Learning from the Oldest: The Impact of Ageing on Healthcare Expenditure

Paolo Berta* Sara Muzzì† Manuel Serrano-Alarcón‡
Stefano Verzillo§

August 15, 2026

Abstract

This paper estimates the causal impact of population ageing on healthcare expenditure. Using linked administrative healthcare records combined with historical demographic data from Lombardy, Italy's largest region, we build a highly granular panel covering all municipalities over 2011-2023. To address endogeneity, we exploit predicted variation in ageing across municipalities based on lagged local age structures and regional cohort-specific survival rates. Our instrumental variable (IV) estimates show that population ageing significantly increases per-capita healthcare expenditure across major domains of care. A 1% increase in the share of the population aged 60 and above raises per-capita expenditure by 0.43% for hospitalizations, 0.75% for pharmaceutical consumption, and 0.30% for outpatient services. IV estimates are generally smaller than their OLS counterparts, by up to 30-48% for the oldest age ratio (75+). This indicates that conventional approaches that do not address endogeneity may overstate the true effect of ageing on healthcare expenditure. Decomposing the ageing effect between older adults (aged 60-74) and the oldest-old (aged 75+) reveals that the latter group is the dominant driver of expenditure growth across all domains. A series of robustness checks and placebo analyses support the validity and consistency of the results.

Keywords: Ageing, Healthcare Expenditure, Instrumental Variables

---

Acknowledgements. The authors gratefully acknowledge the support and the provision of administrative data by the Lombardy Health Authority, within the framework of the Age-Bicocca project, and by the Italian National Institute of Statistics (ISTAT). The authors

*University of Milan-Bicocca. Email: paolo.berta@unimib.it
†University of Milan-Bicocca. Email: sara.muzzi@unimib.it
‡European Commission, JRC. Email: manuel.serrano-alarcon@ec.europa.eu
§European Commission, JRC and IZA@LISER. Email: stefano.verzillo@ec.europa.eu

also gratefully acknowledge Irene Torrini for providing social expenditure data used for robustness purposes. This project is also part of the Research on Administrative Data (RAD) Initiative at the Competence Centre on Microeconomic Evaluation (CC-ME) of the Joint Research Centre (JRC) of the European Commission (EC). The views expressed in this paper are those of the authors and do not necessarily reflect those of the European Commission. During the preparation of this manuscript, the authors used GPT@JRC, Claude and ChatGPT to improve the readability and language of the manuscript and the formatting of Latex tables. After using the tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

1 Introduction

Population ageing is widely regarded as a major driver of long-term healthcare expenditure growth. Rising long-term care expenditure is of particular concern, as it places increasing pressure on the fiscal sustainability of healthcare systems and public finances more broadly (Guthmuller et al., 2021). However, disentangling the causal effect of the ageing component from other confounding trends -such as technological change, economic growth, migration patterns, improvements in medical practice and health status, or policy reforms introduced to face both demographic and technological changes- remains empirically challenging. Standard projections performed by policy institutions typically rely on mechanical age-expenditure profiles and assume that the observed age-specific spending patterns will hold in the future (Commission, 2024). These methods disregard the endogeneity of population ageing in the healthcare expenditure equation, as the observed relationship between ageing and spending may reflect both reverse causality and common confounding trends affecting demographic composition and healthcare needs. Furthermore, conventional projections rarely distinguish between types of care (hospital, outpatient, pharmaceutical) despite the fact that ageing may differentially affect specific services. This creates a crucial knowledge gap that limits healthcare planning and resource allocation.

In this paper, we examine the causal impact of ageing in healthcare expenditure in one of the most aged countries in the world: Italy. We build on Maestas et al. (2023)[1], and use a cohort-survival Bartik-style instrument to estimate the causal impact

[1]Maestas et al. (2023) use a cohort-survival Bartik-style instrument to estimate the causal impact of

of healthcare expenditure and its core dimensions, using microdata from the largest region in Italy, Lombardy. This strategy allows us to address key endogeneity concerns between population ageing and healthcare expenditure, arising from reverse causality, selective migration and other unobserved local trends.

Our analysis relies on rich administrative data covering the universe of healthcare expenditures for the Lombardy population over the period (2011-2023). This data allows us to track not only total healthcare expenditure, but also utilization and spending by major healthcare domains. This level of granularity is particularly relevant for healthcare planning as the burden of ageing on the healthcare system is unlikely to be equally distributed across healthcare domains. Italy provides a unique and informative case study for this analysis: according to World Bank data, with 25% of its population aged 65+, Italy was the fourth most aged country in the world in 2024, behind only Monaco, Japan, and Puerto Rico. As a consequence, this experience offers valuable insights for countries that are at earlier stages of the demographic transition.

Our instrumental variable estimates provide causal evidence that population ageing increases per-capita healthcare expenditure across all major domains of care. For eight-year growth horizons over the pre-COVID period (2011-2019), a 1% increase in the share of the population aged 60 and above raises per-capita expenditure by approximately 0.43% for hospitalizations, 0.75% for pharmaceutical consumption, and 0.30% for outpatient services. These elasticities are economically meaningful and demonstrate substantial heterogeneity across healthcare domains, with pharmaceutical spending exhibiting the strongest response to demographic ageing. The findings remain robust to alternative specifications, including older age ratios, longer instrumental variable lags and extended time periods that include the COVID-19 pandemic.

The obtained instrumental variable estimates are generally smaller than their OLS counterparts, suggesting that simple correlations tend to overstate the effect of population ageing on healthcare expenditure. This pattern is consistent with upward bias in OLS estimates arising from reverse causality, selective migration, or endogenous adjustments

population ageing on economic growth in the US. We extend this methodology for the case of healthcare expenditure.

in healthcare policy. The bias is particularly pronounced for older age thresholds, with IV estimates for the 75+ ratio running 30-48% below their OLS counterparts, suggesting that endogeneity concerns are most severe for the oldest segment of the population.

We estimate that population ageing over the period 2011-2019 contributed to expenditure increases of 4.2% for hospitalizations, 5.2% for pharmaceuticals, and 1.7% for outpatient services. Decomposing this effect between older adults (aged 60-74) and the oldest-old (aged 75+) reveals that, despite starting from a smaller population base, the oldest-old are the dominant driver of ageing-related expenditure growth across all healthcare domains. Given that this group is projected to expand most rapidly in the coming decades, these findings point to an increasing concentration of the fiscal burden of demographic ageing among the very oldest segment of the population.

Our paper contributes to three main strands of the literature. First, we contribute to the literature that studies the determinants of healthcare expenditure using an aggregate panel dataset. Early contributions based on country-level or regional panels document a positive association between population ageing and healthcare spending, (Bech et al., 2011; Piscopo et al., 2024). However these studies are unable to address time-varying unobserved heterogeneity, including policy reforms, medical innovation, migration, etc. As a result, this literature provides limited causal evidence on the effect of ageing *per se*. We depart from this approach by exploiting plausibly exogenous variation in local demographic ageing to avoid confounding from contemporaneous economic and institutional factors. To the best of our knowledge, this paper is the first to provide causal evidence on the effect of population ageing on healthcare expenditure at regional level, and to do so separately by major healthcare domains.

Second, our paper relates to the literature on the relative role of ageing and proximity to death in explaining healthcare expenditure. A large body of research has shown that, once time-to-death is accounted for, the independent effect of ageing on individual healthcare spending is substantially reduced—the so-called red-herring hypothesis (Howdon and Rice, 2018; Werblow et al., 2007; Costa-Font and Vilaplana-Prieto, 2020). While this literature focuses on individual-level expenditure dynamics, our paper examines whether

demographic ageing causally increases healthcare spending at the municipality level. To connect our analysis with this literature, we explicitly investigate the role of end-of-life spending by re-estimating our baseline model using Last-Year-of-Life (LYL) healthcare expenditure as the outcome. This allows us to assess whether the effect of population ageing operates primarily through higher end-of-life costs.

Third, we contribute to the literature on the macroeconomic effects of population ageing. While earlier studies were not conclusive on the relation between population ageing and economic growth (Acemoglu and Restrepo, 2017; Sheiner, 2014), more recent causal studies using cohort-based instruments have found that population ageing reduces output per capita (Maestas et al., 2023). We extend this identification strategy to the case of healthcare utilization on expenditure. Taken together, our results imply that demographic ageing increases the fiscal burden of the healthcare system both directly (through higher healthcare expenditure) and indirectly (through lower GDP and the associated fiscal revenue base).

The remaining of the paper is organised as follows. Section 2 explains the data and the Italian demographic transition. Section 3 details the empirical strategy. Section 4 describes the main results. Section 5 provides robustness checks, while Section 6 translates the estimated elasticities into their fiscal impact, and Section 7 concludes.

2 Data

Our analysis focuses on Lombardy, the most populated Italian region, with approximately 10 million inhabitants. We utilize two main data sources. Population data comes from The Italian National Institute of Statistics (Istat). In particular, we use population data by age, sex and municipality of residence for the period 1952 to 2023[2].

Healthcare data includes individual administrative records provided by the Lombardy Region on healthcare utilization between 2010 and 2023. We consider three health domains: hospitalizations, pharmaceutical prescriptions, and outpatient services (including

[2]The Italian National Institute of Statistics (Istat) is the main provider of official statistics in Italy. Its online database is publicly available at https://demo.istat.it

blood tests, diagnostic procedures, and specialist visits). These records refer to healthcare services reimbursed, either fully or partially, by the Italian National Health Service. Our measure of healthcare expenditure excludes emergency department visits as well as fixed system-level costs, such as medical staff salaries, facility maintenance, and general healthcare administration, because these items are not reported in the available administrative data. For each domain, we aggregate individual-level records to construct municipality-year measures of total real expenditure (inflation-adjusted) and calculate per-capita overall expenditure by dividing by the municipality's population aged 20 and above. In our main specification, we use an eight-year growth period ($p = 8$) covering 2011[3]-2019: the endpoint avoids the confounding effects of the COVID-19 pandemic, while robustness checks extend the sample through 2023.

2.1 Italy's demographic transition

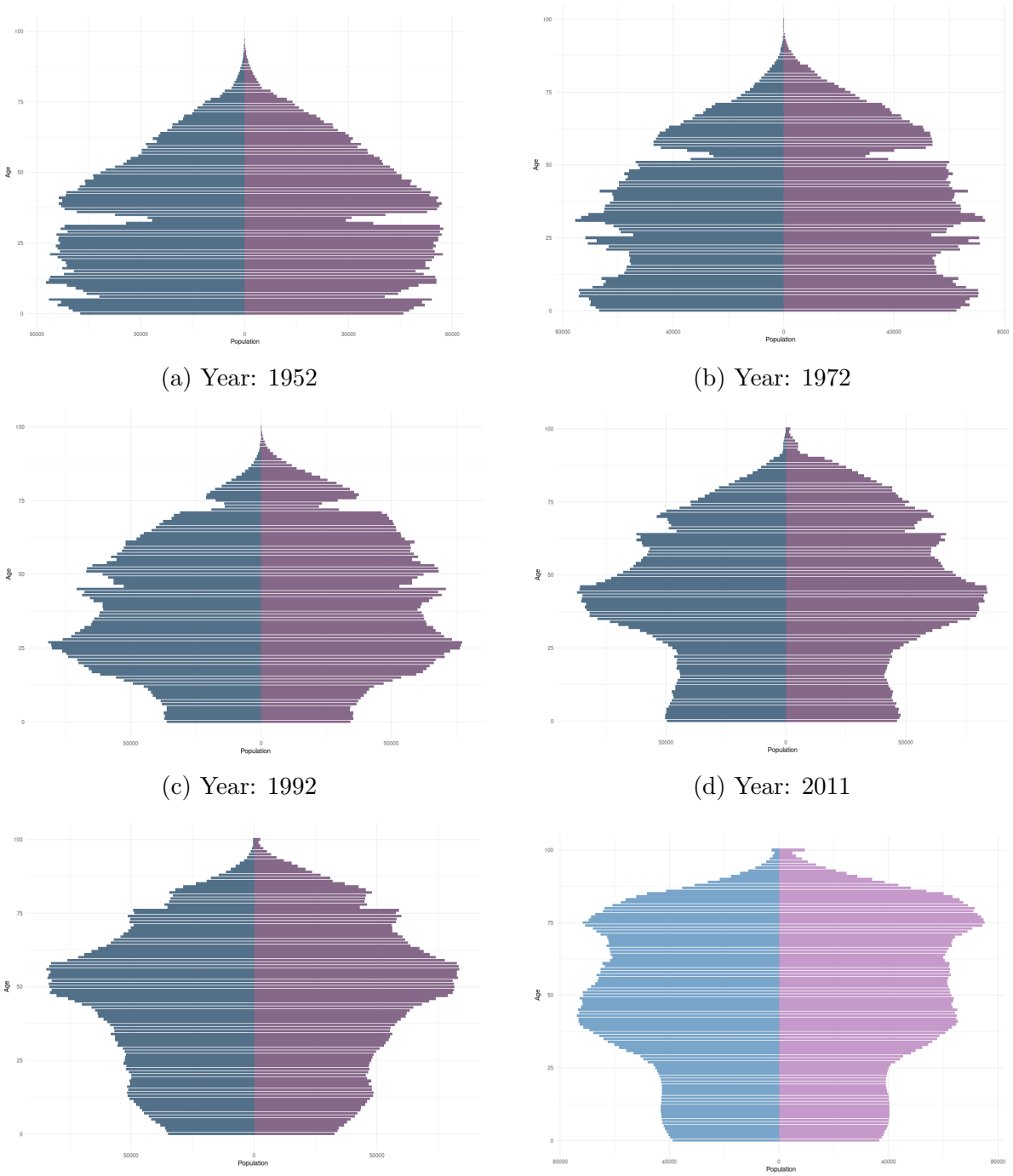
Figure 1 displays population pyramids for Lombardy at selected years from 1952 to 2050, illustrating the dramatic transformation of the population age structure across time.[4] The figure captures the post-World War II Baby Boom (1946-1964), reflected in the broad base of the 1952 and 1972 pyramids, followed by subsequent fertility declines. By 2011, Baby Boom cohorts had begun entering older age groups, accelerating the ageing process. This trend has been reinforced by continued declines in fertility rates and substantial increases in life expectancy driven by healthcare innovations. The share of population aged 65 or more reached 23.3% by 2023 (well above the EU average of 21.3%) and is expected to go well over 30% by 2050.

Figure 2 plots, for each municipality, the standardized 2011-2019 change in per-capita healthcare expenditure by care domain against the standardized observed change in population ageing. Each observation is weighted by the municipality resident population (aged 20+) to account for differences in population size across municipalities. Across all

[3]The starting year is set by the need to construct LYL expenditure, which requires administrative records going back to 2010.

[4]To dynamically visualize the evolution of the population in Lombardy we created a *Shiny App* available at (https://chiarabersani.shinyapps.io/hist_pop_1952_2080/). Importantly, the single age category of 100 years also includes people older than 100.

expenditure categories, the descriptive relationship is positive but limited in magnitude, being this correlation more evident for pharmaceutical than for outpatient and hospital expenditure. Panels (a)-(c) show a slightly upward slope, meaning that municipalities experiencing faster population ageing tend, on average, to exhibit higher growth in per-capita healthcare spending over the same period. However, the estimated associations are relatively weak and characterized by substantial dispersion, suggesting large heterogeneity across municipalities. Overall, these patterns indicate that simple cross-municipality correlations between ageing and healthcare expenditure growth have limited explanatory power.



(a) Year: 1952

(b) Year: 1972

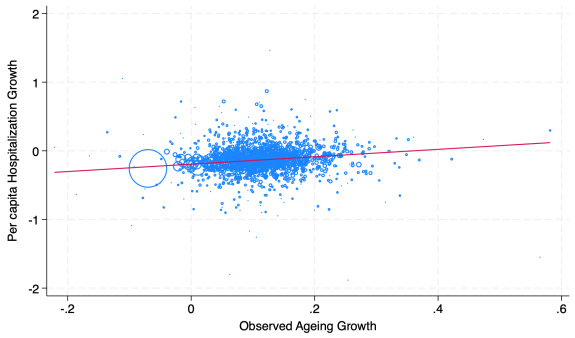
(c) Year: 1992

(d) Year: 2011

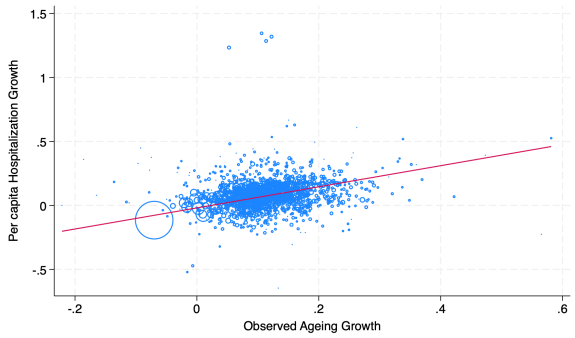
(e) Year: 2023

(f) Year: 2050

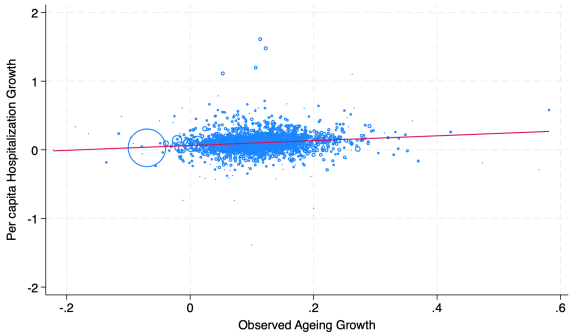
Figure 1: Population pyramids for Lombardy across selected years (1952-2050)



(a) Hospitalizations



(b) Pharmaceutical prescriptions



(c) Outpatient services

Figure 2: Relationship between Observed Ageing and Per-Capita Health Care Expenditure Growth (2011-2019)

Notes: Bubble size reflects municipality population, measured by the number of residents aged 20 and above.

3 Empirical Strategy

Our main model estimates the causal effect of population ageing on the growth of healthcare expenditure. We measure population ageing as the change in the share of the population aged 60 and above (alternatively 65+, 70+, or 75+) between years t and $t+p$. Our main specification examines whether changes in population ageing causally affect the growth rate of real per-capita healthcare expenditure over the same period:

$$\ln\left(\frac{Y_{m,t+p}}{N_{m,t+p}}\right) - \ln\left(\frac{Y_{m,t}}{N_{m,t}}\right) = \beta\left[\ln\left(\frac{\text{P60}_{m,t+p}}{N_{m,t+p}}\right) - \ln\left(\frac{\text{P60}_{m,t}}{N_{m,t}}\right)\right] + \gamma_m + \varepsilon_{m,t} \qquad (1)$$

where $Y_{m,t}$ denotes real healthcare expenditure (i.e., adjusted for inflation) in munic-

ipality m in year t. Specifically, we examine three categories of healthcare expenditure: (i) hospitalizations, (ii) pharmaceutical consumption, and (iii) outpatient services. $N_{m,t}$ is the total population aged 20 years or older[5], and $\text{P60}_{m,t}$ is the older population (aged 60+). We estimate alternative specifications using different age thresholds (65+, 70+, and 75+) to assess how the effect varies with the definition of the elderly population. Our main variable of interest on the right hand side $\left[\ln\left(\frac{\text{P60}_{m,t+p}}{N_{m,t+p}}\right) - \ln\left(\frac{\text{P60}_{m,t}}{N_{m,t}}\right)\right]$, captures the p-year growth in the 60+ share of the population, and serves as our measure of population ageing. By estimating the model in first differences, we account for time-invariant differences across municipalities.

The estimated coefficient $\hat{\beta}$ captures the effect of population ageing on the growth of the specific healthcare use indicator between t and $t + p$, and can be interpreted as an elasticity.

In our main specification, we use an eight-year growth period ($p = 8$) covering 2011-2019, which avoids the disruptive effects of the COVID-19 pandemic on both healthcare utilization and mortality patterns. We also include ATS (Local Health Authority) fixed effects (γ_m) which control for organizational trends during 2011-2019 that are common at the ATS level. Regressions are weighted by municipality population (20+) at time t and standard errors are robust, except in the six-year pooled specification (Table 7), where each municipality contributes two observations and standard errors are clustered at the municipality level to account for within-municipality serial correlation.

As a robustness check, we augment the baseline specification with a vector of time-varying municipal controls. Specifically, $X_{m,t}$ includes the log difference in average municipal income between t and $t + p$, thereby accounting for changes in local economic conditions that may jointly affect population ageing and healthcare expenditure and the log difference in average municipal employment between t and $t + p$, which captures changes in local labour market conditions that may influence both migration patterns and healthcare utilization.

[5]Following Maestas et al. (2023), we use as denominator only the 20+ population to allow the population at t to be predicted by the population structure some decades before, as explained in section 3.1

3.1 Endogeneity of Population Ageing and Instrumental Variable Strategy

Changes in the share of the elderly population may be endogenous to the evolution of healthcare utilization for at least a couple of reasons. First, higher healthcare expenditure may improve health outcomes of the citizens and increase survival, thereby raising the share of older individuals in a municipality (i.e., reverse causality). In this case, conventional OLS estimates would overstate the causal effect of population ageing on healthcare expenditure. Second, selective migration may induce additional endogeneity. Older individuals may decide to relocate to geographic areas to seek better healthcare services, while younger individuals may move in response to better local economic opportunities in urban areas, causing changes in the age structure at the local level that are correlated with healthcare expenditure. Additionally, there may be confounding trends such as technological innovation correlated with both population ageing and healthcare expenditure. Depending on the direction and magnitude of these relationships, OLS estimates of β provided by Equation 1 may be biased either upward or downward.

To address these concerns, we construct an instrumental variable that predicts ageing based on the municipality's past population structure combined with common regional cohort survival rates following Maestas et al. (2023). Specifically, we use the age structure in year $t - x$ to predict changes in the log share of the municipal population aged 60+ between years t and $t + p$ (or aged 65+, aged 70+ and aged 75+ in our alternative specifications).

In our main model, we use an eight-year period to measure growth rates ($p = 8$) and a ten-year lag ($x = 10$) to predict the age structure at both t and $t + 8$. For instance, $x = 10$ and $p = 8$ imply that the age structure observed in 2001 ($t - 10$) is used to predict the elderly share of both 2011 (t) and 2019 ($t + 8$), and consequently the growth rate in the 60+ population during the period 2011-2019. Robustness checks in Section 5, confirm that results remain unchanged when longer instrumental lags ($x = 15$) or extended sample periods (2011-2023 including Covid-19) are used.

Our instruments are constructed as:

$$\ln\left(\frac{\widehat{P60}_{m,t+p}}{\widehat{N}_{m,t+p}}\right) - \ln\left(\frac{\widehat{P60}_{m,t}}{\widehat{N}_{m,t}}\right), \tag{2}$$

where $\widehat{P60}_{m,t+p}$ and $\widehat{P60}_{m,t}$ denote the *predicted* municipal population aged 60+ in years $t+p$ and t, respectively, constructed using regional cohort survival rates and the municipal age structure at time $t-x$.

More precisely for our baseline specification of an eight-year growth period ($p = 8$) and a ten-year lag ($x = 10$):

$$\widehat{P60}_{m,t+p} = \sum_{j \geq 60-x-p} \underbrace{N^j_{m,t-x}}_{\substack{\text{Population age } j \\ \text{in municipality } m \text{ at } t-x}} \times \underbrace{\frac{N^{j+x+p}_{t+p}}{N^j_{t-x}}}_{\substack{\text{Regional survival rate} \\ \text{from } t-x \text{ to } t+p}} = \sum_{j \geq 42} N^j_{m,t-10} \times \frac{N^{j+18}_{t+8}}{N^j_{t-10}}$$

$$\widehat{P60}_{m,t} = \sum_{j \geq 60-x} \underbrace{N^j_{m,t-x}}_{\substack{\text{Population age } j \\ \text{in municipality } m \text{ at } t-x}} \times \underbrace{\frac{N^{j+x}_{t}}{N^j_{t-x}}}_{\substack{\text{Regional survival rate} \\ \text{from } t-x \text{ to } t}} = \sum_{j \geq 50} N^j_{m,t-10} \times \frac{N^{j+10}_{t}}{N^j_{t-10}}$$

The denominator terms $\widehat{N}_{m,t+p}$ and $\widehat{N}_{m,t}$ denote the predicted municipal population aged 20+ and are constructed analogously:

$$\widehat{N}_{m,t+8} = \sum_{j \geq 20-x-p} N^j_{m,t-x} \times \frac{N^{j+x+p}_{t+p}}{N^j_{t-x}} = \sum_{j \geq 2} N^j_{m,t-10} \times \frac{N^{j+18}_{t+8}}{N^j_{t-10}}$$

$$\widehat{N}_{m,t} = \sum_{j \geq 20-x} N^j_{m,t-x} \times \frac{N^{j+x}_{t}}{N^j_{t-x}} = \sum_{j \geq 10} N^j_{m,t-10} \times \frac{N^{j+10}_{t}}{N^j_{t-10}}$$

The resulting predicted ageing measures capture how the municipal older population would evolve if each cohort observed at time $t-10$ were exposed to the same regional cohort-specific survival rates prevailing between t and $t+8$. Intuitively, the instrument leverages variation in municipalities' plausible exogenous and pre-determined historical age structures combined with common regional demographic forces captured by cohort

survival rates.

The predicted change in the old-age population share constructed in Equation (2) serves as an instrument for the endogenous change in the municipal share of the population aged 60+ in a two-stage least squares (2SLS) estimation framework. The second stage equation corresponds to Equation (1), while the first stage equation is given by the Equation (3) below:

$$\ln\left(\frac{\mathrm{P60}_{m,t+8}}{N_{m,t+8}}\right) - \ln\left(\frac{\mathrm{P60}_{m,t}}{N_{m,t}}\right) = \alpha\left[\ln\left(\frac{\widehat{P60}_{m,t+8}}{\widehat{N}_{m,t+8}}\right) - \ln\left(\frac{\widehat{P60}_{m,t}}{\widehat{N}_{m,t}}\right)\right] + \gamma_m + \eta_{m,t}. \tag{3}$$

where the variables and subscripts are the same as those defined in Equation 1. Regressions are weighted by municipality resident population aged 20+ at time t, and robust (except in the six-year pooled specification (Table 7), where standard errors are clustered at the municipality level) standard errors are computed.

The instrument is designed to capture how the municipality's elderly share would have evolved if each initial cohort had experienced the cohort-specific demographic trends observed at the regional level. The identifying assumption is that the past age structure of a municipality affects healthcare expenditure growth only through its impact on subsequent population ageing. Because the age distribution at time t is predetermined, and regional age-specific survival patterns are primarily driven by demographic forces which are common across municipalities, the instrument provides plausibly exogenous variation in predicted ageing across municipalities.

The instrument proposed can be interpreted as a shift-share (Bartik-style) design, with municipal cohort shares at $t - x$ acting as exposure weights and regional cohort-specific survival rates as the regional common shock (Goldsmith-Pinkham et al. 2020; Borusyak et al. 2022). Within each horizon t and $t + p$ separately, the predicted shares $\widehat{P60}_{m,\cdot}/\hat{N}_{m,\cdot}$ are themselves exact linear combinations of cohort shares. However, because the instrument (see Equation 2) is defined as the log-difference of these two ratios across horizons, rather than as a single linear combination of cohort shares as in Goldsmith-Pinkham et al. (2020), the exact Rotemberg-weight decomposition does not apply directly

to our instrument. We assess this diagnostic on a linearized version of the instrument in Section 4.3.

4 Results

This section presents the empirical results from the IV estimation of the causal effect of population ageing on healthcare expenditure. In particular, Section 4.1 examines the validity of the instrumental variable and discusses potential threats to the exclusion restriction assumption while Section 4.2 reports the main estimates of the causal effect of population ageing on healthcare expenditure across the three healthcare dimensions (hospitalizations, pharmaceutical and outpatient services).

4.1 First Stage and Instrument Validity

Identification of the IV estimator relies on two identifying conditions: instrument relevance (a sufficiently strong first stage) and instrument exogeneity, captured by the exclusion restriction. We provide evidence in support of each of these assumptions in turn. Figure 3 illustrates the first-stage relationship underlying our instrumental variable strategy. The predicted change in the share of the population aged 60+, constructed by combining lagged municipal age structures with cohort-specific survival rates at regional level, strongly predicts realized ageing at the municipality level. The figure shows a strong first-stage relationship, with the predicted change in the elderly population share closely aligned with the observed change across municipalities. This provides visual evidence that the instrument explains a substantial share of the cross-municipality variation in population ageing. The corresponding first-stage Kleibergen-Paap Wald F statistic is 109.461 (Table 1), well above the Stock-Yogo critical values, provides strong evidence of instrument relevance and ruling out concerns about weak instruments. Comparable first-stage Kleibergen-Paap Wald F statistics are observed for all alternative age thresholds.

The exclusion restriction requires that the instrument affects healthcare expenditure growth only through its effect on population ageing, and not through any alternative

channel. In other words, the predicted ageing component obtained using the historical age structures should affect current healthcare expenditure only through its impact on the subsequent growth of the elderly population.

A potential threat to the validity of the exclusion restriction assumption would arise if historical age structures were correlated with persistent local economic or demographic trends that independently predict future healthcare expenditure (e.g., long-run patterns in economic activity or population dynamics). To mitigate this concern, we exploit lagged population structures from sufficiently distant periods (15 and 20 years before the outcome period, instead of 10 years) in robustness checks presented in Section 5. Using older age distributions reduces the likelihood that the instrument captures contemporaneous trends that directly affect healthcare spending.

The institutional structure of healthcare provision in Italy further supports the plausibility of the exclusion restriction. Healthcare expenditure is determined primarily at the regional level: regional governments finance healthcare services through regional and national revenue sources, and municipal economic performance does not directly determine healthcare budgets. Any major changes in healthcare expenditure therefore affect all municipalities within a region uniformly and are independent of the cohort-driven demographic variation that our instrument exploits.

A residual concern is that municipality-level population ageing may affect dimensions of local expenditure closely related to healthcare, in which case the instrument would capture broader local spending dynamics rather than ageing per se. We test for this possibility using the two dimensions of local expenditure closest to healthcare expenditure: social and healthcare protection spending. Social expenditure category captures local government interventions in welfare and social services, whereas healthcare protection expenditure includes municipal outlays classified under Mission 13 ("Tutela della salute") of the harmonized local government budget (Legislative Decree 118/2011). Programs 01-06 of this mission relate to regional health service financing and are structurally near-zero at the municipal level, as this is a regional rather than municipal competence. Municipal spending under this mission is concentrated in Program 07 ("Ulteriori spese in

materia sanitaria"), which finances residual health-related expenditure not covered by the other programs of the mission, specifically hygienic-sanitary interventions such as public kennels and public sanitation facilities, and environmental hygiene interventions such as pest control and disinfection. A significant association with either outcome would raise concerns about the exclusion restriction, as these expenditure categories could indirectly affect healthcare expenditure. However, the results reported in Appendix B.1 (social expenditure) and Appendix B.2 (healthcare protection) show no significant association between predicted ageing and either local expenditure dimensions. As a further check, we test whether the instrument spuriously predicts municipal expenditure categories with no plausible link to population ageing or healthcare demand. Using roads expenditure (construction and maintenance of local roads) as a placebo outcome, we find IV effects statistically indistinguishable from zero across all age thresholds (Appendix C). Taken together, these results suggest that the instrument does not proxy for broader local fiscal or health-related policy dynamics, reinforcing the plausibility of the exclusion restriction.



Figure 3: Relationship between Predicted Ageing and Observed Ageing Growth. Size of bubbles reflects municipality population size (20+).

4.2 Main results

In this section, we present the estimates from our main IV-2SLS specification and compare them with the corresponding OLS estimates for each of the healthcare expenditure: hospitalizations, pharmaceutical consumption, and outpatient care (Table 1). To assess the sensitivity of the results to alternative definitions of the treatment variable, we estimate the model using four age thresholds: the share of the population aged 60+, 65+, 70+, and 75+.

The first-stage estimates (last column of Table 1) are positive and statistically significant across all age thresholds (60+/65+/70+). The positive coefficients indicate that municipalities with larger predicted increases in the elderly population share also experience larger observed increases in ageing, consistent with the expected first-stage relationship. The corresponding first-stage Kleibergen-Paap Wald F statistic, reported in section 4.1, indicate that the instrument provides substantial predictive power for the endogenous ageing variable. For the over-60 group, the estimated coefficient is 1.386, implying that a 1% increase in the predicted ageing measure is associated with a 1.386% increase in the observed elderly population share. The magnitude of the first-stage coefficient decreases for older age thresholds, reflecting the increasing importance of mortality differentials at older ages. As cohorts become older, municipality-specific mortality patterns may diverge more substantially from regional cohort-specific survival rates (the shift part of the instrument), reducing the precision with which the instrument predicts ageing among the oldest population groups.

The causal IV estimates reveal economically meaningful and statistically significant causal effects of population ageing on healthcare expenditure across all healthcare domains. A 1% increase in the share of the population aged 60 and above increases per-capita healthcare expenditure by approximately 0.43% for hospitalizations, 0.75% for pharmaceutical prescriptions, and 0.30% for outpatient services. These estimated elasticities highlight substantial heterogeneity in the response of different healthcare domains to demographic ageing, with pharmaceutical expenditure exhibiting the largest elasticity while outpatient care the lowest.

The estimated elasticities decline systematically as the ageing measure is defined using older age thresholds. For example, the elasticity of hospital expenditure decreases from 0.43 for the 60+ population share to 0.39, 0.34, and 0.25 when ageing is measured using the shares of the population aged 65+, 70+, and 75+, respectively. A similar pattern is observed for pharmaceutical and outpatient expenditure (Table 1). This gradient should not be interpreted as evidence of lower healthcare needs among the oldest-old. It reflects a scaling effect: because the elderly share is smaller at higher age thresholds, the same percentage increase corresponds to a smaller absolute change in the population. Using 2011 baseline shares (60+: 33%; 75+: 12%, see Table A1), a 10% increase in the 60+ share implies approximately 3.3 additional percentage points of elderly population, whereas the same relative increase in the 75+ share implies only 1.2 percentage points. The elasticities are therefore not directly comparable across thresholds, and the decline across age groups is a mechanical consequence of the narrowing demographic base rather than a substantive difference in the expenditure response to ageing.

Comparing the IV and OLS estimates provides insights into the potential magnitude and direction of endogeneity bias. The IV estimates are consistently smaller across healthcare domains and ageing threshold indicators. This suggests that reverse causality due to the survival effect of higher healthcare driven-expenditure and healthcare-driven migration may explain the bias of the OLS estimates. Furthermore, the relative difference between OLS and IV estimates increases as the as the definition of ageing shifts from the 65+ to the 70+ and 75+ population shares. For the 75+ specification, the IV estimates are approximately 30-48% smaller than their OLS counterparts. This widening gap is consistent with endogeneity becoming more pronounced when ageing is measured using the oldest population groups. One plausible explanation is that reverse causality and selective migration are relatively more important at advanced ages. Improvements in healthcare quality may have a stronger effect on the survival of the oldest-old, while individuals in these age groups may also be more likely to relocate in response to healthcare availability and perceived quality. As a result, conventional OLS estimates are more likely to overstate the causal effect of population ageing on healthcare expenditure for

the oldest age thresholds.

Importantly, the robustness of these main results is assessed in Section 5, where we show that the estimated elasticities remain stable across a range of alternative specifications, including different instrument lags, additional local economic controls, alternative sample compositions, alternative temporal aggregations, spatially robust standard errors and accounting for migration flows.

Table 1: OLS and IV Elasticities by Care Domain

	Hospitalizations		Pharmaceutical		Outpatient		First Stage
	OLS	IV	OLS	IV	OLS	IV	
Over 60	0.544***	0.431***	0.862***	0.753***	0.401***	0.300***	1.386***
	(0.063)	(0.084)	(0.099)	(0.111)	(0.043)	(0.065)	(0.133)
Over 65	0.460***	0.392***	0.697***	0.604***	0.340***	0.282***	1.298***
	(0.046)	(0.059)	(0.079)	(0.090)	(0.030)	(0.042)	(0.107)
Over 70	0.448***	0.344***	0.622***	0.444***	0.302***	0.199***	1.159***
	(0.051)	(0.064)	(0.098)	(0.108)	(0.038)	(0.052)	(0.102)
Over 75	0.362***	0.253***	0.477***	0.323***	0.208***	0.108***	1.098***
	(0.053)	(0.063)	(0.105)	(0.106)	(0.048)	(0.058)	(0.080)
Observations	1,502		1,502		1,502		1,502

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. All regressions are weighted by the resident population aged 20+ and include ATS fixed effects. The baseline specification uses an eight-year growth window (2011-2019, $p = 8$) and an instrument constructed with a 10-year lag ($x = 10$). K-P Wald F Statistics: 109.46 for 60+, 147.19 for 65+, 128.56 for 70+ and 188.42 for 75+. The Durbin-Wu-Hausman endogeneity test rejects the null hypothesis that the instrumented regressor is exogenous in all care domains and age thresholds ($p < 0.01$), confirming the presence of endogeneity and motivating the IV approach. Robust standard errors are reported in parentheses.

4.3 Instrument Diagnostic

As discussed in Section 3.1, the exact Rotemberg-weight decomposition proposed in Goldsmith-Pinkham et al. (2020) does not apply directly in our setting. We therefore construct a linearized version of the instrument, obtained via a first-order Taylor approximation of the logarithm of each ratio around its sample mean, which recasts it as an exact linear combination of municipal cohort shares weighted by a combined regional survival shock. This linearized instrument tracks the original very closely (correlation of 0.987 with the original, non-linearized instrument), and the resulting 2SLS coefficient (0.4405) differs from our baseline estimate (0.431 reported for Hospitalizations–Over 60 in Table 1) by approximately 2.2%, consistent with the residual approximation error of

a first-order Taylor expansion of the logarithm. Applying the Rotemberg-weight decomposition to this linearized instrument, we verify the exact identity $\hat{\beta}_{IV} = \sum_j \hat{\alpha}_j \hat{\beta}_j$ and we find that all the cohorts aged 60-82 at the base year (2001) account for the large majority of identifying variation, with weights peaking for cohorts aged 69-78 (all associated with strong first-stage F-statistics). Figure A3 plots the Rotemberg-weights against the cohort-specific just-identified estimate for each age cohort: the cohorts carrying the largest positive weights form a tight cluster around $\hat{\beta}_j \approx 0.5$-$0.6$, close to and slightly above the aggregate 2SLS estimate (dashed line), indicating that the identifying variation is not driven by a single outlying cohort but by a cohesive group of similarly-estimated older cohorts. A subset of middle-aged cohorts (50-58 at the base year) instead carry negative Rotemberg-weights, jointly accounting for approximately 10% of total absolute weight; among these, cohorts aged 50-55 have both a non-trivial weight and stable, well-identified estimates systematically above the aggregate ($\hat{\beta}_j$ between 0.56 and 0.82), while cohorts aged 56-58 carry a negligible weight but highly unstable estimates (driven by a near-zero first stage for those individual cohorts) and are omitted from Figure A1 for scale. This pattern is a mechanical consequence of decomposing a growth rate rather than a level: cohorts already above the 60+ threshold at the start of the window mechanically dilute, rather than contribute to, the growth in the 60+ population share. Excluding the negative-weight cohorts shifts the estimated elasticity from 0.4405 to 0.4768, a modest and expected change given that the excluded cohorts carried above-average just-identified coefficients; the baseline estimate remains well within the confidence interval of the alternative specification. Finally, Figure A2 plots these weights directly against age cohorts. In particular, the resulting profile shows positive weights for the youngest cohorts considered, up to age 50 in 2001, after which weights turn negative through age 58. From this point onward, weights turn positive again and rise steadily among the older cohorts, which dominate the aggregate estimate. Similar patterns emerge under alternative age thresholds (65+, 70+, and 75+).[6]

[6]We repeat this shift-share diagnostic for the alternative age thresholds (65+, 70+, and 75+) and find qualitatively similar patterns: identification is always driven by older cohorts with strong first-stage predictive power, and a comparable share of negative Rotemberg-weight can be attributed to middle-aged cohorts for mechanical reasons related to the growth-rate decomposition.

4.4 Inclusion of COVID-19 Period

Table 2 reports the estimated effects of population ageing on expenditure for hospitalizations, pharmaceutical prescriptions, and outpatient services over the extended period 2011-2023 ($p = 12$), which includes the Covid-19 pandemic and the subsequent recovery years. First-stage estimates confirm that predicted ageing remains a strong predictor of the observed change in the elderly population share, with similar coefficients providing further evidence of the instrument relevance over the extended estimation window.

The IV-2SLS estimates show a broadly similar pattern in the responsiveness of healthcare expenditure to population ageing compared with the pre-COVID period 2011-2019 ($p = 8$), particularly for hospitalizations whose coefficients are almost identical. The elasticities are only slightly lower for pharmaceutical and outpatient care when these additional years are included. Overall, these results suggest that the Covid-19 pandemic has not significantly altered the dynamics between population ageing and healthcare expenditure and that our estimates are robust to longer growth periods.

Table 2: Robustness to the Inclusion of the COVID-19 Period (2011-2023, $p = 12$)

	Hospitalizations		Pharmaceutical		Outpatient		First Stage
	OLS	IV	OLS	IV	OLS	IV	
Over 60	0.550***	0.424***	0.755***	0.654***	0.317***	0.221***	1.337***
	(0.044)	(0.067)	(0.057)	(0.071)	(0.031)	(0.049)	(0.127)
Over 65	0.455***	0.382***	0.602***	0.531***	0.265***	0.207***	1.267***
	(0.031)	(0.045)	(0.044)	(0.055)	(0.023)	(0.033)	(0.103)
Over 70	0.422***	0.346***	0.529***	0.433***	0.234***	0.170***	1.190***
	(0.029)	(0.042)	(0.049)	(0.061)	(0.022)	(0.032)	(0.099)
Over 75	0.344***	0.258***	0.420***	0.298***	0.165***	0.084**	1.090***
	(0.040)	(0.049)	(0.067)	(0.073)	(0.033)	(0.043)	(0.081)
Observations	1,502		1,502		1,502		1,502

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. All specifications use $x = 10$. All regressions are weighted by the resident population aged 20+ and include ATS fixed effects. K-P Wald F Statistics: 110.41 for 60+, 152.31 for 65+, 143.49 for 70+ and 181.19 for 75+. Robust standard errors are reported in parentheses.

4.5 Heterogeneity of the Ageing Effect Across Age Groups

The effect of ageing on healthcare expenditure may differ across segments of the older population. Older adults (aged 60-74) and the oldest-old (aged 75+) differ substantially in terms of morbidity, healthcare needs, and proximity to death, suggesting potentially heterogeneous effects on healthcare spending. To examine this heterogeneity, we extend the baseline specification by including both ageing indicators simultaneously:

$$\begin{aligned}\ln\left(\frac{Y_{m,t+p}}{N_{m,t+p}}\right) - \ln\left(\frac{Y_{m,t}}{N_{m,t}}\right) = \beta_1\left[\ln\left(\frac{\text{P6074}_{m,t+p}}{N_{m,t+p}}\right) - \ln\left(\frac{\text{P6074}_{m,t}}{N_{m,t}}\right)\right] \\ + \beta_2\left[\ln\left(\frac{\text{P75}_{m,t+p}}{N_{m,t+p}}\right) - \ln\left(\frac{\text{P75}_{m,t}}{N_{m,t}}\right)\right] \\ + \gamma_j + (\varepsilon_{m,t+p} - \varepsilon_{m,t}). \qquad (4)\end{aligned}$$

where $\text{P6074}_{m,t}$ and $\text{P75}_{m,t}$ denote the population aged 60-74 and 75 and above in municipality m at time t, respectively, and all remaining variables are defined as in Equation (1). The coefficients β_1 and β_2 capture the elasticity of per-capita healthcare expenditure with respect to the older adults and oldest-old population shares, respectively, holding the other group constant.

The two endogenous regressors are jointly instrumented in a system of two first-stage equations. Each first stage includes both excluded instruments -the predicted growth in the young-old share $\left(\ln(\widehat{\text{P6074}}_{m,t+8}/\widehat{N}_{m,t+8}) - \ln(\widehat{\text{P6074}}_{m,t}/\widehat{N}_{m,t})\right)$ and the predicted growth in the oldest-old share $\left(\ln(\widehat{\text{P75}}_{m,t+8}/\widehat{N}_{m,t+8}) - \ln(\widehat{\text{P75}}_{m,t}/\widehat{N}_{m,t})\right)$ - constructed using the cohort-survival approach described in Section 3.1, applied separately to each age group. The two first-stage equations are:

$$\ln\left(\frac{\text{P6074}_{m,t+8}}{N_{m,t+8}}\right) - \ln\left(\frac{\text{P6074}_{m,t}}{N_{m,t}}\right) = \alpha_{11}\left[\ln\left(\frac{\widehat{\text{P6074}}_{m,t+8}}{\widehat{N}_{m,t+8}}\right) - \ln\left(\frac{\widehat{\text{P6074}}_{m,t}}{\widehat{N}_{m,t}}\right)\right] + \alpha_{12}\left[\ln\left(\frac{\widehat{\text{P75}}_{m,t+8}}{\widehat{N}_{m,t+8}}\right) - \ln\left(\frac{\widehat{\text{P75}}_{m,t}}{\widehat{N}_{m,t}}\right)\right] + \mu_{1t} + (\eta_{1,m,t+8} - \eta_{1,m,t}). \tag{5}$$

$$\ln\left(\frac{\text{P75}_{m,t+8}}{N_{m,t+8}}\right) - \ln\left(\frac{\text{P75}_{m,t}}{N_{m,t}}\right) = \alpha_{21}\left[\ln\left(\frac{\widehat{\text{P6074}}_{m,t+8}}{\widehat{N}_{m,t+8}}\right) - \ln\left(\frac{\widehat{\text{P6074}}_{m,t}}{\widehat{N}_{m,t}}\right)\right] + \alpha_{22}\left[\ln\left(\frac{\widehat{\text{P75}}_{m,t+8}}{\widehat{N}_{m,t+8}}\right) - \ln\left(\frac{\widehat{\text{P75}}_{m,t}}{\widehat{N}_{m,t}}\right)\right] + \mu_{2t} + (\eta_{2,m,t+8} - \eta_{2,m,t}). \tag{6}$$

where $\widehat{\text{P6074}}_{m,t}$ and $\widehat{\text{P75}}_{m,t}$ denote the predicted population shares of the respective age groups, constructed by applying the corresponding regional cohort-specific survival rates to the municipal age structure at time $t-10$. Given the presence of two endogenous regressors, we evaluate instrument strength using the Sanderson-Windmeijer conditional first-stage F-statistics rather than relying solely on the conventional first-stage F-statistic. Both conditional F-statistics are several orders of magnitude above the Stock-Yogo critical values. The Anderson canonical correlation LM test rejects underidentification, and the Cragg-Donald Wald F-statistic provides additional evidence against weak identification. Results are then reported in Table 3. Both age groups have positive and statistically significant effects on healthcare expenditure across the different domains, although the magnitude of the effects varies substantially across healthcare categories.

For hospitalizations, the estimated elasticity is larger for the oldest-old aged 75+ (0.245) than for older adults aged 60-74 (0.169), suggesting that, conditional on the share of the other age group, growth in the oldest population drives inpatient expenditure more intensively. This may reflect the fact that hospitalizations become increasingly complex and costly with the age of the patient. In contrast, pharmaceutical expenditure and

outpatient expenditure show a different age profile: the elasticities for the 60-74 group exceed that of the oldest-old, indicating that pharmaceutical and outpatient spending responds more strongly to growth in the younger elderly population.

To decompose the contribution of each age group to aggregate healthcare expenditure growth over the period 2011-2019, we combine the estimated elasticities with the observed changes in the corresponding population shares.[7] Using the estimated IV-2SLS elasticities, demographic ageing alone is estimated to account for a net expenditure increase of 4.43% for hospitalizations, 5.42% for pharmaceuticals, and 1.75% for outpatient services over the period. Importantly, because the average 60-74 population share declined slightly over 2011-2019, this group exerts a small negative contribution across all domains. It should be noted, however, that this decomposition relies on average population shares across Lombardy: municipalities that experienced an increase in their 60-74 share would still have seen their healthcare expenditure rise through this channel. Overall, the entire ageing-driven expenditure increase is therefore attributable to the oldest-old, whose share grew by 17.55% over the period: the 75+ group accounts for 4.30 percentage points increase in hospitalization expenditure, 5.32 percentage points in pharmaceutical expenditure, and 1.74 percentage points in outpatient expenditure.

Taken together, these results reveal that, despite starting from a smaller population base, the oldest-old represents the dominant contributor to ageing-related healthcare expenditure growth across all domains. This reflects both the substantial growth in the size of this population group over the period and its higher estimated healthcare expenditure intensity. Given that the oldest-old population is projected to expand rapidly in the coming decades, these findings suggest an increasing concentration of the healthcare expenditure burden associated with demographic ageing among the oldest segment of the population.

[7]The older-adults share fell slightly from 0.2096 to 0.2089, a log-change of $\ln(0.2089/0.2096) = -0.0033$ (-0.33%), reflecting the fact that large cohorts aged out of the 60-74 bracket into the 75+ group while smaller cohorts entered from below. The oldest-old share increased from 0.1204 to 0.1435, a log-change of $\ln(0.1435/0.1204) = 0.1755$ (17.55%). Multiplying each log-change by the corresponding IV elasticity yields the predicted percentage change in per-capita expenditure attributable to each group, as reported in Panel B of Table 3.

Table 3: Effects of Older Age Groups on Expenditure per Capita and Implied Expenditure Contributions (2011–2019)

	Hospitalizations	Pharmaceutical	Outpatient Services
Panel A: IV Estimates			
ln(Ages 60–74/N)	0.169***	0.394***	0.182***
	(0.048)	(0.059)	(0.042)
ln(Ages 75+/N)	0.245***	0.303***	0.099**
	(0.048)	(0.062)	(0.042)
Obs.	1,502	1,502	1,502
Panel B: Implied Expenditure Contributions from Demographic Change (2011-2019)			
Total ageing contribution	4.24%	5.19%	1.68%
Of which: oldest-old (75+)	4.30%	5.32%	1.74%
Of which: older adults (60–74)	−0.06%	−0.13%	−0.06%

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Panel A reports IV estimates from Equation (4), instrumenting both age-group regressors jointly using predicted cohort-survival growth rates for the 60-74 and 75+ groups ($x = 10$). Robust standard errors are reported in parentheses. Observations are weighted by period-t population aged 20+. Controls include ATS fixed effects. Sanderson-Windmeijer conditional first-stage F-statistics are 5,821.6 and 5,484.7, respectively, well above the Stock-Yogo critical values, indicating no evidence of weak instruments. The Anderson LM test rejects the null of underidentification ($p < 0.001$), while the Cragg-Donald Wald F-statistic equals 2,316.2, providing further evidence against weak identification. Panel B translates the IV elasticities into implied expenditure contributions by multiplying each elasticity by the observed log-change in the corresponding age group's population share between 2011 and 2019 (see Table A1).

4.6 The Role of Proximity to Death in Healthcare Expenditure

The literature on the determinants of healthcare expenditure argues that the relationship between age and healthcare spending largely reflects proximity to death rather than age *per se* - the so-called red herring hypothesis (Howdon and Rice, 2018; Werblow et al., 2007; Costa-Font and Vilaplana-Prieto, 2020; Berta et al., 2021). Under this view, the increase in healthcare expenditure associated with ageing is primarily driven by the higher prevalence of individuals approaching the end of life, who experience a substantial

expenditure spike in their final years.

Our municipality-level approach does not allow us to condition on individual time-to-death, and therefore cannot directly test the red-herring hypothesis. However, this limitation reflects a fundamental difference in the research question rather than a limitation of our contribution. The individual-level literature study whether, conditional on survival, healthcare expenditure is better explained by proximity to death or chronological age. In contrast, our analysis studies whether changes in the population age structure causally increase aggregate healthcare expenditure at the local level, and quantifies the magnitude of this effect by type of health domain. From a fiscal planning perspective, the latter question is directly policy relevant: healthcare authorities cannot condition budgets on residents' individual time-to-death, which is necessarily unpredictable ex ante, but they can observe the age composition of the population and incorporate demographic trends into expenditure planning. Regardless of the underlying individual-level mechanisms - whether end-of-life expenditure, chronic disease management, or age-related healthcare utilization- the proposed IV estimates capture the aggregate causal effect of demographic ageing on healthcare spending, under the identifying assumptions of our instrumental variable strategy.

In our setting, the estimated effect of population ageing naturally incorporates changes in the prevalence of individuals approaching their end of life. When ageing is generated by cohort-size differences, an increase in the elderly population share mechanically implies a larger number of individuals at advanced ages and, therefore, a relatively higher prevalence of individuals closer to the end of life. This is precisely the source of demographic variation exploited by our instrumental variable, which isolates changes in population ageing arising from predetermined municipal age structures combined with common cohort-specific survival patterns. Therefore, if proximity to death contributes to healthcare expenditure growth, this mechanism is part of the aggregate causal effect of demographic ageing estimated by our IV strategy.

To explore this mechanism directly, we re-estimate our main specification in Equation (1) using as outcome variable the growth in the LYL healthcare expenditure, defined

as total real healthcare expenditure incurred during the final 12 months of life by individuals who die in municipality m during the observed period, divided by the municipality's adult resident population aged 20+.

Table 4: OLS and IV Estimates for Last-Year-of-Life (LYL) Expenditure (2011-2019)

	Hospitalizations		Pharmaceutical		Outpatient	
	OLS	IV	OLS	IV	OLS	IV
Over 60	0.793***	0.689***	1.248***	1.043***	1.311***	1.196***
	(0.192)	(0.220)	(0.196)	(0.228)	(0.199)	(0.247)
Over 65	0.640***	0.542***	1.032***	0.919***	1.021***	0.909***
	(0.156)	(0.178)	(0.151)	(0.167)	(0.158)	(0.187)
Over 70	0.717***	0.612***	1.070***	0.894***	1.086***	0.960***
	(0.144)	(0.169)	(0.145)	(0.162)	(0.162)	(0.205)
Over 75	0.604***	0.444***	0.951***	0.780***	1.010***	0.926***
	(0.143)	(0.169)	(0.138)	(0.151)	(0.144)	(0.181)
Obs.	1,070		1,070		1,070	

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. The dependent variable measures the growth in per-capita end-of-life expenditure (last year of life) over the period 2011-2019. Regressions are weighted by the resident population aged 20+. All regressions include ATS fixed effects, and robust standard errors are reported in parenthesis.

Results are reported in Table 4. Note that the LYL sample contains fewer observations (1070 vs 1502) than the baseline sample because GDPR confidentiality requirements prevent the computation of LYL expenditure for small municipalities with fewer than 10 registered deaths in a given year. Table 4 shows that population ageing significantly increases LYL healthcare expenditure across all domains. The persistence of sizeable IV elasticities suggests that expenditure in the final year of life is an important channel linking demographic ageing to aggregate healthcare spending. However, since the estimated effects remain substantial and similar in magnitude to those obtained for total expenditure, the results indicate that ageing-related expenditure growth cannot be attributed exclusively to end-of-life costs.

To quantify the contribution of LYL healthcare expenditure to the overall causal effect of population ageing on healthcare spending, we combine the estimates from Tables 4 and 1 and report the obtained results in Table 5.

Overall, LYL expenditure accounts for approximately 15.8% of the estimated ageing-induced increase in healthcare expenditure over 2011-2019. The contribution of LYL

Table 5: Decomposition of Ageing-Driven Expenditure Growth: Share Attributable to the Last Year of Life

	Ageing Effect (%)		2011 Base (€)		Ageing-Driven Δ€		Share
	Total	LYL	Total	LYL	Total	LYL	due to LYL
Hospitalizations	2.80%	4.53%	581.7	73.5	16.28	3.33	20.4%
Pharmaceutical	4.94%	6.85%	228.8	10.8	11.30	0.74	6.5%
Outpatient	2.00%	7.86%	317.6	16.6	6.36	1.30	20.5%
Total			1,128.1	100.9	33.95	5.37	15.8%

Notes: Ageing effect = IV elasticity (60+, Total vs. LYL model) × log-change in the 60+ population share (6.57%, 2011-2019). Ageing-driven Δ€ = 2011 baseline expenditure × ageing effect, by domain. Share due to LYL = ageing-driven Δ€ (LYL) ÷ ageing-driven Δ€ (Total).

expenditure varies substantially across healthcare domains, accounting for 20.4% of the increase in hospitalization expenditure, 6.5% of pharmaceutical expenditure, and 20.5% of outpatient expenditure. The notably lower share for pharmaceuticals reflects the fact that drug expenditure responds strongly to ageing throughout the population, not only among individuals in their final year of life, consistent with the broader role of medication in managing chronic conditions across the elderly population. Although LYL spending represents a quantitatively relevant component of the ageing effect, the majority of the expenditure increase - around 84% in aggregate - is driven by healthcare consumption occurring outside the final year of life. These results suggest that proximity to death contributes to, but does not fully explain, the aggregate causal effect of demographic ageing on healthcare expenditure identified by our IV strategy.

4.7 Disease-Specific Effects of Population Ageing on Healthcare Expenditure

The aggregate effect of population ageing on healthcare expenditure may reflect not only spending incurred during the last year of life (LYL), but also substantial heterogeneity in the response of different disease categories to demographic change. Different health conditions may exhibit distinct age profiles, treatment patterns, and expenditure dynamics, implying that demographic ageing may affect healthcare spending through changes in the prevalence and management of specific diseases (e.g. chronic conditions such as

oncological or cardiovascular diseases).

To investigate whether the effect of population ageing differs across health conditions, we extend our baseline IV specification by focusing on disease-specific healthcare expenditure. Exploiting administrative records, we identify individuals affected by specific health conditions based on exemption categories. For each disease group, we then measure total healthcare expenditure incurred by these individuals across different care domains, including hospitalizations, pharmaceuticals, and outpatient care, regardless of the specific cause of utilization. We then utilize this expenditure (per total 20+ population per capita), as our dependent variable, while the rest of the model remains similar to Equation 1. This approach allows us to estimate whether ageing has heterogeneous effects on disease-specific healthcare expenditure. We consider two major disease groups: oncological and cardiovascular conditions, which represent key components of morbidity and healthcare utilization among older individuals.

Table A12 reports estimates for expenditure incurred by oncological and Table A13 for expenditure by cardiovascular patients, with Figure 4 summarizing the IV elasticities across care domains and disease groups alongside the baseline estimates for comparison.

For hospitalization expenditure, the estimated elasticities for both oncological and cardiovascular expenditure are quantitatively similar to, and in several cases somewhat larger than, the corresponding baseline elasticities across all age thresholds. This suggests that demographic ageing exerts a comparable, if not slightly stronger, effect on inpatient care utilization among individuals with these two major chronic conditions. This pattern is consistent with hospitalization needs for both disease groups being driven primarily by acute episodes and disease severity rather than by age-specific treatment intensity margins.

The pattern differs for pharmaceutical expenditure. Among oncological patients, the estimated elasticities substantially exceed the baseline at every age threshold, ranging from 1.088 at the Over 60 threshold to 1.185 at Over 70 before declining to 0.754 at Over 75, and remain above unity for three of the four thresholds. This suggests that pharmaceutical spending among oncological patients increases more than proportionally

with demographic ageing over most of the elderly age range.

By contrast, cardiovascular pharmaceutical elasticities are considerably smaller, ranging from approximately 0.35 to 0.60 across the age thresholds. Although their point estimates are broadly comparable to the baseline, they are estimated with substantially greater uncertainty, with confidence intervals generally spanning zero.

Outpatient expenditure exhibits the most pronounced heterogeneity by disease group. While baseline elasticities are modest (ranging from 0.108 to 0.300) and decline monotonically with the age threshold, the oncological subpopulation displays substantially larger point estimates throughout, reaching statistical significance only at the older thresholds, where standard errors are comparatively smaller. The imprecision observed at the Over 60 and Over 65 thresholds may reflect greater heterogeneity in outpatient care pathways among younger oncological patients. Cardiovascular outpatient elasticities, by contrast, remain closer to the baseline magnitudes and reach statistical significance only at the Over 75 threshold, indicating a comparatively modest and more age-concentrated ageing effect on cardiovascular outpatient care.

Taken together, these disease-specific results indicate that the aggregate ageing effect on pharmaceutical and outpatient expenditure documented in Section 4.2 masks considerable heterogeneity across diagnosis-specific expenditure. Oncological expenditure displays substantially larger responses to population ageing than cardiovascular expenditure or overall expenditure, particularly for pharmaceutical spending across all age groups and for outpatient care among the oldest patients.

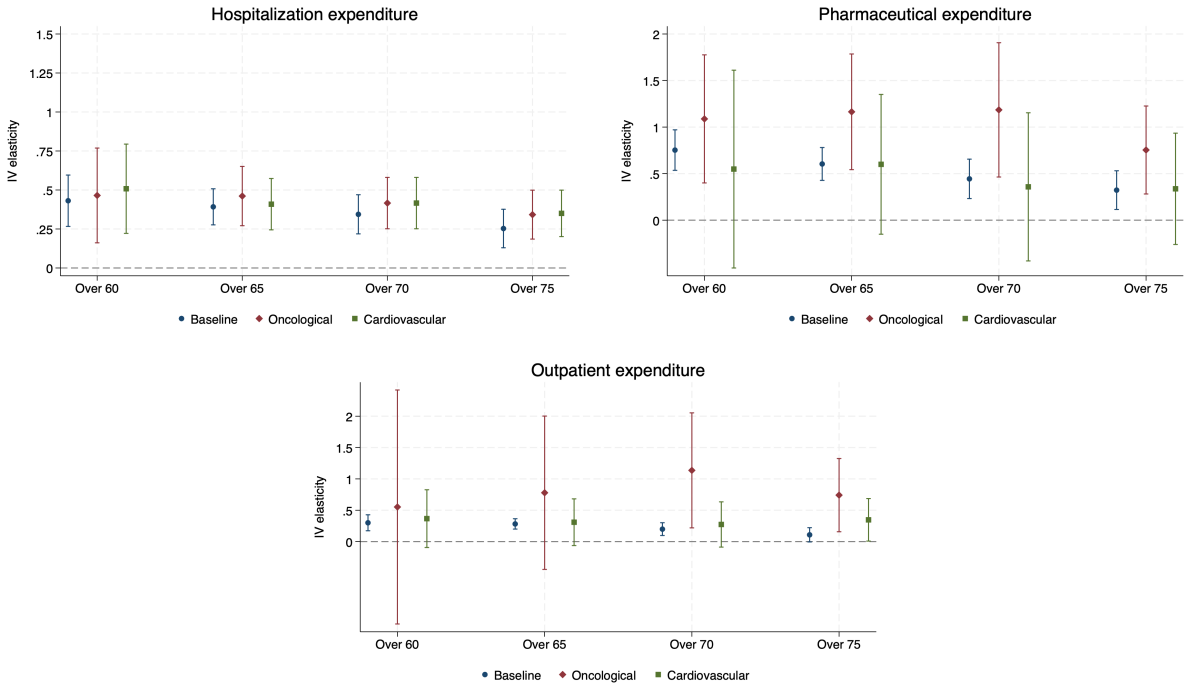


Figure 4: Heterogeneous effects of population ageing by disease category and care domain

5 Robustness and Sensitivity Analysis

This section tests the robustness of our IV estimates to alternative identification choices and sample compositions. In particular, we provide the following robustness checks: i) including economic local factors determined by population ageing; ii) using longer historical lags ($x = 15$ and $x = 20$) to construct our IV; iii) excluding Milan and provincial capitals from the sample; iv) adopting alternative temporal aggregation pooling together two consecutive periods of six-years for each municipality and introducing period fixed effects; v) assess whether the instrument has predictive power for migration flows as an alternative channels that could violate the exclusion restriction; and vi) accounting for spatial correlation across municipalities.

5.1 Controlling for local economic trends

Population ageing may affect local economic outcomes such as employment or income growth (Maestas et al., 2023). As discussed in section 4.1, healthcare expenditure should be independent on those, since this is financed and decided at the regional level. Nevertheless, as robustness check, we augment the baseline specification including $\Delta \ln$(average income) and $\Delta \ln$(employment rate) at local level as controls. IV-2SLS estimates remain virtually unchanged in magnitude and statistical significance (Table 6), confirming that our main results are robust to the inclusion of local economic trends.

Table 6: OLS and IV Elasticities by Care Domain with Time-Varying Controls

	Hospitalizations		Pharmaceutical		Outpatient		First Stage
	OLS	IV	OLS	IV	OLS	IV	
Over 60	0.546***	0.426***	0.869***	0.752***	0.418***	0.305***	1.364***
	(0.060)	(0.080)	(0.092)	(0.103)	(0.038)	(0.059)	(0.121)
Over 65	0.457***	0.385***	0.696***	0.598***	0.347***	0.281***	1.280***
	(0.044)	(0.057)	(0.074)	(0.085)	(0.027)	(0.039)	(0.098)
Over 70	0.442***	0.337***	0.615***	0.434***	0.303***	0.194***	1.144***
	(0.049)	(0.061)	(0.093)	(0.101)	(0.036)	(0.050)	(0.092)
Over 75	0.354***	0.248***	0.466***	0.315***	0.205***	0.105***	1.088***
	(0.050)	(0.059)	(0.098)	(0.099)	(0.045)	(0.055)	(0.073)
Observations	1,502		1,502		1,502		1,502

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. All regressions are weighted by municipality population aged 20+. Controls include ATS fixed effects, $\Delta \ln$(average income), and $\Delta \ln$(employment rate). All regressions adopt an eight-year growth window (2011-2019, $p = 8$) with the standard instrument lag of 10 years ($x = 10$). Robust standard errors are reported in parentheses.

5.2 Alternative IV Lags

To address concerns that predetermined age structures may be correlated with slowly evolving local trends that also affect contemporaneous healthcare expenditure growth, we reconstruct the instrument using age structures from fifteen and twenty years prior ($x = 15$ and $x = 20$), instead of the baseline 10-year lag. Exploiting more distant historical age structures reduces the possibility that the instrument captures municipality-specific trends that directly influence current healthcare expenditure growth (Maestas et al., 2023). Tables A3-A4 in the Appendix report the estimates obtained using these alternative lag structures. First-stage coefficients remain large and statistically significant, confirming that the instrument retains substantial predictive power. Moreover, the IV elasticities remain quantitatively similar to those obtained in the baseline specification. Overall, these results indicate that the estimated effect of population ageing is not sensitive to the choice of the instrumental lag and provide additional support for the plausibility of the exclusion restriction.

5.3 Excluding Milan and Province Capitals

To assess whether our results are driven by the distinctive demographic and healthcare dynamics of large urban centres, we re-estimate the main specification after sequentially excluding the municipality of Milan (the regional capital, accounting for 13.7% of Lombardy's population) and the municipalities of all twelve province capitals[8], which together represent 22.7% of the regional population; the remaining 1490 municipalities within each province are retained in the sample. Results are reported in Tables A5-A6 in the Appendix.

Two findings emerge from this exercise. First, the first-stage coefficients are somewhat smaller when large urban municipalities are excluded. A possible explanation is that urban areas may exhibit more stable age structures over time, allowing historical cohort composition to better predict subsequent demographic ageing. In contrast, smaller

[8]The province capitals excluded in this robustness exercise are: Bergamo, Brescia, Como, Cremona, Lecco, Lodi, Mantova, Milano, Monza, Pavia, Sondrio, and Varese.

municipalities may experience greater population turnover, including selective migration of younger cohorts, which can reduce the predictive power of historical age structures for future population composition.

Second, and more importantly, the IV estimates remain quantitatively very similar across both exercises and are not statistically different from those obtained in the baseline specification. This stability indicates that the estimated effect of population ageing is not driven by specific demographic or healthcare dynamics in large urban centres, but represents a broader feature of the municipal-level relationship between ageing and healthcare expenditure in Lombardy.

5.4 Alternative Temporal Aggregation: Six-Year Growth Windows

As an additional robustness check, we re-estimate the model by pooling two consecutive six-year growth periods, $\tau \in \{2011\text{-}2017, 2017\text{-}2023\}$, and including period fixed effects. This specification has two advantages. First, it more closely replicates the multi-period panel structure adopted by Maestas et al. (2023), in which the same units (i.e. municipalities) are observed across repeated intervals, allowing us to exploit within-municipality variation over multiple periods while controlling for common time shocks through period fixed effects. Including period fixed effects, in fact, controls for common technological progress and other time-varying factors that may affect healthcare expenditure uniformly across municipalities, such as changes in medical practice, treatment availability, and healthcare policies. It also allows us to include initial-period income and employment levels as additional controls both interacted with period fixed effects, rather than their growth rates, which may themselves be partially endogenous to population ageing. Second, pooling multiple periods brings the advantage of reconstructing the instrument independently within each time window: the 2001 age structure is used to predict ageing between 2011 and 2017, while the 2007 age structure is used to predict ageing between 2017 and 2023. The consistency of the IV estimates across these periods therefore provides additional evidence that the results are not driven by the choice of a particular

baseline year for instrument construction.

Table 7 reports the results. The sign and statistical significance of the IV estimates across healthcare domains are broadly consistent with the baseline specification, although some differences emerge when using the shorter aggregation window. In particular, the IV estimates remain generally positive, with the exception of the imprecisely estimated Over75 outpatient coefficient, and the OLS-IV gap continues to increase for older age thresholds, replicating the pattern of upward endogeneity bias documented in Section 4.2.

Quantitatively, the 6-year IV elasticities are generally smaller than the 8-year estimates reported in Table 1. For hospitalizations, the difference is limited and the estimates remain statistically significant, suggesting that the response of inpatient care to demographic ageing materializes relatively quickly. For pharmaceutical expenditure, the attenuation is larger but the estimates remain statistically significant, consistent with medication use and chronic disease management adjusting over a longer horizon. For outpatient expenditure, instead, the shorter aggregation window leads to a more pronounced reduction in the estimated elasticities, with statistical significance no longer preserved. This suggests that outpatient care may require a longer adjustment period to fully capture the expenditure response associated with demographic ageing.

Overall, these findings confirm the robustness of the main conclusions to an alternative temporal aggregation, while also providing suggestive evidence that the timing of the expenditure response differs across healthcare domains.

Table 7: Robustness: Alternative Growth Period ($p = 6$, 2011-2017 and 2017-2023)

	Hospitalizations		Pharmaceutical		Outpatient		First Stage
	OLS	IV	OLS	IV	OLS	IV	
Over 60	0.634***	0.421***	0.798***	0.221***	0.427***	0.038	1.187***
	(0.080)	(0.084)	(0.099)	(0.073)	(0.072)	(0.111)	(0.049)
Over 65	0.519***	0.392***	0.621***	0.172***	0.346***	0.005	1.171***
	(0.060)	(0.060)	(0.075)	(0.055)	(0.053)	(0.094)	(0.042)
Over 70	0.429***	0.276***	0.449***	0.124***	0.253***	0.045	1.081***
	(0.048)	(0.049)	(0.060)	(0.041)	(0.039)	(0.069)	(0.031)
Over 75	0.339***	0.228***	0.347***	0.0002	0.167***	-0.091	1.047***
	(0.045)	(0.048)	(0.057)	(0.044)	(0.038)	(0.066)	(0.030)
Observations	1,502		1,502		1,502		1,502

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. All specifications construct the instrument with a 10-year lag ($x = 10$). All regressions are weighted by the resident population aged 20+ and include ATS fixed effects. Standard errors clustered at municipality level are reported in parentheses. Controls include income at time t interacted with time fixed effects.

5.5 Migration-Related Threats to Identification

A potential threat to our identification strategy arises if the instrument captures migration dynamics that affect healthcare expenditure through channels unrelated to population ageing itself. Following Maestas et al. (2023), who assess this concern by using total population size as outcome variables and find no significant migration response at the state level, we implement a more direct test suited to our municipality-level setting. Specifically, we replace healthcare expenditure per capita with the net migration rate at municipality level as the dependent variable and re-estimate the IV specification separately for each destination age group. The net migration rate is defined as cumulative inflows minus cumulative outflows over the period 2011-2019 divided by the population aged 20 and over in 2011.

Table 8 reports results for age-group specific net migration rates. Panel A considers migration between Lombardy municipalities and all other Italian municipalities, while Panel B restricts migration to movements within Lombardy. The results reveal a consistent and interpretable pattern. Population ageing is associated with positive net migration among individuals aged 0-17, particularly at the higher ageing thresholds. In Panel A, ageing is also associated with significant net outflows of young adults aged 18-34 at the 60+, 65+, and 70+ thresholds. By contrast, in Panel B the estimated coefficients for the

18-34 age group are positive, suggesting that population ageing is associated with net inflows of young adults within Lombardy. The positive net migration observed among individuals aged 0-17 is consistent with family migration, whereby children relocate together with their parents rather than in response to independent economic incentives. The contrasting patterns for young adults suggest that, while older municipalities tend to lose young residents to other Italian regions, they may still attract or retain young adults through migration within Lombardy. Importantly, and following the argument in Maestas et al. (2023), this youth outflow need not represent a threat to our identification strategy: If youth outmigration is a consequence of demographic ageing rather than an independent response to the instrument (i.e. cohort-based predicted ageing), it represents a potential mechanism through which ageing affects healthcare expenditure rather than a violation of the exclusion restriction. Since young adults are below-average users of healthcare, their departure reduces total spending proportionally less than it reduces the adult population, mechanically raising per-capita expenditure independently of any true ageing effect. Thus our IV estimates are expected to be higher than in the case when population ageing does not induce any youth emigration.

For older migrants (65+), we find positive and significant effects on the net migration rate at the 60+ and 65+ thresholds in Panel A and only at the 60+ threshold in Panel B, indicating that municipalities experiencing faster ageing exhibit positive net migration among older residents. Since older net migrants are above-average users of healthcare, this channel would potentially bias our IV estimates upward. The economic magnitude of these effects is, however, very small. For example, the coefficient of 0.0290 for the 60+ threshold in Panel A implies that a 1% increase in the share of individuals aged 60 and over (i.e. a log change of 0.01) is associated with an increase of only 0.00029 net elderly migrants per adult resident. Evaluated at the average municipality in our sample (5,276 adult residents), this corresponds to approximately 1.5 additional net elderly migrants over the entire 2011-2019 period. Relative to the average elderly population of about 1,741 residents, this amounts to less than 0.1%, indicating that the estimated migration response is economically negligible despite its statistical significance. In every

case, these migration effects disappear and become statistically indistinguishable from zero at the 70+ and 75+ thresholds in both panels, precisely the age groups for which our IV estimates are most distinct from OLS and most policy-relevant.

Taken together, these results suggest that migration dynamics are unlikely to represent a major threat to the validity of our IV strategy. The instrument predicts selective outmigration of young adults, which is consistent with migration responses associated with demographic decline and therefore represents, if anything, a potential mechanism through which ageing affects per-capita healthcare expenditure rather than an independent confound. Conversely, positive net migration of older individuals, which would pose a more direct threat to the exclusion restriction by increasing healthcare demand independently of ageing, is limited to the younger elderly thresholds (60+ and 65+) and disappears for the oldest age groups (70+ and 75+). Overall, these findings indicate that the estimated effect of population ageing is unlikely to be driven by migration-related changes in population composition rather than by the ageing process itself.

However, a final concern is that the proposed instrument could proxy for time-invariant local amenities (e.g., healthcare quality and access, cost of living) that attract elderly migrants, so that part of the estimated effect reflects amenity-driven in-migration rather than ageing *per se*. Section 5.5.1 bounds the maximum plausible contribution of this channel using a Conley et al. (2012) plausibly-exogenous exercise calibrated on the observed elderly migration response (Table 8), providing a quantitative -if partial- check on the amenities concern.

Table 8: Net Migration Rate (per Population over 20, 2011)

	0-17	18-34	35-64	Over 65
Panel A: Overall Italy Migrants				
Over 60	0.0117	-0.0842**	-0.0139	0.0290**
	(0.0113)	(0.0398)	(0.0151)	(0.0144)
Over 65	0.0160**	-0.0691**	-0.0032	0.0220*
	(0.0080)	(0.0330)	(0.0109)	(0.0122)
Over 70	0.0230***	-0.0526*	-0.0012	0.0027
	(0.0064)	(0.0310)	(0.0098)	(0.0137)
Over 75	0.0218***	-0.0412	-0.0002	-0.0013
	(0.0056)	(0.0272)	(0.0085)	(0.0120)
Obs.	1,502	1,502	1,502	1,502
Panel B: Within Lombardy Migrants				
Over 60	0.0205*	0.0345**	0.0220	0.0211*
	(0.0107)	(0.0146)	(0.0189)	(0.0118)
Over 65	0.0220***	0.0259**	0.0235	0.0157
	(0.0077)	(0.0100)	(0.0144)	(0.0099)
Over 70	0.0258***	0.0119	0.0164	-0.0011
	(0.0064)	(0.0108)	(0.0136)	(0.0114)
Over 75	0.0235***	0.0083	0.0136	-0.0038
	(0.0057)	(0.0093)	(0.0118)	(0.0099)
Obs.	1,502	1,502	1,502	1,502

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. The dependent variable is the net migration rate by destination age group, defined as net migration (cumulative in-migration minus cumulative out-migration, 2011-2019) divided by the population aged 20 and over in 2011. Panel A refers to migrants from any origin; Panel B restricts the sample to migrants originating from and remaining within Lombardy. Robust standard errors are reported in parentheses.

5.5.1 Economically Motivated Bounds from Elderly Migration

In this section, we quantify the potential relevance of the population ageing-induced migration channel discussed above. We follow Conley et al. (2012), by monetizing the estimated migration coefficient into an implied bias $\tilde{\gamma}_{mig}$ on per-capita healthcare expenditure growth, holding the ageing channel fixed. Every input to this calibration is drawn directly from our administrative data.

We combine the IV-estimated effect of population ageing on net elderly (65+) in-migration over 2011-2019 (Table 8, already expressed in per capita terms of the 2011 population aged 20+) with the actual 2011 per-capita cost differential between an elderly resident and the municipal average. Elderly residents generate average healthcare costs of €2,324 per year compared to €1,158 for the overall population aged 20+, a ratio of roughly 2.0[9]. Because the migration effect is already normalized by the 2011 population aged 20+, no further population rescaling is needed: dividing this effect by the baseline expenditure ratio directly yields $\hat{\gamma}_{mig}$ in the same elasticity units as $\hat{\beta}_{IV}$.

Table 9 reports the resulting identified sets obtained as described in Appendix E, also represented graphically in Figure 5.

The implied violations, measured by $\tilde{\gamma}_{mig}$, are small relative to the estimated IV elasticities, ranging from 0.0012 (75+ threshold) to 0.0211 (60+ threshold). This reflects the fact that, although the ageing instrument predicts a statistically significant increase in net elderly in-migration, the associated inflow represents only a small fraction of the average municipal population and therefore generates only a limited direct effect on healthcare expenditure. Consequently, the identified sets remain tightly centered around the baseline IV estimates, and the corresponding confidence intervals obtained are only slightly wider than the conventional 95% confidence intervals reported in Table 1.

As an additional robustness check, we re-compute $\tilde{\gamma}_{mig}$ assuming a cost differential of 3.5 (rather than 2.0) to account for the possibility that health-motivated elderly migrants

[9]These average per-capita figures (€1,158 for the population aged 20+ and €2,324 for residents aged 65+) may appear low relative to total regional healthcare spending, which amounted to approximately €20 billion in 2019. This is because our expenditure measure captures only the reimbursed cost of hospitalizations, pharmaceutical prescriptions, and outpatient cares. It does not include emergency department visits and fixed system-level costs such as medical staff salaries, facility maintenance, and general healthcare administration, which are not reported in the available data.

are negatively selected on health status relative to the average elderly resident. The resulting bounds widen only modestly and continue to exclude zero across all specifications (see Figure A4), confirming that our conclusions are robust even under a substantially more conservative assumption regarding migrants' healthcare costs.

This calibration is entirely anchored to the administrative data rather than external assumptions and directly quantifies the importance of the specific mechanism identified as a potential threat to identification. The sensitivity parameter $\tilde{\gamma}_{mig}$ is calibrated using the observed magnitude of elderly in-migration and alternative assumptions on the relative healthcare cost of elderly residents; the reported bounds and confidence intervals are therefore conditional on this calibrated value and do not incorporate sampling uncertainty from the calibration step.

Even under the empirically observed magnitude of selective elderly in-migration towards faster-ageing municipalities, potentially reflecting better access to healthcare services or higher-quality care, the implied violation of the exclusion restriction remains economically small. All twelve identified sets and their associated confidence intervals continue to exclude zero by a comfortable margin. The corresponding confidence intervals also exclude zero in all but one case, namely outpatient expenditure for the 75+ threshold, where the interval marginally includes zero. Overall, the results indicate that the empirically observed magnitude of selective elderly in-migration has only a limited impact on inference and does not alter the substantive conclusions of the analysis.

Table 9: Plausibly Exogenous Bounds: Elderly Migration-Based Calibration

Threshold	Outcome	$\hat{\beta}_{IV}$	$\tilde{\gamma}_{mig}$	Identified Set		95% CI	
				Lower	Upper	Lower	Upper
60+	Hosp.	0.431	0.0292	0.402	0.460	0.237	0.625
60+	Pharma	0.753	0.0292	0.724	0.782	0.506	1.000
60+	Outpatient	0.300	0.0292	0.271	0.329	0.143	0.457
65+	Hosp.	0.392	0.0222	0.370	0.414	0.254	0.530
65+	Pharma	0.604	0.0222	0.582	0.626	0.405	0.803
65+	Outpatient	0.282	0.0222	0.260	0.304	0.178	0.386
70+	Hosp.	0.344	0.0027	0.341	0.347	0.216	0.472
70+	Pharma	0.444	0.0027	0.441	0.447	0.230	0.658
70+	Outpatient	0.199	0.0027	0.196	0.202	0.094	0.304
75+	Hosp.	0.253	0.0013	0.252	0.254	0.128	0.378
75+	Pharma	0.323	0.0013	0.322	0.324	0.114	0.532
75+	Outpatient	0.108	0.0013	0.107	0.109	-0.007	0.223

Notes: $\tilde{\gamma}_{mig} = (\hat{\beta}^{net}_{mig}) \cdot (\bar{e}_{65+} - \bar{e}_{tot})/\bar{e}_{tot}$, where $\hat{\beta}^{net}_{mig}$ is the causal (2SLS) effect of realized population ageing (X, instrumented by Z) on net elderly in-migration over 2011-2019 (Table 8, Panel A, column "Over 65"), expressed per capita of the 2011 population aged 20+ (see Appendix E for the derivation). $\bar{e}_{tot} =$ €1,158.13 and $\bar{e}_{65+} =$ €2,324.22 denote actual 2011 per-capita healthcare expenditure for the total and elderly populations, respectively. The Identified Set is computed as $[\hat{\beta}_{IV} - |\hat{\gamma}_{mig}|,\ \hat{\beta}_{IV} + |\hat{\gamma}_{mig}|]$. Confidence intervals are computed using the conservative critical value of 1.96.

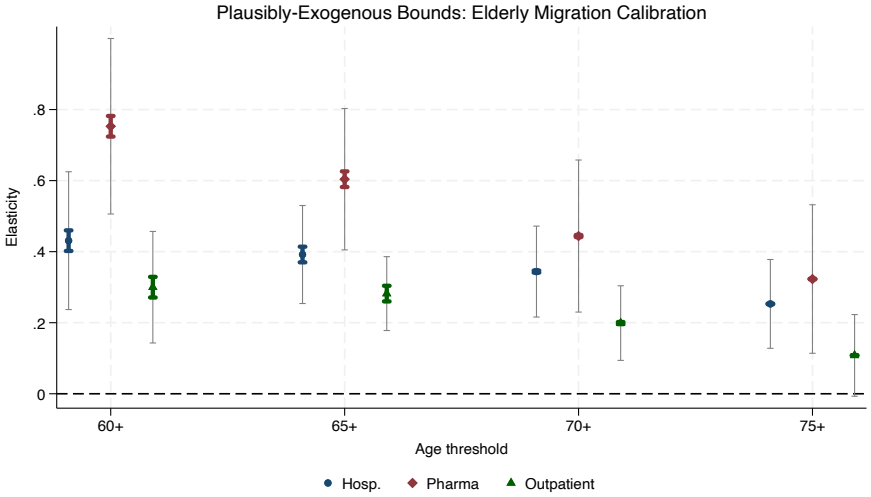


Figure 5: CHR bounds with 95% CI, assuming $\frac{e_{tot}}{e_{65+}} = 2$

5.6 Spatial correlation

As a final robustness check, we address the concern that healthcare expenditure growth may exhibit spatial dependence across neighbouring municipalities, which share common

local shocks, infrastructure, or demographic dynamics and could lead to underestimated standard errors. We re-estimate the baseline specifications using spatially robust standard errors following Conley (1999), allowing the error covariance to decay with geographical distance. As reported in Appendix F, the estimated elasticities remain virtually unchanged relative to the baseline, and statistical significance is largely unaffected across all three healthcare domains and age thresholds, with the only exception being the IV estimate for outpatient expenditure at the 75+ threshold, which becomes imprecisely estimated. Overall, these results confirm that spatial correlation in residuals does not affect our main findings.

6 Quantifying the Fiscal Cost of Ageing

6.1 The Realized Fiscal Impact, 2011-2019

We translate the per-capita ageing elasticities (Table 5) into an aggregate monetary impact on Lombardy's healthcare budget. Table 10 reports the results.

Ageing-driven aggregate contributions were +€129.0 million for hospitalizations, +€89.5 million for pharmaceuticals, and +€50.4 million for outpatient care. For pharmaceuticals this accounts for roughly 70% of the total observed increase (+€127.4 million); for outpatient care, roughly 16% of the observed increase (+€316.6 million); for hospitalizations, aggregate expenditure fell by €545.5 million despite the positive ageing contribution, implying ageing offset about 19% of an underlying structural decline.

Summing across domains, population ageing (60+) added approximately €269 million (95% CI: €209-€329 million) to Lombardy's healthcare expenditure over 2011-2019 - about 1.3% of a single year's €20 billion regional health budget. Using the 2019 population base instead yields €277 million.

This should be read as a lower bound. The 2011-2019 period saw only moderate ageing (the 60+ share rose from 33.0% to 35.2%); the oldest-old, whose expenditure response is largest and most persistent, are set to expand fastest in the coming decades. Part of the estimated impact was also masked by a structural decline in hospitalization spending that may not persist.

Table 10: Aggregate Fiscal Impact of Population Ageing (60+) on Healthcare Expenditure, 2011-2019

	Total Expenditure (€million)		Observed	Ageing-Driven
	2011	2019	Aggregate Δ (€M)	Aggregate (€M)
Hospitalizations	4,608.45	4,062.94	−545.5	+129.0
Pharmaceutical	1,812.52	1,939.87	+127.4	+89.5
Outpatient	2,516.32	2,832.93	+316.6	+50.4

Notes: Total expenditure levels are aggregate real regional healthcare expenditure by domain, constructed by summing individual-level administrative expenditure records, as described in Section 2, in 2010 euros. Observed Aggregate Δ is the difference between the 2019 and 2011 total expenditure columns. Ageing-Driven Aggregate is obtained by multiplying the per-capita ageing-driven effect reported in Table 5 ("Total" column) by the Lombardy population aged 20+ in 2011 (approximately 7.92 million); per-capita expenditure levels underlying this calculation are reported in Table A2. For hospitalizations, where the Observed Aggregate Δ is negative despite a positive Ageing-Driven Aggregate, the latter reflects the extent to which ageing offset an underlying structural decline, not a share of observed growth; see main text for discussion.

7 Conclusions

This paper investigates the causal impact of population ageing on healthcare expenditure using rich administrative data from Lombardy and a municipality-level instrumental variable design exploiting lagged age structures combined with cohort-specific survival rates.

For the pre-COVID period (2011-2019), a 1% increase in the share of the population aged 60 and above raises per-capita expenditure by approximately 0.43% for hospitalizations, 0.75% for pharmaceutical consumption, and 0.30% for outpatient services, with pharmaceutical spending showing more-than-proportional responses. IV estimates are consistently smaller than their OLS counterparts, by 30-48% at older age thresholds, consistent with upward bias from reverse causality and healthcare-driven elderly migration, and indicating that conventional estimates overstate the causal effect of ageing.

Decomposing the effect between older adults (60-74) and the oldest-old (75+) shows that the latter drive hospitalization costs (elasticity 0.245 vs 0.169), while the young-old drive pharmaceutical and outpatient spending more strongly (0.394 vs 0.303, and 0.182 vs 0.099, respectively). Despite starting from a smaller population base, the oldest-old account for the majority of ageing-driven expenditure growth (4.30 of 4.24%, 5.32 of 5.19%, and 1.74 of 1.68% for hospitalizations, pharmaceuticals, and outpatient care), a concentration expected to intensify as this segment becomes the fastest-growing in the coming decades. Re-estimating the model on last-year-of-life (LYL) expenditure shows that proximity to death accounts for only 15.8% of the ageing effect, indicating that ageing operates largely independently of the "red herring" mechanism. Disaggregating by disease, oncological patients drive pharmaceutical elasticities above unity, while cardiovascular responses are smaller and imprecisely estimated; hospitalization elasticities are comparable across both groups, consistent with acute severity rather than age-specific treatment driving inpatient costs.

The instrument's credibility rests on an extensive validation exercise: placebo tests on municipal spending categories unrelated to healthcare (e.g., roads) show no association with predicted ageing; migration responses to predicted ageing were positive and signifi-

cant at the 60+ and 65+ thresholds, motivating a formal sensitivity analysis: plausibly exogenous bounds calibrated on the observed magnitude of elderly migration (Conley et al., 2012) show that this channel is quantitatively too small to threaten the exclusion restriction, excluding a null effect for almost all specifications; and a Rotemberg-weight decomposition of the linearized instrument (Goldsmith-Pinkham et al., 2020) confirms that identification is driven by a cohesive cluster of older birth cohorts rather than any single cohort or municipality. Results are also robust to spatially correlated standard errors and to extending the sample through the COVID-19 period (2011-2023), with hospitalization elasticities virtually unchanged and pharmaceutical/outpatient elasticities only modestly lower.

Taken together, population ageing (60+) is estimated to have added approximately €269 million to Lombardy's healthcare expenditure over 2011-2019, which represents about 1.3% of a single year's €20 billion regional budget. This modest aggregate figure reflects the relatively moderate pace of ageing observed in our sample and should be read as an early signal, not a ceiling: the fiscal pressure documented here is likely to intensify as ageing accelerates and the oldest-old, whose expenditure response we find largest and most persistent, expand most rapidly in the coming decades. Our findings from Italy, at an advanced stage of demographic transition, offer a preview of the fiscal challenges facing healthcare systems in countries earlier in that transition.

References

Acemoglu, D. and Restrepo, P. (2017). Secular stagnation? the effect of aging on economic growth in the age of automation. *American Economic Review*, 107(5):174–179.

Bech, M., Christiansen, T., Khoman, E., Lauridsen, J., and Weale, M. (2011). Ageing and health care expenditure in eu-15. *The European Journal of Health Economics*, 12(5):469–478.

Berta, P., Lovaglio, P. G., and Verzillo, S. (2021). How have casemix, cost and hospital stay of inpatients in the last year of life changed over the past decade? evidence from italy. *Health Policy*, 125(8):1031–1039.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. *The Review of Economic Studies*, 89(1):181–213.

Commission, E. (2024). 2024 ageing report: Economic & budgetary projections for the eu member states (2022-2070). *European Economy, Institutional Paper*, 279.

Conley, T. G. (1999). Gmm estimation with cross sectional dependence. *Journal of Econometrics*, 92(1):1–45.

Conley, T. G., Hansen, C. B., and Rossi, P. E. (2012). Plausibly exogenous. *Review of Economics and Statistics*, 94(1):260–272.

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.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Guthmuller, S., Paruolo, P., and Verzillo, S. (2021). Positive externalities of eu actions on sustainability of health systems. In Baltagi, B. H. and Moscone, F., editors, *The Sustainability of Health Care Systems in Europe*. Emerald Publishing Limited.

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.

Imbens, G. W. and Manski, C. F. (2004). Confidence intervals for partially identified parameters. *Econometrica*, 72(6):1845–1857.

Maestas, N., Mullen, K. J., and Powell, D. (2023). The effect of population aging on economic growth, the labor force, and productivity. *American Economic Journal: Macroeconomics*, 15(2):306–332.

Piscopo, J., Groot, W., and Pavlova, M. (2024). Determinants of public health expenditure in the eu. *Plos One*, 19(3):e0299359.

Sheiner, L. (2014). The determinants of the macroeconomic implications of aging. *American Economic Review*, 104(5):218–223.

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.

Appendices

Appendix A Additional Tables

This section reports additional descriptive statistics and robustness checks. Table A1 presents the evolution of the share of the population belonging to different age groups over time, with each year weighted by the population size at time t. Table A2 reports descriptive statistics on real per capita healthcare expenditure by care domain in 2010 euros, distinguishing between total expenditure and LYL expenditure.

The remaining tables provide robustness checks for the main IV estimates. Tables A3 and A4 assess the sensitivity of the results to alternative values of the instrument lag. Tables A5 and A6 examine whether the results are driven by specific urban areas by excluding, respectively, the municipality of Milan and all provincial capitals. Across these specifications, the estimates remain qualitatively similar to the baseline results.

Year	Share60+	Share65+	Share70+	Share75+	Share60-74
2011	0.3299	0.2506	0.1865	0.1204	0.2096
2012	0.3326	0.2557	0.1897	0.1228	0.2098
2013	0.3343	0.2592	0.1916	0.1255	0.2088
2014	0.3365	0.2634	0.1934	0.1295	0.2070
2015	0.3399	0.2676	0.1958	0.1340	0.2059
2016	0.3430	0.2710	0.1965	0.1378	0.2052
2017	0.3466	0.2737	0.2014	0.1407	0.2060
2018	0.3495	0.2758	0.2048	0.1424	0.2072
2019	0.3523	0.2774	0.2082	0.1435	0.2089
2020	0.3560	0.2796	0.2112	0.1446	0.2114
2021	0.3575	0.2795	0.2113	0.1424	0.2151
2022	0.3625	0.2823	0.2133	0.1461	0.2164
2023	0.3657	0.2834	0.2136	0.1479	0.2179

Table A1: Share of population by year weighted by population at time t

Panel A: Total Expenditure			
Year	Hosp Total	Pharma Total	Outpatient Total
2011	581.65	228.77	317.59
2017	513.99	222.32	335.30
2019	497.53	237.67	346.19
2023	430.10	207.95	320.48
Panel B: Last-Year-of-Life Expenditure			
Year	Hosp LYL	Pharma LYL	Outpatient LYL
2011	73.50	10.79	16.55
2017	82.11	10.80	16.68
2019	90.72	10.11	11.48
2023	60.07	11.69	14.91

Table A2: Real per capita expenditure (in 2010 euros)

Table A3: Robustness: Alternative Instrument Lag ($x = 15$)

	Hospitalizations		Pharmaceutical		Outpatient		First Stage
	OLS	IV	OLS	IV	OLS	IV	
Over 60	0.544***	0.454***	0.862***	0.768***	0.401***	0.305***	1.286***
	(0.063)	(0.083)	(0.099)	(0.113)	(0.043)	(0.065)	(0.125)
Over 65	0.460***	0.401***	0.697***	0.615***	0.340***	0.284***	1.216***
	(0.046)	(0.059)	(0.079)	(0.092)	(0.030)	(0.043)	(0.100)
Over 70	0.448***	0.336***	0.622***	0.438***	0.302***	0.188***	1.072***
	(0.051)	(0.066)	(0.098)	(0.112)	(0.038)	(0.055)	(0.099)
Over 75	0.362***	0.258***	0.477***	0.315***	0.208***	0.104*	1.008***
	(0.053)	(0.063)	(0.105)	(0.108)	(0.048)	(0.059)	(0.078)
Observations	1,502		1,502		1,502		1,502

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. All specifications use an instrument lag of $x = 15$. Regressions are weighted by resident population aged 20+ and include ATS fixed effects. Robust standard errors are reported in parenthesis.

Table A4: Robustness: Alternative Instrument Lag ($x = 20$)

	Hospitalizations		Pharmaceutical		Outpatient		First Stage
	OLS	IV	OLS	IV	OLS	IV	
Over 60	0.544***	0.455***	0.862***	0.761***	0.401***	0.305***	1.236***
	(0.063)	(0.084)	(0.099)	(0.116)	(0.043)	(0.066)	(0.130)
Over 65	0.460***	0.405***	0.697***	0.614***	0.340***	0.290***	1.161***
	(0.046)	(0.059)	(0.079)	(0.094)	(0.030)	(0.043)	(0.098)
Over 70	0.448***	0.342***	0.622***	0.451***	0.302***	0.196***	1.022***
	(0.051)	(0.067)	(0.098)	(0.114)	(0.038)	(0.056)	(0.096)
Over 75	0.362***	0.258***	0.477***	0.323***	0.208***	0.102*	0.954***
	(0.053)	(0.066)	(0.105)	(0.112)	(0.048)	(0.062)	(0.075)
Observations	1,502		1,502		1,502		1,502

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. All specifications use an instrument lag of $x = 20$. Regressions are weighted by resident population aged 20+ and include ATS fixed effects. Robust standard errors are reported in parenthesis.

Table A5: Robustness: Excluding Milan ($N = 1{,}501$)

	Hospitalizations		Pharmaceutical		Outpatient		First Stage
	OLS	IV	OLS	IV	OLS	IV	
Over 60	0.406***	0.268***	0.582***	0.493***	0.310***	0.179***	1.122***
	(0.066)	(0.083)	(0.049)	(0.070)	(0.043)	(0.065)	(0.048)
Over 65	0.366***	0.279***	0.476***	0.378***	0.281***	0.201***	1.075***
	(0.052)	(0.060)	(0.039)	(0.046)	(0.034)	(0.042)	(0.041)
Over 70	0.348***	0.236***	0.366***	0.205***	0.218***	0.102**	0.968***
	(0.050)	(0.057)	(0.042)	(0.046)	(0.031)	(0.042)	(0.030)
Over 75	0.254***	0.142***	0.227***	0.100***	0.104***	0.003	0.951***
	(0.041)	(0.049)	(0.041)	(0.046)	(0.032)	(0.043)	(0.026)
Observations	1,501		1,501		1,501		1,501

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. The sample excludes the municipality of Milan (13.7% of Lombardy's population). Regressions are weighted by resident population aged 20+ and include ATS fixed effects. Robust standard errors are reported in parenthesis.

Table A6: Robustness: Excluding Province Capitals (N = 1,490)

	Hospitalizations		Pharmaceutical		Outpatient		First Stage
	OLS	IV	OLS	IV	OLS	IV	
Over 60	0.360***	0.203**	0.461***	0.365***	0.300***	0.130	0.995***
	(0.071)	(0.095)	(0.047)	(0.075)	(0.055)	(0.082)	(0.031)
Over 65	0.345***	0.247***	0.381***	0.281***	0.283***	0.179***	0.965***
	(0.057)	(0.070)	(0.036)	(0.045)	(0.042)	(0.052)	(0.024)
Over 70	0.318***	0.211***	0.257***	0.117***	0.199***	0.075	0.905***
	(0.053)	(0.064)	(0.031)	(0.040)	(0.037)	(0.048)	(0.019)
Over 75	0.219***	0.109**	0.126***	0.014	0.072*	-0.033	0.904***
	(0.043)	(0.053)	(0.031)	(0.041)	(0.037)	(0.048)	(0.017)
Observations	1,490		1,490		1,490		1,490

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. The sample excludes all 12 province capitals (22.7% of Lombardy's population). Regressions are weighted by resident population aged 20+ and include ATS fixed effects. Robust standard errors are reported in parenthesis.

Appendix B Testing Exclusion Restriction

Appendix B.1 Municipality Social Spending

A potential threat to the exclusion restriction is that predicted ageing could affect healthcare expenditure not only through the demand for medical care, but also through an alternative, non-health channel: municipal social-welfare spending targeted at the elderly population. If our instrument were capturing broader fiscal or policy trends correlated with population ageing, rather than a genuine health-driven channel, we would expect it to also significantly predict elderly-targeted social spending.

We construct our placebo outcome using municipal social expenditure data from ISTAT ("Indagine sulla spesa sociale dei comuni"), which provides spending on social services broken down by user category, including a specific category for services targeted at residents aged 65 and above. We define the placebo outcome as the log change in elderly-targeted social spending per capita between 2016 and 2019:

$$\ln\left(\frac{SS^{elderly}_{m,2019}}{N_{m,2019}}\right) - \ln\left(\frac{SS^{elderly}_{m,2016}}{N_{m,2016}}\right)$$

where $SS^{elderly}_{m,t}$ denotes total municipal social expenditure on services targeted at residents aged 65 and above in municipality m at time t, and $N_{m,t}$ denotes the municipality's population aged 20 and above, consistent with the denominator used throughout the paper (Equation 1). Social spending data are available only for 2016-2019; we therefore restrict this window to the three years for which both endpoints are observed.

For comparability, we re-estimate our baseline healthcare expenditure model (Equation 1) over the same 2016-2019 window, using the instrument constructed analogously to Equation 3 but rescaled to this shorter horizon. This allows us to directly compare, within the same period and sample, whether the instrument predicts healthcare expenditure but not elderly-targeted social spending.

Table A7 reports the results. The instrument continues to significantly predict growth in per-capita healthcare expenditure across all three domains and age thresholds, consistent with our main findings. By contrast, the IV coefficients on elderly-targeted social

spending (rightmost column) are not statistically significant at conventional levels for the Over 60, Over 65, Over 70 and Over 75 thresholds. Taken together, this pattern is broadly consistent with the instrument operating through the healthcare-specific ageing channel rather than through a systematic, confounding social-spending channel.

Table A7: Placebo Test: Elderly-Targeted Social Spending (2016-2019)

	Hospitalizations	Pharmaceutical	Outpatient	Social Spending
Over 60	0.355***	0.669***	0.378***	0.852
	(0.137)	(0.082)	(0.085)	(0.617)
Over 65	0.323***	0.531***	0.403***	0.735
	(0.098)	(0.063)	(0.060)	(0.458)
Over 70	0.236***	0.389***	0.305***	0.256
	(0.086)	(0.057)	(0.052)	(0.428)
Over 75	0.158*	0.253***	0.151**	-0.059
	(0.091)	(0.069)	(0.065)	(0.480)
Observations	1,502	1,502	1,502	1,459

Notes: *** p<0.01, ** p<0.05, * p<0.1. Columns 1-3 report elasticities of per-capita healthcare expenditure with respect to the elderly population share, re-estimated over a three-year growth window (2016-2019) for comparability with the placebo outcome (Column 4), which reports the elasticity of elderly-targeted municipal social spending per capita (resident population aged 20+) with respect to the same instrument. Columns 1-3 and Column 4 are estimated on slightly different samples (1,502 vs. 1,459 municipalities, reflecting missing social-spending data for some municipalities). First-stage F-statistics: Over 60 - 115.82 (Cols 1-3), 115.59 (Col 4); Over 65 - 144.77 (Cols 1-3), 144.64 (Col 4); Over 70 - 191.17 (Cols 1-3), 190.97 (Col 4); Over 75 - 137.62 (Cols 1-3), 137.41 (Col 4). All regressions are weighted by municipality population aged 20+ and include ATS fixed effects. Robust standard errors are reported in parentheses.

Appendix B.2 Municipality Health Protection Spending

As discussed in Section 4.1, healthcare expenditure in Italy is normatively determined at the regional level rather than at the municipal level: regional healthcare budgets depend on regional and national revenues and are not directly linked to the economic or demographic dynamics of individual municipalities. This institutional feature underpins the plausibility of our exclusion restriction, but it also generates a direct, testable prediction: if municipalities genuinely lack discretionary power over healthcare budgets, then predicted ageing should not significantly affect the health-related spending lines that municipalities do control.

We test this prediction using municipal spending on health protection, a discretionary budget category available for a subset of Lombardy municipalities. Municipal expenditure is classified according to the harmonized budget scheme introduced by Legislative Decree 118/2011, which organizes local government spending into "Missioni" and "Programmi." Mission 13 ("Tutela della salute") comprises seven programs: Programs 01-06 correspond to regional health service financing (ordinary and additional funding for LEA, deficit coverage, health investments, and NHS revenue restitution) and are a regional rather than municipal competence, typically showing negligible values in municipal budgets. Program 07 ("Ulteriori spese in materia sanitaria") finances further health-related expenditure not included in the other programs of the mission. According to the official program glossary, this program covers hygienic-sanitary interventions such as public kennels, public sanitation facilities, and similar structures, as well as environmental hygiene interventions such as pest control and disinfestation. It excludes advances for national health system financing from the state treasury, which are instead recorded as pass-through transactions ("partite di giro") under Mission 99 ("Servizi per conto terzi"). Because these functions are locally administered and structurally distinct from the demographic drivers of regional healthcare demand, Program 07 provides a suitable placebo outcome closely adjacent to, but conceptually independent of, the ageing-driven variation captured by our instrument.

Following the same specification as the placebo tests in Appendix C, we use a five-year growth window (2016-2021, $p = 5$) with a correspondingly shortened instrument lag

($x = 5$), and control for ATS fixed effects. Data on this spending category are available for only 119 municipalities, substantially fewer than in our main sample; results should therefore be interpreted with the corresponding loss of statistical power in mind.

Table A8 reports the results. None of the IV coefficients of Health-Related Spending is statistically significant at conventional levels, at any age threshold. This is consistent with the institutional argument in Section 4.1: predicted ageing does not translate into higher municipal discretionary health-related spending, supporting the claim that municipalities lack the budgetary levers to respond to demographic ageing in this domain, and reinforcing the credibility of the exclusion restriction.

Table A8: Placebo Test: Municipal Health-Related Spending, 2016-2021

	Hospitalization	Pharmaceutical	Outpatient	Health-Related Spending
Over 60	0.555**	0.383***	0.450***	-0.757
	(0.225)	(0.124)	(0.125)	(2.425)
Over 65	0.303*	0.311***	0.276***	0.741
	(0.168)	(0.071)	(0.091)	(1.719)
Over 70	0.210	0.241***	0.146*	0.334
	(0.132)	(0.048)	(0.065)	(1.187)
Over 75	0.327*	0.286***	0.033	-0.210
	(0.150)	(0.064)	(0.086)	(1.342)
Observations	119	119	119	119

Notes: *** p<0.01, ** p<0.05, * p<0.1. IV estimates of the elasticity of municipal discretionary health-related spending with respect to the elderly population share. Growth period 2016-2021 ($p = 5$), instrument lag $x = 5$. Controls: ATS fixed effects. Sample restricted to the 119 municipalities for which this spending category is available. Robust standard errors are reported in parentheses.

Appendix C Placebo test: Municipality Spending not related with health

As a further check on the identification strategy, we test whether the instrument spuriously predicts municipal expenditure categories unrelated to healthcare needs. If the instrument were capturing general fiscal capacity or unobserved local trends rather than a genuine ageing-related healthcare demand channel, it would also be expected to predict non-health expenditure. We therefore estimate our baseline specification using municipal spending on Roads (i.e. expenditure for road construction, maintenance, and related local transports) as placebo outcomes. Roads is a category with no plausible direct link to population ageing or healthcare demand. Due to the limited data availability for these expenditure categories, we use a shorter 5-year growth window (2016-2021, $p = 5$) and a correspondingly shorter instrument lag ($x = 5$), while maintaining ATS fixed effects. Table A9 reports the results. The estimated IV effects are statistically indistinguishable from zero across all age thresholds, providing additional evidence that the instrument does not capture broader local expenditure trends unrelated to healthcare ageing pressures.

Table A9: Placebo Test: Healthcare and Non-Healthcare Municipal Spending, 2016-2021

	Hospitalization	Pharmaceutical	Outpatient	Roads
Over 60	0.496***	0.495***	0.192***	1.946
	(0.100)	(0.054)	(0.065)	(1.527)
Over 65	0.428***	0.400***	0.189***	1.541
	(0.068)	(0.036)	(0.045)	(1.274)
Over 70	0.338***	0.260***	0.163***	1.205
	(0.056)	(0.034)	(0.034)	(1.055)
Over 75	0.186**	0.190***	0.043	0.982
	(0.083)	(0.053)	(0.050)	(1.073)
Observations	901	901	901	901

Notes: *** p<0.01, ** p<0.05, * p<0.1. IV estimates of the elasticity of healthcare expenditure and municipal spending on roads with respect to the elderly population share. Growth period 2016–2021 ($p = 5$), instrument lag $x = 5$. Controls: ATS fixed effects. All specifications are estimated on the same sample of 901 municipalities. Robust standard errors are reported in parentheses.

Appendix D Rotemberg-Weights decomposition

This Section reports the Rotemberg-weight diagnostics (Goldsmith-Pinkham et al., 2020) applied to a linearized version of the cohort-survival instrument, for the baseline specification (Hospitalizations, 60+ threshold). In particular, Figure A1 plots, for each active

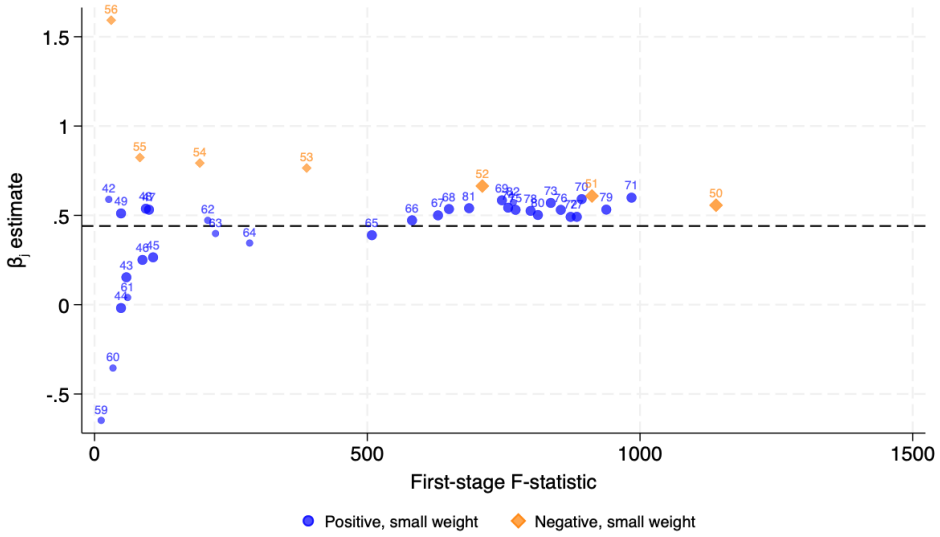


Figure A1: Heterogeneity of β_j: Ageing Instrument (Hospitalizations). Two cohorts with unstable β_j (near-zero first-stage F, negligible α_j weight) are excluded.

age cohort (those who were 42-82 years old in 2001), the cohort-specific just-identified estimate $\hat{\beta}_j$ against its corresponding first-stage F-statistic, distinguishing cohorts carrying positive and negative weights.

Table A10 quantifies the same decomposition: Panel a) reports the sum, mean, and share of positive and negative weights; Panel b) reports correlations among the weights, the regional survival shock, the cohort-specific estimates, the first-stage F-statistics, and the variance of the cohort share; Panel c) lists the five cohorts with the largest and smallest weights in absolute value; Panel d) reports the weighted average of $\hat{\beta}_j$ separately for positive and negative weight cohorts.

Figure A2 displays the same Rotemberg weights as a bar chart across the full age-cohort spectrum, making the underlying pattern graphically clear: positive weights for the youngest cohorts (up to age 50 in 2001) and for the older cohorts (60+), and negative weights for the intermediate range (ages 51-58).

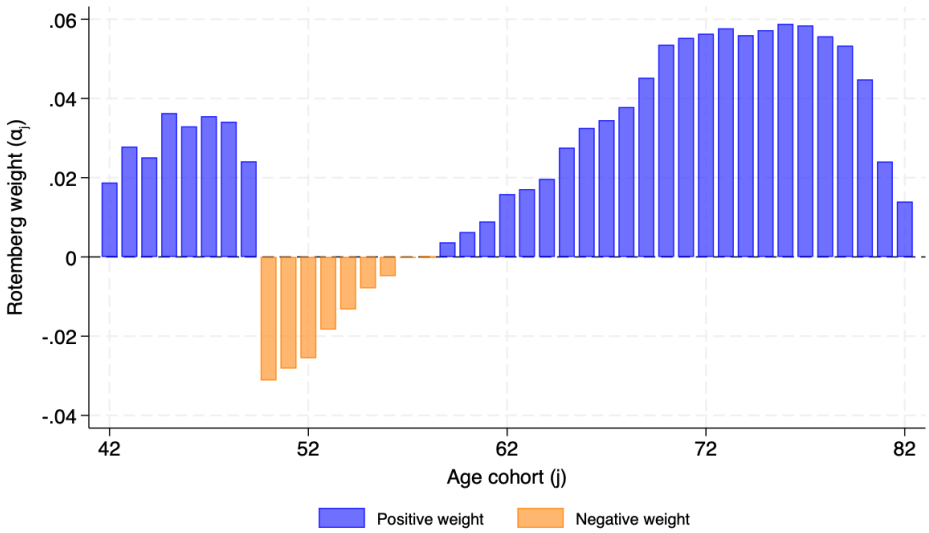


Figure A2: Rotemberg-weights by Age Cohort (as in 2001): Ageing Instrument (Hospitalizations)

Finally, Figure A3 plots each cohort's Rotemberg weight against its just-identified estimate $\hat{\beta}_j$, with a vertical line corresponding to the $\hat{\beta}$ of the aggregate IV estimate: the overall picture is that -as expected- identification is driven by a consistent cluster of older cohorts with similar estimates and strong first-stage F-statistics, rather than by any single outlying cohort.

Table A10: Rotemberg-weights: ageing instrument, Hospitalizations (60+)

Panel a): Negative and positive weights

	Sum	Mean	Share
Negative	−0.1296	−0.0144	0.1029
Positive	1.1296	0.0353	0.8971

Panel b): Correlations of cohort aggregates

	$\hat{\alpha}_j$	$\tilde{g}_j$	$\hat{\beta}_j$	F_j	$\text{Var}(z_j)$
$\hat{\alpha}_j$	1				
$\tilde{g}_j$	−0.091	1			
$\hat{\beta}_j$	−0.170	−0.055	1		
F_j	0.362	−0.606	−0.218	1	
$\text{Var}(z_j)$	0.407	0.844	−0.159	−0.266	1

Panel c): Top-5 Rotemberg-weight cohorts		*Bottom-5 Rotemberg-weight cohorts*	
Age	$\hat{\beta}_j$	Age	$\hat{\beta}_j$
73	0.5635	58	9.9691
76	0.5255	57	7.6343
75	0.5253	56	1.5885
77	0.4865	59	−0.8191
72	0.4847	60	−0.4510

Panel d): Estimates of $\hat{\beta}_j$ for positive and negative weights

	α-weighted Sum	Share of overall $\hat{\beta}$	Mean
Negative	−0.0960	−0.2178	0.8290
Positive	0.5365	1.2178	0.3929

Notes: The table reports statistics on the Rotemberg-weights (Goldsmith-Pinkham et al. 2020) applied to the linearized version of the ageing instrument described in Section 4.3 (Hospitalizations, 60+). Panel a) reports the sum, mean, and share of positive and negative weights; Share is normalized by the sum of absolute weights. Panel b) reports correlations between the weights ($\hat{\alpha}_j$), the combined regional survival shock ($\tilde{g}_j$), the just-identified cohort-specific estimates ($\hat{\beta}_j$), the cohort-specific first-stage F-statistic (F_j), and the cross-municipality variance of the cohort share ($\text{Var}(z_j)$), across all 41 active cohorts (ages 42-82). Restricting to the 39 cohorts with $|\hat{\beta}_j| \leq 5$ leaves the pattern broadly similar, with the exception of the correlation between F_j and $\hat{\beta}_j$, which changes sign (-0.218 to 0.270), reflecting the disproportionate influence of the two cohorts (ages 57-58) with near-zero first-stage F-statistics. Panel c) reports the five cohorts with the largest and smallest absolute Rotemberg-weights; the latter coincide with the cohorts carrying the least-identified, most unstable just-identified estimates. Panel d) reports how $\hat{\beta}_j$ varies with the sign of the Rotemberg-weight; the Mean for negative-weight cohorts excludes the two cohorts (ages 57-58) with near-zero first-stage F-statistics and correspondingly unstable $\hat{\beta}_j$ (unrestricted mean: 12.392).

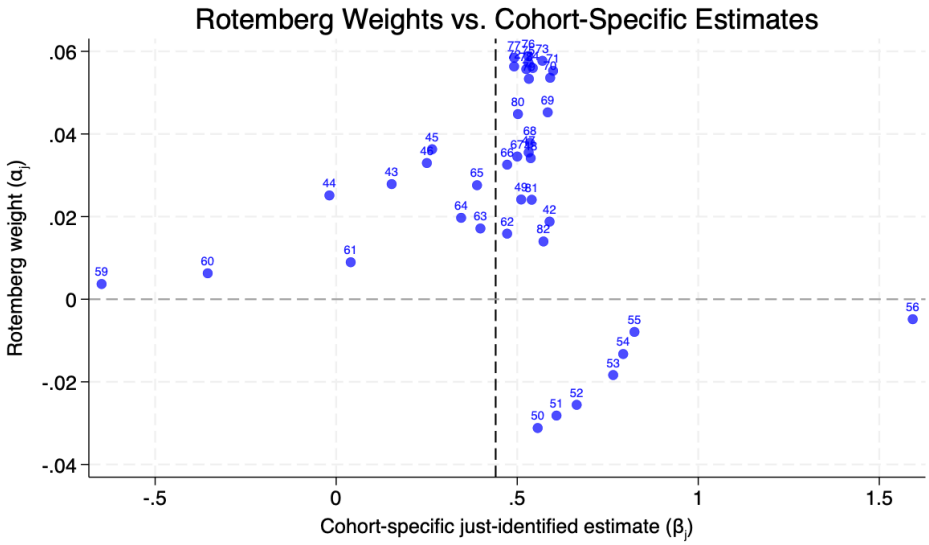


Figure A3: Rotemberg-weights vs. Cohort-Specific Estimates. Vertical line: aggregate IV estimate ($\hat{\beta}$=0.4405). Two cohorts with $\beta_j > 5$ excluded.

Appendix E Plausibly Exogenous Bounds

Appendix E.1 Construction of the Plausibly Exogenous Bounds

Following Conley et al. (2012), we assess the sensitivity of our IV estimates to potential violations of the exclusion restriction by allowing the instrument to have a direct effect on the outcome.

The structural relationship can then be written as:

$$Y = \beta X + \gamma Z + \varepsilon, \tag{7}$$

where γ captures the direct effect of the instrument on the outcome.

The standard IV assumption requires $\gamma = 0$; allowing $\gamma \neq 0$ provides a framework for evaluating the robustness of the estimated causal effect to plausible departures from the exclusion restriction.

The first-stage relationship is given by:

$$X = \pi_1 Z + u, \tag{8}$$

Substituting the first-stage equation (eq. 8) into the structural relationship (eq. 7) yields the reduced-form equation:

$$Y = \beta(\pi_1 Z + u) + \gamma Z + \epsilon \tag{9}$$

or equivalently,

$$Y = \beta\pi_1 Z + \beta u + \gamma Z + \epsilon \tag{10}$$

and

$$Y = (\beta\pi_1 + \gamma)Z + \beta u + \epsilon \tag{11}$$

if the reduced-form regression of Y on Z is given by $Y = \pi_2 Z + v$ then:

$$\pi_2 = \beta\pi_1 + \gamma \tag{12}$$

and

$$v = \beta u + \epsilon \tag{13}$$

Re-arranging gives:

$$\beta\pi_1 = \pi_2 - \gamma \tag{14}$$

and solving for the parameter of interest β yields:

$$\hat{\beta} = \frac{\hat{\pi_2}}{\hat{\pi_1}} - \frac{\gamma}{\hat{\pi_1}}. \tag{15}$$

Since $\frac{\hat{\pi}_2}{\hat{\pi}_1}$ is the standard just-identified IV estimator $\hat{\beta}_{iv}$, equation (15) can be rewritten as:

$$\hat{\beta(\gamma)} = \hat{\beta}_{iv} - \frac{\gamma}{\hat{\pi}_1} \tag{16}$$

From Eq. 16 the conventional IV estimator is recovered ($\hat{\beta}(\gamma) = \hat{\beta}_{iv}$) when $\gamma = 0$. If the instrument has a positive direct effect on healthcare expenditure (γ>0), the IV estimator overstates the causal effect, whereas a negative direct effect implies the opposite.

This expression can be directly computed using the estimates reported in the paper: $\hat{\beta}_{IV}$ is obtained from the baseline IV specification in Table 1, while $\hat{\pi}_1$ is the corresponding first-stage coefficient. We evaluate two calibrated values of γ, corresponding to the observed and a conservative assumption on the ratio of elderly to average per-capita healthcare expenditure, denoted γ_2 and $\gamma_{3.5}$, and for each value of γ we compute the corresponding adjusted estimate $\hat{\beta}(\gamma)$.

To calibrate γ using the elderly migration channel, note that Table 8 reports $\hat{\beta}^{net}_{mig}$, the causal (2SLS) effect of realized ageing (X, instrumented by Z) on net elderly in-migration. The implied direct effect of the instrument on expenditure via this channel

(i.e. γ in Eq. (7), per unit of Z) is therefore:

$$\gamma_{mig} = \hat{\pi}_1 \cdot \hat{\beta}^{net}_{mig} \cdot \frac{\bar{e}_{65+} - \bar{e}_{tot}}{\bar{e}_{tot}} \tag{17}$$

since the reduced-form effect of Z on migration equals $\hat{\pi}_1 \cdot \hat{\beta}^{net}_{mig}$ by construction of the first stage (eq. 8). Substituting (17) into the bias term of Eq. (16) gives:

$$\frac{\gamma_{mig}}{\hat{\pi}_1} = \hat{\beta}^{net}_{mig} \cdot \frac{\bar{e}_{65+} - \bar{e}_{tot}}{\bar{e}_{tot}} \tag{18}$$

where the $\hat{\pi}_1$ terms cancel algebraically. We therefore define the implied bias directly as

$$\tilde{\gamma}_{mig} = \hat{\beta}^{net}_{mig} \cdot \frac{\bar{e}_{65+} - \bar{e}_{tot}}{\bar{e}_{tot}}, \tag{19}$$

which can be computed without separately estimating γ_{mig} or $\hat{\pi}_1$. Here $\hat{\beta}^{net}_{mig}$ is the causal (2SLS) effect of ageing on net elderly (65+) in-migration over 2011-2019, expressed per capita of the 2011 population aged 20+ (Table 8), and $\bar{e}_{65+}$ and $\bar{e}_{tot}$ denote average 2011 per-capita healthcare expenditure for elderly and total residents, respectively.

We consider two values for $\bar{e}_{65+}$ corresponding to alternative assumptions about the relative healthcare expenditure of elderly residents. The first assumes the observed expenditure ratio of approximately 2.0, while the second considers a more conservative ratio of 3.5, allowing for the possibility that elderly migrants are negatively selected on health status relative to the average elderly resident. We denote the resulting calibrated values $\tilde{\gamma}_{mig,2}$ and $\tilde{\gamma}_{mig,3.5}$, respectively.

For each calibrated value of $\tilde{\gamma}_{mig}$, the corresponding identified set is:

$$\mathcal{B}(\tilde{\gamma}_{mig}) = \left[\hat{\beta}_{IV} - |\tilde{\gamma}_{mig}|\,,\ \hat{\beta}_{IV} + |\tilde{\gamma}_{mig}|\right]. \tag{20}$$

Because the direct effect of elderly migration on expenditure is expected to be positive rather than symmetric, an even more stringent alternative would report a one-sided

bound, rather than the symmetric interval.

$$\mathcal{B}(\tilde{\gamma}_{mig}) = \left[\hat{\beta}_{IV}; \hat{\beta}_{IV} + |\tilde{\gamma}_{mig}|\right]. \tag{21}$$

However, we rely on the symmetric (two-sided) bound as the more conservative and standard choice following Conley et al. (2012). These bounds represent the range of causal effects that remain compatible with the data under the assumed magnitude of the violation of the exclusion restriction.

Appendix E.2 Confidence Intervals for the Bounds

The exact variance of the bias-adjusted estimator $\hat{\beta}(\gamma)$, and consequently of the endpoints of the identified set $\mathcal{B}(\gamma_{mig})$, depends on the covariance between the first-stage and reduced-form estimates, or equivalently the joint variance-covariance matrix of $\hat{\beta}_{IV}$ and $\hat{\pi}_1$. Following Conley et al. (2012), when the first stage is estimated with high precision, the additional uncertainty associated with $\hat{\pi}_1$ is negligible relative to the sampling uncertainty of the IV estimator. Therefore, we approximate:

$$SE(\hat{\beta}(\gamma)) \approx SE(\hat{\beta}_{IV}). \tag{22}$$

We further treat γ_{mig} as a fixed calibrated sensitivity parameter rather than an estimated quantity with sampling uncertainty. Therefore, the reported confidence intervals should be interpreted as conditional on the calibrated value of γ_{mig}.

Hence, the confidence interval for each endpoint of the identified set is obtained by expanding the bounds by the standard error of the IV estimator:

$$LB = \hat{\beta}_{IV} - |\tilde{\gamma}_{mig}| - 1.96 \times SE(\hat{\beta}_{IV}), \tag{23}$$

$$UB = \hat{\beta}_{IV} + |\tilde{\gamma}_{mig}| + 1.96 \times SE(\hat{\beta}_{IV}). \tag{24}$$

For completeness, we also compute the exact Imbens-Manski critical values C_α (Im-

bens and Manski 2004). Across all twelve specifications considered, these values are below 1.96 (ranging from 1.68 to 1.95), confirming that the fixed-critical-value approximation used in the main analysis is conservative.

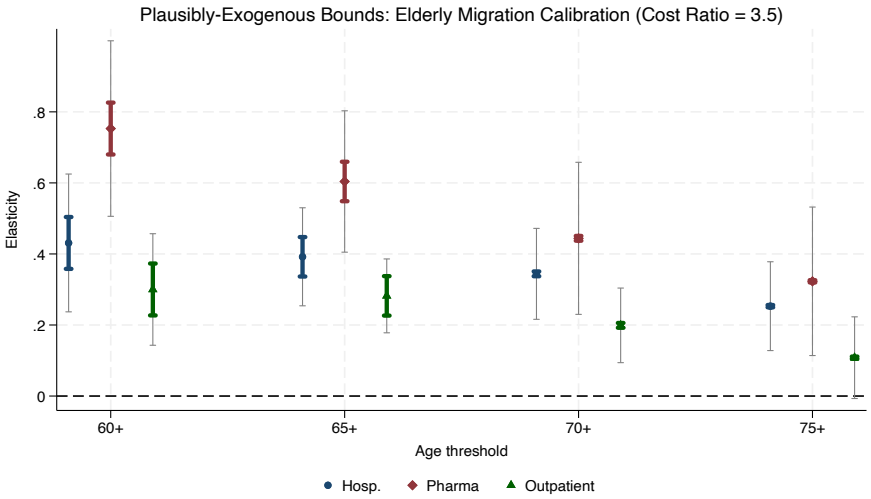


Figure A4: CHR bounds with 95% CI, assuming $\frac{e_{tot}}{e_{65+}} = 3.5$

Appendix F Spatial Correlation

A potential concern in our municipality-level setting is that healthcare expenditure growth may exhibit spatial dependence. Municipalities located in close geographic proximity may share common local shocks, environmental conditions, healthcare infrastructures, labour market conditions or demographic dynamics, leading to correlated regression residuals across neighboring municipalities. Ignoring such spatial correlation could result in underestimated standard errors.

To address this concern, we re-estimate our baseline OLS and IV specifications computing spatially robust standard errors following Conley (1999), allowing the covariance structure of the errors to decay with geographical distance. Table A11 reports the resulting estimates across the three healthcare domains and the four age thresholds considered in the main analysis. The estimated elasticities remain virtually unchanged relative to the baseline specification, indicating that accounting for spatial dependence has a negligible effect on the magnitude of the estimated coefficients. Statistical significance is also largely unaffected, with the only exception being the IV estimate for outpatient expenditure among municipalities with a population share aged over 75, which becomes imprecisely estimated. Overall, these results indicate that our main findings are robust to allowing for spatial correlation in healthcare expenditure growth across municipalities.

Table A11: OLS and IV Elasticities with Spatially Robust Standard Errors

	Hospitalizations		Pharmaceutical		Outpatient	
	OLS	IV	OLS	IV	OLS	IV
Over 60	0.544***	0.431***	0.862***	0.753***	0.401***	0.300***
	(0.075)	(0.108)	(0.097)	(0.096)	(0.038)	(0.068)
Over 65	0.460***	0.392***	0.697***	0.604***	0.340***	0.282***
	(0.056)	(0.076)	(0.077)	(0.086)	(0.031)	(0.048)
Over 70	0.448***	0.344***	0.622***	0.444***	0.302***	0.199***
	(0.057)	(0.082)	(0.085)	(0.111)	(0.041)	(0.066)
Over 75	0.362***	0.253***	0.477***	0.323***	0.208***	0.108
	(0.063)	(0.088)	(0.104)	(0.124)	(0.054)	(0.073)

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. *Notes:* $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. All regressions are weighted by the resident population aged 20+ and include ATS fixed effects. The baseline specification uses an eight-year growth window (2011-2019, $p = 8$) and an instrument constructed with a 10-year lag ($x = 10$). Standard errors reported in parentheses account for spatial correlation across municipalities using Conley (1999) spatial HAC corrections with a distance cutoff of 20 km.

Appendix G Disease Specific Effects of Population Ageing on Healthcare Expenditure

This section reports disease-specific estimates of the effect of population ageing on healthcare expenditure. We replicate the baseline analysis separately for oncological (see Table A12) and cardiovascular patients (see Table A13) to assess the extent to which the estimated elasticities vary across major disease groups.

Table A12: OLS and IV Elasticities by Care Domain- Oncological Patients

	Hospitalizations		Pharmaceutical		Outpatient		First Stage
	OLS	IV	OLS	IV	OLS	IV	
Over 60	0.610***	0.465***	1.241***	1.088***	0.423	0.552	1.393***
	(0.114)	(0.155)	(0.319)	(0.351)	(0.832)	(0.952)	(0.133)
Over 65	0.538***	0.461***	1.163***	1.164***	0.557	0.779	1.301***
	(0.079)	(0.097)	(0.301)	(0.317)	(0.315)	(0.624)	(0.107)
Over 70	0.534***	0.416***	1.185***	1.185***	0.857**	1.136**	1.162***
	(0.069)	(0.084)	(0.329)	(0.368)	(0.426)	(0.468)	(0.102)
Over 75	0.444***	0.342***	0.910***	0.754***	0.657**	0.741***	1.100***
	(0.066)	(0.080)	(0.234)	(0.241)	(0.291)	(0.298)	(0.080)
Observations	1,481		1,481		1,481		1,481

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. All regressions are weighted by the resident population aged 20+ and include ATS fixed effects. The baseline specification uses an eight-year growth window (2011-2019, $p = 8$) and an instrument constructed with a 10-year lag ($x = 10$). K-P Wald F Statistics: 110.35 for Over60, 148.18 for Over65 , 128.07 for Over70 and 188.52 for Over75. Robust standard errors are reported in parentheses.

Table A13: OLS and IV Elasticities by Care Domain- Cardiovascular Patients

	Hospitalizations		Pharmaceutical		Outpatient		First Stage
	OLS	IV	OLS	IV	OLS	IV	
Over 60	0.651***	0.508***	1.043***	0.549	0.402**	0.366	1.393***
	(0.111)	(0.146)	(0.379)	(0.542)	(0.202)	(0.235)	(0.133)
Over 65	0.544***	0.409***	0.831***	0.600	0.386**	0.309	1.301***
	(0.081)	(0.084)	(0.287)	(0.383)	(0.157)	(0.190)	(0.107)
Over 70	0.532***	0.416***	0.768***	0.358	0.344**	0.273	1.162***
	(0.074)	(0.084)	(0.291)	(0.406)	(0.158)	(0.184)	(0.102)
Over 75	0.451***	0.350***	0.664***	0.337	0.365**	0.347***	1.100***
	(0.070)	(0.076)	(0.240)	(0.305)	(0.155)	(0.173)	(0.080)
Observations	1,482		1,482		1,482		1,482

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. All regressions are weighted by the resident population aged 20+ and include ATS fixed effects. The baseline specification uses an eight-year growth window (2011-2019, $p = 8$) and an instrument constructed with a 10-year lag ($x = 10$). K-P Wald F Statistics: 110.45 for Over60, 148.17 for Over65 , 128.14 for Over70 and 188.52 for Over75. Robust standard errors are reported in parentheses.