BLS WORKING PAPERS

U.S. Department of Labor
U.S. Bureau of Labor Statistics
Office of Prices and Living Conditions

Producing Quality Adjusted Hospital Price Indexes

Brett Matsumoto
U.S. Bureau of Labor Statistics

Working Paper 543
November 15, 2021

Producing Quality Adjusted Hospital Price Indexes

Brett Matsumoto†

Bureau of Labor Statistics

November 15, 2021

Abstract

This paper evaluates the relationship between hospital cost and quality and examines the impact of changing hospital quality on measures of hospital price inflation. To construct the official price indexes, government statistical agencies collect the prices of a fixed sample of goods over time. The price indexes can be biased if the quality of the goods is changing over time. In that case, the price index would reflect unobserved quality change in addition to pure price changes. Cost changes associated with the quality change can be used to quality adjust the price index. Measures of hospital quality are published by the Department of Health and Human Services through its Hospital Compare project. This paper estimates the causal relationship between the quality measures and hospital costs using the instrumental variable technique of Doyle et al. (2015), which is based on plausibly exogenous assignment to different ambulance companies. The relationship between cost and quality measures is then used to produce a cost-based quality adjustment for the inpatient hospital producer price index. The quality adjusted inpatient hospital price indexes increase at an average annual rate of 0.19 to 0.26 percentage points less than the unadjusted index from 2010-2016.

1 Introduction

Quality change bias is a major challenge in constructing price indexes. To construct the official price indexes, government statistical agencies collect the prices of a fixed sample of goods over time. These measures can be biased if the quality of the goods is changing over time. In that case, the price index would reflect unobserved quality change in addition to pure price changes. Bias in the measure of prices will also bias measures of real output and

†Contact: matsumoto.brett@bls.gov

productivity. Holding quality constant is particularly challenging when measuring the price of services. For hospitals, publicly available measures of quality could potentially be used to quality adjust price indexes. The Department of Health and Human Services collects quality information from hospitals and produces a number of quality measures that are publicly available through its Hospital Compare project. The purpose of providing these quality measures is to increase transparency and accountability in the healthcare system. Another area where these quality measures could be useful is in quality adjusting hospital prices in the Consumer Price Index (CPI) and the Producer Price Index (PPI) to control for changes in hospital quality over time.

The change in hospital costs associated with the change in quality over time can be used to correct the price indexes for changes in quality. The challenge is in estimating the relationship between hospital costs and quality as this relationship will be complicated by patient selection. The first part of this paper estimates the causal relationship between the quality measures and hospital costs using Medicare claims data for 2009-2015. The causal relationship is identified using the instrumental variable technique of Doyle et al. (2015). They develop an instrument for hospital selection based on plausibly exogenous assignment to different ambulance companies (which have different preferences for hospitals). Doyle et al. (2015) use this instrument to estimate the relationship between Medicare reimbursement and patient outcomes and find that hospitals that have higher Medicare reimbursements have better outcomes. In a follow-up paper, Doyle et al. (2019) use the instrument to study the relationship between the Hospital Compare quality measures and outcome measures and find that the quality measures are associated with better patient outcomes.

Having accurate measures of hospital inflation is important given the size of the hospital industry. In 2018, nominal hospital spending was approximately $1.2 trillion and accounted for 5.8% of GDP. Given the size of hospital spending, even relatively small biases in the hospital price indexes could generate economically significant differences in the calculation of real hospital spending. I find a significant positive relationship between various measures of hospital quality and costs, and use this estimate to construct quality adjusted Producer Price Indexes for hospital inpatient services. Between 2010-2016, the quality adjusted indexes increase at an average annual rate of up to 0.26 percentage points less than the unadjusted index depending on the quality measures used. Over this time period, quality was generally improving, which caused the non-quality adjusted indexes to overstate hospital inflation. In time periods where hospital quality is declining, the unadjusted price indexes will understate hospital inflation.

The relationship between cost and quality is theoretically ambiguous. Even if improving quality requires costly investment, it can potentially lower costs by leading to a reduction in costly errors and medical mistakes. This has led some to hypothesize that the relationship is U-shaped. At low levels of quality, the cost savings from reducing errors exceeds the cost in the investment in improving quality. At higher levels of quality, the relationship reverses. There have been a large number of studies that have attempted to estimate the relationship between hospital costs and quality. Hussey et al. (2013) conduct a survey of the literature on the relationship between provider costs and quality and find the results are inconsistent. However, for hospitals, studies generally find a positive relationship between costs and quality. Examples of papers that find a positive relationship between costs and quality include Romley and Goldman (2011); Picone et al. (2003); Stargardt et al. (2013); Schreyogg and Stargardt (2010). However, many studies find no relationship or an inconsistent relationship depending on how cost and quality are defined (e.g., Kruse and Christensen (2013); Gutacker et al. (2013); Hakkinen et al. (2015)). Finally, some studies do find a negative relationship (e.g., Yasaitis et al. (2009)). A major problem that arises when trying to estimate the relationship between cost and quality is dealing with the non-random selection of patients to hospitals. Exploiting the random assignment to ambulance companies provides a method for overcoming the selection issue.

Medicare uses some of the measures to set reimbursement rates through the Hospital Readmission Reduction Program and the Hospital Value Based Purchasing Program. These programs penalize poor performing hospitals by reducing the Medicare reimbursement rates, so poor performing hospitals have a strong incentive to improve these quality measures. Ideally, this improvement would be accomplished through improving actual quality, though hospitals potentially have some ability to manipulate the measures through selective admissions and coding. The hospital quality measures did improve after the imposition of the programs tying reimbursement to quality, and there has been some debate as to how much the payment programs are responsible for this improvement and how much of the improvement in quality measures reflects actual improvements in quality. For example, Gupta et al. (2018) find that declines in readmission rates were accompanied by increases in mortality rates for heart failure during the start of the Hospital Readmission Reduction Program as hospitals trying to reduce readmissions rates may not readmit patients for whom it would be appropriate. However, Khera et al. (2018) find no evidence of an increase in mortality associated with the decline in readmission rates during the implementation of the Hospital Readmission Reduction Program for heart failure, AMI, or pneumonia admissions. Gupta (2021) examined the onset of the Hospital Readmission Reduction Program and found that about half of the improvements in the quality measures were due to actual improvements in quality.

The remainder of the paper is organized as follows. In Section 2, I provide an overview of the different data sources used in the analysis and discuss the construction of key variables. In section 3, I discuss the empirical strategy for estimating the relationship between hospital costs and quality. The results from this analysis are presented in section 4. Then, in section 5, I use the estimated causal relationship between hospital costs and quality to produce quality adjusted hospital price indexes. Section 6 concludes.

2 Data

This project combines data from a variety of sources. The following section details each of these data sources and discusses the construction of key variables used in the analysis.

2.1 Medicare Claims Data

The primary source of data for estimating the relationship between cost and quality measures is the 5% Medicare Limited Data Set, which contains all Medicare claims for a random sample of 5% of Medicare beneficiaries. The files are divided based on the type of provider (which determines how claims are reimbursed). This paper uses Medicare claims data for the years 2009-2015. Data on hospital admissions is contained in the Inpatient file, and ambulance billing data are contained in the Outpatient and Carrier files (non-institutional provider claims).[1] Professional services are billed as carrier claims regardless of where the service takes place, so physician services performed in a hospital are carrier claims. The inpatient and outpatient payments consist of the facility fee for the hospital. Basic enrollment and demographic information is contained in the Denominator file (which consists of a subset of variables from the Master Beneficiary Summary File).

2.2 Hospital Compare Quality Measures

Medicare collects a variety of data from hospitals in order to construct quality measures that are publicly released on the Hospital Compare website. There are different types of measures that potentially reflect different dimensions of hospital quality. The measures can also be for specific diseases or for specific units of the hospital. I consider several types of measures in this paper: patient satisfaction measures, outcome based measures, and process of care measures. Hospital Compare maintains an archive of past measure releases. The goal is to update the database quarterly, but past years have had as few as 3 and as many as 6 updates

[1] When the ambulance is owned by a hospital, the claim is processed as an outpatient claim. Otherwise ambulance claims are processed as carrier claims.

(with some duplicates). Not all measures are updated in each quarterly update. The length of time covered and the frequency of updates vary from measure to measure.

2.2.1 Patient Satisfaction Measures

The patient satisfaction measures are derived from the responses to the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) Survey. CMS conducts the survey on a random sample of all discharged patients (not just Medicare patients). Patients are asked to rate various aspects of their stay such as whether nurses and doctors communicated well and whether the hospital was clean. They are also asked to provide a global rating for the hospital and whether they would recommend the hospital to others. CMS processes the raw responses and makes adjustments for differences in the mix of patients across hospitals. The adjustment is based on a linear regression on patient demographic characteristics. The coefficients are used to predict the hospital's score if it had the average mix of patients for that quarter. A separate adjustment is made for the survey mode (e.g., telephone, mail, mixed). The mode adjustment is calculated from randomized experiments with the most recent conducted on a random sample of discharges from the first quarter of 2016. The patient satisfaction measures are updated quarterly and are measured over a 1 year window. The survey questions have been largely unchanged since these measures were first reported. One major change in the questionnaire is that the questions related to pain management and the associated quality measure was discontinued as there was concern that it could incentivize the prescription of opioid painkillers. The patient satisfaction measures are used in the Hospital Value Based Purchasing Program.

The patient satisfaction measures used in this paper are presented in table 1. For each question, the measure used is the percent of respondents choosing the most positive response.

2.2.2 Outcome Measures

The outcomes measures include risk adjusted 30 day rates of readmission and mortality for specific conditions. The initial measures were for acute myocardial infarction (AMI), heart failure (HF), and pneumonia (PN). They have expanded these measures to cover additional conditions (stroke, hip fracture, COPD) and procedures (CABG), as well as a readmission measure covering all discharges. In the analysis, I use the readmission and mortality measures for the three original conditions (AMI, HF, and PN). CMS risk-adjusts the measures by estimating a hierarchical generalized linear model on the set of eligible admissions (Medicare FFS patients with a principle diagnosis for one of the conditions). The probability of the outcomes is modeled as a function of patient demographic variables and a hospital specific

Table 1: Patient Satisfaction Measures

Question	Measure
How often did nurses communicate well with patients?	Percent answering "Always"
How often did doctors communicate well with patients?	Percent answering "Always"
How often did patients receive help quickly from hospital staff?	Percent answering "Always"
How often did staff explain about medicines before giving them to patients?	Percent answering "Always"
Were patients given information about what to do during their recovery at home?	Percent answering "Yes"
How often were the patients rooms and bathrooms kept clean?	Percent answering "Always"
How often was the area around patients rooms kept quiet at night?	Percent answering "Always"
Would patients recommend the hospital to friends and family?	Percent answering "Definitely, yes"
How do patients rate the hospital overall? (out of 10)	Percent rating 9 or 10

Notes: Patient satisfaction measures are constructed from the answers to the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) Survey. The measures are the percent giving the most positive possible response. The measures are risk adjusted based on patient characteristics and survey mode.

intercept term. The ratio of the expected number of readmissions/deaths using the hospital specific intercept coefficient to the expected number using the average of the hospital specific intercepts is multiplied by the national readmission/mortality rate. This process controls for differences in patient risk across hospitals. In developing the readmission measures, CMS attempts to identify and exclude planned readmissions. The outcome measures are updated annually and are based on hospital stays within a 2 and a half year window. The longer measure window means that the outcome measures are not as contemporaneous measures of quality as patient satisfaction or process of care measures. For the analysis, this paper uses the 30-day risk adjusted mortality and readmission measures for AMI, HF, and PN.

2.2.3 Process of Care Measures

Early measures were primarily process of care measures that counted the proportion of cases where a specified recommended action was taken in the course of the treatment of certain conditions (heart failure, acute myocardial infarction, and pneumonia). Many of the process of care measures for these 3 conditions have been discontinued in favor of the new measures (particularly outcome measures), and they became largely uninformative as compliance became nearly universal (a large number of hospitals were at or near 100% in these measures). The process of care measures are generally measured over a 1 year window and are updated quarterly. In addition to the disease specific measures, there are also process of care measures for specific areas of the hospital such as emergency room, surgery,

Table 2: Process of Care Measures

Condition	Measure
AMI	Heart Attack Patients Given Aspirin at Discharge
AMI	Heart Attack Patients Given PCI Within 90 Minutes Of Arrival
HF	Heart Failure Patients Given Discharge Instructions
HF	Heart Failure Patients Given an Evaluation of Left Ventricular Systolic (LVS) Function
HF	Heart Failure Patients Given ACE Inhibitor or ARB for Left Ventricular Systolic Dysfunction (LVSD)
PN	Pneumonia Patients Given the Most Appropriate Initial Antibiotic(s)
Surgery	Surgery patients who were given an antibiotic at the right time (within one hour before surgery) to help prevent infection
Surgery	Surgery patients whose preventive antibiotics were stopped at the right time (within 24 hours after surgery)

outpatient procedures, and the use of medical imaging. The process of care measures used in the analysis are presented in table 2.

2.2.4 Quality Measure Discussion

The analysis in this paper uses three categories of quality measures. Later, I consider which measures are most appropriate for the purposes of quality adjusting price indexes. It's important to consider what exactly the different measures are capturing and how they relate to each other. The outcome measures and process of care measures are meant to measure clinical quality. One criticism of the patient satisfaction measures is that may not be measuring clinical quality but instead pick up the effects of non-clinical factors like hospital amenities. For the purposes of quality adjusting price indexes, the potentially broader scope of the patient satisfaction measures beyond strictly measuring clinical quality is not a limitation as the quality adjustment would ideally control for any quality factors that impact the value of the service. Doyle et al. (2017) find a strong relationship between the various quality measures and outcomes, so the patient satisfaction measures also appear to capture clinical quality.

There is also the issue of how well the changes in these measures capture actual changes in quality. The use of the measures in determining hospital reimbursements does provide an incentive for providers to attempt to manipulate the measures. Some researchers have also raised the concern that changes in the quality measures over time simply reflect random variation rather than actual quality change. For example, Joshi et al. (2019) find that most of the decline in readmissions among hospitals with high readmissions rates at the start of the Hospital Readmission Reduction Program can be explained by regression to the mean. If changes in the quality measures over time reflect random variation, then we would expect there to be no relationship between the quality measures and costs. If we find a significant relationship between the quality measures and hospital costs, then the assumption is that changes in the measures are picking up changing characteristics of the hospital transaction that would ideally be held constant in the construction of a price index.

2.3 Medicare Impact Files

Hospital characteristics and cost data come from the Medicare Impact File, which contain hospital-specific values used in the Medicare Inpatient PPS reimbursement. Medicare eligible hospitals are required to submit detailed cost reports to Medicare every year. Medicare uses the cost data in setting reimbursement rates for hospitals through the Prospective Payment System (PPS). The hospital specific cost information is also used in certain situations to determine reimbursement amounts for the hospital. The cost reports are used to calculate cost to charge ratios (CCR), which can be used to calculate the hospital's cost for a given level of charges. The hospital CCRs are taken from the Impact File in the Final Rule of the Inpatient PPS annual update. The cost reports are subject to revision and are not considered final for a couple of years. The CCRs in the current year Impact File are based on the cost reports from 3 years prior. The CCRs from the Impact File can also be based on alternative data the hospital provides to CMS. In addition to the CCR, the Impact file contains other hospital-specific variables that affect reimbursements.

2.4 PPI Microdata

Microdata from the BLS's Producer Price Index (PPI) are used to construct the quality adjusted price indexes. Price data are collected from hospitals by the BLS staff. The sample of prices is selected in two stages. First, hospitals are randomly selected for inclusion in the Producer Price Index (PPI). Then, a set of services at that hospital are selected randomly for the index from recently submitted bills. In both stages the probability a given hospital or service is selected is proportional to its size (defined using revenues). The services are defined based on DRG codes (or similar service code) for inpatient and CPT codes for outpatient services. The measured price is the total amount received by the hospital from all payers for that service. All-payers are eligible to be included in the PPI. The CPI sample only includes out-of-pocket payments and insurance that involves the payment of a premium, so Medicaid and Medicare part A are excluded from the CPI. The characteristics of the service (DRG or similar code for the service, insurance company, length of stay, and other price determining characteristics) are fixed based on the selected bill for that service. In subsequent months, the hospital is instructed to provide the price for the same service with the same characteristics. Periodically, the sample is refreshed with new hospitals and services to reflect the changes in the types of procedures performed over time. Hospitals remain in the sample for approximately 8 years.

3 Sample

3.0.1 Inpatient Admission

To be included in the analysis sample, individuals must be enrolled in Medicare parts A and B at the time of the index admission and continuously enrolled for the 12 prior months. The individuals must be enrolled in part B in order to observe the ambulance claim. The sample is restricted to those who were originally eligible for Medicare due to old age.[2] The index admission must be an emergency admission for a non-deferrable condition where the individual arrived at the hospital via an ambulance (excluding transfers from other hospitals). The hospital sample includes acute care hospitals in the United States (excluding territories). The hospital admission is matched to the ambulance claim from either the Carrier or Outpatient files. The ambulance claims are restricted to those ending at a hospital that start at a non-hospital location.[3] The ambulance claims are excluded from the sample if there are multiple such trips on a single day. Longer distance ambulance trips are also excluded (defined as being over 50 miles). Finally, ambulance companies and hospitals with fewer than 5 observations are dropped.

The sample of index admissions goes from Jan 2010 through Oct 2015. Diagnoses for the 12 months prior to the index admission are used for risk-adjustment, so the first year claims data (2009) is used to define prior diagnoses for index admissions starting in 2010. The sample for the index admission ends in October 2015 with the change over to ICD-10 codes. The hospital cost for the index admission is calculated as the covered charges for the inpatient stay multiplied by the hospital's cost to charge ratio.

The empirical strategy, which is described in detail in the next section, exploits ambulance company preferences for specific hospitals and random assignment to ambulance companies. Transportation for deferrable conditions is more likely to reflect patient hospital preference rather than the ambulance company hospital preference. Conditions are defined based on the 3 digit ICD-9 code for the primary diagnosis code. Admissions for deferrable conditions are more likely to occur during the week. So, an admission for a non-deferrable condition should be as likely during a week day as during a weekend day. Non-deferrable conditions are defined as those conditions with a weekend admission rate closer to 2/7ths than that for hip fractures (which is generally considered to be a non-deferrable reason for admission) over

[2]Individuals can also qualify for Medicare due to disability or End Stage Renal Disease (ESRD).

[3]There are billing codes used to denote the location of pickup and the destination (residence, doctors office, hospital, etc.).

the sample period plus any of the original 29 conditions in Doyle et al. (2015) that failed to meet this criteria for the sample period. The non-deferrable conditions in Doyle are defined for an earlier period, but on a much larger sample of claims.

3.0.2 Constructing the Quality Measures

The quality measures are defined to have non-overlapping measure windows. Outcomes measures have a 2.5 year window. Since the entire sample covers about 5 years, there are 2 sets of outcome measures used. Because of the relatively long measure window, there is limited variation over time for each provider. For the first half of the sample, the providers have one set of outcome measure scores which update once. Patient satisfaction measures and process of care measures have a 1 year window. However, process of care scores are only available through 2014 as many of the measures were discontinued.

Individual quality measures are normalized to have mean zero and standard deviation of 1 in the sample. Measures where lower scores are better are multiplied by negative 1 so that positive values indicate above average quality. The individual measures are averaged to form a quality index for each of the measure types: a patient satisfaction index, a process of care index, and an outcome index (covering readmissions and mortality). Hospitals must not have any missing values for the measures that form the indexes. The index admission date is used to match the corresponding cost and quality data for the hospital stay. The quality measures are matched contemporaneously with the measure window, not the published measures on the Hospital Compare site at the time of admission, as the goal is to estimate the relationship between current quality and costs.

4 Empirical Strategy

4.1 Ambulance Referrals

The challenge in estimating the relationship between hospital quality and costs is patient selection across hospitals. Patient selection arises from the nonrandom assignment of patients to hospitals and will bias the estimated relationship between hospital cost and quality as the unobserved patient characteristics that affect costs are correlated with the hospital characteristics (including quality). One possible source of non-random hospital selection is patient preferences, but the selection can also arise from geographic sorting.

To control for patient selection, the causal relationship between hospital quality and cost is estimated using the ambulance assignment IV approach of Doyle et al. (2015). This instrument has also been used in a number of additional papers (Doyle et al., 2017, 2019;

Hull, 2020). The goal is to estimate the effect of hospital quality on costs, but hospital characteristics are endogenous, so this empirical strategy relies on exogenous assignment to ambulance companies. In areas serviced by multiple ambulance companies, the assignment is generally based on a rotation. So, when individuals call for an ambulance, the specific company assigned is as good as random.

Ambulance companies also have preferences for specific hospitals and some discretion on where to take a given patient. Ambulance preferences for specific hospitals could be based on pre-existing business relationships, and the ambulance could also be owned by the hospital. In terms of discretion, Medicare will generally reimburse trips to the nearest hospital from the point of pickup, although this is defined as being in the service area of the hospital. If the point of pickup is in the service area for multiple hospitals, Medicare will reimburse the trip to any of the hospitals.[4]

Ambulance company preferences are represented as the average hospital characteristics of all the patients the ambulance company transported in the market, denoted $Z_{a(i)}$. The index admission is excluded from the ambulance company average so it is a jackknife IV estimator (JIVE). The ambulance company market is defined using Hospital Referral Regions (HRR). Ambulance companies that operate in multiple HRRs are treated as different companies for the purposes of calculating the instrument. Then, this instrument is used to estimate the following equation:

$$log(C_i) = \alpha X_i + \beta H_{a(i)} + \gamma A_i + \delta^1 textbf1_{diagnosis} + \delta^2 \mathbf{1}_{year} + \delta^3 \mathbf{1}_{county} + \epsilon_i \tag{1}$$

Hospital costs (C) for the index admission i are a function of individual characteristics (X_i), hospital characteristics (H_i), ambulance characteristics (A_i), and indicators for the primary diagnosis code for the index admission, year, and county of residence. The individual characteristics include some basic demographic variables as well as characteristics of the admission such as age, race, gender, Charlson Comorbidity Index for the admission (defined using all diagnosis codes associated with the index admission), and the number of days in the hospital. The hospital characteristics include the quality measures as well as the number of residence to beds (measure of teaching intensity), the average daily census, the share of low income patients, and whether the hospital participates in Medicare's Inpatient Prospective Payment System (PPS). The hospital characteristics are endogenous, so they are instrumented using $Z_{a(i)}$, which is the average characteristic for the ambulance company. Ambulance characteristic include the trip distance (number of loaded miles), whether advanced life support used, whether it's emergency transport, and whether the ambulance is owned by a hospital. The

[4] Medicare will reimburse ambulance trips to a hospital outside of the service area if there is a specific medical need.

indicators for diagnosis are based on 3-digit ICD-9 codes and the geographic indicators are for county of residence. Standard errors are clustered at the hospital referral region level.

4.2 Limitations

The limitations of this method are outlined in detail in Doyle et al. (2015). Some additional limitations arise from using a different source of claims data. Doyle et al. (2015) use respondent identifiable data from a larger sample of total inpatient claims. The respondent identifiable data include the beneficiaries ZIP code of residence. Doyle et al. (2015) also use a larger sample as it contains 100% of inpatient claims. The data I use are the 5% claims Limited Data Set which only includes county of residence. I conduct all of the analysis at the level of county and HRR rather than at the level of ZIP code and Hospital Service Area (HSA) because of the smaller sample size and the lack of ZIP code of residence. This limitation could cause the IV estimation to fail to correct the selection bias as the identifying variation in ambulance company preferences could be operating at a much more local level.

5 Results

5.1 Summary Statistics

Table 3 presents the summary statistics for the characteristics of the index admission, hospital characteristics, and characteristics of the ambulance trip for the analysis sample. The analysis sample includes the index admissions with nonmissing quality data for any of the types of quality measures.

The index admission has an average cost of a little over $10,000 and an average length of stay of a little under 7 days. The hospital characteristics include the average daily census, the disproportionate share percentage (measures the share of patients that are low income), the ratio of residents to beds (measure of hospital teaching intensity), and whether the hospital is a part of Medicare's Inpatient Prospective Payment System (PPS). Also reported are the summary statistics for the ambulance trip. Almost all of the ambulance rides are emergency, which is not surprising given that the admissions are selected on the basis of being nondeferrable emergency admissions. About 72% of the rides used advanced life support, and the average number of loaded miles is 6.7. Only 6.6% of ambulance trips are in ambulances owned by a hospital.

Table 4 presents the summary statistics for the patient characteristics. One limitation of claims data is that there is generally minimal demographic data. The comorbidities are based on any diagnosis on any inpatient, outpatient, or carrier claim for the year preceding the

Table 3: Summary Statistics, Admission characteristics

Variable	Mean	SD
N	371,016	
Characteristics of Index Admission		
Total cost	10,129	12,129
length of stay	6.61	5.11
Charlson Comorbidity index	3.57	2.42
Hospital Characteristics		
Average Daily census	229.9	224.0
Disproportionate Share Percentage	0.286	0.149
Residents to beds ratio	0.101	0.184
Non-PPS hospital	0.144	0.351
Ambulance Characteristics		
Emergency	0.977	0.149
Advanced Life Support	0.718	0.450
Loaded miles	6.71	6.87
Hospital owned	0.066	0.248

index admission. The diagnoses are then grouped based on Hierarchical Condition Codes (HCC), which are diagnosis groupings used by CMS to risk-adjust payments in various programs.

In the analysis sample of index admissions, the average age is almost 82. The sample is about 62% female. Given the age of the sample, it is not surprising that females are overrepresented. Comorbidities are relatively common. Over 90% of the index admissions have a diagnosis of hypertension for some service in the year prior to admission. Other relatively common comorbidities (affecting over a third of the sample) include COPD, diabetes, dementia, and peripheral vascular disease.

5.2 Main Results

Table 5 presents the results of the effect of the quality indexes on total costs. There are 2 sets of results for each of the types of quality measures. For each type of quality measure the results are presented with and without hospital characteristics included. The main parameter of interest is the coefficient on the quality measure index. These results can be interpreted as the change in log cost from a one standard deviation increase in all of the quality measures (since each quality index is the average of a set of normalized measures). The regressions include the individual characteristics, ambulance characteristics, and characteristics of the index admission (see tables 3 and 4 for a list of these additional variables). Also included

Table 4: Summary Statistics, Patient Characteristics

Variable	Mean	SD
N	371,016	
Female	0.619	0.486
Age	81.93	7.81
Black	0.099	0.299
Hispanic	0.019	0.138
Comorbidity: Hypertension	0.902	0.297
Comorbidity: Stroke	0.170	0.376
Comorbility: cerebrovascular disease	0.199	0.399
Comorbidity: Dialysis	0.035	0.185
Comorbidity: Renal failure	0.232	0.422
Comorbidity: COPD	0.389	0.487
Comorbidity: Pneumonia	0.090	0.286
Comorbidity: Diabetes	0.455	0.498
Comorbidity: Malnutrition	0.115	0.320
Comorbidity: Dementia	0.376	0.484
Comorbidity: Paralysis	0.071	0.257
Comorbidity: peripheral vascular disease	0.473	0.499
Comorbidity: Cancer	0.044	0.205
Comorbidity: Injury	0.152	0.359
Comorbidity: Major psych	0.109	0.312
Comorbidity: liver	0.029	0.167
Comorbidity: drugs	0.036	0.186

Comorbidities are defined using Hierarchical Condition Codes (HCC) and are based on any diagnosis code on any Inpatient, Outpatient, or Carrier claim for the year prior to the index admission.

are indicators for the primary diagnosis for the index admission (3 digit ICD-9), year, and county of residence.

The effect of quality on costs is positive and significant for patient satisfaction and outcome measures. Including hospital characteristics causes the estimates to increase slightly for patient satisfaction measures but decrease slightly for outcome measures. The estimated effect of the process of care measures on hospital costs is positive and weakly significant when hospital characteristics are not included. However, these estimates becomes insignificant once hospital characteristics are included. The preferred specification includes the hospital characteristics since they are important for explaining cost differences across hospitals.

In the next section, I use the results for the patient satisfaction and outcome measures to produce quality adjusted price indexes. The process of care measure results are presented for comparison purposes. Since the process of care measures used in the analysis have largely been discontinued, they can not be used as the basis of a quality adjustment moving forward. Also, the results suggest that using process of care measures for a cost based quality adjustment is not appropriate since there is, at best, a weak relationship between process of care quality measures and costs.

6 Implications for Hospital Price Indexes

6.1 Quality Adjustment in the Price Indexes

The producer price index measures the average selling price received by producers. There are 2 forms of PPIs. One is based on grouping of items based on commodity classification and the other is based on the industry of the seller. The hospital industry PPI measures the price of all goods and services sold at hospitals. The hospital industry PPI is disaggregated on the basis of payer and the disease classification of the primary diagnosis. The commodity based hospital PPI is broken down into separate indexes for inpatient and outpatient services. Hospitals are randomly selected and the items that are priced are based on randomly selected recent bills. All price determining characteristics are recorded and in subsequent months, the hospital is asked the price for that exact bill. For inpatient services, the price quote includes all services received during the time of the admission when a bundled payment system is used (DRG or similar system).

There are a variety of methods for quality adjusting prices in price indexes. A quality adjustment is needed when a new version of a product comes out and the original version, which was in the index sample, is no longer available to price. The preferred approach in the PPI when the item changes (e.g. a new feature is added) and the original item is no

Table 5: IV Estimation results for the effect of quality on the log of hospital costs

	Quality Measure					
	Patient Satisfaction		Outcome		Process	
Quality Index	0.058 (0.007)	0.063 (0.006)	0.071 (0.009)	0.054 (0.009)	0.019 (0.010)	0.011 (0.010)
Hospital Characteristics included?	No	Yes	No	Yes	No	Yes
N	370,666	370,666	342,123	342,123	226,566	226,566

All regressions include the individual characteristics in table 4 and the characteristics of the index admission and the ambulance characteristics from table 3. The hospital characteristics include the variables in table 3. All regressions include indicators for the primary diagnosis of the index admission at the 3 digit ICD-9 level, indicators for year, and indicators for county of residence. The hospital characteristics and quality index are instrumented using the ambulance company average. Standard errors are clustered at the county level.

longer available is to make an explicit quality adjustment.[5] The explicit quality adjustment is usually based on the change in producer cost and respondents are asked for the cost information associated with the quality change. The quality adjustment involve adjusting the base price for the cost difference (plus a markup). Another form of explicit quality adjustment uses hedonic regressions. In hedonic regressions, the price is regressed on the price determining characteristics and the parameter estimates are used to adjust for changes in price determining characteristics for individual items over time.

If it is not possible to make an explicit quality adjustment, another option is to just replace the original item with the new version and treat the entire price difference between items as quality change. This can be done in two ways depending on if there is a period where both versions are observed. If both versions are observed in a given period t, then it's possible to use the overlap method. In the overlap method, the old base price is adjusted by the ratio of the price of the items in period t to form the new base price. The overlap method attributes the entire price difference between the items to quality. If there isn't a period where both products are available, then the price change is imputed based on similar products. For example if the old version is available until period t-1 and the new version becomes available in period t, the price change from period t-1 to t is imputed. The imputation method is also referred to as the link method.

The basic methods for addressing quality change in the CPI are similar, though the CPI is based on a cost-of-living framework. In a cost-of-living framework, the value of the quality adjustment is based on the consumer's willingness to pay for the change in quality. However, given the practical difficulty of obtaining willingness to pay estimates, the explicit quality adjustment in the CPI is often also based on changes in producer costs which will generally understate the value of the quality adjustment from the consumer's perspective. There are two sources of quality change bias in the price indexes. One arises if quality is changing but no adjustment is made. This bias is generally thought to lead the price indexes to overstate true price change due to unmeasured quality improvements. The other source of bias is if the quality adjustment made in the index differs from the true value of the quality change.[6]

6.2 Prior Quality Adjustment for Hospital Services in PPI

The PPI had previously used the process of care quality measures for AMI, HF, and PN to do limited quality adjustment. Specifically, they used these measures to perform a quality adjustment for hospital visits for these specific conditions. The process of care measures

[5] See Bureau of Labor Statistics (2013) for more detail on quality adjustment methods in the PPI.

[6] In particular, the link method is generally thought to overstate quality change and could lead to a negative bias in the CPI (Hulton, 1997; Moulton, 1996).

used were all reported as the fraction of successful outcomes (number successfully performed divided by the total number of possible times). These were combined to form disease specific measures by combining the numerators and denominators for the selected measures for a given condition. The quality adjustment assumed that a change in this composite quality measure has a proportional effect on the price (i.e., a 1% increase in the composite quality measure for the condition is associated with a 1% increase in price). The quality adjustment was made once a year, and 2008 was the first year it was performed. The process of care measures used in the adjustment have all been discontinued, and the last time the quality adjustment was made was in 2016. This quality adjustment had a minimal effect on the overall index as it only affected a small set of items though the quality adjustment for a specific item could be substantial. The major limitation of this method was that the relationship between costs and quality measures was assumed to be proportional as the true relationship was unknown.

6.3 Quality Adjustment Method

This section describes how the estimates of the relationship between hospital quality measures and costs are used to construct quality adjusted hospital price indexes using patient satisfaction and outcome measures. I use the microdata from the PPI inpatient hospital index. A quality adjustment is performed on all items in the PPI inpatient sample where hospital quality data are available. Not all hospitals are required to report quality information and Hospital Compare will not report scores for hospitals if the sample size is not sufficiently large. Also, it is not always possible to cleanly match hospitals in the PPI sample with hospitals in Hospital Compare.[7]

When estimating the relationship between cost and quality, the quality measures were matched to admissions based on the measure window in order to estimate the relationship between contemporaneous quality and costs. For quality adjustment purposes, the quality measures are incorporated based on when they are published. The measures are updated in the quality adjustment at the start of the quarter after they are published. Satisfaction measures are updated quarterly and the outcome measures are updated annually. Changes in the measures are smoothed over the entire period to avoid discontinuities in the measure update months. So, the change in the patient satisfaction index is smoothed over 3 months (quarterly update) and the change in the outcome index is smoothed over 12 months (annual update). The quality indexes are formed using the same values from the estimation sample to perform the normalization for each individual quality measure. Then, the patient satisfaction

[7]Hospitals are matched manually based on name and address.

and outcome quality indexes are formed as the average of the normalized individual measures.

When an item enters the sample, the price is recorded for the base period. I match this to the quality measure for the base period. The base period cost is estimated using the CCR corresponding to the base period (base period price times the base period CCR). In the PPI, the price index is calculated by dividing the item price in period t with the base period price in period t. For simplicity, assume that the base period is period 0. The base period price is adjusted over time to reflect changes in the item that makes the item price non-comparable over time. So, the quality adjustment occurs through an adjustment to the base period price. The value of the quality adjustment (denoted VQA) is calculated from the regression results in the previous section. Let β^Q denote the parameter on the quality index from the hospital cost regression, Q_t denote the quality index in period t, and Q_0 denote the quality index in the base period. Then, the value of the quality adjustment is given by:

$$VQA_t = (1 - exp(\beta^Q * (Q_t - Q_0))) * Cost_0 * \frac{1}{CCR_t} \tag{2}$$

The change in quality indexes, regression parameter, and base period costs are used to estimate the change in cost associated with the change in quality. The CCR in period t is used to markup this cost change to generate the value of the quality change. The VQA is used to create an adjusted base period price for period t. Let BP denote the base price, then the adjusted base price in period t is given by:

$$BP_t^{adj} = BP_t * \frac{P_t}{P_t - VQA_t} \tag{3}$$

6.4 Quality adjusted price indexes

Figure 1 presents the results from the quality adjustment using satisfaction measures and using outcome measures from December 2009 to December 2016. The outcome measures were not published in a comparable format at the start of the sample period. The values from the year 2009 update are used as the base period quality, so the first change in the quality measures (and first quality adjustment) does not occur until the values from the 2010 annual update are incorporated (in Oct 2010). The quality adjustment causes the index to grow less slowly over the time period. The unadjusted index grows by about 16 percent over the period, while the quality adjusted indexes grow by 14 and 14.5 percent for outcome measures and satisfaction measures respectively. The quality adjustment using outcome measures has a minimal impact until 2014. The impact of the patient satisfaction quality adjustment occurs earlier and reaches a peak by around 2014.

From Dec 2009 - Dec 2016, the hospital inpatient PPI grew at an average annual rate of

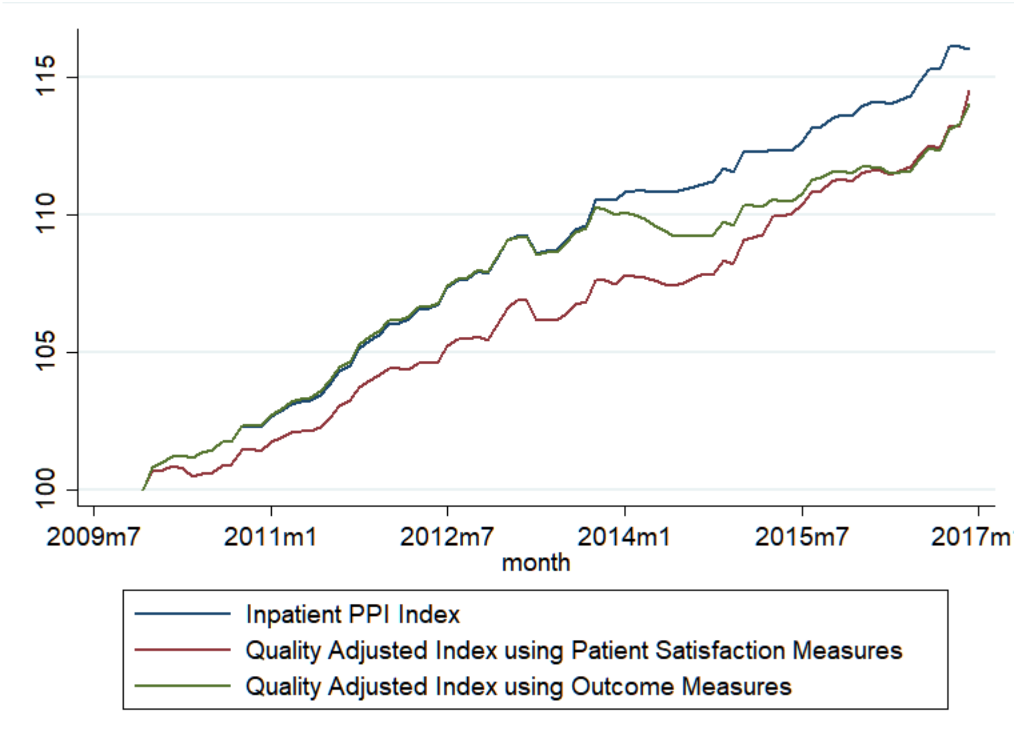

Figure 1: Quality adjusted and non-quality adjusted inpatient PPI (Dec 2009 = 100)

2.15 percent while the quality adjusted index using outcome measures grew at an average annual rate of 1.89% and the quality adjusted index using patient satisfaction measures grew at an average annual rate of 1.96%. The impact of quality adjustment is 0.19 percentage points lower average annual inflation using the patient satisfaction measures and 0.26 percentage points lower using the outcome measures over this period. The outcome measure quality adjustment does not begin until the end of 2010, so a better comparison is over the Dec 2010-Dec 2016 period. Over this period, the hospital inpatient PPI increases at an average annual rate of 2.13 percent while the quality adjusted index using outcome measures increases at an average annual rate of 1.82 percent (for a difference of 0.3 percentage points). Although the quality adjustment may seem to have a relatively small impact on the overall inpatient hospital price index, even small differences in hospital inflation can meaningfully impact the calculation of real hospital output given the size of the hospital sector.[8]

The impact of the quality adjustment is not constant over this time period. Figure 2 shows the year over year percentage change in the quality adjusted indexes and the unadjusted index. The patient satisfaction based adjustment had a larger impact early in the period compared to the adjustment using the outcome measures. In late 2015, the satisfaction measure based quality adjustment actually leads the quality adjusted prices to rise faster

[8] In 2016, nominal hospital spending was $1.1 trillion and accounted for 5.8% of nominal GDP.

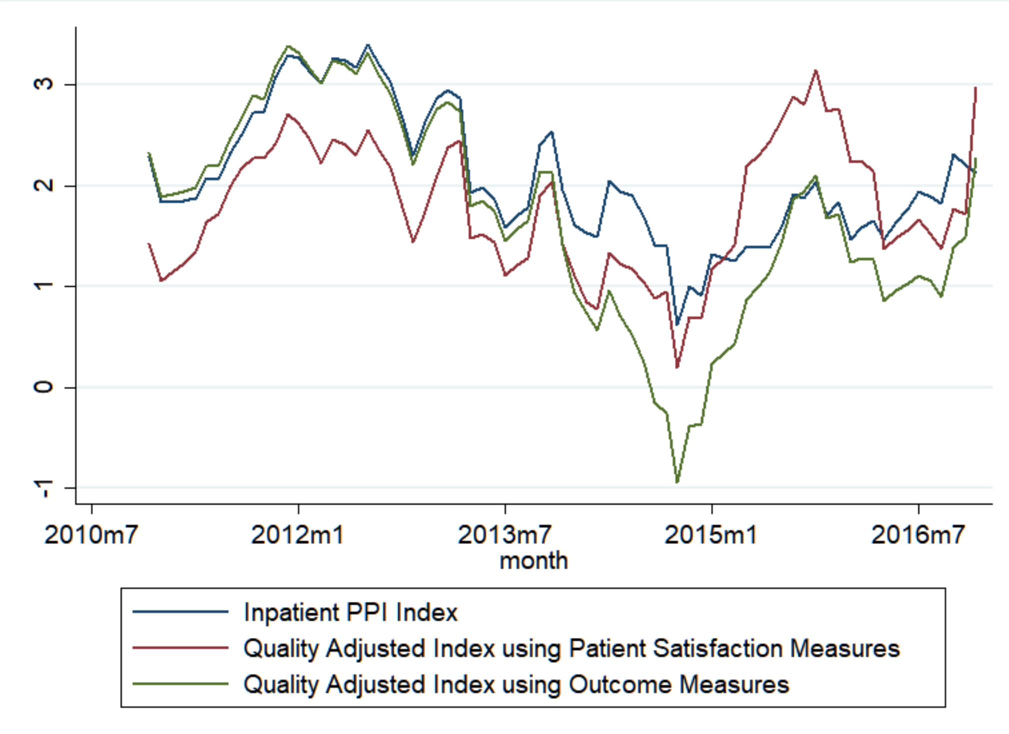

Figure 2: Quality adjusted and non-quality adjusted inpatient PPI: 12 month percentage change

than the unadjusted prices. The outcome measure based quality adjustment has minimal impact early, but starts to have a bigger impact in 2014 and beyond. Overall, the impact of quality change on prices varies over time and quality change is not monotonic. It is possible for hospital quality to decline over periods of time which causes the quality adjusted indexes to rise faster than the unadjusted indexes.

7 Conclusion

In this paper, I develop a feasible approach for implementing a cost-based quality adjustment for hospital price indexes. The approach uses publicly available measures of hospital quality. The challenge is in estimating the relationship between the quality measures and costs. In order to overcome the challenge of patient selection, I use the ambulance assignment IV of Doyle et al. (2015) that exploits plausibly exogenous assignment to ambulance company. This instrument is used to estimate the relationship between hospital quality and costs using Medicare claims data. The relationship between hospital quality and costs is then used to construct quality adjusted hospital price indexes using microdata from the PPI. Overall, from Dec 2009 - Dec 2016 the quality adjustment lowers the average annual hospital inpatient inflation rate by between 0.19 and 0.26 percentage points. The impact is not constant over time. There are periods where the impact of the quality adjustment is greater, and there are also periods where quality declines which leads the quality adjustment to show a larger price increase than the unadjusted index.

Quality measures are available for other types of medical services, so it may be possible to use them to quality adjust other medical price indexes. The challenge in extending the methods of this paper to other medical services is that the empirical strategy for estimating the relationship between cost and quality (ambulance assignment) is specific to inpatient hospitals.

References

Doyle, J., J. Graves, and J. Gruber (2017). Uncovering waste in us healthcare: Evidence from ambulance referral patterns. *Journal of Health Economics 54*, 25–39.

Doyle, J., J. Graves, and J. Gruber (2019). Evaluating measures of hospital quality: Evidence from ambulance referral patterns. *Review of Economics and Statistics 101*(5), 841–852.

Doyle, J., J. Graves, J. Gruber, and S. Kleiner (2015). Measuring returns to hospital care: Evidence from ambulance referral patterns. *Journal of Political Economy 123*(1), 170–214.

Gupta, A. (2021). Impacts of performace pay for hospitals: The readmissions reduction program. *American Economic Review 111*(4), 1241–1283.

Gupta, A., L. Allen, D. Bhatt, M. Cox, A. DeVore, P. Heidenreich, A. Hernandez, E. Peterson, R. Matsouaka, C. Yancy, and G. Fonorow (2018). Association of the hospital readmissions reduction program implementation with readmission and mortality outcomes in heart failure. *JAMA Cardiology 3*(1), 44–53.

Gutacker, N., C. Bojke, S. Daidone, N. Devlin, D. Parkin, and A. Street (2013). Truly inefficient or providing better quality of care? analysing the relationship between risk-adjusted hospital costs and patients' health outcomes. *Health Economics 22*, 931–947.

Hakkinen, U., G. Rosenqvist, T. Iverson, C. Rehnberg, and T. Seppala (2015). Outcome, use of resources and their relationship in the treatment of ami, stroke and hip fracture at european hospitals. *Health Economics 24*, 116–139.

Hull, P. (2020). Estimating hospital quality with quasi-experimental data. Working Paper.

Hulton, C. (1997). Quality change in the cpi. *Federal Reserve Bank of St. Louis Review 79*(3), 87–100.

Hussey, P., S. Wertheimer, and A. Mehrotra (2013). The association between health care quality and cost. *Annals of Internal Medicine 158*(1), 27–34.

Joshi, S., T. Nuckols, J. Escarce, P. Huckfeldt, I. Popescu, and N. Sood (2019). Regression to the mean in the medicare hospital readmission reduction program. *JAMA Internal Medicine 179*(9), 1167–1173.

Khera, R., K. Dharmarajan, Y. Wang, Z. Lin, S. Bernheim, Y. Wang, S. Normand, and H. Krumholz (2018). Association of the hospital readmissions reduction program with mortality during and after hospitalization for acute myocardial infarction, heart failure, and pneumonia. *JAMA Network Open 1*(5), 1–15.

Kruse, M. and J. Christensen (2013). Is quality costly? patient and hospital cost drivers in vascular surgery. *Health Economics Review 3*(22), 1–13.

Moulton, B. (1996). Bias in the consumer price index: What is the evidence? *Journal of Economic Perspectives 10*(4), 159–177.

Picone, G., F. Sloan, S. Chou, and D. Taylor (2003). Does higher hospital cost imply higher quality of care. *Review of Economics and Statistics 85*(1), 51–62.

Bureau of Labor Statistics (2013). Quality adjustment in the ppi. https://www.bls.gov/ppi/qualityadjustment.pdf.

Romley, J. and D. Goldman (2011). How costly is hospital quality? a revealed-preference approach. *Journal of Industrial Economics 54*(4), 578–608.

Schreyogg, J. and T. Stargardt (2010). The trade-off between costs and outcomes: The case of acute myocardial infarction. *Health Services Research 45*, 1585–1601.

Stargardt, T., J. Schreyogg, and I. Kondofersky (2013). Measuring the relationship between costs and outcomes: The example of acute myocardial infarction in german hospitals. *Health Economics 23*(6), 653–669.

Yasaitis, L., E. Fisher, J. Skinner, and A. Chandra (2009). Hospital quality and intensity of spending: Is there an association? *Health Affairs 28*(4), w566–572.