"Using Insurance Claims Data in the Medical Price Indexes"

Authors: John Bieler, Caleb Cho, Brett Matsumoto, Brian Parker, Daniel Wang[1]

November 18th 2019

1. Introduction

The use of medical insurance claims data in the construction of the medical price indexes presents many opportunities and challenges for the Bureau of Labor Statistics. These data are insurance claims that have been approved by insurance companies and include the transaction price for a corresponding medical service(s). This project seeks to develop a feasible methodology for supplementing manual price collection in the CPI medical indexes using insurance claims data. As part of an initial feasibility study, we construct price indexes using data purchased from an insurance company for a large city and compare them to the CPI medical indexes for that city. Some of the practical issues we consider are:

(1) the effect of the time lag involved in using claims data;
(2) weighting issues related to the use of both claims data and traditional manually collected data; and
(3) the high variability of prices in the claims data at the level of specific provider and service (and insurer).

The results of our preliminary analysis show promise for the use of claims data, and we are currently analyzing additional data to assess the feasibility of using claims data at a larger scale.

The BLS has previously investigated the use of claims data and determined that the benefits were not sufficient to justify the higher cost, even if the practical issues related to using claims data in the production of the official price statistics could be addressed (Song, et al. 2004). There are growing challenges to the manual collection of prices for medical services that threaten to introduce bias into the official statistics and have led the BLS to reconsider the use of claims data. Respondent cooperation among medical providers has declined and the response rates for the household surveys used to select providers have fallen. Declining response rates either lower the sample sizes which increases the variance of the published statistics or increase the cost of data collection in order to maintain a stable sample size. Providers are also increasingly refusing to provide prices for services covered by private

[1] Corresponding author: Brett Matsumoto matsumoto.brett@bls.gov. The views and opinions expressed are those of the authors and do not represent those of the BLS. This work is preliminary and no changes to existing methodology will be implemented on the basis of the results presented here. John Bieler, Brian Parker, and Daniel Wang work in the CPI division at BLS. Caleb Cho and Brett Matsumoto work in the Division of Price and Index Number Research at BLS. The authors would like to thank Erin Trish for providing private insurance enrollment data.

insurance leading to an overrepresentation of prices for the uninsured and non-covered services. If these trends continue, it will become increasingly challenging to generate accurate and representative medical price indexes using current methods.

For the initial stage of this project, we purchased claims data from a single insurance company for one city covering a period of two years, 2009 and 2010. This company is able to provide data with a two month lag, which is likely the smallest possible lag given that it takes some time to process and adjudicate claims. In order to alleviate the burden on BLS infrastructure from storing and processing very large claims databases as well as to minimize cost, we worked with the company to develop a sampling procedure. The dataset consists of the highest expenditure providers in a reference period (calendar year 2008). Then, unique procedures are randomly selected based on the reference period expenditure for these providers. A unique procedure is defined using Current Procedure Terminology (CPT) codes for physicians and hospital outpatient and Diagnostic Related Group (DRG) codes for hospital inpatient. In order to reduce the possibility of a doctor or hospital having no claims in a given month for these procedures, we restricted the sampling frame based on a quantity threshold for total number of claims in the reference period. The data consist of the average allowed amount or total payment to the provider, standard deviation of the average allowed amount, and the number of claims for every selected provider/procedure unit each month from January 2009 through December of 2010.

Using the data, we construct price indexes for physician services and hospital services using different weighting structures for aggregating and then compare the indexes to the corresponding CPI indexes for the area. Overall, the indexes formed from the claims data perform reasonably well once outliers are removed. Outliers were removed on a case by case basis considering the price history of each quote and reasonableness of price level for the corresponding service. We experimented with a number of weighting structures, including weighting the service price relatives by base period expenditure, expenditure over the entire period, base period quantity, quantity over the entire period, and a hybrid approach that weights within provider by current period quantity and across providers by base period expenditure. The quantity weight and hybrid approach generally perform similarly, while the expenditure weighted index exhibits larger month to month variation. The hybrid approach has some advantageous features. First, using expenditure to weight across outlets is similar to how aggregation occurs in the CPI. Second, using quantities to weight within provider gives more weight to services where the average is calculated from more claims, and therefore have lower sampling variability. Unfortunately, the hybrid approach is not, to the best of our knowledge, grounded in price index theory. Determining the exact properties of this index is ongoing work.

We believe that this project demonstrates the practical feasibility of incorporating claims data, but some methodological issues remain to be solved. First, there can be substantial variation in prices for a given service. Changes in average price for a given provider/service can simply

reflect redrawing a sample from the same underlying price distribution rather than a shift in the distribution. This is particularly problematic when the number of claims used to calculate the average price is small. The fundamental problem is that there are additional price determining characteristics that are either not observed or not controlled for in our methodology. A downside of adding additional controls (e.g., holding diagnosis code constant for a given provider/service) is that it increases the likelihood of no claims for a given provider/service in a month. The hybrid approach minimizes this problem by giving more weight to price relatives that are calculated from a larger number of claims. The change in average price is more likely to be a true price change if the average price is calculated from large samples. A pure unit value approach which calculates the average price at the provider/service level and aggregates with equal weight will interpret sampling variability as a real price change. There has been some theoretical work on the issue of the effect of price dispersion on price indexes (e.g. Baye (1985), Reinsdorf (1994)), but this work assumes that it's an identical product sold at different prices due to market frictions. This is different from the case of medical care where differences in prices for a given provider/service are due to unobserved quality differences. For example, the average price of a service could increase in a given month because sicker patients seek treatment for a given condition. In other cases where "big data" are used to calculate price indexes (e.g., scanner data) the products are defined at a very granular level and the calculation of average price does not involve averaging over what are essentially different products.

We are currently analyzing the data for the next stage of this project, which involves the purchase of a more comprehensive set of data. An advantage of purchasing data directly from the insurance company is that the data are generally available with a shorter lag and the insurance companies are known. The manual data collection can therefore focus on other private insurers and other payer types. Claims databases are generally less expensive than purchasing directly from insurance companies, but are only available after a longer lag, require more BLS resources to use and process, and are not representative of the entire private insurance market (which becomes a problem if the insurance companies are not identified in the data). If BLS continues manual collection of prices for private insurance, then the insurers who are also in the claims database will be overrepresented in the sample. For this next stage of the project we are evaluating the feasibility of using large claims databases for supplementing the CPI data.

2. Medical Care in the CPI

Existing methodology for collecting medical price data has some important advantages and disadvantages compared to the use of claims data. Currently, the BLS collects price data directly from providers in a manual collection process. Providers are selected randomly based on the results of a household survey.[2] The CPI goes to that provider and selects a procedure to price

[2] The Telephone Point of Purchase Survey (TPOPS). Hospital providers are selected using a different data set.

based on the billing history of the provider. A bill for the selected procedure is chosen and price determining characteristics such as insurance provider and plan type are recorded. In subsequent months the provider is asked to provide a price for the same service, holding the characteristics of the original bill constant. The price is defined as the total amount paid (or is expected to be paid) to the provider regardless of whether the payment comes from the patient or the insurance company.

Within a given geographic area for a given item category (called Entry Level Items (ELI)), the collected prices are used to form a price relative which measures the one period change in price for that item/area combination. This research focuses on the ELI's for physician services and hospital services (both inpatient and outpatient). Let P_j,t be the price of good j in period t. The price relative for the ELI is:

$$R_{\{t,t-1\}} = \frac{\sum_{\{j \in ELI\}} \frac{W_{\{j,0\}} \ P_{\{j,t\}}}{P_{\{j,0\}}}}{\sum_{\{j \in ELI\}} \frac{W_{\{j,0\}} P_{\{j,t-1\}}}{P_{\{j,0\}}}} \quad ,$$

where W is the item weight. The item weight is proportional to the share of expenditures on that particular good or service in the base period.[3]

The current data collection process has some important advantages for the purposes of constructing price indexes. Characteristics of the service are held fixed in the repricing, which fits with the fixed-basket matched model of the CPI and PPI. Also, the data are timely and are designed to be representative of the universe of medical goods and services consumed by households. These features are generally not found in claims data. Claims databases only represent a subset of private insurance claims and are only available after a lag. Another limitation of claims data is that there is a missing price if the provider does not bill any claims for a given service in a given month. Since insurance claims data are not fully representative of the market, it cannot serve as a full replacement for manual collection. This mitigates any potential cost savings from switching to the use of claims data. Therefore, using claims data to supplement existing data collection could possibly require an increase in cost for the BLS. One area where claims data has a clear advantage is the vastly larger sample size. An advantage of the larger sample size from the use of claims data is that the BLS could potentially produce medical price indexes at a finer level (e.g. regional medical price indexes, disease specific indexes, physician indexes by specialty, etc.).

3. Data

Data were purchased from a large insurance company hereafter referred to in this paper as the insurer. The data covers medical claims data for the insurer's plans for2009 and 2010. The data were sampled from a large metropolitan area. The sample included the top 10 hospital outlets

[3] See the BLS Handbook of Methods.

and top 25 physician outlets by total expenditure.[4] Ten medical services were sampled at each hospital and five medical services were sampled at each physician office via probability proportional to sales/expenditure (PPS).[5] Doctor's office and hospital outpatient services are defined using Current Procedure Terminology (CPT) codes. Hospital inpatient services are defined using Diagnosis Related Group (DRG) codes.

Each medical service and provider combination can be thought of as a single quote. Each quote is supplied with an average price, averaging across all transactions at the outlet for the specified medical service. These average prices are the prices that enter the price index calculation, which makes these unit value indexes.[6] There can be considerable variation in price at the level of provider and service (and insurance company) that is driven by unobserved factors. A limitation of a unit value index is that it assumes that the average price is calculated across homogeneous goods. Changes in the mix of unobserved characteristics will be measured as a price change in a unit value index.

Also included is a monthly quantity value which reveals the number of transactions used in creating the average price. The average price and quantity values are updated each month between 2009 and 2010, for a total of 24 months. Outlets and items were sampled based on expenditure in the year 2008.

We bought a second dataset from Marketscan to act as a benchmark for comparison. In Marketscan the items are defined in a similar way to the insurer by identifying a particular medical service-provider combination. Average prices and quantity values accompanied each quote, updated monthly. Each quote was given a static weight over the 2006-2010 period using its quantity share during this period. A quote's summed quantity during 2006-2010 was divided by the total 2006-2010 quantity of all quotes to get its quantity share.[7] Given that it includes a broader set of payers (at least private insurance payers), we consider the MarketScan indexes to be the benchmark.

Two methodologies were used to weight and calculate two indexes from the insurer's data:

- Insurer Quantity Share: Quantity shares over the 2009-2010 period were used to create static weights for each quote. The insurer provided monthly quantity values for each quote. A quote's summed quantity during 2009-2010 was divided by total 2009-2010 quantity of all quotes to get its quantity share.

[4] Given the small sample size of this initial feasibility study, we selected the highest expenditure providers to minimize the number of months with missing data. Since the providers are not randomly sampled, the resulting indexes will not be representative.

[5] Cost considerations dictated the sample size of this pilot study.

[6] See "Accounting for Health" for a discussion of unit value indexes and claims data (National Research Council, 2010).

[7] The MarketScan indexes were produced as part of a prior research project and we do not currently have an active subscription for the data for this time period. Therefore, we are unable to alter the methodology used to produce the MarketScan indexes.

Quantity weighting minimizes price volatility seen in quotes with low quantity. This is helpful because average prices provided by the insurer do not control for the International Classification of Diseases (ICD) code, a price determining characteristic. Large changes in average price for low quantity quotes are likely due to the changing composition of ICD codes between months or other unobserved factors. Weighting by quantity minimizes this effect. Additionally, this methodology best replicates how the MarketScan indexes were created, allowing for comparability.

- Insurer Hybrid: This is a 2-step weighting methodology. First, quotes are aggregated within outlets and weighted by their monthly quantity share to get an outlet relative. Each quote's quantity share weight is updated every month. Outlet relatives are then aggregated using outlet expenditure shares from the 2008 base period provided by the insurer. These outlet weights are static.

  This methodology maintains the benefits of quantity weighting described above while also attempting to replicate CPI methodology. CPI indexes weight quotes using outlet expenditure share. A hybrid method minimizes the effect of changing ICD codes while still creating indexes comparable to the CPI.

In addition to the insurer only indexes, we construct all payer indexes by combining the insurer indexes with the corresponding CPI Physicians' Services and Hospital Services indexes. Such all payer indexes show what a hypothetical CPI supplemented with alternative data might look like.

4. Results:

Private Insurance Only Indexes:

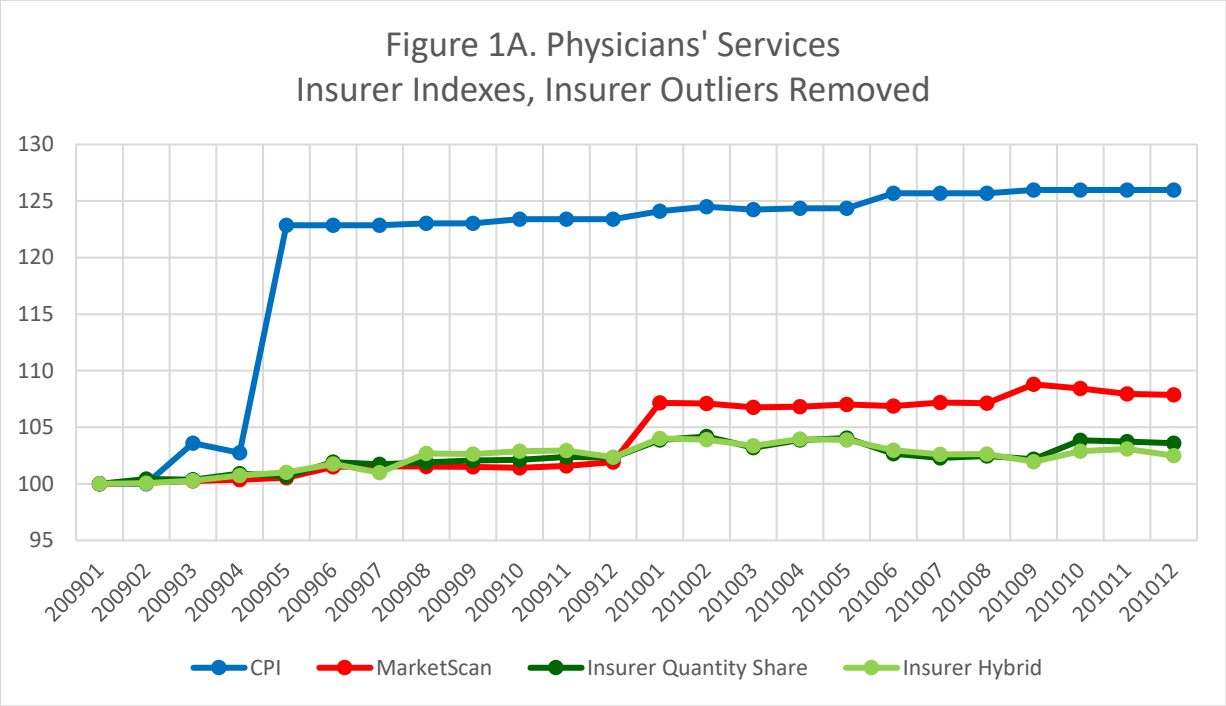


Figure 1A plots the two insurer indexes, CPI, and MarketScan for physician services. The insurer indexes track relatively closely to MarketScan, our presumed true price index. Due to the CPI's level shift in 200905, it is difficult to draw comparisons to CPI. The level shift was caused by an office visit quote (for dizziness) that increased nearly 80% and carried nearly 25% of the weight for this item area for a self-pay price. Figure 1B below, removes this one outlier.

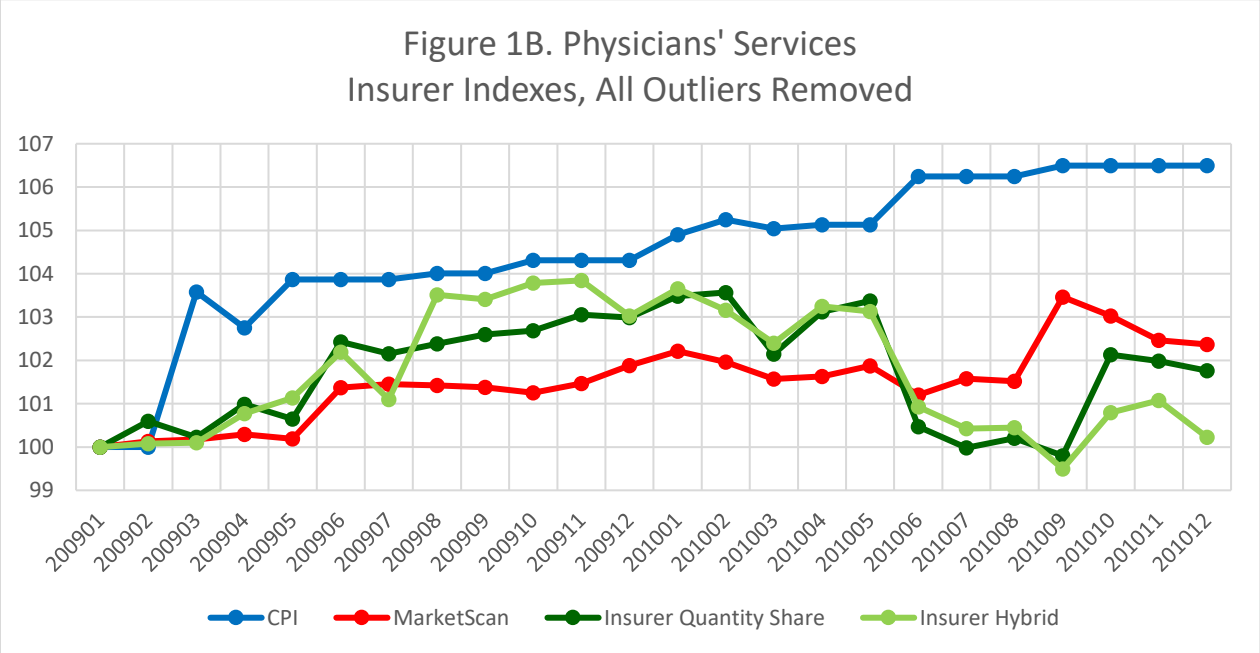


Figure 1B removes outliers from CPI and MarketScan indexes (the insurer outliers are already removed). A large level shift in the CPI is removed at 200905, as well as a large level shift in the MarketScan data in 201001.

The 200905 level shift was replaced with the CPI's average monthly relative from 200901 to 201012. Both the insurer indexes and the CPI track similarly to MarketScan. The goodness of fit measure that we use to track how closely the indexes move together over time is the sum of squared errors (SSE) between the aggregate indexes. The SSE was calculated between MarketScan, the true price index reference, and each of the three other indexes:

- CPI-MarketScan SSE = **257**
- Insurer Quantity Share-MarketScan SSE = **39**
- Insurer Hybrid-MarketScan SSE = **62**

The differences between SSE's is small; all four indexes are close. These results show that the insurer tracks closer to MarketScan than CPI tracks to MarketScan, regardless of weighting method.

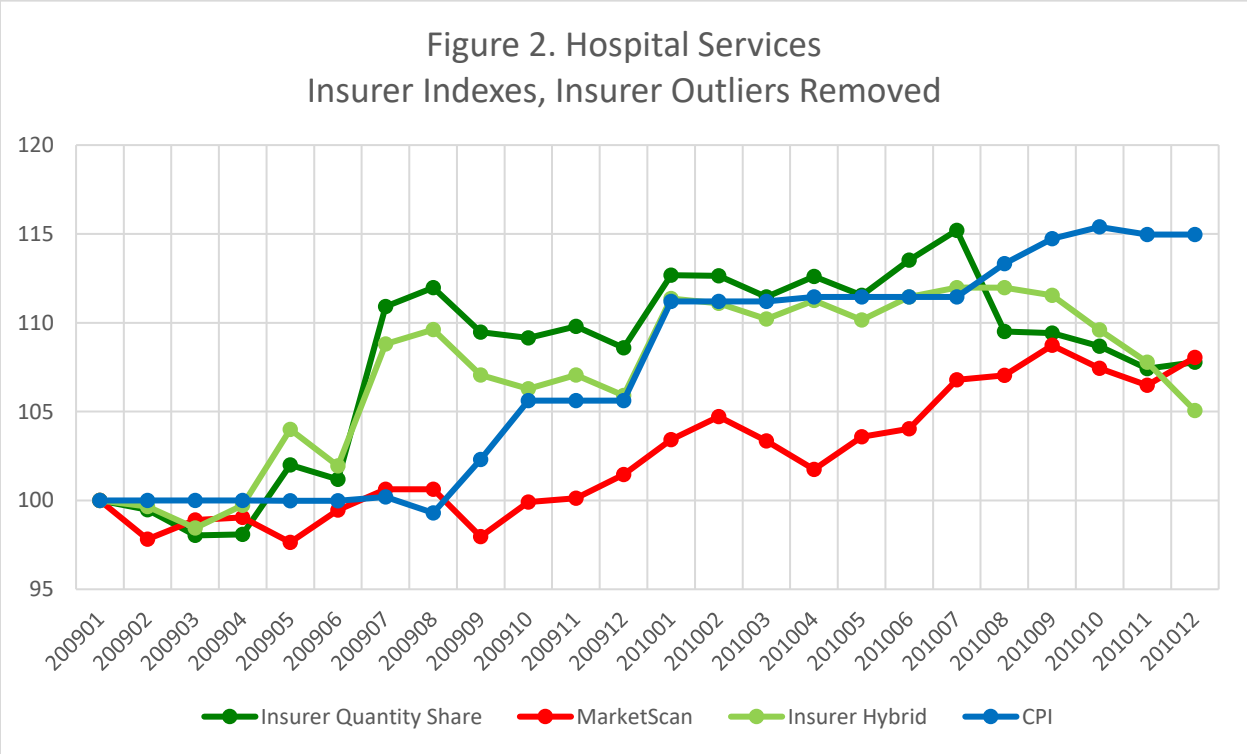


Figure 2's Hospital data shows both indexes for the insurer tracking closer to the CPI than MarketScan in terms of index trend. Outliers from the insurer's data were removed, no outliers in the CPI or MarketScan data were identified. SSE's were also calculated between MarketScan and the three other indexes:

- CPI-MarketScan SSE = **770**

- Insurer Quantity Share-MarketScan SSE = **1189**
- Insurer Hybrid-MarketScan SSE = **2444**

Unlike Physician Services, CPI tracks closest to MarketScan. This is not unexpected. For Hospital Services, CPI has only been able to collect insurance quotes in the Primary Sampling Unit (PSU), making it essentially a private insurance index which is what MarketScan is.[8] The insurer's data only includes insurance from the insurer.

Combining the Insurer and CPI Data

This section combines the insurer indexes with the corresponding CPI Physicians' Services and Hospital Services indexes. For example, the insurer Quantity Share Hospital Services index is blended with the CPI Hospital Services index to create a hypothetical CPI index that is supplemented with alternative data. The insurer Hybrid Hospital Services index is also separately combined with the CPI Hospital Services index. Blended indexes were then compared to MarketScan and CPI. The MarketScan index is combined with other payers or insurers from the CPI to represent all CPI-elgible payers.

The insurer and CPI indexes were combined using payer type market share data at the region level and private insurance market share data for the MSA. First, market share for private insurance was determined for the region in the Medical Expenditure Pannel Survey (MEPS). Market share of the insurer among private insurance payers was determined using enrollment data from InterStudy. These two market share proportions were multiplied to estimate the insurer's market share of the entire healthcare market in the MSA.

The insurer index and CPI index were combined using the insurer's market share to weight the insurer index. This required removing any quotes pricing the insurer from the CPI indexes before combining. Using these market share proportions to combine the indexes assumes that all insurer quotes are represented by the insurer's data only.

The figures below blend the insurer indexes — with outliers removed — with corresponding CPI indexes. Outliers were also removed for the CPI Physicians' Services index and MarketScan Physicians' Services index. No outliers were identified for the CPI Hospital Services index.

[8] Primary sampling units (PSU) are the smallest geographical unit in the CPI. They generally correspond to the Metropolitan Statistical Areas (MSA) for major MSAs. BLS defines the PSU boundaries for non-metropolitan areas.

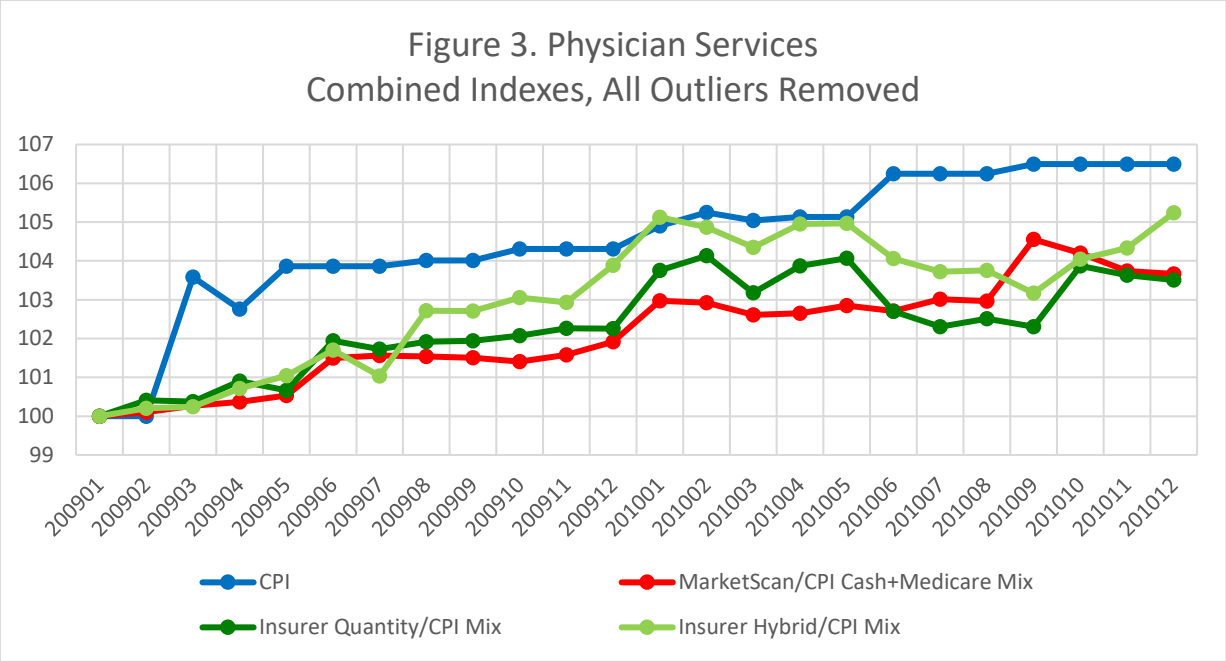


Figure 3 indexes show the insurer and CPI blended indexes. The MarketScan index is also combined with CPI cash quotes and Medicare quotes to create a representative market, rather than commercial insurance only. MarketScan and CPI data are mixed by payer-type market share using MEPS regional data. There were no private insurance quotes used for Physicians' services CPI for this metro area during this time period.

The insurer/CPI blended indexes track more closely to MarketScan than stand-alone CPI. Assuming MarketScan to be the true price index, this suggests supplementing CPI data with alternative data can improve the accuracy of the Physician Services CPI index. SSEs were also calculated for these indexes:

- CPI – MarketScan SSE = **157**
- Insurer Quantity Share and CPI Combined – MarketScan SSE = **15**
- Insurer Hybrid and CPI Combined – MarketScan SSE = **24**

The SSEs further support using alternative data in the Physician Services CPI index.

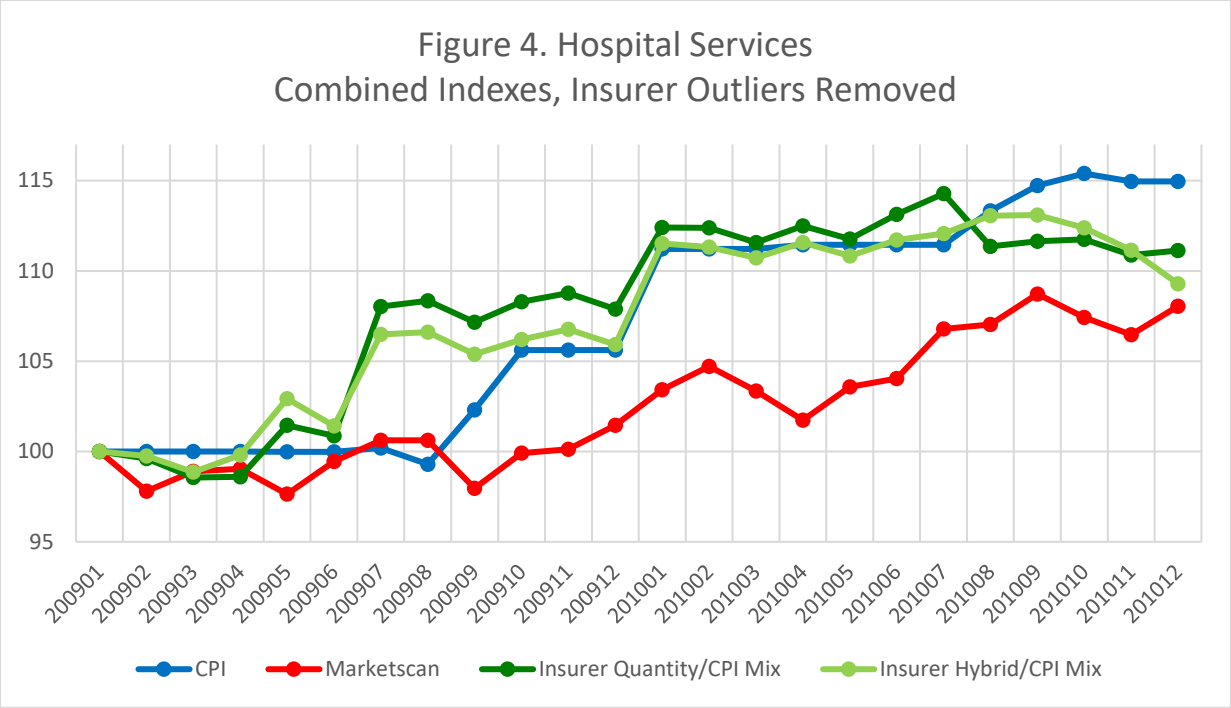


Figure 4 shows the insurer and CPI hospital indexes combined. However, in the PSU there are only insurance quotes for CPI hospital data for this time period. This means that MarketScan is not mixed with any cash prices to provide an insurance-only reference point. Both the insurer/CPI blended indexes track more closely to CPI than MarketScan. SSEs were also calculated for these indexes:

- CPI – MarketScan SSE = **770**
- Insurer Quantity Share and CPI Combined – MarketScan SSE = **1009**
- Insurer Hybrid and CPI Combined – MarketScan SSE = **767**

Combined insurer and CPI indexes track very closely to the CPI Hospital Services index. However, this means that the combined insurer and CPI indexes do not dramatically improve the CPI Hospital Services index in the PSU. This may suggest that the insurer data is not necessarily as helpful in areas like such as this PSU where CPI has many quotes that are routinely repriced. But in areas where CPI is completely unproductive, area-specific insurance company data may be more accurate than entire area imputations.

These findings indicate potential improvements from using alternative data, but further research in other PSUs are necessary.

5. Lag Analysis

A key characteristic of alternative medical claims data is its lag. Any alternative data that is used to supplement CPI indexes will require processing on the part of the vendor. Alternative data are also based on fully adjudicated claims, which can take several months to finalize. Relative to

CPI's current methodology, this is a weakness of using alternative data. This results in data that are lagged by some months time before it can be delivered to the CPI. In its market research, CPI has found multiple data sources of lag time ranging from two months to nine months. Specifically, this lag time is counted from the date of service of the medical claim. For example, a doctor's office visit with a date of service in January will not be delivered to the CPI until early April. BLS could use this January data to calculate March indexes, showing December to January price relatives in the April to March index. This section examines the behavior of CPI indexes when combined with lagged alternative data.

Lagged and Combined Indexes for the MSA

The figures displayed in this section examine the insurer, MarketScan, and CPI data at the MSA level. It is important to acknowledge that this analysis is for one MSA over a relatively short 22 month period, potentially limiting the generalizability of these findings. The next section analyzes supplemented, lagged indexes on a national level over a longer time frame.

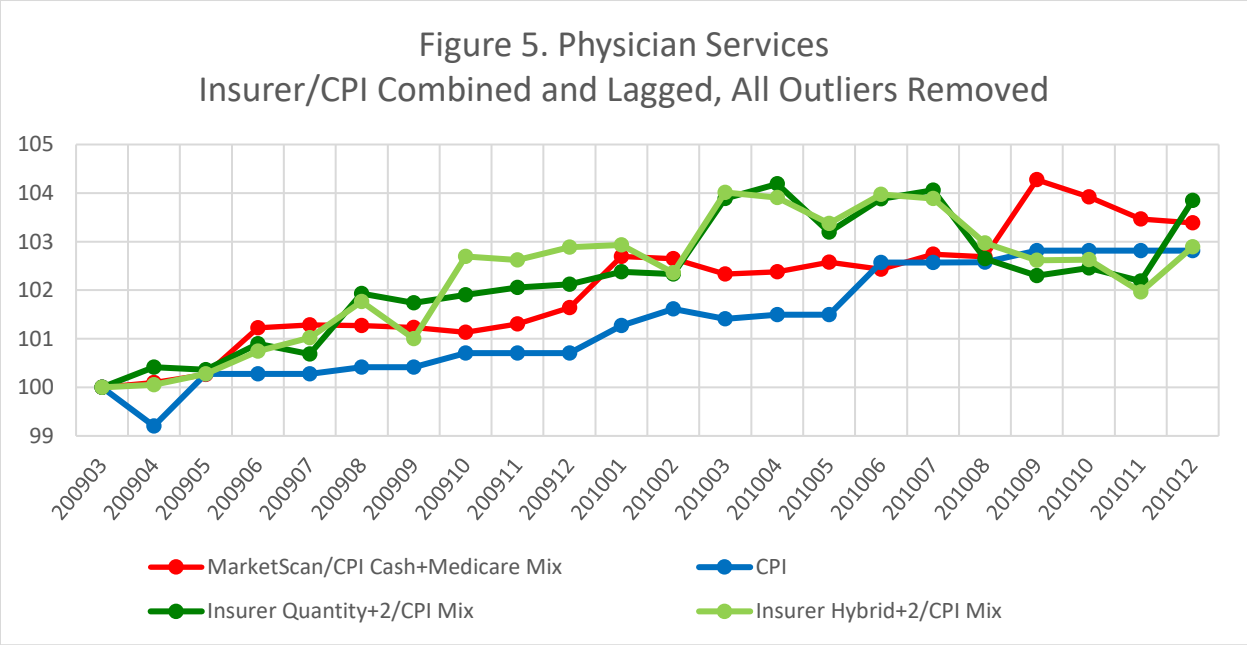
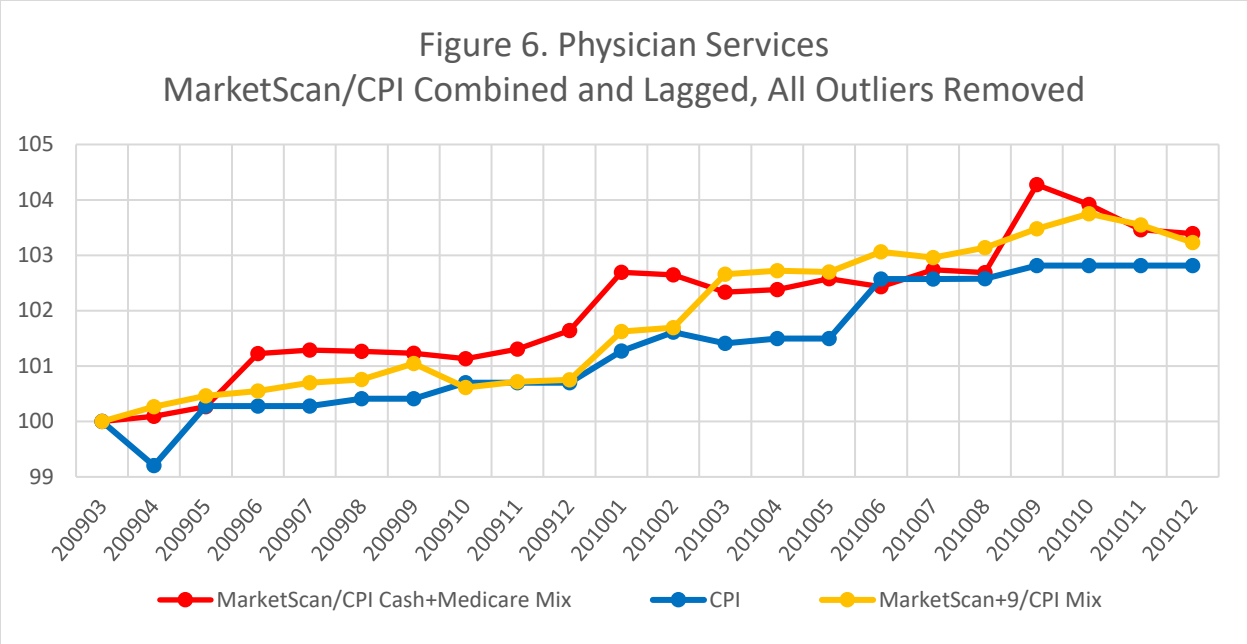


Figure 5 presents the CPI Physician Services index combined with the insurer Quantity Share index, and the insurer Hybrid index with outliers removed. It seems that CPI tracks more closely to MarketScan in its month-to-month behavior (other than MarketScan's dramatic rise in 201009) than the supplemented, lagged indexes. However, looking at the SSEs shows all three indexes are close to MarketScan. Given how close the three indexes are, it is difficult to draw a definitive conclusion here.

- CPI – MarketScan SSE = **15.6**
- CPI and Insurer Quantity Share Combined and Lagged – MarketScan SSE = **20.6**
- CPI and Insurer Hybrid Combined and Lagged – MarketScan SSE = **39.8**



To see how a longer lag time would affect supplemented indexes, MarketScan data were first lagged nine months then combined with CPI. Nine months is the shortest lag time offered by MarketScan. Like the unlagged MarketScan true price index, the lagged MarketScan and CPI data are mixed by payer-type market share using MEPS regional data.

Figure 6 shows that all three indexes are close, however the SSEs suggest that the MarketScan/CPI combined and lagged index is closest to the MarketScan true index:

- CPI and MarketScan Combined and Lagged – MarketScan SSE = **6.2**

This is the smallest SSE of the various supplemented and lagged CPI indexes (this includes the SSEs calculated for figure 5). This may be expected insofar as combining CPI and MarketScan data would result in an index that tracks closer to MarketScan (i.e. lagging MarketScan and comparing it to unlagged MarketScan resulted in the small difference). However, the supplemented index uses MarketScan lagged by nine months, suggesting that a significant lag time does not significantly skew the combined index.

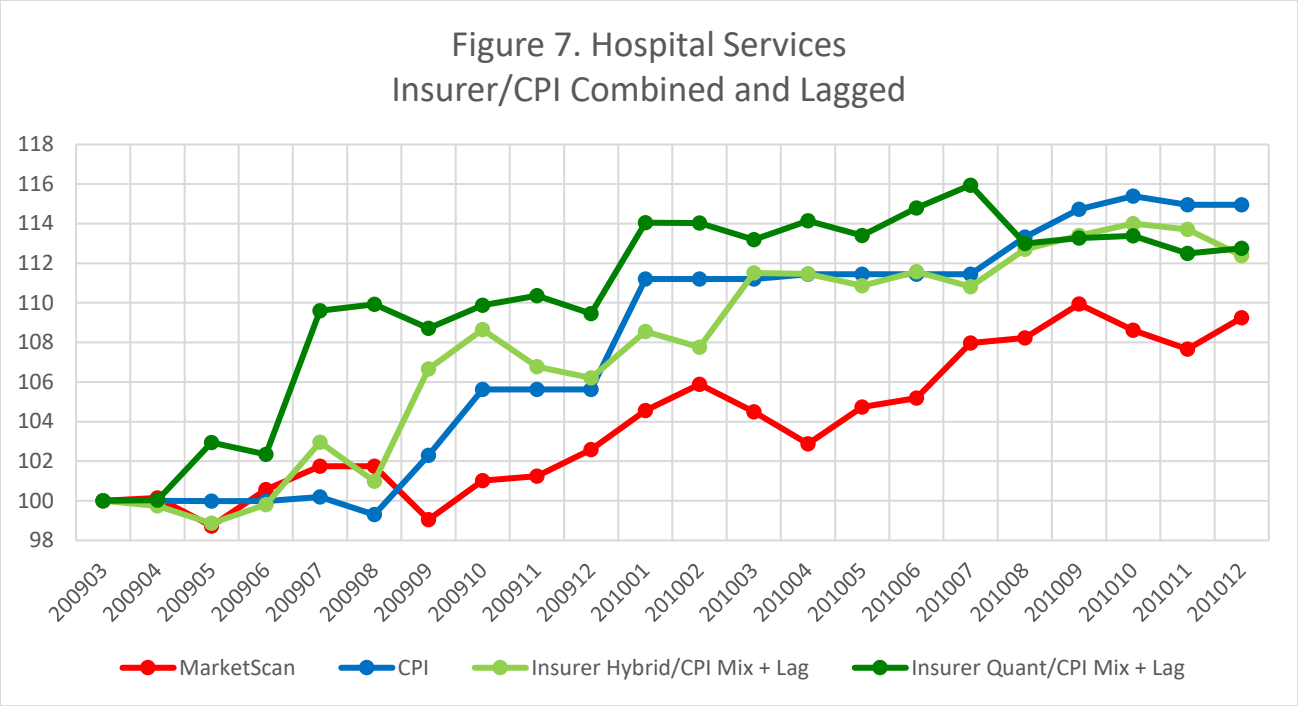
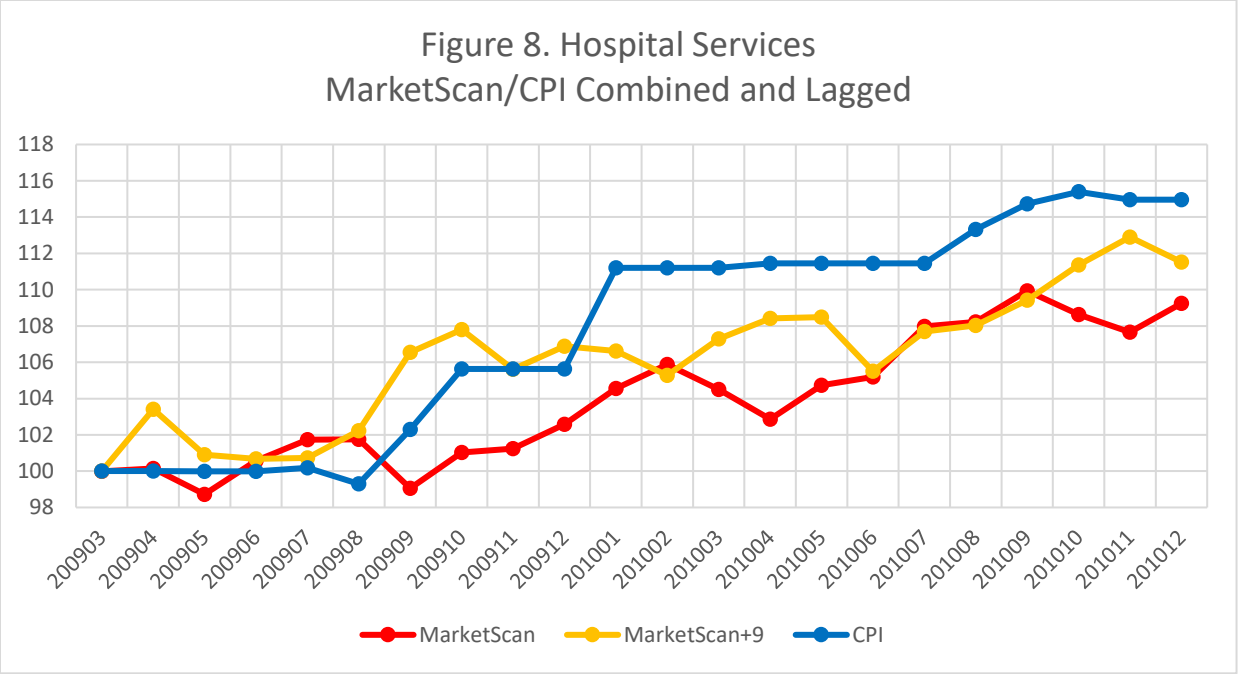


Figure 7 shows CPI Hospital Services indexes supplemented with lagged insurer data. MarketScan and true CPI indexes are included for comparison. Again, CPI hospital data in the PSU has only insurance quotes. Thus, MarketScan is not mixed with any cash prices to provide an insurance-only reference point.

The supplemented, lagged indexes seem to track closer to the CPI in terms of behavior and trend. But SSEs between the indexes show that the CPI index supplemented with the insurer Hybrid index tracks closest to MarketScan, though all are similar:

- CPI – MarketScan SSE = **539**
- CPI and Insurer Quantity Share Combined and Lagged – MarketScan SSE = **617**
- CPI and Insurer Hybrid Combined and Lagged – MarketScan SSE = **499**



CPI Hospital Services supplemented with nine-month lagged MarketScan data tracks closer to MarketScan. Again, this may be expected but suggests that a nine month lag may be reasonable for production. Similar to Physician Services, the SSE shows a MarketScan/CPI combined and lagged nine months tracks closest to true MarketScan than the Insurer/CPI combined and lagged indexes:

- CPI and MarketScan Combined and Lagged – MarketScan SSE = **254**

This analysis shows that when CPI is combined with MarketScan lagged nine months, the resulting supplemented index tracks closest to real-time MarketScan. Even with a much shorter lag, Insurer/CPI combined and lagged indexes do not track as closely to MarketScan's true price index. This may suggest that a more representative and exhaustive alternative source like MarketScan should be prioritized over lag time.

National Data Lag Analysis

This section looks at how lagged data affects CPI indexes on a national level, rather than a specific PSU. This intends to remove any PSU-specific abnormalities that may have been observed in the PSU data. MarketScan hospital and physician national indexes are combined with national level CPI Hospital and Physicians' Services indexes, respectively. It is important to note that CPI Medical is not considering using a national level alternative claims index to supplement the CPI. If alternative claims data is to be used in the CPI, it will be done so at a PSU index level. This analysis is for research purposes only, to gain a better understanding of how lagged data may look at a national level.

The MarketScan and CPI data are combined using national MEPS payer-type market shares. The MarketScan data is weighted by MEPS share of commercial insurance. The CPI cash and Medicare data is weighted using MEPS share of self-pay and Medicare. These portions are then aggregated into a larger index.

All data in this section is at the national level, and observed over a 52 month period.

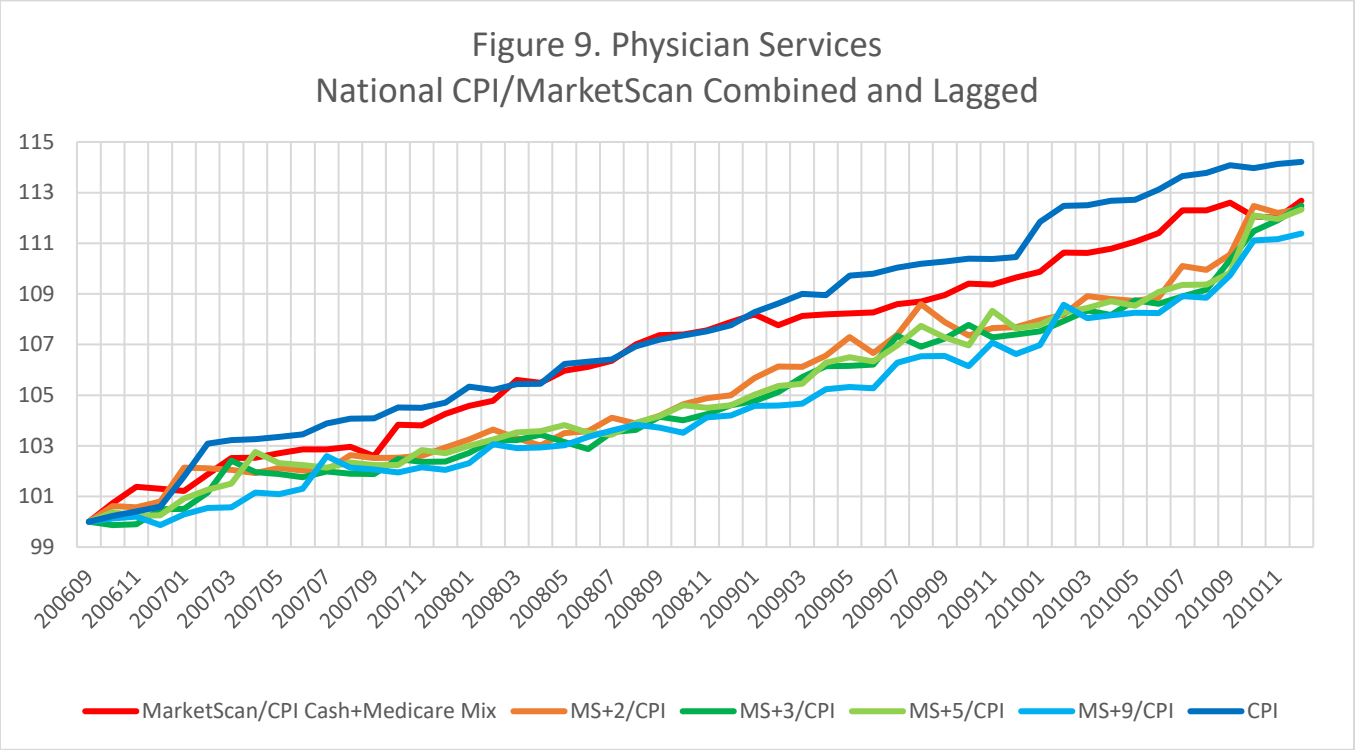


Figure 9 shows multiple CPI indexes for Physician Services supplemented with MarketScan data lagged at various time intervals. The true MarketScan index, combined with CPI cash and Medicare quotes, is also included as a reference. SSEs between the various CPI supplemented indexes and the true MarketScan are:

- CPI and MarketScan Combined, Lagged 2 Months – MarketScan SSE = **20.3**
- CPI and MarketScan Combined, Lagged 3 Months – MarketScan SSE = **11.7**
- CPI and MarketScan Combined, Lagged 5 Months – MarketScan SSE = **46.3**
- CPI and MarketScan Combined, Lagged 9 Months – MarketScan SSE = **78.3**
- CPI – MarketScan SSE = **331.4**

SSEs generally show that as the lag increases, the further the supplemented CPI index tracks from the MarketScan true price index. Also reinforcing the findings above, Figure 9 shows that CPI indexes supplemented with MarketScan data, lagged by up to 9 months outperforms the true CPI itself. This is due to the CPI's oversampling of self-pay quotes.

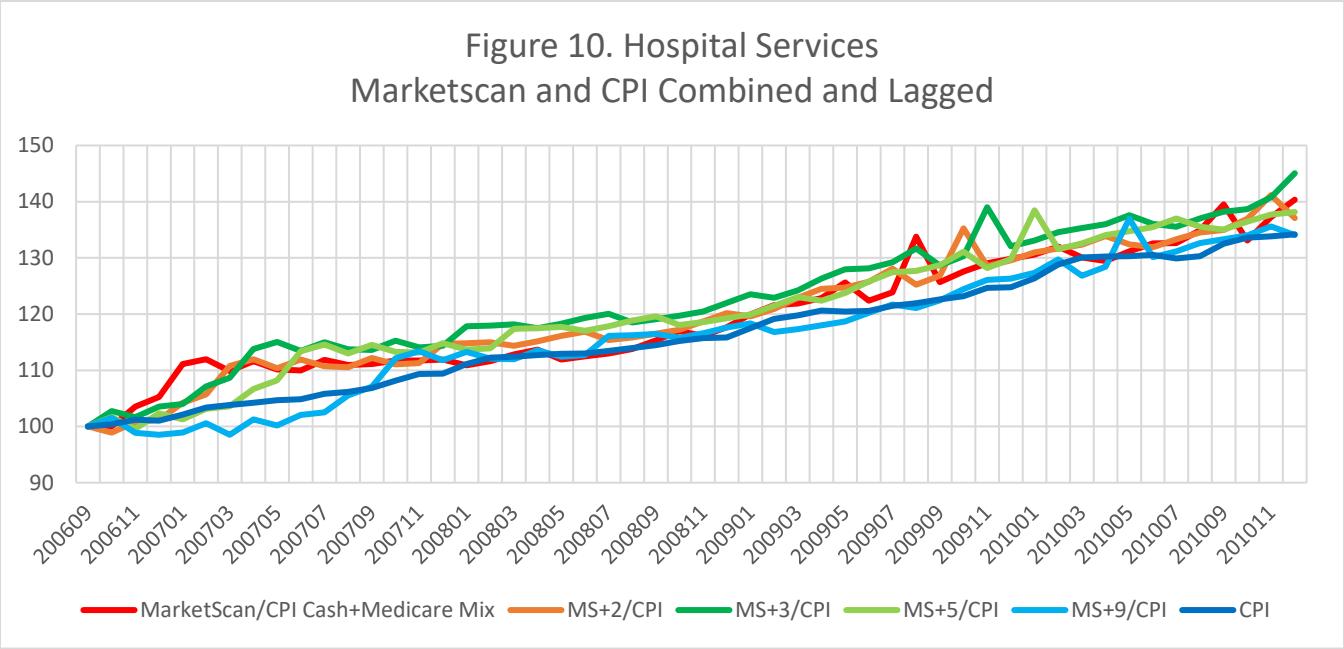


Figure 10 replicates Figure 9 using Hospital Services data rather than Physician Services. The true MarketScan index, combined with CPI cash and Medicare quotes, is also included as a reference. SSEs between the various CPI supplemented indexes and MarketScan are:

- CPI and MarketScan Combined, Lagged 2 Months – MarketScan SSE = **479**
- CPI and MarketScan Combined, Lagged 3 Months – MarketScan SSE = **972**
- CPI and MarketScan Combined, Lagged 5 Months – MarketScan SSE = **753**
- CPI and MarketScan Combined, Lagged 9 Months – MarketScan SSE = **1,391**
- CPI – MarketScan = **869**

Again, we see that generally as lag increases the accuracy decreases. For hospital data, it appears that CPI supplemented with MarketScan lagged by two months outperforms true CPI. But using a longer lag may result in a less accurate CPI index.

Lag analysis on a national level indicates that as lag time increases, the accuracy of the supplemented index decreases. This is an intuitive finding, and CPI Medical recommends using the shortest possible lag time when supplementing with alternative data.

6. Conclusion

Based on the three sections of analysis above, CPI Medical has drawn the following conclusions:

1. **Even with lag, CPI indexes can be improved by supplementing with alternative data.** This is observed in Figure 5, where the insurer/CPI combined and lagged indexes

outperform the CPI index. Figures 5 and 6 show that MarketScan lagged nine months combined with CPI outperform any insurer/CPI combination. Figures 7 and 8 show the same. We believe that the improvement would likely be greatest in areas with lower provider response rates than the PSU studied here.

On a national level, this is also observed in Figures 9 and 10. Figure 9 shows CPI supplemented with MarketScan is more accurate than the published CPI for physician services, up to nine months of lag. However, for hospital services the results are less definitive.

2. **Supplementing using a more exhaustive and representative alternative data source can be more accurate than supplementation with alternative data of a single payer.** This holds true even if the more exhaustive data source has a significantly longer lag. Figures 5 and 6 show that MarketScan lagged nine months combined with CPI outperform any insurer/CPI combination with the insurer data lagged a much shorter two months. Figures 7 and 8 show the same.

   However, this finding may likely be a result of the limited sample size collected from the insurer rather than MarketScan being a multi-payer source. This can be mitigated by using a national source that allows access to the full database, rather than a limited sample size like that of the insurer.

3. **As the lag time increases, the accuracy of supplemented CPI indexes may decrease.** This is observed in Figures 9 and 10, where a general trend of longer lag is correlated with less accuracy.

4. **There are discrepancies between PSU-level data and national data as to how long of a lag is acceptable**. In figures 5-8, we observe that MarketScan data lagged by nine months combined with CPI outperform the true CPI index at the PSU level. However, in Figure 10 the MarketScan/CPI combined and lagged index only outperforms true CPI on a two month lag at the national level. This suggests further research in multiple PSUs is required to confirm whether the MSA data is an outlier.

7. Next Steps

We have obtained claims data on a national level. It includes data from all CPI areas and more medical providers in a given area (and more services per provider) than this initial study. It also includes a greater number of insurance providers. In order to reduce the price variability within a given item, we condition on additional price determining characteristics. For physician services, we condition on the place of service (since reimbursement amounts depend on whether the service is provided in a facility or office setting). For inpatient hospital services, we

condition on primary diagnosis (at the ICD chapter level). The tradeoff of conditioning on additional characteristics is that it increases the likelihood that a given item (defined as a combination of characteristics) will not appear in a given month. The sampling procedure will otherwise remain the same, except that providers will be selected randomly (proportional to base year expenditures) rather than selecting the highest expenditure providers.

The data include many different insurance companies and since the insurance company is a price determining characteristic, we condition on the insurance company when sampling the procedures. The insurance companies in the data are de-identified by the data provider, so unlike the insurer data we do not know the specific insurance company associated with each sampled item. This can complicate the creation of the blended indexes with the manually collected CPI data since we cannot determine which insurance company quotes should be excluded.

References:

Baye, Michael. (1985), "Price Dispersion and Functional Price Indexes," *Econometrica*, Vol. 53, No. 1, p 217-224.

Bureau of Labor Statistics (2016). Handbook of Methods. Division of BLS Publishing, Office of Publications and Special Studies.

National Research Council. (2010). *Accounting for Health and Health Care: Approaches to Measuring the Sources and Costs of Their Improvement.* Panel to Advance a Research Program on the Design of National Health Accounts, Committee on National Statistics. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Reinsdorf, Marshall. (1994), "The Effect of Price Dispersion on Cost of Living Indexes," *International Economic Review*, Vol 35, No. 1, p 137-149.

Song, X., W. Marder, R. Houchens, J. Conklin, and R. Bradley (2009). "Can a disease based price index improve the estimation of the medical CPI?" In E. Diewert, J. Greenless, and C. Hulten (Eds.), Price Index Concepts and Measurement, pp. 329{372. National Bureau of Economic Research.