The Big Three Medical Price Indexes: A Comparative Review and Analysis*

Robert Hovakimyan

May 2021

Abstract

This paper attempts to carefully and clearly describe the data and methodology behind the Big Three medical price indexes (the Medical Consumer Price Index (MCPI), the Disease Based Price Index (DBPI), and the Medical Care Expenditure Index (MCE)) in an effort to understand why they tell different stories of medical price inflation. I discuss the impact of hypothetical and practical scenarios on the differences between the three indexes and pose some hypotheses to explain these differences. The MCPI and DBPI were rising together until around 2008 when the DBPI slowed down, arguably due to utilization shifts. The stark difference between the MCE and other indexes is driven by the price per encounter rising faster than the price per service. While no hypotheses are actually tested in this paper, its insights and comparisons serve as a jumping off point for doing so.

1 Introduction

Since 2009, healthcare spending has contributed to about 17-18% of the United States' Gross Domestic Product (GDP) – almost quadruple its share in 1960[1]. This fact alone suggests that accurately measuring healthcare expenditure is a prerequisite for accurately measuring all expenditure (GDP).

*I am grateful to Emi Nakamura for her time and patience with me. I would not be here without her support.

[1]CMS.gov

Healthcare expenditure has been rising since 1960, in both absolute and per-capita terms. Americans spent $27 billion on healthcare in 1960 ($146 per capita) and $3,795.4 billion on healthcare in 2019 ($11,582 per capita), implying that Americans spent 80 times more per capita in 2019 than they did in 1960 on healthcare *in nominal terms*. Is this a valid comparison? The widely cited consumer price index (CPI), published by the Bureau of Labor Statistics (BLS), suggests that prices were 9 times higher in 2019 than they were in 1960 on average (meaning a dollar in 1960 got you 9 times more stuff than that same dollar in 2019). Taking this into account, 2019 healthcare expenditure was $421.7 billion 1960 dollars ($1,284 per capita), so Americans spent closer to 8.8 times more on healthcare per capita in 2019 than 1960, not 80.

However, prices across vastly different sectors do not change by the same amount from year to year. To deflate healthcare expenditure, inflation for healthcare goods and services should be used. Cue medical price indexes! The BLS publishes a Medical CPI which attempts to track healthcare inflation[2]. The MCPI allows us to compare 2019 expenditure in 1960 healthcare dollars with 1960 expenditure in 1960 healthcare dollars, giving us a better sense of how much more medical *stuff* was actually purchased (and therefore produced). According to the MCPI, medical prices increased 23 fold between 1960 and 2019, suggesting even less medical stuff was actually produced than the overall CPI has us believe.

Nominal healthcare expenditure can be broken down into two components: quantity and price. Medical price indexes allow us to understand which portion of the rise in expenditure is attributable to price growth and thus which portion is attributable to quantity growth. When it comes to healthcare expenditure, quantity growth can be thought of as a combination of disease prevalence growth[3] and per capita output growth. As diseases become more prevalent, we expect to see more healthcare consumed *to treat the same diseases of the same severity with the same quality of healthcare services*. But the MCPI does not track the prices of services used to treat the same disease of the same severity, giving rise to alternative price indexes which attempt to more accurately measure true healthcare price inflation and decompose expenditure growth into more detailed factors.

[2]The MCPI time series can be found here.

[3]Prevalence is usually defined as the percentage of the population with the disease, so population growth is also a factor here.

These indexes are known as disease-based indexes. Instead of holding a basket of medical goods and services fixed, they hold the diseases fixed and allow for a mix in the goods and services used to treat them. Both the BLS and Bureau of Economic Analysis (BEA) have acted on recommendations from the literature and constructed their own indexes (Roehrig (2017) provides an excellent summary). I dub the BLS's Disease Based Price Index (DBPI), the BEA's Medical Care Expenditure Index (MCE), and the Medical Consumer Price Index (MCPI) the "Big Three" medical price indexes and this paper reviews them in isolation before comparing their differences both on paper and in practice.

1.1 Lay of the Land

Sections 2 through 4 discuss the Big Three medical price indexes in turn, formatted as follows: Subsection 1 lays out the purpose or goal of the index, Subsection 2 discusses from where and how the data is gathered, Subsection 3 explains the methodology behind the calculation of the index, and Subsection 4 plots the index and points out noteworthy trends.

Section 5 emphasizes key conceptual differences between the indexes and shows how they materialize in comparative scenarios. Section 6 lists hypotheses for the observed differences between the indexes and discusses how they might be tested. Finally, Section 7 concludes.

2 BLS' Medical CPI (MCPI)

This section's information comes from the BLS' MCPI fact sheet[4] and the latest CPI Handbook of Methods.

2.1 Purpose

Price indexes generally come in two flavors: cost of goods indexes (COGI) and cost of living indexes (COLI). A COGI answers the question "how do the prices of specific goods (and services) change over time?" while a COLI answers the question "how does the price of maintaining the same standard

[4] Found here

of living change over time?". Answering the second question is rightly considered much more important and relevant, so the consumer price index (CPI) aims to *approximate* a COLI[5].

The CPI is split into eight major groups, one of which is medical care (two other major groups, for example, are housing and transportation), which are then aggregated to create the influential overall CPI. The MCPI is the medical care portion of the overall CPI. It's purpose is to construct an approximation of a COLI for medical care by "track[ing] retail prices of good[s] and service[s] of constant quality and quantity over time" via "out of pocket household spending" on medical care services and commodities.

While the CPI is not mathematically explicit about what the ideal COLI for medical care looks like, a COLI for medical care should answer the question "how does the price of obtaining the same amount of health change over time?". As the BLS admits, the MCPI is trying to approximate an answer to this question rather than answer it directly.

2.2 Data

Most price indexes use two core pieces of data: price data and weight data, which are often gathered from some expenditure survey–the MCPI is no different. Using census data, the BLS first determines from which geographic locations (called areas) they will collect data[6]. The Consumer Expenditure Survey (CE)[7] is then used to determine which retailers (called outlets) the BLS' data collectors will visit to collect price and weight data on specific goods and services (called items). After this process, the BLS is left with price and weight data on item-area combinations (one such combination might be physicians services in the Chicago area). Figure 1 (taken from the BLS) lists all 10 items that are included in the MCPI alongside their relative importance to the overall MCPI.

The CE gathers out-of-pocket medical expenditure data annually for the MCPI weights. The MCPI updates these every two years (for example, 2013 and 2014 weights were used for the index in 2015 through 2017). Price information is collected by the data collectors on a monthly basis. While

[5]It is an approximation because "Cost of living is affected by many things not captured in market transactions, and the cost of achieving a living standard cannot be observed directly" (BLS, 2020)

[6]More information can be found here.

[7]More information here.

Figure 1: Definitions of published medical care indexes and relative importance as of December 2019.

Item	Definition	Relative importance (percent)	Percentage of the Medical Care Index
Medical care	Medical care commodities and medical care services	8.833	100%
A. Medical care commodities	Prescription drugs, nonprescription over-the-counter-drugs, and other medical equipment and supplies	1.643	19%
1. Medicinal drugs	All prescription and over-the-counter drugs	1.569	18%
a. Prescription drugs	All drugs dispensed by prescription. Mail order outlets are included. Prices reported represent transaction prices between the pharmacy, patient, and third party payer, if applicable.	1.184	13%
b. Nonprescription drugs	All nonprescription drugs, including topicals	0.385	4%
2. Medical equipment and supplies	Nonprescription medicines and dressings used externally, contraceptives, and supportive and convalescent medical equipment (e.g., adhesive strips, heating pads, athletic supporters, and wheelchairs)	0.074	1%
B. Medical care services	Professional medical services, hospital services, nursing home services, adult day care, and health insurance	7.190	81%
1. Professional services	Physicians, dentists, eye care providers, and other medical professionals	3.643	41%
a. Physicians' services	Services by medical physicians in private practice, including osteopaths, which are billed by the physician. Includes house, office, clinic, and hospital visits. (Excludes independent lab work and ophthalmologists. See Eyeglasses and eye care.)	1.811	21%
b. Dental services	Services performed by dentists, oral or maxillofacial surgeons, orthodontists, periodontists, or other dental specialists in group or individual practice. Treatment may be provided in the office or hospital.	0.990	11%
c. Eyeglasses and eye care	Services and goods provided by opticians, optometrists, and ophthalmologists. Includes eye exams, dispensing of eyeglasses and contact lenses, office visits, and surgical procedures in the office or hospital.	0.369	4%
d. Services by other medical professionals	Services performed by other professionals such as psychologists, chiropractors, physical therapists, podiatrists, social workers, and nurse practitioners in or out of the office. Also, includes independent lab work and imaging services.	0.474	5%
2. Hospital and related services	Services provided to inpatients and outpatients. Includes emergency room visits, nursing home care and adult day care.	2.378	27%
a. Hospital services	Services provided to patients during visits to hospitals, ambulatory surgical centers, or other similar settings.	2.186	25%
i. Inpatient hospital services (1)	Services for inpatients. Includes a mixture of itemized services, Diagnosis Related Group -based services, per diems, packages, or other bundled services.	N/A	
ii. Outpatient hospital services (1)	Services provided to patients classified as outpatients in hospitals, free standing services facilities, ambulatory surgery, and urgent care centers.	N/A	
b. Nursing home and adult day care services	Charges for residential care at nursing homes, nursing home units of retirement homes, and convalescent or rest homes. Also includes non-residential adult day care.	0.122	1%
c. Care of invalids, elderly and convalescents in the home	Fees paid to individuals or agencies for the personal care of invalids, elderly or convalescents in the home including food preparation, bathing, light house cleaning, and other services	0.069	1%
3. Health Insurance	Indirect approach based on retained earnings method. See Health Insurance section.	1.170	13%

Footnotes:
(1) Substratum index: a special index published below the typical item level. Relative importance is not available for these indexes.
(2) CPI pricing and weighting excludes institutionalized populations such as those living in nursing homes.

Source: BLS MCPI Factsheet

only out-of-pocket expenditure is used to determine weights, prices reflect both out-of-pocket payments and third party (i.e. insurance) payments.

2.3 Methodology

The MCPI is constructed in two stages. First, basic indexes are constructed for each item-area combination. For the 9 non-insurance items, traditional methods are used, while the health insurance index is constructed differently. Second, these item-area indexes are combined into an aggregate MCPI using a modified Laspeyres formula. See Figure 1 for a refresher of the 10 medical care items and what percentage of the aggregate MCPI they make up.

2.3.1 Non-Insurance Item Indexes

Most of the general CPI item indexes are calculated using a geometric mean formula, but the 9 non-insurance medical item indexes are different. Prescription drugs (since 2016), physician's services, hospital services, dental services, services by other medical professionals, and nursing homes/adult day care all use Laspeyres formulas.

A geometric mean index for an item-area combination in month t relative to base month b looks like this:

$$I_{b,t}^{i,a} = \Pi_{n\in(i,a)}(\frac{P_{n,t}}{P_{n,b}})^{\frac{W_{n,b}}{\sum_{n\in(i,a)} W_{n,b}}}$$

And a Laspeyres index for an item-area combination in month t:

$$I_{b,t}^{i,a} = \frac{\sum_{n\in(i,a)} \frac{W_{n,b}}{P_{n,b}} P_{n,t}}{\sum_{n\in(i,a)} \frac{W_{n,b}}{P_{n,b}} P_{n,b}}$$

In the above equations, there are some number N items within each item-area combination whose prices, P, are used in both the current month, t, and base year, b. The base period weights for each specific item, $W_{n,b}$, are constructed using the relative proportion of expenditures on that item as observed in the CE[8]. Therefore, these weights divided by prices can be thought of as quantities, making the Laspeyres formulation look more "textbook".

[8]The weights are also a factor of duplication, geography, nonresponse adjustments, estimated total daily expenditures, and more. See here.

It is crucial to note that only base period expenditure weights are taken into consideration. Imagine that patients in the Bay Area substitute away from a dental service which has become more expensive since the base period towards another one of the same quality which has become cheaper. This substitution has lowered the cost of obtaining the same level of dental care between the two periods, yet the Laspeyres $I_t^{DentalServices,BayArea}$ captures none of this substitution effect, instead the index in month t rises relative to the base period, indicating a rise in the cost of living. This is a poor approximation of a COLI (as the true COLI has arguably not risen, while the index has).

Although a geometric mean also uses base period expenditure weights, it does not hold the quantity fixed like the Laspeyres. Instead, it holds the share of total dental expenditure on each specific dental service constant. So, if a service becomes twice as expensive and patients buy half as much of it, the share of their total expenditure on that service would stay the same and the geometric formula would perfectly capture the substitution. However, the demand for medical services is widely accepted to be very inelastic[9] so using a geometric mean formulation probably overstates substitution effects in this sector and thus a Laspeyres (which ignores them entirely) is seen as more fitting for most medical service indexes.

These item-area indexes are then aggregated over all the geographical areas and items to create the overall MCPI. Regardless of how the item-area index is constructed (geometric mean or Laspeyres or another method), the aggregation from item-area index to MCPI is Laspeyres. But before discussing aggregation, there is a special medical care item index we have left out.

2.3.2 Health Insurance Index

As aforementioned, the MCPI wants to track goods of constant quantity and quality over time. One challenging price to keep quality-constant is that of health insurance. Simply tracking the prices of premiums month to month would lead to misleading results as the benefits and risks of any given plan are constantly in flux and there are certainly quality differences between different insurance plans not captured entirely by differences in premiums.

[9]In a case more extreme than dental services: It is unlikely that patients would buy half as many heart surgeries if the price of heart surgeries doubled. They would buy less, perhaps, but probably by a very little amount.

Explicitly adjusting for differences in quality and changes in these benefits and risks is very costly and near impossible, so the BLS has developed an indirect method called the "Retained Earnings Method". In short, this method splits the insurance premiums into "benefits paid out" (BPO), which are the medical care reimbursements by the insurance plan, and "retained earnings" (RE), which is the premiums minus benefits paid out. The BLS views RE as the cost of providing medical care reimbursements and thus assumes that if BPO increases while RE stays the same, or if BPO stays constant while RE decreases, the quality of insurance has increased relative to its cost. In other words, the change in the ratio of RE:BPO, called the Retention Benefits Ratio (RBR), indirectly determines quality change. A decrease in the RBR indicates an increase in quality, and an increase in the RBR indicates a decrease in quality. Using annual RE and BPO data, the BLS captures month to month changes (or inflation) in the RBR, called the Retained Earnings Relative (RER).

$$RER_t = \frac{RBR_t}{RBR_{t-1}}$$

For each non-insurance medical item, i, a Health Insurance Product (HIP), is created for month t. $I_{i,t}$ is simply $I_{b,t}^{i,a}$ aggregated over all areas.

$$HIP_{i,t} = RER_t * \frac{I_{i,t}}{I_{i,t-1}}$$

These 9 index products are then aggregated into the Health Insurance index for month t (HI_t):

$$HI_t = \frac{\sum_{i=1}^{9} w_{i,t-1} * HIP_{i,t}}{\sum_{i=1}^{9} w_{i,t-1} * HIP_{i,t-1}}$$

Unlike the equations provided for the 9 non-insurance indexes, this index is not between t and a base period, but rather between t and t-1. But the index between t and b could be backed out.

I *think* the weights, $w_{i,t}$, are determined using the following process[10]:

1. The BLS determines what proportion of health insurance premiums go towards benefits paid out. For example, if \$100 in premiums are collected and \$90 of benefits are paid out, then 0.9 of premiums go to

[10] This process is described by the BLS in words here.

BPO and the other 0.1 go to RE. The 0.9 BPO weight will later be redistributed into the 9 non medical insurance indexes, while the 0.1 RE weight will stay in the health insurance index.

2. This 0.1 RE weight is then distributed to each of the 9 index products using some unlisted method which creates a $w_{i,t}$ for each $HIP_{i,t}$.

Finally, the 0.9 BPO weight is then distributed to the 9 non-insurance indexes. This means that the non-insurance indexes are tracking out of pocket and insurance reimbursement price relatives. This process is described in the BLS webpage but without any math explicitly describing what they're doing, it is difficult to understand what this weight distribution process entails.

2.3.3 Aggregate Index Calculation

Once all 320 item-area indexes (9 non-insurance items and 1 insurance item for 32 areas) are constructed, they are aggregated up to an overall MCPI. The only missing piece is the aggregation weights. Each item-area index is assigned a weight as follows:

$$AW_{\beta}^{i,a} = \frac{(\hat{P}_b \hat{Q}_{\beta})^{i,a}}{100}$$

Here, b is the base period of the item-area index as in 1.3.1, while β is the reference period of expenditures (this is the part of the weight that is updated every 2 years as described in 1.2). This weight is a dollar amount made up of an estimated price, $\hat{P}$ and estimated quantity, $\hat{Q}$. The reasons estimated prices and quantities are used are 1) to allow the BLS to use more than just one month of data (rather, they use two years worth), and 2) to keep the item area expenditures smoother by averaging them with item regional expenditures (the 32 areas make up 8 regions). However, this process puts disproportionate weight on larger areas. Since β is updated every two years, the quantities in the weights are also updated at this frequency. However, b for most item-area indexes is 1982-1984, so it is changes in estimated quantities that drive changes in the aggregate weights. Because β is adjusted every two years, there could be a jump in the index from December before the adjustment and January after the adjustment. To circumvent this jump, the last month using old weights (every other December) is called the "pivot" month (labeled v) and the aggregate weights are constructed via this pivot month such that the overall index is identical in the pivot month (every other December) and the month after (January). For example, from July 2017 to December 2019 (the latest pivot month), β was 2015-2016. Starting January 2020, β shifted to 2017-2018.

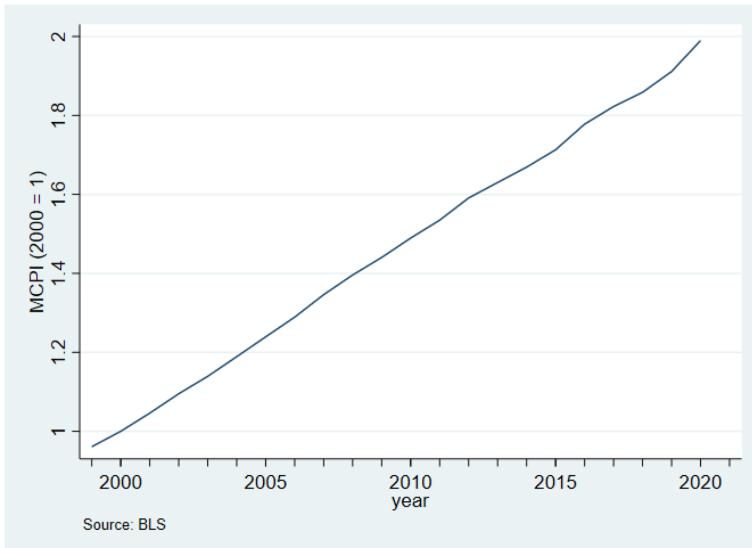
The aggregate index in time t with base period z is constructed as follows:

$$MCPI_{z,t} = \frac{\sum_{i \in I, a \in A} AW_{\beta}^{i,a} * I_{b,t}^{i,a}}{\sum_{i \in I, a \in A} AW_{\beta}^{i,a} * I_{b,v}^{i,a}} * MCPI_{z,v}$$

2.4 Results

The MCPI from 1999-2020 is shown below in Figure 2. According to the MCPI, the approximate cost of obtaining the same amount of health has doubled since the turn of the century. All else equal, consumers' expenditure on healthcare would need to double for them to obtain just as much health in 2020 as they did in 1999. However, all else is not equal. For example, the quality of medical care has hardly been constant. The 21st century saw the advent of artificial organs, the functional MRI, and the cyberknife, to list only a few examples. Some of the MCPI's depicted price increase surely captures this increase in quality. Additionally, the MCPI holds the weights which make up its goods and services constant, meaning some substitutions away from expensive services to cheaper ones will be completely missed and others will only be caught with a two year lag. While quality change is not something the other big medical price indexes address, they do greatly improve on capturing substitution effects. Section 5 expands on both points further.

Figure 2: The BLS's Medical Consumer Price Index (2000 = 1)



3 BLS' Disease Based Price Index (DBPI)

All information in this section is from Bradley et al. 2015, Roehrig 2017, the PINR DBPI webpage[11] and Bradley 2017 [12].

3.1 Purpose

Unlike the MCPI, the DBPI is concerned with "measur[ing] the change in the average price level to treat an episode of a specific disease." Instead of holding the "basket" (or items) of medical goods and services fixed over time, it holds the disease fixed. While the MCPI can be broken down into 10 item indexes, the DBPI can be broken down into 19 disease indexes. The advocates of a disease-based index emphasize that indexing by disease makes it easier to uncover both what diseases are contributing to rising healthcare costs and what is driving this rise in healthcare costs (inflation, prevalence, or utilization).

Like the MCPI, the DBPI also takes a COLI approach. While someone may go to the grocery store to buy just apples and nothing else, its unrealistic

[11] Found here.
[12] See here.

that someone walks into a hospital to buy an X-ray and nothing else. Patients usually purchase medical services with the goal of treating some condition, so they care about the total price of treating that condition and not the isolated price of an X-Ray. The X-Ray alone does not contribute to increases in their standard of living; only when it is bundled with surgery, therapy, and drugs all working in tandem to improve their condition does their standard of living also improve. If a medical COLI is answering the question “how does the price of obtaining the same amount of health change over time?”, it should take into consideration the fact that health is obtained by treating diseases, not by consuming X-Rays. These ideas are further developed in 3.3.

3.2 Data

The BLS draws from Schultze and Mackie’s lengthy 2002 report which recommends that the BLS choose a subset of diagnoses from the International Classification of Diseases (ICD) and match them with existing medical treatment data from large retrospective medical claims. They also suggest that the BLS reprice the items every month and update the quantity weights every year or two. This is precisely what the BLS now does! They reprice the items every month using the CPI and Producer Price Index (PPI) and use the Medical Expenditure Panel Survey (MEPS) as a close substitute for retrospective claims data (although MEPS is far from “large”).

3.2.1 Weights

MEPS is a yearly survey of household, insurance, and Medicaid/Medicare expenditure in healthcare across the US[13]. Households are selected randomly and interviewed over the course of two years. They are asked questions about any diseases contracted and which providers were used in treating those disease(s). Those providers are then surveyed as well. MEPS has many files available in its public database, but the two used are the “events” and “conditions” files.

The “conditions” file links individuals to specific conditions as defined by their 3 digit ICD. The “event” file links individuals to specific events like doctor’s visits and drug purchases, and notes the total payment with a breakdown of sources (out-of-pocket, private insurance, medicare/medicaid,

[13]Learn more about MEPS here.

etc.). By linking the conditions and events to individuals, the DBPI is able to get annual utilizations, which they use as weights. Because MEPS has a 3 year lag, these utilitizations are updated yearly with a 3 year lag. Thus, the weights used for the DBPI in 2016 reflect 2013 utilization data.

3.2.2 Prices

The initial base period price of service k used to treat disease d is taken from MEPS, but all future prices are constructed using CPI/PPI information. More precisely, the price of service k in treating disease d in a given period t is:

$$P_{k,d,t} = P_{k,d,b} * \frac{I_{k,t}}{I_{k,b}}$$

Depending on k, I refers to either the CPI or PPI. If it is the CPI, like in the case of pharmaceuticals, then I is exactly the item index familiar from the previous section (just replace the i subscript with k). $P_{k,d,b}$ is the initial price from MEPS. While MEPS does not actually list prices, it does list total expenditures for a service or item and the quantity purchased, so prices can be backed out.

3.3 Methodology

The methodology behind the DBPI comes explicitly from the goal of constructing a COLI. First, nominal expenditures are decomposed, revealing the price inflation component the DBPI tracks. Next, the cost of living is explicitly computed from an individual's constant elasticity of substitution (CES) utility function, uncovering the COLI under specific conditions. Then, these specific conditions are argued to not hold, which gives rise to an alternate method drawing on recommendations from Schultze and Mackie (2002).

3.3.1 Nominal Expenditure Decomposition

Bradley et al. (2015) decompose nominal expenditure in treating disease d during time t, $E_{d,t}$, into a few key components: Price (P), population (Pop), prevalence rate (r), and output per patient (Q).

$$E_{d,t} = P_{d,t} * Pop_t * r_{d,t} * Q_{d,t}$$

Then, expenditure growth for disease d between two periods, say t and $t-1$, is a combination of price inflation, population growth, prevalence rate growth, and output per patient growth. However, there is no reason why the quality of treatment should stay the same between the two periods. Bradley et al. introduce the concept of quality change from a COLI perspective. Let $h_{d,t}$ be the health outcome of treating the disease in period t. This could be different from $h_{d,t-1}$. More importantly, let $v(h_{d,t})$ be the value consumers place on this health outcome. Decomposing nominal expenditure growth now looks like this:

$$\frac{E_{d,t}}{E_{d,t-1}} = \frac{P_{d,t}v(h_{d,t-1}) * Pop_t * r_{d,t} * Q_{d,t}v(h_{d,t})}{P_{d,t-1}v(h_{d,t1}) * Pop_{-1} * r_{d,t-1} * Q_{d,t-1}v(h_{d,t-1})}$$

In this decomposition, the right hand side is multiplied by $v(h_{d,t-1}) * v(h_{d,t})/v(h_{d,t}) * v(h_{d,t-1}) = 1$, so the same definition of nominal expenditure is being used. But under this decomposition, the changes in quality (or the value consumers place on the health outcome of treating disease d in period t) will not be conflated with changes in price or quantity. Controlling for quality in this way will lead to capturing the "true" price inflation, which is the goal of these medical price indexes. There have been several methods proposed for measuring or proxying $v(h_{d,t})$, but the DBPI is not concerned with directly measuring this yet; they pose it as a "future goal".

3.3.2 Demand Side COLI Approach

This subsection is concerned with an *individual's* demand for medical care products used to treat a disease. Aggregates will be considered in the next subsection.

If all consumption in a period t is grouped into medical consumption, $H_{d,t}$, defined by the acquisition of some "health stock" to treat disease d, and non-medical consumption, c_t, we might imagine the following utility function for an individual with constant elasticity of substitution (CES) between medical consumption and non-medical consumption:

$$U(c_t, H_{d,t}) = [(a_c c_t)^{1/\rho} + (a_H H_{d,t})^{1/\rho}]^{1/\rho}$$

A CES function is used here for two purposes: 1) To go from an item index to an overall index (like the MCPI does and DBPI will do shortly), the utility function needs to be separable (which a CES is) and 2) A CES form is easy to work with. While it does not capture the many nuances of real life, it may make more sense in the case of medical goods (where substituting away from medical consumption is very difficult regardless of where you are on the isoquant) than in other contexts.

The COLI *for an individual* can then be derived from the associated expenditure functions in period t and base period b. Here, σ, the elasticity of substitution, equals $\frac{-1}{\rho-1}$, $P_{c,t}$ is the price of non-medical consumption in time t, and $P_{h_{d,t}}$ is the price of extra health stock from treating disease d in time t.

$$COLI_{b,t} = \frac{[(P_{c,t}/a_c)^{1-\sigma} + (P_{h_{d,t}}/a_H)^{1-\sigma}]^{1/1-\sigma}}{[(P_{c,b}/a_c)^{1-\sigma} + (P_{h_{d,b}}/a_H)^{1-\sigma}]^{1/1-\sigma}}$$

While this is what the DBPI would love to track on an individual level, the fundamental issue is that $P_{h_{d,t}}$ does not exist because "health stock" is not bought and sold on any market. Additionally, estimating σ, a_c, and a_H is a difficult task, especially for the aggregate economy. However, thanks to Sato and Vartia, it is possible to measure a COLI without having to estimate these three parameters. So all that needs to be dealt with is the mysterious price of health stock.

Bradley et al. take a creative approach to this problem, which is the source of a fundamental difference between the DBPI and MCPI perspectives. They let $h_{d,t}$ be a function of K inputs of medical goods and services (like prescription drugs, physician's services, and hospital services) captured in the Kx1 vector $z_{d,t}$. The prices of these K inputs are precisely what the MCPI tracks as though they are final goods and services. The DBPI creators, however, treat these as inputs into the final product, so $h_{d,t} = f_{d,t}(z_{d,t})$. Here, $f(.)$ is also assumed to be CES and the health stock production function to treat disease d at time t looks like this:

$$h_{d,t} = f_{d,t}(z_{d,t}) = [\sum_{k=1}^{K}(a_{k,d}z_{k,d,t})^{\gamma}]^{1/\gamma}$$

Each $z_{k,d,t}$ is a specific component (good or service) in the vector of K services. The coefficients on each service, $a_{k,d}$, must be constant over time for the next steps to mathematically follow. A CES assumption is even more reasonable here because medical inputs have substitution elasticises close to a constant 0, meaning they are perfect complements of each other[14]. It then follows that the unit price of obtaining the health stock, $P_{h_{d,t}}$, can be written as:

$$P_{h_{d,t}} = [\sum_{k=1}^{K} (\frac{P_{k,t}}{a_{k,d}})^{1-\omega}]^{1/1-\omega}; \omega = -1/1 - \gamma$$

A Sato Vartia price index can then be computed without knowing γ or $a_{k,d}$. This Index would measure $P_{h_{d,t}}/P_{h_{d,b}}$ allowing us to decompose expenditure growth *for an individual* into price growth and output growth as follows:

$$\frac{E_{d,t}}{E_{d,t-1}} = \frac{P_{h_{d,t}}}{P_{h_{d,t-1}}} * \frac{f_{d,t}(z_{d,t})}{f_{d,t-1}(z_{d,t-1})}$$

However, for this line of reasoning to hold, $a_{k,d}$ in the production function must be constant over time for all k services used to treat disease d. Intuitively, there's no reason why this should be true (especially over longer periods of time). For example, if one of the factors in the production of health stock to treat Alzheimer's, $h_{Alzheimer's,t}$, is prescription drugs, then as prescriptions get more effective at treating the disease (i.e. producing more health stock per drug), $a_{drug,Alzheimer's}$ will increase. Many papers which introduce alternative medical price indexes for specific diseases document an increase in this coefficient for most services used to treat diseases over time. This is one reason why "superlative index theory" (using the Sato-Vartia index to construct a COLI) fails for medical goods.

One might imagine approximating $a_{k,d,t}$ to overcome this problem. For example, let it be a function of some characteristics $c_{k,t}$. The most popular method in index number theory to estimate $a_{k,d,t}(c_{k,t})$ is called hedonics, a regression-based technique which links changes in an item's price with changes in its characteristics[15]. However, hedonics fails here for three reasons: 1) Not all $c_{k,t}$ are observed, 2) For those $c_{k,t}$ which are observed, hedonics does not find a strong link between prices of services and their characteristics

[14]Recall: an X-Ray by itself does no good; health stock is produced only when it is combined with other items like drugs and therapy.

[15]See Aizcorbe's excellent A Practical Guide to Price Index & Hedonic Techniques for more on hedonics.

because price is not the driving factor in medical decision making [16], and 3) since most medical payments are not out-of-pocket, but rather reimbursed by insurance, the price of a medical service is not a reflection of consumers' "willingness to pay", rendering hedonic results are un-interpretable.

Overall, a traditional COLI approach is no good for computing a medical index at the individual level (as we have seen in this subsection). And to compute one at the aggregate level, at minimum $h_{d,t}$ must be measured but we have seen this is not yet possible. This leads Bradley et al to their current formulation.

3.4 The Current DBPI Formulation

As mentioned in Section 3.2, the DBPI follows the recommendation of Schultze & Mackie (2002) for both the sources of their data and the timing of their updates. The current DBPI for disease d at time t relative to a base period b looks like this:

$$DBPI_{d,b,t} = \frac{\sum_{k=1}^{K} P_{k,d,t} z_{k,d,y(t)}}{\sum_{k=1}^{K} P_{k,d,b} z_{k,d,y(b)}}$$

Recall from Section 3.2 how the price component is constructed and that the weights, z, are updated yearly. So $y(t)$ is simply a function which takes in a year-month and outputs the corresponding year. Bradley et al. prove that $DBPI_{d,b,t}$ is a COLI if the following two conditions hold:

1. $h_{d,t} = f_{d,t}(z_{d,t})$ has a Leontieff production function with coefficients $a_{k,d,y(t)}$

   A Leontieff function is simply a CES production function with constant elasticity of substitution equal to 0. As mentioned in Section 3.3, this assumption lies close to the truth. While there is no such thing as a constant zero substitution elasticity in real life, we can view $h_{d,t}$ as approximately Leontieff. More importantly, the coefficient, a, is assumed to only change yearly. While progress in quality and effectiveness is made month to month, we should not be missing much by assuming it changes year to year.

[16] If you ask a doctor or patient what influences their decision to choose one treatment over another, price will not be first (or even second) on that list.

2. $h_{d,t} = h_{d,b} \forall t, b$ and $P_{d,k,t} = P_{k,t} \forall d$

The first half of this assumption requires that the health function's output be the same between the two periods. This must hold for months within the same year where both the coefficient, a, and quantity of inputs, z, are fixed. However, when comparing months more than a year apart, it is difficult to square this assumption with the previous one. If a is allowed to change year to year, then the quantity of inputs, z, must adjust perfectly for the health outcome to stay the same. While it makes sense for less drugs and X-Rays to be consumed as their quality increases, there is no reason why the quantity should adjust perfectly to the quality increase. The second part of this assumption states that the price of a service, k, is not different across diseases. For example, the drug Vicodin should not be more expensive for a patient treating arthritis than for one using it to relieve pain after knee surgery. This assumption has not always held in practice (see Section 5.1).

Bradley et al. point out that one key difference between the DBPI and MCPI methodology is the time subscript of this quantity weight. If the BLS were to construct the MCPI by disease (which they theoretically could do if they gathered disease-level information[17]), their disease based index would look like this:

$$MCPI_{d.b.t} = \frac{\sum_{k=1}^{K} P_{k,d,t} z_{k,d,y(b)}}{\sum_{k=1}^{K} P_{k,d,b} z_{k,d,y(b)}}$$

The quantities, or weights, are not updated yearly in the MCPI methodology to reflect utilization shifts. In the DBPI's case, $y(t)$ and $y(t-1)$ are the same for all months besides January, where $y(t)$ is a year after $y(t-1)$. This causes the index to jump every January, when it reflects price and utilization change instead of just price change. The DBPI comes in many flavors, one of which smooths out this utilization change over all 12 months.

[17]In fact, Bradley constructs an MCPI index using the same data that the DBPI uses to compare the two. See Section 3.5 for the graph.

3.4.1 DBPI Flavors

The DBPI comes in 4 flavors[18]: Smoothed without comorbidities adjustment, Smoothed with comorbidities adjustment, Non-smoothed without comorbidities, and Non-smoothed with comorbidities adjustment. Smoothing refers to smoothing out the aforementioned jump every January. This is done by allocating 1/12th of the yearly quantity change to each month. Comorbidities is a fancy word for a medical event in MEPS, like a doctor's visit, used to treat multiple conditions at once. Not adjusting for comorbidities means viewing the visit as though its multiple separate visits, one for each condition. Adjusting for comorbidities involves assigning some percentage of that visit's weight to each condition (where the percent weights add up to 100%). The DBPI does this using a "pro rationing method". Bradley et al. give the following helpful example: "If the average quantity of office visits to treat heart disease is 3 and the average quantity to treat diabetes is 2, then if an office visit treats both diabetes and heart disease, then 3/5 of the visit is allocated to heart disease and 2/5 is allocated to diabetes." The implications of changes in comorbidity trends will be discussed in Section 5.

3.5 Results

Figure 3 shows the MCPI next to Bradley's "Traditional" Index and the adjusted (for comorbidities) DBPI. The traditional index is constructed using the same data sources as the DBPI (MEPS for weights and CPI/PPI for price information) but with the MCPI methodology (fixed weights, or what's called a Lowe index). The Traditional index is comorbidity adjusted, and the DBPI is smoothed.

The MCPI is less volatile than the other two because the sample size it uses is huge compared to MEPS. However, MEPS' sample size is large enough to render these observed differences statistically significant.

The first interesting difference is between the MCPI and the Traditional index. Keeping methodology constant, simply shifting to using MEPS for weight data and the PPI for some price data lowers the price level. The DBPI uses PPI prices for physician's services and hospital services because they include Medicaid/Medicare while their CPI counterparts do not. Its possible that this price difference is the primary driver of the differences

[18]This is actually a small fib. The DBPI also has a dental seperated option which is not relevant to our discussion and has very small practical effects.

between the indexes since physician services and hospital services account for 46% of the MCPI. However, its more likely that the weights are driving this difference. This begs the question of whether the CE or MEPS is more representative of actual utilization in the country.

The second important difference is between the Traditional Index and the DBPI Adjusted Index. This difference is what sparked the desire to investigate these medical price indexes in the first place. The two indexes are rising at about the same clip from 1999-2007, but in 2007 there is a very visible divergence between the two. In 2008, 2009, and 2010, the DBPI barely increases and even falls in 2011, while the traditional index keeps rising. Section 6.1 lays out hypotheses for why this divergence may have occurred.

Figure 3: MCPI, Traditional, and DBPI Adjusted (2000 = 1)

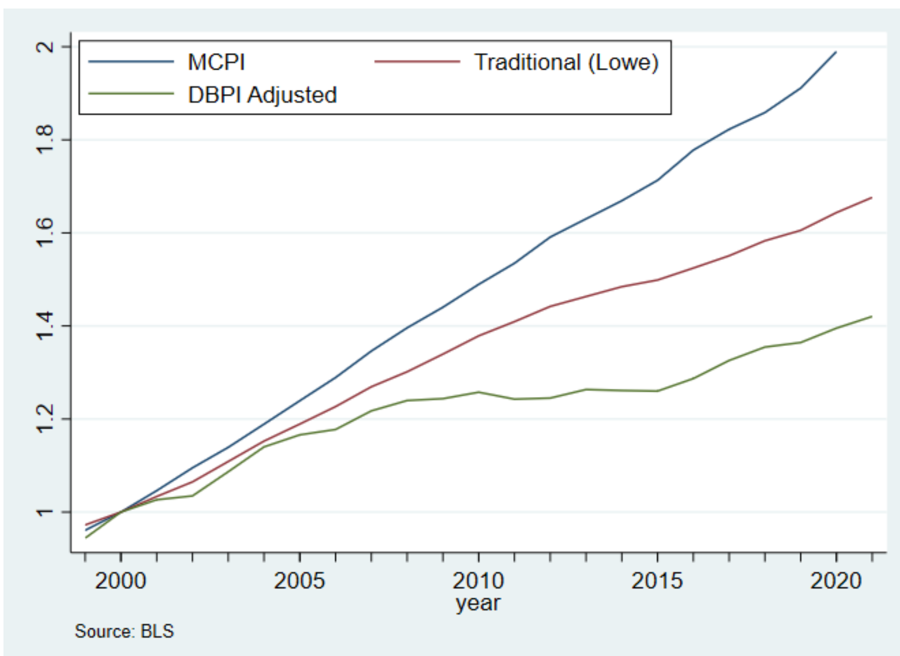
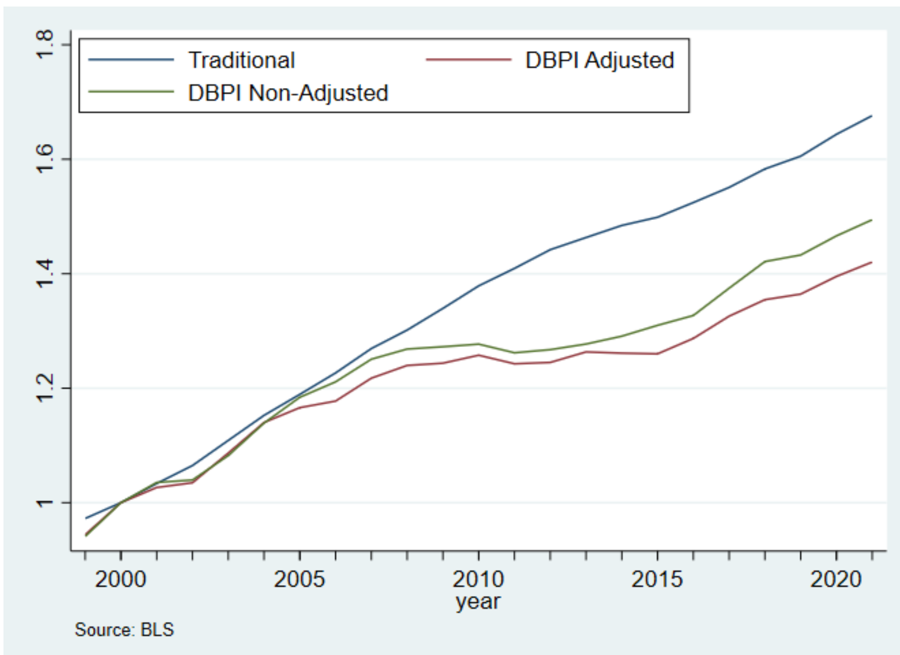


Figure 4 drops the MCPI and adds in the non-comorbidity adjusted DBPI. The difference between the ajdusted and non adjusted DBPI is minor but the gap seems to be increasing over time. This is probably a result of comorbidities increasing over time. The adjusted DBPI fell between 2013 and 2015 while the non-adjusted rose. This suggests the average price to treat a disease increased over this period, but the same visit or encounter was utilized more efficiently by treating multiple diseases in one go, thus lowering the weights in the adjusted index.

Figure 4: The DBPI Indexes and the Traditional Index (2000 = 1)



4 BEA's Medical Care Expenditure Index (MCE)

All the information in this section is from Aizcorbe & Nestoriak (2011) and the BEA's Healthcare Satellite Account website[19].

4.1 Purpose

Unlike both the MCPI and DBPI, the MCE wants to "track the overall cost of care", not the price of individual services, to "better measure spending trends and treatment prices." The BEA believes the MCE is the "first step toward... better assessment of value in healthcare spending". This may seem like a minor difference from the DBPI, but it leads to large differences in practice (as depicted in Section 4.4). Crucially, the MCPI and DBPI use price changes for specific goods and services, while the MCE tracks the overall cost of the encounter (or visit) by disease. The MCE is undoubtedly a COLI; so much so, in fact, that its growth is not yet be decomposed into price and quantity (or utilization) growth.

4.2 Data

The BEA has 2 main healthcare datasets and it constructs an MCE for each. The first dataset comes directly from MEPS, while the second blends MEPS with other claims data. For comparisons between the indexes to be meaningful, the focus here will be on the MEPS MCE, even though the blended MCE is a tighter measure of inflation since it uses more data. The BEA follows the methodology of Aizcorbe & Nestoriak (2011) closely. While the authors do not use MEPS in their paper, they use a similar claims database and it is easy to imagine how their weight and price collection method can be applied to MEPS.

Because the MCE tracks the total cost of care and is literally just a ratio of expenditures between two periods (as 3.3 will make clear), thinking about the index in terms of a weight and price component is a little strange. As aforementioned, the MCE cannot currently be decomposed into these two components. However, identifying a "weight" and "price" component helps us compare it with the other indexes.

[19] Found here.

4.2.1 Weights

Each service k to treat a given disease d is assigned a weight at time t which is equal to the number of service encounters for k divided by the number of cases of d treated: $x_{k,d,t}/N_{d,t}$. In other words, each service is given a weight equal to its share of all services used to treat the disease. Both x and N are readily available in MEPS by aggregating up from individual level claims to nationally representative numbers for a given quarter (e.g. prescription drugs used to treat Alzheimer's in Q1 2018). See 2.2.1 for a refresher on MEPS.

4.2.2 Prices

For each service k used to treat disease d at time t, the nationally representative cost is denoted $c_{k,d,t}$. Imagine that in MEPS, there were 10,000 patients who were treated for Epilepsy in Q1 2018. Say 7,000 of these visits involved an MRI, and the average cost of each MRI encounter was \$2,000 (this includes all payments: out-of-pocket and third party reimbursements). In this example, $d = Epilepsy$, $k = MRI$, $x_{k,d} = 7,000$, $N_d = 10,000$, and $c_{k,d} = \$2,000$.

4.3 Methodology

Aizcorbe & Nestoriak (2011) derive the MCE formulation from a comparison with what they call "Service Price Indexes" (or traditional indexes, as Bradley would put it). The MCE for a disease d is simply the ratio of per case expenditure between a period t and base period b. A "period" in the MCE is a quarter, while the MCPI and DBPI publish monthly. The MCE looks like this:

$$MCE_{d,b,t} = \frac{c_{d,t}}{c_{d,b}} = \frac{\sum_k (c_{k,d,t} x_{k,d,t})/N_{d,t}}{\sum_k (c_{k,d,b} x_{k,d,b})/N_{d,b}}$$

Here, $c_{k,d,t}$ and $x_{k,d,t}$ are grouped because their product is the total expenditure for a service k used to treat d. Summing these over all services gives us total expenditure to treat the disease. Dividing this by the number of cases treated is the per case expenditure. $c_{k,d,t}$ and $x_{k,d,t}$ are grouped because their product is the total expenditure for a service k used to treat d. Summing these over all services gives us total expenditure to treat the disease. Dividing this by the number of cases treated is the per case expenditure.

This is contrasted with what the authors call a Service Price Index (SPI). They claim that the MCPI is very close to an SPI aggregated over all diseases. An SPI looks like this:

$$SPI_{d,b,t} = \frac{\sum_k (c_{k,d,t} x_{k,d,b})/N_{d,b}}{\sum_k (c_{k,d,b} x_{k,d,b})/N_{d,b}}$$

Notice the different subscript for the "weight" term in the numerator. Instead of being a ratio of expenditures to treat disease d between the two periods, the SPI's numerator is the hypothetical cost of services provided to patients in the base period at period t prices. This is a Laspeyres price index like the MCPI. Aizcorbe and Nestoriak show that the difference between the MCE and SPI at the disease-level is driven by utilization changes. Specifically,

$$MCE_d = SPI_d = \sum_k [SPI_{d,k}(dU_{d,k} - 1)]$$

where

$$SPI_{d,k} = \frac{c_{d,k,t} x_{d,k,b}}{\sum_k c_{d,k,t} x_{d,k,b}}$$

and

$$dU_{d,k} = \frac{x_{d,k,t}}{N_{d,t}} / \frac{x_{d,k,b}}{N_{d,b}}$$

Changes in utilization are just changes in the "weight" component of the index. It makes sense that this is the driver of any difference between the SPI and MCE, because that is their only difference. The $SPI_{d,k}$ can be thought of as a weight for each service: the more utilized a specific service is in treating a disease, the larger the differences that changes in the utilization of that service will create. This difference is explored more in Section 5.

4.3.1 Comorbidities

Recall from Section 3.4.1 that the DBPI deals with comorbidities (a single medical visit in which multiple diseases are treated) using a "pro rationing method". The MCE believes they use a more accurate method by employing more information. They use "episode groupers" which allocate spending from individuals claims records to a distinct condition using a computer algorithm. These algorithms also take into account the patient's medical history when assigning proportions of a visit to each disease. Aizcorbe and

Nestoriak say that the largest disadvantage to their comorbidity method is that the algorithms are a “black box”. The DBPI’s pro rationing method, on the other hand, is simple and transparent.

4.4 Results

Figure 5 plots the MCE against the two DBPIs and the Traditional index. At last, the Big 3 Indexes all in one picture![20] The MCE is consistently higher than the other indexes (and is even higher than the MCPI for a few years as Figure 6 shows). The literature thinks of the MCE as tracking the “service price per encounter” while the other indexes all track the ”service price per procedure”. Historically, the service price per encounter has risen faster than the service price per procedure. Why this has happened is explored in Sections 5 and 6.

The most notable trend in the MEPS MCE is the huge jump in 2016. Figure 6 reveals that MEPS is the culprit. Figure 6 plots the MCE from the Blended account, which is is the preferred MCE as it uses more data than just MEPS and is therefore less volatile. We compare the MEPS MCE with the Traditional and DBPI indexes because they all use the exact same source for weights. However, because the Blended MCE uses more data, it is less volatile. When the volatility of the smaller MEPS sample size is removed, the MCE is consistently larger than even the MCPI.

The huge jump in 2016 for the MEPS MCE is not even a blip in the Blended MCE. This suggests something very particular went on in the MEPS dataset to greatly increase the average expenditure per encounter, but not the average price per service[21].

4.4.1 Inflation Rate Differences

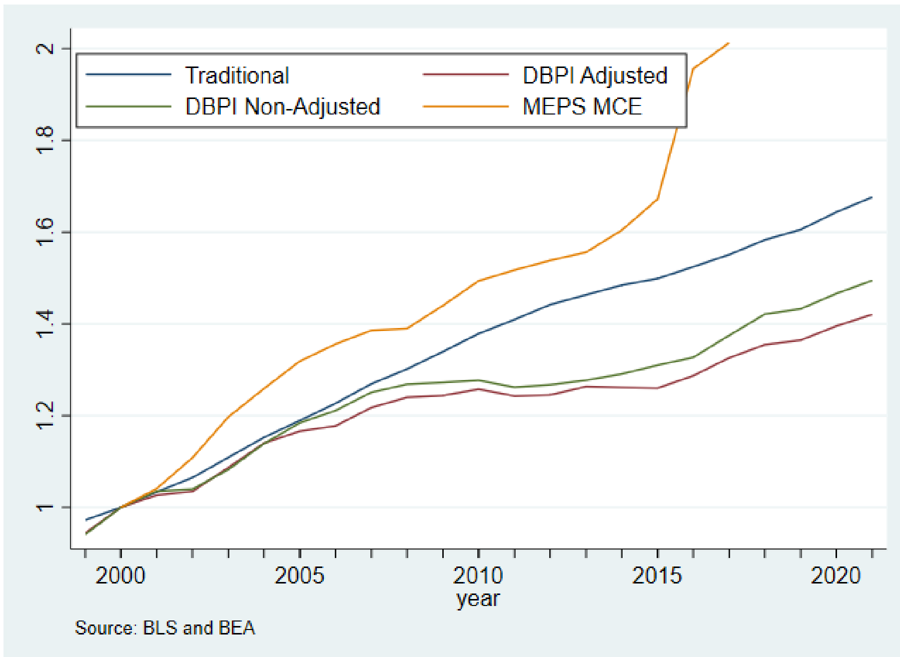
So far, all the graphs have depicted price levels. Figure 7 plots the year to year inflation rate for the different indexes, where the inflation rate in percent at time t for index i, $\pi_{t,i}$, is defined as

$$100\% * \frac{\pi_{t,i} - \pi_{t-1,i}}{\pi_{t-1,i}}$$

[20] While the MCPI, not the traditional index, is what I dubbed one of the big three, the traditional takes the place of the MCPI in these comparisons.

[21] The analysis in Section 5 suggests the intensity of services increased in MEPS.

Figure 5: The Big Three (2000 = 1)



Looking at the price indexes from this perspective reveals that the MCPI approximates the COLI for healthcare as rising at a constant clip of about 3.5% per annum, while the Traditional index marks growth at about 2.5% yearly over this period. The adjusted DBPI's average yearly inflation is 1.9 % (with a low of less than 0 and a high of over 5) and the volatile MEPS MCE averages 4.2% year[22]. These are four very different stories of how medical prices have changed in the last two decades. What we must do is identify which story best fits reality, both on paper and in practice. The following section is devoted to this analysis.

[22]The Blended MCE, not shown here, averages 3.7%.

Figure 6: The MCPI and the Two Flavors of MCE

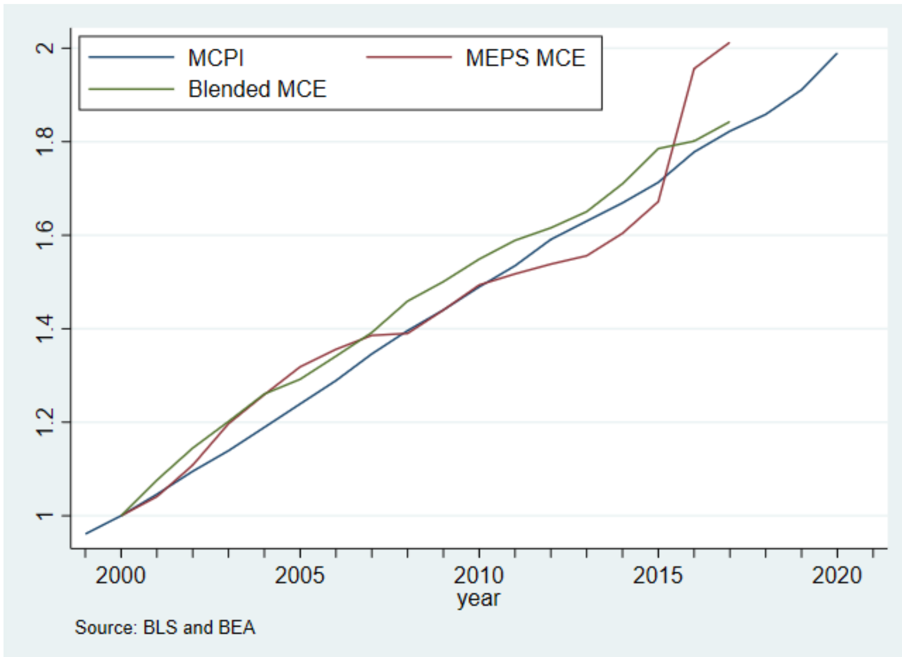
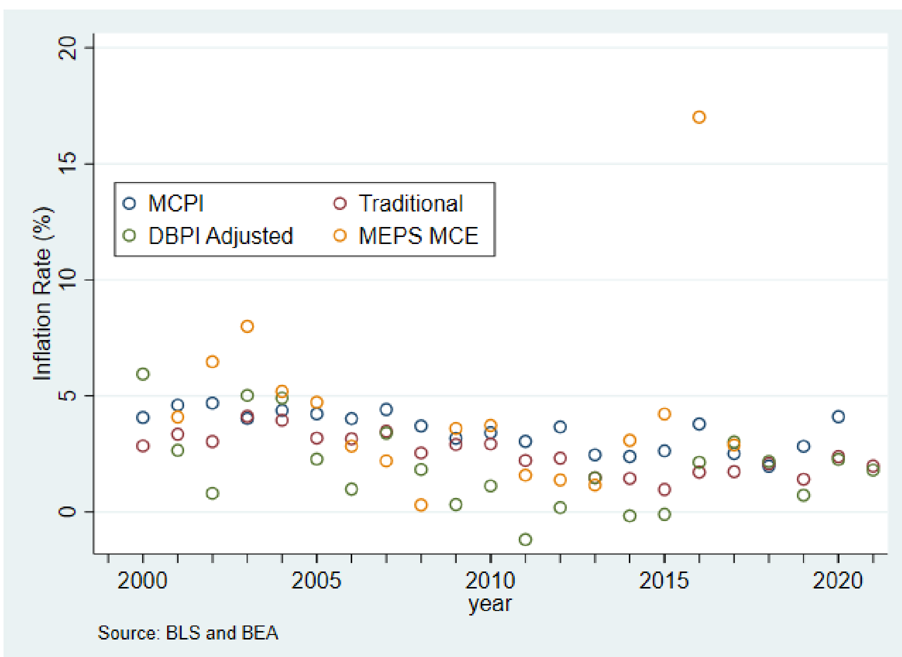


Figure 7: Inflation Rates for the Big Three



5 Key Differences in Theory and Practice

This section contains various scenarios and compares the effects these scenarios would have on the big three indexes.

5.1 Theoretical Scenarios:

1. Intra-industry substitutions

   Say Alzheimer's patients substitute away from a brand name drug towards a generic one because the price of the brand name drug is more expensive in period t than it was in the base period b. This leads to the quantity of the brand name purchased in t, $Q_{Brand,t}$ to be less than the quantity of the drug purchased in b, $Q_{Brand,b}$. And on the flip side, $Q_{generic,t} > Q_{generic,b}$.

   MCPI: The MCPI for drugs was computed using a geometric mean formula before 2016. This formulation assumes $Weight_{Brand,t} = Weight_{Brand,b}$, where $Weight_{Brand}$ is the share of expenditure towards the brand name drug relative to expenditure on all drugs. If the quantity purchased falls exactly the right amount to keep the expenditure share constant, then the geometric formulation perfectly captures this within industry substitution. If, however, $Weight_{Brand,t} < Weight_{Brand,b}$ (quantity falls so much that even with the higher price, share of expenditure is lower) then the brand name drug's price relative will be larger than it should be in a true COLI. The "price relative" is: $\frac{P_{Brand,t}}{P_{Brand,b}}^{\frac{W_{Brand,b}}{\sum W_{Drugs,b}}}$. The constant share assumption is violated for the generic drug's price relative as well, but in this case the violation leads the price relative to be smaller than it should be in a COLI. What determines the overall MCPI drug index's distance to the truth is the difference in these two price relatives. In reality, however, it was discovered that $Weight_{Brand,t} > Weight_{Brand,b}$. Even though Q_{Brand} falls, it does not fall enough for the new weight to be less than the old weight. Thus, the MCPI was deeemed to understate the true COLI by assuming constant expenditure shares and since 2016, a Laspeyres formulation is used. A Laspeyres formulation assumes no quantity change, so it ignores all within industry substitutions. As such, it overstates a COLI (but the BLS believes the Laspeyre's overstatement to be closer to the COLI than the geometric mean's understatement).

DBPI: Because the DBPI updates its weights annually, it is identical to a Laspeyres in this regard from month-to-month (no substitution is taken into account). However, it captures yearly changes in quantity that would arise from within industry substitutions.

MCE: The MCE updates its "weights" quarterly and is in fact only published quarterly. Like the DBPI, it captures substitution effects when its weights are updated. In this case, every quarter as opposed to each year.

Assuming t and b are months in different years, the MCPI will show faster growth than the DBPI and MCE. Because the DBPI gets its pricing information from the CPI/PPI, it should show faster growth than the MCE.

2. Across-industry substitutions

   To treat depression, patients have been shifting from therapy to prescription drugs (antidepressants), which are much cheaper. This is an across industry substitution.

   MCPI: The MCPI deals with multiple industries only in its aggregate index. As usual, the weight component is what matters here. Because the aggregate index is Laspeyres, the index assumes no substitution across industries and thus captures none of the lowered COLI in a 2 year period. Recall from 2.3 that the aggregate weights are updated every 2 years.

   DBPI and MCE: The exact same logic applies here as in the above example. The DBPI and MCE will grow slower than the MCPI, but the MCE should be the slowest grower of the three because it will catch this substitution sooner.

3. Intensity Shifts

   Imagine that the intensity of treatment increases over time. In a single doctor's visit to treat arthritis in year b, 2 X-Rays were taken on average. In a similar visit in year t, 4 X-Rays are taken on average. The intensity has doubled and undoubtedly, the cost of treating arthritis has increased.

   MCPI: The MCPI tracks the price of the same X-Ray from the same doctor's office year to year, it does not take intensity into account. So it completely ignores this shift.

DBPI: Because the DBPI takes prices straight from the PPI/CPI and weights based on number of encounters to treat a disease (which don't change with intensity), it also fails to reflect intensity shifts.

MCE: The MCE, unlike the other two indexes, is not concerned with tracking the price of specific services over time. It directly measures expenditures per encounter. So if patients are paying for double the X-Rays in period t than they were in period b, the MCE views this (correctly) as an increase in the cost of living, growing faster than the MCPI and DBPI.

4. Price of specific procedure different depending on disease

Roehrig (2017) mentions that when a specific drug's patent expired, prescription prices for patients treating that specific disease (hyperlipidemia) fell while average drug prices were rising. This meant the drug price used to treat hyperlipidemia was different in both level and growth than the average price of prescription drugs.

MCPI: The MCPI makes no effort to index by disease, so of course this fact is not taken into consideration by the MCPI. If we were to manually index the MCPI by disease like this (where d = hyperlipidemia and K is the set of all drugs used to treat it),

$$MCPI_{d.b.t} = \frac{\sum_{k=1}^{K} P_{k,d,t} z_{k,d,y(b)}}{\sum_{k=1}^{K} P_{k,d,b} z_{k,d,y(b)}},$$

this index would grow slower than the overall MCPI for drugs all else constant. However, such an index is not published.

DBPI: Because the DBPI grabs its inflation measure for prescription drugs from the CPI-Rx (the CPI for pharmaceuticals)in the drug component of its disease based index, the DBPI for hyperlipidemia would grow faster than the true COLI.

MCE: The MCE is not victim to the whims of the CPI/PPI and so it would grow slower than the DBPI and MCPI in this instance as the expenditure for encounters used to treat hyperlipidemia decreased (all else constant).

5. Severity Shifts

    If the severity of a disease increases from the base period, that means more inputs are required to achieve the same level of health stock. From one perspective, the COL has not changed. The quality of the inputs is the same and $h_{d,t} = f(z_{d,t}) = h_{d,b}$. All else equal, the cost of obtaining $h_{d,t}$ is the same.

    However, it has changed from another perspective. Let 0 be the level of health at which humans feel like they're really living– define achieving this level of health stock as the cost of living. At time b, when the disease was less severe, a patient with disease d started off with H=-5. But patients at time t, when the disease is more severe, have a baseline level of, say, -6. Let $h_{d,t} = h_{d,b} = 5$. After being treated in b, the patient is at H=0, while after being treated in period t the patient is at H = -1. So even more inputs z are required to get to H=0, and from this perspective the cost of living has increased. We would observe an increase in the quantity of medical goods and services purchased to treat the disease as patients chase down H=0.

    From a third perspective, the COL has decreased. Instead of defining the cost of living as obtaining the same amount of $h_{d,t}$, define it as the utility gained from obtaining $h_{d,t}$, or $v(h_{d,t})$. Assuming diminishing returns, even though $h_{d,t} = h_{d,b} = 5$, $v(h_{d,t}) > v(h_{d,b})$ because the patients are going from a level of -6 to -1 health as opposed to -5 to 0 as they did in the base period. All 3 indexes disagree with either perspective and view the cost of living as having remained unchanged; they want to track the price of obtaining $h_{d,t}$.

    MCPI: From the perspective that the COLI is unchanged, the MCPI actually does the best job. It does not update its weights using real expenditure data, and the numerator of a Laspeyres index is precisely the hypothetical expenditure on period b quantities at period t prices. If the severity increases, the MCPI won't budge.

    DBPI and MCE: Because both the DBPI and MCE use MEPS to calculate weights and MEPS will observe more encounters to treat a disease in period t relative to b as severity increases, these indexes will both grow all else equal. They are overstating the true COLI as they define it.

6. Treatment quality improves

This one is easy– none of the big 3 even attempt to adjust for quality change. Usually, increases in quality lead to higher prices which do not necessarily imply that the cost of living has increased. In many cases, the cost of living undoubtedly decreases because the health output per dollar is higher. However, all these indexes would perceive quality change as a price increase and overstate the COLI.

What has happened in practice:

1. Utilization shifts (inpatient to outpatient).

   Utilization shifts can be thought of as across industry substitutions. Bradley points to the shift from inpatient to outpatient care as the key culprit in the trend break, and the MCE's methodology is born out of the implications of utilization shifts on medical price indexes. So this fact must be important.

2. Comorbidities have increased over time.

   Comorbidities are only relevant when indexing by disease. So the MCPI is unbothered by this fact. However, both the DBPI and MCE account for comorbidities. An increase in comorbidities indicates that more diseases are being treated per encounter over time. Both disease based indexes assign proportions of each visit to each disease that was treated. This means that the weights, or quantities, are decreasing. Instead of 2 visits, one to treat anemia and another to treat hyperlipidemia, we observe 1 visit to treat both. Both the DBPI and MCE will assign less than 100% weight to each disease, so their growth will be slower than the MCPI all else equal.

3. Treatment intensities have increased.

6 Hypotheses

6.1 What are sources of the divergence between the DBPI and the Traditional Index?

Historically, patients are shifting to cheaper services (e.g. from inpatient to outpatient). Additionally, the number of uninsured people was rising until 2006 (which meant more expensive emergency room visits before 2006 than after). This slow down of uninsured growth would be captured in the 2009 DBPI (MEPS has a 3 year lag). It is important to note that only quantity weights can be driving the difference between these two.

6.1.1 Hypothesis: Obamacare mattered.

Perhaps Obamacare reduced the number of emergency room visits by increasing the number of insured individuals. Obamacare was not phased in until 2011, so its effects on MEPS and thus the DBPI would not show up until 2014, but the divergence seems to have occurred closer to 2008. So while this hypothesis is a natural one to consider, its difficult to square with the timing of the divergence.

6.1.2 Hypothesis: Price transparency making patients behave "smarter".

Maybe patients were not utilizing cheaper services because they were unaware of the price difference. Were there price transparency policies or price comparison resources provided by private insurers around 2008? If research discovers this to be the case, we could run the following regression of the difference between the DBPI index and Traditional index.

$$(DBPI - Traditional)_t = a + bD_t + cT(t) + e_t$$

Here, every t is a year-month between 1999 and 2021. D is a dummy which takes the value of 1 for all year-months 3 years after which the price transparency policies and/or services were implemented (this 3 year lag insures any effects of the policy are in MEPS and therefore affect the DBPI); let 3 years after implementation be time s. $T(t)$ is a time trend which takes on the functional form best suited for the trend of the difference between the indexes. The difference looks roughly linear before 2008, so the function T could very well be $T(t) = t$ where t is the number of months since January 1999. The coefficient of interest is b, which tells us the difference in the average difference before and after time s. If this hypothesis is correct, we expect a positive and significant b.

6.1.3 Hypothesis 3: Insurer Incentives

Was there some change to the incentives faced by insurers that made cheaper services more attractive to insurers? MEPS lists the expenditures faced by insurers for each encounter, so it should be possible to see if inpatient services suddenly cost insurers more than they had previously, leading them to push customers towards outpatient care. If this is the case, then it is also worth investigating why the prices faced by insurers rose.

6.2 Why is the BEA's MEPS MCE so much larger than the other two?

The intensity of services per encounter is increasing over time. As a result, the average price of a service encounter in MEPS grows much more rapidly than the BLS' service price index. Both prices and quantity weights can be driving the differences between the MCE and the other two indexes, but the literature has traced the difference between the MCE and DBPI to the price component.

6.2.1 Hypothesis: Hours per encounter are decreasing.

As more Americans per capita get sick, if the number of hours worked by medical professionals does not correspondingly increase, every patient visit or encounter must be allotted less time. This may be associated with higher intensity treatments as medical professionals have less time to understand the patient's condition (leading to, for example, more bottles of prescription drugs per patient over time even though each individual bottle's price is the same year to year). This assumes the efficiency of diagnosis/assessment remains the same over the years. I'm not sure yet how to proxy total medical hours worked, or if there are current measures of this. In an ideal world, say we have access to the time sheets of various hospitals. For each hospital, i, during year-month t, we have a record of how many hours medical care professionals worked, $H_{i,t}$, and how many patients they serviced, $P_{i,t}$. We would then be able to track the patients per hour ratio, $R_{i,t} = P_{i,t}/H_{i,t}$, and run the following regression to test the hypotheses:

$$Y_{i,t} = \alpha + \beta R_{i,t} + \vec{\gamma}\vec{X}_i + \delta SEV_t + \epsilon_{i,t}$$

Y is a proxy for average intensity of treatment; it is hard to imagine a good proxy that works accross all diseases and services, so perhaps we should analyze each specific disease or service individually. For example, Y might be the number of pills prescribed and R would be the ratio of hours worked by doctors who prescribe drugs to the number of patients who are prescribed drugs. X is a vector of hospital specific characteristics (like the size of the hospital, whether or not the hospital is in an urban area, and anything else that is deemed relevant). SEV is a measure of the average severity for patients with the disease at time t. If we can get time specific measures of hospital characteristics and/or hospital specific measures of severity, then we can index $\vec{X}$ and SEV accordingly. If we see a significant and negative β, it would provide evidence of an associative link between the hours per patient and intensity. Understanding if the underlying mechanism is doctor's trading off time for more intense procedures will involve more analysis and good controls to ensure omitted variables don't throw a wrench into the interpretation of β.

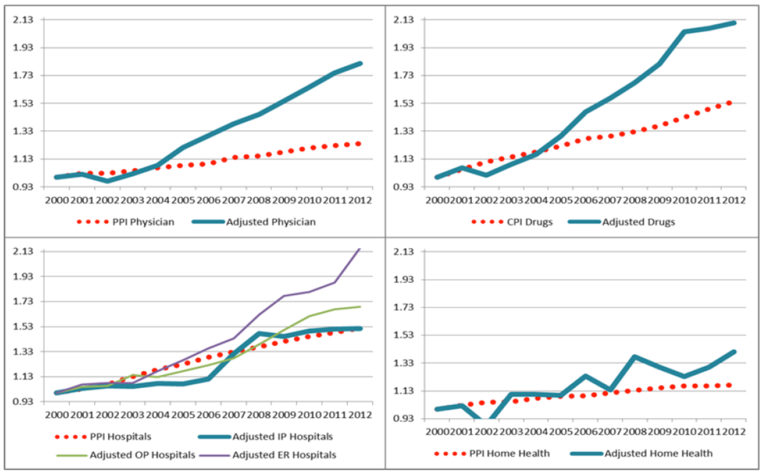
6.2.2 Hypothesis: A specific disease is driving the majority of this difference.

The best diseases to look into first are diabetes and dementia, as they have exhibited the largest rates of treatment growth over the last 2 decades. How much of the expenditure growth described by the MCE is due to prevalence growth in Diabetes/Dementia? The BEA has a readily available breakdown by disease type where diabetes falls into the large (and most rapidly growing) "endocrine; nutritional; and metabolic diseases and immunity disorders" category, for example.

6.2.3 Hypothesis: A specific service is driving the majority of this difference.

The best service to look into first is prescription drugs, as the MEPS encounter price for prescription drugs is alarmingly higher than its CPI counterpart (see Figure 8). How much of the difference between the DBPI and the MCE can be explained by the difference in the MEPS average price and the CPI-Rx price?

Figure 8: Comparison of BLS Prices Indexes and MEPS Encounter Prices for Selected Services



Note: Adjusted results are based on MEPS encounter prices. Estimates are from Bradley, Dunn and Rittmueller (2016).

Source: Roehrig (2017)

7 What We Have Learned and What is Left to be Done

The Big Three medical price indexes are all ultimately trying to answer the same question: How has the cost of obtaining the same amount of health changed? The MCPI approaches this question by tracking the prices of medical goods and services of the same quantity over time. The DBPI builds upon the MCPI by using price inflation from the MCPI and its sister index, the Medical Producer Price Index, but indexes by disease as opposed to item categories like the MCPI. The DBPI also updates its weights using expenditure data from MEPS, while the MCPI does not update weights in its item indexes and only updates them every two years in its aggregate index. Additionally, the DBPI is able to break down expenditure growth into detailed components by disease like output growth, prevalence growth, population growth, and inflation. The MCE arguably provides the best answer to the question because it tracks actual expenditures and does not rely on a fixed basket index like the CPI or PPI for its price information. However, it cannot break down expenditure growth into detailed components like the DBPI since it has no explicit quantity and price components.

Comparing the Big Three reveals that the MCE has remained above the two indexes since 2000, and it exhibits larger growth year to year. If MCE is truly the best approximation of a COLI for medical care, then the fact that its larger than even the MCPI is alarming. Before raising too much alarm, however, it is important to remember than the MCE is tracking the prices of slightly different units. The MCPI and DBPI track the prices of specific services, while the MCE tracks prices of encounters. The literature has linked the difference between the MCE and the other indexes to the increases in intensity per encounter. Understanding why the intensity has increased should be our first step to finding a targeted solution to the increase in healthcare expenditure. Additionally, none of these indexes control for quality change. Its possible that quality has increased so much in the last 20 years that Americans are actually paying less per unit of health in real terms today. The next important thing to do is to try and measure quality change. Thus far, there have been a few disease specific studies trying to measure quality controlled prices. Some examples of disease specific analysis include Cutler et al (1998), Frank et al (2004), Lucarelli and Nicholson (2009), and Howard et al (2015), while Dauda et al (2019) take a comprehensive approach.

Personally, my next steps are to test some of the hypotheses presented in Section 6 and try to find causal links between the proposed mechanisms driving the utilization shifts from in patient to out patient care (like price transparency policies/services) and the increases in intensity (like minutes per encounter). For the readers interested in exploring the topic of medical price indexes further, Aizcorbe and Highfill's *Price Indexes for US Medical Care Spending, 1980-2006* (2020) is an excellent contemporary starting point.

References

[1] Aizcorbe. "A practical guide to price index and hedonic techniques". Practical Econometrics. (2014).

[2] Aizcorbe, Highfill. "Price Indexes for US Medical Care Spending, 1980–2006." Review of Income and Wealth 66, no. 1 (2020): 205-226. (2020).

[3] Aizcorbe, Nestoriak. "Changing mix of medical care services: Stylized facts and implications for price indexes". Journal of Health Economics. (2011).

[4] Bradley, Junjan, Rozental. "Experimental Disease based Price Indexes". US Bureau of Labor Statistics. (2015)

[5] Bureau of Labor Statistics. "Consumer Prixe Index". Handbook of Methods. (2020)

[6] Cutler, McClellan, Newhouse, Remler. "Are Medical Prices Declining? Evidence from Heart Attack Treatments". The Quarterly Journal of Economics. (1998).

[7] Dauda, Seidu, Abe Dunn, and Anne Hall. "Are medical care prices still declining? A systematic examination of quality-adjusted price index alternatives for medical care." (2019).

[8] Frank, Berndt, Busch, Lehman. "Quality-Constant Price Indexes for the Ongoing Treatment of Schizophrenia: an Exploratory Study". NBER. (2004).

[9] Howard, Bach, Berndt, Conti. "Pricing in the Market for Anticancer Drugs". The Journal of Economic Perspectives. (2015).

[10] Lucarelli, Claudio, and Sean Nicholson. A quality-adjusted price index for colorectal cancer drugs. No. w15174. National Bureau of Economic Research. (2009).

[11] Roehrig, Charles. "A Comparison of Bureau of Economic Analysis and Bureau of Labor Statistics Disease-Price Indexes". Center for Sustainable Health Spending. (2017).

[12] Schultze, Mackie. “At What Price?” Conceptualizing and Measuring Cost-of-Living and Price Indexes”. The National Academies Press. (2002).