HEALTH CARE
COST INSTITUTE

2019 Health Care Cost and Utilization Report

Analytic Methodology

2019 V1.0

September 2021

Note: This analytic methodology is appropriate for the *2019 Health Care Cost and Utilization Report*, as our methods are continually refined. Interested parties are encouraged to refer to the appropriate methodology and report.

HEALTH CARE COST INSTITUTE

Table of Contents

HEALTH CARE
COST INSTITUTE

1. Introduction

For the *2019 Health Care Cost and Utilization Report*, the Health Care Cost Institute (HCCI) presented national estimates of health care spending, utilization, prices, and service-mix for the population of individuals younger than 65 and covered by employer-sponsored private health insurance (ESI). The data behind these estimates came from a national, multipayer, commercial health care claims database containing information provided by three data contributors – Aetna, Humana, and Blue Health Intelligence. The HCCI dataset contains over 1 billion commercial medical and pharmacy claims per year, representing the health care activity of more than 55 million individuals per year for the years 2012 through 2019. This document describes in detail the methods used to transform raw claims into descriptive statistics presented in the annual report.

For the annual *Health Care Cost and Utilization* reports HCCI produces an analytic subset of its database, consisting of all non-Medicare claims for beneficiaries younger than age 65, covered by ESI and whose claims were filed with a contributing health plan between 2015 and 2019. Figure 1 shows the process used to clean the ESI claims data. The process included categorizing claims, calculating utilization by service category, and adjustments to make the data representative of the national population younger than 65 with ESI. Additional adjustments were made as noted in the report for age and sex, inflation, and service-mix.

The data are made representative of the national population younger than 65 with ESI using population weights based on U. S. Census Bureau data. A completion method was used to estimate the components of claims that were incomplete at the end of the reporting period for 2019 data. In most cases, no adjustment was performed for inflation, so the estimated dollars in these reports are nominal unless otherwise stated.

HEALTH CARE COST INSTITUTE

FIGURE 1: PROCESS FLOW

A note on premiums

HCCI does not report on premiums or their determinants. For more information on health insurance premiums and the multiple factors that affect them (including health care expenditures; insured, group, and market characteristics; benefit design; and the regulatory environment), see Congressional Budget Office, Private Health Insurance and Federal Policy,[i] and Kaiser Family Foundation and Health Research & Education Trust, 2018 Employer Health Benefits Survey.[ii]

Changes in the methodology

Compared to earlier versions, HCCI's updated analytic methodology included the following changes.

- An analysis decomposing spending growth into various factors is included. Details on the methodology are included in Section 3.3 of this document.
- For the 2019 analytic dataset, 2015-2018 were considered complete, and no actuarial adjustment was performed. The 2019 claims were actuarially completed using completion factors based on 2018 claims data.
- Service-mix changes were estimated based on changes in the DRGs, CPT codes, and NDC codes submitted on claims from 2015 to 2019. This method replaces previous method that relied on average intensity weights of diagnosis-related groups (DRGs), relative value units (RVUs), and ambulatory payment classifications (APCs) reported by the Centers for Medicare and Medicaid Services (CMS). Further details on the methodology used to estimate changes in service mix are included in Section 3.2 of this document.

HEALTH CARE
COST INSTITUTE

2. Methods

2.1 Data collection

HCCI has access to health care claims data for approximately 55 million Americans in every year between 2012 and 2019 who have commercial health insurance coverage. This dataset was developed from de-identified claims data that were compliant with the Health Insurance Portability and Accountability Act (HIPAA) and included the allowed amounts (actual prices paid) to providers for services. To produce the findings in the *2019 Health Care Cost and Utilization Report*, HCCI used an analytic subset of its data consisting of all eligible claims for insured individuals younger than age 65, covered by either fully-insured or self-insured employer-sponsored health insurance (ESI).

The final analytic subset consisted of approximately 52 million covered lives per year, for the years 2015 through 2019 (Table 1). The claims used in the 2019 report include over 5 billion claim lines and represent the health care activity of 35% of all individuals younger than 65 covered by ESI, making this one of the largest data sources on the privately insured available.

TABLE 1: ANALYTIC SUBSET FOR 2019 REPORT – TOTAL COVERED LIVES BY CALENDAR YEAR

Year	Covered Lives
2015	52,500,000
2016	52,500,000
2017	52,300,000
2018	51,500,000
2019	53,400,000

Source: HCCI, 2021. Notes: Data refer only to HCCI holdings of claims for beneficiaries covered by employer-sponsored health insurance and younger than age 65. Data rounded to the nearest 100,000.

Between July 2020 and November 2020, each data contributor updated the 2018 claims data they previously submitted in addition to providing new data from 2019. HCCI's data manager confirmed the data integrity of each claims file (membership, medical, and pharmacy) in each year with the appropriate data contributor.

From these base datasets, a single analytical dataset was constructed for analysis using the process shown in Figure 1. Analysis of the analytic dataset is described in Section 3.

2.2 Claims categorization

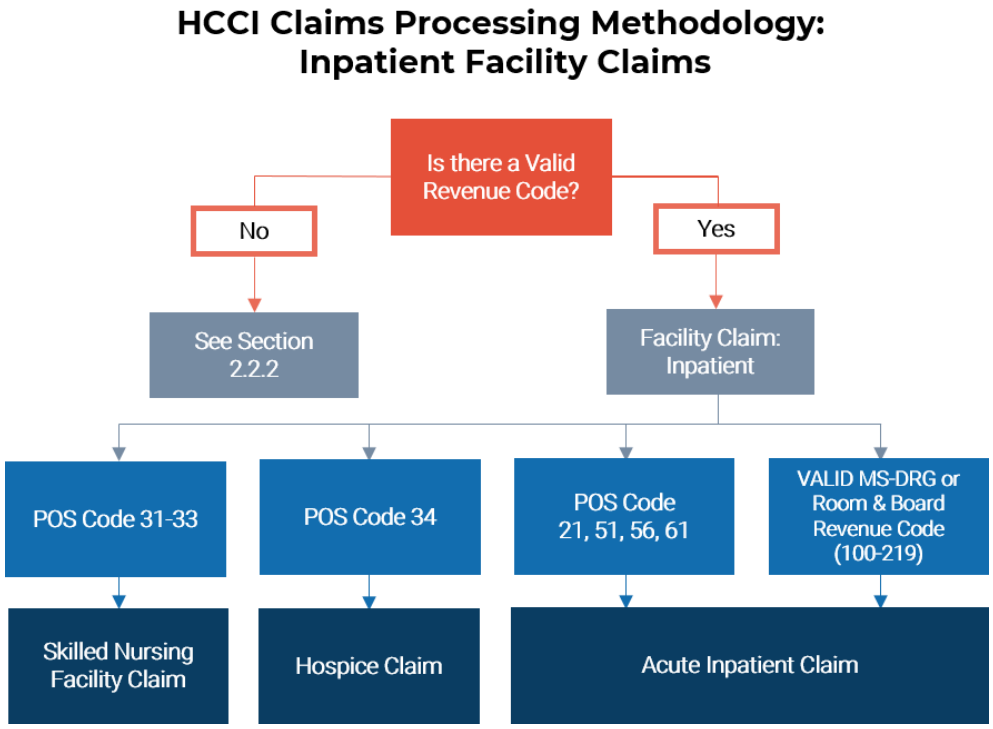
At the highest level, claims data were grouped into four major service categories: inpatient facility, outpatient facility, professional procedure, and prescription drugs and devices.

Inpatient facility claims were from hospitals, skilled nursing facilities (SNFs), and hospices, where there was evidence that the insured stayed overnight (Figure 2). The outpatient facility category contained claims that did not include an overnight stay but included observation and emergency room claims as well as claims for other outpatient services (Figure 3). Both outpatient and inpatient claims were for only the facility charges associated with such claims. HCCI classified services as professional procedure services if claims did not include valid revenue codes (i.e., were not billed by facilities).

2.2.1 Facility claims

Medical claims with a valid revenue code were assumed to be facility claims. In absence of that, claims were assumed to be professional procedure claims. Once processed, facility claims were grouped into two major service categories—inpatient and outpatient based on place of service (POS) codes (Figure 2 and Figure 3).

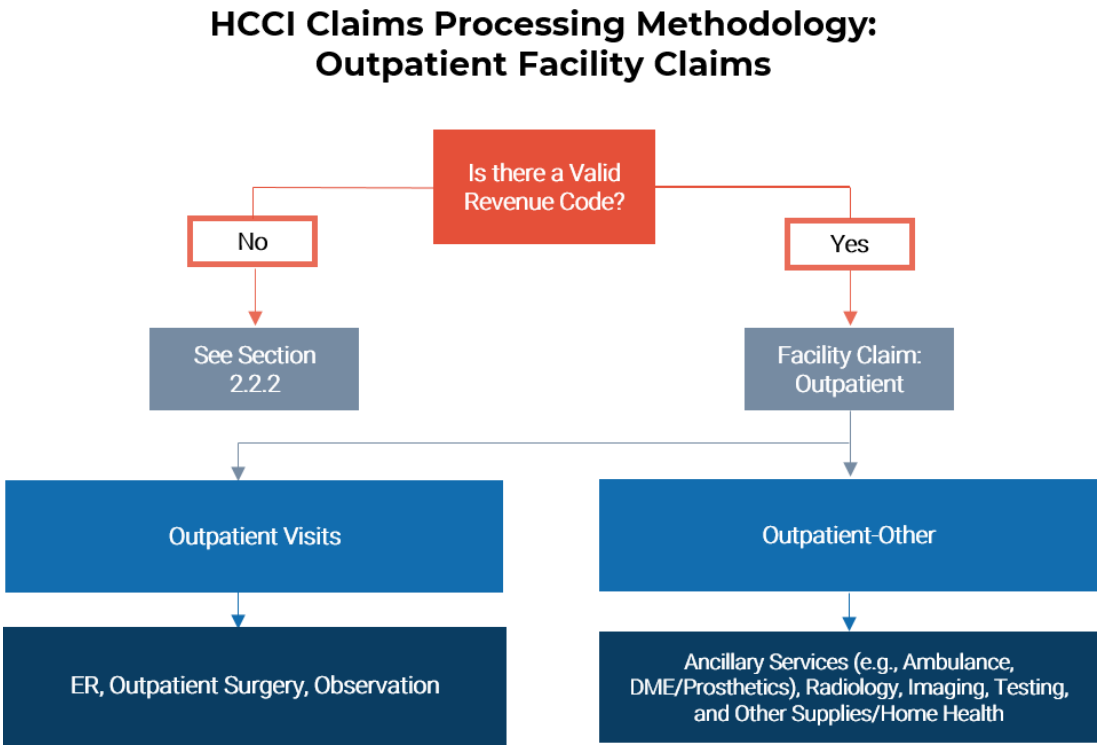
FIGURE 2: FACILITY CLAIMS PROCESS, INPATIENT



2.2.1.1 Inpatient facility claims

Inpatient services are rendered when patients are kept overnight for treatment but not observation (Figure 2). The inpatient services category included claims with the following criteria: place of service (POS) codes 21, 51, 56, and 61; a valid Medicare Severity Diagnosis-Related Group (MS-DRG) code (V32); or a room and board revenue code of 100-219. This category also included skilled nursing facility (SNF) and hospice claims.

FIGURE 3: FACILITY CLAIMS PROCESS, OUTPATIENT



2.2.1.2 Outpatient facility claims

Outpatient services are rendered by sections of a hospital providing medical services that do not require an overnight stay or hospitalization (e.g., emergency room (ER), outpatient surgery, observation). These services can also be provided at freestanding outpatient facilities, including free-standing surgical centers, ambulatory surgical centers (ASCs), and clinics with certain diagnostic testing technologies (e.g., MRIs). These outpatient facilities all file Health Care Financing Administration (HCFA) 1500 form with insurers. The outpatient category was used for all facility claims not characterized as inpatient (Figure 3).

HEALTH CARE
COST INSTITUTE

2.2.2 Professional procedure and prescription claims

2.2.2.1 Professional procedure claims

Professional procedure claims are claims filed by a health care professional for medical services provided. These services included those provided in both hospital and non-hospital settings. Claims with no valid revenue code were assumed to be a professional procedure claim, unless otherwise noted.

The professional procedure category also includes facility claims for some independent clinics, such as small private practices, and multi-specialty clinics (e.g., offering primary care and x-rays). Clinics included in the professional procedure category did not file a HCFA 1500 with insurers.

2.2.2.2 Prescription drug claims

Prescription drug claims are claims submitted by retail, mail-order, and specialty pharmacies for prescription drugs and devices.

Administered drugs and any devices rendered by a physician or facility were identified as professional procedures, outpatient services, or part of an inpatient admission.

HEALTH CARE
COST INSTITUTE

2.3 Adjustment methodologies

2.3.1 Claims completion methodology

Claims data reflect health care services performed (i.e., claims incurred) in the year noted. Claims generally require time for submission to the payer, processing, and payments to the provider (sometimes called the claim payment lag time or run-out period).

Completion is a standard actuarial practice designed to allow for the calculation of utilization, prices, spending, and out-of-pocket spending of health care services when a full set of claims is not available. Services that have outstanding claims may have a missing or incomplete record. Completion allows for the estimation of the cost impact of the outstanding claims to avoid undercounting or under-projecting trends.

Completion factors varied by type of measure (i.e., spending, utilization, out-of-pocket spending) and service category (i.e., inpatient, outpatient, professional, prescription drug). The factors were based on historical claims payment patterns specific to the HCCI dataset. Annual completion factors were calculated by measure and service and applied to aggregate estimates.

For the *2019 Health Care Cost and Utilization Report*, claims incurred from January 1, 2019 through December 31, 2019 and paid through June 30, 2020 were collected. An adjustment was needed to account for the changes to 2019 claims that would occur after June 30, 2019. Claims from 2015-2018 were assumed to be fully adjudicated.

2.3.2 Population weighting methodology

We weighted spending, utilization, and enrollment using ESI weights to develop estimates that were representative of the national ESI population younger than 65. ESI weights were calculated using the American Community Survey (ACS) 2019 5-year estimates Public Use Microdata Sample (PUMS).

Data Processing Steps

1) We subset the raw ACS data to records with private health insurance coverage and generated age band codes (=1 if AGE between 0 and 17, =2 if AGE between 18 and 24, =3 if AGE between 25 and 34, =4 if AGE between 35 and 44, =5 if AGE between 45 and 54, =6 if AGE between 55 and 64, and =7 if age >= 65).
    a. We then collapsed the records by age band code, sex, PUMA code and state code for total ESI enrollees. This enabled the development of weights using the survey-based targets. 27,417 age-band, sex, PUMA code groups were created per calendar year with corresponding weights.

2) From the HCCI enrollment data, we collapsed the records by calendar year, age band code, sex, and zip code for total enrollees and total enrollees' member months.

3) We used geographic crosswalk from the Missouri Census DataCenter at the University of Missouri to merge data from ACS ESI (step 1) to HCCI 2.0 ESI (step2) by zip codes/ZCTA and PUMA-state code. Note, standard residential zip codes and ZCTAs are identical.

4) We then calculated the ESI weights by age band code, sex, PUMA-state and calendar year using the following formula:

$$ESI\ Weights_{y,s,g,a} = \frac{Number\ of ESI\ enrollees\ from\ ACS_{y,s,g,a}}{Number\ of\ ESI\ enrollees\ from\ HCCI\ 2.0_{y,s,g,a}}$$

Where:

- *y denotes calendar year*
- *s denotes sex*
- *g denotes PUMA-state geography*
- *a denotes age band code.*

5) ESI Weights were capped at the 95th percentile of the weights in each calendar year so as to not overly weight a given age-band, sex, PUMA code group with few enrollees in the HCCI database.

   a. Capping ESI weights at the 95th percentile affected 5.3% of the age-bad, sex, PUMA code groups containing 6.1% of the total ESI population and 0.5% of the HCCI enrollment.
   b. Sensitivity analysis was performed using the 90th and 99th percentiles.
6) ESI Weights were applied by multiplying the spending, utilization, and member months in each age-band, sex, PUMA code group by the corresponding weights. Weighted spending, utilization, and member months were then summed to create weighted national estimates.
7) A separate set of weights for prescription drug data were created following the same method with the restriction on the HCCI enrollment that members have prescription drug benefits.

HEALTH CARE
COST INSTITUTE

3. Analysis

The analytic dataset contains estimates of the key measures of the *2019 Health Care Cost and Utilization Report*--spending, out-of-pocket spending, utilization, and prices--for people younger than 65 and covered by ESI. The statistics were weighted by geography-age-sex to be nationally representative as described in Section 2.3.2 of this document.

All estimates presented in the report can be found in the *2019 Health Care Cost and Utilization Report Downloadable Dataset*.

3.1 Key Measures

Key Measures and corresponding figures in the *2019 Health Care Cost and Utilization Report*:

- Spending per person (Figures 1-9)
- Out-of-pocket spending per person (Figures 9-11)
- Utilization per 1,000 people (Figures 2, 4, 7, and 8)
- Average price per service (Figures 2, 4, 7, and 8)
- Service mix (Figure 8)

3.1.1 Population membership

Membership in the ESI population is calculated using the total number of months individuals are insured. The average number of people with ESI are calculated using total member months divided by 12, to estimate 12 months of coverage or the cost for a year of health care. This method of estimation counts two people with 6 months of coverage the same as one person with 12 months of insurance coverage.

HEALTH CARE
COST INSTITUTE

3.1.2 Spending per Person

Per capita health care spending on people with ESI is calculated by summing in each year all the dollars directly spent on health care services for filed claims and dividing that amount by the average number of people with ESI (total months of ESI coverage divided by 12). Both dollars and people are weighted to be representative of the national ESI population (Section 2.3.2). By this method, the per person spending in the report estimates the cost per person, even for people who did not use health care services.[iii] This metric is a subset of overall national health care spending and may not be comparable to other metrics of national spending because it covers only persons having group ESI and younger than 65 years.

Similar methods were used to calculate out-of-pocket spending per person (the dollars paid by members for health services through copayments, co-insurance, and deductibles).

3.1.3 Utilization per 1,000 people

In the annual *Health Care Cost and Utilization Report*s, HCCI calculated utilization rates per 1,000 insured individuals. The utilization measure was produced by summing for each service category the admissions, procedures, and filled prescription days. The resulting amount was divided by the average number of people with ESI. This provided a per-person utilization count by service category, which was then multiplied by 1,000. Total utilization reflects the spending weighted sum of inpatient, outpatient, professional procedures, and prescription utilization. We use the average spending share in the previous year as the spending weight.

To determine the utilization count, reimbursements for claims were analyzed. In the following rules, *reimbursement* refers to any monetary payment to a provider, whether a professional procedure provider, facility, or pharmaceutical vendor.

- If the reimbursement dollars for an admission, visit, or procedure were equal to 0, the utilization count was set at 0.

- If the reimbursement dollars for an admission, visit, or procedure were less than 0, the utilization count was set at minus 1. Negative reimbursement amounts occur from claim reversals, making it important to reverse the utilization count as well.
- If the reimbursement dollars for an admission, visit, or procedure were greater than 0, the utilization count was set at 1.

Service category-specific rules are as follows:

- Inpatient facility: acute, SNF, and hospice
  - If multiple claims had the same patient identification, DRG, and provider with overlapping or contiguous admission or discharge dates, they were grouped into one admission.
- Outpatient facility
  - If multiple claims and/or claim lines had the same patient identification, CPT code, and service dates they were grouped into one procedure.
- Professional services
  - If multiple claims and/or claim lines had the same patient identification, CPT code, and service dates they were grouped into one procedure.
- Prescriptions

  Prescription drug utilization counts were the number of filled days of a prescription dispensed by retail, mail-order, and specialty pharmacies for prescription drugs and devices. This provides a standard unit, since differing classes of scripts may be for different lengths of time, which could obscure changes in prescription utilization if the number of filled scripts was instead counted. For example, one month of birth control

is 28 filled days, while a round of antibiotics might be 14 filled days.

3.1.4 Price

In the annual *Health Care Cost and Utilization Reports*, HCCI calculated prices as the average price per service by dividing total spending by total utilization per service or subservice category. By this method, the derived calculation includes the "prices" paid by the payer and the patient out of pocket.

3.2 Service Mix

Trends in spending are attributable to shifts in the prices and use of services, as well as changes in the types of services provided. Take, for example, the hypothetical scenario below of a given provider who performs two types of chest X-rays in their clinic: a less expensive 2-view X-ray and a more expensive 4-view X-ray with computer aid detection. In this hypothetical, X-ray prices and the total volume of X-rays performed can both remain constant over time but spending on X-rays still increased by $3,000 due to shifts in the types of services used.

	2015			2019		
Service	Price	Use	Spend	Price	Use	Spend
2-view Chest X-ray	$200	35	$7,000	$200	20	$4,000
4-view Chest X-ray (CAD)	$400	15	$6,000	$400	30	$12,000
Total		50	$13,000		50	$16,000

To estimate the impact of service mix composition on spending, we calculated a service mix factor, which measures how spending would be different in a given year if people utilized the same services in the same proportions relative to the initial year of the study.

A service mix factor was calculated separately for each year during the study period (comparing each year from 2016 to 2019 to the base year, 2015), and for each service

category (inpatient, outpatient, professional services, and prescription drugs). The calculation involved two key steps. First, we calculated the change in spending attributable to the change in utilization for a given service (DRG, CPT code, or NDC code), in each year. In other words, what portion of 2019 spending on a given service was due to the change in the use of that service.

For each service (s), we multiplied its price in a given year (t) by the difference in utilization per capita in year t minus the utilization per capita in the initial year (2015). If a service was no longer provided, we derived an imputed price by inflating (or deflating) price from when the service most recently (or first) appeared.

$$\text{Cost of Change in Service Mix}_{s,t} = \text{Price}_{s,t} \cdot (\text{Use per capita}_{s,t} - \text{Use per capita}_{s,2015})$$

In the hypothetical example above, the cost of change in service mix for the 2-veiw chest x-ray would be calculated as follows:

- $\text{Price}_{\text{2-viewxray,2019}}$=200
- Use per $\text{capita}_{\text{2-viewxray,2019}}$=20
- Use per $\text{capita}_{\text{2-viewxray,2015}}$=35
- Therefore, Cost of Change in Service $\text{Mix}_{\text{2-viewxray,2019}}$=200(20-35) = -$3,000

The cost of change in service mix for the 4-view chest X-ray with computer aid detection would be $6,000.

Second, we computed adjusted spending in each year and service category. To do this, we calculated a sum of the cost of the change in service mix on a service (DRG, CPT code, or NDC code) across all services in a given service category, in each year. Continuing with the hypothetical example above, if the two X-ray procedures were considered their own category of service, the total cost of change due to service mix would be $3,000 (-$3,000+$6,000).

When summed across all services, this measure can be interpreted as the difference in spending due to a change in use, weighted by the price (here used as a proxy for intensity) of each service. In our hypothetical x-ray examples, as both the prices and

total volume of services utilized did not change over time in our example, the shift in service use accounts for the entirety of the growth in spending.

This sum was subtracted from total unadjusted spending for that service category and year, resulting in an adjusted spending metric. In the X-ray example, adjusted spending = $13,000 (unadjusted spending of $16,000 minus the total cost of change due to service mix of $3,000).

Within each year (*t*) and service category (*c*), adjusted spending was divided by actual spending to create the service mix factor:

$$\text{Service Mix Factor}_{c,t} = \frac{\text{Adjusted Total Spending}_{c,t}}{\text{Unadjusted Total Spending}_{c,t}}$$

Continuing with our hypothetical X-ray example, the service mix factor would equate to 0.81 (adjusted spending of $13,000 divided by unadjusted spending of $16,000).

The service mix factor, when applied to spending, price, or use, accounts for the impact of the change in composition of services provided from each measure. For instance, if more expensive outpatient services were utilized year-over-year, service mix would make up for the larger share of higher-cost services in outpatient spending. Alternatively, if the use of generic prescription drugs increased over time relative to brand drugs, service mix would offset any decrease in spending that resulted from a shift towards lower price services (NDC codes for prescription drugs).

3.3 Spending growth decomposition, 2015-2019

To measure the contribution of different factors on cumulative per person spending growth from 2015 to 2019, all dollar amounts were first converted to 2019 dollars using CPI-U to account for inflation. Next, the change in average spending per person was decomposed into four components: service price, quantity of services, changes in the mix of services (service mix), and demographics (age band and sex composition). We performed separate calculations for each year and sex-age band group. Age bands

were defined as younger than 17 years of age, 18–24, 25–34, 35-44, 45-54, and 55–64 (Individuals older than age 64 were excluded. This value was then multiplied by the proportion of the total population in each age band-sex group and summed. Spending growth decomposition can be described in the following illustrative formula:

$$PCSpend_t - PCSpend_{t-1} = Quantity\ of\ Services + Demographics\ mix + Price\ mix + Service\ mix$$

Where average spending per member, $PCSpend_t$, is defined as sum of spending per person (defined in section 3.1.2) for all service categories of interest.

Quantity of Services: To account for spending growth attributed to quantity of services, we used the following formula:

$$Quantity\ of\ Services = \frac{(Util_{t,c,a} - Util_{t-1,c,a})}{1000} * (Share\ of\ ESI\ population_{t-1,a}) * Real\ Price\ in\ 2019\ dollars_{t,c,a}$$

Where:

- *t denotes calendar year*
- *c denotes service category*
- *a denotes age band-sex*
- *Util denotes weighted utilization per 1,000 (section 3.1.3)*
- *Real Price in 2019 dollars defined as price per service (section 3.1.4) inflated to 2019 dollars.*
- *Share of ESI population (section 3.1.1)*

Demographics mix: To account for spending growth attributed to change in demographics mix, we used the following formula:

$$Demographics = \left(Share\ of\ ESI\ population_{t,a} - Share\ of\ ESI\ population_{t-1,a}\right) * Real\ Price\ in\ 2019\ dollars_{t-1,c,a} * \frac{Util_{t-1,c,a}}{1000}$$

Where:

- *t denotes calendar year*
- *c denotes service category*
- *a denotes age band-sex*
- *Util denotes weighted utilization per 1,000 (section 3.1.3)*
- *Real Price in 2019 dollars defined as price per service (section 3.1.4) inflated to 2019 dollars.*
- *Share of ESI population (section 3.1.1)*

Price: To account for spending growth attributed to change in prices, we used the following formula:

$$Price\ mix = (Real\ Price\ in\ 2019\ dollars_{t,c,a} - Real\ Price\ in\ 2019\ dollars_{t-1,c,a}) * \frac{Util_{t-1,c,a}}{1000} * \left(Share\ of\ ESI\ population_{t-1,a}\right)$$

Where:

- *t denotes calendar year*
- *c denotes service category*
- *a denotes age band-sex*
- *Util denotes weighted utilization per 1,000 (section 3.1.3)*
- *Real Price in 2019 dollars defined as price per service (section 3.1.4) inflated to 2019 dollars.*
- *Share of ESI population (section 3.1.1)*

Service Mix: To account for spending growth attributed to change in service mix, we used the following formula:

$$Service\ mix = (Service\ Mix\ Spend_{t,c,a} - Service\ Mix\ Spend_{t-1,c,a}) * (Share\ of\ ESI\ population_{t-1,a})$$

Where:

- *t denotes calendar year*
- *c denotes service category*
- *a denotes age band-sex*
- *Share of ESI population (section 3.1.1)*
- *Service Mix spend$_{t,c}$ = Real spending per member in 2019 dollars$_{t,c}$ - Adjusted spending per member$_{t,c}$*
    - *Adjusted spending per member$_{t}$= Real spending per member$_{t,c}$ * Service mix weight$_{t,c}$*
        - *Service mix weight$_{t,c}$ (section 3.2)*

HEALTH CARE
COST INSTITUTE

Notes

[i] Congressional Budget Office. Private Health Insurance Premiums and Federal Policy. February 11, 2016. Available from: https://www.cbo.gov/publication/51130.

[ii] Kaiser Family Foundation and Health Research & Educational Trust, "2018 Employer Health Benefits Survey." https://www.kff.org/health-costs/report/2018-employer-health-benefits-survey/.

[iii] To calculate total prices paid for total expenditures per capita, the insured (co-payments, coinsurance, and deductibles) and payer expenditures per capita are summed. For inpatient, outpatient, and professional claims, prices paid are calculated for all members who have medical insurance. For prescription claims, prices paid are calculated for all members with medical and prescription insurance.

[10] Calculated using data from the Center for Health Information Analysis, State of Massachusetts, "Annual Report on the Performance of the Massachusetts Health Care System" for 2017 through 2019. Available from: http://www.chiamass.gov/annual-report/?_ga=2.88182555.1443192041.1576609192-1511595466.1576609192

[11] Altarum, "The Impact of Prescription Drug Rebates on Health Plans and Consumers," April 2018. Available at: https://altarum.org/sites/default/files/Altarum-Prescription-Drug-Rebate-Report_April-2018.pdf.

[12] Department of Managed Health Care, State of California, "Prescription Drug Cost Transparency Report (SB 17), Measurement Year 2017." Available at: https://www.dmhc.ca.gov/Portals/0/Docs/DO/SB17Report.pdf.