



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.

Appendices

Contents

Appendix 1. Additional tables and a figure	2
Appendix 2. Instrumental variable diagnostics	12
Appendix 2.1. Descriptive Statistics of the IV and Shares	12
Appendix 2.2. Covariate balance test	12
Appendix 2.3. Placebo IV test	14
Appendix 2.4. Shares analysis	18
Appendix 2.5. Local Average Treatment Effect (LATE) profiling	21
Appendix 2.6. Bias analysis	25
Appendix 3. IV sensitivity analyses	26
Appendix 4. Healthcare expenditure in real terms and PPP international dollars	32
Bibliography	35

Appendix 1. Additional tables and a figure

Table A1. Age-sex distributions of samples in the Korean Community Health Survey

Year	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Total samples	220,258	230,715	229,229	229,226	228,921	228,781	228,712	228,558	228,452	228,381	228,340	229,099
Sex												
Male	101,358 (46.0)	107,080 (46.4)	104,575 (45.6)	103,017 (44.9)	102,898 (44.9)	102,722 (44.9)	103,537 (45.3)	102,829 (45.0)	103,048 (45.1)	102,484 (44.9)	102,241 (44.8)	102,572 (44.8)
Female	118,900 (54.0)	123,635 (53.6)	124,654 (54.4)	126,209 (55.1)	126,023 (55.1)	126,059 (55.1)	125,175 (54.7)	125,729 (55.0)	125,404 (54.9)	125,897 (55.1)	126,099 (55.2)	126,527 (55.2)
Age												
19-29	26,074 (11.8)	30,733 (13.3)	28,169 (12.3)	26,045 (11.4)	25,370 (11.1)	24,794 (10.8)	25,617 (11.2)	24,740 (10.8)	25,355 (11.1)	24,328 (10.7)	23,440 (10.3)	23,383 (10.2)
30-39	40,047 (18.2)	41,590 (18.0)	40,135 (17.5)	38,018 (16.6)	36,629 (16.0)	34,471 (15.1)	34,477 (15.1)	32,460 (14.2)	31,390 (13.7)	29,887 (13.1)	28,235 (12.4)	26,712 (11.7)
40-49	44,837 (20.4)	47,905 (20.8)	46,033 (20.1)	45,280 (19.8)	45,062 (19.7)	43,247 (18.9)	44,085 (19.3)	41,895 (18.3)	40,955 (17.9)	39,484 (17.3)	36,905 (16.2)	35,911 (15.7)
50-59	39,078 (17.7)	41,684 (18.1)	42,820 (18.7)	44,891 (19.6)	45,260 (19.8)	45,909 (20.1)	45,644 (20.0)	45,471 (19.9)	45,862 (20.1)	45,144 (19.8)	44,522 (19.5)	44,171 (19.3)
60-69	35,766 (16.2)	35,329 (15.3)	36,138 (15.8)	36,247 (15.8)	36,014 (15.7)	36,252 (15.8)	35,706 (15.6)	38,365 (16.8)	38,788 (17.0)	40,748 (17.8)	42,611 (18.7)	44,941 (19.6)
70-	34,456 (15.6)	33,474 (14.5)	35,936 (15.7)	38,745 (16.9)	40,586 (17.7)	44,108 (19.3)	43,183 (18.9)	45,627 (20.0)	46,102 (20.2)	48,790 (21.4)	52,627 (23.0)	53,981 (23.6)

Unit: the number of samples (%)

Source: Korean Community Health Survey Guidebook (2019)

Table A2. The negotiation breakdown history of selected conversion factors in South Korean healthcare system

Year	Clinics	Hospitals	Dentistry	Korean medicine	Pharmacies
2024	X				X
2023	X			X	
2022		X	X		
2021	X	X	X		
2020	X				
2019	X		X		
2018					
2017					
2016		X	X		
2015			X	X	
2014					
2013	X		X		
2012		X			
2011	X				
2010	X	X			
2009	X				
2008	X	X			

X: breakdown

* Two healthcare providers were excluded due to their insignificant shares in medical expenditures (less than 1%).

Notes: Conversion factors are a key component of the pricing formula for healthcare services, and they significantly affect healthcare expenditure. The seven provider categories are as follows:

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. 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 category of healthcare provider is set by the Health Insurance Policy Deliberation Committee, which consists of representatives of payers, providers, and public agencies.

Table A3. First-stage regression results of FEIV estimator models

Model	(5)	(6)
VARIABLES	Healthcare expenditure per capita (1 thousand USD)	Healthcare expenditure per capita (1 thousand USD)
The weighted index of conversion factors (IV)†	2.03e-06*** (4.97e-07)	1.46e-06*** (4.74e-07)
Medical aid users		-1.129*** (0.383)
College graduates		1.138*** (0.390)
Marital status		0.456* (0.249)
KNHIS premium per capita (1 thousand USD)		-0.00756** (0.00314)
Office workers		0.0167 (0.0661)
Injury incidence		0.0635 (0.0505)
Age-sex-structure	Yes	Yes
Region fixed effects	Yes	Yes

Year Fixed effects	Yes	Yes
Socioeconomic covariates‡	No	Yes
Observations	2,880	2,880
R-squared	0.984	0.985

Notes: The table shows first-stage regression results from Fixed Effects Instrumental Variable (FEIV) models (Models 3-4) in Table 3. 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 Korean Community Health Survey (KCHS). The dependent variable is healthcare expenditure per capita, the endogenous variable, encompassing payments for insured health benefits from the Korean National Health Insurance Service (KNHIS), legal copayment, and medical aid programs.. 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.

† 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.

Abbreviations: FEIV = Fixed Effects Instrumental Variable, IV = Instrumental Variable, KNHIS = Korean National Health Insurance Service.

Robust standard errors in parentheses

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

Table A4. First and second stage regression results of the base model

Model	(5)	(7)
VARIABLES	Healthcare expenditure per capita (1 thousand USD)	QALY loss per capita
The weighted index of conversion factors (IV)†	2.03e-06*** (4.97e-07)	
Healthcare expenditure per capita (1 thousand USD)		-0.0778*** (0.0254)
Year 2008	0.734** (0.300)	-0.0125 (0.0132)
Year 2009	0.691** (0.280)	-0.00942 (0.0122)
Year 2010	0.629** (0.259)	-0.00885 (0.0116)
Year 2011	0.556** (0.241)	-0.00992 (0.0118)
Year 2012	0.422* (0.217)	-0.0172 (0.0128)
Year 2013	0.317 (0.193)	-0.0188 (0.0128)
Year 2014	0.241 (0.165)	-0.0176 (0.0118)
Year 2015	0.153 (0.136)	-0.0200* (0.0108)

Year 2016	0.130 (0.106)	-0.0143* (0.00819)
Year 2017	0.0664 (0.0699)	-0.0132** (0.00588)
Year 2018	0.0240 (0.0350)	-0.00714** (0.00317)
Age-sex-structure	Yes (Joint F-stat: 11.8, p<0.001)	Yes (Joint F-stat: 407.82, p<0.001)
Region fixed effects	Yes	Yes
Year Fixed effects	Yes (Joint F-stat: 70.98, p<0.001)	Yes (Joint F-stat: 91.45, p<0.001)
Socioeconomic covariates†	No	No
Observations	2,880	2,880
R-squared	0.984	0.092

Notes: The table shows first- and second-stage regression results from Fixed Effects Instrumental Variable (FEIV) models (Model 3) in Table 3. 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 Korean Community Health Survey (KCHS). The dependent variable is healthcare expenditure per capita, the endogenous variable, encompassing payments for insured health benefits from the Korean National Health Insurance Service (KNHIS), legal copayment, and medical aid programs.. 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.

† 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.

Abbreviations: FEIV = Fixed Effects Instrumental Variable, IV = Instrumental Variable, KNHIS = Korean National Health Insurance Service.

Robust standard errors in parentheses

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

Table A5. Mortality and morbidity results of FEIV estimator models

Model	(8)	(9)
VARIABLES	Mortality-related QALY loss per capita‡	Morbidity-related QALY loss per capita§
Healthcare expenditure per capita (1 thousand USD)†	-0.0313*** (-0.0544, -0.00822)	-0.0465** (-0.0851, -0.00795)
Instrumental variable (IV)	Yes	Yes
Age-sex-structure	Yes	Yes
Region fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Socioeconomic covariates	No	No
Observations	2,880	2,880

R-squared	0.291	0.043
IV Diagnostics ^{††}		
First-stage F-statistics	16.71	16.71
Anderson-Rubin Wald p-value	0.00190	0.0106
Endogeneity test p-value	0.00485	0.0214

Notes: The table shows regression results of mortality- and morbidity-related quality-adjusted life years (QALY) loss per capita from the Fixed Effects Instrumental Variable (FEIV) model (Model 3) in Table 3. The analysis uses a panel of 2,880 observations (240 municipalities over 12 years from 2008 to 2019). All data were sourced from the Korean Statistics Information Service (KOSIS) and the Korean Community Health Survey (KCHS). The dependent variables are mortality- and morbidity-related QALY loss per capita. 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), legal copayment, and medical aid programs. In models (7) and (8), 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.

[‡] 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 Korean 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.

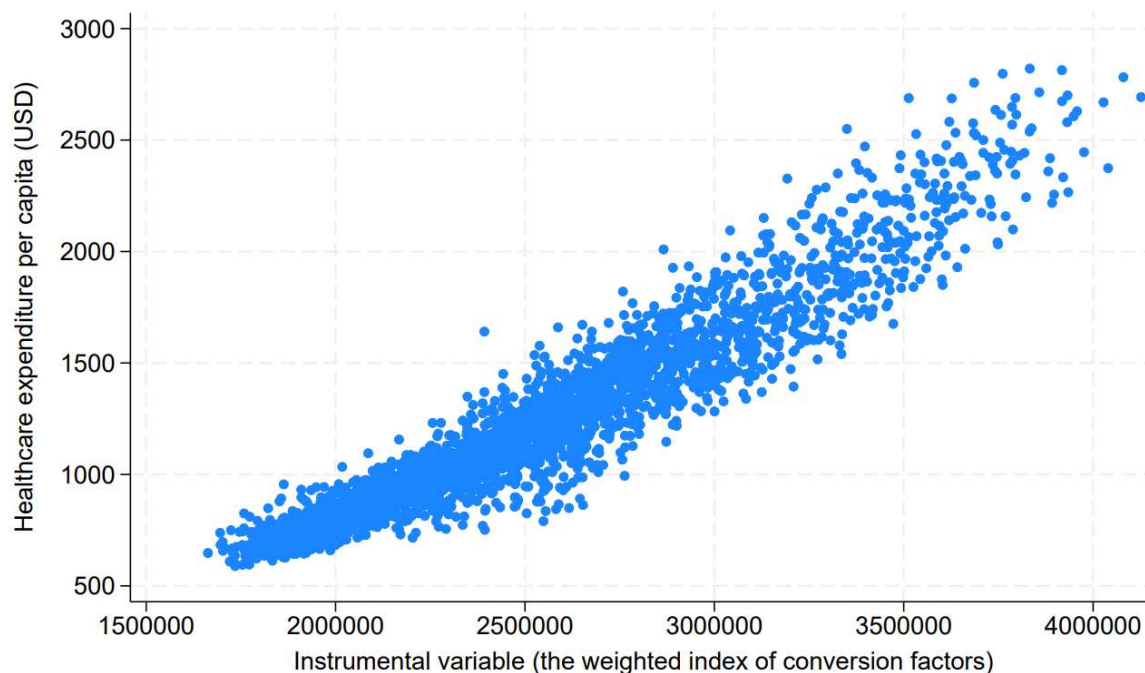
¶ IV diagnostic tests assess instrument validity. The Kleibergen and Paap first-stage F-statistic indicates instrument relevance; a value greater than 10 is the conventional threshold, suggesting relevance. The Anderson-Rubin confidence interval is a weak-instrument robust confidence interval. It is 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: 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

Appendix Figure 1. The monotonicity assumption



Notes: Healthcare expenditure per capita, the endogenous variable, 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. 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.

Source: Korean Statistical Information System and Korean National Health Insurance Service

Appendix 2. Instrumental variable diagnostics

This study uses the Anderson-Rubin confidence interval to rule out the weak instrument bias. The strong instrumental variable (IV) is a basic requirement for the IV regression. Furthermore, the bias from the violation of the exclusion restriction assumption is exacerbated by the weak instrument (Felton & Stewart, 2022). The conventional yardstick of the IV weakness is whether the F statistic in the first stage is larger than 10, but recently, many researchers have doubted the credibility of the rule of thumb (Felton & Stewart, 2022). Anderson-Rubin confidence interval, the most powerful unbiased test in a just-identified IV model, is robust regardless of the IV strength and heteroscedasticity (Andrews, Stock, & Sun, 2019).

Appendix 2.1. Descriptive Statistics of the IV and Shares

We report descriptive statistics for the instrument and the shares in Table A6. The shares are annual visit numbers per 1,000 residents by provider types, conditional on regional age-sex structures. The IV is the sum of the product of conversion factors and shares.

Table A6. Descriptive Statistics: Instrumental Variable and Shares

VARIABLES	Mean	SD	Min	Max
IV ($\times 10^3$)	2337.6	381.5	1663.1	4126.7
Share of Clinics	10606.4	907.1	9108.8	15107.7
Share of Hospitals	5295.2	927.9	3755.2	10564.8
Share of Dental Clinics	1251.1	52.5	1136.8	1477.8
Share of Korean Medicine	2021.9	309.9	1382.9	3458.6
Share of Midwifery	0.1	0.0	0.0	0.1
Share of Pharmacies	9745.1	789.6	8449.8	13651.5
Share of Public Health Centres	307.4	103.8	131.7	846.8

Notes: Statistics are weighted by population weight.

Appendix 2.2. Covariate balance test

We conduct a covariate balance test to check the exogeneity of the IV. This test examines the relationship between the IV and the covariates included in the models.

To carry out the diagnostic test, we compute a residualised Instrumental Variable (IV) ([Dobbie, Goldin, & Yang, 2018](#)) using a two-way fixed effects model that regresses the IV on age-sex structures as follows.

$$IV_{jt} = \gamma ASG_{jt} + \tau_t + \alpha_j + ResidIV_{jt}$$

Where IV_{jt} denotes the instrumental variable, ASG_{jt} age-sex structures, τ_t year fixed effects, α_j region fixed effects, and $ResidIV_{jt}$ residualised IV.

This residualised IV is unaffected by region fixed effect, year fixed effect and age-sex structures, and yields the same results as the original IV in the base model.

The residualised IV is approximately normally distributed and exhibits sufficient variation in the histogram in Figure A2.

We regress each observable control variable on the residualised IV using a two-way fixed effects model with clustered standard errors. Note that all models are univariate analyses to avoid multicollinearity, which could significantly increase standard errors. A coefficient plot (Figure A3) with a 95% confidence interval is provided, with all variables standardised to allow for a direct comparison of correlation strength. While two observable control variables, the proportion of agri-livestock-fishery workers and the number of pharmacists, are correlated with the residualised IV, all standardised coefficients for observable control variables are less than 0.05. This means that a one standard deviation (SD) change in the IV is associated with a change of less than 0.05 SD in any of the control variables. Furthermore, the inclusion of these two variables in the sensitivity analyses (Table 4) yields estimates very similar to our main model. It is worth noting that the predicted shares are not correlated with the residualised IV.

In conclusion, the results of the covariate balance test suggest that the IV is unlikely to be significantly influenced by potential confounders.

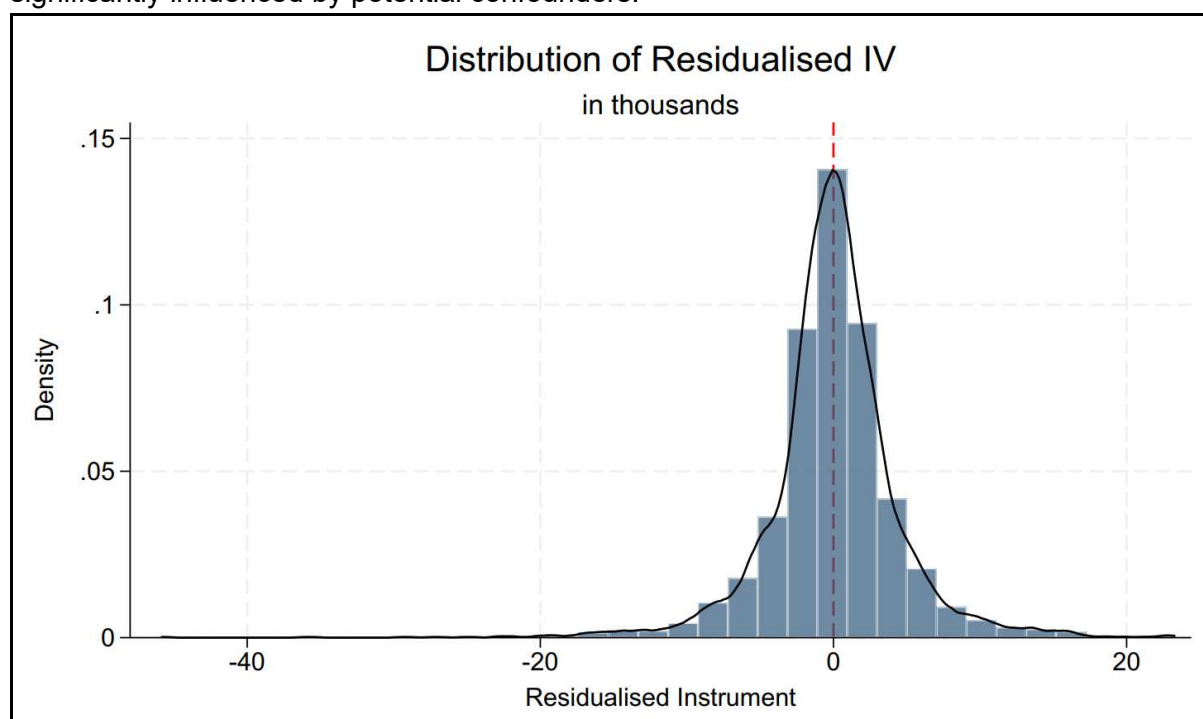


Figure A2. The Distribution of Residualised IV

Notes: The instrumental variable (IV) is residualised using a two-way fixed effects model that regresses the IV on the age-sex structure.

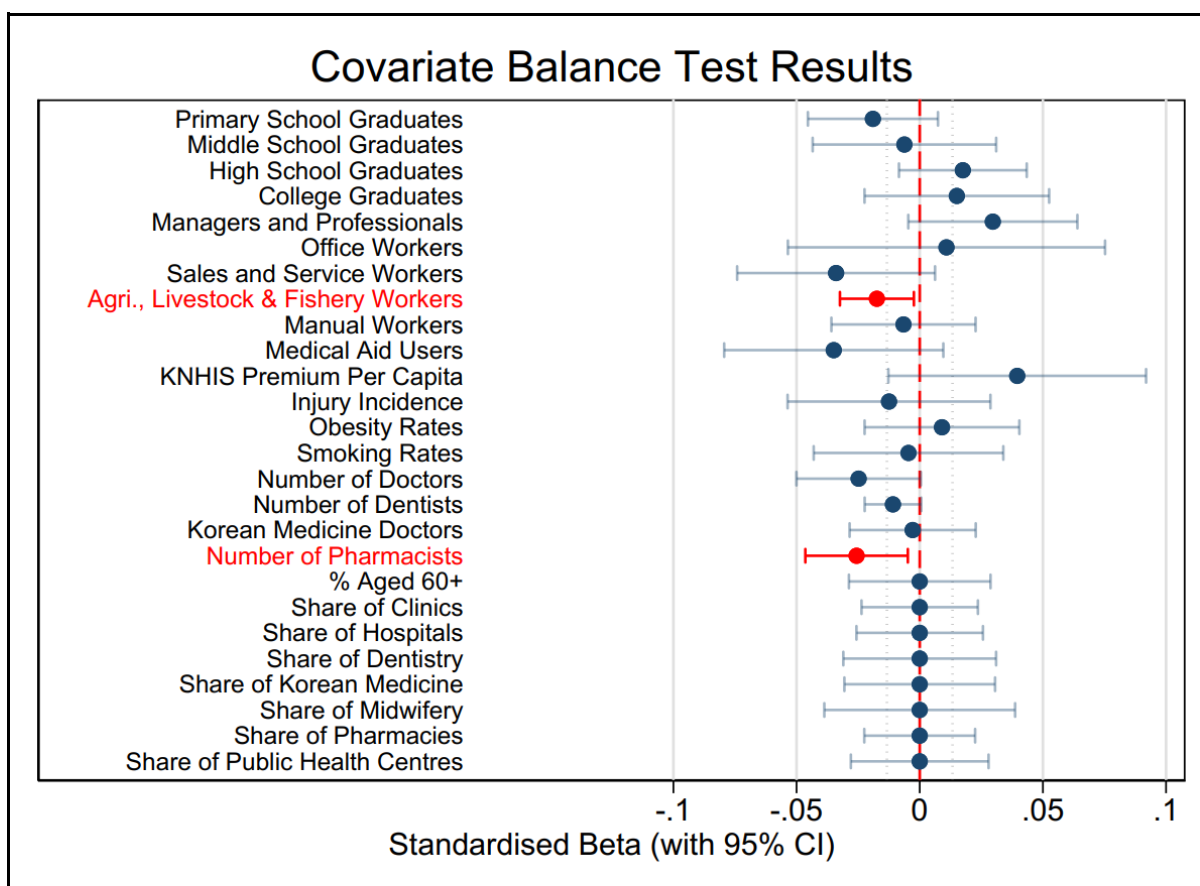


Figure A3. Covariate Balance Test Results

Notes: The instrumental variable (IV) is residualised using a two-way fixed effects model that regresses the IV on the age-sex structure. Each observable control variable is then regressed on the residualised IV using a two-way fixed effects model with clustered standard errors. All models are univariate to avoid multicollinearity. All variables are standardised to allow for a direct comparison of correlation strength.

Appendix 2.3. Placebo IV test

The weighted index of conversion factors is used as the IV. Although conversion factors are determined at the national level, they have different impacts on regional healthcare expenditures. 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. This volume share is a function of the age-sex structure within each municipality, since different age-sex groups exhibit 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 healthcare utilisation rates of age-sex groups by healthcare providers are incorporated into the IV. The IV structure is:

$$IV_{jt} = CF_t \times VN \times ASG_{jt}$$

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 and serve as frequency weights for conversion factors. VN is averaged across regions and years to ensure that it is uncorrelated with region- or time-specific confounders, and thus it does not contribute to IV variation. Finally, let ASG_{jt} be a 38×1 vector representing the proportion of each age-sex group in municipality j in year t . Although these age-sex structures could drive IV variation, their potential confounding effects could be controlled by including the exact same age-sex structures as covariates.

However, 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:

$$PIV_{jt} = 1_{1 \times 7} \times VN \times ASG_{jt}$$

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. The first-stage regression and reduced form regression results of the placebo IV models are reported. Also, the results of the baseline model with the placebo IV as a covariate could reinforce the robustness.

Furthermore, we present the partial R-squared values of the original IV and the placebo IV in the first stage in Table A7. In the base model, the partial R-squared of the original IV is 2.9%, whereas that of the placebo IV in Model 10 is 0.05%. Even if the total R-squared values of the two models are approximately the same, the partial R-squared value of the original IV is far larger than that of the placebo IV. The 2.9% is not a small number, given that the canonical IV presented by Angrist and Krueger (Angrist & Krueger, 1991), the individual's season of birth, had partial R-squared values between 0.01% and 0.05% (Bound, Jaeger, & Baker, 1995). Therefore, we believe our IV has sufficient variation and explanatory power to be exploited.

Table A7. First-stage regression results of placebo IV tests

Model	(10)	(11)	(12)
VARIABLES	Healthcare expenditure per capita (1 thousand USD)	Healthcare expenditure per capita (1 thousand USD)	Healthcare expenditure per capita (1 thousand USD)
The weighted index of conversion factors (IV)†	2.03e-06*** (1.58e-06, 2.49e-06)		2.03e-06*** (1.57e-06, 2.48e-06)
Placebo IV‡		-2.089 (-5.572, 1.393)	-1.568 (-5.007, 1.871)
Age-sex-structure	Yes	Yes§	Yes§
Region fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Socioeconomic covariates	No	No	No
Observations	2,880	2,880	2,880
R-squared	0.984	0.984	0.984
Partial R ² of the IV	2.888%	Not applicable	2.867%
Partial R ² of the Placebo IV	Not applicable	0.0534%	0.0311%

Notes: The table shows first-stage regression results with placebo IV from the Fixed Effects Instrumental Variable (FEIV) model (Model 3) in Table 3. 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 Korean Community Health Survey (KCHS). The dependent variable is healthcare expenditure per capita, the endogenous variable, encompassing payments for insured health benefits from the Korean National Health Insurance Service (KNHIS), legal copayment, and medical aid programs. 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. All regressions are weighted by municipal population.

† 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.

‡ The placebo IV is constructed without conversion factors, the main drivers of the IV variations.

§ 26 age-sex group variables are dropped due to the multicollinearity when including the placebo IV in models (11) and (12).

Abbreviations: FEIV = Fixed Effects Instrumental Variable, IV = Instrumental Variable.

95% Confidence intervals in parentheses

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

Table A8. Reduced-form regression results of placebo IV tests

Model	(13)	(14)	(15)
VARIABLES	QALY loss per capita	QALY loss per capita	QALY loss per capita
The weighted index of conversion factors (IV)†	-1.58e-07*** (-2.26e-07, -9.01e-08)		-1.59e-07*** (-2.27e-07, -9.08e-08)
Placebo IV‡		-0.121 (-0.638, 0.395)	-0.161 (-0.677, 0.354)
Age-sex-structure	Yes	Yes§	Yes§
Region fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Socioeconomic covariates	No	No	No
Observations	2,880	2,880	2,880
R-squared	0.285	0.279	0.285

Notes: The table shows reduced-form regression results with placebo IV from the Fixed Effects Instrumental Variable (FEIV) model (Model 3) in Table 3. 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 Korean 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. All regressions are weighted by municipal population.

† 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.

‡ The placebo IV is constructed without conversion factors, the main drivers of the IV variations.

§ 26 age-sex group variables are dropped due to the multicollinearity when including the placebo IV in models (14) and (15).

Abbreviations: FEIV = Fixed Effects Instrumental Variable, IV = Instrumental Variable, QALY = quality-adjusted life year.

95% Confidence intervals in parentheses

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

Appendix 2.4. Shares analysis

In the shift-share IV scheme, shares need to be included as a control if they do not sum to one, as this could be a source of confounding (Borusyak, Hull, & Jaravel, 2025). However, the constructed shares, $VN \times ASG$, cannot be included in our base model due to perfect multicollinearity with existing age-sex structures (ASG). That is, since the matrix VN is constant across all observations, the shares are linear combinations of the ASG variables. Therefore, controlling for ASG already achieves the same effect as controlling for $VN \times ASG$. Instead, we constructed a placebo IV that assumes identical conversion factors across all providers and years.

As a further validation, we replicated the placebo IV test, replacing the placebo IV with seven provider shares as covariates (Table A9). The results are identical to the original placebo IV test. Note that Two-Stage Least Squares (2SLS) method is not viable for the model that includes these shares due to perfect multicollinearity.

Table A9. First-stage and reduced-form regression results with shares of the IV

Model	(10)	(16)	(13)	(17)
VARIABLES	Healthcare expenditure per capita (1 thousand USD)	Healthcare expenditure per capita (1 thousand USD)	QALY loss per capita	QALY loss per capita
The weighted index of conversion factors (IV) [†]	2.03e-06*** (1.05e-06, 3.01e-06)	2.03e-06*** (1.05e-06, 3.01e-06)	-1.58e-07*** (-2.45e-07, -7.13e-08)	-1.58e-07*** (-2.45e-07, -7.13e-08)
Share of clinics [‡]		-0.0404 (-0.127, 0.0458)		-0.00760 (-0.0173, 0.00212)
Share of hospitals		-0.00382 (-0.0104, 0.00279)		-0.000592 (-0.00138, 0.000200)
Share of dentistry		-0.0106 (-0.143, 0.121)		-0.00926 (-0.0298, 0.0113)

Share of Korean Medicine		0.0244 (-0.0239, 0.0727)		0.00408 (-0.00132, 0.00947)
Share of midwifery		-5,864 (-15,174, 3,446)		-105.1 (-1,402, 1,192)
Share of pharmacies		0.0265 (-0.0548, 0.108)		0.00740 (-0.00201, 0.0168)
Share of public health centres		0.126* (-0.00963, 0.261)		0.0111 (-0.00460, 0.0267)
Age-sex-structure	Yes	Yes§	Yes	Yes§
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Socioeconomic covariates	No	No	No	No
Observations	2,880	2,880	2,880	2,880
R-squared	0.984	0.984	0.285	0.285

Notes: The table shows first-stage and reduced-form regression results of Fixed Effects Instrumental Variable (FEIV) model (Model 3) in Table 3, with exposure shares as covariates. 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 Korean Community Health Survey (KCHS). The dependent variable is healthcare expenditure per capita, the endogenous variable, encompassing payments for insured health benefits from the Korean National Health Insurance Service (KNHIS), legal copayment, and medical aid programs. 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. All regressions are weighted by municipal population.

† 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.

‡ 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 IV is the sum of the product of conversion factors and shares.

§ 8 age-sex group variables are dropped due to the multicollinearity when including the shares in models (16) and (17).

Abbreviations: FEIV = Fixed Effects Instrumental Variable, IV = Instrumental Variable.

95% Robust confidence intervals in parentheses

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

Appendix 2.5. Local Average Treatment Effect (LATE) profiling

The LATE profiling is carried out to address a concern that the LATE could not represent the whole population due to the limited compliance with the IV. In a continuous treatment variable and IV setting, compliers could be found by comparing the IV coefficients of different subgroups in the first-stage regression (Nicoletti, Salvanes, & Tominey, 2023; von Hinke, Rice, & Tominey, 2022). That is, the subgroup where the medical spending is more responsive to the IV than other subgroups is more likely to be compliers. In this study, the type of municipality is used for dividing the observations into subgroups because boroughs, urban cities, and rural cities have quite different socioeconomic and demographic characteristics. Also, the leave-one-out test is conducted to explore the heterogeneous LATEs in different types of municipalities, as the plain subgroup analysis cannot secure enough observations for the IV estimation. This test repeats the baseline analysis, excluding one subgroup and compares the effects. By combining the results of the LATE profiling and the leave-one-out test, the gap between the average treatment effect and the LATE could be revealed. Also, the effects of medical spending on the YLD and YLL per capita are estimated separately to explore the treatment effect heterogeneity.

Table A10. LATE profiling results

Model	(18)	(19)	(20)
VARIABLES	Healthcare expenditure per capita (1 thousand USD)	Healthcare expenditure per capita (1 thousand USD)	Healthcare expenditure per capita (1 thousand USD)
The weighted index of conversion factors (IV) [†]	2.73e-06*** (1.28e-06, 4.17e-06)	2.03e-06*** (5.06e-07, 3.54e-06)	1.46e-06*** (4.95e-07, 2.43e-06)
Age-sex-structure	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Socioeconomic covariates	No	No	No
Municipality type [‡]	Boroughs	Urban cities	Rural cities
Observations	1,164	792	924
R-squared	0.989	0.991	0.991

Notes: The table shows first-stage regression results with different types of municipality from the Fixed Effects Instrumental Variable (FEIV) model (model 3) in Table 3. The analysis uses subsets of 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 Korean Community Health Survey (KCHS). The dependent variable is healthcare expenditure per capita, the endogenous variable, encompassing payments for insured health benefits from the Korean National Health Insurance Service (KNHIS), legal copayment, and medical aid programs. 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. All regressions are weighted by municipal population.

[†] 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.

‡ In this research, municipalities are defined as the lowest-level administrative units that have a basic local government. South Korea consists of 9 provinces ("Do") and 7 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").

Abbreviations: FEIV = Fixed Effects Instrumental Variable, IV = Instrumental Variable.

95% Robust confidence intervals in parentheses

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

Table A11. Leave-one-out test results

Model	(21)	(22)	(23)
VARIABLES	QALY loss per capita	QALY loss per capita	QALY loss per capita
Healthcare expenditure per capita (1 thousand USD)†	-0.127*** (-0.224, -0.0305)	-0.0540** (-0.100, -0.00771)	-0.0676*** (-0.114, -0.0215)
Instrumental variable (IV)	Yes	Yes	Yes
Age-sex-structure	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Socioeconomic covariates	No	No	No
Municipalities included‡	Urban and rural cities	Boroughs and rural cities	Boroughs and urban cities
Observations	1,716	2,088	1,956
Cost to avert one QALY loss (USD)§	7,845 (4,455, 32,800)	18,524 (9,974, 129,674)	14,798 (8,796, 46,586)
IV Diagnostics¶			
First-stage F-statistics	15.69	17.48	11.73
Anderson-Rubin Wald p-value	0.000182	0.0175	0.000973

Endogeneity test p-value	0.00191	0.0308	0.00501
--------------------------	---------	--------	---------

Notes: The table shows leave-one-out test results from Fixed Effects Instrumental Variable (FEIV) models (Model 3) in Table 3. The analysis uses subsets of a panel of 2,880 observations (240 municipalities over 12 years from 2008 to 2019). All data were sourced from the Korean Statistics Information Service (KOSIS) and the Korean 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), legal copyment, and medical aid programs. In models (7) and (8), 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.

‡ In this research, municipalities are defined as the lowest-level administrative units that have a basic local government. South Korea consists of 9 provinces ("Do") and 7 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 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: FEIV = Fixed Effects Instrumental Variable, IV = Instrumental Variable, QALY = quality-adjusted life year.

95% Robust confidence intervals in parentheses

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

Appendix 2.6. Bias analysis

The bias analysis and the placebo IV test are carried out to investigate the possible violations of the exclusion restriction assumption. The bias analysis is a grid-based approach analysing the treatment effects with a set of hypothetical direct effects of the IV on the outcome (Claxton et al., 2015; Felton & Stewart, 2022; Wang, Jiang, Zhang, & Small, 2018). The method is to report the treatment effects of the IV regressions with the adjusted outcomes subtracted by the hypothetical direct effects multiplied by the instrument (Felton & Stewart, 2022). The direct effects are standardised using the standard error of the reduced form regression (Wang et al., 2018). The direct effects which make the Anderson-Rubin 95% confidence interval include zero or reduce the treatment effect to zero are reported.

Table A12. Bias analysis result

	Direct effect	t-value†
Anderson-Rubin CI includes 0‡	-7.30e-08	-1.66
Treatment effect = 0	-1.59e-07	-3.61

Notes: The robust standard error in the reduced form regression in the baseline FEIV model (Model 13) (4.41e-08) is utilised to compute the t-values.

† Conventionally, the absolute value of the t-value of 2 or higher suggests a statistically significant effect.

‡ 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). In general, Anderson-Rubin confidence intervals are wider than usual confidence intervals.

Abbreviations: CI = Confidence Interval.

Appendix 3. IV sensitivity analyses

IV sensitivity analyses are conducted with alternative specifications of the IV. Using contemporaneous regional age-sex structures could be a source of 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. In the baseline, the IV structure is as follows.

$$IV_{jt} = CF_t \times VN \times ASG_{jt}$$

An alternative IV with one-year lagged age-sex structures is specified as:

$$IV_{jt}^l = CF_t \times VN \times ASG_{j\ t-1}$$

An alternative IV with time-invariant age-sex structures is specified as:

$$IV_{jt}^{ASG2007} = CF_t \times VN \times ASG_{j\ 2007}$$

Another concern is the temporal dependence of conversion factors. 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 are included as a control in a sensitivity analysis. The control with one-year lagged conversion factors (CIV) is specified as:

$$CIV_{jt} = CF_{t-1} \times VN \times ASG_{jt}$$

Finally, we conduct additional IV sensitivity analyses by constructing IVs using only VN and ASG in 2008, rather than those from 2008-2019 (Table A14).

$$IV_{jt}^{VN2008} = CF_t \times VN_{2008} \times ASG_{jt}$$

$$IV_{jt}^{2008} = CF_t \times VN_{2008} \times ASG_{j\ 2008}$$

Our constructed shares (VN×ASG) are designed to capture expected healthcare needs rather than endogenous actual utilisation. Specifically, the weighting matrix VN is constructed by averaging national mean visit frequencies across 38 age-sex groups over a 12-year period (2008-2019). By using long-term national averages, we effectively remove regional variations and year-specific shocks. We then interact these fixed weights with each municipality's age-sex population proportions (ASG) to derive the final shares. As this measure relies exclusively on national biological profiles and local demographics, it remains orthogonal to local supply-side factors.

While these constructed shares should be, in principle, exogenous, we tested this assumption by using VN and ASG in 2008, rather than those from 2008-2019.

Table A13. IV sensitivity analysis results 1

Model	(3) Baseline	(24)	(25)	(26)
VARIABLES	QALY loss per capita	QALY loss per capita	QALY loss per capita	QALY loss per capita
Healthcare expenditure per capita (1 thousand USD)†	-0.0778*** (-0.128, -0.0280)	-0.0801*** (-0.135, -0.0247)	-0.0625*** (-0.0990, -0.0259)	-0.0832*** (-0.131, -0.0350)
Instrumental variable‡				
Conversion factor (CF)	Current	Current	Current	Current
Visit numbers (VN)	Time-invariant national average	Time-invariant national average	Time-invariant national average	Time-invariant national average
Age-sex structure (ASG)	Current (2008-2019)	One-year lagged (2007-2018)	Time-invariant (2007)	Current (2008-2019)
Controls				
Age-sex-structure	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Socioeconomic covariates	No	No	No	No
IV with one-year lagged conversion factors (2007-2018)§	No	No	No	Yes
Observations	2,880	2,880	2,880	2,880

Cost to avert one QALY loss (USD)	12,847 (7,831, 35,738)	12,488 (7,384, 40,460)	16,008 (10,098, 38,598)	12,019 (7,608, 28,599)
IV Diagnostics†				
First-stage F-statistics	16.71	12.64	30.54	39.87
Anderson-Rubin Wald p-value	0.000288	0.00225	0.000179	0.000103
Endogeneity test p-value	0.00157	0.00748	0.000970	0.000217

Notes: The table shows IV sensitivity analysis results. The baseline is a Fixed Effects Instrumental Variable (FEIV) model (model 3) in Table 3. The analysis uses a panel of 2,880 observations (240 municipalities over 12 years from 2008 to 2019). All data were sourced from the Korean Statistics Information Service (KOSIS) and the Korean Community Health Survey (KCHS). The dependent variable is QALY loss per capita. 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. All regressions are weighted by municipal population. The Cost to avert one QALY loss is calculated using the formula: $-1 \div (\text{coefficient on healthcare expenditure}) \times 1000$.

† 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. 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.

‡ Alternative specifications are applied to the IV. In model (24), one-year lagged age-sex structures replaced the current ones. In model (25), time-invariant age-sex structures in 2007 replaced the current ones.

§ In model (26), the IV remains the same, but an alternative version of the IV with one-year lagged conversion factors is included as a control.

¶ 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: 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 A14. IV sensitivity analysis results 2

Model	(3) Baseline	(27)	(28)
VARIABLES	QALY loss per capita	QALY loss per capita	QALY loss per capita
Healthcare expenditure per capita (1 thousand USD)†	-0.0778*** (-0.128, -0.0280)	-0.0723*** (-0.118, -0.0267)	-0.0566*** (-0.0898, -0.0233)
Instrumental variable‡			
Conversion factor (CF)	Current	Current	Current
Visit numbers (VN)	Time-invariant national average	National average in 2008	National average in 2008
Age-sex structure (ASG)	Current (2008-2019)	Current (2008-2019)	Time-invariant (2008)
Controls			
Age-sex-structure	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Socioeconomic covariates	No	No	No

Observations	2,880	2,880	2,880
Cost to avert one QALY loss (USD)	12,847 (7,831, 35,738)	13,823 (8,476, 37,444)	17,674 (11,130, 42,894)
IV Diagnostics¶			
First-stage F-statistics	16.71	18.83	35.35
Anderson-Rubin Wald p-value	0.000288	0.000322	0.000317
Endogeneity test p-value	0.00157	0.00176	0.00188

Notes: The table shows IV sensitivity analysis results. The baseline is a Fixed Effects Instrumental Variable (FEIV) model (model 3) in Table 3. The analysis uses a panel of 2,880 observations (240 municipalities over 12 years from 2008 to 2019). All data were sourced from the Korean Statistics Information Service (KOSIS) and the Korean Community Health Survey (KCHS). The dependent variable is QALY loss per capita. 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. All regressions are weighted by municipal population. The Cost to avert one QALY loss is calculated using the formula: $-1 \div (\text{coefficient on healthcare expenditure}) \times 1000$.

† 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. 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.

‡ Alternative specifications are applied to the IV. In model (27), national average visit numbers in 2008 replaces the national mean visit numbers averaged over the period 2008-2019. In model (28), time-invariant age-sex structures in 2008 replaced the current ones.

¶ 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: FEIV = Fixed Effects Instrumental Variable, IV = Instrumental Variable, QALY = quality-adjusted life year.

95% Robust confidence intervals in parentheses

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

Appendix 4. Healthcare expenditure in real terms and PPP international dollars

We present the replication results of our base model 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. GDP deflator and PPP data are sourced from the World Bank's World Development Indicators. The estimates derive from alternative healthcare expenditure adjusted for purchasing power parity (PPP) and inflation are broadly consistent with the baseline estimate (Table A15). Consequently, we have maintained our original base model.

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.

Table A15. Additional analyses in real terms and PPP international dollars

Model	(3) Baseline	(29)	(30)
VARIABLES	QALY loss per capita	QALY loss per capita	QALY loss per capita
Healthcare expenditure per capita†	-0.0778*** (-0.128, -0.0280)	-0.0702*** (-0.120, -0.0200)	-0.0829*** (-0.142, -0.0236)
Unit of Healthcare expenditure‡			
Currency	USD	USD	KRW (X 1000)
Exchange rate	Average exchange rate 2008-2019	2019 PPP international dollars	Not applicable
GDP deflator (base year: 2019)	Not applied (Nominal)	Applied (Real)	Applied (Real)
Controls			
Age-sex-structure	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes

Socioeconomic covariates	No	No	No
Observations	2,880	2,880	2,880
Cost to avert one QALY loss	\$12,847 (7,831, 35,738)	\$14,240 (8,299, 50,102)	12,069 (7,034, 42,464) (1000 KRW)
IV Diagnostics¶			
First-stage F-statistics	16.71	10.69	10.69
Anderson-Rubin Wald p-value	0.000288	0.000288	0.000288
Endogeneity test p-value	0.00157	0.00152	0.00152

Notes: The baseline is a Fixed Effects Instrumental Variable (FEIV) model (model 3) in Table 3. The analysis uses a panel of 2,880 observations (240 municipalities over 12 years from 2008 to 2019). All data were sourced from the Korean Statistics Information Service (KOSIS) and the Korean Community Health Survey (KCHS). The dependent variable is QALY loss per capita.

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. All regressions are weighted by municipal population. The Cost to avert one QALY loss is calculated using the formula: $-1 \div (\text{coefficient on healthcare expenditure}) \times 1000$.

† 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. 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.

‡ Baseline healthcare expenditure is calculated using the average exchange rate of 1,134.71 KRW/USD for the 2008–2019 period. We apply the GDP deflator to the expenditure in Model 29 and convert this into 2019 PPP international dollars. In Model 30, we apply the GDP deflator to the expenditure to express it in 2019 Korean Won. GDP deflator and PPP data are sourced from the World Bank's World Development Indicators.

¶ 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: FEIV = Fixed Effects Instrumental Variable, IV = Instrumental Variable, QALY = quality-adjusted life year.

95% Robust confidence intervals in parentheses

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

Bibliography

- Andrews, I., Stock, J. H., & Sun, L. (2019). Weak Instruments in Instrumental Variables Regression: Theory and Practice. *Annual Review of Economics*, 11(Volume 11, 2019), 727–753. Retrieved April 1, 2024 from
- Angrist, J. D., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics*, 106(4), 979–1014. Retrieved December 25, 2025 from
- Borusyak, K., Hull, P., & Jaravel, X. (2025). A practical guide to shift-share instruments. *The Journal of Economic Perspectives: A Journal of the American Economic Association*, 39(1), 181–204. Retrieved July 16, 2025 from
- Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90(430), 443. Retrieved December 25, 2025 from
- 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 Technology Assessment*, 19(14), 1–503, v–vi.
- Dobbie, W., Goldin, J., & Yang, C. S. (2018). The effects of pre-trial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review*, 108(2), 201–240. Retrieved December 22, 2025 from
- Felton, C., & Stewart, B. M. (2022, August). *Handle with Care: A Sociologist's Guide to Causal Inference with Instrumental Variables*. Retrieved from <https://doi.org/10.31235/osf.io/3ua7q>
- Nicoletti, C., Salvanes, K. G., & Tominey, E. (2023). Mothers working during preschool years and child skills: Does income compensate? *Journal of Labor Economics*, 41(2), 000–

000. Retrieved August 2, 2024 from
- von Hinke, S., Rice, N., & Tominey, E. (2022). Mental health around pregnancy and child development from early childhood to adolescence. *Labour Economics*, 78, 102245.
- Wang, X., Jiang, Y., Zhang, N. R., & Small, D. S. (2018). Sensitivity analysis and power for instrumental variable studies. *Biometrics*, 74(4), 1150–1160.