Stephen T. Parente (USA)

Development of a Medical Productivity Index for health insurance beneficiaries

Abstract

In this paper, a Medical Productivity Index (MPI) is proposed as a metric to capture the value of care received by patients from medical providers. For the health sector, such a metric could address the growing concern that medical care expenditures are sapping the economic vitality of a nation if these outlays show a productivity gain. The two primary components of the MPI are a measure of health outcomes and a measure of medical care effort. The MPI is applied to a national sample of Medicare 2007-2009 claims data. Application of the MPI shows both a cyclical and long-term trend in medical care productivity. There are substantial regional variations in MPI as well. Extensions of the MPI could provide disease and insurance contract specific sub-sector component comparisons in future applications. The use of MPI to retrospective claims and contemporary claims data provides a technology to track changes in medical productivity to gauge the impact of future health reform and medical technologies as well as an aging society to patients and the health care industry.

Keywords: health insurance, productivity, health economics, physician payment, Medicare.

Introduction

The consistent rise in health care expenditures has been a source of growing concern across the developed world. In the United States, the recently passed Patient Protection and Affordable Care Act (PPACA) was developed in part to reduce the year over year increase in health insurance expenditures that is well above the general inflation rate. Economists have stated for decades that the increase has been driven by an aging population and the pace of new medical technologies that can be more expensive per capita despite being more effective than existing therapies for a condition (Weisbrod, 1991; Pauly, 2008; Pauly, 1986; Culter and McClellan, 2001).

With trillions of dollars in publicly financed health insurance projected, policymakers will benefit from knowing the value of this expenditure to society on an ongoing basis (Kleinke, 2001; McClellan, Tunis, Engl, Med, 2005). Specifically, a metric of the welfare gains of care financed by insurance on individual consumers is needed. In this paper, a Medical Productivity Index (MPI) is proposed as such a metric to capture the value of care received by patients from medical providers. Measuring consumer productivity from insurance financed medical care has great appeal. In general, productivity takes into account the output of goods or service as well as the inputs required to create such an output. For example, in the manufacturing sector, productivity indices inform us of the changes in industrial output given labor or capital inputs. For the health sector, such a metric could address the growing concern that medical care expenditures are sapping the economic vitality of a nation if these outlays show a productivity gain. The two primary components of the MPI are a measure of health outcomes and a measure of medical care effort. The final attribute of the MPI is that it can be applied to health insurance claims data retrospectively and on an ongoing basis with manageable computational effort. This will allow historic comparisons with current trends as health reform legislation is executed to slow the rate of health care expenditure in the United States.

This paper proceeds as follows. First, the comparative landscape of national health care metrics and consumer indices relevant to health care are examined to identify what novel features an MPI should deploy. Second, the conceptual model of an MPI is discussed. Third, the health insurance data and methodological approach to create the MPI are presented. Fourth, the results of applying the MPI to a national sample of Medicare claims are shown. Finally, the implications, caveats and extensions of the MPI are discussed.

1. Background

Economic indices have a long history in market economies. They range from stock market indices of prices to weekly, monthly, quarterly and annual surveys of consumer and purchaser behavior. The first stock index was created by Charles Dow in 1884 through an average of 11 liquid stocks[1]. The most famous of market indexes is the Dow Jones Industrial Average (DJIA) and its sub-sector industry components ranging from manufacturing to banking to health care. Thanks to advances in telecommunications and computers, the DJIA and similar indices managed in other nations such as the FTSE100 (UK), DAX (Germany), Hang Seng (China) and Nikkei (Japan) are now a real-time metrics of market activity and perceived value of companies in a market economy.

[1] http://www-stat.wharton.upenn.edu/~steele/Courses/434/434Context/Indices/DowFlaws.pdf.

In contrast to stock market indices, productivity indices are not generally real time averages. This is largely due to lags in the time for collection of data that vary significantly in their availability from seconds to days. In addition, these indices need to be adjusted by their certified organizations to be properly calibrated. The Chinese have been credited with using the first productivity index 40 centuries ago[1]. Productivity indices are commonly used for agriculture and energy production. Typically indices are ratios, with products produced in a given time interval as the numerator and labor, or capital input, as the denominator. In general, they refer to industry or firm specific output.

Another form of index used in market economies is consumer sentiment. The most widely used of this type of index is the Michigan Consumer Sentiment index. The consumer confidence measures were devised in the late 1940's by George Katona at the University of Michigan. They have now developed into an ongoing, nationally representative survey based on telephonic household interviews. The Index of Consumer Sentiment (ICS) is comprised of these interviews. The Index of Consumer Expectations (a sub-index of ICS) is included in the Leading Indicator Composite Index published by the U.S. Department of Commerce, Bureau of Economic Analysis[2]. The index represents an average of individual consumer responses from a set of telephone surveys.

In the healthcare industry, the most commonly used economic index is the medical price component of the Consumer Price Index (CPI). The US Department of Labor's CPI program produces monthly data on changes in the prices paid by urban consumers for a representative basket of goods and services[3]. Medical care goods and services are measures as part of the CPI program. The medical CPI has recorded the historic pace of medical care prices compared to general inflation for over three decades. Over that period, the medical care inflation rate has been significantly greater, sometimes double, the general inflation rate in the United States. While the medical CPI records the increase in prices, there has been no routine recording of the welfare gain from continually rising prices in medical care.

Within the last twenty years, there have been periodic measures of national health care utilization trends and medical care outcomes. The most cited of these measures are produced by the Dartmouth Atlas and are based on Medicare health insurance claims data[4]. The two most widely cited uses of the Dartmouth Atlas are the recording of region specific differences in health care utilization per capita and the hospital discharge mortality rate per capita. The range of geographic differences in health care and care-related mortality have driven national efforts in the United States to systematically improve clinical outcomes through the development of clinical guidelines, and more recently, with federally financed health improvement pilots from the Medicare program as part of PPACA legislation. One of the major advances from the Dartmouth Atlas is the acceptance of routine reporting of health care utilization and outcomes based on health insurance claims data. Another feature of the Dartmouth Atlas is the summary of consumer/patient-level information to aggregate measures by region for comparison. The Dartmouth Atlas successfully demonstrates a routine use of health insurance claims data to measure health outcomes.

Based on a review of effective economic indices, the attributes selected to develop the MPI were the routine and consistent reports from stock indices, the relationship between input and output over a defined period of time from a productivity index and the consumer to regional aggregation possible from health insurance claims data. These features constitute the requisite parameters for the MPI conceptual model development process.

2. Conceptual model

The conceptual model behind a productivity metric is based on an economic production function for health outcomes at a patient level. For such a model we introduce the following expression:

$$H_{i,t} = f(M_{i,t-1}(l_{i,t-1}, k_{i,t-1}), P_{i,t}(G_{i,t}, AG_{i,t}, S_{i,t})), \qquad (1)$$

where H is the health level for person i and time period t, M is the medical care received by person i and time period t-1, l is the medical care labor by physicians for M at time period t-1 for patient i, k is the capital and other attributes part of M at time period t-1 for patient i, P is the patient i underlying health status at time period t, G is the genetic predisposition of patient i at time period t, AG is the age and gender of patient i at time period t, S is the unexpected health shock to patient i at time period t.

In equation (1), a patient's health at time t will be a function of the medical care they receive, M, and the patients underlying health status, P. One significant feature of this model is that the medical care input into a patient's health production function is lagged by a prior time period (e.g., one quarter of a year) than the output measure of health. Measured in the

[1] http://wvuscholar.wvu.edu:8881//exlibris/dtl/d3_1/apache_media/L2V 4b GlicmlzL2R0bC9kM18xL2FwYWNoZV9tZWRpYS82MDMy.pdf.

[2] http://www.sca.isr.umich.edu/.

[3] http://www.bls.gov/cpi/.

[4] http://www.dartmouthatlas.org/.

current time period, health contributing, personal attributes such as genetics (G), age and gender (AG), and any unexpected health shocks (S) such as a car accident or major acute care event are underlined. The rationale behind lagging the medical care input is that the impact of medical care can take time to have an appreciable impact on health and would be measured prior to the contemporary period. Therefore, for a patient who suffered a mild heart attack, measuring the health outcomes after the medical care expended for the heart attack would be a more valid way of measuring the last result of medical care on health.

The expression for health contained in equation (1) provides the conceptual foundation for the creation of a medical productivity index. In equation (1) the unit of analysis is the patient at time t. The MPI represents the average of patient health at time t divided by the medical care effort applied to that patient at time t-1 and is represented as equation (2).

$$MPI_t = \frac{\sum H_{it}}{M_{it-1}} \text{ over all patients } i. \quad (2)$$

Note in equation (2) that the expression for a patient's health status in period t (P_{it} in equation (1)) has been omitted from the MPI. The rationale is that much of a person's genetic disposition, gender and age would not vary sufficiently from t-1 to t, creating an impact on H_{it}. Also, a person's health shock (S_{it}), as a separate measurement, would likely be too highly correlated to H_{it} for practical inclusion.

In creating the MPI expressed as equation (2) the goal was to produce a statistical metric with four essential attributes. First, it was designed to be used on a timely basis through the use of recent secondary data sources with the potential of being updated in real time. The second attribute was for the MPI to be repeatable over multiple periods of time in a consistent manner. For example, the same MPI could be used for measurement of medical productivity over several years using existing data sources. The third attribute was for the MPI to be able to decompose by the two component parts (contemporary health outcomes and medical care provided) to understand which of the two may be driving changes in productivity. The final attribute was for the MPI to be applicable to different patient populations identified by region, medical condition, treatment and demographics. This is accomplished by running the MPI on a subset of patients with a given population characteristic such as the presence of diabetes or receipt of a knee replacement within the last year.

3. Data

The data used for the application of the MPI presented is the Medicare National Claims History File (NCH). Specifically, a random sample of 5 percent of all beneficiaries of the Medicare population in the claims data for services in calendar years 2007 through 2009 is used for this analysis. Only Medicare beneficiaries 65 years of age or older were included in the analysis. The data used was de-identified and provided for health care efficiency analytic prototype development. There was no personal health information required for use in this analysis. The value of using the Medicare data is that its format is commonly found in other US health insurance data architectures, so that the MPI could be repeatable with Medicare, but also in the Medicaid and commercial private insurance populations as well (Parente, Weiner, Garnick et al., 1995). Furthermore, the public policy importance of the Medicare data is underscored by the impact of the program on future US debt projections if left on its current trajectory for health care expenditures.

To increase the opportunity to apply the MPI to health insurance claims data sources, only a handful of variables from the data were used to generate the MPI. These include the quarter of service, the encrypted beneficiary ID, CPT4/HCPCS procedure code, patient geocodes (e.g., zip, county and state) and ICD9 diagnosis code. Beneficiary age and gender are also extracted from the data, but only for subset analysis as necessary. All of these variables can also be found on commercial insurance data from large private insurers as well as Medicaid data. Furthermore, there is great consistency in the use of this data with the only major future change anticipated in 2013 being the transition from ICD9 to ICD10.

The other database required for this analysis is the Resource Based Relative Value Scale (RBRVS) information published for the Center for Medicare and Medicaid Services web site. The RBRVS database is an annual publication of the relative physician, capital and office components embedded in the approximate cost of the application of a medical procedure. Each CPT4/HCPCS code that is present on the physician claims for services rendered as part of the NCH can be mapped to a component of the RBRVS system to identify the amount of physician labor used per procedure. This combination of the RBRVS and NCH databases provides the essential information necessary to derive the medical care input component of MPI. A detailed description of how the RBRVS is executed follows.

3.1. Medical Productivity Index estimation approach. Creation of the MPI from health insurance claims data required two essential components: a patient level measure of health and a patient level measure of medical care expended. The measure of health from claims data is developed using the John Hopkins Adjusted Clinical Group system (ACG)

(Weiner, Starfield, Steinwachs, Mumford, 1991). The ACGs use only patient ID, diagnosis code, age and gender to create a vector of 34 binary variables representing different levels of illness burden. To create a metric of health status, a patient level summation of the vector of 34 Adjusted Diagnostic Groups (ADGs) was used to calculate overall illness burden during a contemporary 90 day window. The inverse of this illness burden measure derived from ADGs is used to measure a patient's health level at time *t*.

The medical care input metric is developed by identifying the RBRVS physician work effort associated with the CPT/HCPCS procedures performed by medical providers and summing up for 90 days prior to the 90 day health level metric developed above. In practice, this means the second quarter of one year's index represents the average patient ratio of health (calculated in the second quarter) over the summed medical care physician labor effort per patient as calculated in the first quarter of the year. It is important to note that the physician labor component is from effort in all settings a physician operates in, including inpatient hospital, outpatient hospital, office and long-term care settings.

The index presented in this paper is based on 12 quarters of data with the last quarter representing the health of patients in the 4th quarter of 2009 given the medical effort of 3rd quarter of 2009. Although the index is used in this context on only a quarterly basis, a moving avarice updated version could be updated in real time moving the 90 day windows minute by minute based on claims data received and analyzed in real time or close to real time.

The index is based on the claims data of 1.875 million beneficiaries receiving care from 2007 to 2009. The beneficiary had to be eligible for Medicare for at least four quarters. The index calculation allows for new Medicare beneficiaries at age 65 when they become eligible, as well as those who will exit traditional Medicare to join private insurance operated by a Medicare Advantage health plan or die within the program. In constructing the database, the decision was made for the first use of the index to have an unbalanced number of participants per quarter since that would be the likely use of the index technology if it were eventually applied as a real time reporting metric.

4. Results

The results produced for application of the index on Medicare data are presented in Figures 1 through 3 and Table 1 (see Appendix). Figure 1 presents the average MPI over three years. The values range from 0.109 to 0.120 during that period. There is a distinct seasonal pattern where MPI is the lowest in the first half of the year and then picks up in the second half of the year. The other major observation is the general decreasing trend over time. This appears to be mostly the large drop in MPI between the 4th quarter of 2008 and 1st quarter of 2009. This drop reflects a 5.9% drop in MPI. The index decreases a percentage point more by the 2nd quarter of 2009, which also coincides with the worse period of the US's recent economic recession. By the 3rd quarter of 2009, MPI has increased to 0.110, but not nearly to the level of the prior 3rd quarter of 2008 when MPI equaled 0.117.

In Figure 2, a comparison of the MPI by state for the 3rd quarter of 2009 is presented. The MPI values for each state are categorized by 3 equally sized quintiles to represent a high, medium and low placement on the MPI distribution. The states with the highest MPI are in the Southern part of the country and include Georgia, Florida, Alabama, Tennessee and North Carolina. Large states Texas, California, and New York also have low MPI. States with moderate MPI are mostly concentrated in the Midwest and Northeast. Similar to the results from the Dartmouth Atlas, there is significant variation in health metric reported by the states.

Figure 3 presents a decomposition of the component parts of the MPI, physician labor at period *t*-1 and health level at period *t*. This result is for the 3rd quarter of 2009. MPI, physician labor and health level are all normalized for a 0 to 100 scale in the figure. The state with the lowest MPI is New Jersey. New Jersey has the lowest health level and the highest level of prior quarter physician effort. In contrast, the state with the best MPI is Wyoming. This appears to be driven by the highest health level among the states and one of the lowest uses of physician services in the prior period.

To get an exact measure of the differences shown graphically in Figure 3, Table 1 presents the normalized metrics for MPI, physician level effort and health index. The Table is sorted in descending order from the highest recorded MPI in the 3rd quarter of 2009 to the lowest one. States with average physician effort like Georgia end up with a lower MPI because of lower health outcomes compared to other states with similar physician effort. This state specific decomposition provides additional insight as to which of the two components of the MPI might be driving a state's resulting index and order among other states.

Conclusion

The creation of the MPI provides a new metric to examine the value health insurance beneficiary's receive for their medical care. The central finding of this paper is that health insurance claims data can be used to generate a recurring metric for medical productivity. Since the metric can be generated at a beneficiary level, creating subset populations by disease or treatment provided, as well as region or demographics, should be a fairly straightforward process. An interesting policy-related finding from application of the MPI is potential evidence of substantial reduction in productivity related to the economic recession in the United States during 2009. Whether the decrease in MPI was the product of genuine reduction in health will require additional analysis with 2010 data in a future analysis.

As the MPI is considered for use with other databases, it will be important to understand how to interpret the statistics. For example, how should a state like Wyoming interpret a high MPI? A high CPI could reflect either a combination of low medical care effort /good patient health status or average medical care effort/great patient health status. Likewise a low MPI could reflect either relatively high medical care use/average patient health status or average medical care use/low health status. If a state is concerned by their MPI, they can design interventions to improve either patient health status or lower wasteful medical care use when compared to peer states.

Caveats

There is no perfect metric for measuring economic activity, and this MPI has three notable caveats. The first caveat is the accuracy of an arbitrary 90 day lagged window reflecting medical effort. 90 days is the number of days in a quarter and so it has a logical relationship with economic indicators that are reported quarterly such as business earnings. In addition, 90 days are used for Medicare's proposed bundled payment rules to construct an episode of care. This suggests events that are a part of major hospital event such as knee replacement need only 90 day post hospital discharge to capture the medical effort used to improve a person's health status. If a 90 day window introduces a bias, at least it will be consistently applied to all patient care. Over time the window can be adjusted following an understanding of the extent to which a 90 day look back window for medical care effort contributes to a biased estimate of productivity.

Another caveat is the lack of accounting for case-mix differences across patients. This problem is accounted for at a gross level by measuring illness burden by ACG when generating the health status metric for the numerator of the CPI. In future applications of the MPI, one can condition the use of the index on a subset of patients with a major chronic or acute condition to see how the MPI would vary for the population. The MPI is designed for these subset analyses to be completed on an ongoing basis. For example, one could compare overall medical care productivity for the over 65 aged population to a subset of the same senior population who are 5 year cancer survivors or beneficiaries with two or more chronic conditions. Although both subset populations can be quite expensive to treat, the medical profession prides itself on high quality outcomes for populations such as these with significant health needs.

The final caveat is the application of MPI exclusively on traditional Medicare claims data. Given the successful application of the MPI to Medicare, there is no reason why it cannot be applied to other claims databases from private insurers and other public insurers such as Medicaid. If another nation has a similar RBRVS physician effort recording system by procedure code, the technology should be exportable as well to compare productivity across nations. For international comparison, a physician effort equivalent could be developed using ICD10 surgical and treatment procedure codes in the future and joins the rest of the industrialized world in the use of the most recent diagnostic and procedure coding system.

Extensions

Application of the MPI shows both a cyclical and long-term trend in medical care productivity. Many extensions of the MPI can be developed in a fashion similar to that of the different industry-specific Dow Jones Stock Indices such as one for transportation and another for energy. A focus on several major condition-specific applications of the MPI can be easily constructed. One can focus on patients with specific conditions and then aggregate their component health status and medical effort into a condition-specific MPI. Likely conditions to model include chronic diseases, diabetes, depression, congestive heart failure, and acute illnesses such as injury from an accident. Besides condition, one can also focus on a set of people receiving certain medical technologies. For medical device manufacturers that clearly identify health status improvement as one of their central aims, a subset of MPI by patients who received cardiac or orthopedic technologies could be useful.

Applying MPI by different health insurance programs would be useful for public policy analysis. For example, MPI could be applied to patients who receive their Medicare benefit though Medicare Advantage private health insurance programs. This analysis would provide insights into the different medical productivity the two programs achieve for seniors. One concern about Medicare Advantage is that participating health plans are paid extra for medical management of complex diseases but little value (at the margin) is recorded when the program is compared to traditional Medicare. A comparison of the MPI for both populations over time as well as by region would be valuable to see, in fact, which program is actually more useful to patients.

Application of the MPI shows both a cyclical and long-term trend in medical care productivity. There are substantial regional variations in MPI as well. Extensions of the MPI could provide disease and insurance contract specific sub-sector component comparisons in future applications. The use of MPI to retrospective claims and contemporary claims data provides a valuable technology to track changes in medical productivity to gauge the impact of future health reform, medical technologies and an aging society to patients and the health care industry.

References

1. Culter D. and McClellan, M. (2001). "Is Technological Change in Medicine Worth It?" *Health Affairs*, 20 (5), pp. 11-29.
2. Kleinke, J.D. (2001). "Price of Progress: Prescription Drugs in the Health Care Market", *Health Affairs*, 20 (5), pp. 43-60.
3. McClellan M.B. and Tunis S.R., N. Engl J. Med (2005). "Medicare coverage of ICDs", 352, pp. 222-224.
4. Pauly, M.V. (2008). "Markets Without Magic: How Competition Might Save Medicare", AEI Press.
5. Pauly M.V. (1986). "Taxation, Health Insurance, and Market Failure", *Journal of Economic Literature*, 24 (6), pp. 629-675.
6. Parente S.T., Weiner, J.P., Garnick, D., Richards, T. Fowles, J., Lawthers, A., Chandler, P. and Palmer, R.H. (1995). "Developing a Quality Improvement Database Using Health Insurance Data: A Guided Tour with Application to Medicare's National Claims History File", *Journal of American Medical Quality*, November.
7. Weiner, J.P., B.H. Starfield, D.M. Steinwachs and L.M. Mumford (1991). "Development and Application of a Population-Oriented Measure of Ambulatory Care Case-Mix", *Medical Care*, 29 (5), May, pp. 452-472.
8. Weisbrod, B.A. (1991). "The Health Care Quadrilemma: An Essay on Technological Change, Insurance, Quality of Care, and Cost Containment", *Journal of Economic Literature*, 29 (2), pp. 523-552.

Appendix

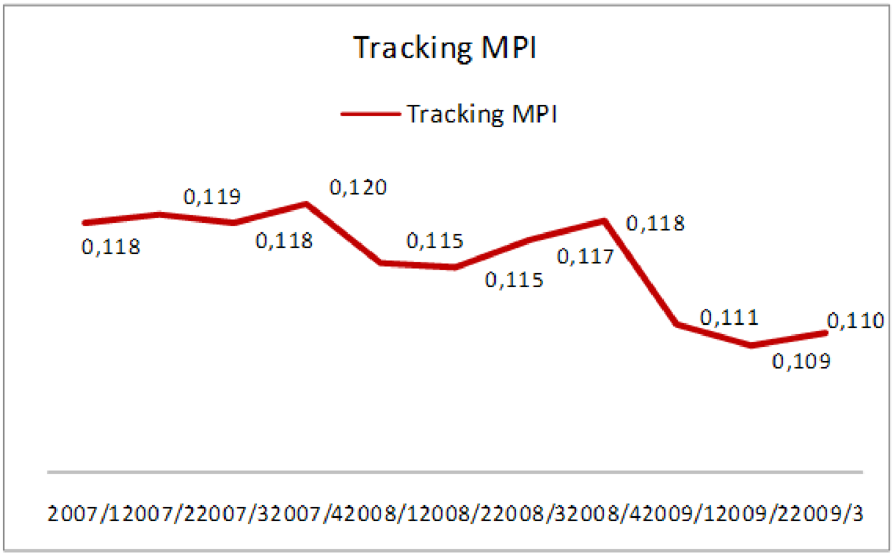

Fig. 1. MPI – tracking trend (2007-2009), Traditional Medicare

Fig. 2. MPI, state variation, 3rd quarter 2009, Traditional Medicare

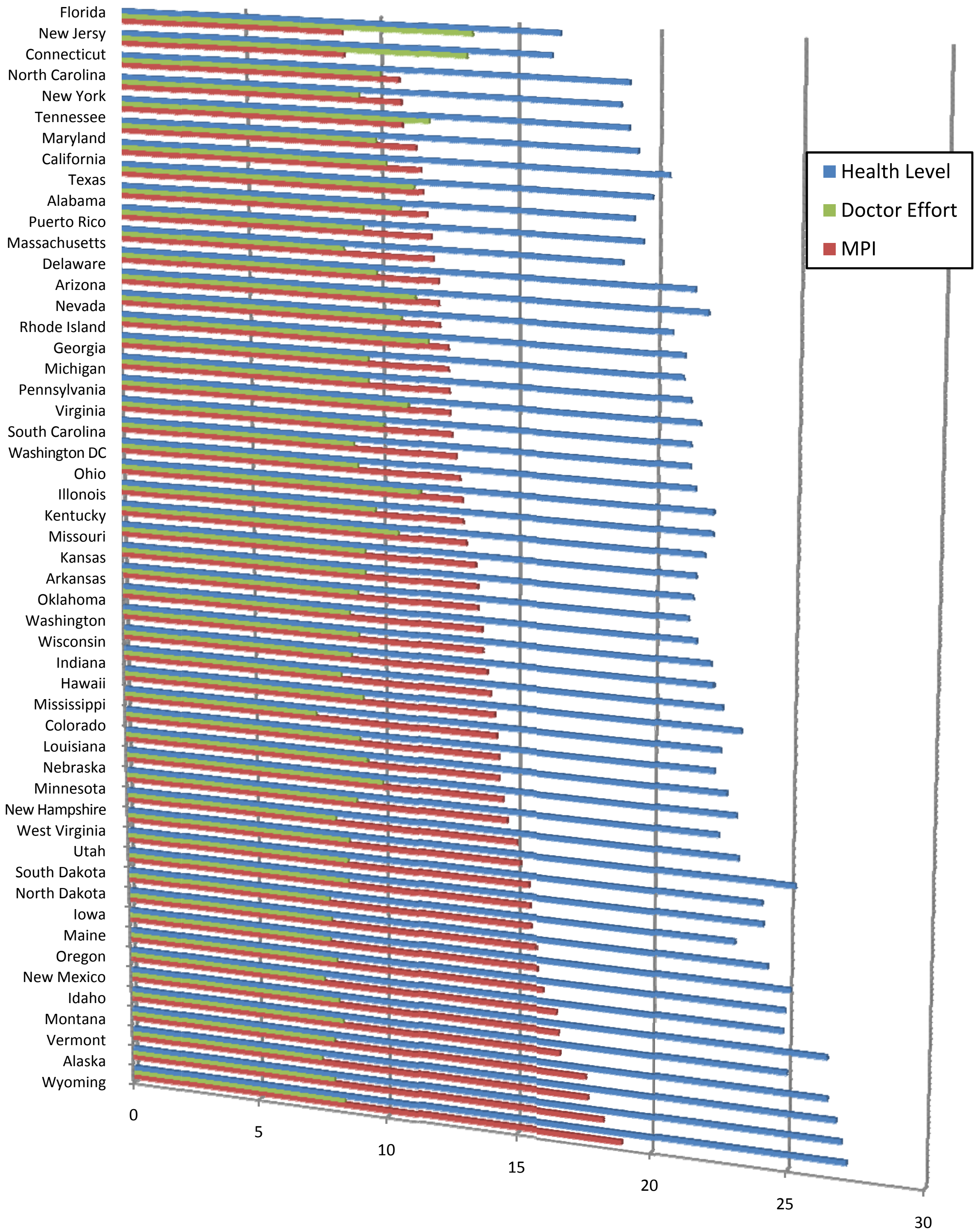

Fig. 3. State variation in MPI component parts, 3rd quarter 2009, Traditional Medicare

Table 1. MPI state ranking, 3rd quarter 2009, Traditional Medicare

MPI	Physician effort	Health level	State
18,89	8,32	27,17	Wyoming
18,20	7,91	26,97	Alaska
17,62	7,43	26,76	Vermont
17,52	7,92	26,42	Montana
16,50	8,13	26,39	New Mexico
15,09	8,58	25,15	New Hampshire
15,72	7,79	25,05	Iowa

Table 1 (cont.). MPI state ranking, 3rd quarter 2009, Traditional Medicare

MPI	Physician effort	Health level	State
16,55	8,28	24,93	Idaho
15,92	8,05	24,84	Maine
16,43	7,55	24,77	Oregon
15,67	7,86	24,22	North Dakota
15,45	8,53	24,04	Utah
15,41	8,52	23,99	West Virginia
14,13	9,17	23,17	Indiana
14,99	8,06	23,10	Minnesota
14,44	9,89	23,01	Louisiana
15,50	7,78	23,01	South Dakota
14,28	9,31	22,65	Colorado
13,97	8,33	22,46	Wisconsin
14,20	7,35	22,42	Hawaii
14,58	8,90	22,38	Nebraska
14,28	9,03	22,19	Mississippi
13,84	8,73	22,13	Washington
12,94	11,37	22,07	District of Columbia
12,95	9,70	22,05	Ohio
13,67	9,03	22,04	Oklahoma
12,10	11,27	21,81	Delaware
13,07	10,55	21,76	Illinois
12,48	10,98	21,55	Michigan
13,64	8,68	21,51	Arkansas
13,41	9,30	21,45	Kentucky
12,84	9,07	21,42	South Carolina
13,49	9,32	21,36	Missouri
12,09	9,82	21,34	Massachusetts
12,56	10,05	21,23	Pennsylvania
12,47	9,48	21,23	Georgia
12,71	8,92	21,22	Virginia
13,49	9,00	21,20	Kansas
12,44	11,71	21,00	Nevada
12,44	9,48	20,96	Rhode Island
12,12	10,74	20,55	Arizona
11,47	10,23	20,40	Maryland
11,55	11,25	19,81	California
11,83	9,33	19,50	Alabama
11,33	9,86	19,29	Tennessee
11,66	10,74	19,16	Texas
10,71	10,03	18,99	Connecticut
10,84	11,78	18,97	New York
11,91	8,61	18,76	Puerto Rico
10,81	9,22	18,68	North Carolina
8,62	13,37	16,53	Florida
8,69	13,16	16,23	New Jersey