no. 704863, was carried out under the Agreement 2016-051-000001, sequential number 1447, between The Danish Data Protection Agency and Aarhus University and has been approved by Statistics Denmark. There is no Danish institutional review board for studies based on standard register data at Statistics Denmark. Individual-level information from the administrative registers is confidential to the Danish Administrative Procedures (Section 27) and the Danish Criminal Code (Section 152), and data cannot be made publicly available. The data used for this analysis are available upon submission of an application to Statistics Denmark, <https://www.dst.dk/en>. Restrictions may apply to the availability of data. Upon request, the authors will assist in replicating study results.

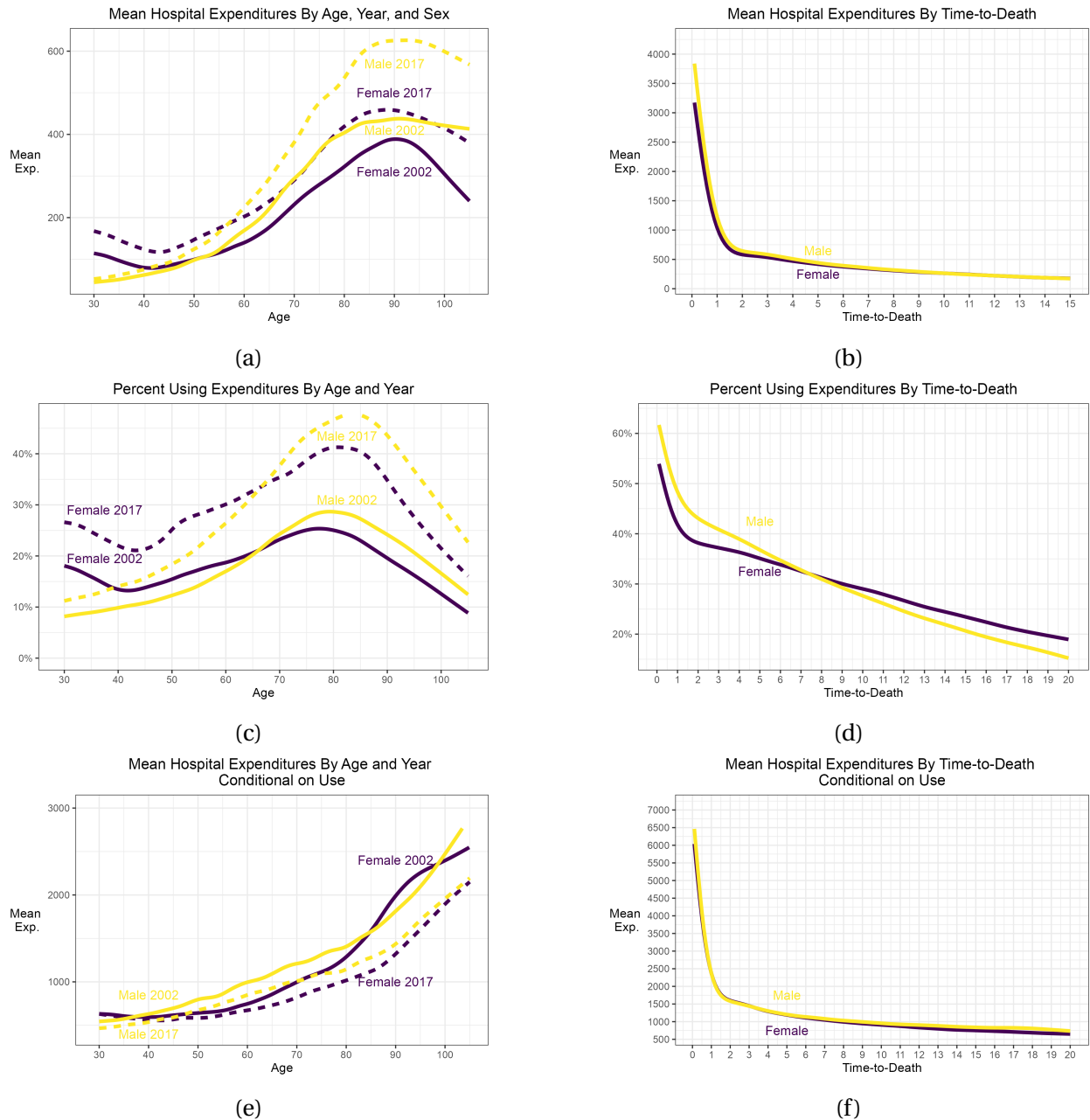


Figure 1: Summary statistics by age, time-to-death, year, and sex. (a) Mean monthly hospital expenditures by age for all individuals; (b) mean monthly hospital expenditures by time-to-death for all individuals; (c) percent using hospital expenditures within a month by age for all individuals; (d) percent using hospital expenditures within a month by time-to-death for all individuals; (e) mean monthly hospital expenditures by age for individuals with positive hospital expenditures; and (f) mean monthly hospital expenditures by time-to-death for individuals with positive hospital expenditures. All amounts are expressed in 2020 U.S. dollars.

second part of equation (1), shown in 1e —reveals a different trend. Among individuals with positive hospital expenditures, mean expenditures rise consistently with age, indicating that older individuals incur higher expenditures in months with admissions.

We also examine the relationship between hospital expenditures and time-to-death for deceased individuals, which is a key focus of the Red Herring hypothesis. Figure 1b demonstrates that mean hospital expenditures peak within a month of death and then decline sharply until approximately 1.5 years prior to death. After this point, expenditures decrease more gradually in an almost linear fashion. This trend holds true for both sexes and aligns with findings from previous studies, such as French et al. (2017). For the first stage of the two-part model in equation (1)—the probability of incurring hospital expenditures— Figure 1d shows that this probability is highest close to death, with many individuals receiving hospital care during their final months. The trend in Figure 1d mirrors that in Figure 1b, with an initial steep decline followed by a more gradual reduction, albeit with a different outcome measure. Notably, more than 50% of females and over 60% of males use hospital care in their last month of life. A pronounced sex disparity is observed, with males exhibiting a higher likelihood of hospitalization up to approximately 7.5 years prior to death compared to females. Finally, Figure 1f presents mean hospital expenditures by time-to-death for the subsample of individuals who use hospital care, corresponding to the second part of the two-part model in equation (1). This plot exhibits a similar rapid decline in expenditures until 1.5 years prior to death, but with larger mean values compared to Figure 1b, as individuals with zero hospital expenditures are excluded.

By comparing age-specific hospital expenditures from 2002 to the averages in 2017, we derive descriptive insights into the Steepening Hypothesis for hospital expenditures in absolute terms. Comparing the solid line for 2002 data with the dashed line for 2017 in Figure 1a, we observe a widening gap in expenditures for males aged 30 to 90. At older ages, however, this difference narrows. For females, the trend is less pronounced, suggesting potential differences in Steepening patterns by age and sex. A similar pattern emerges in the proportion of observations utilizing hospital care, as shown in Figure 1c, which aligns with the first component of equation (1). However, when examining mean expenditures among individuals with positive hospital care expenditures conditional on use, Figure 1e reveals a contrasting trend: for both sexes, mean hospital expenditures decreased between 2002 and 2017. This finding challenges conventional expectations about hospital spending trends and suggests that the Steepening Hypothesis may not consistently hold if hospital expenditures decrease in age and time. Specifically, in the second component of the two-part model, average hospital expenditures per user have declined over time across most age groups for both sexes.

4 Results

This section presents the findings of our study for the full population of individuals. Table 1 displays the regression results derived from the model specified in equation (1), utilizing four distinct parameterizations of the Probit and Poisson models, as outlined in equations (2) to (5). Additionally, we provide estimates of the nominal marginal association, E_{ψ} , for the two-part model, computed using the Delta Method, as detailed in Section 2.2. Our computational resources enable the estimation of our models for a random subsample of 10,000,000 observations.

The first three columns report estimates from the Naïve specifications, which remain agnostic to the parameters of the Red Herring and Steepening Hypotheses by excluding time-to-death, age interacted with time, and time itself, incorporating only age and other control variables as regressors. We find that age is associated with an increase in the mean probability of incurring any hospital expenditure in the Probit model, as indicated by a positive parameter estimate of 0.014. Moreover, the p-value reported in the squared brackets indicates that this parameter estimate is statistically significant at the 5% level. Similarly, the Poisson estimates reveal a positive relationship with age, yielding a parameter estimate of 0.017. This translates into a statistically significantly higher mean monthly hospital expenditure of \$6.94 per year of age. Comparing these estimates to other studies is inherently challenging due to differences in econometric specifications, healthcare systems, the scope of healthcare expenditures considered, and reporting methods. Nevertheless, our findings closely align with [Seshamani and Gray \(2004b\)](#), who estimate that hospital expenditures in England increase by approximately \$6.88 per month between ages 65 and 70.¹⁴ Similarly, [Hyun et al. \(2016\)](#) report a monthly increase of \$7.08 in inpatient hospital expenditures for individuals aged 60–64 compared with those aged 65–70.¹⁵

In columns four to six, the Red Herring parameterization introduces time-to-death (TTD) as a regressor. Consistent with the existing literature ([Seshamani and Gray, 2004b](#)), TTD exhibits large, negative parameter estimates in both the Probit and Poisson models. Consequently, in the full two-part model, each additional year further from death is associated with a \$177.00 reduction in average monthly hospital expenditures. The age coefficients decrease across all three Red Herring columns, corroborating the findings of [Zweifel et al. \(1999\)](#) that the age effect diminishes upon accounting for time-to-death. The age parameter estimates, although reduced, remain statistically significant, similar to existing literature using larger datasets ([Seshamani and Gray, 2004a](#); [Atella and Conti, 2014](#)). Additionally, the introduction of time-to-death as a

¹⁴This calculation is based on a £250 increase over five years, converted to a monthly rate using the January 1999 exchange rate of \$/£1.65 ([Bank of England, 2025](#)): \$6.88 per month = \$/£1.65 · £250 / 5 years / 12 months per year.

¹⁵This is derived from a 484,000 KRW difference in inpatient expenditures in 2012, converted to USD using the January 2012 exchange rate of 1 USD = 1,140 KRW ([Federal Reserve \(US\), 2025a](#)) \$7.08 per month = \$/KRW1/1,140 · KRW 484,000 / 5 years / 12 months per year.

Table 1: GLM model estimates

Model:	Naïve			Red Herring			Steepening			Joint Model		
	Probit	Poisson	E_{ψ}	Probit	Poisson	E_{ψ}	Probit	Poisson	E_{ψ}	Probit	Poisson	E_{ψ}
Age	0.014 (0.004) [0.000]	0.017 (0.000) [0.000]	6.94 (1.890) [0.000]	0.012 (0.003) [0.000]	0.003 (0.000) [0.000]	3.43 (0.517) [0.000]	0.011 (0.003) [0.000]	0.018 (0.000) [0.000]	6.88 (1.896) [0.000]	0.008 (0.002) [0.000]	0.003 (0.000) [0.000]	3.32 (0.504) [0.000]
TTD	- (-) [-]	- (-) [-]	- (-) [-]	-0.101 (0.019) [0.000]	-0.470 (0.002) [0.000]	-117.00 (42.194) [0.006]	- (-) [-]	- (-) [-]	- (-) [-]	-0.103 (0.020) [0.000]	-0.470 (0.002) [0.000]	-117.41 (41.915) [0.005]
Age-Time	- (-) [-]	- (-) [-]	- (-) [-]	- (-) [-]	- (-) [-]	- (-) [-]	0.033 (0.006) [0.000]	-0.011 (0.001) [0.000]	11.44 (1.659) [0.000]	0.037 (0.006) [0.000]	0.003 (0.000) [0.000]	9.15 (0.785) [0.000]
Controls	Yes			Yes			Yes			Yes		
DoF	9,999,972			9,999,968			9,999,968			9,999,964		
Log Lik.	-3,196,833,760			-2,906,760,856			-3,194,966,130			-2,906,648,274		

E_{ψ} refers to the marginal effects of the full model. Standard errors in parenthesis. P-values in squared brackets. TTD is short for time-to-death. To ensure convergence in the numerical optimization (Greene, 2017), Age-Time is divided by 100 in the Probit and Poisson regressions but not in E_{ψ} which is the average marginal effect with respect to age and time. Controls refer to an intercept, month dummies, and a male dummy. The Red Herring and Joint Model also include \tilde{X}^* as a control while time is also included in the Steepening and Joint Model. DoF is shorthand for degrees of freedom. Models are estimated with 10,000,000 observations and 2,191,490 observations in the Poisson part. Amounts in 2020 dollars.

regressor increases the log-likelihood function, indicating improved model fit.

The Steepening parameterization in columns seven through nine incorporates time and the age-time interaction into the Naïve specification. Notably, the coefficient for the age-time interaction, estimated at 0.033 in the Probit model, suggests that the probability of incurring hospital expenditures increases more rapidly with age in later calendar years than in earlier years. Conversely, in the Poisson model, the negative coefficient of -0.011 reflects a counter-Steepening effect, indicating that among those who do incur hospital expenditures within a month, expenditures growth decrease over time for older individuals, a pattern observed in Figure 1e. Meanwhile, simply interpreting the parameter estimates can be misleading, especially for nonlinear functions and interaction effects, thus marginal associations are preferred for the direction and size of effects (see Appendix B.1 for the calculation of marginal associations). The marginal association estimate, accounting for the full set of the Probit and Poisson coefficients, results in a net increase of \$11.44 in expenditures. Thus, each additional year of age amplifies the increase in hospital spending over time by \$11.44 on average. Gregersen (2014) report a comparable figure for Norwegians aged 50-54 in 2014, amounting to \$11.05.¹⁶ The inclusion of Steepening parameters increases the log-likelihood function, though to a lesser extent than time-to-death.

In the final three columns of Table 1, we present our primary results, examining the Red Her-

¹⁶An average annual hospital expenditure Steepening effect of €183.3 for 50-54-year-olds corresponds to a monthly rate of \$11.05 = \$/€1/1.38 · 183.3€ / 12 months, using the exchange rate of 1.38 in January 2014 (Federal Reserve (US), 2025b).

ring Hypothesis and the Steepening Hypothesis simultaneously using the new parameterization specified in equation (5), which incorporates time-to-death, time, and age interacted with time as regressors. Compared to the estimates under the Red Herring parameterization, the inclusion of the Steepening parameters (time and the age-time interaction) produces minimal changes in the estimates for age and time-to-death. The age effects remain small, while the time-to-death effects are substantial. When comparing the Steepening parameterization estimates to those from the Joint Model, we observe that the introduction of time-to-death effects greatly influences the parameter estimates. Notably, the interaction effect between age and time in the Poisson component changes from a negative to a positive value of 0.003, suggesting a Steepening effect where mean hospital expenditures increase more rapidly for older individuals. The overall effect in the Joint Model, E_ψ , is slightly reduced to \$9.15.¹⁷ Incorporating both time-to-death from the Red Herring Hypothesis and the Steepening parameters yields the highest log-likelihood value, indicating the best fit to data. A complete regression table, including additional controls and marginal effects in the Probit and Poisson components, is provided in Appendix C.1. We analyze the impact of the Red Herring and the Steepening Hypotheses on aggregate hospital expenditures in Section 5.

Hypotheses Tests

In Table 1, the full-population parameter estimates are all highly statistically significant at the 5% level, as shown by the p-values in square brackets.¹⁸ For each single parameter, a test of a zero-effect null hypotheses on time-to-death, and the interaction of age with time, can all be rejected irrespective of the model specification. Meanwhile, we are also interested in testing multiple parameter simultaneously, i.e., in both parts of our two-part model as well as the joint test of both the Red Herring and the Steepening parametrizations. To test whether multiple parameters are simultaneously zero in the full two-part model in equation (1), Table 2 tests the three hypotheses on the parameter estimates from equation (6) using the Wald test.

Table 2: Wald Hypothesis Test of Parameters

	#Res.	Test Statistic	P-value
H_0^1	6	106,217	0.000
H_0^2	6	20,762	0.000
H_0^3	2	39,247	0.000

Hypothesis tests and p-values. #Res. refers to the number of parameter restrictions.

¹⁷The full marginal association of \$9.15 is lower than the \$11.44 observed under the Steepening parameterization, despite the positive shift in the age-time interaction parameter in both the Probit and Poisson parts. This reduction is due to the inclusion of additional regressors and the associated changes in other regression coefficients, which affect the estimated effect.

¹⁸We expect highly significant estimates due to our huge sample size.

The first hypothesis, H_0^1 , maintains the Naïve model as the true specification; the second hypothesis, H_0^2 , proposes a strict Red Herring setup with only time-to-death as regressor; and, lastly, H_0^3 posits the Steepening specification, excluding time-to-death as a regressor, as discussed in Section 2.2. Each Wald-test returns a p-value of zero, providing strong evidence against H_0^1 , H_0^2 , and H_0^3 , all of which are rejected. Thus, we fail to reject the alternative hypothesis that all parameters are nonzero. This leaves the Joint Model as the preferred specification. Wald tests of the marginal effects and Bayesian p-value tests confirm this finding (Appendix B.2), and we draw similar conclusions when analyzing the female and male populations separately, as reported in Appendix C.2.

4.1 Heterogeneity by Age and Sex

The impact of a changing population composition on hospital expenditures depends on which age and sex groups experience changes in size. This variation is particularly relevant when effect sizes differ across subpopulations, thus necessitating a precise understanding of these disparities. Existing studies provide valuable insights into age-related differences: [Seshamani and Gray \(2004a\)](#) and [Atella and Conti \(2014\)](#) explore the Red Herring Hypothesis, while [Kollerup et al. \(2022\)](#) and [Gregersen \(2014\)](#) examine age differences in the context of the Steepening Hypothesis. However, these analyses are limited, providing only partial evidence, as they consider each hypothesis in isolation, thus overlooking their potential coexistence. Formal statistical tests remain absent from the literature for coexistence as well as group-differences. Finally, heterogeneity by sex has received minimal attention in both strands of literature ([Kallestrup-Lamb et al., 2024](#)).

In this section, we present and implement a method to address these limitations. Traditionally, the effects of age-groups and sex are modeled using binary indicators, leading to complex, high-dimensional models that are difficult to interpret. We take a more parsimonious approach by analyzing each group separately, which allows for greater flexibility as each group has distinct parameter estimates for controls and key variables. Additionally, estimating separate parameters enables the use of Welch tests ([Welch, 1947](#)) to assess whether aging patterns—measured by age, time-to-death, and Steepening—significantly differ across age and sex subgroups. While the Steepening Hypothesis is classically formulated as a global statement, i.e., that the growth rate of healthcare expenditures increases across the entire age range, we argue that estimating age-specific patterns can provide more nuanced insight into how and where Steepening occurs. By restricting estimation to narrower age bands, we allow for the possibility that the rate of expenditure growth is not uniform over the age-distribution. From a policy perspective, these insights helps identify which age-sex subpopulations may exhibit the strongest hospital

expenditure acceleration under Red Herring and Steepening effects.

Table 3 presents these estimates. We consider sex-specific estimates from our Joint Model across 10-year age groups, including those aged 90 and above. In Panel A, we report the full marginal effects of the two-part model, E_ψ , accounting for the competing contributions of the Probit and Poisson components via the Delta Method.¹⁹ Our estimates confirm findings from the literature that considers each hypothesis separately for hospital expenditures as the outcome.²⁰ The time-to-death (TTD) effects generally increase in absolute magnitude across age groups, indicating that end-of-life hospital expenditures are highest for older individuals (Seshamani and Gray, 2004b). Additionally, Steepening effects (Age · Time) are statistically significant for all age groups and increase with age (Gregersen, 2014). Our approach, however, also provides several new insights, made possible by accounting for coexistent hypotheses and subgroup analyses by age and sex. First, we find simultaneous Red Herring and Steepening effects for all age and sex groups, as shown in Panel A and confirmed by Wald tests (see Table A6, Appendix C.3), suggesting that evaluating each hypothesis in isolation may be misleading. Second, time-to-death and Steepening effects are generally statistically different across age groups, as implied by Welch tests (Table A8, Appendix C.3), demonstrating that population change affects each subgroup differently. Third, the magnitude of the effects differ by sex and are statistically significantly different from each other (Table A9, Appendix C.3), with males generally exhibiting higher time-to-death and Steepening effects than females. Fourth, not all age groups experience Steepening effects. Among females aged 30–40 and individuals aged 90 and above, the Steepening coefficient is negative, suggesting that hospital expenditure growth is lower among older individuals in these groups. Instead, Steepening effects are most pronounced between ages 50 and 80, indicating a steepening of the age-profile in this interval.

To further disentangle the relative contributions of the Probit and Poisson components, we replicate the analysis from Panel A separately for the Probit part in Panel B and the Poisson part in Panel C, corresponding to the extensive and intensive margins, respectively. For time-to-death, coefficients are negative in both Panel B and Panel C, indicating that both the probability of hospital utilization and the amount of expenditures in a given month decrease the further an individual is from death. While both components contribute to the overall pattern, their relative magnitudes vary across age groups. For the Steepening effect, the patterns across the two margins are more nuanced. At the extensive margin (Panel B), all Steepening effects are

¹⁹Specifically, the change in hospital expenditures with respect to a variable, x , is the expectation of $\partial\psi/\partial x = \partial F/\partial x \cdot g + F \cdot \partial g/\partial x$ across individuals for x being age or time-to-death. For the age-time interaction, accounting for the full influence of parameter estimates on age, time, and the age-time interaction is critical, as derived in Appendix B.1.

²⁰For long-term care expenditures Karlsson and Klohn (2014) finds a different pattern with a few instances of lower expenditures among older individuals. Moreover, in a non-regression based descriptive statistic, Tanuseputro et al. (2015) find hospital expenditures to be larger in the last single year of life for individuals above age 75.

positive, except for females aged 90+, with the largest effects observed between ages 50 and 80. Thus, the probability of hospital utilization contributes positively to the full effect in Panel A. On the intensive margin (Panel C), this pattern diverges. Hospital expenditures grow more slowly over time for individuals aged 30–40 and those aged 80+, with the latter effect particularly pronounced among males. The largest absolute Steepening parameter is -\$26.41 for males at ages 90+, while for females, the most notable effect appears in the 30–40 age group, likely reflecting childbirth-related expenditures. The negative Steepening effects for the oldest age groups suggest that hospital expenditure growth is contained among those already utilizing hospital services, mitigating some of the fiscal pressures associated with population change at older ages. However, Steepening effects remain positive between ages 40–80 for females and 40–70 for males, indicating a significant expenditure burden from demographic expansion in these age ranges among hospitalized individuals.

Table 3: Marginal Association by Age Groups

Age Group	[30;40)		[40;50)		[50;60)		[60;70)		[70;80)		[80;90)		90+	
Sex	F	M	F	M	F	M	F	M	F	M	F	M	F	M
Panel A: Full Model, E_ψ														
Age	-5.34 (0.560)	1.71 (0.296)	1.92 (0.215)	3.72 (0.621)	2.79 (0.386)	5.76 (0.895)	4.98 (0.407)	9.23 (1.046)	3.15 (0.443)	-0.76 (0.378)	-6.33 (0.264)	-17.29 (1.355)	-18.02 (0.587)	-26.66 (0.808)
TTD	-92.48 (8.377)	-34.00 (4.461)	-91.99 (8.985)	-61.06 (9.203)	-124.18 (13.674)	-117.10 (19.229)	-163.51 (22.332)	-193.40 (32.393)	-217.17 (32.398)	-278.30 (39.701)	-215.68 (30.377)	-305.32 (32.730)	-176.62 (15.762)	-304.04 (13.984)
Age-Time	-1.92 (0.704)	2.55 (0.070)	5.49 (0.351)	5.88 (0.607)	8.23 (0.624)	10.11 (1.166)	11.37 (0.706)	17.90 (1.494)	12.80 (1.161)	11.54 (0.489)	5.77 (1.002)	1.23 (0.460)	-7.65 (0.164)	-8.56 (0.495)
Panel B: Extensive Margin, $\Pr(H>0)$, E_F														
Age	-0.006 (0.001)	0.002 (0.000)	0.002 (0.000)	0.003 (0.000)	0.003 (0.000)	0.006 (0.001)	0.005 (0.000)	0.009 (0.001)	0.002 (0.000)	0.005 (0.000)	-0.008 (0.000)	-0.008 (0.000)	-0.011 (0.000)	-0.012 (0.000)
TTD	-0.052 (0.005)	-0.017 (0.002)	-0.060 (0.006)	-0.032 (0.003)	-0.065 (0.006)	-0.049 (0.005)	-0.064 (0.005)	-0.054 (0.004)	-0.050 (0.003)	-0.051 (0.003)	-0.030 (0.001)	-0.037 (0.001)	-0.017 (0.000)	-0.030 (0.000)
Age-Time	0.000 (0.000)	0.005 (0.001)	0.008 (0.001)	0.007 (0.001)	0.011 (0.001)	0.011 (0.001)	0.014 (0.001)	0.017 (0.002)	0.013 (0.001)	0.017 (0.001)	0.005 (0.000)	0.006 (0.000)	-0.004 (0.000)	0.000 (0.000)
Panel C: Intensive Margin, $H H>0$, E_g														
Age	-8.24 (1.130)	5.66 (1.290)	2.28 (0.517)	11.58 (1.402)	4.10 (0.853)	6.42 (1.779)	3.36 (0.262)	3.78 (0.679)	3.25 (0.502)	-16.81 (1.213)	10.65 (2.320)	-17.80 (1.332)	7.43 (1.300)	-28.00 (0.584)
TTD	-292.01 (9.461)	-228.19 (6.570)	-303.68 (5.048)	-295.32 (3.901)	-353.21 (7.498)	-422.43 (23.169)	-410.47 (26.442)	-540.81 (45.908)	-519.21 (46.169)	-618.17 (56.547)	-562.93 (48.918)	-672.05 (50.171)	-623.33 (34.884)	-812.45 (25.090)
Age-Time	-9.85 (1.912)	-0.53 (0.517)	2.43 (0.852)	8.83 (0.900)	4.60 (1.274)	5.16 (1.361)	1.75 (0.253)	6.04 (1.240)	1.59 (0.586)	-19.06 (1.711)	-0.88 (1.383)	-17.34 (1.733)	-7.96 (1.055)	-26.41 (0.935)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. Obs.	1,036,274	1,056,546	1,107,349	1,133,078	1,041,301	1,049,897	912,058	883,170	610,754	520,022	337,275	204,593	80,406	27,277
Num. Obs. E_g	210,671	112,176	200,262	150,483	242,061	191,218	254,788	231,560	198,025	186,890	106,966	78,561	19,133	8,696

E_F , E_g , and E_ψ refer to the marginal association of the Probit, Poisson, and full two-part model, respectively. Standard errors in parenthesis. P-values in squared brackets. TTD is short for time-to-death. Controls refer to an intercept, month dummies, time, and \tilde{X}^* . DoF is shorthand for degrees of freedom. Amounts in 2020 dollars.

4.2 Heterogeneity by Time-to-Death and Sex

The decomposition proposed in Section 2.3 suggests, mathematically, that a new interaction effect combining the Red Herring and the Steepening Hypotheses might exist. To assess the empirical standpoint on this, we estimate our Joint Model for three distinct time-to-death groups. If the Steepening effect differs across time-to-death groups, there is evidence favoring this new Red Herring-Steepening interaction. The decomposition does not address differences by sex, but since Section 4.1 found marked differences in Steepening effects by sex, we conduct this analysis for each sex separately. The three time-to-death groups are: less than 1.5 years to death (around the kink in the time-to-death plot in Figure 1b), between 1.5 and 4 years, and more than 4 years to death, respectively.

In Panel A of Table 4, we present the marginal associations of the two-part model, E_{ψ} . Comparing our estimates of time-to-death effects with the existing literature that focuses exclusively on the Red Herring hypothesis without addressing the Steepening Hypothesis, our findings corroborate those of previous studies (Seshamani and Gray, 2004a). Specifically, the time-to-death effect exhibits a nonlinear relationship with proximity to death, characterized by escalating expenditures as death approaches. For individuals closest to death, our time-to-death effects are substantial: -\$2,254.46 for females and -\$2,928.96 for males. These effects diminish considerably to -\$113.99 and -\$146.07, respectively, for those with 1.5 to 4 years remaining until death. For individuals with more than four years to death, the time-to-death parameter cannot be estimated due to insufficient variation in time-to-death within this subsample.

Variability in the Steepening effects among individuals with different time-to-death has not been previously examined in the literature. Large and positive Steepening effects are concerning for both individuals still far from death and those nearing the end of life. If healthcare expenditures steepen in the years just before death, future aggregate hospital expenditures could increase greatly, especially if the number of individuals near death increase as seen, e.g., with the aging of the baby boomer cohorts. For females with less than 1.5 years to death, the negative Steepening effect of -\$5.84 indicates that older individuals are experiencing flatter (or declining) spending trajectories over time, thus reducing the expenditure pressure of expanding older populations. Meanwhile, for the remaining two time-to-death groups, the Steepening effect is positive, with magnitudes of \$3.88 and \$6.92, respectively. Males, however, have exhibited a substantially larger degree of Steepening across all three time-to-death groups: \$40.06, \$14.74, and \$8.11, respectively. Consequently, a surge in the older population may present a markedly greater economic concern for males than it does for females in the last years of life. Welch tests of equal Steepening parameters across age groups reject the null hypothesis of equal parameters and support the alternative that parameter estimates differ significantly for both females and males (see Table A10, Appendix C.4). This heterogeneity provides empirical support for the Red

Table 4: Marginal Association by Time-to-Death Groups

Sex	TTD \in [0;1.5)		TTD \in [1.5;4)		TTD $>$ 4	
	F	M	F	M	F	M
Panel A: Full Model, E_ψ						
Age	-59.40 (4.623) [0.000]	-36.57 (2.682) [0.000]	-14.93 (1.313) [0.000]	-7.68 (1.000) [0.000]	2.87 (1.014) [0.005]	4.97 (1.490) [0.001]
TTD	-2,254.46 (192.133) [0.000]	-2,928.96 (241.446) [0.000]	-113.99 (13.018) [0.000]	-146.07 (15.171) [0.000]	- (-) [-]	- (-) [-]
Age·Time	-5.84 (0.735) [0.000]	40.06 (2.953) [0.000]	3.88 (0.062) [0.000]	14.74 (0.376) [0.000]	6.92 (0.505) [0.000]	8.11 (1.221) [0.000]
Panel B: Extensive Margin, $\Pr(H>0)$, E_F						
Age	-0.010 (0.000) [0.000]	-0.002 (0.000) [0.000]	-0.006 (0.000) [0.000]	0.001 (0.000) [0.000]	0.002 (0.001) [0.000]	0.005 (0.001) [0.000]
TTD	-0.137 (0.004) [0.000]	-0.144 (0.003) [0.000]	-0.017 (0.001) [0.000]	-0.022 (0.001) [0.000]	- (-) [-]	- (-) [-]
Age·Time	0.000 (0.000) [0.000]	0.010 (0.000) [0.000]	0.006 (0.000) [0.000]	0.014 (0.000) [0.000]	0.010 (0.002) [0.000]	0.011 (0.002) [0.000]
Panel C: Intensive Margin, $H H>0$, E_g						
Age	-61.94 (4.050) [0.000]	-59.16 (3.179) [0.000]	-17.70 (1.331) [0.000]	-22.73 (1.777) [0.000]	5.64 (2.071) [0.006]	8.00 (3.208) [0.013]
TTD	-4,070.09 (267.076) [0.000]	-4,693.04 (303.080) [0.000]	-234.39 (19.763) [0.000]	-269.15 (19.981) [0.000]	- (-) [-]	- (-) [-]
Age·Time	-13.09 (1.066) [0.000]	14.73 (1.316) [0.000]	-14.22 (0.821) [0.000]	-17.47 (1.918) [0.000]	0.50 (0.290) [0.085]	1.52 (0.856) [0.076]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Num. Obs.	115,256	110,539	189,158	185,651	4,821,003	4,578,393
Num. Obs. E_g	52,070	57,501	71,025	77,266	1,108,811	824,817

E_F , E_g , and E_ψ refer to the marginal association of the Probit, Poisson, and full model, respectively. Standard errors in parenthesis. P-values in squared brackets. TTD is short for time-to-death. Controls include an intercept, month dummies, time, and \check{X}^* . Amounts in 2020 dollars.

Herring-Steepening interaction. Thus, the Red Herring and Steepening Hypotheses demonstrate specific nuances not previously considered in the literature. Panel A, Figure 4 yields several additional findings. First, we accept the alternative hypothesis of coexistence of the Red Herring and Steepening effects within each time-to-death group for both sexes, as Wald tests reject all null hypotheses detailed in Section 2.2 (Table A11, Appendix C.4). Second, the Steepening effects differ significantly between females and males for each time-to-death group (Table A11, Appendix C.4). These results demonstrate that hospital expenditure patterns are substantially influenced by the sex composition of population changes.

To assess the extent to which extensive and intensive margins contribute to these findings, we estimate the marginal associations for the Probit and Poisson models in Panels B and C, respectively. At the extensive margin, all subgroups exhibit positive Steepening effects, indicating that over time, older individuals had a higher probability of incurring any hospital expenditures within a given month. However, for females with the shortest time-to-death, this effect size approaches zero. Combined with a negative Steepening effect at the intensive margin of -\$13.09, the net effect for this subgroup is negative in Panel A. At the intensive margin, Steepening is also negative for both males and females with 1.5 to 4 years to death, suggesting that older individuals approaching death have been utilizing fewer monthly hospital expenditures over time. Males with less than 1.5 years to death represent an exception, with a parameter estimate of \$14.73. Notably, all demographic groups with more than four years to death show insignificant Steepening effects at the intensive margin, indicating that age-specific growth in hospital expenditures has occurred predominantly near end-of-life. Consequently, the overall Steepening effects observed for most groups in Panel A can be primarily attributed to Steepening occurring at the extensive margin.

5 Decomposition Results

In this section, we apply the decomposition method introduced in Section 2.3 to the full dataset described in Section 3, with monthly hospital expenditures aggregated to annual amounts for each individual. Table 5 presents the results of the analysis for the years $t_1 = 2017$ and $t_2 = 2002$. As shown in the first row, total hospital expenditures increased by \$3,594.1 million from 2002 to 2017. For the two-component decomposition in Panel A, only 20.8% of this increase was attributed to population changes, while a substantial 79.2% resulted from intertemporal changes in mean hospital expenditures. We also observe notable differences by sex. For females, just 12.6% of the expenditure increase was attributable to population changes, whereas for males, this figure was nearly three times larger at 30.9%. Conversely, changes in mean hospital expenditures played a more significant role for females, accounting for 87.4% of the total increase.

Turning to the four-component decomposition in Panel B, the Naïve component exerted an upward pressure on expenditure growth, accounting for 40.8% of the change. Meanwhile, the Red Herring error had the expected mitigating effect, reducing the increase by 19.3%, aligning closely with previous estimates that do not account for Steepening (Breyer and Lorenz, 2021). While the Red Herring error exhibited little variation by sex, the Naïve age component was more pronounced among males, exceeding half of their total hospital expenditure growth from 2002 to 2017. Steepening effects, representing intertemporal changes in hospital expenditures, contributed significantly to expenditure growth, accounting for 59.9% of the total increase, three times larger than the Red Herring error in absolute terms. This effect was particularly pronounced for females, where it explained 69.6% of the change. In the final two columns of Table 5, we examine the newly identified Red Herring-Steepening effect, derived in Section 2.3 and empirically validated in Section 4.2. This effect contributed 19.3% to the total hospital expenditure increase, roughly one-fifth of the total growth. Its magnitude was nearly identical for both sexes, underscoring its importance in explaining rising hospital expenditures. Crucially, the Red Herring-Steepening component was of the same absolute size as the time-to-death prediction error (19.3%), reinforcing the significance of this phenomenon. A robustness analysis using the years $t_1 = 2015$ and $t_2 = 2005$ confirms our main decomposition estimates presented in Table 5. The robustness estimates exhibit minimal deviation from the primary findings, as detailed in Appendix C.6.

Table 5: Decomposition of Hospital Care Expenditure Differences between 2002 and 2017

	Panel A: Two-Component Decomposition					Panel B: Four-Component Decomposition							
	Total	Pop. Change		Healthc. Change		Naïve		RH Error		Steepening		RH Steep.	
	\$mio.	\$mio.	%	\$mio.	%	\$mio.	%	\$mio.	%	\$mio.	%	\$mio.	%
All	3594.1	747.6	20.8	2846.5	79.2	1440.6	40.1	-693.0	-19.3	2153.5	59.9	693.0	19.3
Female	1740.1	218.9	12.6	1521.2	87.4	528.1	30.4	-309.2	-17.8	1212.0	69.6	309.2	17.8
Male	1854.0	573.2	30.9	1280.8	69.1	954.0	51.5	-380.7	-20.5	900.0	48.5	380.7	20.5

Pop. Change is the change due to population changes. Healthc. Change is the change due to changes in mean hospital expenditures. RH Error is the Red Herring error. RH Steep. is the change in hospital expenditures due to intertemporal changes in the Red Herring related estimation error. Estimates are based on the entire Danish population in 2002 and 2017. Amounts in 2020 dollars.

6 Discussion

Several potential mechanisms may underlie the observed Steepening effect and its interactions with the Red Herring Hypothesis, including technological advancements in hospital care (Chandra and Skinner, 2012; Laudicella et al., 2022), and increasing multimorbidity prevalence among older individuals (Rosella et al., 2018). However, shifts in hospital utilization and prices also ex-

plain the differences in hospital expenditures (Dieleman et al., 2025). Our analysis indicates that the extensive margin—measured by the number of months with hospital care utilization—plays a key role in the Steepening pattern, as older individuals are increasingly likely to use hospital expenditures within a given month as they grow older. This suggests that both treatment practice and the demand for care may be underlying mechanisms inducing Steepening. Prior research suggests that the Red Herring effect appears to be mediated by health status (Seshamani and Gray, 2004b; Polder et al., 2006; Howdon and Rice, 2018; Moore et al., 2017; Maynou et al., 2023). Indeed, Carreras et al. (2018) found that time-to-death becomes statistically insignificant when severity-based health status measures are incorporated into the analysis. While our findings illuminate the broader implications of traditional population aging metrics (age, time-to-death, and steepening), further research disentangling these health drivers of both Red Herring and Steepening effects is warranted, particularly with regard to sex-specific patterns and differences across healthcare settings.

A longstanding concern in the Red Herring literature relates to the plausible endogeneity between hospital expenditures and time-to-death (Salas and Raftery, 2001), which potentially influences our analysis. If hospital services successfully prolong life, then higher expenditures may be both a consequence and a cause of longer time-to-death. Several studies have addressed this issue partially using instrumental variable approaches (Felder et al., 2010; Karlsson and Klohn, 2011), with parental age at death serving as an instrument for time-to-death (Kolodziejczyk, 2020; Costa-Font and Vilaplana-Prieto, 2020). Others have treated past healthcare expenditures and time-to-death as predetermined exogenous factors influencing current time-to-death (Felder et al., 2010). However, the direct effect of hospital expenditures on time-to-death remains largely unexamined in this context. While endogeneity concerns have not been explicitly raised for the Steepening Hypothesis, they could still influence the combined Red Herring-Steepening framework discussed in this paper through similar mechanisms.

7 Conclusion

The healthcare expenditure effects of an increasing number of older individuals remain one of the most pressing challenges for healthcare systems globally (Christensen et al., 2009). Previous research has addressed this phenomenon through two distinct concepts: the Red Herring Hypothesis (Zweifel et al., 1999; Seshamani and Gray, 2004b) and the Steepening Hypothesis (Buchner and Wasem, 2006). This paper makes a significant contribution by investigating both hypotheses concurrently. Through a novel model parameterization and analysis of comprehensive register-based data capturing monthly hospital expenditures for the entire Danish population aged 30+, we provide robust evidence for the concurrence of both mechanisms.

Results from rigorous statistical testing demonstrate that neither hypothesis should be rejected independently, nor should their joint effect. We find substantial heterogeneity in the Red Herring and Steepening effects across age groups and time-to-death categories. As a novel contribution to this literature, our sex-stratified analysis demonstrates that the effects of population change differ across sexes, generally being more pronounced for males than females,

When analyzed together, the magnitude of both time-to-death and Steepening effects diminishes compared to when each is considered in isolation, suggesting important interactions between these phenomena at the three dimensional intersection of age, time-to-death, and time. Nevertheless, the empirical effects attributable to each hypothesis remain substantial. A novel decomposition analysis shows that from 2002 to 2017, Steepening effects accounted for 59.9% of the total increase in hospital expenditures, three times the size of the mitigating Red Herring effect. Furthermore, the decomposition identifies a previously unrecognized Red Herring-Steepening interaction effect, which contributed an additional 19.3% to the total expenditure increase. Overall, our findings emphasize that when the Red Herring and Steepening Hypotheses are analyzed together, their interaction reveals a more complex and concerning picture of hospital expenditure dynamics than much of the existing literature suggests.

References

- Atella, V. and Conti, V. (2014). The effect of age and time to death on primary care costs: The Italian experience. *Social Science & Medicine*, 114:10–17. [5](#), [9](#), [17](#), [20](#)
- Bank of England (2025). A Millennium of Macroeconomic Data for the UK: U.S. / U.K. Foreign Exchange Rate. Retrieved from FRED, Federal Reserve Bank of St. Louis. Original data from the Bank of England’s Three Centuries of Macroeconomic Data project. [17](#)
- Belotti, F., Deb, P., Manning, W. G., and Norton, E. C. (2015). twopm: Two-part models. *The Stata Journal*, 15(1):3–20. [5](#)
- Bjørner, T. B. and Arnberg, S. (2012). Terminal costs, improved life expectancy and future public health expenditure. *International Journal of Health Care Finance and Economics*, 12(2):129–143. [3](#), [8](#), [11](#)
- Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, pages 436–455. [10](#)
- Breyer, F. and Lorenz, N. (2021). The “red herring” after 20 years: Ageing and health care expenditures. *The European Journal of Health Economics*, 22(5):661–667. [3](#), [7](#), [27](#)
- Buchner, F. and Wasem, J. (2006). “Steeping” of health expenditure profiles. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 31(4):581–599. [3](#), [7](#), [9](#), [12](#), [28](#)
- Carreras, M., Ibern, P., and Inoriza, J. M. (2018). Ageing and healthcare expenditures: Exploring the role of individual health status. *Health Economics*, 27(5):865–876. [5](#), [9](#), [12](#), [28](#), [A15](#)
- Chandra, A. and Skinner, J. (2012). Technology growth and expenditure growth in health care. *Journal of Economic Literature*, 50(3):645–80. [27](#)
- Christensen, B. J., Gørtz, M., and Kallestrup-Lamb, M. (2016). Medical spending in Denmark. *Fiscal Studies*, 37(3-4):461–497. [13](#)
- Christensen, K., Doblhammer, G., Rau, R., and Vaupel, J. W. (2009). Ageing populations: The challenges ahead. *The Lancet*, 374(9696):1196–1208. [3](#), [28](#)
- Costa-Font, J. and Vilaplana-Prieto, C. (2020). ‘more than one red herring’? Heterogeneous effects of ageing on health care utilisation. *Health Economics*, 29:8–29. [28](#)
- Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica: Journal of the Econometric Society*, pages 829–844. [4](#)
- Danish Health Data Authority (2024). Dødsårsagsregisteret – Rapport 2024. Technical report, Sundhedsdatastyrelsen. Datagrundlag: Dødsårsagsregisteret pr. 30. juni 2024. [14](#)
- Danmarks Nationalbank (2024). Yearly exchange rate. Data retrieved on 05/02/2024. [13](#)

- Deb, P. and Norton, E. C. (2018). Modeling health care expenditures and use. *Annual Review of Public Health*, 39:489–505. 4
- Dieleman, J. L., Weil, M., Beauchamp, M., Bisignano, C., Crosby, S. W., DeJarnatt, D., Lescinsky, H., Mokdad, A. H., Ostroff, S., Paul, H., et al. (2025). Drivers of variation in health care spending across US counties. *JAMA Health Forum*, 6(2):e245220–e245220. 28
- Dorfman, R. (1938). A note on the δ -method for finding variance formulae. *Biometric Bulletin*. 3, 6
- Engle, R. F. (1984). Wald, likelihood ratio, and Lagrange multiplier tests in econometrics. *Handbook of Econometrics*, 2:775–826. 10, A1
- Federal Reserve (US) (2025a). South korean won to u.s. dollar spot exchange rate. H.10 Foreign Exchange Rates Release. Retrieved from FRED, Federal Reserve Bank of St. Louis, March 13, 2025. 17
- Federal Reserve (US) (2025b). U.s. dollars to euro spot exchange rate [dexuseu]. <https://fred.stlouisfed.org/series/DEXUSEU>. Retrieved from FRED, Federal Reserve Bank of St. Louis on March 13, 2025. 18
- Felder, S. and Werblow, A. (2008). Does the age profile of health care expenditure really steepen over time? New evidence from Swiss Cantons. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 33:710–727. 3, 7
- Felder, S., Werblow, A., and Zweifel, P. (2010). Do red herrings swim in circles? Controlling for the endogeneity of time to death. *Journal of Health Economics*, 29(2):205–212. 12, 28
- French, E. B., McCauley, J., Aragon, M., Bakx, P., Chalkley, M., Chen, S. H., Christensen, B. J., Chuang, H., Côté-Sergent, A., and De Nardi, M. (2017). End-of-life medical spending in last twelve months of life is lower than previously reported. *Health Affairs*, 36(7):1211–1217. 16
- Geue, C., Briggs, A., Lewsey, J., and Lorgelly, P. (2014). Population ageing and healthcare expenditure projections: New evidence from a time to death approach. *The European Journal of Health Economics*, 15(8):885–896. 3, 8, 11, 36, A15
- Gourieroux, C., Monfort, A., and Trognon, A. (1984). Pseudo maximum likelihood methods: Theory. *Econometrica*, pages 681–700. 5
- Greene, W. H. (2017). *Econometric Analysis*. Pearson, 7th edition. 18
- Gregersen, F. A. (2014). The impact of ageing on health care expenditures: A study of steepening. *The European Journal of Health Economics*, 15(9):979–989. 3, 7, 8, 9, 12, 18, 20, 21, A1, A15
- Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. In *Annals of Economic and Social Measurement, volume 5, number 4*, pages 475–492. NBER. 6
- Howdon, D. and Rice, N. (2018). Health care expenditures, age, proximity to death and morbidity: Implications for an ageing population. *Journal of Health Economics*, 57:60–74. 6, 28, A15

- Hyun, K.-R., Kang, S., and Lee, S. (2016). Population aging and healthcare expenditure in Korea. *Health Economics*, 25(10):1239–1251. [17](#), [A15](#)
- Insurance and Pension (FP) (2023). Sundhedsforsikringerne finansierer hovedsageligt, hvor der er egenbetaling på sundhedsområdet. Report authored by Sophie Orebo Wenzel. Published on 29.11.2023. [13](#)
- Kallestrup-Lamb, M., Marin, A. O., Menon, S., and Søgaard, J. (2024). Aging populations and expenditures on health. *The Journal of the Economics of Ageing*, page 100518. [3](#), [4](#), [20](#)
- Karlsson, M. and Klohn, F. (2011). Some notes on how to catch a red herring: Ageing, time-to-death & care costs for older people in Sweden. Technical report, Darmstadt Discussion Papers in Economics. [28](#)
- Karlsson, M. and Klohn, F. (2014). Testing the red herring hypothesis on an aggregated level: ageing, time-to-death and care costs for older people in sweden. *The European Journal of Health Economics*, 15(5):533–551. [21](#), [A15](#)
- Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795. [4](#), [10](#), [A2](#), [A3](#)
- Kitagawa, E. M. (1955). Components of a difference between two rates. *Journal of the American Statistical Association*, 50(272):1168–1194. [10](#)
- Kollerup, A., Kjellberg, J., and Ibsen, R. (2022). Ageing and health care expenditures: The importance of age per se, steepening of the individual-level expenditure curve, and the role of morbidity. *The European Journal of Health Economics*, pages 1–29. [3](#), [5](#), [6](#), [7](#), [8](#), [9](#), [13](#), [14](#), [20](#), [A15](#)
- Kolodziejczyk, C. (2020). The effect of time to death on health care expenditures: Taking into account the endogeneity and right censoring of time to death. *The European Journal of Health Economics*, 21:945–962. [28](#)
- Laudicella, M., Li Donni, P., Olsen, K. R., and Gyrd-Hansen, D. (2022). Age, morbidity, or something else? A residual approach using microdata to measure the impact of technological progress on health care expenditure. *Health Economics*, 31(6):1184–1201. [6](#), [27](#)
- Liang, K.-Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22. [10](#)
- Manning, W. G. and Mullahy, J. (2001). Estimating log models: To transform or not to transform? *Journal of Health Economics*, 20(4):461–494. [5](#)
- Maynou, L., Street, A., García, A., et al. (2023). Living longer in declining health: Factors driving healthcare costs among older people. *Social Science & Medicine*, 327:115955. [6](#), [28](#)
- Moore, P. V., Bennett, K., and Normand, C. (2017). Counting the time lived, the time left or illness? age, proximity to death, morbidity and prescribing expenditures. *Social Science & Medicine*, 184:1–14. [5](#), [7](#), [28](#)

- Norton, E. C. and Dowd, B. E. (2018). Log odds and the interpretation of Logit models. *Health Services Research*, 53(2):859–878. 6
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, pages 693–709. 10
- Polder, J. J., Barendregt, J. J., and van Oers, H. (2006). Health care costs in the last year of life—the dutch experience. *Social Science & Medicine*, 63(7):1720–1731. 3, 8, 11, 28
- Ricketts, T. C. (2011). The health care workforce: Will it be ready as the boomers age? A review of how we can know (or not know) the answer. *Annual Review of Public Health*, 32:417–430. 3
- Rosella, L., Kornas, K., Huang, A., Bornbaum, C., Henry, D., and Wodchis, W. P. (2018). Accumulation of chronic conditions at the time of death increased in Ontario from 1994 to 2013. *Health Affairs*, 37(3):464–472. 27
- Salas, C. and Raftery, J. P. (2001). Econometric issues in testing the age neutrality of health care expenditure. *Health Economics*, 10(7):669–671. 28
- Seshamani, M. and Gray, A. (2004a). Ageing and health-care expenditure: The red herring argument revisited. *Health Economics*, 13(4):303–314. 3, 4, 6, 8, 17, 20, 24
- Seshamani, M. and Gray, A. M. (2004b). A longitudinal study of the effects of age and time to death on hospital costs. *Journal of Health Economics*, 23(2):217–235. 5, 6, 14, 17, 21, 28
- Statens Serum Institut (2014). Takstsystem vejledning. 13
- Statistics Denmark (2023). Danish register data for research – access via statistics denmark. Accessed through Statistics Denmark’s research services. <https://www.dst.dk/en/TilSalg/Forskningsservice>. 12
- Statistics Denmark (2024a). Consumer price index. Data retrieved on 05/02/2024. 13
- Statistics Denmark (2024b). Statistikdokumentation for Befolkningen 2024. Technical report, Danmarks Statistik, Copenhagen, Danmark. 14
- Statistics Denmark (2025). SHA1: Udgifter til sundhed efter funktion, aktør, finansieringskilde og prisenhed. <https://www.statistikbanken.dk/SHA1>. Accessed: 2025-06-18. Enhed: Mio. kr. (løbende priser). 13
- Tanuseputro, P., Wodchis, W. P., Fowler, R., Walker, P., Bai, Y. Q., Bronskill, S. E., and Manuel, D. (2015). The health care cost of dying: A population-based retrospective cohort study of the last year of life in Ontario, Canada. *PLoS one*, 10(3):e0121759. 21
- van Baal, P. H. and Wong, A. (2012). Time to death and the forecasting of macro-level health care expenditures: Some further considerations. *Journal of Health Economics*, 31(6):876–887. 6, 8, 11, 35
- von Wyl, V. (2019). Proximity to death and health care expenditure increase revisited: A 15-year panel analysis of elderly persons. *Health Economics Review*, 9(1):1–16. 5, 8, 9

- Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54(3):426–482. 3
- Welch, B. L. (1947). The generalization of student's' problem when several different population variances are involved. *Biometrika*, 34(1/2):28–35. 20
- 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. 9, 12
- White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, pages 1–25. 5
- WHO (2024). Current health expenditure (% of gdp) - united states, euro area. Data retrieved from the WHO Global Health Expenditure database on April 15, 2024. Available at: apps.who.int/nha/database. 3
- 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. A15
- Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT press. A2
- Zweifel, P., Felder, S., and Meiers, M. (1999). Ageing of population and health care expenditure: A red herring? *Health Economics*, 8(6):485–496. 3, 6, 8, 9, 17, 28

A A Model for Decomposing Changes in Aggregate Hospital Expenditures

This section outlines our approach to decomposing changes in aggregate hospital expenditures between $t_2 = 2002$ and $t_1 = 2017$ into distinct components. These include: (i) the Naïve Model, which assumes expenditures vary only by age; (ii) the Red Herring Error, capturing the role of time-to-death in expenditure variation; (iii) the Steepening Hypothesis, which accounts for intertemporal changes in age-specific hospital expenditures; and (iv) a newly identified component, the Red Herring-Steepening effect, which isolates Steepening effects specifically within the last years of life.

Following [van Baal and Wong \(2012\)](#), let annual aggregate hospital expenditures, \mathcal{H} , in year, T , be expressed as the sum of mean hospital expenditures, \bar{H} , across groups of individuals at age, A , and time-to-death, TTD , categories

$$\mathcal{H}_T = \sum_A \sum_{TTD} N_{T,A,TTD} \bar{H}_{T,A,TTD}, \quad (11)$$

where $N_{T,A,TTD}$ presents the number of individuals of age. We consider ages as integers from 30 to 110+, with time-to-death rounded to the first decimal. A key advantage of this formulation is its reliance on directly observable data without imposing restrictive functional assumptions on the relationships between expenditures, age, time-to-death, or calendar time.

The change in aggregate hospital expenditures from t_2 to t_1 can be decomposed as follows:

$$\begin{aligned} \Delta_{t_1,t_2}^{\mathcal{H}} &:= \mathcal{H}_{t_1} - \mathcal{H}_{t_2} \\ &= \sum_A \sum_{TTD} N_{t_1,A,TTD} \bar{H}_{t_1,A,TTD} - N_{t_2,A,TTD} \bar{H}_{t_2,A,TTD} \\ &= \underbrace{\sum_A \sum_{TTD} [N_{t_1,A,TTD} - N_{t_2,A,TTD}] \bar{H}_{t_1,A,TTD}}_{\text{Population Change}} + \underbrace{\sum_A \sum_{TTD} [\bar{H}_{t_1,A,TTD} - \bar{H}_{t_2,A,TTD}] N_{t_2,A,TTD}}_{\text{Expenditure Change}}. \end{aligned} \quad (12)$$

The Population Change component isolates the impact of demographic shifts, $N_{t_1,A,TTD} - N_{t_2,A,TTD}$, while holding mean expenditures constant at their t_1 level, $\bar{H}_{t_1,A,TTD}$. If expenditures were unchanged between t_1 and t_2 , this term would fully explain the variation in total spending. However, when mean expenditures also evolve over time, individuals in t_2 ($N_{t_2,A,TTD}$) exhibit different levels of hospital consumption, generating the Expenditure Change component, which captures shifts in per capita spending, $\bar{H}_{t_1,A,TTD} - \bar{H}_{t_2,A,TTD}$.

While this two-part decomposition is attractive, it does not explicitly separate the contributions of the Red Herring Hypothesis and the Steepening Hypothesis. No prior work in the literature has provided such a decomposition. The Red Herring Hypothesis posits that hospital expenditures vary not just by age but also by proximity to death. If this were ignored, aggregate projections of hospital expenditures would mistakenly attribute time-to-death effects to aging alone. To disentangle these effects, we expand the Population Change term using a

time-to-death-insensitive expenditure measure, $\bar{H}_{T,A}$, yielding

$$\begin{aligned}
& \underbrace{\sum_A \sum_{TTD} [N_{t_1,A,TTD} - N_{t_2,A,TTD}] \bar{H}_{t_1,A,TTD}}_{\text{Population Change}} \\
&= \sum_A \sum_{TTD} [N_{t_1,A,TTD} - N_{t_2,A,TTD}] \left[\bar{H}_{t_1,A,TTD} + \bar{H}_{t_1,A} - \bar{H}_{t_1,A} \right] \\
&= \underbrace{\sum_A \sum_{TTD} [N_{t_1,A,TTD} - N_{t_2,A,TTD}] \bar{H}_{t_1,A}}_{\text{Naïve}} \\
&+ \underbrace{\sum_A \sum_{TTD} [N_{t_1,A,TTD} - N_{t_2,A,TTD}] \left[\bar{H}_{t_1,A,TTD} - \bar{H}_{t_1,A} \right]}_{\text{Red Herring Error}} \tag{13} \\
&= \underbrace{\sum_A [N_{t_1,A} - N_{t_2,A}] \bar{H}_{t_1,A}}_{\text{Naïve}} \\
&+ \underbrace{\sum_A \sum_{TTD} [N_{t_1,A,TTD} - N_{t_2,A,TTD}] \left[\bar{H}_{t_1,A,TTD} - \bar{H}_{t_1,A} \right]}_{\text{Red Herring Error}}
\end{aligned}$$

The Naïve component assumes expenditures depend solely on age, whereas the Red Herring Error quantifies the misattribution error, $\bar{H}_{t_1,A,TTD} - \bar{H}_{t_1,A}$, due to overlooking time-to-death effects. This decomposition provides a direct estimate of the impact of proximity to death on expenditure projections. Others have also quantified the Red Herring Error, but were unable to use exact decompositions of the effect, relying instead on predictions from regression models, see, e.g., [Geue et al. \(2014\)](#)

In addition to examining the influence of the Red Herring hypothesis on aggregate hospital expenditures, we also aim to understand intertemporal changes in mean hospital expenditures over time. Specifically, we assess the role of the Steepening Hypothesis in driving aggregate expenditure changes. Since the Steepening Hypothesis pertains solely to age-related changes in hospital expenditures over time, we isolate this effect by expanding the Expenditure Change component of equation (12). To do so, we introduce age- and time-specific hospital expenditures that are independent of time-to-death, denoted as $\bar{H}_{T,A}$

$$\begin{aligned}
& \underbrace{\sum_A \sum_{TTD} \left[\bar{H}_{t_1,A,TTD} - \bar{H}_{t_2,A,TTD} \right] N_{t_2,A,TTD}}_{\text{Expenditure Change}} \\
&= \sum_A \sum_{TTD} \left[\bar{H}_{t_1,A,TTD} + \left(\bar{H}_{t_1,A} - \bar{H}_{t_1,A} \right) - \bar{H}_{t_2,A,TTD} + \left(\bar{H}_{t_2,A} - \bar{H}_{t_2,A} \right) \right] N_{t_2,A,TTD} \\
&= \underbrace{\sum_A \sum_{TTD} \left[\bar{H}_{t_1,A} - \bar{H}_{t_2,A} \right] N_{t_2,A,TTD}}_{\text{Steepening effect}} + \underbrace{\sum_A \sum_{TTD} \left[\left(\bar{H}_{t_1,A,TTD} - \bar{H}_{t_1,A} \right) - \left(\bar{H}_{t_2,A,TTD} - \bar{H}_{t_2,A} \right) \right] N_{t_2,A,TTD}}_{\text{Red Herring-Steepening}} \\
&= \underbrace{\sum_A \left[\bar{H}_{t_1,A} - \bar{H}_{t_2,A} \right] N_{t_2,A}}_{\text{Steepening effect}} + \underbrace{\sum_A \sum_{TTD} \left[\left(\bar{H}_{t_1,A,TTD} - \bar{H}_{t_1,A} \right) - \left(\bar{H}_{t_2,A,TTD} - \bar{H}_{t_2,A} \right) \right] N_{t_2,A,TTD}}_{\text{Red Herring-Steepening}}
\end{aligned} \tag{14}$$

The first term, $\left[\bar{H}_{t_1,A} - \bar{H}_{t_2,A} \right] N_{t_2,A,TTD}$, captures Steepening-related changes in hospital expenditures. It represents the number of individuals, N , at age A , experiencing an age-specific hospital expenditure change of $\bar{H}_{t_1,A} - \bar{H}_{t_2,A}$ between t_1 and t_2 .

The second component of equation (14) introduces a new effect documented in this paper. This effect accounts for intertemporal changes in the time-to-death estimation error—specifically, how hospital expenditures evolve as time-to-death-related expenditures develop over time. The term $\bar{H}_{t_1,A,TTD} - \bar{H}_{t_1,A}$ represents the estimation error at time t_1 , while $\bar{H}_{t_2,A,TTD} - \bar{H}_{t_2,A}$ is its counterpart at t_2 . The difference, $\left(\bar{H}_{t_1,A,TTD} - \bar{H}_{t_1,A} \right) - \left(\bar{H}_{t_2,A,TTD} - \bar{H}_{t_2,A} \right)$, captures the time-dependent change in this error. We interpret this term as a Steepening-related development in time-to-death estimation errors, affecting $N_{t_2,A,TTD}$ individuals at age A with TTD years to death.

By integrating these components, we arrive at the final decomposition of aggregate hospital expenditure changes

$$\begin{aligned}
\Delta_{t_1,t_2}^{\mathcal{H}} &:= \mathcal{H}_{t_1} - \mathcal{H}_{t_2} \tag{15} \\
&= \underbrace{\sum_A \sum_{TTD} \left[N_{t_1,A,TTD} - N_{t_2,A,TTD} \right] \bar{H}_{t_1,A,TTD}}_{\text{Population Change}} + \underbrace{\sum_A \sum_{TTD} \left[\bar{H}_{t_1,A,TTD} - \bar{H}_{t_2,A,TTD} \right] N_{t_2,A,TTD}}_{\text{Expenditure Change}} \tag{16} \\
&= \underbrace{\sum_A \sum_{TTD} \left[N_{t_1,A,TTD} - N_{t_2,A,TTD} \right] \bar{H}_{t_1,A}}_{\text{Naïve}} + \underbrace{\sum_A \sum_{TTD} \left[N_{t_1,A,TTD} - N_{t_2,A,TTD} \right] \left[\bar{H}_{t_1,A,TTD} - \bar{H}_{t_1,A} \right]}_{\text{Red Herring Error}} \\
&+ \underbrace{\sum_A \sum_{TTD} \left[\bar{H}_{t_1,A} - \bar{H}_{t_2,A} \right] N_{t_2,A,TTD}}_{\text{Steepening effect}} + \underbrace{\sum_A \sum_{TTD} \left[\left(\bar{H}_{t_1,A,TTD} - \bar{H}_{t_1,A} \right) - \left(\bar{H}_{t_2,A,TTD} - \bar{H}_{t_2,A} \right) \right] N_{t_2,A,TTD}}_{\text{Red Herring-Steepening}} \tag{17}
\end{aligned}$$

A key advantage of our decomposition framework is its exact breakdown of aggregate hospital expenditure changes into distinct components widely discussed in the literature on hospital expenditures and population aging. Namely, a Naïve component, the Red Herring Error, the Steepening Effects, and the new Red Herring-Steepening effect. Notably, our approach does not

rely on any assumptions regarding population sizes, N , or mean expenditures, \bar{H} . It remains robust even when certain age groups are sparsely populated (e.g., extreme ages), ensuring its applicability across different demographic structures.

A.1 Proof of Equality Between the Red Herring Error and Red Herring-Steepening Terms

We show that the Red Herring Error term is exactly equal to the negative of the Red Herring-Steepening term in equation (15), such that their sum is zero.

Let $\bar{H}_{T,A} = \frac{1}{N_{T,A}} \sum_{TTD} \bar{H}_{T,A,TTD} N_{T,A,TTD}$ denote the average hospital expenditure at age A in period T . By construction, this implies the identity: $\bar{H}_{T,A} N_{T,A} = \sum_{TTD} \bar{H}_{T,A,TTD} N_{T,A,TTD}$. Hence, aggregate hospital expenditures can be expressed as: $\mathcal{H}_T = \sum_A \sum_{TTD} \bar{H}_{T,A,TTD} N_{T,A,TTD} = \sum_A \bar{H}_{T,A} N_{T,A}$.

The change in hospital expenditures between two periods t_1 and t_2 is then:

$$\begin{aligned} \Delta_{t_1, t_2}^{\mathcal{H}} &:= \mathcal{H}_{t_1} - \mathcal{H}_{t_2} = \sum_A \bar{H}_{t_1, A} N_{t_1, A} - \sum_A \bar{H}_{t_2, A} N_{t_2, A} \\ &= \underbrace{\sum_A (N_{t_1, A} - N_{t_2, A}) \bar{H}_{t_1, A}}_{\text{Naïve}} + \underbrace{\sum_A (\bar{H}_{t_1, A} - \bar{H}_{t_2, A}) N_{t_2, A}}_{\text{Steepening effect}}. \end{aligned} \quad (18)$$

On the other hand, equation (15) decomposes the same difference using a four-part breakdown that explicitly includes terms associated with time-to-death (TTD). Since both decompositions are mathematically equivalent, any residual terms not appearing in equation (18) must cancel. Specifically, the sum of the Red Herring Error term and the Red Herring-Steepening term in equation (15) must be zero. Therefore, the Red Herring Error term is equal to the negative of the Red Herring-Steepening term. ■

B Online Methodology Appendix

B.1 Steepening effects Despite Null Age-Time Interactions

In this section, we show that Steepening effects may exist event though interaction effects, δ_{AT} and γ_{AT} , are zero. Consider the full model $\psi = F \cdot g$, under the joint model specification where we suppress the dependence on linear arguments, $X'\delta$ and $X'\gamma$, in the Joint Model in equation (5). The second order derivative of ψ with respect to age, A , and calendar time, T , considered the Steepening effect (Gregersen, 2014), is

$$\begin{aligned}
 \partial^2 \psi / (\partial A \partial T) &= F'' \cdot [\delta_T + A\delta_{AT}] \cdot [\delta_A + T\delta_{AT}] \cdot g \\
 &\quad + F' \cdot \delta_{AT} \cdot g \\
 &\quad + F' \cdot [\delta_A + T\delta_{AT}] \cdot g' \cdot [\gamma_T + A\gamma_{AT}] \\
 &\quad + F' \cdot [\delta_T + A\delta_{AT}] \cdot g' \cdot [\gamma_A + T\gamma_{AT}] \\
 &\quad + F \cdot g'' \cdot [\gamma_T + A\gamma_{AT}] \cdot [\gamma_A + T\gamma_{AT}] \\
 &\quad + F \cdot g' \cdot \gamma_{AT},
 \end{aligned} \tag{19}$$

for $'$ and $''$ indicating a first and second order derivative, respectively, of a function.

In the case where the interaction terms are zero, $\delta_{AT} = \gamma_{AT} = 0$, the expression simplifies to

$$\begin{aligned}
 \partial^2 \psi / (\partial A \partial T) &= F'' \cdot \delta_T \cdot \delta_A \cdot g \\
 &\quad + F' \cdot \delta_A \cdot g' \cdot \gamma_T \\
 &\quad + F' \cdot \delta_T \cdot g' \cdot \gamma_A \\
 &\quad + F \cdot g'' \cdot \gamma_T \cdot \gamma_A.
 \end{aligned} \tag{20}$$

Although the age-time interaction parameters are zero, Steepening effect could persist, i.e., $\partial^2 \psi / (\partial A \partial T) > 0$, depending on the relative size of the remaining terms in equation (20). Similar arguments can be made for the first part function, F , and the second part function, g , where the second order derivatives are

$$\partial^2 F / (\partial A \partial T) = F'' \cdot [\delta_T + A\delta_{AT}] \cdot [\delta_A + T\delta_{AT}] + F' \cdot \delta_{AT} \tag{21}$$

$$\partial^2 g / (\partial A \partial T) = g'' \cdot [\gamma_T + A\gamma_{AT}] \cdot [\gamma_A + T\gamma_{AT}] + g' \cdot \gamma_{AT}. \tag{22}$$

Our approach for addressing these cases of steepening is to test the average marginal effects of the estimated models with the Delta Method and apply Wald tests of joint significance of the marginal effects as discussed in Section 2.2.

B.2 Robustness: Statistical Tests

Several statistical tests exist to test zero-restrictions on parameters of interest (Engle, 1984). So far, we have utilized the Wald test, to test the three hypotheses of Section 2.2. While a test on the parameter estimates themselves is of interest, much of our research has focused on the marginal association on more interpretable outcomes, i.e., the probability of using hospital care, and the numerical dollar-amount of hospital expenditures, as calculated with the Delta Method. To

assess the robustness of our findings towards this different outcome, we redefine the hypotheses to reflect the marginal association

$$\begin{aligned}
H_0^{me,1} &: \mathbb{E} \left[\frac{\partial f(H_i, \theta)}{\partial TTD} \quad \frac{\partial f(H_i, \theta)}{\partial T} \quad \frac{\partial f(H_i, \theta)}{\partial A \cdot T} \right] = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}, \\
H_0^{me,2} &: \mathbb{E} \left[\frac{\partial f(H_i, \theta)}{\partial A} \quad \frac{\partial f(H_i, \theta)}{\partial T} \quad \frac{\partial f(H_i, \theta)}{\partial A \cdot T} \right] = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}, \\
H_0^{me,3} &: \mathbb{E} \left[\frac{\partial f(H_i, \theta)}{\partial TTD} \right] = \begin{bmatrix} 0 \end{bmatrix},
\end{aligned} \tag{23}$$

where the function, f , which can be (i) the joint model, ψ , (ii) the first part, F , or (iii) the second part, g , in equation (1). As each of the marginal associations from the Delta Method are Gaussian under maximum likelihood estimation, a Wald test can also be used to test the hypotheses in equation (7) (Wooldridge, 2010, Chapter 3).

Traditional hypothesis testing, e.g., the Wald test, has a tendency to reject most null-hypotheses in large samples due to small p-value, i.e., the probability of observing data given a hypothesis, $Pr(D | H_0)$. To add robustness on this issue, we employ a second testing framework to the hypotheses in equation (6) which does not share this feature. This approach focuses on the probability of the null hypothesis, H_0 , against the alternative, H_A , given the data, D , i.e., $Pr(H_0 | D)$. These p-values may align more closely with a researchers testing objective by evaluation evidence in favor of a null hypothesis. As advocated by Kass and Raftery (1995), such *Bayesian p-values* can be derived from Bayes formula and require Bayes factors and a prior odds ratio for calculation

$$Pr(H_0 | D) = \frac{\frac{Pr(D|H_0)}{Pr(D|H_A)} \times \frac{Pr(H_0)}{Pr(H_A)}}{1 + \frac{Pr(D|H_0)}{Pr(D|H_A)} \times \frac{Pr(H_0)}{Pr(H_A)}} = \frac{\text{Bayes factor} \times \text{prior odds}}{1 + \text{Bayes factor} \times \text{prior odds}}. \tag{24}$$

Bayes factors can be interpreted as likelihood ratios, i.e., how good the model fits under the null compared to the alternative. The prior odds signifies the researchers prior beliefs regarding the null hypothesis compared to its alternative, prior odds = $Pr(H_0) / Pr(H_A)$. A critique of the Bayesian p-values are their dependence on subjective beliefs about the prior odds. Specifically, the prior odds could be hacked ex-post if a certain p-value is desired. A common practice is, however, to assess the robustness of the Bayesian p-values resulting from various prior odds levels. We derive the derivation of equation (24) and the estimation of the Bayes factor in Appendix B.3.

We apply these two tests to the full sample estimates from Section 4, and report the results in Table A1. The p-values of the Wald-tests on the marginal effects are close to zero and reject all null-hypotheses at conventional significance-levels.²¹ The Bayesian p-value tests provides a similar conclusion.²² with p-values of zero offering no support for the null hypotheses.

²¹We note that H_0^3 has a p-value of 0.005 which is larger than the remaining p-values. An investigation reveals that relatively large marginal effects of time-to-death on the Poisson part of the model is the source of the low test-statistic.

²²For the Bayesian tests we set a prior belief in the null of 20%. Due to the size of the Bayes factors, the Bayesian p-values remain zero even if we raise our beliefs in the null as much as up to 99%.

Table A1: Hypothesis test results

	Wald Test Marginal Effect			Bayesian Test Parameter		
	#Res.	Test Statistic	P-value	#Res.	log Bayes Factor	P-value
H_0^1	3	164,426	0.000	6	-84,883,437	0.000
H_0^2	3	8,081	0.000	6	-2,349,166	0.000
H_0^3	1	8	0.005	2	-84,613,510	0.000

Hypothesis tests and p-values. #Res. refers to the number of parameter restrictions. In the Bayesian test we set a 20% prior belief in the null.

B.3 Testing with Bayes Factor

This section discusses testing of our hypothesis in Section 2.2, but rather than using the classical p-values signifying the probability of observing data, D , given a null hypothesis, H_0 , we consider the reverse, i.e., the probability of the null hypothesis given data. To this end, we take point of departure in [Kass and Raftery \(1995\)](#) concerning Bayes Factors.

Considering a null hypothesis, H_0 , and its alternative, H_A , we obtain from Bayes's theorem

$$\frac{Pr(H_0 | D)}{Pr(H_A | D)} = \frac{Pr(D | H_0) Pr(H_0)}{Pr(D | H_A) Pr(H_A)} \quad (25)$$

posterior odds = Bayes factor \times prior odds,

where Bayes factor, informally, has the interpretation of the likelihood ration of the two hypothesis and the prior odds ratio signifies the prior relative belief towards the null and its alternative of the analyst. Note the restrictions that $1 = Pr(H_0 | D) + Pr(H_A | D)$ and $1 = Pr(H_0) + Pr(H_A)$. Using these relationships and reworking equation (25) we obtain an expression for the *Bayesian p-values*

$$Pr(H_0 | D) = \frac{\text{Bayes factor} \times \text{prior odds}}{1 + \text{Bayes factor} \times \text{prior odds}}. \quad (26)$$

the Bayesian p-values increases in Bayes factor and a higher likelihood of data stemming from a model under the null hypothesis increases evidence in favor of the null hypothesis. Reversely a low Bayes factor implies poor relative evidence in favor of the null, i.e., the alternative fits data better.

In large samples, Bayes factor is easily approximated as the exponential of the Schwarz criterion, S , ([Kass and Raftery, 1995](#)). Let $\hat{\ell}_{H_0}$ and $\hat{\ell}_{H_A}$ be the maximized log-likelihood functions under the null and its alternative, respectively, and $|\cdot|$ be the length of its argument, then

$$S = \hat{\ell}_{H_0} - \hat{\ell}_{H_A} - \frac{1}{2} (|\theta_0| - |\theta_A|) \log(N), \quad (27)$$

which we use to calculate the Bayes p-values.

C Online Results Appendix

C.1 Full Parameter Estimates

Table A2 presents the complete regression results corresponding to Section 4. Unlike Table 1, this table includes additional explanatory variables: the intercept, a male dummy, month dummies, and an indicator for whether time-to-death exceeds four years. Additionally, it reports the marginal effect on the probability of hospital expenditure utilization (E_F) and the extensive margin conditional on expenditure use (E_g).

Table A2: GLM model estimates

Model:	Naïve					Red Herring					Steepening					Joint Model				
	Probit	Poisson	E_F	E_g	E_ψ	Probit	Poisson	E_F	E_g	E_ψ	Probit	Poisson	E_F	E_g	E_ψ	Probit	Poisson	E_F	E_g	E_ψ
Intercept	-1.545 (0.569)	5.864 (0.008)	-0.444 (0.001)	5,313.659 (4.170)	773.987 (527.275)	-1.434 (0.541)	6.486 (0.005)	-0.410 (0.004)	5,877.241 (4.499)	945.912 (540.010)	-1.556 (0.585)	5.873 (0.010)	-0.445 (0.004)	5,321.162 (5.278)	774.038 (529.684)	-1.430 (0.555)	6.518 (0.005)	-0.406 (0.008)	5,905.710 (4.173)	955.257 (542.393)
Male	-0.131 (0.035)	0.127 (0.003)	-0.038 (0.004)	115.303 (3.021)	-7.691 (8.840)	-0.137 (0.036)	0.073 (0.003)	-0.039 (0.004)	66.231 (2.458)	-17.873 (3.731)	-0.133 (0.035)	0.128 (0.003)	-0.038 (0.004)	116.132 (3.053)	-7.942 (8.809)	-0.139 (0.037)	0.073 (0.003)	-0.040 (0.004)	66.316 (2.472)	-18.250 (3.628)
Age	0.014 (0.004)	0.017 (0.000)	0.004 (0.000)	15.465 (0.083)	6.944 (1.890)	0.012 (0.003)	0.003 (0.000)	0.003 (0.000)	3.002 (0.026)	3.430 (0.517)	0.011 (0.003)	0.018 (0.000)	0.004 (0.000)	15.577 (0.083)	6.882 (1.896)	0.008 (0.002)	0.003 (0.000)	0.003 (0.000)	3.045 (0.027)	3.324 (0.504)
TTD	-	-	-	-	-	-0.101 (0.019)	-0.470 (0.002)	-0.029 (0.005)	-425.949 (2.469)	-117.003 (42.194)	-	-	-	-	-	-0.103 (0.020)	-0.470 (0.002)	-0.029 (0.005)	-425.715 (2.447)	-117.407 (41.915)
Time	-	-	-	-	-	-	-	-	-	-	0.005 (0.003)	-0.002 (0.000)	0.007 (0.001)	-8.448 (0.212)	4.416 (0.255)	0.004 (0.003)	-0.004 (0.000)	0.007 (0.001)	-1.722 (0.151)	5.715 (0.266)
Age·Time	-	-	-	-	-	-	-	-	-	-	0.033 (0.006)	-0.011 (0.001)	0.011 (0.001)	7.028 (0.163)	11.435 (1.659)	0.037 (0.006)	0.003 (0.000)	0.010 (0.001)	1.354 (0.145)	9.147 (0.785)
TTD>4	-	-	-	-	-	0.593 (0.121)	1.912 (0.006)	0.169 (0.028)	1,732.858 (7.499)	519.551 (176.522)	-	-	-	-	-	0.619 (0.128)	1.911 (0.006)	0.176 (0.028)	1,731.149 (7.392)	524.220 (175.611)
Feb	0.004 (0.001)	-0.182 (0.006)	0.001 (0.000)	-165.054 (4.868)	-35.050 (15.439)	0.005 (0.001)	-0.178 (0.005)	0.001 (0.000)	-161.524 (4.478)	-34.227 (13.816)	0.002 (0.001)	-0.181 (0.006)	0.001 (0.000)	-164.063 (4.866)	-35.369 (15.367)	0.002 (0.001)	-0.178 (0.005)	0.001 (0.000)	-161.372 (4.476)	-34.714 (13.749)
Mar	0.032 (0.009)	-0.107 (0.006)	0.009 (0.001)	-97.192 (5.748)	-13.186 (8.619)	0.033 (0.009)	-0.105 (0.006)	0.009 (0.001)	-95.208 (5.766)	-13.112 (7.713)	0.028 (0.008)	-0.106 (0.006)	0.008 (0.001)	-95.656 (5.726)	-13.832 (8.531)	0.029 (0.008)	-0.105 (0.006)	0.008 (0.001)	-94.906 (5.765)	-14.018 (7.699)
Apr	0.023 (0.006)	-0.200 (0.008)	0.007 (0.001)	-181.627 (7.336)	-34.008 (16.499)	0.023 (0.007)	-0.196 (0.008)	0.007 (0.001)	-177.252 (7.336)	-33.251 (14.654)	0.017 (0.005)	-0.198 (0.008)	0.005 (0.000)	-179.357 (7.305)	-34.983 (16.383)	0.017 (0.005)	-0.195 (0.008)	0.005 (0.000)	-176.847 (7.331)	-34.604 (14.625)
May	0.041 (0.012)	-0.166 (0.005)	0.012 (0.001)	-150.510 (4.833)	-22.571 (13.817)	0.042 (0.012)	-0.160 (0.005)	0.012 (0.001)	-145.127 (4.467)	-21.854 (12.424)	0.034 (0.010)	-0.163 (0.005)	0.010 (0.001)	-147.697 (4.833)	-23.858 (13.662)	0.034 (0.010)	-0.160 (0.005)	0.010 (0.001)	-144.585 (4.450)	-23.607 (12.427)
Jun	0.048 (0.013)	-0.180 (0.004)	0.014 (0.001)	-162.966 (3.434)	-23.496 (15.242)	0.049 (0.014)	-0.174 (0.004)	0.014 (0.001)	-157.459 (3.439)	-22.825 (13.718)	0.039 (0.011)	-0.176 (0.004)	0.011 (0.001)	-159.221 (3.410)	-25.098 (15.046)	0.039 (0.011)	-0.173 (0.004)	0.011 (0.001)	-156.730 (3.413)	-25.048 (13.735)
Jul	0.022 (0.006)	-0.241 (0.006)	0.006 (0.001)	-218.463 (5.040)	-42.226 (20.565)	0.023 (0.007)	-0.242 (0.005)	0.006 (0.001)	-218.872 (4.444)	-42.496 (18.945)	0.010 (0.003)	-0.236 (0.006)	0.003 (0.000)	-214.210 (5.079)	-44.259 (20.360)	0.011 (0.003)	-0.241 (0.005)	0.003 (0.000)	-218.012 (4.458)	-45.201 (18.955)
Aug	0.055 (0.016)	-0.203 (0.006)	0.016 (0.001)	-183.896 (5.132)	-26.378 (16.714)	0.056 (0.016)	-0.201 (0.005)	0.016 (0.001)	-182.185 (4.724)	-26.635 (15.215)	0.041 (0.012)	-0.197 (0.006)	0.012 (0.001)	-178.674 (5.174)	-28.722 (16.454)	0.042 (0.012)	-0.200 (0.005)	0.012 (0.001)	-181.174 (4.733)	-29.839 (15.274)
Sep	0.074 (0.020)	-0.165 (0.006)	0.021 (0.002)	-149.856 (5.266)	-14.102 (12.688)	0.075 (0.020)	-0.160 (0.006)	0.022 (0.002)	-145.079 (5.170)	-13.900 (11.273)	0.059 (0.016)	-0.159 (0.006)	0.017 (0.002)	-144.145 (5.317)	-16.754 (12.412)	0.059 (0.016)	-0.159 (0.006)	0.017 (0.002)	-143.963 (5.185)	-17.494 (11.368)
Okt	0.074 (0.020)	-0.150 (0.005)	0.021 (0.002)	-136.130 (4.012)	-11.118 (11.549)	0.075 (0.020)	-0.148 (0.005)	0.021 (0.002)	-133.917 (4.023)	-11.489 (10.472)	0.057 (0.015)	-0.143 (0.005)	0.016 (0.002)	-129.609 (4.069)	-14.108 (11.241)	0.057 (0.016)	-0.146 (0.005)	0.016 (0.002)	-132.638 (4.038)	-15.548 (10.600)
Nov	0.080 (0.021)	-0.136 (0.009)	0.023 (0.002)	-123.210 (7.705)	-6.928 (9.389)	0.080 (0.022)	-0.135 (0.008)	0.023 (0.002)	-121.971 (7.111)	-7.633 (8.642)	0.060 (0.016)	-0.128 (0.009)	0.017 (0.002)	-115.928 (7.800)	-10.205 (9.059)	0.061 (0.016)	-0.133 (0.008)	0.017 (0.002)	-120.505 (7.163)	-12.097 (8.798)
Dec	0.041 (0.011)	-0.182 (0.003)	0.012 (0.001)	-164.713 (2.935)	-25.630 (15.092)	0.042 (0.012)	-0.181 (0.003)	0.012 (0.001)	-163.762 (2.783)	-25.843 (13.760)	0.020 (0.006)	-0.173 (0.003)	0.006 (0.001)	-156.697 (2.987)	-29.284 (14.735)	0.020 (0.006)	-0.179 (0.003)	0.006 (0.001)	-162.167 (2.800)	-30.776 (13.920)
DoF	9,999,972					9,999,968					9,999,968					9,999,964				
Log Lik.	-3,196,833,760					-2,906,760,856					-3,194,966,130					-2,906,648,274				

Note: TTD is short for time-to-death. Standard errors in parenthesis. P-values in squared brackets. Models are estimated with 10,000,000 observations and 2,191,490 in the Poisson part. E_F denote to the marginal probability of using hospital expenditures, E_g is the marginal amount used conditional on using any hospital expenditures while E_ψ refers to the marginal effects of the full two-part model. Age·Time is divided by 100 in the Probit and Poisson regressions but not in E_F , E_g , or E_ψ . Amounts in 2020 dollars.

C.2 Robustness by Sex

We examine whether the Red Herring and Steepening hypotheses differ by sex. Figures 1.(a) and 1.(b) illustrate minor variations in hospital expenditures across sexes. To investigate this further, Table A3 presents the full parameter estimates from the joint model in equation (5), including results from both the Probit and Poisson regressions. It also reports the marginal associations for (i) the probability of using hospital expenditures (E_F), (ii) the conditional amount spent (E_g), and (iii) the overall two-part model (E_ψ). Additionally, the table includes estimates for the intercept, month dummies, and an indicator for whether time-to-death exceeds four years.

Overall, the parameter estimates for males and females in Table A3 are largely consistent with the main population estimates in Table 1. Age remains positively associated with the probability of using hospital expenditures, as indicated by Probit estimates of 0.004 for females (column 1) and 0.015 for males (column 6). Regarding time-to-death, being one year further from death reduces hospital expenditures, conditional on usage, with parameter estimates of -0.454 for females (column 2) and -0.484 for males (column 7). For the Steepening Hypothesis, the age-time interaction term is positive in the Probit model for both sexes, suggesting a Steepening effect. However, while males also exhibit a positive and significant estimate in the Poisson model, the parameter for females is insignificant. This suggests that steepening is evident on both margins for males but is absent on the extensive margin for females.

To formally assess whether our main findings hold across sexes, we replicate the three hypothesis tests from equation (6) using sex-disaggregated data (Table A4). The results confirm the main conclusions in Table 2: all three hypotheses are rejected for both sexes at the conventional 5% significance level. As in the main results, the Wald test for the marginal effects under hypothesis H_0^3 exhibits the highest p-value (2.1% for males), largely due to the high variance in time-to-death estimates.

A key question in health economics is whether population aging affects males and females differently, particularly given sex-specific differences in mortality rates. If changes in mortality rates vary by sex, hospital expenditure growth could also differ. To test whether age, time-to-death, and their interaction have equal associations for both sexes, we conduct a Welch test Welch on the parameter estimates in Table A3. The results, presented in Table A5, yield p-values below 5% for all marginal associations and key regressors. This indicates that we cannot reject the null hypothesis that the marginal effects are identical across sexes. Thus, at a population level, the association between population aging and hospital expenditures appears to be similar for males and females.

Table A4: Hypothesis tests of hypothesis in equation (6) by sex

Sex	Hypothesis	Wald Test Parameter		Bayesian Test Parameter		Wald Test Marginal Effect			
	#Res.	Test Statistic	P-value	log Bayes Factor	P-value	#Res.	Test Statistic	P-value	
Female	H_0^1	6	1,776,411	0.000000	-37,683,552	0.000000	3	8,352	0.000000
	H_0^2	6	497,093	0.000000	-1,311,301	0.000000	3	5,690	0.000000
	H_0^3	2	169,884	0.000000	-37,464,562	0.000000	1	6	0.0115288
Male	H_0^1	6	3,515,548	0.000000	-47,334,375	0.000000	3	3,810	0.000000
	H_0^2	6	522,947	0.000000	-1,024,901	0.000000	3	1,328	0.000000
	H_0^3	2	2,280,650	0.000000	-47,261,128	0.000000	1	5	0.0206969

Note: Hypothesis tests statistics and p-values. #Res. refers to the number of parameter restrictions. In the Bayesian test we set a 50% prior belief in the null.

Table A3: Estimates of Joint Model by Sex

	Female					Male				
	Probit	Poisson	E_F	E_g	E_ψ	Probit	Poisson	E_F	E_g	E_ψ
Intercept	-1.220 (0.281) [0.000]	6.475 (0.164) [0.000]	-0.372 (0.011) [0.000]	5,473.67 (293.255) [0.000]	1,021.79 (340.653) [0.003]	-1.881 (0.290) [0.000]	6.648 (0.074) [0.000]	-0.488 (0.051) [0.000]	6,542.20 (147.030) [0.000]	858.62 (371.777) [0.021]
Age	0.004 (0.000) [0.000]	0.003 (0.002) [0.032]	0.002 (0.000) [0.000]	2.80 (1.291) [0.030]	2.24 (0.760) [0.003]	0.014 (0.000) [0.000]	0.003 (0.001) [0.023]	0.005 (0.001) [0.000]	3.26 (1.560) [0.037]	4.75 (1.307) [0.000]
TTD	-0.092 (0.006) [0.000]	-0.454 (0.082) [0.000]	-0.028 (0.005) [0.000]	-383.79 (58.775) [0.000]	-114.45 (45.303) [0.012]	-0.117 (0.007) [0.000]	-0.484 (0.085) [0.000]	-0.030 (0.006) [0.000]	-476.58 (99.734) [0.000]	-120.14 (51.931) [0.021]
Age · Time	0.025 (0.002) [0.000]	-0.001 (0.001) [0.328]	0.010 (0.002) [0.000]	0.30 (0.453) [0.514]	8.10 (1.405) [0.000]	0.049 (0.004) [0.000]	0.009 (0.004) [0.045]	0.011 (0.002) [0.000]	2.67 (1.480) [0.071]	10.36 (2.442) [0.000]
Time	0.011 (0.000) [0.000]	-0.002 (0.001) [0.114]	0.008 (0.002) [0.000]	-2.50 (0.916) [0.006]	5.79 (0.628) [0.000]	-0.005 (0.001) [0.000]	-0.006 (0.003) [0.036]	0.006 (0.001) [0.000]	-0.68 (0.318) [0.034]	5.47 (1.093) [0.000]
$\mathbb{1}\{\text{TTD} > 4\}$	0.536 (0.032) [0.000]	1.913 (0.474) [0.000]	0.163 (0.033) [0.000]	1,617.17 (354.790) [0.000]	518.36 (222.148) [0.020]	0.697 (0.041) [0.000]	1.908 (0.433) [0.000]	0.181 (0.036) [0.000]	1,878.03 (489.042) [0.000]	527.35 (232.996) [0.024]
Feb	0.006 (0.002) [0.000]	-0.158 (0.083) [0.058]	0.002 (0.000) [0.000]	-133.63 (66.651) [0.045]	-30.63 (25.910) [0.237]	-0.002 (0.001) [0.164]	-0.199 (0.089) [0.025]	-0.001 (0.000) [0.037]	-196.20 (94.335) [0.038]	-39.00 (29.669) [0.189]
Mar	0.032 (0.003) [0.000]	-0.082 (0.033) [0.014]	0.010 (0.002) [0.000]	-69.45 (26.352) [0.008]	-8.97 (10.581) [0.397]	0.025 (0.000) [0.000]	-0.129 (0.060) [0.033]	0.006 (0.002) [0.000]	-126.82 (63.748) [0.047]	-19.35 (18.527) [0.296]
Apr	0.015 (0.001) [0.000]	-0.170 (0.083) [0.042]	0.005 (0.001) [0.000]	-143.47 (66.521) [0.031]	-30.78 (26.078) [0.238]	0.020 (0.001) [0.000]	-0.222 (0.096) [0.021]	0.005 (0.001) [0.000]	-218.96 (102.020) [0.032]	-38.59 (31.640) [0.223]
May	0.032 (0.002) [0.000]	-0.132 (0.056) [0.018]	0.010 (0.002) [0.000]	-111.48 (43.804) [0.011]	-19.05 (17.621) [0.280]	0.037 (0.000) [0.000]	-0.189 (0.082) [0.020]	0.010 (0.003) [0.000]	-186.44 (86.591) [0.031]	-28.32 (25.766) [0.272]
Jun	0.041 (0.002) [0.000]	-0.154 (0.064) [0.015]	0.012 (0.003) [0.000]	-130.15 (50.043) [0.009]	-21.43 (20.171) [0.288]	0.038 (0.001) [0.000]	-0.193 (0.074) [0.009]	0.010 (0.002) [0.000]	-190.23 (79.408) [0.017]	-28.83 (24.683) [0.243]
Jul	0.009 (0.001) [0.000]	-0.215 (0.102) [0.035]	0.003 (0.001) [0.000]	-181.58 (80.826) [0.025]	-41.52 (32.434) [0.201]	0.013 (0.001) [0.000]	-0.268 (0.116) [0.020]	0.003 (0.001) [0.001]	-264.18 (122.737) [0.031]	-49.04 (38.579) [0.204]
Aug	0.040 (0.003) [0.000]	-0.182 (0.087) [0.037]	0.012 (0.002) [0.000]	-153.76 (69.560) [0.027]	-27.22 (26.710) [0.308]	0.043 (0.001) [0.000]	-0.219 (0.093) [0.018]	0.011 (0.003) [0.000]	-215.83 (98.697) [0.029]	-32.63 (29.688) [0.272]
Sep	0.065 (0.004) [0.000]	-0.124 (0.054) [0.023]	0.020 (0.004) [0.000]	-104.45 (42.971) [0.015]	-9.42 (15.889) [0.553]	0.053 (0.001) [0.000]	-0.197 (0.078) [0.011]	0.014 (0.003) [0.000]	-194.12 (82.794) [0.019]	-26.05 (24.748) [0.292]
Okt	0.061 (0.003) [0.000]	-0.116 (0.060) [0.052]	0.019 (0.004) [0.000]	-98.47 (47.981) [0.040]	-8.94 (16.593) [0.590]	0.053 (0.002) [0.000]	-0.179 (0.083) [0.032]	0.014 (0.003) [0.000]	-175.91 (87.933) [0.045]	-22.50 (25.035) [0.369]
Nov	0.065 (0.004) [0.000]	-0.112 (0.044) [0.010]	0.020 (0.004) [0.000]	-94.71 (34.247) [0.006]	-6.96 (12.928) [0.590]	0.056 (0.001) [0.000]	-0.155 (0.075) [0.038]	0.014 (0.003) [0.000]	-153.00 (78.911) [0.053]	-17.42 (21.564) [0.419]
Dec	0.017 (0.002) [0.000]	-0.132 (0.055) [0.017]	0.005 (0.001) [0.000]	-111.73 (43.589) [0.010]	-22.78 (18.329) [0.214]	0.024 (0.000) [0.000]	-0.230 (0.110) [0.037]	0.006 (0.002) [0.000]	-226.50 (115.990) [0.051]	-39.16 (34.608) [0.258]
N	5,125,417	1,231,906	5,125,417	1,231,906	5,125,417	4,874,583	959,584	4,874,583	959,584	4,874,583
log Lik.										

E_F denote to the marginal probability of using hospital expenditures, E_g is the marginal amount used conditional on using any hospital expenditures while E_ψ refers to the marginal effects of the full two-part model. Standard errors in parenthesis. P-values in squared brackets. TTD is short for time-to-death. Age·Time is divided by 100 in the Probit and Poisson regressions but not in E_F , E_g , or E_ψ . Amounts in 2020 dollars.

Table A5: P-values of Welch Test of Similar Estimates for Females and Males

	E_F	E_g	E_ψ
Age	0.007	0.822	0.097
TTD	0.798	0.423	0.934
Age·Time	0.724	0.125	0.422

Note: Welch test of similar parameter estimates based on estimated in Table A3 for the full population of females and males, respectively. E_F denote to the marginal probability of using hospital expenditures, E_g is the marginal amount used conditional on using any hospital expenditures while E_ψ refers to the marginal effects of the full two-part model.

C.3 Robustness by Age and Sex: Tests

Table A6 shows the tests on marginal associations from Section B.2 for each age-sex-group. Table A7 performs the Wald test of Section 2.2 and the Bayesian p-value calculations from Section B.2. All tests reject the null hypotheses providing evidence in favor of the joint model as the alternative hypothesis including both time-to-death and Steepening effects.

Table A8 show Welch tests of equal means parameter estimates between two age-sex-groups for the three marginal effects, E_F , E_g , and E_ψ , respectively. The parameter estimates are as reported in Table 3. Lower diagonal elements of the tables hold the tests for the time-to-death parameter while the upper diagonal elements report the Welch test for the age-time interaction. Small p-values indicated that the null hypothesis of equal parameter estimates is strongly rejected, indicating that each group has a unique effect. Adjacent age-groups is more often not rejected. Table A9 perform the similar Welch tests but compares the parameter estimates between sexes for each age group. For most parameter estimates, females and males differ significantly, suggesting that population aging impacts genders in distinct ways.

Table A6: Hypothesis tests of hypothesis in equation (7) by Age and Sex

Ages	Hypothesis		Wald Test, E_F		Wald Test, E_g		Wald Test, E_ψ		
	#Res.	DoF	Test Stat.	P-value	Test Stat.	P-value	Test Stat.	P-value	
Female									
[30;40)	H_0^1	3	1,036,240	537.940	0.000000	14,310.151	0.000000	75,930.582	0.000000
	H_0^2	3	1,036,240	168.989	0.000000	1,052.755	0.000000	75,080.338	0.000000
	H_0^3	1	1,036,240	127.536	0.000000	775.019	0.000000	121.870	0.000000
[40;50)	H_0^1	3	1,107,315	1,039.280	0.000000	10,483.417	0.000000	7,213.309	0.000000
	H_0^2	3	1,107,315	271.779	0.000000	116.103	0.000000	7,539.340	0.000000
	H_0^3	1	1,107,315	101.100	0.000000	977.815	0.000000	104.820	0.000000
[50;60)	H_0^1	3	1,041,267	981.961	0.000000	53,678.173	0.000000	11,510.002	0.000000
	H_0^2	3	1,041,267	331.410	0.000000	218.302	0.000000	12,932.345	0.000000
	H_0^3	1	1,041,267	121.549	0.000000	25,143.413	0.000000	82.469	0.000000
[60;70)	H_0^1	3	912,024	1,144.813	0.000000	613.369	0.000000	6,793.449	0.000000
	H_0^2	3	912,024	427.169	0.000000	275.364	0.000000	1,443.997	0.000000
	H_0^3	1	912,024	158.383	0.000000	601.445	0.000000	53.610	0.000000
[70;80)	H_0^1	3	610,720	923.526	0.000000	865.385	0.000000	4,270.132	0.000000
	H_0^2	3	610,720	1,184.193	0.000000	512.779	0.000000	1,506.373	0.000000
	H_0^3	1	610,720	239.697	0.000000	179.026	0.000000	44.934	0.000000
[80;90)	H_0^1	3	337,241	1,700.001	0.000000	7,189.324	0.000000	11,773.575	0.000000
	H_0^2	3	337,241	1,238.208	0.000000	2,511.840	0.000000	11,074.921	0.000000
	H_0^3	1	337,241	402.685	0.000000	145.778	0.000000	50.413	0.000000
90+	H_0^1	3	80,372	7,839.253	0.000000	8,566.295	0.000000	26,915.946	0.000000
	H_0^2	3	80,372	8,980.677	0.000000	7,323.404	0.000000	7,321.818	0.000000
	H_0^3	1	80,372	1,635.790	0.000000	327.514	0.000000	125.564	0.000000
Male									
[30;40)	H_0^1	3	1,056,512	271.730	0.000000	20,210.133	0.000000	11,154.165	0.000000
	H_0^2	3	1,056,512	208.034	0.000000	302.235	0.000000	7,035.333	0.000000
	H_0^3	1	1,056,512	89.836	0.000000	767.329	0.000000	58.104	0.000000
[40;50)	H_0^1	3	1,133,044	436.138	0.000000	73,924.226	0.000000	4,776.713	0.000000
	H_0^2	3	1,133,044	228.542	0.000000	197.512	0.000000	7,619.332	0.000000
	H_0^3	1	1,133,044	84.943	0.000000	17,298.578	0.000000	44.020	0.000000
[50;60)	H_0^1	3	1,049,863	882.230	0.000000	9,801.600	0.000000	10,711.800	0.000000
	H_0^2	3	1,049,863	354.013	0.000000	35.313	0.000001	9,169.842	0.000000
	H_0^3	1	1,049,863	107.597	0.000000	1,117.930	0.000000	37.088	0.000000
[60;70)	H_0^1	3	883,136	1,302.186	0.000000	18,368.798	0.000000	5,226.744	0.000000
	H_0^2	3	883,136	412.629	0.000000	113.877	0.000000	5,576.746	0.000000
	H_0^3	1	883,136	148.848	0.000000	246.752	0.000000	35.645	0.000000
[70;80)	H_0^1	3	519,988	1,678.999	0.000000	1,184.672	0.000000	38,271.921	0.000000
	H_0^2	3	519,988	1,886.165	0.000000	549.226	0.000000	1,001.603	0.000000
	H_0^3	1	519,988	304.198	0.000000	159.484	0.000000	49.138	0.000000
[80;90)	H_0^1	3	204,559	1,835.428	0.000000	1,165.655	0.000000	15,055.872	0.000000
	H_0^2	3	204,559	5,092.486	0.000000	1,115.840	0.000000	1,827.501	0.000000
	H_0^3	1	204,559	772.172	0.000000	207.242	0.000000	87.015	0.000000
90+	H_0^1	3	27,243	11,264.858	0.000000	8,205.288	0.000000	97,796.000	0.000000
	H_0^2	3	27,243	9,837.948	0.000000	10,810.903	0.000000	41,467.054	0.000000
	H_0^3	1	27,243	4,967.937	0.000000	1,117.005	0.000000	472.675	0.000000

Note: Wald test of marginal association by age and sex with null hypotheses in equation (7) based on the age- and sex-specific estimates in Table 3. #Res. refers to the number of parameter restrictions and DoF is the degrees of freedom.

Table A7: Hypothesis Tests of Age and Sex Specific Estimates

Ages	Hypothesis		Wald Test		Bayes P-value	
	#Res.	Test Stat.	P-value	Test Stat.	P-value	
Female						
[30;40)	H_0^1	6	227,188	0.000000	-550,367	0.000000
	H_0^2	6	199,989	0.000000	-108,886	0.000000
	H_0^3	2	17,225	0.000000	-534,808	0.000000
[40;50)	H_0^1	6	302,189	0.000000	-1,681,246	0.000000
	H_0^2	6	77,010	0.000000	-10,187	0.000000
	H_0^3	2	189,637	0.000000	-1,678,097	0.000000
[50;60)	H_0^1	6	2,503,715	0.000000	-4,799,431	0.000000
	H_0^2	6	251,567	0.000000	-44,148	0.000000
	H_0^3	2	183,711	0.000000	-4,781,440	0.000000
[60;70)	H_0^1	6	1,125,613	0.000000	-8,629,363	0.000000
	H_0^2	6	18,949	0.000000	-23,793	0.000000
	H_0^3	2	350,321	0.000000	-8,611,648	0.000000
[70;80)	H_0^1	6	3,207,222	0.000000	-11,795,648	0.000000
	H_0^2	6	84,520	0.000000	-16,466	0.000000
	H_0^3	2	1,004,722	0.000000	-11,766,660	0.000000
[80;90)	H_0^1	6	1,415,500	0.000000	-9,053,124	0.000000
	H_0^2	6	533,393	0.000000	-135,431	0.000000
	H_0^3	2	711,582	0.000000	-8,870,034	0.000000
90+	H_0^1	6	1,954,533	0.000000	-2,002,936	0.000000
	H_0^2	6	144,568	0.000000	-64,413	0.000000
	H_0^3	2	203,080	0.000000	-1,907,422	0.000000
Male						
[30;40)	H_0^1	6	120,456	0.000000	-427,175	0.000000
	H_0^2	6	94,158	0.000000	-98,405	0.000000
	H_0^3	2	32,477	0.000000	-349,514	0.000000
[40;50)	H_0^1	6	685,141	0.000000	-1,513,077	0.000000
	H_0^2	6	171,119	0.000000	-133,759	0.000000
	H_0^3	2	81,655	0.000000	-1,489,476	0.000000
[50;60)	H_0^1	6	1,857,623	0.000000	-5,531,109	0.000000
	H_0^2	6	515,335	0.000000	-44,705	0.000000
	H_0^3	2	1,322,665	0.000000	-5,519,590	0.000000
[60;70)	H_0^1	6	1,415,140	0.000000	-13,112,830	0.000000
	H_0^2	6	412,511	0.000000	-28,891	0.000000
	H_0^3	2	487,213	0.000000	-13,108,051	0.000000
[70;80)	H_0^1	6	7,700,791	0.000000	-16,484,131	0.000000
	H_0^2	6	204,314	0.000000	-213,487	0.000000
	H_0^3	2	2,836,682	0.000000	-16,401,831	0.000000
[80;90)	H_0^1	6	2,021,666	0.000000	-9,949,476	0.000000
	H_0^2	6	117,942	0.000000	-94,051	0.000000
	H_0^3	2	462,470	0.000000	-9,917,107	0.000000
90+	H_0^1	6	6,159,660	0.000000	-1,567,170	0.000000
	H_0^2	6	1,518,007	0.000000	-43,961	0.000000
	H_0^3	2	1,668,405	0.000000	-1,516,911	0.000000

Note: Wald test from Section 2.2 and Bayesian p-values from Section B.2 of parameter estimates, $\hat{\theta}$ by age and sex. Tests are based on the age- and sex-specific estimates in Table 3. #Res. refers to the number of parameter restrictions and DoF is the degrees of freedom.

Table A8: P-values of Equal Means Welch Tests for Time-to-Death (lower diagonal) and Age-Time Interaction (upper diagonal) by Age and Sex Groups

		A·T, Female							A·T, Male						
		[30;40]	[40;50]	[50;60]	[60;70]	[70;80]	[80;90]	90+	[30;40]	[40;50]	[50;60]	[60;70]	[70;80]	[80;90]	90+
Panel A: Extensive Margin, $\Pr(H>0)$, E_F															
TTD	[30;40]		0.000	0.000	0.000	0.000	0.000	0.000	0.028	0.000	0.000	0.000	0.009	0.000	
	[40;50]	0.306		0.048	0.000	0.001	0.000	0.000	0.000	0.003	0.000	0.000	0.494	0.000	
	[50;60]	0.080	0.518		0.067	0.278	0.000	0.000	0.000	0.003	0.003	0.000	0.000	0.000	
	[60;70]	0.090	0.618	0.844		0.310	0.000	0.000	0.000	0.452		0.900	0.000	0.000	
	[70;80]	0.764	0.166	0.028	0.027		0.000	0.000	0.000	0.729	0.578		0.000	0.000	
	[80;90]	0.000	0.000	0.000	0.000	0.000		0.000	0.161	0.015	0.000	0.000		0.000	
	90+	0.000	0.000	0.000	0.000	0.000	0.000		0.467	0.000	0.000	0.000		0.000	
Panel B: Intensive Margin, $H H>0$, E_g															
TTD	[30;40]		0.000	0.000	0.000	0.000	0.000	0.388	0.000	0.000	0.000	0.000	0.000	0.000	
	[40;50]	0.277		0.156	0.443	0.417	0.042	0.000	0.000	0.025	0.069	0.000	0.000	0.000	
	[50;60]	0.000	0.000		0.028	0.032	0.004	0.000	0.000		0.634	0.000	0.000	0.000	
	[60;70]	0.000	0.000	0.037		0.806	0.062	0.000	0.000	0.021		0.000	0.000	0.000	
	[70;80]	0.000	0.000	0.000	0.041		0.100	0.000	0.000	0.001	0.288		0.483	0.000	
	[80;90]	0.000	0.000	0.000	0.006	0.516		0.000	0.000	0.000	0.000	0.054	0.476		0.000
	90+	0.000	0.000	0.000	0.000	0.072	0.315		0.000	0.000	0.000	0.000	0.002	0.012	
Panel C: Full Model, E_ψ															
TTD	[30;40]		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000	
	[40;50]	0.968		0.000	0.000	0.000	0.797	0.000	0.008	0.001	0.000	0.000	0.000	0.000	
	[50;60]	0.048	0.049		0.001	0.001	0.037	0.000	0.000	0.009		0.000	0.259	0.000	0.000
	[60;70]	0.003	0.003	0.133		0.293	0.000	0.000	0.000	0.000	0.043		0.000	0.000	0.000
	[70;80]	0.000	0.000	0.008	0.173		0.000	0.000	0.000	0.000	0.000	0.098		0.000	0.000
	[80;90]	0.000	0.000	0.006	0.166	0.973		0.000	0.000	0.000	0.000	0.015	0.599		0.000
	90+	0.000	0.000	0.012	0.632	0.260	0.254		0.000	0.000	0.000	0.000	0.002	0.541	0.971

Note: Welch test of equal parameter estimates by age and sex groups for estimates in Table 3. P-values for time-to-death on lower diagonal elements and the age-time interaction are reported in upper diagonal elements.

Table A9: P-values of Equal Means Welch Tests for Females and Males by Age Group

	[30;40)	[40;50)	[50;60)	[60;70)	[70;80)	[80;90)	90+
Panel A: Extensive Margin, $\Pr(H>0)$, E_F							
Age	0.000	0.124	0.000	0.000	0.000	0.290	0.505
TTD	0.000	0.000	0.037	0.165	0.828	0.000	0.000
Age·Time	0.000	0.311	0.901	0.144	0.002	0.000	0.000
Panel B: Intensive Margin, $H H>0$, E_g							
Age	0.000	0.000	0.239	0.562	0.000	0.000	0.000
TTD	0.000	0.190	0.004	0.014	0.175	0.119	0.000
Age·Time	0.000	0.000	0.765	0.001	0.000	0.000	0.000
Panel C: Full Model, E_ψ							
Age	0.000	0.006	0.002	0.000	0.000	0.000	0.000
TTD	0.000	0.016	0.764	0.448	0.233	0.045	0.000
Age·Time	0.000	0.585	0.154	0.000	0.317	0.000	0.082

Note: Welch test of equal parameter estimates for females and males by age groups for estimates in Table 3.

C.4 Robustness by Time-to-death and Sex

Table A11 presents the results of Wald tests for the hypotheses outlined in Section B.2. Table A12 performs the Wald test of Section 2.2 and the Bayesian p-value calculations from Section B.2. When considering individuals with $TTD < 4$, we exclude the zero-restrictions on time-to-death in the hypotheses. Across all cases, we reject the hypotheses, providing strong evidence in favor of the alternative hypothesis, i.e., the joint model and the coexistence of the Red Herring and Steepening Hypotheses.

Table A10 examines whether the marginal effects differ across time-to-death groups for each sex using Welch tests. A p-value of zero indicates a rejection of the null hypothesis, confirming that the effects are significantly different between groups. However, for some comparisons, p-values exceed conventional significance levels of 5%, suggesting that differences may not be statistically meaningful. Notably, for males, the Steepening effect ($A \cdot T$) on the extensive margin does not differ significantly between the [0, 1.5) and 4+ time-to-death groups, as indicated by a high p-value of 69.1%. This implies that we cannot reject the null hypothesis that these groups exhibit identical Steepening effects. Table A13 perform the similar Welch tests but compares the parameter estimates between sexes for each age group. For most parameter estimates, females and males differ significantly, suggesting that population aging impacts them in distinct ways.

Table A10: P-values of Equal Means Welch Tests for Time-to-Death (lower diagonal) and Age-Time Interaction (upper diagonal) by Time-to-death and Sex Groups

		A·T, Female			A·T, Male		
		[0;1.5)	[1.5;4)	4+	[0;1.5)	[1.5;4)	4+
Panel A: Extensive Margin, $\Pr(H>0)$, E_F							
TTD	[0; 1.5)		0.000	0.000		0.000	0.691
	[1.5; 4)	0.000		0.071	0.000		0.090
Panel B: Intensive Margin, $H H>0$, E_g							
TTD	[0; 1.5)		0.398	0.000		0.000	0.000
	[1.5; 4)	0.000		0.000	0.000		0.000
Panel C: Full Model, E_ψ							
TTD	[0; 1.5)		0.000	0.000		0.000	0.000
	[1.5; 4)	0.000		0.000	0.000		0.000

Note: Welch test of equal parameter estimates by age and sex groups for estimates in Table 4. P-values for time-to-death on lower diagonal elements and the age-time interaction are reported in upper diagonal elements.

Table A11: Hypothesis Tests of Hypothesis in Equation (7) by Time-to-death and Sex

Group	Hypothesis		Wald Test, E_F		Wald Test, E_g		Wald Test, E_ψ		
	#Res.	DoF	Test Stat.	P-value	Test Stat.	P-value	Test Stat.	P-value	
Female									
TTD $\in [0;1.5)$	H_0^1	3	115,224	13,833.874	0.000000	271.134	0.000000	3,451.547	0.000000
	H_0^2	3	115,224	15,568.198	0.000000	286.944	0.000000	5,107.139	0.000000
	H_0^3	1	115,224	1,481.035	0.000000	250.342	0.000000	137.683	0.000000
TTD $\in [1.5;4)$	H_0^1	3	189,126	3,033.856	0.000000	1,577.919	0.000000	9,539.253	0.000000
	H_0^2	3	189,126	2,393.684	0.000000	662.802	0.000000	7,036.283	0.000000
	H_0^3	1	189,126	730.559	0.000000	145.214	0.000000	76.673	0.000000
TTD > 4	H_0^1	2	4,820,973	45.680	0.000000	64.677	0.000000	15,459.218	0.000000
	H_0^2	3	4,820,973	736.895	0.000000	347.148	0.000000	31,191.826	0.000000
	H_0^3	NA	NA	NA	NA	NA	NA	NA	NA
Male									
TTD $\in [0;1.5)$	H_0^1	3	110,507	5,989.685	0.000000	6,017.486	0.000000	2,358.600	0.000000
	H_0^2	3	110,507	6,088.672	0.000000	5,653.367	0.000000	657.069	0.000000
	H_0^3	1	110,507	1,914.655	0.000000	250.019	0.000000	147.159	0.000000
TTD $\in [1.5;4)$	H_0^1	3	185,619	2,159.883	0.000000	714.859	0.000000	8,355.548	0.000000
	H_0^2	3	185,619	2,851.027	0.000000	1,116.048	0.000000	37,239.445	0.000000
	H_0^3	1	185,619	882.331	0.000000	178.660	0.000000	92.700	0.000000
TTD > 4	H_0^1	2	4,578,363	80.817	0.000000	7.477	0.0581444	7,705.467	0.000000
	H_0^2	3	4,578,363	3,548.025	0.000000	89.428	0.000000	7,932.998	0.000000
	H_0^3	NA	NA	NA	NA	NA	NA	NA	NA

Note: Wald test of marginal association by age and sex with null hypotheses in equation (7) based on the time-to-death- and sex-specific estimates in Table 4. #Res. refers to the number of parameter restrictions and DoF is the degrees of freedom.

Table A12: Hypothesis Tests of Age and Time-to-death Specific Estimates

Age Group	Hypothesis		Wald Test		Bayes P-value	
	#Res.	DoF	Test Statistic	P-value	Test Statistic	P-value
Female						
TTD $\in [0;1.5)$	H_0^1	6	1,127,764	0.000000	-21,430,154	0.000000
	H_0^2	6	6,063,297	0.000000	-5,292,636	0.000000
	H_0^3	2	218,270	0.000000	-21,091,715	0.000000
TTD $\in [1.5;4)$	H_0^1	6	3,416,159	0.000000	-843,593	0.000000
	H_0^2	6	11,924,339	0.000000	-1,326,430	0.000000
	H_0^3	2	389,942	0.000000	-668,344	0.000000
TTD > 4	H_0^1	4	118,728	0.000000	-489,283	0.000000
	H_0^2	6	5,280,376	0.000000	-6,563,522	0.000000
	H_0^3	NA	NA	NA	NA	NA
Male						
TTD $\in [0;1.5)$	H_0^1	6	2,225,400	0.000000	-29,634,778	0.000000
	H_0^2	6	7,516,677	0.000000	-4,833,063	0.000000
	H_0^3	2	12,963	0.000000	-28,818,441	0.000000
TTD $\in [1.5;4)$	H_0^1	6	1,208,504	0.000000	-1,109,253	0.000000
	H_0^2	6	4,394,037	0.000000	-1,959,967	0.000000
	H_0^3	2	145,172	0.000000	-938,478	0.000000
TTD > 4	H_0^1	4	210,291	0.000000	-489,257	0.000000
	H_0^2	6	5,648,288	0.000000	-7,416,481	0.000000
	H_0^3	NA	NA	NA	NA	NA

Note: Wald test from Section 2.2 and Bayesian p-values from Section B.2 of parameter estimates, $\hat{\theta}$ by time-to-death and sex. Tests are based based on the time-to-death- and sex-specific estimates in Table 4. #Res. refers to the number of parameter restrictions and DoF is the degrees of freedom.

Table A13: P-values of Equal Means Welch Tests for Females and Males by Time-to-death Group

	[0;1.5)	[1.5;4)	4+
Panel A: Extensive Margin, $\Pr(H>0)$, E_F			
Age	0.000	0.124	0.000
TTD	0.000	0.000	0.037
Age·Time	0.000	0.311	0.901
Panel B: Intensive Margin, $H H>0$, E_g			
Age	0.000	0.000	0.239
TTD	0.000	0.190	0.004
Age·Time	0.000	0.000	0.765
Panel C: Full Model, E_ψ			
Age	0.000	0.006	0.002
TTD	0.000	0.016	0.764
Age·Time	0.000	0.585	0.154

Note: Welch test of equal parameter estimates for females and males by age groups for estimates in Table 4.

C.5 Robustness to Expanded Parametrization

A potential concern with our testing of the coexistence of Red Herring and Steepening effects is that misspecification of the functional form may lead to biased inference. Our Joint Model, specified in equation (5), adopts a parsimonious structure centered on key regressors—age, time-to-death, and an age-time interaction. To assess the robustness of our findings under more flexible specifications, we re-estimate the models in Sections 4, 4.1, and 4.2 using an expanded parameterization. Following established practice in the literature, we augment the Joint Model with several regressors. These include: age squared (Carreras et al., 2018; Howdon and Rice, 2018); a binary indicator being one in the month of death, D , and its interactions with age (Carreras et al., 2018; Hyun et al., 2016) and calendar time (Gregersen, 2014; Kollerup et al., 2022); age interacted with time-to-death (Geue et al., 2014; Karlsson and Klohn, 2014); and time-to-death squared (Wong et al., 2011). This results in the following Expanded Joint Model

$$\begin{aligned}
 \text{Expanded Joint Model: } \quad X_i \delta^R &= \delta_{TTD}^R TTD_i + \delta_{TTD^2}^R TTD_i^2 + \delta_{TTD \cdot A}^R TTD_i A_i + \delta_A^R A_i + \delta_{A^2}^R A_i^2 \\
 &\quad + \delta_T^R T_i + \delta_{A \cdot T}^R A_i T_i + \delta_{D \cdot T}^R D_i T_i + \delta_D^R D_i + \delta_{A \cdot D}^R A_i D_i + \delta^R \check{X}_i^R \\
 X_i \gamma^R &= \gamma_{TTD}^R TTD_i + \gamma_{TTD^2}^R TTD_i^2 + \gamma_{TTD \cdot A}^R TTD_i A_i + \gamma_A^R A_i + \gamma_{A^2}^R A_i^2 \\
 &\quad + \gamma_T^R T_i + \gamma_{A \cdot T}^R A_i T_i + \gamma_{D \cdot T}^R D_i T_i + \gamma_D^R D_i + \gamma_{A \cdot D}^R A_i D_i + \gamma^R \check{X}_i^R.
 \end{aligned} \tag{28}$$

Specifically, we augment the model with: Age squared (Carreras et al., 2018; Howdon and Rice, 2018); an indicator for the final time-period (month) of death, D , and its interaction with age (Carreras et al., 2018; Hyun et al., 2016); age interacted with time-to-death (Geue et al., 2014; Karlsson and Klohn, 2014); calendar time interacted with a death-indicator (Gregersen, 2014; Kollerup et al., 2022); and time-to-death squared (Wong et al., 2011). Equation (28) presents the expanded parametrization of the Probit and Poisson models

This expanded parametrization modifies the Jacobian and Hessian matrices of the maximum likelihood estimates, which in turn affect the calculation of cluster-robust standard errors used in hypothesis testing. While we incorporate these changes into our estimation procedure, we omit the explicit expressions for brevity.

The null hypotheses for the Wald tests under this expanded model are redefined as follows:

$$\begin{aligned}
 H_0^{1.R}: \quad &\delta_{TTD} = \gamma_{TTD} = \delta_{TTD^2} = \gamma_{TTD^2} = \delta_{TTD \cdot A} = \gamma_{TTD \cdot A} = \delta_T = \gamma_T \\
 &= \delta_{A \cdot T} = \gamma_{A \cdot T} = \delta_{D \cdot T} = \gamma_{D \cdot T} = \delta_D = \gamma_D = \delta_{A \cdot D} = \gamma_{A \cdot D} = 0, \\
 H_0^{2.R}: \quad &\delta_A = \gamma_A = \delta_{A^2} = \gamma_{A^2} = \delta_T = \gamma_T = \delta_{A \cdot T} = \gamma_{A \cdot T} = 0, \\
 H_0^{3.R}: \quad &\delta_{TTD} = \gamma_{TTD} = \delta_{TTD^2} = \gamma_{TTD^2} = \delta_{TTD \cdot A} = \gamma_{TTD \cdot A} \\
 &= \delta_{D \cdot T} = \gamma_{D \cdot T} = \delta_D = \gamma_D = \delta_{A \cdot D} = \gamma_{A \cdot D} = 0.
 \end{aligned} \tag{29}$$

We treat the month-of-death indicator D and its interactions as capturing Red Herring effects, and therefore constrain these terms to zero in $H_0^{1.R}$ and $H_0^{3.R}$.

Using the hypotheses in equation (29), we perform Wald tests and compute Bayesian p-values for: (i) the full population (Table A14); (ii) the age- and sex-disaggregated analyses (Table A16); and (iii) the time-to-death and sex analyses (Table A18).

We also reassess the marginal associations from equation (7) using the Delta Method for the full two-part model E_ψ , the extensive margin E_F , and the intensive margin E_g . Table A15 reports results for the full population, with Tables A17 and A19 reporting results for the age/sex and time-to-death/sex analyses, respectively.

In conclusion, the results strongly reject the null hypotheses in all cases. This confirms the simultaneous presence of Red Herring and Steepening effects, even under the more flexible functional form. Therefore, our main findings are robust to alternative model specifications.

Table A14: Hypothesis Tests of Hypothesis in Equation (29) under Expanded Joint Model

Hypothesis	Wald Test			Bayesian Test	
	#Res.	Test Statistic	P-value	log Bayes Factor	P-value
H_0^1	16	1,118,767	0.0000000	-328,067,347	0.0000000
H_0^2	8	48,749	0.0000000	-7,831,510	0.0000000
H_0^3	12	814,038	0.0000000	-326,240,933	0.0000000

Note: Hypothesis tests statistics and p-values. #Res. refers to the number of parameter restrictions. In the Bayesian test we set a 50% prior belief in the null.

Table A15: Hypothesis Tests of Hypothesis in Equation (7) under Expanded Joint Model

Hypothesis	Wald Test, E_F				Wald Test, E_g		Wald Test, E_ψ	
	#Res.	DoF	Test Stat.	P-value	Test Stat.	P-value	Test Stat.	P-value
H_0^1	3	9,999,952	346,768.317	0.0000000	3,039,579.914	0.0000000	118,084.354	0.0000000
H_0^2	3	9,999,952	92,528.466	0.0000000	29,031.123	0.0000000	27,188.324	0.0000000
H_0^3	1	9,999,952	21,394.488	0.0000000	2,959,613.833	0.0000000	30,890.132	0.0000000

Note: Wald test of marginal association by age and sex with null hypotheses in equation (7) based on the time-to-death- and sex-specific estimates in the Expanded Joint Model. #Res. refers to the number of parameter restrictions and DoF is the degrees of freedom.

Table A16: Hypothesis Tests of Age and Sex Specific Estimates under Expanded Joint Model

Ages	Hypothesis		Wald Test		Bayes P-value	
	#Res.	Test Stat.	P-value	Test Stat.	P-value	
Female						
[30;40)	H_0^1	16	625,110	0.000000	-4,432,753	0.000000
	H_0^2	8	267,949	0.000000	-109,286	0.000000
	H_0^3	12	208,661	0.000000	-4,408,259	0.000000
[40;50)	H_0^1	16	910,055	0.000000	-13,806,931	0.000000
	H_0^2	8	265,047	0.000000	-22,810	0.000000
	H_0^3	12	577,317	0.000000	-13,767,837	0.000000
[50;60)	H_0^1	16	3,288,241	0.000000	-30,680,584	0.000000
	H_0^2	8	715,628	0.000000	-86,115	0.000000
	H_0^3	12	1,494,231	0.000000	-30,564,493	0.000000
[60;70)	H_0^1	16	2,443,508	0.000000	-47,107,619	0.000000
	H_0^2	8	165,068	0.000000	-78,510	0.000000
	H_0^3	12	2,102,277	0.000000	-46,859,294	0.000000
[70;80)	H_0^1	16	14,111,603	0.000000	-39,080,544	0.000000
	H_0^2	8	411,176	0.000000	-48,459	0.000000
	H_0^3	12	7,341,043	0.000000	-38,848,315	0.000000
[80;90)	H_0^1	16	13,295,886	0.000000	-21,849,346	0.000000
	H_0^2	8	589,433	0.000000	-167,940	0.000000
	H_0^3	12	12,244,362	0.000000	-821,545,574	0.000000
90+	H_0^1	16	15,558,576	0.000000	-4,137,558	0.000000
	H_0^2	8	629,804	0.000000	-79,414	0.000000
	H_0^3	12	11,091,372	0.000000	-4,017,525	0.000000
Male						
[30;40)	H_0^1	16	451,512	0.000000	-3,126,854	0.000000
	H_0^2	8	101,273	0.000000	-130,180	0.000000
	H_0^3	12	180,004	0.000000	-3,015,781	0.000000
[40;50)	H_0^1	16	1,213,452	0.000000	-12,138,141	0.000000
	H_0^2	8	246,400	0.000000	-133,759	0.000000
	H_0^3	12	592,620	0.000000	-12,032,795	0.000000
[50;60)	H_0^1	16	5,654,526	0.000000	-33,033,256	0.000000
	H_0^2	8	912,401	0.000000	-70,631	0.000000
	H_0^3	12	3,493,114	0.000000	-32,890,146	0.000000
[60;70)	H_0^1	16	6,400,946	0.000000	-56,041,409	0.000000
	H_0^2	8	802,795	0.000000	-42,454	0.000000
	H_0^3	12	3,895,920	0.000000	-55,899,972	0.000000
[70;80)	H_0^1	16	16,587,405	0.000000	-50,661,974	0.000000
	H_0^2	8	160,869	0.000000	-113,685	0.000000
	H_0^3	12	9,874,884	0.000000	-50,227,724	0.000000
[80;90)	H_0^1	16	17,166,450	0.000000	-23,349,059	0.000000
	H_0^2	8	123,795	0.000000	-42,295	0.000000
	H_0^3	12	11,668,831	0.000000	-23,278,599	0.000000
90+	H_0^1	16	17,657,174	0.000000	-2,832,435	0.000000
	H_0^2	8	1,420,596	0.000000	-40,389	0.000000
	H_0^3	12	14,435,009	0.000000	-2,773,288	0.000000

Note: Wald test from equation (29) and Bayesian p-values from Section B.2 of parameter estimates, $\hat{\theta}$ by age and sex from the Expanded Joint Model. #Res. refers to the number of parameter restrictions.

Table A17: Hypothesis Tests of Hypothesis in equation (7) by Age and Sex under Expanded Joint Model

Ages	Hypothesis		Wald Test, E_F		Wald Test, E_g		Wald Test, E_ψ		
	#Res.	DoF	Test Stat.	P-value	Test Stat.	P-value	Test Stat.	P-value	
Female									
[30;40)	H_0^1	3	1,036,228	2,118.141	0.000000	41,288.498	0.000000	143,823.488	0.000000
	H_0^2	3	1,036,228	223.948	0.000000	1,473.026	0.000000	54,711.307	0.000000
	H_0^3	1	1,036,228	1,582.922	0.000000	14,197.605	0.000000	14,704.945	0.000000
[40;50)	H_0^1	3	1,107,303	6,399.278	0.000000	53,320.083	0.000000	42,751.564	0.000000
	H_0^2	3	1,107,303	465.923	0.000000	291.910	0.000000	4,962.143	0.000000
	H_0^3	1	1,107,303	1,779.493	0.000000	42,919.724	0.000000	29,976.391	0.000000
[50;60)	H_0^1	3	1,041,255	2,408.184	0.000000	50,888.055	0.000000	60,977.516	0.000000
	H_0^2	3	1,041,255	351.239	0.000000	252.543	0.000000	17,721.526	0.000000
	H_0^3	1	1,041,255	2,185.619	0.000000	44,447.489	0.000000	37,109.120	0.000000
[60;70)	H_0^1	3	912,012	3,924.966	0.000000	170,682.218	0.000000	37,085.658	0.000000
	H_0^2	3	912,012	529.292	0.000000	965.505	0.000000	9,799.301	0.000000
	H_0^3	1	912,012	3,314.338	0.000000	167,819.359	0.000000	18,652.898	0.000000
[70;80)	H_0^1	3	610,708	5,276.992	0.000000	3,614,664.421	0.000000	97,711.373	0.000000
	H_0^2	3	610,708	1,309.773	0.000000	640.359	0.000000	1,524.370	0.000000
	H_0^3	1	610,708	4,971.440	0.000000	3,102,102.022	0.000000	34,390.849	0.000000
[80;90)	H_0^1	3	337,229	3,503.570	0.000000	1,771,828.371	0.000000	207,744.482	0.000000
	H_0^2	3	337,229	1,289.984	0.000000	1,333.324	0.000000	10,034.419	0.000000
	H_0^3	1	337,229	3,364.055	0.000000	1,302,267.971	0.000000	181,120.534	0.000000
90+	H_0^1	3	80,360	11,537.804	0.000000	461,392.766	0.000000	114,792.760	0.000000
	H_0^2	3	80,360	18,759.534	0.000000	14,507.675	0.000000	4,221.451	0.000000
	H_0^3	1	80,360	6,754.282	0.000000	453,783.059	0.000000	114,317.975	0.000000
Male									
[30;40)	H_0^1	3	1,056,500	1,014.451	0.000000	51,861.849	0.000000	31,201.542	0.000000
	H_0^2	3	1,056,500	441.090	0.000000	1,912.917	0.000000	9,598.894	0.000000
	H_0^3	1	1,056,500	776.931	0.000000	35,483.666	0.000000	16,525.549	0.000000
[40;50)	H_0^1	3	1,133,032	4,488.343	0.000000	98,795.724	0.000000	17,911.900	0.000000
	H_0^2	3	1,133,032	344.971	0.000000	746.480	0.000000	9,918.820	0.000000
	H_0^3	1	1,133,032	400.352	0.000000	39,391.042	0.000000	5,800.959	0.000000
[50;60)	H_0^1	3	1,049,851	3,838.470	0.000000	479,845.930	0.000000	38,440.381	0.000000
	H_0^2	3	1,049,851	359.506	0.000000	165.977	0.000000	10,657.866	0.000000
	H_0^3	1	1,049,851	1,857.579	0.000000	347,911.177	0.000000	18,816.454	0.000000
[60;70)	H_0^1	3	883,124	4,600.196	0.000000	4,899,210.744	0.000000	23,364.738	0.000000
	H_0^2	3	883,124	961.886	0.000000	364.596	0.000000	9,452.976	0.000000
	H_0^3	1	883,124	4,561.055	0.000000	1,974,072.205	0.000000	20,091.659	0.000000
[70;80)	H_0^1	3	519,976	20,743.832	0.000000	1,111,247.113	0.000000	202,675.625	0.000000
	H_0^2	3	519,976	1,416.525	0.000000	2,377.588	0.000000	22,721.047	0.000000
	H_0^3	1	519,976	6,995.211	0.000000	845,818.193	0.000000	168,789.781	0.000000
[80;90)	H_0^1	3	204,547	34,629.140	0.000000	1,134,279.542	0.000000	145,606.976	0.000000
	H_0^2	3	204,547	3,459.459	0.000000	4,326.261	0.000000	2,487.632	0.000000
	H_0^3	1	204,547	20,518.342	0.000000	987,192.789	0.000000	145,276.026	0.000000
90+	H_0^1	3	27,231	76,094.181	0.000000	3,981,435.865	0.000000	442,225.455	0.000000
	H_0^2	3	27,231	29,048.400	0.000000	15,787.851	0.000000	53,467.864	0.000000
	H_0^3	1	27,231	57,916.018	0.000000	3,165,851.992	0.000000	365,881.426	0.000000

Note: Wald test of marginal association by age and sex with null hypotheses in equation (7) based on the age- and sex-specific estimates in the Expanded Joint Model. #Res. refers to the number of parameter restrictions and DoF is the degrees of freedom.

Table A18: Hypothesis Tests of Age and Time-to-death Specific Estimates under Expanded Joint Model

Age Group	Hypothesis		Wald Test		Bayes P-value	
	#Res.		Test Statistic	P-value	Test Statistic	P-value
Female						
TTD $\in [0;1.5)$	H_0^1	16	66,199,153	0.000000	-60,404,639	0.000000
	H_0^2	8	2,215,737	0.000000	-1,781,561	0.000000
	H_0^3	12	60,404,639	0.000000	-1,781,561	0.000000
TTD $\in [1.5;4)$	H_0^1	12	1,781,561	0.000000	-901,071	0.000000
	H_0^2	8	896,212	0.000000	-505,229	0.000000
	H_0^3	6	2,329,817	0.000000	-708,541	0.000000
TTD > 4	H_0^1	4	22,766	0.000000	-469,675	0.000000
	H_0^2	8	99,150	0.000000	-10,159,673	0.000000
	H_0^3	NA	NA	NA	NA	NA
Male						
TTD $\in [0;1.5)$	H_0^1	16	111,276,343	0.000000	-35,286,791	0.000000
	H_0^2	8	4,168,894	0.000000	-2,185,768	0.000000
	H_0^3	12	95,300,994	0.000000	-34,559,923	0.000000
TTD $\in [1.5;4)$	H_0^1	12	5,913,072	0.000000	-1,219,566	0.000000
	H_0^2	8	2,180,218	0.000000	-758,601	0.000000
	H_0^3	6	2,976,652	0.000000	-1,077,573	0.000000
TTD > 4	H_0^1	4	11,139	0.000000	-487,458	0.000000
	H_0^2	8	189,505	0.000000	-8,384,418	0.000000
	H_0^3	NA	NA	NA	NA	NA

Note: Wald test from equation (29) and Bayesian p-values from Section B.2 of parameter estimates, $\hat{\theta}$ by time-to-death and sex. Tests are based on the time-to-death- and sex-specific estimates of the Expanded Joint Model. #Res. refers to the number of parameter restrictions and DoF is the degrees of freedom.

Table A19: Hypothesis Tests of Hypothesis in Equation (7) by Time-to-death and Sex under Expanded Joint Model

Group	Hypothesis		Wald Test, E_F		Wald Test, E_g		Wald Test, E_ψ		
	#Res.	DoF	Test Stat.	P-value	Test Stat.	P-value	Test Stat.	P-value	
Female									
TTD \in [0;1.5)	H_0^1	3	115,212	10,462,566.678	0.0000000	1,321.115	0.0000000	97,032.573	0.0000000
	H_0^2	3	115,212	15,875,212.220	0.0000000	974.639	0.0000000	20,788.132	0.0000000
	H_0^3	1	115,212	1,108,896.386	0.0000000	866.128	0.0000000	91,030.012	0.0000000
TTD \in [1.5;4)	H_0^1	3	189,120	4,287,850.570	0.0000000	370.818	0.0000000	30,652.950	0.0000000
	H_0^2	3	189,120	4,896,409.767	0.0000000	587.887	0.0000000	84,068.219	0.0000000
	H_0^3	1	189,120	101,223.585	0.0000000	194.668	0.0000000	590.985	0.0000000
TTD $>$ 4	H_0^1	3	4,820,971	25,397.142	0.0000000	6.843	0.0326709	80.133	0.0000000
	H_0^2	3	4,820,971	60,029.854	0.0000000	20.859	0.0003377	12,892.384	0.0000000
	H_0^3	NA	NA	NA	NA	NA	NA	NA	NA
Male									
TTD \in [0;1.5)	H_0^1	3	110,495	8,742,001.315	0.0000000	1,395.730	0.0000000	39,927.922	0.0000000
	H_0^2	3	110,495	22,091,912.522	0.0000000	461.846	0.0000000	68,867.218	0.0000000
	H_0^3	1	110,495	367,506.988	0.0000000	1,100.961	0.0000000	29,835.457	0.0000000
TTD \in [1.5;4)	H_0^1	3	185,613	6,020,325.074	0.0000000	566.711	0.0000000	34,796.059	0.0000000
	H_0^2	3	185,613	4,332,540.643	0.0000000	511.055	0.0000000	35,323.126	0.0000000
	H_0^3	1	185,613	117,799.965	0.0000000	406.020	0.0000000	466.633	0.0000000
TTD $>$ 4	H_0^1	3	4,578,361	125,356.597	0.0000000	9.031	0.0109355	418.503	0.0000000
	H_0^2	3	4,578,361	187,259.492	0.0000000	23.679	0.0000926	8,390.307	0.0000000
	H_0^3	NA	NA	NA	NA	NA	NA	NA	NA

Note: Wald test of marginal association by age and sex with null hypotheses in equation (7) based on the time-to-death- and sex-specific estimates of the Expanded Joint Model. #Res. refers to the number of parameter restrictions and DoF is the degrees of freedom.

C.6 Robustness: Decomposition

Table A20 show the decomposition of hospital expenditure changes for years 2005 to 2015. Results are close to the main estimates for years 2002 and 2017 in Table 5.

Table A20: Decomposition of Hospital Care Expenditure Differences between 2005 and 2015

	Total	Pop. Change		Healthc. Change		Naïve		TTD Error		Steepening		TTD Steep.	
	\$mio.	\$mio.	%	\$mio.	%	\$mio.	%	\$mio.	%	\$mio.	%	\$mio.	%
All	2359.9	541.2	22.9	1818.7	77.1	1027.1	43.5	-485.9	-20.6	1332.8	56.5	485.9	20.6
Female	1131.9	155.8	13.8	976.1	86.2	373.5	33.0	-217.7	-19.2	758.4	67.0	217.7	19.2
Male	1228.0	411.7	33.5	816.3	66.5	680.8	55.4	-269.2	-21.9	547.2	44.6	269.2	21.9