Estimates of Physician Productivity: An Evaluation

Joseph P. Newhouse, Ph.D. and Anna D. Sinaiko, M.P.P.

Fisher (2007-2008) has carried out an analysis of physician productivity that is the most thorough to date. Because of limitations in the data, however, he is forced to make many assumptions to arrive at his estimates. This evaluation describes six assumptions to which the results are likely to be relatively sensitive as well as a number of other assumptions to which the results are likely to be less sensitive.

INTRODUCTION

The article by Fisher (2007-2008) clearly is by far the best attempt yet to measure physician productivity. Unfortunately, his estimates are sensitive to the numerous assumptions he necessarily had to make because of inadequacies in the data. In this comment we examine several of those assumptions in order to form a judgment about the overall amount of uncertainty surrounding his estimates.

ASSUMPTION 1

Fisher has two sources of physician income, but neither covers the entire period nor splits physician income into a labor component and a return-on-equity (ROE) in the physician's practice. From his sources he constructs a single measure of annual income over the 1982-2004 period. His principal source is the gross receipts data from the National Income and Product Accounts (NIPA). The NIPA figure includes not only data on gross receipts in physicians offices, but also receipts of ambulatory surgery centers (ASCs) and health maintenance organizations (HMOs) centers. The inclusion of ASCs and HMOs would not distort Fisher's results with respect to physician offices if physician offices were a constant share for all his calculations over the period, but we know that is not the case since ASCs have grown. Because they are still a rather small share of total receipts (most likely under 5 percent), however, the error from this assumption is probably modest.

From the NIPA figure on gross receipts Fisher deducts estimated amounts for other inputs, including wages of non-physician employees, certain intermediate inputs, business taxes, and depreciation. Not all intermediate inputs, however, are included in the NIPA figures on intermediate inputs, so the residual from this calculation includes both physician income and some remaining intermediate inputs.

To remove the remaining intermediate inputs from the residual NIPA values, Fisher uses a second source of physician income, three surveys carried out in 1995, 1999, and 2003 (Tu and Ginsburg, 2006). Ginsburg's income values are purely physician income, and Fisher uses them to calculate the proportion of the NIPA numbers for those 3 years that is physician income and the proportion that is unaccounted for by intermediate inputs such as the salary of the office nurse. He creates similar estimates for 2 additional years by using data from U.S. Bureau of Economic Analysis (BEA) input/output tables that give intermediate input

The authors are with Harvard University. The statements expressed in this article are those of the authors and do not necessarily reflect the views or policies of Harvard University or the Centers for Medicare & Medicaid Services (CMS).

expenses for 1987 and 1992 for physician and dentist offices combined. He uses these data to establish the percentage of total revenues that are intermediate inputs in those 2 years and nets that out of gross revenue to give an estimate of physician income. Thus, he has estimated values of physician income for 5 different years between 1987 and 2003, and he interpolates between and extrapolates from those 5 years so as to cover each year of the 1982-2004 period.

There are at least two potential problems with this procedure in addition to the issue of the inclusion of the ASC and HMO data in the NIPA values. First, the BEA input/output data include dentists as well as physicians. If dentists' share of income and intermediate inputs relative to gross revenues differs from physicians', the 1987 and 1992 values are not comparable to the values from the Ginsburg surveys since the Ginsburg values pertain solely to physicians. Second, the interpolations and extrapolations pose a potential source of error, especially since the Ginsburg data are from a sample of physicians and thus will contain some measurement error which will enter both measured endpoints. Because these endpoints are used to compute the trend over a relatively short period (1995-2003), the measurement error could be material. Overall, how much these factors affect the estimates is unclear.

The resulting income values also contain a potential ROE, which Fisher wishes to net out to obtain a pure measure of labor compensation. He does so by using net profit rates for incorporated physicians (from Statistics of Income), and imputing that rate to all other types of physicians, thus making the assumption that profit rates are similar across all types of physicians.

Fisher runs some sensitivity tests by obtaining data from the BEA on the proportion of ambulatory health services that are intermediate inputs and from the U.S. Census Bureau on the same proportion for physician offices. These additional data are reproduced in Table 1. The R^2 between the BEA data and Fisher's data is only 0.13 (and would be lower if the BEA data extended back to 1982 since Fisher's data are constant between 1982 and 1987 [data not shown in table]), so this comparison does not give us much comfort. The U.S. Census Bureau data are available for only 3 years. In 2 of the 3 years the values are similar, but in the most recent year, 2002, they are reasonably discrepant. In any event, we find it hard to draw much information from a comparison of three data points. The 1992 and 1997 BEA data are benchmarked to the Economic Census data and, therefore, the BEA and Census data should be similar. At the same time of Fisher's study, however, the 2002 BEA data had not yet been benchmarked and so were not linked to the 2002 Economic Census data, which may be why the 2002 data differ.

ASSUMPTION 2

To obtain an estimate of changes in real output, Fisher must deflate expenditures on physician services by a price index. To do so he uses the Producer Price Index (PPI) for 1998 and later, but he must use the Consumer Price Index (CPI) before 1997. (He uses a BEA transition value for 1997.) Although the PPI is based on transaction prices, the CPI is based on list prices, so Fisher needs an estimate of transaction prices in the pre-1997 period. Thus, the uncertainty about income is compounded by uncertainty about the magnitude of physician discounts or allowances in the pre-1997 period.

Table 1

Alternative Intermediate Input Expense Proportions, by Year

Year	U.S. Bureau of Economic Analysis Ambulatory Health Services	Fisher Physician Offices	U.S. Census Bureau Physician Offices
		Percent	
1987	20.6	29.1	—
1988	24.4	29.1	—
1989	22.3	29.0	—
1990	22.1	29.0	—
1991	22.0	29.0	—
1992	23.2	28.9	27.1
1993	24.2	28.8	—
1994	24.5	28.6	—
1995	26.7	28.4	—
1996	28.0	27.0	—
1997	29.9	27.4	27.1
1998	32.1	28.1	—
1999	32.0	28.4	—
2000	31.8	28.4	—
2001	30.8	29.9	—
2002	31.0	32.2	37.4
2003	32.1	34.4	—
2004	32.1	35.7	—

SOURCE: Fisher, C., 2007.

To convert the CPI list prices to transaction prices, Fisher uses an estimated discount (allowance) rate for Medicare plus private payers based on annual data on the Medicare allowance rate. He ignores Medicaid, but Medicaid comprises a relatively small share of physician income, 7 percent in 2005.[1] He assumes physician discounts for all payers on average equaled the Medicare discount in 1975. He also has a general population (i.e., Medicare plus private) allowance rate from the 1996 Medical Expenditure Panel Survey (MEPS), which is roughly 10 percentage points higher than the 1996 Medicare-only value. To obtain general population figures for the years between 1975 and 1996, he uses the values for the two endpoints previously described and fits a quadratic form for the intermediate years (because the Medicare-only data, which are available annually, seem to follow a quadratic form). To fit the quadratic for all payers, Fisher has three unknowns, namely the three coefficients of a quadratic, but only two data points (1975 and 1996). To identify the equation he therefore imposes the first derivative of the quadratic from the change in the Medicare-only allowance from 1975 to 1976. His logic is presumably that both Medicare and private discounts (allowances) were small in both 1975 and 1976 and hence the Medicare change in those 2 years equaled the change in private allowances. He then applies the predicted values of the allowance rate to the CPI to get an estimated index of transaction prices.

There are three strong and ultimately untestable assumptions in this method that:

- Medicare plus private discounts followed a quadratic form over the 1975-1996 period.

[1] The 7 percent figure is from http://www.cms.hhs.gov/NationalHealthExpendData/02_NationalHealthAccountsHistorical.asp#TopOfPage. It includes clinical services, but excluding them would not change the conclusion that Medicaid is a small proportion of physician income.

- The first derivative of the quadratic can be imposed from the 1975-1976 change.
- The Medicare discounts and private discounts were essentially equal in the mid-1970s.

There are some data bearing on the second assumption. Any independent measurement error in levels will be compounded in measures of change, and any measurement error in the 1975 to 1976 change will affect all of Fisher's calculated values. Table 2 shows the year-to-year changes in the Medicare allowance rate for the years 1975-1979. Clearly the amount of annual change is a noisy variable, with the largest single change being the one that Fisher used.

In addition to this method of computing a price index, Fisher calculates an alternative price index based on the ratio of Medicare physician rates to the commercial rates of two national insurers. These data, collected for the Medicare Payment Advisory Commission (MedPAC), cover the periods from 1989-1996 and 1999-2004. To extrapolate them backward from 1989 and to impute values for 1997 and 1998, Fisher observes that there is a near perfect correlation between these data and CMS data on the proportion of physician revenue from private sources. He uses this relationship to impute values for the years that are missing from the MedPAC data.

Table 2

Annual Changes in the Medicare Allowance Rate: 1975-1979

Year	Level	Change from Prior Year
1975	92.2	—
1976	91.0	-1.2
1977	90.7	-0.3
1978	90.2	-0.5
1979	89.5	-0.7

SOURCE: Fisher, C., 2007.

Table 3

Fisher's Output Price Index Compared with an Output Price Index Based on the MedPAC Data on Private Prices: 1985-1995

Year	Fisher's Price Series	MedPAC Data
	Percent Change	
1985	4.6	10.4
1986	5.7	(4.4)
1987	5.6	0.6
1988	5.3	10.0
1989	5.2	(2.5)
1990	4.8	3.2
1991	3.5	3.1
1992	3.5	5.7
1993	2.5	0.2
1994	1.1	2.0
1995	0.9	2.8

NOTE: MedPAC is Medicare Payment Advisory Commission.
SOURCE: Fisher, C., 2007.

Table 3 compares the values Fisher uses based on the Medicare allowances with those calculated from the MedPAC data. Over the 1985-1997 period Fisher's series is going up annually 1 percentage point faster than the series from the MedPAC data (3.3 versus 2.3 percent). Moreover, there are some substantial differences in certain years. Because the multifactor productivity (MFP) measure is a trailing 10-year average, extreme values in any one year can make a non-trivial difference to the 10-year average as they enter or exit the moving average. Despite these potential problems, we find the congruence between the 7 data points from 1989 to 1996 in which Fisher's "backcasted" PPI overlaps with the MedPAC data comforting (R^2 = 0.87). We only wish the overlap period were longer.

Before leaving this point, we agree with Fisher that the PPI should be used for 1998 and later, but in appraising the confidence one should have in the resulting

MFP values, it would be helpful to know how often the U.S. Bureau of Labor Statistics is actually able to obtain transaction prices. In 1997 less than one-half the intended number of practices were included in the sample (Berndt et al., 2000), although the situation may have improved since that time.

ASSUMPTION 3

The physician services billed under the Medicare fee schedule include those in both inpatient and outpatient settings. No data source estimates the capital applied in the inpatient setting, so Fisher can hardly be faulted for omitting it. Nonetheless, some of the observed change in physician productivity is likely attributable to changes in capital inputs in that setting. Given the volume of physician services carried out in the hospital, including in the outpatient department, this omission could be material.

Omitting hospital capital has one further implication. To the degree the site of care has shifted from the hospital to the office or the ASC, measured productivity is biased down, because the share of unmeasured inputs is falling. Shifting the site of care among the outpatient department, inpatient care, and the ASC also is a problem to the degree intermediate input and labor mixes differ among these sites, since Fisher's data on those inputs are from all those sites combined.

ASSUMPTION 4

To estimate the labor input, Fisher must estimate changes in physician hours. To do so he uses Current Population Survey data on office-based physicians, but hospital-based physicians are not included. According to American Medical Association data, office-based physicians are about 75 percent of all patient care physicians. Thus, Fisher does not count some attending physicians in teaching hospitals who bill Medicare.[2] More importantly, his calculations do not include residents who do not bill, but who augment productivity of the senior physicians who do bill Medicare. This is a problem to the degree that the share of hospital-based physician hours is changing or the share of resident hours is changing. During the period in question legislated limits on resident hours were enacted, so it is plausible that the share of resident hours fell. This would have caused measured physician productivity to fall.

ASSUMPTION 5

The available data permit only a crude adjustment for changes in specialty mix, which matters to the degree the specialty mix changes and productivity is differential between specialties. Specialties with a higher proportion of new procedures may have greater true productivity change because of learning-by-doing, and in practice changes in the relative value units may not fully account for this change. Specialists may also be more productive in ways that are not measured for the same Healthcare Common Procedure Coding System code, such as more accurate diagnoses in the same length visit or fewer tests (or vice versa) to reach a diagnosis. This last point is related to the issue of whether the relative value unit defines a homogeneous product across specialties.

ASSUMPTION 6

Fisher calculates the ROE from the return to incorporated physician offices,

[2] Hospital-based physicians are employed under contract with hospitals to provide direct patient care. Office-based physicians are in solo or group practice or other patient care employment.

but these data are available only for 1995-2002. As a result, they must be imputed for the periods 1982-1994 and 2003-2004.

The previous six points discuss Fisher's assumptions and imputations that seem to us the most important in determining his estimates. We list here several other assumptions or imputations that are individually likely to be immaterial, but that cumulatively add to the uncertainty of his estimates.

- The fringe benefit data apply to all health and hospitals rather than to physician offices.
- Values for the intermediate input proportion of total spending for the period 1982-1986 must be imputed.
- Business taxes must be imputed for physician income from 1982 to 1997.
- The number of osteopathic physicians must be imputed for 16 of 23 years.
- Physician weeks worked must be imputed for 2004.
- HMO centers are included in the NIPA values, but physicians in those centers are not reimbursed under the Medicare fee schedule.

CONCLUSION

We close with two general comments on any effort to estimate physician MFP, including Fisher's. First, the theory underlying productivity measurement assumes competitive pricing. Given the prevalence of administered prices in the physician marketplace, most notably in the Medicare Program itself, this is a strong assumption. For example, in the 1990s the creation of networks by managed care companies is widely thought to have reduced physician rents in the commercial market. To the degree any reduction in rents reduced physician income and was not picked up in the price indices, there will be a downward bias in measured productivity. This is also true to the degree the Sustainable Growth Rate mechanism, part of the formula used to determine annual updates in physician fees in the Medicare Program, reduced any rents in Medicare prices in a way that was not picked up in the price indices, especially in 2002 when fees were cut. New procedures may also enjoy rents, and rents may exist in smaller markets where competition is limited.

Second, consider the following quote from Griliches (1992) on the importance of accounting for quality in any measurement of productivity in service sectors:

> "…one needs to consider...that difficulties in measuring output and prices in services may have resulted in a mismeasurement of productivity growth in these sectors, a mismeasurement that accounts for some or even much of the observed contrast with commodities … Why is the problem [of productivity measurement] more serious in some of the service sectors? …The conceptual problem arises because in many service sectors it is not clear what is being transacted, what is the output, and what services correspond to the payments made to the providers …Over all this hangs the ubiquitous issue of quality change. The problem is general and pervasive …in general, because of the underlying heterogeneity of transactions, the difficulty of making comparisons across time and space is even greater."

One conclusion from this quote is that our task of pointing out uncertainties in Fisher's estimates is much easier than his task of estimating physician productivity.

REFERENCES

Berndt, E.R., Cutler, D.M., Frank, R.G., et al.: Medical Care Prices and Output. In Culyer, A. and Newhouse, J. (eds.): *Handbook of Health Economics*. Elsevier-Amsterdam Publishing. Amsterdam, Holland. 2000.

Fisher, C.: Multifactor Productivity in Physicians' Offices. *Health Care Financing Review* 29(2):15-32, Winter 2007-2008.

Griliches, Z.: Output Measurement in the Service Sector. University of Chicago Press. Chicago, IL. 1992.

Tu, H.T. and Ginsburg, P.B.: Losing Ground: Physician Income 1995-2003. Center for Health System Change, Washington, DC. June 2006.

Reprint Requests: Joseph P. Newhouse, Ph.D. Harvard University, 180 Longwood Avenue, Boston, MA 02115. E-mail: Joseph_Newhouse@harvard.edu