Sam Houston State University
Department of Economics and International Business
Working Paper Series

Is Health Care a Necessity or a Luxury? New Evidence from a Panel of U.S. State-Level Data

Donald G. Freeman
Sam Houston State University

SHSU Economics & Intl. Business Working Paper No. 12-03
August 2012

Abstract:

This paper estimates the income elasticity of health care expenditures using annual data on health spending by state in the U.S. from 1966-2009. Panel stationarity tests incorporating structural breaks in the levels and trends in Health Care Expenditures and Disposable Personal Income yield inconsistent results, with stationarity rejected for HCE but not for DPI. Regression results using levels estimation robust to orders of integration differed considerably depending on time period, and the cross-state variation in elasticity estimates was quite large. Results of the first difference models provide more consistent estimates across time periods, whether expressed as averages of individual state estimates or as pooled time series. Income elasticities for the full sample fall in the range 0.21-0.22, below that of recent research.

Is Health Care a Necessity or a Luxury?
New Evidence from a Panel of U.S. State-Level Data

by

Donald G. Freeman, Ph.D.
Professor of Economics
Department of Economics & International Business
Sam Houston State University
P.O. Box 2118
Huntsville, Texas 77341-2118
tel: 936-294-1264
fax: 936-294-3488
e-mail: freeman@shsu.edu
url: http://www.shsu.edu/~eco_www/faculty/freeman.html

Paper originally presented to the 41st Annual Conference of Regional Science Association International: British and Irish Section. August, 2012

Is Health Care a Necessity or a Luxury?
New Evidence from a Panel of U.S. State-Level Data

Abstract

This paper estimates the income elasticity of health care expenditures using annual data on health spending by state in the U.S. from 1966-2009. Panel stationarity tests incorporating structural breaks in the levels and trends in Health Care Expenditures and Disposable Personal Income yield inconsistent results, with stationarity rejected for HCE but not for DPI. Regression results using levels estimation robust to orders of integration differed considerably depending on time period, and the cross-state variation in elasticity estimates was quite large. Results of the first difference models provide more consistent estimates across time periods, whether expressed as averages of individual state estimates or as pooled time series. Income elasticities for the full sample fall in the range 0.21-0.22, below that of recent research.

Is Health Care a Necessity or a Luxury?
New Evidence from a Panel of U.S. State-Level Data

I. Introduction

A major concern of many of the world's developed economies is the increase in health expenditures as a share of national income. Health care expenditures (HCE) now account for over 10% of GDP in most western economies, and in the United States will approach 20% within the next decade.[1] Aging populations will increase the demand for health delivery, with consequent pressure on capacity and prices of health services and products.

An important question for public policy regarding the growth of HCE is the role of income as a driver of health spending. It is clear that HCE has grown faster than incomes, as evidenced by its increasing share of total expenditure, but whether this is a consequence of higher incomes or of other factors is still an open question. If increases in income, *ceteris paribus*, occasion a more than proportionate increase in HCE, the elasticity of HCE is greater than unity, and HCE is a luxury good. In this case, policy concerns are lessened, as HCE can be viewed as a good where consumers have considerable discretion over spending levels.

If however the observed increasing share of HCE in total expenditures is driven more by cost factors, with upward shifting supply and price-inelastic demand, then questions of affordability and access become more important to policy makers. Unlike housing or travel, two other categories with increasing expenditure shares over time, HCE is less amenable to postponement and more often of a critical nature. Complicating matters is the fact that "health", the desired outcome of HCE, is itself an input into income generation. "Wealthier is Healthier," as Pritchett

[1] "Health Expenditures" include both out-of-pocket and government financed spending.

and Summers (1996) titled their article, but healthier societies are also wealthier and more productive.

Accurately measuring the response of HCE to income levels has been the objective of a large body of research in economics, beginning with Kleiman (1974) and Newhouse (1977) and with many contributions since.[2] Most of this literature has focused on international comparisons of HCE and income, with increasing emphasis on panel data techniques of estimation as longer time series of country-specific data become available. An advantage of panel data is in the multiple realizations of the relationship under study, with potential gains in power, but the increased variability in cross-country data comes at a cost of unobservable factors that can bias the results. Strong heterogeneity is one explanation given for the observed increase in HCE income elasticity as one moves from individual level to state or country level analysis; see for example Hansen and King (1996).

Because health care systems differ substantially across countries, one way to control for heterogeneity and still benefit from the advantages of multiple time series is to examine HCE for multiple jurisdictions within a single country. The U.S. states are a natural candidate for this type of study. States obviously differ in many dimensions, but the hybrid system of part-private, part-public health financing is common to all. Studies of HCE and income using U.S. states are relatively rare. Freeman (2003) assumes homogeneous responses of HCE to income across states in a panel over 1966-1998, finding that HCE is a necessity good with elasticity in the range 0.8-0.85. Wang and Rettenmaier (2006) estimate separate elasticities for each state, finding income elasticity of HCE below unity in 16 states and above unity in 32 states. Moscone and

[2] A complete survey would require an article in itself. Literature most relevant to this paper is cited.

Tosetti (2009) employ recent developments in controlling for cross-section dependence in panel data to estimate a much lower elasticity for the group of states of between 0.36 and 0.45.

A critical question that these time series studies of HCE and income have had to answer is whether the series are stationarity. A time series is strictly stationary if arbitrary sub-samples have a common distribution; it is weakly or covariance stationary if arbitrary sub-samples have common first and second moments. If a series is stationary, it is said to be integrated of order zero, or I(0), if its first difference is stationary, it is said to be integrated of order one, or I(1). As is well known, using OLS to estimate relationships among non-stationary series or among series of different orders of integration can result in spurious outcomes.

Hansen and King (1996) were among the first to examine the issue of stationarity in the series commonly used in cross-country analyses of HCE, using standard augmented Dickey Fuller (ADF) tests to conclude that many series, including HCE and GDP, were non-stationary. Relationships among I(1) series can still be estimated in levels, provided that they are cointegrated,[3] but Hansen and King were unable to reject the null of no cointegration between HCE and GDP for 19 of the 20 OECD countries in their sample. Based on their findings, Hansen and King questioned previous research that had estimated income elasticities of HCE in excess of unity.

McCoskey and Selden (1998) revisited the question of stationarity in HCE and GDP by applying developments in panel unit root tests to the OECD data. Because the individual country time series were relatively short (27 years), McCoskey and Selden argued that unit root tests

[3] A classic cause of non-stationarity in a time series is the existence of a unit root in the data generating process (DGP), such that $y_t = y_{t-1} + DT + e_t$. In this case, it can be shown that the series has both a deterministic trend, a function of the "DT" term, and a stochastic trend, essentially the sum of all past error terms. If two series share a stochastic trend, they are said to be cointegrated.

would lack power against highly persistent but stationary alternatives. Applying the Im, Pesaran, and Shin (1996) "IPS" test to the panel of countries,[4] McCoskey and Selden were able to reject the unit root null for HCE and GDP, thereby mitigating concerns that using standard regression techniques to estimate the elasticity of HCE resulted in misspecification.

Gerdtham and Löthgren (2000) included time trends in ADF tests of individual countries and the IPS panel test to conclude that both HCE and GDP were in fact I(1), in contrast to McCoskey and Selden (1996). Gerdtham and Löthgren also employed Kwiatowski, Phillips, Schmidt, and Shin (KPSS, 1992) stationarity tests to confirm their findings. The null of the KPSS test is stationarity, in contrast to the ADF test with its null of a unit root. Extending their tests to the question of cointegration, Gerdtham and Löthgren were able to conclude that HCE and GDP are cointegrated for the panel of OECD countries using either the IPS or the KPSS type test.

Freeman (2003) employed an approach similar to Gerdtham and Löthgren to estimate the elasticity of HCE with respect to income using a panel of U.S. states over the time period 1966-1998. In addition to the IPS test of a panel unit root, Freeman (2003) also used a Fisher-type test using the statistic $P_\lambda = -2\sum_i^n \ln(p_i)$, where p_i is the p-value of the ADF unit root test or Engle-Granger cointegration test for state i . P_λ is distributed as χ^2_{2n} assuming cross-section independence, but critical values must be simulated when this assumption is violated. Freeman found that HCE and income were I(1) and cointegrated, with estimated income elasticity of HCE between 0.817 and 0.844 using preferred specifications.

More recent developments in the time series literature emphasize a middle ground between a purely stationary series and a non-stationary series characterized by a unit root. Beginning with

[4] The IPS test is essentially a t-type test of the average of all the ADF t-tests for the panel of time series. Published critical values for the IPS test assume cross-section independence, an assumption unlikely to be met in practice.

Perron (1989), it has been realized that the existence of large shocks could cause structural changes to the deterministic components of a time series, resulting in too-frequent non-rejection of the unit root hypothesis. Applying structural break tests to health care expenditures, Jewell, et al. (2003) allow two level shifts in HCE and GDP and find in both cases that they are able to reject the unit root hypothesis for OECD data. Carrion-i-Silvestre (2005) allows for multiple shifts in levels and trends with cross-section dependence and also finds that HCE and GDP are stationary, though his conclusions for HCE are not as strong as those for GDP.

Thus far, only Wang and Rettenmaier (2007) have applied structural break tests to U.S. state-level HCE and gross state product (GSP) data. Using a single break over the period 1980-2000, they find that both HCE and GSP have unit roots and are cointegrated, with income elasticities exceeding one in a majority of states. Wang and Rettenmaier's estimated elasticities are much higher than in most recent research, and as they note, the precision of their estimates may be affected by the relatively short time series.

This paper uses a much longer sample of state health expenditures, from 1966-2009, to conduct stationarity tests. Data availability is enhanced by recently released updates from the Centers for Medicare & Medicaid Services (CMS), and by the use of earlier data published in the Health Care Financing Review (1985). We also incorporate the multiple break analysis of Carrion-i-Silvestre (2005) to allow the number of breaks to vary across HCE and income, and across states. This greater flexibility proves very useful, as we find states are heterogeneous with respect to the number and the timing of breaks in both HCE and income.

Prior results in the literature have invariably concluded that either both HCE and income are stationary or both are non-stationary. This paper finds a mixed result, with evidence after accounting for structural breaks that HCE is non-stationary and income is stationary. This result does not rule out the existence of a long-run relationship between HCE and income, but it does suggest that estimation of the relationship must be handled with care. An approach robust to uncertainty about orders of integration in the variables based on Pesaran, Shin and Smith (2001) produces inconclusive evidence of a long-run relationship between HCE and DPI in levels. Elasticities calculated using first-difference models on the other hand produce consistent estimates in the range .017 to 0.26.

The following section describes the data and the stationarity and structural break tests. Section III reports the results of the tests on the individual series and on the panels. Section IV provides estimates of the elasticity of HCE with respect to disposable personal income (DPI). Section V concludes.

II. Data and Stationarity Tests

HCE is taken from the Centers for Medicare and Medicaid Services *Health Expenditures by State of Residence*. These data consist of total personal health care spending by state and by service, and are expressed as spending per capita. DPI is taken from the U.S. Department of Commerce, Bureau of Economic Analysis (BEA). These data are personal income, net of taxes. Both series are deflated by the Consumer Price Index (CPI). The national CPI is used as state-level price indexes do not exist.[5]

[Table 1 about here]

[5] Moscone and Tosetti (2010) construct state-level CPI from city and regional indexes published by the Bureau of Labor Statistics, using an approach from Holly, Pesaran, and Yamagata (2010). These local indexes by their nature are subject to more variation and measurement error than the national average. Some of the associations, moreover, are questionable: The Atlanta city price index is used for Alabama and South Carolina, San Diego for Arizona, and Boston for virtually all of New England.

Table 1 provides summary statistics for the HCE and DPI series. Average trend growth in HCE has been more than twice that of DPI, though HCE growth in later years has dropped about one percentage point. DPI growth has also slowed, but the differential from early to later years is less than 0.1 percent. Citizens of northeastern states generally have higher incomes and pay more in health costs than citizens of southern and western states. And while there are extremes in both growth rates and levels, the majority of the sample falls within a relatively narrow range for both HCE and DPI, as evidenced by standard deviations that are less than 20 percent of mean in all cases.

We first conduct unit root and stationarity tests for the two series by individual state and by panel of states with no consideration of structural breaks. Given that each type of test has size and power problems with time series with near unit roots and when heteroscedasticity and autoregressive errors are present, it is often useful to employ both as confirmatory analysis.

The unit root test uses the augmented Dickey-Fuller framework:

$$\Delta y_{i,t} = \alpha_i + \beta_i t + \rho_i y_{i,t-1} + \sum_{j=1}^{k} \delta_{i,j} \Delta y_{i,t-j} + \epsilon_{i,t} \ , \qquad (1)$$

with $y_{i,t}$ either HCE or DPI for state i at time t, $\varepsilon_{i,t}$ a stationary process and the number of lags k sufficient to control for serial correlation. The unit root null hypothesis for each individual state and time series is H_0: $\rho_i = 0$ and the alternative is H_1: $\rho_i < 0$; for the panel of states, the unit root null is H_0: $\rho_i = 0 \ \forall i$ and the alternative is H_1: $\rho_i < 0$ for some $n = N_1/N$, where N_1 is a subset of the total number of states N. The test statistic for the individual series is the ADF t-test of ρ_i, measured against finite sample critical values as compiled by MacKinnon (1994). For the panel test, we use the aforementioned Fisher-type statistic $P_\lambda = -2\sum_i^n \ln(p_i)$ as suggested by Maddala and Wu (1998), with $P_\lambda \sim \chi^2_{2N}$ under the assumption of cross section independence.

Because this assumption is unlikely to be met in the present case, the critical values for P_λ are bootstrapped using a method suggested by Maddala and Wu that maintains the cross-section temporality of the residuals.

The stationarity test uses the KPSS framework:

$$y_{i,t} = \alpha_i + \beta_i t + r_{i,t} + \epsilon_{i,t} \quad (2)$$

$$r_{i,t} = r_{i,t-1} + \mu_{i,t} \quad (3)$$

with $\mu_{i,t}$ i.i.d. (0, $\sigma_{u,i}^2$) and independent of $\epsilon_{i,t}$. The stationarity of $y_{i,t}$implies that $r_{i,t} = r_{i,0}$, a constant, or equivalently that $\sigma_{u,i}^2 = 0$. The KPSS test uses the statistic

$$LM = \left(T^{-2} \sum_{t=1}^{T} \hat{S}_{i,t}^2\right) / \widehat{\omega}_i^2 \quad (4)$$

where $\hat{S}_{i,t} = \sum_{j=1}^{t} \hat{\epsilon}_{i,j}$ is the partial sum of the residuals from regressing $y_{i,t}$on a constant and a time trend, as in (2) above, and $\widehat{\omega}_i^2$ a consistent estimate of the long run variance of $\epsilon_{i,t}$. Under the null hypothesis *H₀:* $\sigma_{u,i}^2 = 0$, *LM* converges to a function of standard Brownian motion, with critical values that depend on the number of deterministic terms in (2). In the heterogeneous panel version, as developed by Hadri (2000), the test statistic is given by:

$$LM_N = N^{-1} \sum_{i=1}^{N} \left(T^{-2} \sum_{t=1}^{T} \widehat{S_{i,t}^2}\right) / \widehat{\omega}_i^2. \quad (5)$$

In the case of cross-section independence, $LM_N$is shown to converge asymptotically to a standard normal distribution, but when this condition is not met, critical values must be simulated.

[Tables 2 & 3 about here]

Table 2 presents the results of the unit root and stationarity tests for HCE and Table 3 the results for DPI. In Panel A of each table the values of the KPSS and ADF statistics are provided, along with p-values and optimal lags for the ADF test. The rightmost columns provide the answer to the question of whether the null hypothesis of each test is rejected at the 10% level or less, the largest rejection region normally used in empirical work, and whether the results of the tests are in agreement. Given the opposite null hypotheses, the tests are in agreement when one indicates rejection and the other indicates non-rejection; thus rejection in the KPSS and non-rejection in the ADF tests are consistent evidence of non-stationarity in the data.

The evidence is overwhelmingly on the side of non-stationarity for the HCE series, both individually and in panel. The KPSS test indicates rejection of stationarity for all 50 series individually; the ADF test indicates non-rejection of the null of non-stationarity in 47 states. There are no cases where both tests agree on stationarity of HCE.

For the panel of states, the Pesaran (2004) "*CD*" test of cross-section independence, specified as $CD = \sqrt{\frac{2T}{N(N-1)}}\left(\sum_{i=1}^{N-1}\sum_{j=I+1}^{N}\hat{r}_{ij}\right)$, where $\widehat{r_{ij}}$ are the contemporaneous cross correlations of the residuals from equation (2), is distributed as a standard normal variant. A *CD* statistic exceeding the critical value indicates that the states' HCEs are cross-dependent. The critical values for the panel unit-root and stationarity tests must then be estimated via the bootstrap.

As shown in Panel B of Table 2, the panel tests agree with the tests from the individual series. The Maddala-Wu unit root test fails to reject the null of non-stationarity for the panel and the Hadri test soundly rejects the null of stationarity.

The evidence of non-stationarity of state DPI in Table 3 is strong but not as overwhelming as that for HCE. The KPSS test indicates rejection of stationarity for 47 states' DPI, but the ADF test rejects the unit root hypothesis in 14 instances. The tests agree on non-stationarity in 36 states; there are three instances in which the two tests agree on stationarity. For the panel, the *CD* test indicates cross-dependence here as well, so the bootstrap is again used to generate critical values for the panel tests. The Maddala-Wu test statistic is well out of the rejection region, and the Hadri LM_N test statistic indicates rejection of the stationarity null at less than the 5 percent critical value.

The results of these tests for the U.S. states agree with Gerdtham and Lothgren (2000) using international data and with Wang and Rettenmaier (2007) and Moscone and Tosetti (2010) for the U.S. states using a shorter time interval. In the above cases only Wang and Rettenmaier consider structural breaks in the testing framework, and in that case only a single break was incorporated, as the time series, at 21 annual observations, were relatively short. Carrion-i-Silvestre (2005), on the other hand, extends stationarity testing on international HCE and income to a framework allowing multiple structural breaks. Carrion-i-Silvestre finds up to five structural breaks for HCE and four for GDP in a panel of OECD countries. After accounting for the breaks and for cross-country dependence in the sample, Carrion-i-Silvestre is able to conclude using a Hadri-type stationarity test that HCE and GDP are stationarity.[6]

In the following section, Carrion-i-Silvestre's methodology allowing multiple structural breaks is incorporated in stationarity tests. Using this methodology produces results that cast doubt on the premise that U.S. states' HCE and DPI have similar orders of integration.

[6] Strictly speaking, this is true only at the 5 percent level; Carrion-i-Silvestre's preferred test rejects the null of stationarity for HCE at the 10 percent level, a finding that has implications for what follow in in the succeeding section.

III. Stationarity Tests with Structural Breaks

The unit root versus trend-stationary problem in the previous section is essentially the question of whether the trend is changing every period or never. By introducing the possibility of structural breaks, the question can be changed to one of whether the trend is changing every period or "infrequently".[7] These infrequent changes are large enough to cause a significant change in some feature of the data, such as the level or trend, and are likely to stay in effect for a significant subsample of the data. There is no settled approach to identifying these changes, but a number of advances have been made since the original paper by Perron (1989) demonstrating how incorporating the 1970s oil shock could reverse the conclusion of a unit root in many U.S. macroeconomic time series.

In this paper, we use the Carrion-i-Silvestre (2005) test that considers breaks in both levels and trends. To operationalize the test, expand the specification of $r_{i,t}$ in (3) to include additional deterministic terms in levels and trends:

$$r_{i,t} = r_{i,t-1} + \sum_{k=1}^{m_i} \theta_{i,k} D(T_{b,k}^{i})_t + \sum_{k=1}^{m_i} \gamma_{i,k} DU_{i,k,t} + \mu_{i,t} \qquad (6)$$

where $\mu_{i,t} \sim iid(0, \sigma_{u,i}^2)$ and $r_{i,0} = r_i$, a constant, with i =1, . . . , N states and t = 1, . . ., T time periods. The dummy variables $D(T_{b,k}^{i})_t$ and $DU_{i,k,t}$ are defined as $D(T_{b,k}^{i})_t$= 1 for $t = T_{b,k}^{i} + 1$ and zero elsewhere, and $DU_{i,k,t}$ = 1 for $t > T_{b,k}^{i}$ and 0 elsewhere, with $T_{b,k}^{i}$ denoting the k-th date of the break for the i-th individual for k = 1, . . ., m_i. Hence the null hypothesis of stationarity is equivalent to H_0: $\sigma_{u,i}^2 = 0$, in which case combining (2) and (6) gives:

$$y_{i,t} = a_i + \beta_i t + \sum_{k=1}^{m_i} \theta_{i,k} DU_{i,k,t} + \sum_{k=1}^{m_i} \gamma_{i,k} DT_{i,k,t}^{*} + \epsilon_{i,t} \qquad (7)$$

[7] Of course, "infrequently" in this context can also mean "never".

with $DT^{*}_{i,k,t} = t - T^{i}_{b,k}$ for $t > T^{i}_{b,k}$ and 0 elsewhere and $a_i = \alpha_i + r_i$. The model in (7) allows for shifts in both levels and trends for each state in the panel, and is completely flexible regarding the number of breaks and the timing.

As a practical matter, some limit must be placed on the number of breaks, m_i, and for the length of the interval $(T^{i}_{b,k+1} - T^{i}_{b,k})$. Too many breaks or too short intervals call into question the meaning of "stationarity". Given the length of the sample, 44 years, we restrict the maximum number of breaks to 3, resulting in at most 4 intervals. We also restrict the minimum interval length to 7 years, which is approximately 16 percent of the sample. Zivot and Andrews (1992) and Perron (2005) cite 0.15 as a popular choice for ε in defining the range [ε, 1 - ε] to search for possible breaks.

To identify the location and number of breaks, we use the sequential search procedure of Bai and Perron (1998). Equation (7) is estimated for each state for HCE and DPI for no breaks and for all possible breaks in the interval [1973, 2002]. The Modified Information Criterion (MIC) of Liu, Wu, and Zidek (LWZ, 1997) is used to compare the no-break case with the minimum MIC from the single break case. If $MIC_i(0) < MIC_i(1)$ for all $T^{i}_{b,1}$ in [1973, 2002], then no structural break is identified. Otherwise, the $T^{i}_{b,1}$ that minimizes $MIC_i(1)$ is identified as a structural break, and the search for a second break commences on either side of $T^{i*}_{b,1}$, provided that the minimum interval limit is available. If a $T^{i}_{b,2}$ is found such that $MIC_i(2) <$ min $MIC_i(1)$, then the series is adjudged to have two breaks, and the search for a third break commences within the interval(s) of minimum length 7 among the three available .

After identifying the optimum number of breaks, we then test for stationarity using the residuals from (7) in the KPSS test for the individual states and the Hadri-type test for the HCE and DPI panels. The asymptotic critical values for the no-break cases clearly do not apply, so critical values must again be computed via the bootstrap. Carrion-i-Silvestre et al. (2005) provide the details of the Hadri-type test for multiple structural breaks and the methodology for calculating the expected value and variance of the test statistic.

[Tables 4 and 5 about here]

Tables 4 and 5 provide the results of structural break test and stationarity tests for HCE and DPI for the fifty states, individually and in panel. In Table 4 the sequential break procedure identifies structural breaks for HCE in 48 states, with 45 states identified as having two breaks and 14 identified with three breaks. Only in Massachusetts and Montana were no breaks identified. The most common year of a structural break is 1979 with 21 occurrences, followed by 1991 with 14, and 2002 with 11.

The KPSS test statistic fell into the 90^{th} percentile or above in 32 states, with 19 exceeding the 95^{th} percentile and 2 exceeding the 99^{th} percentile of the empirical distribution. Thus in the majority of the states the null hypothesis of stationarity is rejected even when accounting for structural breaks. As shown in Carrion-i-Silvestre, et al. (2005), the expected value of the KPSS statistic with structural breaks depends on the number of breaks and their position within the sample. The greater the number of breaks, the smaller will be the asymptotic expectation of the KPSS statistic.

For the Hadri-type test in the panel setting, the critical values were computed using a bootstrap methodology that draws the residuals while maintaining the contemporaneous error structure. For each draw, an index *s* was randomly chosen from 1 through 41. The index was used to select residuals in blocks of 4:

$\hat{\epsilon}_{1,s}, \ldots, \hat{\epsilon}_{50,s};\ \hat{\epsilon}_{1,s+1}, \ldots, \hat{\epsilon}_{50,s+1}; \hat{\epsilon}_{1,s+2}, \ldots, \hat{\epsilon}_{50,s+3}; \hat{\epsilon}_{1,s+4}, \ldots, \hat{\epsilon}_{50,s+4}$.[8] In this way the contemporaneous cross-correlation of the residuals is preserved, and any autocorrelation specific to the states is preserved as well. This process was repeated 11 times with replacement for each replication to produce 50 series of 44 residuals. The values of the pseudo-sample $\hat{y}_{i,t}$ were generated from these residuals. The critical values were computed as percentiles of 5000 replications of the procedure. As shown in Panel B of Table 4, the panel stationarity test indicates rejection of the null at the one percent level.

For the comparable analysis of state DPI, Table 5 shows that 43 states had at least one structural break, 23 with at least two breaks, and 11 with three breaks. State DPI averages exactly 0.5 fewer breaks than state HCE. The modal break year for DPI is 1980, with 15 occurrences, followed by 1991, with 9 occurrences, and 1974, with 8 occurrences. Unsurprisingly, all three correspond to national business cycle peaks.

Along with fewer identified breaks, an even larger difference between DPI and HCE behavior at the state level is reflected in the results of the stationarity tests. The KPSS test statistic indicates rejection at the 10 percent level or less in only 14 cases, with 6 states falling into the 95th percentile and 2 states into the 99th percentile, using the bootstrapped critical values. The panel stationarity test falls well short of the critical region for rejection, indicating that the panel of states' DPI is stationary around changing levels and trends.

These results are similar to those of Carrion-i-Silvestre's (2005) for an international panel of HCE and GDP for 20 OECD countries. Carrion-i-Silvestre found an average of 2.65 breaks for HCE versus 2.3 breaks for GDP, and the panel test statistic for stationarity was much larger

[8] There is no sure way to choose optimal block size. Too large a block and randomness is lost; too small and the features of the data are lost. In a Monte Carlo simulation, van der Plas (2010) found that block size *b = T/10* performed reasonably well.

for HCE than for GDP while the critical values from the bootstrapped distribution were much smaller. If a test size of 0.10 were used instead of 0.05, Carrion-i-Silvestre would have been able to reject the null of stationarity for the panel test of international HCE, but not for GDP, the result that we find using the U.S. states and DPI.

Thus we are confronted with findings that are contradictory to previous research: HCE and DPI may be of different orders of integration. If so, this presents something of a dilemma when estimating the relationship between the two. If both are stationary, regression in levels is consistent and efficient; if both are non-stationary and not cointegrated, regression in first differences is in order; if both are non-stationary and cointegrated, regression in first differences with the addition of the cointegrating term is in order.

Given the uncertainty about the order of integration between HCE and DPI, we turn to a methodology developed by Pesaran, Shin and Smith (PSS, 2001) to deal with just this eventuality. This approach uses a "bounds test" with two sets of critical values providing a band of coverage for all possible classifications of the regressors with respect to orders of integration. We employ this methodology in the following section to estimate the level relationship between HCE and DPI.

IV. Estimates of HCE Income Elasticity

In this section we employ a number of specifications for estimating the response of HCE to income so as to compare the present research with previous research and to provide confirmatory analysis of our results.

The model for estimation is the conditional error correction model (ECM) developed in PSS (2001):

$$\Delta y_{it} = \boldsymbol{\alpha}_i'\boldsymbol{d}_t + \gamma y_{it-1} + \boldsymbol{\beta}_i'\boldsymbol{x}_{\boldsymbol{it-1}} + \sum_1^k \boldsymbol{\delta}_{\boldsymbol{ik}}\boldsymbol{\Delta z}_{\boldsymbol{it-k}} + \boldsymbol{\omega}'\Delta\boldsymbol{x}_{it} + e_{it}, \qquad \textbf{(8)}$$

where y_{it} is HCE per capita in state i in year t; $\boldsymbol{d_t}$ is a vector of observed deterministic effects, such as state-specific intercepts, time trends, or seasonal effects; $\boldsymbol{x_{it}}$ is a $k \times 1$ vector of regressors including DPI; $\boldsymbol{z_{it}} = (y_{it}, \boldsymbol{x_{it}})$; and the errors, e_{it}, may be contemporaneously correlated.[9]

In addition to income, the regressor set includes the public share of health care expenditures, calculated as the sum of Medicare and Medicaid payments; the share of the population under 15; and the share of the population over 65. This set of regressors is similar to that used by Baltagi and Moscone (2010) in their panel analysis of international health expenditures.

The null hypothesis of no long-term relationship between HCE and the regressors is the joint hypothesis $H_0 : \gamma_i = 0$ and $\boldsymbol{\beta}_i' = 0$. PSS have published asymptotic critical values for the F-statistic testing exclusion of the lagged levels of the dependent variable and the regressors from (8). These values differ according to the dimension of $\boldsymbol{x_{it}}$, but not according to the number of time periods T. PSS do show, however, that the critical values are not particularly sensitive to T in their empirical example.

Table 6 presents the results of the test of a long-run relationship between HCE and DPI using DPI only. The test is conducted over two sample periods, 1974-2009 and 1980-2004. The first period corresponds to the full dataset available, allowing for the inclusion of lags in (8); the

[9] Identification of the model requires that contemporaneous income levels affect health expenditures, but not the reverse.

second period corresponds to the dataset used by Moscone and Tosetti (2010), allowing for both a comparison with their results, and a test of the sensitivity of the estimates to a change in the sample.

[Table 6 about here]

The results in Table 6 provide very little evidence of a stable bivariate relationship for either time period between HCE and DPI at the state level. The hypothesis of no long-term relationship between HCE and DPI cannot be rejected for 45 states in the full sample or for 38 states in the shorter sample. The average elasticities differ markedly from the larger to the smaller sample. The large standard deviations of the distribution of state estimates reflect the broad dispersion of the individual estimates, including nearly half the elasticities as negative.

The wide variation of individual state estimates is a common phenomenon in measuring HCE response to income. The standard deviation of Moscone and Tosetti's (2010) individual state estimates as reported in their Table VII is 0.567, and that of Yang and Rettenmaier (2007) in their Table V is 0.719. While it is true that the average itself can be measured with great precision, as for example, when pooling cross sections, it is still an open question as to whether that average is conveying useful information regarding individual states.

Because of the presence of extreme values, a trimmed average of the middle 80 percent of estimated elasticities is also presented. The gap between the averages of the full and the abbreviated samples is diminished but remains relatively large. Nor can the gap be explained by a simple scaling up of the coefficients from the larger sample; the correlation coefficient for the individual state estimates between the larger and smaller samples is only 0.34, statistically significant but small in magnitude considering the addition of only 11 additional years to the smaller sample.[10]

Turning to the full set of regressors in Table 7, we see much stronger evidence for the existence of a long-run relationship, especially in the shorter sample. For the full sample, in 20 states the null hypothesis of no relationship can be rejected, and is at least in doubt in another 12. For the shorter sample, the null of no relationship can be rejected in 46 states and is in doubt in 2 more. The addition of the controls for public funding and demographics clearly makes a substantive difference in the results.

[Table 7 about here]

The average elasticities in Table 7 are considerably larger than their counterparts in Table 6. In the case of the smaller sample, Colorado and Texas are outliers, with estimated elasticities of +94.1 and -17.9, respectively. Trimming the sample to the middle 80 percent results in a more reasonable average for the shorter sample but not much change for the full sample. In contrast to the outcome in Table 6, the average income elasticity is smaller for the shorter sample. Again, the average estimates are in agreement with previous research concluding that health care is a necessity good, but the averages come with a great deal of variation. Agreement across the two samples is even less when using the full set of regressors: the correlation coefficient for the two sets of individual state elasticities is only 0.191.

Given the instability in elasticity estimates using the levels approach, we employ a version of equation (8) above using only first differences of the variables:

[10] Agreement in health income elasticity estimates from previous research is even less: the correlation between estimates for the 48 states in common to both samples in Wang and Rettenmaier (2007) and Moscone and Tosetti (2010) is only 0.20.

$$\Delta y_{it} = \boldsymbol{\alpha}_i' \boldsymbol{d}_t + \sum_1^{\boldsymbol{k}} \boldsymbol{\delta}_{\boldsymbol{ik}} \boldsymbol{\Delta z}_{\boldsymbol{it-k}} + \boldsymbol{\omega}' \Delta \boldsymbol{x}_{it} + e_{it}, \qquad \textbf{(8')}$$

where, as before, y_{it} is HCE per capita in state i in year t; $\boldsymbol{d_t}$ is a vector of observed deterministic effects, such as state-specific intercepts, time trends, or seasonal effects; $\boldsymbol{x_{it}}$ is a $k \times 1$ vector of regressors including DPI; $\boldsymbol{z_{it}} = (y_{it}, \boldsymbol{x_{it}})$. Estimation will be state-by-state and pooled with time effects. Results are reported in Table 8.

[Table 8 about here]

In contrast to the levels results, the estimates for the first difference regressions are close in magnitude. Income elasticity falls in the range 0.177-0.260, depending on specification, and the dispersion of the individual state estimates is much reduced. In addition, the correlation between state estimates across the two samples is more than twice as large as the comparable levels comparison.

V. **Conclusion**

This paper estimates the income elasticity of health care expenditures using annual data on health spending by state in the U.S. from 1966-2009. Because of the variation in specifications and results in previous research, there is considerable importance in establishing the orders of integration for the time series used in the estimation. Panel stationarity tests incorporating structural breaks in the levels and trends in HCE and DPI yield inconsistent results, with stationarity rejected for HCE but not for DPI.

Because the results of the stationarity tests differed from those of previous research, we employed an additional test developed by Pesaran, Smith, and Shin (2001) for the presence of a long-run relationship in the event of uncertainty regarding the stationarity of the regressors. The PSS test indicates little evidence of a long-run bivariate relationship between HCE and DPI, but the addition of controls for the share of health care financed by public spending and for demographics strengthens the case for a multivariate relationship, especially in a sub-sample using the years 1981-2004. The results differed considerably depending on time period, and the cross-state variation in elasticity estimates was quite large.

Given the lack of consistency in results using levels specification, the model was re-estimated using first-differences in state-by-state and in pooled regressions. The results of the first difference models provide more consistent estimates across time periods, whether expressed as averages of individual state estimates or as pooled time series. Elasticities for the full sample fall in the range 0.21-0.22, below that of recent estimates by Moscone and Tosetti (2010) and much below that of Yang and Rettenmaier (2007).

The main conclusion to draw from this analysis is that the evidence is preponderant that health care is a necessity good, and that increases in HCE are probably driven less by income than by advances in treatments for difficult-to-cure maladies and by increases in the cost of health care inputs, including labor and research. If the increasing share of health care in the nation's budget is a concern, controlling these costs and finding more efficiency in health care delivery will be of paramount importance.

References

Bai J and P Perron (1998). Estimating and testing linear models with multiple structural breaks. *Econometrica* 66: 47-78.

Baltagi B and F Moscone (2010). Health care expenditure and income in the OECD reconsidered: evidence from panel data. IZA Discussion Paper 4851.

Bureau of Economic Analysis (2012). Regional economic accounts. Accessed June 1, 2012 at http://bea.gov/regional/index.htm.

Carrion-i-Silvestre J (2005). Health care expenditure and GDP: Are they broken-stationary? *Journal of Health Economics* 24: 839-854.

Carrion-i-Silvestre J, T Del Barrio, and E Lopez-Bazo (2003). Breaking the panels: an application to GDP per capita. University of Barcelona Working Paper.

Centers for Medicare and Medicaid Services (2011). Health expenditures by state of residence. Retrieved June 1, 2012 at http://www.cms.gov/NationalHealthExpendData/downloads/resident-state-estimates.zip.

Freeman D. (2003). Is health care a necessity or a luxury? Pooled estimates of income elasticity from U.S. state-level data. *Applied Economics* 35: 495-502.

Gerdtham U and M Lothgren (2000). On stationarity and cointegration of international health expenditures and GDP. *Journal of Health Economics* 19: 461-475.

Hadri, K. (2000). Testing for stationarity in heterogeneous panel data. *Econometrics Journal* 3:148-161.

Hansen P and A King (1996). The determinants of health care expenditure: a cointegration approach. *Journal of Health Economics* 15: 127-137.

Holly, S, M Pesaran and T Yamagata (2010). A spatio-temporal model of house prices in the USA. *Journal of Econometrics* 158(1): 160-173.

Im K, H Pesaran and Y Shin *(*2003*)*. Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115: 53-74.

Kleiman E. (1974). *The Determinants of National Outlay on Health.* Macmillan: London.

Kwiatowski D, P Phillips, P Schmidt, Y Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* 54: 159-178.

Levit K. (1985). Personal health care expenditures by state: 1966-1982. *Health Care Financing Review* 6: 1-49.

Liu J, S Wu, J Zidek (1997). On segmented multivariate regressions. *Statistica Sinica* 7: 497-525.

MacKinnon J. (1994). Approximate asymptotic distribution functions for unit root and cointegration tests. *Journal of Business and Economic Statistics* 12: 167-176.

Maddala G and S Wu (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics* 61: 631-652.

McCoskey S and T Selden (1998). Health care expenditure and GDP:panel data unit root tests. *Journal of Health Economics* 17: 369-376.

Moscone F. and E. Tosetti (2010). Health expenditure and income in the United States. *Health Economics* 19: 1385-1403.

Newhouse J. (1977). Medical care expenditure: a cross-national study. *Journal of Human Resources* 12: 115-125.

Perron P (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica* 57: 1361-1401.

Perron P (2005). Dealing with structural breaks. Boston College Working Paper.

Pesaran H (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica* 74: 967-1012.

Pesaran H and E Tosetti (2009). Large panels with common factors and spatial correlations. Mimeo, University of Cambridge.

Pesaran H, Y Shin, and R Smith (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics* 16: 289-326.

Pritchett L and L Summers (1996). Wealthier is healthier. *Journal of Human Resources* **31** (4): 841–868.

Van der Plas M (2010). Bootstrapping the KPSS stationarity test: a comparative analysis. Master Thesis, Maastricht University.

Wang Z. and A Rettenmaier (2007). A note on cointegration of health expenditures and income. *Health Economics* 16: 559-578.

Zivot E. and D Andrews (1992). Further evidence on the great crash, the oil price shock, and the unit root hypothesis. *Journal of Business and Economic Statistics* 10: 251-270.

Table 1. Summary Statistics of U.S. State-Level Health Care Expenditures (HCE) and Disposable Personal Income (DPI). Annual Data, 1966-2009

	Mean	Standard Deviation	Maximum	Minimum
HCE				
Levels ($2009)	6,970	937	9,277 (Massachusetts)	5,030 (Utah)
CAGR 1966-2009[a]	3.9	0.4	4.9 (Mississippi)	2.7 (California)
DPI				
Levels ($2009)	34,507	4,388	45,725 (Connecticut)	28,037 (Mississippi)
CAGR 1966-2009[a]	1.6	0.3	2.0 (New Hampshire)	0.8 (Alaska)

[a]Measured as the slope of a logarithmic trend

Table 2. Results of Stationarity and Unit-Root Tests for Per Capita Health Care Expenditure, 50 states, 1966-2009

Panel A: Tests of Individual Series	KPSS stat.	ADF Stat.	ADF *p*-val	Opt. Lag, ADF	KPSS reject at < 10%?	ADF reject at < 10%?	Tests agree: non-stationarity?	Tests agree: stationarity?
Alabama	0.734^{c}	-1.038	0.939	3	Y		Y	
Alaska	0.298^{c}	-3.509^{b}	0.038	10	Y	Y		
Arizona	0.631^{c}	-1.025	0.941	10	Y		Y	
Arkansas	0.783^{c}	-1.061	0.935	2	Y		Y	
California	0.642^{c}	-2.600	0.280	3	Y		Y	
Colorado	0.464^{c}	-2.046	0.576	5	Y		Y	
Connecticut	0.422^{c}	-2.093	0.550	3	Y		Y	
Delaware	0.318^{c}	-2.577	0.291	4	Y		Y	
Florida	0.872^{c}	-1.547	0.812	3	Y		Y	
Georgia	0.822^{c}	1.486	1.000	10	Y		Y	
Hawaii	0.687^{c}	-1.882	0.664	2	Y		Y	
Idaho	0.221^{b}	-2.037	0.581	3	Y		Y	
Illinois	0.547^{c}	-2.190	0.495	3	Y		Y	
Indiana	0.415^{c}	-1.875	0.667	3	Y		Y	
Iowa	0.247^{c}	-1.839	0.686	3	Y		Y	
Kansas	0.365^{c}	-2.507	0.325	3	Y		Y	
Kentucky	0.505^{c}	-0.893	0.957	3	Y		Y	
Louisiana	0.815^{c}	-1.337	0.878	3	Y		Y	
Maine	0.323^{c}	-2.386	0.387	3	Y		Y	
Maryland	0.497^{c}	-2.743	0.219	4	Y		Y	
Massachusetts	0.262^{c}	-2.751	0.215	3	Y		Y	
Michigan	0.814^{c}	-2.294	0.438	3	Y		Y	
Minnesota	0.627^{c}	-1.527	0.820	2	Y		Y	
Mississippi	0.552^{c}	-2.253	0.460	3	Y		Y	
Missouri	0.754^{c}	-2.244	0.465	3	Y		Y	
Montana	0.167^{b}	-2.579	0.290	3	Y		Y	
Nebraska	0.219^{b}	-3.064	0.115	3	Y		Y	
Nevada	0.712^{c}	-2.692	0.240	4	Y		Y	
New Hampshire	0.278^{c}	-3.130^{b}	0.099	5	Y	Y		
New Jersey	0.518^{c}	-1.655	0.770	3	Y		Y	
New Mexico	0.353^{c}	-2.269	0.451	3	Y		Y	
New York	0.280^{c}	-2.790	0.201	5	Y		Y	
North Carolina	0.377^{c}	-1.182	0.914	9	Y		Y	
North Dakota	0.573^{c}	-2.893	0.165	3	Y		Y	
Ohio	0.744^{c}	-1.735	0.735	3	Y		Y	
Oklahoma	0.191^{b}	-2.400	0.379	2	Y		Y	
Oregon	0.233^{c}	-4.788^{c}	0.000	9	Y	Y		
Pennsylvania	0.844^{c}	-0.810	0.965	8	Y		Y	
Rhode Island	0.279^{c}	-2.930	0.153	3	Y		Y	
South Carolina	0.484^{c}	-1.163	0.918	3	Y		Y	
South Dakota	0.322^{c}	-1.710	0.746	3	Y		Y	
Tennessee	0.860^{c}	-0.421	0.986	2	Y		Y	
Texas	0.773^{c}	-1.536	0.816	3	Y		Y	
Utah	0.503^{c}	-2.545	0.306	3	Y		Y	
Vermont	0.684^{c}	-2.464	0.346	3	Y		Y	
Virginia	0.783^{c}	-0.673	0.975	10	Y		Y	
Washington	0.270^{c}	-2.653	0.256	3	Y		Y	
West Virginia	0.405^{c}	-1.800	0.705	3	Y		Y	
Wisconsin	0.249^{c}	-2.526	0.315	7	Y		Y	
Wyoming	0.777^{c}	-2.114	0.538	2	Y		Y	
State Totals				4.1*	50	3	47	0

Panel B: Tests of Panels of States	Test Statistic	Value	Critical Values, Empirical Distribution (%)		
Pesaran test of cross-section independence	*CD*	5.00^{c}	90	95	99
Maddala-Wu Unit Root	P_{λ}	98.96	292.24	318.52	380.81
Hadri Stationarity	LM_{N}	75.01^{c}	41.08	47.40	58.23

*Average. $^{a, b, c}$ reject at 10%, 5%, and 1% levels respectively. The asymptotic critical values for the KPSS statistic at the 90th, 95th, and 99th percentiles are 0.120, 0.149, and 0.229, respectively.

Table 3. Results of Stationarity and Unit-Root Tests for Per Capita Disposable Income, 50 states, 1966-2009

Panel A: Tests of Individual States	KPSS stat.	ADF Stat.	ADF *p*-val	Opt. Lag, ADF	KPSS reject?	ADF reject?	Tests agree: non-stationarity?	Tests agree: stationarity?
Alabama	0.359[c]	-3.841	0.015	8	Y	Y		
Alaska	0.516[c]	-2.480	0.338	4	Y		Y	
Arizona	0.279[c]	-3.085	0.110	3	Y		Y	
Arkansas	0.307[c]	-2.872	0.172	7	Y		Y	
California	0.281[c]	-2.929	0.153	4	Y		Y	
Colorado	0.250[c]	-2.638	0.263	8	Y		Y	
Connecticut	0.297[c]	-1.917	0.646	3	Y		Y	
Delaware	0.166[b]	-3.209	0.083	6	Y	Y		
Florida	0.330[c]	-2.926	0.154	3	Y		Y	
Georgia	0.309[c]	-1.621	0.784	3	Y		Y	
Hawaii	0.232[c]	-3.215	0.081	7	Y	Y		
Idaho	0.273[c]	-2.885	0.167	2	Y		Y	
Illinois	0.123[a]	-2.501	0.327	8	Y		Y	
Indiana	0.128[a]	-2.724	0.226	8	Y		Y	
Iowa	0.242[c]	-2.612	0.274	2	Y		Y	
Kansas	0.291[c]	-3.726	0.021	2	Y	Y		
Kentucky	0.228[b]	-3.615	0.029	8	Y	Y		
Louisiana	0.403[c]	-2.354	0.404	3	Y		Y	
Maine	0.110	-3.900	0.012	8		Y		Y
Maryland	0.229[c]	-2.765	0.210	8	Y		Y	
Massachusetts	0.177[b]	-4.126	0.006	6	Y	Y		
Michigan	0.129[b]	-2.049	0.574	3	Y		Y	
Minnesota	0.111	-3.606	0.029	8		Y		Y
Mississippi	0.240[c]	-2.982	0.137	2	Y		Y	
Missouri	0.075	-4.396	0.002	8		Y		Y
Montana	0.419[c]	-1.809	0.700	8	Y		Y	
Nebraska	0.158[b]	-2.502	0.327	8	Y		Y	
Nevada	0.178[b]	-3.524	0.037	3	Y	Y		
New Hampshire	0.318[c]	-2.722	0.227	4	Y		Y	
New Jersey	0.175[b]	-2.361	0.400	8	Y		Y	
New Mexico	0.392[c]	-2.591	0.284	8	Y		Y	
New York	0.195[b]	-2.028	0.586	7	Y		Y	
North Carolina	0.370[c]	-1.585	0.798	3	Y		Y	
North Dakota	0.244[c]	-2.761	0.212	2	Y		Y	
Ohio	0.155[b]	-4.264	0.004	8	Y	Y		
Oklahoma	0.467[c]	-3.141	0.097	8	Y	Y		
Oregon	0.195[b]	-2.762	0.211	3	Y		Y	
Pennsylvania	0.144[a]	-4.299	0.003	8	Y	Y		
Rhode Island	0.192[b]	-2.873	0.171	3	Y		Y	
South Carolina	0.292[c]	-2.502	0.327	4	Y		Y	
South Dakota	0.288[c]	-2.102	0.545	8	Y		Y	
Tennessee	0.317[c]	-2.068	0.564	3	Y		Y	
Texas	0.428[c]	-2.306	0.431	2	Y		Y	
Utah	0.368[c]	-2.618	0.272	8	Y		Y	
Vermont	0.188[b]	-3.054	0.118	8	Y		Y	
Virginia	0.510[c]	-2.270	0.451	3	Y		Y	
Washington	0.303[c]	-3.007	0.130	2	Y		Y	
West Virginia	0.244[c]	-3.470	0.043	8	Y	Y		
Wisconsin	0.201[b]	-2.942	0.149	8	Y		Y	
Wyoming	0.507[c]	-2.380	0.390	6	Y		Y	
State Totals				5.5[*]	47	14	36	3

Panel B: Tests of Panels of States	Test Statistic	Value	Critical Values, Empirical Distribution (%)		
Pesaran test of cross-section independence	CD	4.59[c]	90	95	99
Maddala-Wu Unit Root	P_{λ}	201.16	311.08	335.49	398.25
Hadri Stationarity	LM_N	33.84[b]	27.90	31.48	41.08

[*]Average. [a, b, c] reject at 10%, 5%, and 1% levels respectively. The asymptotic critical values for the KPSS statistic at the 90th, 95th, and 99th percentiles are 0.120, 0.149, and 0.229, respectively.

Table 4 Structural Break and Stationarity Tests for Per Capita Health Care Expenditures, 50 states, 1966-2009

Panel A: Individual State Results	m	Break 1	Break 2	Break 3	KPSS	Critical Values, Empirical Distribution (%)		
						90	95	99
Alabama	2	1991	1979		0.059[a]	0.057	0.064	0.078
Alaska	3	1991	1980	2001	0.051[b]	0.044	0.046	0.052
Arizona	2	1995	1974		0.072[a]	0.067	0.075	0.091
Arkansas	2	1991	1974		0.081[b]	0.072	0.079	0.092
California	2	1975	1995		0.057[a]	0.054	0.061	0.073
Colorado	2	1994	1979		0.058[b]	0.047	0.051	0.060
Connecticut	3	1988	1980	1973	0.034	0.045	0.052	0.071
Delaware	2	1986	1979		0.065	0.074	0.086	0.111
Florida	2	1993	1979		0.059[b]	0.051	0.056	0.065
Georgia	2	1993	1979		0.062[b]	0.057	0.062	0.074
Hawaii	2	1997	1974		0.047	0.050	0.056	0.069
Idaho	2	1991	1979		0.064[a]	0.058	0.067	0.084
Illinois	3	1991	1974	2002	0.071[b]	0.060	0.066	0.077
Indiana	2	1990	1980		0.069[a]	0.065	0.076	0.097
Iowa	2	1991	1980		0.063[c]	0.046	0.052	0.063
Kansas	2	1974	1991		0.057[a]	0.052	0.058	0.071
Kentucky	3	1990	2002	1980	0.053[b]	0.048	0.052	0.059
Louisiana	2	1995	1974		0.089[b]	0.079	0.088	0.104
Maine	2	1979	2002		0.062	0.067	0.080	0.103
Maryland	3	1974	1986	1995	0.042[b]	0.037	0.041	0.052
Massachusetts	0				0.262	0.273	0.324	0.433
Michigan	2	1988	1979		0.055[a]	0.055	0.066	0.087
Minnesota	3	1991	2002	1975	0.029[a]	0.027	0.030	0.034
Mississippi	2	1979	1991		0.045[a]	0.045	0.050	0.059
Missouri	2	1978	1991		0.041	0.044	0.049	0.061
Montana	0				0.167	0.220	0.265	0.381
Nebraska	2	1979	2001		0.055	0.058	0.065	0.079
Nevada	3	1979	1995	2002	0.045[b]	0.038	0.043	0.052
New Hampshire	1	1985			0.091	0.094	0.105	0.127
New Jersey	3	1990	1980	1973	0.033	0.039	0.047	0.070
New Mexico	2	1988	1979		0.064[a]	0.056	0.066	0.083
New York	2	1987	1974		0.067	0.074	0.084	0.101
North Carolina	3	1989	1980	2002	0.056[c]	0.047	0.050	0.056
North Dakota	2	1978	1985		0.046	0.056	0.067	0.096
Ohio	3	1993	1979	2002	0.041[b]	0.036	0.039	0.045
Oklahoma	2	1974	1991		0.079[b]	0.068	0.075	0.093
Oregon	2	1979	1994		0.047[b]	0.038	0.041	0.050
Pennsylvania	1	1994			0.114	0.137	0.154	0.184
Rhode Island	2	1986	1979		0.062	0.068	0.082	0.107
South Carolina	3	1991	1980	2002	0.058[c]	0.046	0.050	0.056
South Dakota	3	1990	1980	2002	0.025	0.030	0.032	0.038
Tennessee	3	1990	1979	2002	0.037[a]	0.037	0.040	0.044
Texas	2	1994	1979		0.058[b]	0.051	0.054	0.062
Utah	2	1994	1979		0.055[a]	0.051	0.056	0.067
Vermont	2	1979	2002		0.075[a]	0.073	0.084	0.106
Virginia	2	1993	1979		0.055[b]	0.049	0.055	0.067
Washington	2	1995	1979		0.060[b]	0.050	0.054	0.062
West Virginia	2	1991	1977		0.064[b]	0.051	0.056	0.066
Wisconsin	1	1978			0.111	0.128	0.156	0.207
Wyoming	3	1982	1991	1974	0.057[c]	0.029	0.034	0.043
# states with breaks at least		48	45	14	# states > crit. value	32	19	4

Panel B: Panel Stationarity test			Test Statistic	Critical Values, Empirical Distribution (%)		
				90	95	99
	LM(λ)		8.214[c]	6.212	6.876	8.124

[a, b, c] reject at 10%, 5%, and 1% levels respectively.

Table 5 Structural Break and Stationarity Tests for Per Capita Disposable Income, 50 states, 1966-2009

Panel A: Individual State Results	m	Break 1	Break 2	Break 3	KPSS	Critical Values, Empirical Distribution (%)		
						90	95	99
Alabama	1	1979			0.145[a]	0.139	0.162	0.218
Alaska	2	1979	1998		0.042	0.048	0.052	0.066
Arizona	2	1974	1991		0.042	0.047	0.053	0.066
Arkansas	1	1979			0.053	0.078	0.093	0.131
California	1	1991			0.097	0.104	0.116	0.138
Colorado	3	1974	2000	1991	0.031	0.034	0.039	0.049
Connecticut	3	1984	1991	1974	0.029	0.035	0.042	0.054
Delaware	1	1984			0.099	0.107	0.120	0.147
Florida	3	1974	1991	1983	0.041[a]	0.039	0.047	0.058
Georgia	1	2002			0.100	0.165	0.198	0.264
Hawaii	1	1972			0.163	0.210	0.249	0.338
Idaho	1	1979			0.079	0.086	0.097	0.125
Illinois	2	1980	1990		0.050[b]	0.036	0.042	0.053
Indiana	2	1980	2002		0.042	0.051	0.060	0.083
Iowa	1	1980			0.077	0.078	0.095	0.131
Kansas	2	1974	1998		0.049	0.065	0.075	0.098
Kentucky	1	1980			0.116[b]	0.088	0.105	0.144
Louisiana	3	1986	1978	1999	0.020	0.024	0.026	0.031
Maine	0				0.110	0.223	0.266	0.371
Maryland	1	1974			0.141	0.190	0.224	0.286
Massachusetts	2	1984	1991		0.067[a]	0.063	0.070	0.084
Michigan	0				0.129	0.181	0.212	0.310
Minnesota	0				0.111	0.198	0.238	0.336
Mississippi	2	1980	1972		0.044[a]	0.040	0.047	0.062
Missouri	0				0.075	0.174	0.212	0.295
Montana	3	1980	1998	1972	0.024[b]	0.020	0.023	0.028
Nebraska	1	1979			0.059	0.084	0.101	0.143
Nevada	1	1980			0.057	0.074	0.087	0.120
New Hampshire	1	1984			0.098	0.099	0.122	0.156
New Jersey	0				0.175	0.244	0.286	0.369
New Mexico	3	1980	2001	1987	0.033	0.033	0.035	0.041
New York	3	1984	1991	1974	0.058[c]	0.037	0.043	0.054
North Carolina	1	2001			0.105	0.153	0.179	0.247
North Dakota	1	1976			0.191[c]	0.112	0.134	0.188
Ohio	0				0.155	0.203	0.248	0.344
Oklahoma	3	1987	1998	1973	0.017	0.026	0.029	0.036
Oregon	1	1980			0.105[b]	0.082	0.100	0.133
Pennsylvania	1	1979			0.122	0.138	0.165	0.227
Rhode Island	2	1984	1991		0.045[a]	0.042	0.049	0.061
South Carolina	3	1974	1984	2000	0.036[a]	0.036	0.039	0.045
South Dakota	2	1980	1972		0.067[a]	0.057	0.072	0.105
Tennessee	3	1972	1984	2002	0.036	0.038	0.041	0.050
Texas	2	1980	1998		0.036	0.042	0.046	0.054
Utah	2	1980	1997		0.040	0.043	0.047	0.053
Vermont	0				0.188	0.229	0.285	0.407
Virginia	1	1991			0.091	0.094	0.104	0.125
Washington	1	1980			0.108	0.111	0.135	0.176
West Virginia	1	1980			0.040	0.056	0.067	0.085
Wisconsin	2	1980	1998		0.048[a]	0.043	0.049	0.061
Wyoming	3	1983	1998	1978	0.027	0.034	0.037	0.042
# states with breaks at least		43	23	11	# states > crit. value	14	6	2

Panel B: Panel Stationarity test		Test Statistic	Critical Values, Empirical Distribution (%)		
			90	95	99
	LM(λ)	6.171	11.931	13.294	16.602

[a, b, c] reject at 10%, 5%, and 1% levels respectively.

Table 6. Bounds Test of a Long-Term Bivariate Relationship Between HCE and DPI

	1974-2009					1980-2004				
			Long-term relationship					Long-term relationship		
	Income Elasticity	F-Statistic	No	Yes	?	Income Elasticity	F-Statistic	No	Yes	?
Alabama	-2.646	0.608	1	0	0	1.608	1.605	1	0	0
Alaska	0.919	4.332	1	0	0	1.550	28.444	0	1	0
Arizona	2.210	3.434	1	0	0	2.757	5.111	1	0	0
Arkansas	-3.568	1.079	1	0	0	0.094	0.738	1	0	0
California	2.142	9.362	0	1	0	14.745	8.747	0	1	0
Colorado	0.873	1.919	1	0	0	4.740	2.714	1	0	0
Connecticut	1.919	7.268	0	0	1	1.771	5.792	1	0	0
Delaware	0.996	2.172	1	0	0	0.959	3.179	1	0	0
Florida	0.921	1.013	1	0	0	1.692	2.912	1	0	0
Georgia	1.066	0.780	1	0	0	1.627	2.208	1	0	0
Hawaii	0.182	1.249	1	0	0	1.009	4.461	1	0	0
Idaho	-0.059	2.694	1	0	0	-2.719	6.713	0	0	1
Illinois	-0.078	1.308	1	0	0	0.350	1.884	1	0	0
Indiana	-0.566	2.290	1	0	0	-1.163	2.915	1	0	0
Iowa	-0.301	2.242	1	0	0	-0.856	15.442	0	1	0
Kansas	-0.821	4.779	1	0	0	-0.953	9.098	0	1	0
Kentucky	-1.857	0.802	1	0	0	0.149	0.826	1	0	0
Louisiana	-0.568	0.771	1	0	0	-2.252	3.870	1	0	0
Maine	0.996	3.516	1	0	0	0.320	36.821	0	1	0
Maryland	1.530	3.117	1	0	0	1.157	3.449	1	0	0
Massachusetts	1.865	13.904	0	1	0	2.141	8.758	0	1	0
Michigan	-0.231	1.796	1	0	0	-2.205	4.249	1	0	0
Minnesota	0.213	0.878	1	0	0	1.190	12.672	0	1	0
Mississippi	-0.359	1.413	1	0	0	-6.914	11.100	0	1	0
Missouri	0.228	3.525	1	0	0	0.991	2.810	1	0	0
Montana	-0.357	3.933	1	0	0	-0.543	2.030	1	0	0
Nebraska	0.613	5.181	1	0	0	0.471	3.859	1	0	0
Nevada	-0.555	3.781	1	0	0	-2.746	4.094	1	0	0
New Hampshire	1.280	8.719	0	1	0	1.628	7.390	0	1	0
New Jersey	2.108	2.103	1	0	0	2.719	3.948	1	0	0
New Mexico	-0.783	2.232	1	0	0	0.452	0.975	1	0	0
New York	0.758	3.613	1	0	0	0.933	3.609	1	0	0
North Carolina	1.797	6.093	1	0	0	1.857	5.085	1	0	0
North Dakota	-0.058	5.615	1	0	0	1.006	1.515	1	0	0
Ohio	0.010	1.536	1	0	0	0.726	2.993	1	0	0
Oklahoma	-0.492	5.206	1	0	0	-0.537	5.377	1	0	0
Oregon	0.263	3.424	1	0	0	-0.011	1.749	1	0	0
Pennsylvania	-6.500	0.908	1	0	0	0.725	5.271	1	0	0
Rhode Island	1.226	4.807	1	0	0	1.227	5.334	1	0	0
South Carolina	2.434	1.462	1	0	0	2.113	2.130	1	0	0
South Dakota	-0.241	3.373	1	0	0	-0.582	9.115	0	1	0
Tennessee	-1.834	0.233	1	0	0	1.430	2.476	1	0	0
Texas	-0.553	1.109	1	0	0	-0.415	1.240	1	0	0
Utah	-0.415	3.738	1	0	0	0.015	2.154	1	0	0
Vermont	1.088	4.819	1	0	0	0.602	9.713	0	1	0
Virginia	1.636	3.448	1	0	0	1.459	5.702	1	0	0
Washington	-0.428	3.788	1	0	0	0.810	2.437	1	0	0
West Virginia	-2.825	2.659	1	0	0	0.383	0.984	1	0	0
Wisconsin	0.486	2.397	1	0	0	0.176	2.287	1	0	0
Wyoming	-0.470	9.325	0	1	0	-0.129	6.212	1	0	0
State Totals			45	4	1			38	11	1
Mean	0.064					0.711				
Std Dev	1.610					2.678				
Trim Mean	0.240					0.634				
Trim Std Dev	0.834					0.844				

Notes to Tables 6 and 7: The null hypothesis of no long-term relationship between HCE and the regressors is the joint hypothesis $H_0 : \gamma_i = 0$ and $\boldsymbol{\beta}_i' = 0$ in the specification $\Delta y_{it} = \boldsymbol{\alpha}_i' \boldsymbol{d}_t + \gamma y_{it-1} + \boldsymbol{\beta}_i' \boldsymbol{x}_{it} + \sum_1^k \boldsymbol{\delta}_{ik} \boldsymbol{\Delta z}_{it-k} + \boldsymbol{\omega}' \Delta \boldsymbol{x}_{it} + e_{it}$. For the bivariate case in Table 6, the bounds for the F-statistic are (6.56, 7.30): less than the lower bound indicates no long term relationship, greater than the upper bound indicates a long term relationship, within the bounds indicates uncertainty about the relationship.

Table 7. Bounds Test of a Long-Term Relationship Between Health Care Expenditures and Disposable Personal Income, Public Funding, and Demographics

	1974-2009					1980-2004				
			Long-term relationship					Long-term relationship		
	Income Elasticity	*F*-Statistic	No	Yes	?	Income Elasticity	*F*-Statistic	No	Yes	?
Alabama	0.699	1.949	1	0	0	0.732	12.013	0	1	0
Alaska	0.565	0.899	1	0	0	1.575	22.250	0	1	0
Arizona	0.866	3.887	0	0	1	1.657	13.181	0	1	0
Arkansas	3.393	2.428	1	0	0	-0.854	16.532	0	1	0
California	1.340	4.388	0	1	0	-0.583	14.856	0	1	0
Colorado	1.446	2.819	1	0	0	94.135	6.266	0	1	0
Connecticut	1.459	4.616	0	1	0	1.087	53.572	0	1	0
Delaware	0.776	5.305	0	1	0	0.539	2.849	1	0	0
Florida	0.394	2.233	1	0	0	-1.819	11.366	0	1	0
Georgia	1.619	3.850	0	0	1	1.061	9.289	0	1	0
Hawaii	0.717	1.793	1	0	0	1.013	3.339	0	0	1
Idaho	0.529	6.070	0	1	0	-0.757	12.981	0	1	0
Illinois	-0.784	2.696	1	0	0	-1.056	4.873	0	1	0
Indiana	1.002	10.194	0	1	0	0.399	12.056	0	1	0
Iowa	1.757	5.169	0	1	0	-0.183	12.746	0	1	0
Kansas	-0.118	3.930	0	0	1	-1.501	4.930	0	1	0
Kentucky	0.537	2.658	1	0	0	0.132	7.951	0	1	0
Louisiana	-0.925	1.757	1	0	0	-1.157	1.823	1	0	0
Maine	0.670	3.570	0	0	1	0.293	12.859	0	1	0
Maryland	1.039	7.017	0	1	0	0.638	9.192	0	1	0
Massachusetts	1.689	6.660	0	1	0	-0.348	17.992	0	1	0
Michigan	0.210	2.955	1	0	0	-0.627	13.141	0	1	0
Minnesota	-2.375	3.797	0	0	1	0.805	3.291	0	0	1
Mississippi	1.172	1.849	1	0	0	-0.143	20.233	0	1	0
Missouri	0.114	3.034	1	0	0	0.647	4.718	0	1	0
Montana	0.061	6.216	0	1	0	0.291	6.959	0	1	0
Nebraska	-0.002	3.658	0	0	1	1.193	41.464	0	1	0
Nevada	-0.728	3.454	0	0	1	-2.001	7.540	0	1	0
New Hampshire	2.243	3.330	0	0	1	2.669	10.793	0	1	0
New Jersey	-0.574	3.057	1	0	0	0.567	9.688	0	1	0
New Mexico	0.834	3.382	0	0	1	1.545	5.359	0	1	0
New York	0.840	5.348	0	1	0	1.320	12.502	0	1	0
North Carolina	1.093	2.150	1	0	0	-0.461	9.304	0	1	0
North Dakota	0.295	8.228	0	1	0	0.451	5.076	0	1	0
Ohio	0.071	3.818	0	0	1	-0.657	10.546	0	1	0
Oklahoma	0.281	3.301	0	0	1	1.559	26.524	0	1	0
Oregon	1.383	2.836	1	0	0	0.489	5.546	0	1	0
Pennsylvania	0.904	5.283	0	1	0	0.199	7.429	0	1	0
Rhode Island	1.464	7.435	0	1	0	1.042	10.569	0	1	0
South Carolina	1.284	5.832	0	1	0	-0.206	25.561	0	1	0
South Dakota	0.315	6.472	0	1	0	0.344	11.214	0	1	0
Tennessee	0.917	0.782	1	0	0	0.368	7.916	0	1	0
Texas	-1.333	2.874	1	0	0	-17.869	4.632	0	1	0
Utah	1.210	4.234	0	0	1	0.306	5.236	0	1	0
Vermont	0.880	1.504	1	0	0	0.332	8.441	0	1	0
Virginia	1.057	4.755	0	1	0	0.511	8.899	0	1	0
Washington	0.554	6.439	0	1	0	0.441	7.098	0	1	0
West Virginia	-0.358	12.013	0	1	0	0.148	21.820	0	1	0
Wisconsin	1.450	5.471	0	1	0	0.289	18.931	0	1	0
Wyoming	-0.448	4.437	0	1	0	-0.283	15.426	0	1	0
State Totals			18	20	12			2	46	2
Mean	0.630					1.765				
Std Dev	0.956					13.605				
Trim Mean	0.672					0.307				
Trim Std Dev	0.546					0.654				

Notes to Tables 6 and 7: The null hypothesis of no long-term relationship between HCE and the regressors is the joint hypothesis $H_0 : \gamma_i = 0$ and $\boldsymbol{\beta}_i' = 0$ in the specification $\Delta y_{it} = \boldsymbol{\alpha}_i'\boldsymbol{d}_t + \gamma y_{it-1} + \boldsymbol{\beta}_i'\boldsymbol{x}_{it} + \sum_1^k \boldsymbol{\delta}_{ik}\Delta\boldsymbol{z}_{it-k} + \boldsymbol{\omega}'\Delta\boldsymbol{x}_{it} + e_{it}$. For the bivariate case in Table 6, the bounds for the *F*-statistic are (6.56, 7.30): less than the lower bound indicates no long term relationship, greater than the upper bound indicates a long term relationship, within the bounds indicates uncertainty about the relationship.

Table 8. First Difference Estimates of Health Care Expenditures on Disposable Personal Income, 50 States

Time Period	1974-2009		1980-2004	
	Income Elasticity	*t*-Statistic	Income Elasticity	*t*-Statistic
Alabama	0.2230	0.8542	0.0426	0.1047
Alaska	0.0318	0.1654	0.3765	0.7028
Arizona	0.1817	1.2111	0.0155	0.0516
Arkansas	0.1004	0.5413	0.0267	0.0847
California	0.2713	1.3957	0.3174	1.3362
Colorado	0.1155	0.6775	-0.0112	-0.0416
Connecticut	0.4561	2.5336	0.4481	2.0461
Delaware	0.3144	1.6942	0.0669	0.3357
Florida	0.0552	0.2739	0.1285	0.5146
Georgia	0.3789	2.2425	0.3160	1.4143
Hawaii	-0.2435	-1.1851	-0.2797	-1.1630
Idaho	-0.0416	-0.1187	-0.3646	-0.5338
Illinois	0.1464	0.6557	0.0425	0.1540
Indiana	0.2238	1.1027	-0.0500	-0.1473
Iowa	0.1639	0.6511	0.0981	0.2505
Kansas	0.0114	0.0403	-0.0583	-0.1389
Kentucky	0.1774	1.4380	0.1289	0.6240
Louisiana	0.3737	2.1040	0.4889	1.8198
Maine	0.2934	1.3867	0.0911	0.2701
Maryland	0.3300	1.4957	-0.0146	-0.0514
Massachusetts	0.4155	2.8692	0.3996	2.0841
Michigan	0.2236	1.7119	0.0174	0.1047
Minnesota	0.2086	1.2263	0.0990	0.3724
Mississippi	0.4058	0.9790	0.3572	0.4828
Missouri	0.3607	2.0114	0.3715	1.4506
Montana	0.2424	0.8092	0.6900	2.0205
Nebraska	0.1997	1.1775	0.2098	0.6611
Nevada	0.3812	1.9962	0.5262	1.4281
New Hampshire	0.3417	2.2695	0.2937	1.4573
New Jersey	0.2590	1.1863	0.3679	1.3866
New Mexico	0.4374	2.1535	0.5171	2.2029
New York	0.1654	1.0772	0.1485	1.2219
North Carolina	0.5528	2.8389	0.5124	1.5625
North Dakota	0.1016	0.6276	0.2918	2.8771
Ohio	0.1296	0.7704	-0.1118	-0.5369
Oklahoma	-0.1523	-0.8204	-0.0476	-0.1424
Oregon	0.3304	1.6040	0.3622	1.2512
Pennsylvania	0.2032	0.7942	-0.1820	-0.9902
Rhode Island	0.5035	2.8933	0.6123	2.2081
South Carolina	0.3416	1.1824	0.1115	0.2797
South Dakota	0.1496	0.8583	0.2358	0.7225
Tennessee	0.4671	2.8392	0.3298	1.6226
Texas	-0.0122	-0.0723	0.0129	0.0477
Utah	0.2215	0.8351	0.2978	0.6971
Vermont	0.3036	1.5817	0.1370	0.4759
Virginia	0.3580	1.8186	0.1074	0.5317
Washington	0.0314	0.1467	0.1464	0.4269
West Virginia	0.1303	0.7641	0.1561	0.5987
Wisconsin	0.4076	2.0311	0.3081	1.0598
Wyoming	-0.1094	-0.7190	-0.2569	-1.2595
State Average:	0.2230		0.1776	
	(0.173)*		(0.232)*	
Pooled Estimate:	0.2122		0.2601	
	(0.030)**		(0.032)**	

*Standard Deviation of state estimates **Standard error of pooled estimate