

Reply to “Pitfalls in the Analysis of Regional Variations in Health Care:  
A Response to Hadley, Berenson, Waidmann and Zuckerman,”

by Amitabh Chandra, Elliott Fisher, and Jonathan Skinner (August 23, 2007)

Jack Hadley, Ph.D., George Mason University  
Robert Berenson, M.D., The Urban Institute  
Timothy Waidmann, Ph.D., The Urban Institute  
Stephen Zuckerman, Ph.D., The Urban Institute

September 5, 2007

Before responding to specific points raised in your commentary, we'd like to start with something you said in your cover e-mail, specifically, that you plan to send this "...to the policy makers who wondered whether all of our Dartmouth work was now made irrelevant by your work." If this was in fact the way Mark Miller and CBO described the implications of our analysis in their communication with you, it is most unfortunate, since an inference like this is a gross exaggeration of anything our analysis does or implies. We have not estimated any outcome models and draw no conclusions about the extent of inefficiency in medical care received by the elderly. Moreover, our study analyzes variations in elderly individuals' medical spending over a three-year observation period for most of the cases in our sample, not regional variations in average annual spending per elderly person. As such, we draw no inferences about the sources or causes of regional variation in average Medicare spending per Medicare beneficiary.

At this point, the primary difference between the implications of our analysis and your work is the role of the supplies of local medical resources, which we do not find to be a significant determinant of variation in individual spending. However, we do not address the issue of the local organization of medical resources and, as such, make no inferences regarding the potential difference in the efficiency of obtaining care from an organized system of providers as opposed to a set of fragmented and independent providers.

With regard to the four "claims" about your work that you address so extensively in your commentary, our report describes these as potential caveats that serve as a rationale for a study of the relationship between medical spending and health outcomes at the individual level for elderly Medicare beneficiaries. We certainly do not say that your work is either wrong or irrelevant. Rather, our goal in raising these concerns was to argue that policy makers would benefit from knowing whether the fundamental relationship you have identified between variations in regional spending and individuals' outcomes also applies to the relationship between variations in individuals' spending and their outcomes. The shorter version of our report, which we are preparing to submit for publication and will present at the HCFO briefing, omits these concerns and motivates the analysis by emphasizing the relevance to policy makers

of understanding the determinants of variations in individual Medicare beneficiaries' spending, regardless of the ultimate inferences about the relationship between spending and outcomes.

While it is our intent to proceed to estimating outcome models using an IV methodology, we have not yet done so and, therefore, find it both extraordinary and inappropriate to be so severely criticized for an analysis we have not yet done. We are well aware of the pitfalls and difficulties of IV analysis, and clearly state that this current analysis of variations in medical spending for individual Medicare beneficiaries is the first step in the search for potential exogenous identifying variables that meet the criterion of being significantly related to individuals' medical spending.

The report also explicitly states the possibility that some of the variables we identify as potential instruments may in fact be correlated with actual or expected health, and that a major part of the next phase of our study will be to assess this issue, as well as the sensitivity of any IV results to the specification of the set of exogenous identifying variables.

At the end of the day, we will have one of three possible general results.

- None of our potential identifying variables may satisfy the tests for instrument validity, and we will draw no inferences about the relationship between individuals' spending, their health outcomes, and the efficiency of medical care use.
- We may have satisfactory instruments, but not find a positive relationship between spending and outcomes. This finding will reinforce the results of your work.
- Or, we may have satisfactory instruments and find that individuals who receive more care have better outcomes. In this case, we would call for more research to identify and resolve the apparent differences between our findings.

The remainder of our reply addresses some of the specific points you made about our analysis. (These comments refer to a revised version of the analysis, which we are currently preparing, that uses dummy variables to represent quintiles rather than deciles of areas grouped by the end-of-life expenditure index.)

1. Log transformation minimizes the apparent difference in the (true) means across expenditure classes.

We agree that the log transformation reduces the apparent difference in the means. In our sample, mean total spending per person is \$862 per month in the 1<sup>st</sup> quintile and \$1,298 in the 5<sup>th</sup> quintile, a difference of 50%. When we estimate the models without the log transformation, the difference in total spending per month between the first and fifth quintiles is reduced to 29%, from \$436 to \$249, which is greater than the 19% difference estimated from the log transformation model.

However, the log transformation does not affect our estimate of the relative contributions of area and personal characteristics to the amount of variation explained by the model. While it is true that the  $R^2$ s are smaller using untransformed data, it is still the case that personal characteristics account for most of the explained variation. The full set of variables has a  $R^2$

of 0.09, while a model with area variables, age, race and gender has a  $R^2$  of 0.02. Estimating the model with only personal characteristics generates a  $R^2$  0.087. Thus, the relative amounts of variation explained by area vs. personal characteristics are essentially the same regardless of whether the dependent variable is measured in logarithms.

## 2. Including urban dummy variables is the reason for the modest role of geographic variation.

The coefficient estimates of the end-of-life expenditure quintile dummy variables are quite robust with respect to whether any other geographic variables, including the urban dummies, are included in the model. Excluding all of the other area variables increases the coefficient of the 5<sup>th</sup> quintile dummy from .29 to .339 in the log version of the dependent variable, and from 248 to 263 in the unlogged version of the dependent variable. This is one of the reasons we conclude that the variation captured by the end-of-life expenditure quintile dummy variables does not appear to be correlated with variations in the local supplies of medical providers. By way of speculation, it may capture differences in how providers are organized, in providers' financial incentives, and/or in unmeasured personal characteristics.

## 3. Use of flawed instruments

Since we have not conducted an IV analysis and have not tested the potential set of identifying variables from the first-stage analysis, these criticisms are both premature and inappropriate. We are aware of the relationships you identify, but think it important to point out that most of them are based on cross-sectional analyses of annual spending and various health outcomes. Our analysis is based on spending over (up to) three years and will analyze the change in health over a time period of up to three years (between a person's entry into the MCBS and their exit), as a function of baseline or pre-determined characteristics. It is essentially longitudinal, following the spending and health changes of the same set of individuals. Thus, it is by no means certain that the findings from other studies using different research designs necessarily apply.

For example, although many studies find that marriage is associated with better health and presumably lower spending, we find that beneficiaries who were never married spend less than married people. This suggests that our specific measure may be capturing differences in access to care or attitudes about the use of medical care, not the protective effects of marriage on health.

Similarly, Kronick's commentary (2006) about prior health insurance in another study by Hadley and Waidmann (2006a) is largely irrelevant to this analysis. That study focused on the non-elderly and any type of insurance coverage over time. Our current analysis focuses on a very specific type of private retiree health insurance, i.e., with very low or not premium cost to the beneficiary, on the grounds that retention of this type of private supplementary coverage is less likely to be related to health status. Moreover, our reply (Hadley and Waidmann 2006b) to Kronick pointed out the flaws in his observations, and other studies (Dor, Sudano and Baker 2006; Baker et al. 2006; McWilliams et al. 2004) reinforced our findings using different methods. In sum, while some, or possibly all of the potential

exogenous identifying variables from the first-stage analysis may ultimately fail to pass muster as valid instruments, we will investigate and determine this in our forthcoming analysis.

4. Use of the medical care CPI to adjust for inflation is problematic

We are aware of the problems of the medical care CPI as an accurate measure of true inflation in medical care prices. However, inaccuracies from the use of the medical care CPI result in measurement error in the dependent variable in this study and, therefore, should not bias in the coefficient estimates of the independent variables. Moreover, all of our models include year dummy variables that should control for the effects of inaccurate adjustments for changes in prices over time.

5. Use of county fixed effects

You are correct that the models we report were not estimated with county fixed effects. The paragraph that states this is wrong and should have referred to a sensitivity test we conducted in which we used county fixed effects in place of the other geographic variables to check whether the estimated effects of the personal characteristics variables were sensitive to county fixed effects. They were not.

6. Results are the same for Medicare and total spending

We don't necessarily disagree with this conclusion, but believe that it would be useful to demonstrate it.

7. The adjustments for variations in health status across HRRs are adequate

Since we do not analyze regional variations in spending, this issue is not directly relevant to our analysis. We do find, however, that when our sample is stratified by initial (baseline) self-reported general health status, the difference in spending between people living in the 1<sup>st</sup> and 5<sup>th</sup> quintile areas of the end-of-life expenditure index is much smaller for people in fair or poor health initially than those in excellent, very good, or good health, 8% compared to 21%. Of course, this result says nothing about inefficiency, nor does it imply that your adjustments are incomplete. It is simply another way of looking at the relationships between health, medical spending, and geography that may, or may not, provide useful information about the extent and sources of inefficiency in medical care.

## REFERENCES

Baker, D. et al. 2006. "Health Insurance Coverage and the Risk of Decline in Overall Health and Death among the Near-Elderly," *Medical Care* 44 (3):277-282

Dor, A., J. Sudano and D. Baker. 2006. "The Effect of Private Insurance on the Health of Older Working Age Adults: Evidence from the Health and Retirement Survey," *Health Services Research* 41 (3): 759-787.

Hadley, J. and T. Waidmann. 2006a. "Health Insurance and Health at Age 65: Implications for Medical Care Spending on New Medicare Beneficiaries," *Health Services Research* 41 (2): 429-451.

Hadley, J. and T. Waidmann. 2006b. "Commentary – Response to Richard Kronick," *Health Services Research* 41 (2): 461-466.

Kronick, R. 2006. "Sophisticated Methods but Implausible Results: How Much Does Health Insurance Improve Health?" *Health Services Research* 41 (2): 452-460.

McWilliams, J. et al. 2004. "Health Insurance Coverage and Mortality among the Near-Elderly," *Health Affairs* 23 (4): 223-233.