



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/244254/>

Version: Accepted Version

Article:

Kim, Jinhyun (2026) How Much Does It Cost the South Korean Public Healthcare System to Produce One Healthy Life Year? Health Economics. ISSN: 1057-9230

<https://doi.org/10.1002/hec.70134>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

RESEARCH ARTICLE

How Much Does It Cost the South Korean Public Healthcare System to Produce One Healthy Life Year?

Jinhyun Kim

¹Department of Economics and Related Studies, University of York, York, UK

Correspondence

Email: jinhyun.kim@york.ac.uk

Present address

Abstract

Investigating the impact of healthcare expenditure on health is crucial for optimising public resource allocation. This can inform the value of expanding healthcare budgets and provide a benchmark for allocations within healthcare systems. However, empirical evidence remains mixed. The potential causes of mixed results—reverse causality and unobservable confounders—need to be addressed through robust empirical methods. This study assesses the South Korean public healthcare cost of producing one quality-adjusted life year (QALY) by using conversion factors—elements in the pricing formula of healthcare services determined by negotiations—as the instrument, with two-way fixed effects. Moreover, since morbidity significantly affects health, it is crucial to estimate the impact of healthcare expenditure on morbidity using credible data. We use a unique dataset of directly measured regional EQ-5D indices from 240 municipalities for the period 2008–2019, which is unavailable in other countries. We estimate that \$12,847 of public healthcare expenditure is required to gain one QALY. Sensitivity analyses using alternative calculation methods, data, and models produce estimates ranging from \$8,850 to \$17,674 per QALY. Our findings suggest that the healthcare cost of producing one QALY is substantially lower than the GDP per capita and the willingness-to-pay for health in South Korea.

KEYWORDS

Healthcare Productivity, Healthcare Expenditure, Quality-Adjusted Life Year, Instrumental Variable, Panel Data

1 | INTRODUCTION

Healthcare expenditure represents a considerable proportion of economic activity in high-income countries. The OECD countries' average ratio of total health expenditure to GDP is projected to reach 11.2% in 2040, rising from 8.8% in 2018 (OECD, 2024). Among these countries, between 2000 and 2018, South Korea had the fastest-growing health expenditure per capita, with an annual growth rate of 7.5%. South Korea also exhibits the highest projected growth rate between 2019 and

2040 (3.9% annually) (OECD, 2024). This is attributed to its rapidly ageing population alongside it having the lowest fertility rate in the world (Kirkegaard, 2021).

As the financial burden of healthcare expenditure in high-income countries threatens fiscal sustainability, effective and efficient spending on healthcare has become increasingly significant for policymakers. This has led many researchers to investigate the marginal impact of healthcare expenditure on health to inform resource allocation in public spending and healthcare. Estimating the effect of healthcare expenditure on health can inform the value of expanding or contracting healthcare budgets and provide a benchmark for the allocation of existing healthcare system finances.

However, the extent to which health outcomes are causally attributable to healthcare remains a subject of much debate. Although numerous medical technologies have demonstrated their effectiveness through randomised controlled studies, empirical investigations into the impact of healthcare expenditure have produced mixed results (Baltagi, Moscone, & Tosetti, 2012; Claxton, Lomas, & Martin, 2018; Filmer &

Abbreviations: CET, Cost-Effectiveness Threshold; DALY, Disability-Adjusted Life Year; FEIV, Fixed Effects Instrumental Variable; GDP, Gross Domestic Production; OECD, The Organisation for Economic Co-operation and Development; HRQoL, Health-Related Quality of Life; IV, Instrumental Variable; KCHS, Korea Community Health Survey; KNHIS, Korean National Health Insurance Service; KOSIS, Korean Statistics Information Service; KRW, Korean Won; LATE, Local Average Treatment Effect; QALY, Quality-Adjusted Life Year; RVU, Relative Value Unit; USD, United States Dollar.

Pritchett, 1999; Thornton, 2011). The estimated effects of healthcare expenditure range from significant health improvements to substantial detrimental impacts. These heterogeneous results could stem from two empirical challenges that researchers have to address: reverse causality and unobservable confounders (Laura C Edney et al., 2022). Poor health leads to high healthcare expenditure, which inhibits the identification of the health-improving effect of healthcare expenditure. Moreover, many unobservable factors could simultaneously affect both healthcare expenditure and health outcomes, thereby creating significant biases in treatment effect estimates. Another important issue is data availability, because obtaining precise and consistent mortality and morbidity data may be challenging.

This paper aims to assess the causal effect of healthcare expenditure on health outcomes using instrumental variable (IV) estimation. The IV estimation isolates the portion of the endogenous regressor that is uncontaminated by unobservable confounders. This study employs seven provider-specific conversion factors from the South Korean public healthcare system as an IV. Conversion factors are a component of the pricing formula of healthcare services, which significantly affects healthcare expenditure. The validity of conversion factors as an instrument rests on the process through which they are determined, complex annual negotiations between the government (a monopsonistic buyer) and seven government-approved healthcare provider associations (monopolistic sellers), a structure known as a bilateral monopoly. The government is legally mandated to negotiate with these seven associations, with each association representing a specific type of provider (clinics, hospitals, dentists, Korean Medicine, midwives, pharmacies, and public health centres). In this setting, negotiation outcomes depend on the relative bargaining power and strategic interactions of the parties rather than socioeconomic and health-related factors, making negotiation outcomes inherently difficult to predict (Gravelle & Rees, 2004) and plausibly exogenous to underlying population health trends and other unobservable confounders.

One of the first studies to explore the causal impact of healthcare expenditure on health, Filmer and Pritchett (1999) assessed the effect of the share of health spending in GDP on under-5-year mortality in a cross-national study of 100 countries. They found that the health impact of healthcare expenditure was not statistically significant. In contrast, more recent cross-national studies have found statistically significant mortality-improving effects of healthcare expenditure (Baltagi et al., 2012; Berger & Messer, 2002; Bokhari, Gai, & Gottret, 2007). However, cross-national studies might not assess the effect of healthcare expenditure precisely due to a lack of comparability; heterogeneous population characteristics and healthcare systems could undermine analytical accuracy, and inconsistent measurement methods could lead to a reliability issue (Farahani, Subramanian, & Canning, 2010).

Within-country studies could capture the causal impact of healthcare expenditure on health more accurately than cross-national studies. However, the results are mixed. Crémieux, Ouellette, and Pilon (1999) estimated the effect of healthcare spending on life expectancy using Canadian regional panel data and found a beneficial effect. More recent

studies in England and India have also found a significant protective effect of healthcare expenditure on mortality (Claxton et al., 2018; Farahani et al., 2010; Martin, Rice, & Smith, 2008). Conversely, a United States state-level panel study found that healthcare expenditure had no beneficial effect on mortality (Thornton, 2011).

A major limitation of the studies discussed above is that the effect of healthcare expenditure on morbidity has not been empirically assessed. This issue is critical as many non-fatal diseases incur substantial healthcare expenditure. Recently, a growing body of literature has estimated the effect of healthcare expenditure on quality-adjusted life years (QALY) or disability-adjusted life years (DALY) using regional mortality and morbidity data (Laura Catherine Edney, Haji Ali Afzali, Cheng, & Karnon, 2018; Martin, Lomas, Claxton, & Longo, 2021; Moler-Zapata et al., 2022; Siverskog & Henriksson, 2019; Vallejo-Torres, García-Lorenzo, & Serrano-Aguilar, 2018). These studies have found significant health-improving effects of healthcare expenditure. However, many of them have used incomplete or approximated proxy morbidity data rather than directly measured regional morbidity data.

Few studies have utilised complete morbidity data. D. M. Cutler et al. (2022) used individual mortality and health-related quality of life data to examine the effect of Medicare expenditure on health outcomes and found a beneficial impact on quality-adjusted life expectancy. Weaver et al. (2022) and Park, Dieleman, Weaver, Bae, and Eggleston (2025) obtained morbidity data from the Global Burden of Disease Study. They found a protective health impact of healthcare expenditure and computed the cost to improve one DALY. However, these three studies rely on a selection on observables approach, specifically motivated by understanding differences in health outcomes that can be attributed to observable factors, where interpretation of the resulting estimates as causally attributable could be compromised by unobservable confounders (D. M. Cutler et al., 2022; Park et al., 2025; Weaver et al., 2022).

This study assesses the marginal public healthcare cost required to produce one QALY using IV estimation with two-way fixed effects. IV estimation and two-way fixed effects are two common empirical strategies to deal with unobservable confounders and reverse causality (Laura C Edney et al., 2022). A valid IV must not affect the outcome variable except through its impact on the treatment variable, which requires a strong assumption. However, the security of this assumption is improved in the context of a two-way fixed effects model, which removes time-invariant individual or regional characteristics as well as purely time-dependent characteristics from the individuals or regions. That is, the IV could be unrelated to the unobservable time-invariant historical factors (Wooldridge, 2010) and solely time-dependent factors when using the two-way fixed effects. In addition, South Korean regional panel data provide precise measures of morbidity, enabling the estimation of the effect of healthcare expenditure on morbidity. The Korean Community Health Survey (KCHS) provides annually measured regional health-related quality of life (HRQoL) data for 250 municipalities from 2008 to 2019, data that are not available in other countries.

The present study is the first attempt to investigate the South Korean healthcare cost for one QALY improvement using IV estimation. To our knowledge, this study is also the first to use South Korean conversion factors as the instrument for healthcare expenditure. By using rich South Korean morbidity data, this approach contributes robust and credible evidence to the existing literature on the causal impact of healthcare expenditure on health. Moreover, investigating the cost of producing one QALY through spending on public healthcare is crucial to inform healthcare resource allocation decisions for South Korea, which is one of the fastest-ageing countries in the world and has the highest healthcare expenditure growth rate among OECD countries. An estimate of the cost per QALY could help determine both the overall level of healthcare expenditure and the optimal allocation of the healthcare budget.

2 | INSTITUTIONAL BACKGROUND

South Korea features a compulsory publicly-funded single-payer health insurance, the Korean National Health Insurance Service (KNHIS), which covers the entire population except for the medical aid users. Approximately 3% of the population with earnings below 40% of the national median income is covered by the medical aid scheme financed by taxes. The KNHIS and medical aid users share the same healthcare providers and services (WHO, 2015). The services reimbursed by the KNHIS, the medical aid, and legal copayment, referred to as insured health benefits, have accounted for more than 80% of the total healthcare expenditure over the last two decades. Legal copayments are set at 20% to 30% of the total public healthcare cost, with fixed rates varying by patient age and type of providers, and annual individual caps.

The present study uses public healthcare expenditure, which encompasses expenditures from the KNHIS, legal copayment, and medical aid program, on insured health benefits as the treatment variable; this focuses on the public healthcare system, a critical element for fiscal sustainability. Non-insured health benefits are determined not to be reimbursed by the KNHIS because the national economic evaluation process identifies them as not being cost-effective. This reimbursement decision-making is conducted by an expert committee and subsequently approved by the Health Insurance Policy Deliberation Committee, made up of representatives of payers, providers, and public agencies (Kwon, 2019). Therefore, non-insured health benefits are financed by out-of-pocket expenditure and private health insurance. According to the statistics from private health insurance companies, the top 5 non-insured health benefits reimbursed by private health insurance were not cost-effective or even without proven effectiveness, such as manual therapy for musculoskeletal diseases, intravenous fluid therapy, and vertebral surgery for non-severe cases (G.-S. Kim, 2023). South Korean regulations prohibit healthcare providers from offering KNHIS-covered services as non-insured benefits (i.e., charging patients the full price). This ensures a clear demarcation, preventing any overlap between the

services classified as insured and non-insured health benefits. Consequently, the impact of non-insured health benefits expenditure would not be significant in this study.

Annually measured, municipality-level panel data on healthcare expenditures (the sum of KNHIS expenditure, legal copayment, and medical aid expenditure) and health outcomes are constructed. In this research, municipalities are defined as the lowest-level administrative units that have a basic local government. South Korea consists of nine provinces ("Do") and seven metropolitan cities ("Teukbyeol-si" and "Gwangyeok-si"). Provinces comprise urban cities ("Si") and rural cities ("Gun"), and metropolitan cities are composed of boroughs ("Jachi-gu"). The number of municipalities is approximately 250, comprising boroughs, urban cities, and rural cities.

3 | EMPIRICAL STRATEGY

3.1 | Empirical Model

A regional health panel model can be written as:

$$H_{jt} = \beta \text{Exp}_{jt} + \gamma X_{jt} + \tau_t + \alpha_j + \varepsilon_{jt} \quad (1)$$

where H_{jt} denotes regional health outcomes, β the effect of regional public healthcare expenditure, Exp_{jt} regional public healthcare expenditure, X_{jt} a vector of time-varying regional attributes, τ_t year fixed effects, α_j region fixed effects, and ε_{jt} the idiosyncratic error. The subscript j indicates municipality, and t the year.

Two empirical issues could impede the identification of the treatment effect β . First, healthcare expenditure can be affected by historical health outcomes (Claxton et al., 2018). A region with poor health outcomes in the previous year tends to generate high healthcare expenditure in the current year. This reverse causality hampers the identification of the treatment effect, and the South Korean health panel data are also likely to suffer from this issue. Second, unobservable confounders could create biases. Regional socioeconomic and demographic characteristics can affect both healthcare expenditure and health outcomes. Individuals with low wealth, low income, and limited education tend to be less healthy than others (D. Cutler, Deaton, & Lleras-Muney, 2006). In addition, socioeconomically disadvantaged individuals tend to consume more publicly funded healthcare (Cookson, Propper, Asaria, & Raine, 2016). In the South Korean context, expenditure on non-insured health benefits could be a confounder. That is, a large portion of the non-insured health benefits is reimbursed by private health insurance, and having private health insurance tends to increase the public healthcare expenditure because the copayment of the public healthcare is also reimbursed by private health insurance (M. J. Kim, 2023). Although including these confounding factors as covariates might help remove biases, data on many socioeconomic confounders and on spending for non-insured health benefits are not available. Moreover, there could be more unobservable confounders that have not yet been explored, considering the complexity of health determinants.

To obtain an unbiased estimate, instrumental variable (IV) estimation is exploited. This method utilises the exogenous variation of the endogenous regressor affected by the IV, which must not be related to the outcome—either directly or indirectly—except through the treatment variable (Martin et al., 2021). Therefore, the IV employed in this study must affect South Korean public healthcare expenditure (the relevance assumption). In addition, the IV must not have a separate direct effect on regional mortality and morbidity. The IV affects mortality and morbidity only through its impact on healthcare expenditures (the exclusion restriction assumption). Furthermore, the treatment variable needs to increase monotonically as the IV increases (the monotonicity assumption). Meeting these assumptions is challenging; thus, rigorous diagnostic procedures are conducted to justify the validity of the IV in the present study. Be aware that the IV estimate is the local average treatment effect rather than the average treatment effect, because the IV estimation relies solely on the exogenous part of the regressor.

In addition, a two-way fixed effects model is used to ensure the validity of the IV. This model controls for τ_t , the year fixed effects and α_j , the region fixed effects. The year fixed effects capture all factors that depend solely on the year, such as national economic growth, inflation and KNHIS policy reforms. The region fixed effects account for time-invariant regional attributes, such as regional socioeconomic inequalities. The elimination of both year and region fixed effects could help ensure that the IV is not correlated with year-specific and region-specific factors, such as economic growth, KNHIS policy reforms, and regional socioeconomic inequalities.

3.2 | Using a Component of the Pricing Formula as the Instrumental Variable

The IV in this study is defined as the weighted index of the conversion factors. Conversion factors denote the prices paid per relative value unit (RVU). KNHIS employs RVUs to set a fee schedule for each healthcare intervention. Conversion factors are selected as the IV because they are determined through annual negotiations between the government and healthcare providers in a bilateral monopoly setting, thereby having an unpredictable nature (see Figure 1 and Table A2) and plausibly being unrelated to health outcomes except through their effect on healthcare expenditure (the exclusion restriction assumption). Conversion factors are strong predictors of healthcare expenditure, since healthcare service prices equal the product of RVUs and conversion factors (the relevance assumption). This direct mechanical link also ensures that regional healthcare expenditure increases monotonically as the IV increases (the monotonicity assumption) (see Figure A1). It should be noted that conversion factors are determined approximately six months before they are implemented, ensuring they are not affected by the contemporaneous healthcare expenditure and health outcomes of the application year.

RVUs assign numerical values to interventions based on workload, risks, and resources required to deliver them. There are approximately 9,000 RVUs, which reflect three core components: work (i.e., providers'

expertise, time, and effort); practice expense (i.e., overhead costs such as equipment, supplies, and staffing); and risks (covering the cost associated with potential legal risks). This creates a common "unit" for comparing and pricing diverse interventions. The RVU schedule was initially set in 2001 and subsequently updated in 2008, 2017, and 2024.

To calculate the final reimbursement rate for an intervention, each RVU is multiplied by a conversion factor (KRW per RVU). Seven distinct conversion factors correspond to the following provider categories: clinics, hospitals, dentistry, Korean Medicine, midwives, pharmacies, and public health centres. Representatives of each provider category negotiate their respective conversion factors annually with the government.

The government aims to adjust conversion factors annually to account for inflation, economic growth, and demographic changes. However, negotiations often fail because healthcare providers tend to demand higher conversion factor growth rates while the government seeks to contain overall healthcare expenditure. If a negotiation fails, the conversion factor for that type of healthcare provider is set by the Health Insurance Policy Deliberation Committee, consisting of representatives of payers, providers, and public agencies (Kwon, 2019). As a result of the conflict of interest, conversion factor growth rates are highly unpredictable (Figure 1). Despite the unpredictability, national trends such as inflation and economic growth could affect conversion factors due to the government's intention. This could result in the violation of the assumption of exclusion restriction. Including year fixed effects would address this concern, but raw conversion factors lack regional variation and are therefore perfectly collinear with year fixed effects.

To employ year-fixed effects, a shift-share instrument is employed. The shift-share instrument, commonly used in labour, trade, and macroeconomics studies, is defined as the sum of a common set of shifts weighted by heterogeneous exposure shares (Borusyak, Hull, & Jaravel, 2025). In this study, the national conversion factor shock is the common set of exogenous shifts, and shares of the providers in regional healthcare expenditures are various exposure shares. Although conversion factors affect healthcare expenditure at the national level, they have different impacts on regional healthcare expenditures depending on municipalities' pre-existing conditions. That is, the regional impact of a provider-specific conversion factor is determined by the regional healthcare volume share of the specific provider, which varies across regions.

To avoid contemporaneous endogeneity, constructed shares are used rather than observed regional healthcare volumes. National mean healthcare visit counts by age-sex group are derived by averaging 12 years of national statistics (2008–2019), thereby excluding regional and temporal variation. These visit numbers are multiplied by the proportion of each age-sex group in the regional population to compute exposure shares. The constructed shares thus capture expected healthcare needs based on demographic composition alone, independent of regional supply-side factors. The constructed shares exhibit sufficient regional variation, as different age-sex groups display distinct healthcare utilisation patterns across provider types. For example, municipalities with a

higher proportion of elderly residents experience greater expenditure sensitivity to hospital and Korean Medicine conversion factors, given that older individuals more frequently use these services. Accordingly, municipality- and year-specific age-sex structures and national health-care utilisation rates of age-sex groups by healthcare providers are incorporated into the IV. The IV structure is:

$$IV_{jt} = CF_t \times VN \times ASG_{jt} \quad (2)$$

The structure of the IV at municipality j in year t is defined as a product of three matrices. Let CF_t be a 1×7 vector of conversion factors in year t , corresponding to the seven provider categories: clinics, hospitals, dentistry, Korean Medicine, midwives, pharmacies, and public health centres. Let VN be a 7×38 matrix of national average visit numbers for each of 38 age-sex groups (in five-year age intervals) across provider types. These visit numbers are the sum of outpatient visits and admission days. Finally, let ASG_{jt} be a 38×1 vector representing the proportion of each age-sex group in municipality j in year t . In the shift-share IV scheme, the sum of shares needs to be included as a control if shares do not sum to one, as this could be a source of confounding (Borusyak et al., 2025). For this reason, and to control for demographic confounding effects, the age-sex variables used in constructing the IV are included as covariates.

3.3 | FEIV Estimator and IV Diagnostics

The weighted index of conversion factors could identify the causal effect of healthcare expenditure via the fixed effects IV (FEIV) estimator. Region fixed effects can remove time-invariant regional characteristics from the IV. Consequently, potential time-invariant regional confounding effects from incorporating regional demographic differences into the IV could be removed. Year fixed effects can remove national trends, such as economic growth and the KNHIS policy reforms, from the IV. Including age-sex structures as covariates can control the time-varying confounding effects of the age-sex structures included in the IV. Then, the IV can be truly exogenous, and the local average treatment effect (LATE) of healthcare expenditure on health outcomes can be estimated without being affected by confounders and reverse causality.

In the base model, only regional age-sex structures are included as covariates. Therefore, health outcome and spending variables are not standardised. Controlling for age-sex structures is essential for the validity of the IV because the IV contains them. Although socioeconomic variables (e.g., regional income and education) and behavioural factors (e.g., smoking and obesity rates) were collected, they are reserved for sensitivity analyses. Including various covariates in IV models requires caution due to the risk of collider bias and post-instrumental bias (Glynn, Rueda, & Schuessler, 2023). Regional age-sex structures are likely to affect other covariates. Thus, including both age-sex structures and other covariates could produce a collider bias, creating a false relationship between the treatment variable and the outcome variable. Conversion factors are theoretically orthogonal to regional characteristics because

negotiators do not consider them. National confounders, such as economic growth and inflation, could be dealt with by using year fixed effects.

Two-stage least squares (2SLS) is exploited for the FEIV estimator. The first-stage regression equation is specified as:

$$Exp_{jt} = \theta Z_{jt} + \delta X_{jt} + \tau_t + \alpha_j + \eta_{jt} \quad (3)$$

where Exp_{jt} denotes regional healthcare expenditure, Z_{jt} the instrumental variable, X_{jt} the vector of covariates, τ_t year fixed effects, α_j region fixed effects, and η_{jt} the idiosyncratic error.

The second-stage regression equation is:

$$H_{jt} = \beta \widehat{Exp}_{jt} + \gamma X_{jt} + \tau_t + \alpha_j + \varepsilon_{jt} \quad (4)$$

where H_{jt} is the regional QALY loss per capita, β the effect of healthcare expenditure, ε_{jt} the idiosyncratic error.

Residual variation in visit numbers and age-sex structures not captured by fixed effects or covariates may persist. To address this, a placebo IV test is conducted. The test uses a placebo IV, which has no impact on the treatment (Felton & Stewart, 2022). The placebo IV (PIV) is constructed by assuming identical conversion factors, the main drivers of the IV variations, across all providers and years:

$$PIV_{jt} = \mathbf{1}_{1 \times 7} \times VN \times ASG_{jt} \quad (5)$$

If the placebo IV neither predicts healthcare expenditure nor health outcomes, it implies that residual influences of visit numbers and age-sex structures are limited enough to be ignored. When performed instead with exposure shares for seven types of providers, this test further examines whether the results are driven by the composition of provider types rather than conversion factors. The details of additional IV diagnostics—including Anderson–Rubin confidence intervals, bias analysis, covariate balance test, and LATE profiling—are provided in Appendix 2.

Furthermore, a series of IV sensitivity analyses are carried out with alternative specifications of the IV. Using contemporaneous regional age-sex structures could be a source of collider bias because they might be affected by health outcomes. To address this, the baseline model is re-estimated using one-year lagged age-sex structures or time-invariant age-sex structures. To further check robustness, we re-estimate the main model using visit numbers from a single baseline year (2008), rather than the average across the entire 2008–2019 period. Additionally, the temporal dependence of conversion factors could be a concern. For example, high conversion factor growth rates in one year might suppress growth in the subsequent year. To account for this possibility, the weighted index of one-year lagged conversion factors is included as a control in a sensitivity analysis. Details of IV sensitivity analysis are provided in Appendix 3.

All analyses are performed in Stata version 18, weighting observations by average municipal populations (2008–2019). All standard

errors are clustered by municipality. To ensure robustness, we also report bootstrapped 95% exposure-robust confidence intervals, as recommended by Borusyak et al. (2025). Detailed procedures for calculating these confidence intervals are provided in Appendix 2.7. Year fixed effects are included via dummy variables, and no models are discounted.

4 | DATA

Regional aggregated panel data are constructed using municipal-level healthcare expenditure and health outcomes data. The time window is from 2008 to 2019, which is determined based on the data availability. Administrative reforms between 2008 and 2019 led to the merger of several municipalities. The total number of municipalities is 240 after the mergers. A sensitivity analysis is conducted by removing these merged municipalities.

4.1 | The Treatment Variable

Regional healthcare expenditure data from 2008 to 2019 were collected from the Korean Statistics Information Service (KOSIS) for the treatment variable. These data include the KNHIS expenditure, copayment, and medical aid spending on insured health benefits. The healthcare expenditure is complete administrative data, not sampled data. All expenditures are converted to US dollars (USD) using the average exchange rate of 1,134.71 KRW/USD for 2008–2019.

We present the replication results of our base model in the appendices, using two alternative measures to aid policymakers in making informed decisions: healthcare expenditures converted to 2019 constant Korean Won using the GDP deflator, and the corresponding values converted to 2019 Purchasing Power Parity (PPP) international dollars. Conversion factors, which represent prices for a relative value unit, are determined in nominal terms. Converting healthcare expenditure to real terms could lead to information loss in the relationship between these variables. Furthermore, year fixed effects in our models account for temporal changes in the price level throughout the study period.

4.2 | Outcome Variables

This study employs the quality-adjusted life year (QALY) loss as the health outcome, one of the most widely used generic health measures. QALY loss comprises the sum of mortality-related QALY loss and morbidity-related QALY loss.

Mortality-related QALY loss captures years of life lost, adjusted for health-related quality of life (HRQoL). It is computed by summing expected EQ-5D index values from the age at death up to the conditional life expectancy for each deceased person, then averaged over the municipal population. For this calculation, expected EQ-5D index values are set to the national average for the corresponding age-sex group in a given year. Complete life tables and regional mortality counts were

sourced from the Korean Statistical Information Service (KOSIS). Year-, age-, and sex-specific HRQoL values are derived from national average EQ-5D indices in the Korea Community Health Survey (KCHS). Mortality-related QALY loss for an expired individual i , sex s , age group a , and year t is defined as:

$$MortQALYL_{ijt} = \sum_{a^d}^{LE(a,s,t)} EQ(a,s,t) \quad (6)$$

where $MortQALYL$ denotes mortality-related QALY loss and a^d the age of death and $LE(a,s,t)$ is life expectancy conditional on age group a , sex s , and year t and $EQ(a,s,t)$ is mean EQ-5D-3L index conditional on age group a , sex s , and year t . Mean mortality-related QALY loss for a municipality j and year t is defined as:

$$\overline{MortQALYL}_{jt} = \sum_i^{N_{jt}^d} MortQALYL_{ijt} \div N_{jt} \quad (7)$$

where N_{jt}^d denotes the number of deaths and N_{jt} the municipal population. The mean mortality-related QALY loss is assumed to equal regional mortality-related QALY loss per capita. An alternative sensitivity analysis assumes unconditional life expectancies of 90.32 years for men and 94.01 years for women—projections for 2072 under the government's best-case scenario (2022 projection).

Morbidity-related QALY loss is calculated from annual municipal mean EQ-5D-3L indices obtained from the KCHS. The KCHS employs a two-tier stratified systematic sampling design, surveying approximately 900 adults per municipality with in-person interviews by trained interviewers (total $N \approx 225,000$; $\sim 0.5\%$ of the population; age-sex distributions in Table A1). Individual sampling weights—reflecting municipal age-sex and household-type structures—are applied to estimate the regional mean EQ-5D index, which is assumed to approximate true mean HRQoL. Tariffs for EQ-5D index calculation follow Lee et al. (2009), which is the latest large-scale research studying EQ-5D-3L tariffs of South Korea.

Mean morbidity-related QALY loss for municipality j in year t is specified as:

$$\overline{MorbQALYL}_{jt} = \sum_i^{N_{jt}^A} (1 - EQ_{ijt}) \div N_{jt}^A \quad (8)$$

where $MorbQALYL$ denotes morbidity-related QALY loss and EQ_{ijt} is individual EQ-5D-3L index and N_{jt}^A is the municipal adult population. Be aware that the KCHS surveys only adults. This formulation assumes that the mean HRQoL of individuals under age 19 is the same as that of individuals over age 19. An alternative sensitivity analysis assumes EQ-5D-3L = 1 for those under 19.

Several additional sensitivity analyses are performed. First, observations with highly unstable QALY loss—indicative of substantial population migration—are excluded. Second, a model with regional mortality counts as the dependent variable is estimated in a sensitivity analysis to assess robustness.

4.3 | Covariates

Regional population data by 37 age-sex groups, annually investigated administrative data, were obtained from the KOSIS. These populations are divided by the municipal total population to compute the proportions of age-sex groups, or the age-sex structures. In the base model, the effects of age-sex structure and year fixed effects are assumed to be homogeneous across all municipalities. In a sensitivity analysis, this assumption is relaxed: age-sex structure effects and year fixed effects are allowed to vary by municipality type (boroughs, urban cities, and rural cities).

Regional injury incidence, defined as the proportion of individuals who visited the emergency department due to injuries in the past 12 months, is included in all sensitivity analyses with extended covariates. Various socioeconomic variables are included in a sensitivity analysis with extended covariates: KNHIS premium per capita (as a proxy for income), proportion of medical aid users, educational attainment, marital status, and occupational composition. Another sensitivity analysis further includes behavioural factors alongside the socioeconomic covariates: regional smoking rates and obesity rates. Furthermore, the numbers of medical doctors, dentists, Korean Medicine doctors, and pharmacists per capita are included as covariates in a sensitivity analysis.

4.4 | Descriptive Statistics

Descriptive statistics are presented for four-year intervals to explore longitudinal patterns (Table 1) because within-municipality variance over time is the main source of the variations for the FEIV estimator model. Healthcare expenditure per capita has increased steadily, while several indicators have trended toward more favourable conditions for better health outcomes. Such socioeconomic improvements and declines in smoking rates imply considerable confounding potential. In the absence of IV estimation, the effect of healthcare expenditure would be indistinguishable from the confounding effects.

5 | RESULTS

This section first presents the results of computing the overall QALY loss. Next, the results of two-way fixed effects models (for comparison) and FEIV estimator models (the main results) are reported. Finally, the sensitivity analysis results are presented.

The national trends in QALY loss per capita, mortality-related QALY loss per capita, and morbidity-related QALY loss per capita remain stable over the study period (Table 2). Although slight upward trends are observable, they are substantially smaller than the corresponding standard deviations. In contrast, healthcare expenditures significantly increase over time. Given South Korea's rapid population ageing, this rise in healthcare expenditure may have contributed to the health improvement.

Table 3 presents the results of fixed effects and FEIV estimator models. In the fixed effects models (Models 1-2), the coefficients for healthcare expenditure are not statistically significant, even after the inclusion of socioeconomic covariates and injury incidence. The proportions of medical aid users and injury incidence both exert detrimental impacts on health. Conversely, higher shares of college graduates and office workers have a protective effect on QALY loss. Marital status and KNHIS premium per capita have no statistically significant coefficients. Log transformation of healthcare expenditure or health outcome variable does not yield significant estimates of the treatment effect.

In contrast to the fixed effects models, all healthcare expenditure coefficients of FEIV estimator models (Models 3-4) are statistically significant and negative, indicating that additional spending reduces QALY loss (Table 3). The bootstrapped exposure-robust confidence intervals of the FEIV estimator models (Models 3 and 4) do not include zero. Incorporating year fixed effects and socioeconomic covariates increases the estimated healthcare productivity but decreases the first-stage F-statistics. Marginal productivity for the KNHIS, expressed as additional cost to avert one QALY loss, ranges from USD 11,020 to USD 12,847 when year fixed effects are included to ensure internal validity. The first-stage F-statistic of Model 3 exceeds the conventional threshold of 10, while that of Model 4 is slightly smaller than 10. The Anderson-Rubin confidence intervals of the two models do not include 0, indicating that the IV remains relevant. Note that the Anderson-Rubin confidence interval is a more robust method for detecting weak instrument bias than the conventional threshold (Andrews, Stock, & Sun, 2019). The p-values of the endogeneity tests are all less than 0.05, implying the endogeneity of the non-IV models. The first-stage results of FEIV estimator models can be found in Tables A3 and A4. Despite the inclusion of numerous covariates in Model 4, Model 3 is preferred as the baseline model because it has a higher first-stage F-statistic and is less vulnerable to collider bias. Thus, Model 3 serves as the baseline for calculating the cost to avert one QALY loss, for IV diagnostics, and for sensitivity analyses. Healthcare expenditure per capita shows a protective effect on both mortality- and morbidity-related QALY loss per capita, with comparable effect sizes (Table A5).

The results of IV diagnostics show no major issue with the IV (Appendix 2). The descriptive statistics and the distribution of the IV can be found in Table A6 and Figure A2, respectively. Nearly all observable covariates and exposure shares show no statistically significant relationship with the instrumental variable. The two variables that show statistically significant relationships exhibit correlations with the IV that are negligible in magnitude (Figure A3). The placebo IV test indicates that variation arises from conversion factors rather than age-sex structures (Tables A7 and A8). The models replacing the placebo IV with constructed exposure shares yield almost identical results to the placebo IV test results (Table A9). The bias analysis results imply that the IV's direct effects on health outcomes would need to be significant at the 10% level for the Anderson-Rubin CI to include zero (Table A12). The complier profiling identifies boroughs as most responsive to the IV and rural cities as least responsive (Table A10). IV sensitivity analyses

using alternative IV specifications (including lagged conversion factors, lagged age-sex structures, and time-invariant age-sex structures) yield results similar to the baseline model (Model 3) estimate (Tables A13). As anticipated, since visit numbers do not vary regionally or temporally, the re-estimated main model using visit numbers from a single baseline year (2008), rather than the average across the entire 2008–2019 period, yields similar estimates (Table A14). The first-stage F-statistics consistently exceed 10 and estimated treatment effects remain statistically significant. The estimated cost to avert one QALY loss across IV sensitivity analyses ranges from USD 12,019 to USD 17,674 (Tables A13 and A14).

Sensitivity analyses appear in Table 4. All healthcare expenditure coefficients remain statistically significant. The estimated cost to avert one QALY loss across sensitivity analyses altering non-IV features spans USD 8,850–17,513, falling within the 95% robust confidence interval of the baseline model (Model 3) estimate (Table 4). The first-stage F-statistics all exceed 10, except for the model with smoking and obesity rates, ranging from 9.375 to 22.39. Alternative calculations of mortality-related QALY loss yield an increased treatment effect, whereas alternative morbidity-related QALY loss computations decrease it. Excluding merged municipalities increases the estimated effect but lowers the first-stage F-statistic. Dropping outliers with high temporal variability diminishes both the treatment effect and the first-stage F statistic modestly. Adding extended covariates—including smoking rates and obesity prevalence alongside socioeconomic variables—increases the treatment effect; however, inclusion of physician density attenuates it. The model with heterogeneous year fixed effects and age-sex structure effects increases the first-stage F-statistic, but the treatment effect diverges markedly from other estimates. Furthermore, this model produces negative predicted healthcare expenditures in about 40% of municipalities in the first stage. Regional mortality rates are highly negatively correlated with healthcare expenditure, as is QALY loss per capita. The estimates derived from alternative healthcare expenditure adjusted for purchasing power parity (PPP) and inflation are broadly consistent with the baseline model (Model 3) estimate (Table A15).

6 | DISCUSSION

This study estimates the local average treatment effect (LATE) of South Korean public healthcare expenditure using regional panel data and the FEIV estimator, providing the first estimate of the marginal productivity of the South Korean public healthcare system. Our key finding is that an additional one thousand USD in healthcare expenditure reduces QALY loss by 0.0778 per capita (95% robust CI: -0.128 to -0.0280) (Table 3). This is equivalent to a required healthcare expenditure of 12,847 USD to avert one QALY loss (95% robust CI: 7,831 to 35,738 USD). The first-stage F-statistic for the baseline model (Model 3) is greater than 10, exceeding the conventional relevance threshold. In contrast to the FEIV models, fixed effects models appear insufficient to fully overcome bias, potentially due to remaining reverse causality and time-varying

confounding effects (Table 3). The placebo IV test indicates that the IV variation is driven by exogenous conversion factors because the placebo IV is not statistically significant in both cases of replacing the original IV and being included as a control with the original IV (Tables A7 and A8). The bias analysis indicates that the estimates are robust because the required t-value of the IV direct effects to make the estimate zero is substantial (Table A12). IV sensitivity analyses suggest robustness of the IV against potential biases from either using contemporaneous age-sex structures or the temporal dependence of conversion factors (Table A13). Across all sensitivity analyses, which alter IV and non-IV features, the required healthcare cost to avert one QALY loss ranges from 8,850 USD to 17,674 USD, reflecting potential treatment effect heterogeneity (Tables 4, A13 and A14). All sensitivity analysis results are statistically significant and within the 95% robust confidence interval of the baseline model (Model 3) estimate, showing the robustness of the model.

The model (3) is a baseline model for the calculation of the cost to avert one QALY loss, IV diagnostics, and sensitivity analyses. Unlike fixed effects models, IV models do not require the inclusion of all possible observable confounders. Covariates in IV models are required to ensure the conditional unconfoundedness for the IV, not for the treatment variable itself. Given that conversion factors are determined at the national level without considering regional characteristics, most regional covariates, except for age-sex structures, may be unnecessary. Moreover, including irrelevant covariates is not recommended owing to the risk of introducing endogenous “bad controls”, which could exacerbate bias beyond that of ordinary least square models with unobservable confounders (Felton & Stewart, 2022; Glynn et al., 2023). Behavioural factors are highly expected to be colliders, and even socioeconomic status could be colliders since they could be affected by both health outcomes and demographic factors.

This study does not examine the specific mechanisms through which conversion factors influence health outcomes via expenditure. However, we offer a tentative hypothesis. Conversion factors may improve health by enabling higher-quality healthcare services through their effect on prices. Healthcare service prices are primarily determined by the conversion factor over the study period. Healthcare providers can plan adjustments to labour and material inputs based on anticipated increases in conversion factors. If conversion factor increases are lower than expected, providers may reduce their labour costs by treating more patients with fewer doctors and nurses, or reduce material costs by substituting lower-cost alternatives for existing materials. From the data and analysis in this study, it is not possible to identify the exact mechanism underpinning our results, and this warrants investigation in future research, but it is plausible that the significant estimated effects of expenditure on health are driven mainly by exogenous variations in quality.

The estimated marginal productivity of healthcare has critical policy implications for the resource allocation policy in the South Korean healthcare system. One of the implications is that the estimate could inform the health opportunity cost, meaning the health gain forgone which displaced investments would have generated. In South Korea, the

government subsidises and oversees the KNHIS to contain healthcare expenditure growth, which means that funding for new health technologies displaces investment in existing ones. Thus, economic evaluations of healthcare interventions ought to reflect the health opportunity cost, which can be proxied by the marginal productivity of the public healthcare system. Consequently, the estimated healthcare cost to avert one QALY loss from this study can inform the cost-effectiveness threshold (CET) in South Korea.

Until 2021, South Korea used the Gross Domestic Product (GDP) per capita as an implicit cost-effectiveness threshold for economic evaluation of pharmaceuticals. Since 2022, this has been replaced by the incremental cost-effectiveness ratios of drugs previously determined to be reimbursed (HIRA, 2022). The practice of using GDP per capita as the CET has been criticised for its lack of a theoretical basis or empirical evidence (Bae, 2018).

The estimated healthcare cost to avert one QALY loss in the baseline model (Model 3) is approximately half of South Korea's GDP per capita in 2008 and about one-third of that in 2018. This finding is consistent with the previous IV estimation studies, which have found that the empirically estimated healthcare expenditure to improve one QALY is typically less than the GDP per capita (Claxton et al., 2018; Laura Catherine Edney et al., 2018; Lomas, Martin, & Claxton, 2019; Siverskog & Henriksson, 2019). The consistency of these IV estimation results from England, Australia, Sweden, and now South Korea could support the validity of the strategy proposed by Claxton et al. (2015).

Furthermore, this study offers an advantage over previous research by employing a consistent methodology and data source. Specifically, it utilises IV estimation with two-way fixed effects and directly measured regional mortality and morbidity data spanning a 12-year period. Although Laura Catherine Edney et al. (2018) exploited individual-level 12-year panel data to compute HRQoL in Australia, they used fixed effects for morbidity-related QALY gain estimation with panel data, while mortality-related QALY gain estimation employed IV two-stage least-squares regression with cross-sectional data. Vallejo-Torres et al. (2018) employed 5-year panel data but relied on predicted, rather than directly measured, EQ-5D data, as HRQoL data were only available for a single year (2011/12). Therefore, this study contributes more robust evidence to the existing literature by precisely capturing regional mortality and morbidity variations.

The estimated healthcare cost to gain one healthy life year in this study is lower than the 20,687 USD per one DALY estimated by Park et al. (2025), a relevant study in South Korea with a similar study period (2010-2019). However, Park et al. (2025) base their estimate on the assumption that healthcare expenditure contributed to 50% of the observed health improvement; if the contribution was assumed to be 80%, the cost per DALY fell to 12,924 USD. Their study included long-term care expenditure, which is not included in the present study, and they exploited decomposition analysis, rather than IV estimation or fixed effects model. However, the cost range of improving one healthy life year is similar to that of the current study.

Furthermore, the baseline model (Model 3) estimate from this study is substantially lower than empirical estimates of willingness-to-pay for one life year, which ranged from 26,209 USD to 35,498 USD at the average exchange rate between 2008 and 2019 (Ahn, Kim, Shin, & Park, 2012). This aligns with literature suggesting that willingness-to-pay estimates are generally much larger than the estimated required healthcare expenditure to gain one QALY (Wouterse, van Baal, Versteegh, & Brouwer, 2023). For example, willingness-to-pay is estimated to be 4.67 times higher than the required healthcare expenditure to gain one QALY in the UK and 2.5 times higher in the US (Woods, Lomas, Sculpher, Weatherly, & Claxton, 2024).

The analysis in this study may either under- or overestimate the required healthcare cost to avert one QALY loss. Underestimation could arise if the IV is positively correlated with the non-insured health benefit expenditure, although non-insured health benefits spending is not considered when determining conversion factors. However, this effect is likely limited due to the low productivity of non-insured health benefits. In this study, we assumed that regional healthcare expenditure, the explanatory variable, is not correlated with spending on primary prevention that is excluded from the explanatory variable. Excluding primary prevention services could also lead to underestimation due to confounding. For example, economic growth could affect both primary prevention spending and the conversion factors. Due to the fragmented nature of funding sources of primary prevention services, data to test this assumption are unavailable. This is a limitation of this study. However, controlling for two-way fixed effects should mitigate this confounding bias. Thus, any remaining bias is likely to be insignificant.

Conversely, the required healthcare cost to avert one QALY loss could be overestimated. The leave-one-out test suggests heterogeneous productivity across municipalities, with a significantly smaller marginal effect observed in boroughs, which have the wealthiest residents and most developed healthcare infrastructure (Table A11). A probable cause is the diminishing marginal returns of healthcare. As boroughs are the most responsive to the IV in the complier profiling test (Table A10), the average treatment effect for the entire population could be larger than the baseline model (Model 3) LATE. Nevertheless, this potential underestimation does not alter the policy implication: the current and previous cost-effectiveness thresholds in South Korea could have been set considerably higher than the health opportunity cost.

To our knowledge, this is the first study to assess the impact of healthcare spending on HRQoL-adjusted life year lost in South Korea using an IV estimation with two-way fixed effects. The results could therefore help policymakers in diverse settings make decisions, such as determining KNHIS premium rates or healthcare-related budgets. In addition, the findings could provide an empirical reference range for the cost-effectiveness threshold in the economic evaluation of healthcare interventions. Given that the GDP per capita was used as the cost-effectiveness threshold until 2021, a significant amount of QALY might have been lost as a result of using a high cost-effectiveness threshold. Hence, this study's policy impact could be considerable.

The strengths of this study are its empirical strategy and high-quality data. The use of an IV helps to mitigate concerns of reverse causality and unobservable confounders, significant obstacles in ascertaining health-care productivity. Incorporating regional fixed effects and year fixed effects could further strengthen the validity of the IV. The 12-year panel dataset consists of precise administrative data for mortality, population, and healthcare expenditure, supplemented by nationally representative, directly surveyed EQ-5D data. With the exception of HRQoL data, all variables in the baseline model (Model 3) are complete administrative data, thereby avoiding sampling bias.

One of the main limitations of this research is its reliance on a static model. Healthcare productivity may change over time, and the effect of healthcare expenditure could persist for more than one year. Thus, a dynamic model might analyse the treatment effect more precisely. However, this study aims to estimate average healthcare productivity over 12 years as a foundational step. Future research could exploit a dynamic model to investigate time trends in healthcare productivity or the duration of healthcare expenditure effects. In addition, future studies could explore regional differences in healthcare productivity, which is not addressed in the present study. The linearity and additivity assumptions of the models are another limitation. Nonetheless, there is no possible empirical strategy to deal with those issues in a model with a continuous IV to our knowledge. Future studies could employ new empirical strategies robust to these issues. Finally, future studies using individual-level data could confirm the marginal healthcare productivity identified in this study.

7 | CONCLUSION

This study is the first to assess the effect of South Korean public health-care expenditure on QALYs using instrumental variable estimation. Our findings indicate that the South Korean public healthcare system requires 12,847 USD to produce one QALY. This study adds robust empirical evidence to the literature on the effect of healthcare expenditure on morbidity by utilising credible health-related quality of life data. Moreover, the estimated cost of gaining one QALY is significantly lower than the conventional benchmarks for valuing healthcare expenditure, such as the GDP per capita or the willingness-to-pay for health in South Korea. This suggests that resource allocation policies in South Korea may need to be reformed in light of this new empirical evidence.

TABLE 1 Descriptive Statistics

Variables	2008-2011		2012-2015		2016-2019	
	Mean	SD	Mean	SD	Mean	SD
Healthcare expenditure per capita (USD)	817	154	1023	202	1426	299
Residents whose age ≥ 60	14.8%	5.7%	17.1%	6.0%	20.7%	6.5%
Residents whose age ≥ 80	1.8%	0.9%	2.3%	1.3%	3.1%	1.7%
Residents who had no education	6.7%	5.1%	5.9%	4.5%	4.8%	3.6%
College graduates	30.6%	12.0%	34.1%	12.0%	38.2%	11.6%
Office workers	10.3%	2.9%	11.5%	3.0%	12.1%	3.2%
Agro-livestock-fisheries workers	4.2%	7.8%	4.0%	8.0%	3.4%	7.1%
KNHIS premium per capita (USD)	496	650	693	434	903	264
Medical aid users	3.4%	2.0%	2.9%	1.5%	2.9%	1.3%
Number of doctors per capita	0.2	0.1	0.2	0.2	0.2	0.2
Obesity rate	22.3%	2.8%	24.8%	2.6%	29.9%	4.1%
Smoking rate	25.1%	3.0%	22.6%	2.9%	21.8%	3.7%
Injury incidence	5.3%	2.0%	7.0%	2.3%	6.4%	2.7%
Observations	960		960		960	

Notes: The statistics reported are calculated from a panel dataset of 240 South Korean municipalities [boroughs ("Gu"), urban cities ("Si"), and rural cities ("Gun")] over three distinct four-year intervals from 2008 to 2019. The number of observations for each period is 960 (240 municipalities $\times$ 4 years). All data were sourced from the Korean Statistics Information Service (KOSIS) and the Korea Community Health Survey (KCHS). All mean calculations are weighted by municipal populations. Healthcare expenditure includes payments for insured health benefits from the Korean National Health Insurance Service (KNHIS), legal copayment, and medical aid programs. Education levels, measured for individuals over age 30, are categorised into five groups: college graduates, high school graduates, middle school graduates, primary school graduates, and individuals with no education. Similarly, there are five occupation categories: managers and professionals, office workers, sales and service workers, agro-livestock-fisheries workers, and manual workers. Although only a selection of education and occupation variables are presented in this table, the full set of categories is included as socioeconomic covariates in the study's regression models. The KNHIS premium per capita is a proxy for income. Medical aid users are defined as individuals with earnings below 40% of the national median income. The number of doctors per capita excludes dentists and Korean Medicine doctors. The obesity rate is the proportion of residents with a BMI of 25 or higher, and the smoking rate is the proportion of residents who currently smoke. Injury incidence represents the proportion of residents who visited an emergency department due to an injury in the past 12 months. All monetary values are expressed in U.S. dollars (USD), converted using the average exchange rate of 1,134.71 KRW/USD for the 2008–2019 period.

Abbreviations: KNHIS = Korean National Health Insurance Service, SD = Standard deviation, USD = United States Dollar.

TABLE 2 Mortality- and Morbidity-related QALY Loss Calculation Results

Year	QALY loss per capita (A) [†]		Mortality-related QALY loss per capita (B) [‡]		Morbidity-related QALY loss per capita (C) [§]		Healthcare expenditure per capita (USD) [¶]		Residents whose age ≥ 60	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
2008	0.141	0.030	0.071	0.021	0.070	0.010	714	78	14.0%	5.5%
2009	0.142	0.029	0.072	0.021	0.070	0.009	792	112	14.5%	5.6%
2010	0.143	0.029	0.073	0.021	0.071	0.009	857	177	15.0%	5.7%
2011	0.144	0.028	0.071	0.021	0.073	0.008	903	159	15.5%	5.8%
2012	0.143	0.028	0.072	0.021	0.071	0.009	935	168	16.1%	5.8%
2013	0.145	0.028	0.071	0.020	0.073	0.009	981	179	16.7%	5.9%
2014	0.146	0.029	0.071	0.020	0.075	0.010	1,053	196	17.4%	6.0%
2015	0.143	0.028	0.071	0.021	0.072	0.009	1,123	213	18.2%	6.1%
2016	0.144	0.029	0.071	0.021	0.073	0.009	1,244	232	19.2%	6.2%
2017	0.142	0.028	0.071	0.021	0.071	0.008	1,346	250	20.1%	6.3%
2018	0.144	0.028	0.073	0.021	0.071	0.009	1,479	269	21.2%	6.4%
2019	0.145	0.028	0.073	0.021	0.071	0.008	1,635	289	22.3%	6.6%
Observations	240		240		240		240		240	

Notes: This table presents the Quality-Adjusted Life Year (QALY) loss, a composite health outcome measure that combines mortality and morbidity, using data from a panel of 240 municipalities. All underlying data were sourced from the Korean Statistics Information Service (KOSIS) and the Korea Community Health Survey (KCHS). All mean calculations are weighted by the municipal population.

[†] QALY loss per capita (A) is the sum of mortality-related QALY loss per capita (B) and morbidity-related QALY loss per capita (C), where $A = B + C$.

[‡] Mortality-related QALY loss per capita (B) measures the years of life lost from premature death, adjusted for health-related quality of life (HRQoL). It is computed by summing expected EQ-5D index values from the age at death up to the conditional life expectancy for each deceased person, then averaged over the municipal population. For this calculation, expected EQ-5D index values are set to the national average for the corresponding age-sex group in a given year.

[§] Morbidity-related QALY loss per capita (C) quantifies the loss of quality of life from health conditions within the living population. It is calculated as one minus the municipality's mean EQ-5D-3L index, which is derived from Korea Community Health Survey data of the adult population. The calculation assumes that the HRQoL of individuals under 19 is the same as that of the adult population.

[¶] Healthcare expenditure includes payments for insured health benefits from the Korean National Health Insurance Service (KNHIS), legal copayment, and medical aid programs.

Abbreviations: QALY = Quality-Adjusted Life Year, SE = Standard error, USD = United States Dollar.

TABLE 3 Fixed Effects (FE) and Fixed Effects Instrumental Variable (FEIV) Regression Results on QALY Loss

Model	(1)	(2)	(3)	(4)
VARIABLES	QALY loss per capita	QALY loss per capita	QALY loss per capita	QALY loss per capita
Healthcare expenditure per capita [†] (1 thousand USD)	-0.00233 (-0.00926, 0.00460)	-0.000434 (-0.00757, 0.00670)	-0.0778*** (-0.128, -0.0280)	-0.0907*** (-0.167, -0.0146)
Medical aid users		0.0998*** (0.0254, 0.174)		-0.0189 (-0.146, 0.108)
College graduates		-0.0897** (-0.169, -0.0100)		0.0302 (-0.108, 0.168)
Office workers		-0.0219*** (-0.0365, -0.00732)		-0.0203** (-0.0384, -0.00219)
Marital status		-0.0131 (-0.0606, 0.0344)		0.0141 (-0.0464, 0.0746)
KNHIS premium per capita (1 thousand USD)		-0.000225 (-0.000943, 0.000493)		-0.000864* (-0.00192, 0.000190)
Injury incidence		0.0551*** (0.0398, 0.0704)		0.0603*** (0.0438, 0.0767)
Instrumental variable (IV)	No	No	Yes	Yes
Age-sex-structure	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Socioeconomic covariates [‡]	No	Yes	No	Yes
Observations	2,880	2,880	2,880	2,880
R-squared	0.279	0.311	0.092	0.068
Cost to avert one QALY loss (USD) [§]	\$428,604	\$2,303,709	\$12,847 (\$7,831, \$35,738)	\$11,020 (\$5,993, \$68,425)
Bootstrapped exposure-robust CI			(-0.0876, -0.0682) (\$11,413, \$14,660)	(-0.1050, -0.0765) (\$9,527, \$13,068)
IV Diagnostics [¶]				
First-stage F-statistics	Not applicable	Not applicable	16.71	9.52
Anderson-Rubin Wald p-value	Not applicable	Not applicable	0.000288	0.00225
Endogeneity test p-value	Not applicable	Not applicable	0.00157	0.00353

Notes: The table shows results from two-way Fixed Effects (FE) models (Models 1-2) and Fixed Effects Instrumental Variable (FEIV) models (Models 3-4) using a two-stage least squares (2SLS) approach. The analysis uses a panel of 2,880 observations (240 municipalities over 12 years from 2008–2019). All data were sourced from the Korean Statistics Information Service (KOSIS) and the Korea Community Health Survey (KCHS). The dependent variable is quality-adjusted life year (QALY) loss per capita, a composite health outcome measure that combines regional mortality and morbidity. All monetary values are expressed in U.S. dollars (USD), converted using the average exchange rate of 1,134.71 KRW/USD for the 2008–2019 period. Age-sex structure includes controls for the proportions of the 37 five-year age-sex groups in the municipal population, as recorded annually on July 1st. Year fixed effects are implemented using 12 dummy variables for the period from 2008 to 2019. Injury incidence is defined as the proportion of residents who visited an emergency department due to an injury in the past 12 months. All regressions are weighted by municipal population.

[†] Healthcare expenditure per capita, the endogenous variable, encompasses payments for insured health benefits from the Korean National Health Insurance Service (KNHIS) and medical aid programs. In model (3) and (4), it is instrumented using a weighted index of conversion factors. Conversion factors denote the prices paid per relative value unit (RVU). The KNHIS employs RVUs to set a fee schedule for each healthcare intervention. Conversion factors are determined through annual negotiations between the government and healthcare providers, thereby having an unpredictable nature. Seven distinct conversion factors correspond to the following provider categories: clinics, hospitals, dentistry, Korean Medicine, midwives, pharmacies, and public health centres. The instrumental variable (IV) is constructed using these conversion factors, weighted by national average healthcare utilisation patterns and regional age-sex demographics. Although conversion factors are determined at the national level, their impacts on regional healthcare expenditures differ. This variation arises because the regional impact of a provider-specific conversion factor is determined by that provider's share of healthcare volume, which varies across regions. This volume share is a function of the age-sex structure within each municipality, as different age-sex groups exhibit distinct healthcare utilisation patterns across provider types.

[‡] Socioeconomic covariates include the following. Medical aid users are individuals with earnings below 40% of the national median income. Education levels, measured for individuals over age 30, are categorised into five groups: college graduates, high school graduates, middle school graduates, primary school graduates, and individuals with no education. Similarly, there are five occupation categories: managers and professionals, office workers, sales and service workers, agro-livestock-fisheries workers, and manual workers. While only a selection of education and occupation variables are shown in this table, the complete set of categories is included as socioeconomic covariates in the regression models. Marital status is the proportion of residents who are legally married and cohabiting with their partner. The KNHIS premium per capita is a proxy for income.

[§] The Cost to avert one QALY loss is calculated using the formula: $-1 \div (\text{coefficient on healthcare expenditure}) \times 1000$.

[¶] IV diagnostic tests assess instrument validity. The Kleibergen and Paap first-stage F-statistic indicates instrument relevance; a value greater than 10 is the conventional threshold suggesting relevance. The Anderson-Rubin confidence interval is a weak-instrument robust confidence interval. It is the most powerful unbiased test in a just-identified IV model and robust regardless of the IV strength and heteroscedasticity (Andrews, Stock, & Sun, 2019). The Wald p-value indicates whether the confidence interval includes zero. The endogeneity test determines whether healthcare expenditure is endogenous using the difference of two Sargan-Hansen statistics (one with the IV and one without); a low p-value confirms that an IV approach is necessary.

Abbreviations: FE = Fixed Effects, FEIV = Fixed Effects Instrumental Variable, IV = Instrumental Variable, QALY = Quality-Adjusted Life Year.

95% Robust confidence intervals in parentheses

*** p<0.01, ** p<0.05, * p<0.1

TABLE 4 Sensitivity Analysis Results

Alternatives [‡]	Coefficients on healthcare expenditure per capita (1 thousand USD) [†]	Cost to avert one QALY loss (USD) [§]	Bootstrapped exposure-robust CI	First-stage F-statistics [¶]	Observations
Baseline [model (3)]	-0.0778*** (-0.128, -0.0280)	12,847 (7,831, 35,738)	(-0.0876, -0.0682) (\$11,413, \$14,660)	16.71	2,880
Alternative mortality-related QALY loss calculation	-0.113*** (-0.175, -0.0514)	8,850 (5,724, 19,460)	(-0.1287, -0.0974) (\$7,773, \$10,272)	16.71	2,880
Alternative morbidity-related QALY loss calculation	-0.0698*** (-0.115, -0.0244)	14,335 (8,690, 40,904)	(-0.0785, -0.0611) (\$12,733, \$16,368)	16.71	2,880
Alternative mortality- and morbidity-related QALY loss calculation	-0.105*** (-0.162, -0.0476)	9,524 (6,160, 21,014)	(-0.1197, -0.0901) (\$8,351, \$11,096)	16.71	2,880
Excluding merged municipalities	-0.0856** (-0.158, -0.0133)	11,680 (6,331, 75,260)	(-0.1002, -0.0712) (\$9,984, \$14,053)	12.46	2,760
Excluding outliers with high variability over time	-0.0746** (-0.134, -0.0147)	13,405 (7,436, 67,987)	(-0.0845, -0.0648) (\$11,841, \$15,435)	15.58	2,796
Extended covariates including smoking and obesity	-0.0933** (-0.171, -0.0153)	10,716 (5,836, 65,408)	(-0.1088, -0.0780) (\$9,193, \$12,824)	9.38	2,880
Extended covariates including smoking, obesity, and the number of doctors	-0.0836** (-0.153, -0.0145)	11,963 (6,548, 69,132)	(-0.0946, -0.0724) (\$10,574, \$13,810)	12.63	2,880
Heterogeneous year fixed effects and age-sex structure effects	-0.0571*** (-0.0975, -0.0167)	17,513 (10,256, 59,844)	(-0.0632, -0.0510) (\$15,821, \$19,626)	22.39	2,880
With regional mortality as the outcome	-0.00285*** (-0.00484, -0.000852)	Not available	(-0.0035, -0.0022)	16.71	2,880

Notes: This table examines the robustness of the primary finding by re-estimating the baseline model [Model (3) from Table 3] under alternative calculation methods, data samples, and model specifications. The analyses use a panel of 2,880 observations from 240 municipalities over 12 years (2008–2019), except in analyses where outliers are excluded. All data were sourced from the Korean Statistics Information Service (KOSIS) and the Korea Community Health Survey (KCHS). All monetary values are expressed in U.S. dollars (USD), converted using the average exchange rate of 1,134.71 KRW/USD for the 2008–2019 period.

[†] All regressions employ the Fixed Effects Instrumental Variable (FEIV) estimation using a two-stage least squares (2SLS) approach and are weighted by municipal population. The treatment variable, healthcare expenditure per capita, encompasses payments for insured health benefits from the Korean National Health Insurance Service (KNHIS), legal copayment, and medical aid programs. The instrumental variable is a weighted index of conversion factors. Conversion factors denote the prices paid per relative value unit (RVU). The KNHIS employs RVUs to set a fee schedule for each healthcare intervention. The outcome variable is quality-adjusted life year (QALY) loss per capita, and all models include region fixed effects, year fixed effects and control for age-sex structures. Age-sex structure includes controls for the proportions of the 37 five-year age-sex groups in the municipal population, as recorded annually on July 1st. Year fixed effects are implemented using 12 dummy variables for the period from 2008 to 2019.

[‡] Each row represents a distinct sensitivity analysis. Explanations are as follows:

Alternative mortality-related QALY loss calculation: Uses projected, unconditional life expectancies (90.32 years for males and 94.01 years for females), adjusted for HRQoL, to calculate life year losses. In the baseline, conditional (on age and sex) life expectancies, adjusted for HRQoL, are used.

Alternative morbidity-related QALY loss calculation: Assumes individuals under 19 have perfect health (EQ-5D-3L index = 1). In the baseline, the mean EQ-5D-3L index of individuals under 19 is assumed to equal that of adults.

Alternative mortality- and morbidity-related QALY loss calculation: Combines the two alternative calculation methods described above.

Excluding merged municipalities: Removes municipalities that were merged due to administrative reforms during the study period.

Excluding outliers with high variability over time: Excludes municipalities with highly unstable QALY loss measures, possibly due to significant population changes.

Extended covariates, including smoking and obesity: Adds regional smoking and obesity rates as control variables alongside socioeconomic covariates (the proportion of medical aid users, education levels, occupation, marital status, and the KNHIS premium per capita) and injury incidence. The obesity rate is the proportion of residents with a BMI of 25 or higher, and the smoking rate is the proportion of residents who currently smoke.

Extended covariates including smoking, obesity, and the number of doctors: Further adds the number of medical professionals as controls: the number of doctors per capita, the number of dentists per capita, the number of Korean Medicine doctors per capita, and the number of pharmacists per capita.

Heterogeneous year fixed effects and age-sex structure effects: Allows year fixed effects and age-sex structure effects to vary by type of municipality (borough, urban city, or rural city).

With regional mortality as the outcome: Replaces the QALY loss per capita variable with regional mortality, defined as the number of regional deaths divided by the municipal population.

[§] The Cost to avert one QALY loss is calculated using the formula: $-1 \div (\text{coefficient on healthcare expenditure}) \times 1000$.

[¶] The Kleibergen and Paap first-stage F-statistic indicates instrument relevance; a value greater than 10 is the conventional threshold suggesting relevance.

Abbreviations: QALY = Quality-Adjusted Life Year.

95% Robust confidence intervals in parentheses

*** p<0.01, ** p<0.05, * p<0.1

References

- Ahn, J. H., Kim, Y. H., Shin, S. J., & Park, J. Y. (2012). A joint asian study on cost-effectiveness in health care decision making. (pp. 17–20).
- Andrews, I., Stock, J. H., & Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annu. Rev. Econom.*, 11(Volume 11, 2019), 727–753.
- Bae, E. Y. (2018). What should the cost-effectiveness threshold reflect? *The Korean Journal of Health Economics and Policy*, 24(2).
- Baltagi, B. H., Moscone, F., & Tosetti, E. (2012). Medical technology and the production of health care. *Empir. Econ.*, 42(2), 395–411.
- Berger, M. C., & Messer, J. (2002). Public financing of health expenditures, insurance, and health outcomes. *Appl. Econ.*, 34(17), 2105–2113.
- Bokhari, F. A. S., Gai, Y., & Gottret, P. (2007). Government health expenditures and health outcomes. *Health Econ.*, 16(3), 257–273.
- Borusyak, K., Hull, P., & Jaravel, X. (2025). A practical guide to shift-share instruments. *J. Econ. Perspect.*, 39(1), 181–204.
- Claxton, K., Lomas, J., & Martin, S. (2018). The impact of NHS expenditure on health outcomes in england: Alternative approaches to identification in all-cause and disease specific models of mortality. *Health Econ.*, 27(6), 1017–1023.
- Claxton, K., Martin, S., Soares, M., Rice, N., Spackman, E., Hinde, S., ... Sculpher, M. (2015). Methods for the estimation of the national institute for health and care excellence cost-effectiveness threshold. *Health Technol. Assess.*, 19(14), 1–503, v–vi.
- Cookson, R., Propper, C., Asaria, M., & Raine, R. (2016). Socio-Economic inequalities in health care in england: Socio-economic inequalities in health care in england. *Fisc. Stud.*, 37(3–4), 371–403.
- Crémieux, P. Y., Ouellette, P., & Pilon, C. (1999). Health care spending as determinants of health outcomes. *Health Econ.*, 8(7), 627–639.
- Cutler, D., Deaton, A., & Lleras-Muney, A. (2006). The determinants of mortality. *J. Econ. Perspect.*, 20(3), 97–120.
- Cutler, D. M., Ghosh, K., Messer, K. L., Raghunathan, T., Rosen, A. B., & Stewart, S. T. (2022). A satellite account for health in the united states. *Am. Econ. Rev.*, 112(2), 494–533.
- Edney, L. C. [Laura C], Lomas, J., Karnon, J., Vallejo-Torres, L., Stadhouers, N., Siverskog, J., ... Ochalek, J. (2022). Empirical estimates of the marginal cost of health produced by a healthcare system: Methodological considerations from Country-Level estimates. *Pharmacoeconomics*, 40(1), 31–43.
- Edney, L. C. [Laura Catherine], Haji Ali Afzali, H., Cheng, T. C., & Karnon, J. (2018). Estimating the reference incremental Cost-Effectiveness ratio for the australian health system. *Pharmacoeconomics*, 36(2), 239–252.
- Farahani, M., Subramanian, S. V., & Canning, D. (2010). Effects of state-level public spending on health on the mortality probability in india. *Health Econ.*, 19(11), 1361–1376.
- Felton, C., & Stewart, B. M. (2022). *Handle with care: A sociologist's guide to causal inference with instrumental variables*.
- Filmer, D., & Pritchett, L. (1999). The impact of public spending on health: Does money matter? *Soc. Sci. Med.*, 49(10), 1309–1323.
- Glynn, A. N., Rueda, M. R., & Schuessler, J. (2023). Post-Instrument bias in linear models. *Sociol. Methods Res.*, 00491241231156965.
- Gravelle, H., & Rees, R. (2004). 10D: Bilateral monopoly. In H. Gravelle & R. Rees (Eds.), *Microeconomics* (3rd ed., pp. 223–226). Harlow, England: Financial Times Prentice Hall.
- HIRA. (2022). The first open to the public of the ICERs of drugs subjected to economic evaluation after introducing the positive list system [press release]. *Health Insurance Review & Assessment Service*.
- Kim, G.-S. (2023). Current status and issues of supplementary health insurance. Korea Insurance Research Institute Seminar.
- Kim, M. J. (2023). Unintended consequences of healthcare reform in south korea: Evidence from a regression discontinuity in time design. *Health Res. Policy Syst.*, 21(1), 60.
- Kirkegaard, J. F. (2021). The pandemic's long reach: South korea's fiscal and fertility outlook. *Policy Briefs*.
- Kwon, S. (2019). Case study: Republic of korea. In S. L. Barber, L. Lorenzoni, P. Ong, et al. (Eds.), *Price setting and price regulation in health care: Lessons for advancing universal health coverage*. World Health Organization.
- Lee, Y.-K., Nam, H.-S., Chuang, L.-H., Kim, K.-Y., Yang, H.-K., Kwon, I.-S., ... Kim, Y.-T. (2009). South korean time trade-off values for EQ-5D health states: Modeling with observed values for 101 health states. *Value Health*, 12(8), 1187–1193.
- Lomas, J., Martin, S., & Claxton, K. (2019). Estimating the marginal productivity of the english national health service from 2003 to 2012. *Value Health*, 22(9), 995–1002.
- Martin, S., Lomas, J., Claxton, K., & Longo, F. (2021). How effective is marginal healthcare expenditure? new evidence from england for 2003/04 to 2012/13. *Appl. Health Econ. Health Policy*, 19(6), 885–903.
- Martin, S., Rice, N., & Smith, P. C. (2008). Does health care spending improve health outcomes? evidence from english programme budgeting data. *J. Health Econ.*, 27(4), 826–842.
- Moler-Zapata, S., Kreif, N., Ochalek, J., Mirelman, A. J., Nadjib, M., & Suhrcke, M. (2022). Estimating the health effects of expansions in health expenditure in indonesia: A dynamic panel data approach. *Appl. Health Econ. Health Policy*, 20(6), 881–891.
- OECD. (2024). *Fiscal sustainability of health systems*. OCED Publishing.
- Park, S., Dieleman, J., Weaver, M. R., Bae, G., & Eggleston, K. (2025). Health care spending increases and value in south korea. *JAMA Health Forum*, 6(1), e245145–e245145.
- Siverskog, J., & Henriksson, M. (2019). Estimating the marginal cost of a life year in sweden's public healthcare sector. *Eur. J. Health Econ.*, 20(5), 751–762.
- Thornton, J. A. (2011). Does more medical care improve population health? new evidence for an old controversy. *Appl. Econ.*, 43(24), 3325–3336.
- Vallejo-Torres, L., García-Lorenzo, B., & Serrano-Aguilar, P. (2018). Estimating a cost-effectiveness threshold for the spanish NHS. *Health Econ.*, 27(4), 746–761.
- Weaver, M. R., Joffe, J., Ciarametaro, M., Dubois, R. W., Dunn, A., Singh, A., ... Dieleman, J. L. (2022). Health care spending effectiveness: Estimates suggest that spending improved US health from 1996 to 2016. *Health Aff.*, 41(7), 994–1004.
- WHO. (2015). *Republic of korea health system review*. Manila: WHO Regional Office for the Western Pacific.
- Woods, B., Lomas, J., Sculpher, M., Weatherly, H., & Claxton, K. (2024). Achieving dynamic efficiency in pharmaceutical innovation: Identifying the optimal share of value and payments required. *Health Econ.*, 33(4), 804–819.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data: Second edition (the MIT press)* (second edition). MIT Press.
- Wouterse, B., van Baal, P., Versteegh, M., & Brouwer, W. (2023). The value of health in a Cost-Effectiveness analysis: Theory versus practice. *Pharmacoeconomics*, 41(6), 607–617.

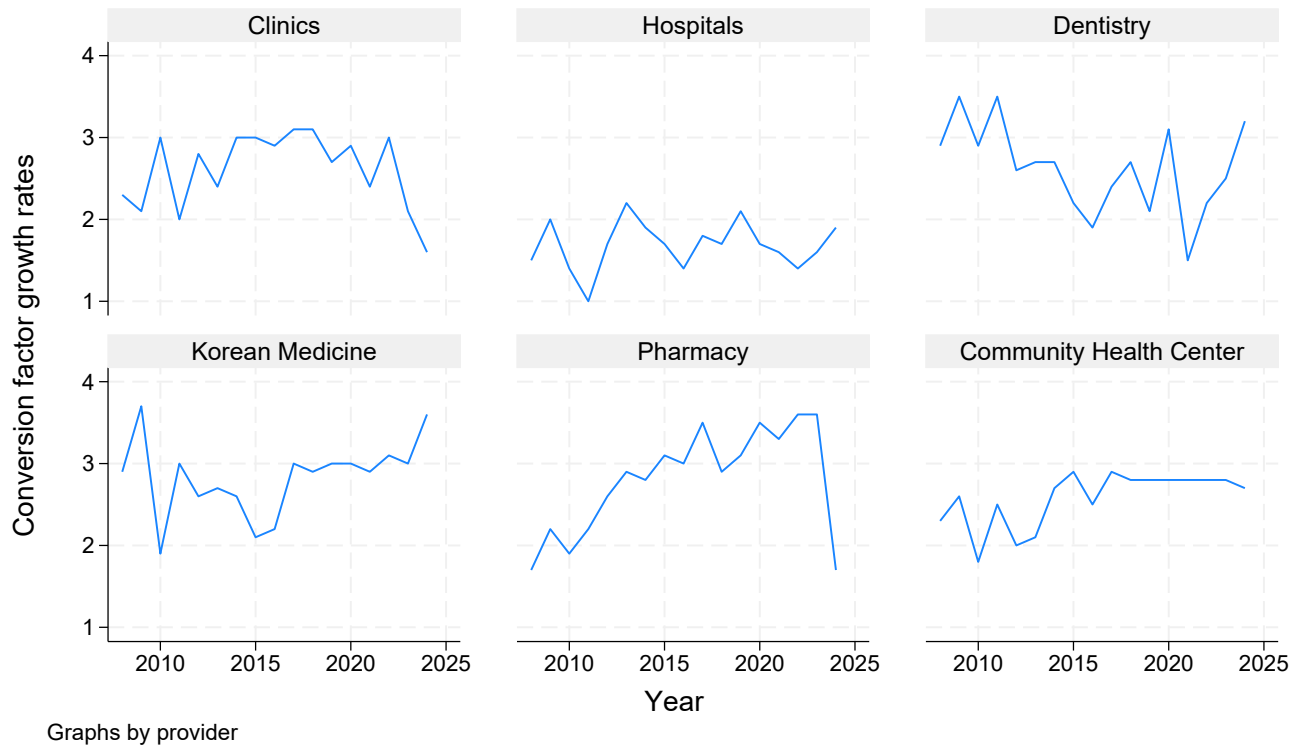


FIGURE 1 Conversion factor growth rates from 2008 to 2024

Note: Conversion factors denote the prices paid per relative value unit (RVU). The KNHIS employs RVUs to set a fee schedule for each healthcare intervention. Conversion factors are determined through annual negotiations between the government and healthcare providers, thereby having an unpredictable nature. Seven distinct conversion factors correspond to the following provider categories: clinics, hospitals, dentistry, Korean Medicine, midwives, pharmacies, and public health centres. The midwives' conversion factors are omitted due to its small share in the total health-care expenditure (less than 0.1%).

Source: Korean National Health Insurance Service