Stationarity of health expenditures: a re-examination using panel unit root tests with heterogeneous structural breaks

Todd Jewell [a], Junsoo Lee [b], Margie Tieslau [a, *], Mark C. Strazicich [a]

[a] *Department of Economics, University of North Texas, P.O. Box 311457, Denton, TX, 76203, USA*

[b] *Department of Economics, College of Business Administration, University of Central Florida, P.O. Box 161400, Orlando, FL 32816, USA*

September 20, 2001

1. Introduction

This paper analyzes the time series properties of health care expenditures (hereafter HE) using recently advanced econometric techniques in panel unit root tests. In recent decades, health care economists have devoted a great deal of attention to the time-series pattern of national HE, and their findings have had important policy implications. In particular, much of the early research on this subject found a strong positive correlation between HE and GDP in developed countries, with income elasticity greater than one. This suggests that health care is a "luxury good" and the share of HE in GDP will increase with per capita income. In general, these studies find that the overwhelming majority of variation in HE can be explained by variations in per capita GDP. This may imply that the government's ability to manage HE is severely limited (Culyer, 1990). Furthermore, if HE cannot be affected by public policy, then health care cost containment may be a nearly impossible task for government influence, especially since the growth in HE is likely a result of new health care technology at higher prices rather than increasing prices of existing technology (Weisbrod, 1991).

The introduction of international longitudinal data from the OECD has made it possible to analyze the relationship between HE and GDP over time while simultaneously controlling for individual-country differences. This has led to new inferences about HE in more recent literature. Hitiris and Posnett (1992), for example, use a panel of time-series data from 20 OECD countries covering the years 1960 to 1987 and find support for a *unitary* income elasticity of HE. Their analysis, however, ignores one of the major issues of time series estimation: the potential for nonstationarity and spurious regression. That is, if HE or GDP is nonstationary then a regression of HE on GDP is "spurious" and the results from estimating such a model are invalid. As such, the estimated income elasticity of HE and the finding of a significant and positive relationship between HE and GDP may be meaningless results. This is an important issue from a public policy perspective since the government's ability to manage expenditures on health care hinges, at least in part, on obtaining an accurate estimate of the determinants of HE. Accurate knowledge of the time series properties of HE and GDP is a prerequisite to empirical testing. Only if HE and GDP are each stationary (or there exists a stationary linear combination implying they are cointegrated) can researchers be confident that a regression of HE on GDP is meaningful.[1]

Following the above line of reasoning, researchers have recently begun to examine whether HE and GDP are stationary. The bulk of this research suggests that these variables are nonstationary. Hansen and King (1996), for example, use augmented Dickey-Fuller tests to perform univariate unit root tests for several time series, including HE and GDP. Their results indicate little support for stationarity. Blomqvist and Carter (1997) perform univariate unit root tests using a nonparametric modification of the Dickey-Fuller test as suggested by Phillips and Perron (1988). In general, they find that both series are nonstationary. In addition, Blomqvist and Carter employ the panel unit root test of Levin and Lin (1993) and are again unable to reject the null of non-stationarity in HE and GDP. Taken together, these findings indicate that previous regression results that assume stationarity may be invalid.

More recently, researchers have begun to favor panel-based unit root tests over univariate tests. The use of panel data when testing for a unit root allows for more powerful tests due to increased sample size and inclusion of heterogeneous cross-country information not available in single-country tests. McCoskey and Selden (1998), for example, use the panel data unit root test of Im, Pesaran, and Shin (1997, hereafter IPS) to test for stationarity in HE and GDP. They find

[1] In the case where both HE and GDP are nonstationary, one may consider a possible stationary linear combination of these variables within the context of a cointegrating regression. See Gerdtham and Lothgren (2000) for a discussion of this issue. In this paper, we focus on the stationarity of HE. However, if either HE or GDP is stationary, then the issue of cointegration becomes a moot point.

evidence in favor of stationarity for HE and GDP. In their comment, Hansen and King (1998) question the findings of McCoskey and Selden due to their omitting a time trend in their empirical tests. As Hansen and King note, excluding a time trend imposes the restriction that HE cannot be trend-stationary under the alternative, which may lead to bias results in unit root tests. This outcome may be especially important given the positive trend of HE over time. Gerdtham and Lothgren (2000) challenged the findings of McCoskey and Selden by using a battery of tests, which include a time trend, and conclude that both HE and GDP are non-stationary. The inconsistency of previous results concerning the stationarity of HE and GDP, despite the use of similar data, may be due to differences in specification of the respective unit root tests; in particular, the dissimilar results could be due to the inclusion or exclusion of a time trend.[2]

One important potential drawback of previous research is the failure to allow for structural breaks in unit root tests. Both McCoskey and Selden (1998) and Gerdtham and Lothgren (2000) recognize this potential drawback, but both papers note that a suitable panel unit root test that allows for heterogeneous breaks was not available at the time of publication. Further analysis of the stationarity of HE and GDP allowing for structural breaks is required to properly investigate the relationship between the two variables. In this paper, we examine the issue of stationarity in HE, while leaving the subject of stationarity in GDP and the cointegration of HE and GDP for future research. We make use of Lagrange multiplier (LM) unit root tests in the univariate and panel setting. To fill in a gap in previous research, our unit root tests allow for the possibility of structural breaks.

---

[2] The papers described above use slightly different data sets in their unit root tests, but the effect of this on unit root tests is expected to be minor. For example, Hansen and King (1996) and McCoskey and Selden (1998) use data from 20 OECD countries for the period 1960 to 1987. Blomqvist and Carter (1997) use data from 20 OECD countries for the period 1960 to 1991. Gerdtham and Lothgren (2000) use data from 21 OECD countries for the period 1960 to 1997.

2. Structural breaks in unit root tests

At the present time, no published studies exist that address the stationarity of HE in the presence of structural breaks. Following Perron (1989), it is well known that a structural break can be mistaken for non-stationarity in time series. Perron shows that failure to account for an existing structural break will lead to a substantial loss of power to reject the null hypothesis in unit root tests. Therefore, it is quite possible that the inability to reject the unit root null hypothesis in previous HE research has been due, at least partially, to the failure to allow for structural breaks.

The task of allowing for structural breaks in the framework of existing Dickey-Fuller-type panel unit root tests, such as those proposed by Levin and Lin (1993) and IPS, is difficult to implement. This is due to the fact that the distribution of these tests will critically depend on nuisance parameters indicating the location of breaks (see, e.g., Im and Lee, 2001). Given this situation, it would be extremely difficult, if not impossible, to control for the numerous possible combinations of heterogeneous structural breaks that can occur in each country using Dickey-Fuller-type panel unit root tests. To provide a remedy, we apply the panel unit root test recently developed by Im and Lee (2001). Their panel test is derived from the univariate one-break LM unit root test developed by Amsler and Lee (1995), which is an extension of the no-break LM unit root test initially developed by Schmidt and Phillips (1992).

Although it is extremely difficult (computationally? in terms of estimation?) to allow for heterogeneous structural breaks in a Dickey-Fuller-type panel unit root test, the distribution of the LM panel unit root test does not depend on such break point nuisance parameters. This advantage of the LM panel test follows from invariance results, which are similar to those in the univariate test. Amsler and Lee (1995) show that the limiting distribution of the univariate LM unit root test with structural breaks is the same as that of the test without breaks. In other words, the distribution of the LM unit root test is unaffected by breaks. Im and Lee (2001) extend the LM unit root test with breaks to the panel framework and show that this same "invariance" property carries to the panel test. As such, the same critical values that apply to the LM panel unit root test *without* breaks are valid for the LM panel unit root test *with* breaks, regardless of the location and number of breaks in each cross-section of the panel. In addition, for the benchmark case where no breaks are allowed, the LM panel unit root test is shown to be more powerful than the IPS Dickey-Fuller-type panel test. By allowing for structural breaks in the panel test, as in the univariate case, a significant gain in power should result.

3. The LM panel unit root test with breaks

3.1. The panel data model with breaks

The question of interest in our analysis is whether or not a panel data series on health expenditures (HE_{it}) is stationary. In order to answer this question, we first consider the panel data model:

$$HE_{it} = \delta_i' Z_{it} + e_{it}, \text{ where: } e_{it} = \beta_i e_{i,t-1} + \varepsilon_{it} , \qquad (1)$$

where "i" indexes cross-section units (i.e., countries); "t" indexes time periods (i.e., years); Z_{it} is a vector of exogenous variables that describe how HE_{it} is generated; δ_i is the corresponding parameter vector for these exogenous variables; and e_{it} is the error term of the process. Within e_{it}, the parameter β_i tests the null hypothesis of a unit root, and ε_{it} is a zero-mean error term that allows for heterogeneous variance structure across cross-section units but assumes no cross-correlations.

Now consider the case where the underlying model for HE_{it} allows for one structural break in level.[3] In this case, the vector of exogenous variables, Z_{it}, takes the following form: [1 t D_{it}]′ where D_{it} is a dummy variable that marks the time when a structural break occurs. That is, if we define the location at which the break occurs for country "i" as "T_{Bi}", then the dummy variable takes the form: $D_{it} = 1$ if $t > T_{Bi}$, and 0 otherwise.[4] Note that the model has the advantage of allowing for heterogeneous structural breaks that vary by country as distinguishable from common shocks that uniformly apply to all countries. Re-writing equation (1) to incorporate the above gives:

$$HE_{it} = \delta_{1i} + \delta_{2i}\, t + \delta_{3i}\, D_{it} + e_{it}\,, \text{ where: } e_{it} = \beta_i\, e_{i,\,t-1} + \varepsilon_{it}\,. \qquad (1')$$

In this framework, the unit root null hypothesis tests whether or not $\beta_i = 1$. As such, the critical values of this test will be invariant to the estimated parameters in the vector δ_i. An additional feature that enhances the appeal of this model, and similar to the IPS panel test, is that β_i allows for heterogeneous measures of persistence.

An important feature of the model is that it allows for a break under both the null and alternative hypotheses. To see this, suppose that $\beta_i = 1$ for all values of "i" so that the null model is given by:

$$HE_{it} = \delta_{1i} + \delta_{2i}\, t + \delta_{3i}\, D_{it} + e_{it}\,,\ e_{it} = e_{i,\,t-1} + \varepsilon_{it}\,. \qquad (2)$$

Solving this equation for e_{it} and applying the quasi-differencing transformation, we obtain:

$$\Delta e_{it} = HE_{it} - HE_{i,\,t-1} - \delta_{2i} - \delta_{3i}\,[D_{it} - D_{i,\,t-1}]\,. \qquad (3)$$

[3] We consider only the “crash” model that allows for a (permanent) change in level (see Im and Lee, 2001).

[4] Of course, this framework can easily be extended to allow for two structural breaks by defining Z_{it} as [1 t D_{1it} D_{2it}]′, where D_{1it} is a dummy variable that captures the first break and D_{2it} is a dummy variable that captures the second break. Note that “1” and “t” allow for country-specific intercepts and time trend terms, respectively, in the panel test.

Suppose we define $[D_{it} - D_{i,\,t-1}] \equiv B_{it}$ (so that B_{it} takes on a value of 1 for the time period $T_{Bi} + 1$, and 0 at all other times) and let $\Delta e_{it} \equiv \nu_{it}$. In this case, equation (3) can be re-written so that the model under the null becomes:

$$HE_{it} = \delta_{2i} + \delta_{3i} B_{it} + HE_{i,\,t-1} + \nu_{it}\,. \tag{4}$$

Allowing for a break under the unit root null is important since omitting this term may lead to invalid results (Perron, 1989).

3.2. The LM panel test statistic

The first step in computing the LM panel unit root test statistic is to compute the conventional univariate LM unit root test statistics for each country. To that end, we begin with the following regression:

$$\Delta HE_{it} = \delta_i' \,\Delta Z_{it} + \phi_i \,\tilde{S}_{i,t-1} + u_{it}\,, \tag{5}$$

where ΔHE_{it} and ΔZ_{it} are the first-differenced values of HE_{it} and Z_{it}, respectively; $\tilde{S}_{i,t-1}$ is the detrended value of $HE_{i,t-1}$ (following Schmidt and Phillips, 1992); and u_{it} is the model's stochastic disturbance term. The presence of a unit root in HE_{it} (for country "i") implies that $\phi_i = 0$. Then, it follows that the univariate LM test statistic can be computed using the *t*-test that tests the null and alternative hypotheses: H_0: $\phi_i = 0$ (unit root and non-stationary) *versus* H_1: $\phi_i < 0$ (no unit root and stationary) for each country. We denote the *t*-test for each country as $\tilde{\tau}_i$.

To endogenously determine the location of breaks, we employ the univariate minimum LM unit root tests as suggested in Lee and Strazicich (1999a and 1999b). They propose a grid search procedure that determines the break point location for each country "i" ("T_{Bi}") where $\tilde{\tau}_i$ is at a minimum (i.e., the most negative) and, hence, their test is referred to as a "minimum LM test."[5] The minimized value of $\tilde{\tau}_i$ is denoted as "LM_i^τ" and is the break point(s) least favorable to the unit root null hypothesis. Since critical values are invariant to the location of the break(s), we do not need to simulate new critical values for different combinations of the break points. Furthermore, unlike the competing Dickey-Fuller-type endogenous break tests, the minimum LM unit root test is not subject to spurious rejections in the presence of a break under the null (see, e.g., Nunes, Newbold, and Kuan, 1997, and Lee and Strazicich, 2001).

To correct for the presence of autocorrelated disturbances in the model, we include augmentation terms $\Delta\tilde{S}_{i,t-j}$ in (5), where j = {1, 2, ..., k_i} and "k_i" is selected so as to eliminate autocorrelation in the errors. The optimal value of k_i is determined for each country at each combination of break point(s) using the general-to-specific methodology suggested by Perron (1989). In this regard, the optimal lag length "k_i" and break location(s) "T_{Bi}" are jointly and endogenously determined by the data.

Following Im and Lee (2001), the null and alternative hypotheses, respectively, in the panel test are given by: H_0: $\phi_i = 0$, i.e., unit root for all countries *versus* H_1: $\phi_i < 0$, i.e., one or more countries reject the unit root. The LM panel test statistic is computed by averaging the univariate LM unit root *t*-test statistics estimated for each country, LM_i^τ, as follows:

$$\overline{LM}_{NT} = \frac{1}{N}\sum_{i=1}^{N} LM_i^\tau \ . \qquad (6)$$

A standardized panel LM unit root test statistic is then constructed by letting $E[L_T]$ and $V(L_T)$ denote the expected value and variance of LM_i^τ, respectively, under H_0, and computing:

[5] The grid search is performed over the interval [0.1T, 0.9t] to eliminate end points. In addition, this methodology can easily be extended to the case allowing for two breaks, T_{1Bi} and T_{2Bi}, where the grid search is repeated at each possible combination of two break points to locate the minimum test statistic.

$$\Gamma_{LM} = \frac{\sqrt{N}\left[\overline{LM}_{NT} - E[L_t]\right]}{\sqrt{V(L_T)}} . \qquad (7)$$

Numerical values for $E[L_T]$ and $V(L_T)$ are provided in Table 1 of Im and Lee (2001). In the presence of heterogeneous autocorrelated errors, we use the weighted averages of $E[L_T(k_i)]$ and $V(L_T(k_i))$ to be consistent with the country-specific optimal values of k_i.

The panel LM unit root test has several attractive features. First, the distribution of the test statistic depends on N and T but does not depend on any other parameters under the null hypothesis. Im and Lee derive the asymptotic properties of their panel test statistic and, similar to the IPS panel unit root test, the asymptotic distribution of the LM panel unit root test is standard normal. Most importantly, the distribution of the panel LM test is unaffected by the presence of break(s). This so called "invariance result" holds for any finite number of breaks and regardless of whether or not the break locations are correctly determined. This is convenient, since it makes it unnecessary to simulate new critical values for $E[L_T]$ and $V(L_T)$ depending on the number and location of the breaks identified in each country.

4. Empirical results

4.1. Data

The source of data for our analysis is the OECD Health Data File (1998). This data file contains information on all OECD countries for the period 1960 to 1997. Our analysis makes use of logged data on health expenditures from 20 OECD countries, consistent with those in

Gerdtham and Lothgren (2000).[6] The health expenditure data is measured in per capita terms in constant 1990 dollars.

4.2. Test results

To implement our testing, we must first determine the number of structural breaks in each country. We begin by employing the univariate two-break minimum LM unit root test of Lee and Strazicich (1999b). First, the optimal value of k_i is determined, as in Section 3.2, for each possible combination of two breaks. Following this, the location of the breaks is determined to be where the *t*-statistic testing the unit root null is minimized. After endogenously determining the location of two breaks, the *t*-statistic of each estimated break coefficient is examined for significance at the 10% level in an asymptotic normal distribution (i.e., absolute value greater than 1.645). If less than two breaks is significant, the procedure is repeated using the one-break minimum LM unit root test of Lee and Strazicich (1999a). If no break is significant, then the no-break LM unit root test of Schmidt and Phillips (1992) is employed. In this manner, the location of breaks, the number of breaks, and the number of lagged augmented terms (k_i) are jointly determined for each country.[7] One final consideration is to allow for the possibility of country-specific and time-specific fixed effects in the panel test. Country-specific fixed effects are included, since the *t*-statistics used to calculate the panel test statistic are estimated with heterogeneous intercept terms in the univariate tests. To allow for common time-specific effects,

[6] To maintain a balanced panel, we omit New Zealand and Portugal, which have missing observations, and we include Japan.

[7] One might suggest extending the above procedure to allow for more than two structural breaks in each cross-section. Although the invariance results for the minimum LM unit root test should hold for any finite number of breaks, the finite sample properties with more than two breaks have not been examined. In addition, the computational burden would significantly increase when searching all combinations of more than two break points and identifying the number of lagged augmented terms. However, allowing for more than two breaks may not be a concern here, as we find strong rejections of the unit root null with one break.

we "demean" the data by subtracting the average value of HE_t from each HE_i in each year t.[8] We are now ready to compute the LM panel test statistic described in equation (7).

Results of the estimation procedure are shown in Tables 1 and 2. To examine the robustness of our results, testing is performed with and without demeaning for common time effects. To allow for the possibility of a trend stationary alternative, all unit root tests include both an intercept and time trend term. The estimated univariate LM unit root test statistics are shown in the second column of each Table. The optimal number of breaks are shown in the third column, the optimal number of lagged differenced terms that correct for serial correlation are shown in the fourth column, and the location of the breaks are shown in the last column.

Table 1 displays results without allowing for common time effects. Examining the univariate test results, we see only two rejections of the unit root null at the 10% level of significance. One structural break is identified in 12 out of 20 countries, with most breaks occurring in an 11-year period from 1972 to 1982. The only exceptions are Norway (1986) and Spain (1992). The preponderance of estimated break points during the 1970s may reflect technological advances during this period that lead to large increases in the cost of medical care. (Do we need to discuss the implications of these breaks?) We see a marked contrast in the panel test. Contrary to the univariate tests, the panel test rejects the unit root null at less than the 1% level of significance (test statistic is –2.334). (Is this due to the increased power of the test? Should we drop Table 1 and simply discuss Table 2?)

Table 2 displays the results with adjustment for common time-specific effects. In the univariate test results, we see a significant increase in the number of rejections of the unit root

[8] It should be noted that including time-specific effects has a distinct meaning in regards to structural breaks. In essence, the inclusion of time-specific effects in the panel setting is equivalent to allowing for a structural break in each year that is common (homogeneous) to all OECD countries. After including

null as compared to Table 1. More than half of the countries (11 out of 20) reject the unit root null at less than 10% as compared to two countries in Table 1. These results suggest the importance of allowing for time-specific effects. The number of (heterogeneous) breaks identified is (again) one or zero in each country, with 14 countries having one structural break. The estimated break points for most of the sample now occur during the period from 1970 to 1981, roughly the same period as in Table 1. The exceptions are Norway (1986), Greece (1986), and Finland (1992). (Do we need to discuss the fact that some of the estimated break points change a lot from Table 1 to Table 2? For instance, Finland changes from 1974 to 1992.) Moving to the panel test, compared to Table 1 we see an even stronger rejection of the unit root null hypothesis at much less than the 1% level of significance (test statistic is –8.695).

5. Conclusion

Whether health care expenditures are stationary or nonstationary can have critical implications for researchers and policy makers desiring to model, explain, and impact this important economic sector. Previous research has, in general, concluded that health care expenditures follow a nonstationary process with a unit root. The policy implication from such a result is that government intervention to manage health expenditures may be thwarted by the random nature of aggregate health expenditure. In terms of econometric estimation, if health care expenditures are nonstationary, then a regression of health expenditures on possible determinants will essentially be meaningless. A potential drawback to previous unit root tests is that they do not allow for the possibility of structural breaks in unit root tests. As initially noted by Perron

common time effects, the remaining structural breaks are those specific (heterogeneous) to individual countries.

(1989), failure to allow for an existing structural break makes rejection of the unit root null less likely. We re-examine the time series properties of health care expenditures in twenty OECD countries for 1960 to 1997 by using a LM panel unit root test that allows for heterogeneous structural breaks. The combination of panel data with structural breaks significantly increases the power of the test to reject a false null.

Our panel results strongly reject the unit root null hypothesis at less than the 1% level of significance and imply that health expenditures follow a stationary process with, in general, one (permanent) change in level. In addition, structural breaks are identified most often during the 1970s. The major implication of our findings is that modeling health care expenditures in a regression framework can give meaningful results. The findings in this paper are in sharp contrast to those of Hansen and King (1996), Blomqvist and Carter (1997), and Gerdtham and Lothgren (2000) who do not consider structural breaks in their unit root tests. In addition, whereas McCoskey and Selden (1998) reject the unit root null hypothesis using panel data, their model may not be correctly specified as they omit a time trend term from their unit root tests and do not consider structural breaks. Contrary to this, our panel test results are robust to inclusion of a time trend.

References

Amsler, C., Lee, J., 1995. An LM test for a unit-root in the presence of a structural change. Econometric Theory 11, 359–368.

Blomqvist, A. G., Carter, R. A. L., 1997. Is health care really a luxury? Journal of Health Economics 16, 207–229.

Culyer, A. J., 1990. Cost containment in Europe, in Health Care Systems in Transition, OECD, Paris, France, 29–40.

Gerdtham, U.-G., Jonsson, B., 1991. Conversion factor instability in international comparisons of health care expenditure. Journal of Health Economics 10, 227–234.

Gerdtham, U.-G., Lothgren, M., 2000. On stationarity and cointegration of international health expenditure and GDP. Journal of Health Economics 19, 461–475.

Hansen, P., King, A., 1996. The determinants of health care expenditure: a cointegration approach. Journal of Health Economics 15, 127–137.

Hitiris, T., Posnett, J., 1992. The determinants and effects of health expenditure in developed countries. Journal of Health Economics 11, 173–181.

Im, K. S., Lee, J., 2001. LM unit root test with panel data: a test robust to structural changes. Mimeo. Department of Economics, University of Central Florida.

Im, K. S., Pesaran, M. H., Shin, Y., 1997. Testing for unit roots in heterogeneous panels. Mimeo. Department of Applied Economics, University of Cambridge.

Lee. J., Strazicich, M., 1999a. Minimum LM unit root test. Mimeo. Department of Economics, University of Central Florida.

Lee. J., Strazicich, M., 1999b. Minimum LM unit root test with two structural breaks. Mimeo. Department of Economics, University of Central Florida.

Lee. J., Strazicich, M., 2001. Break point estimation with minimum unit root tests and spurious rejections. Mimeo. Department of Economics, University of Central Florida.

Levin, A. Lin, C-F., 1993. Unit root tests in panel data: new results. Mimeo. Department of Economics, University of California at San Diego.

Kwiatkowsi, D., Phillips, P. C. B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity *versus* the alternative of a unit root. Journal of Econometrics 54, 159–178.

McCoskey, S. K., Selden, T. M., 1998. Health care expenditures and GDP: panel data unit root test results. Journal of Health Economics 17, 369–376.

Newhouse, J. P., 1977. Medical care expenditures: a cross-national survey. Journal of Human Resources 12, 115–125.

Ng, S., Perron, P., 1995. Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. Journal of the American Statistical Association 90, 269–281.

Nunes, L., Newbold, P., Kuan, C, 1997. Testing for unit roots with breaks: evidence on the great crash and the unit root hypothesis reconsidered. Oxford Bulletin of Economics and Statistics 59, 435–448.

OECD Health Data File (1998). (Need to complete this reference.)

Parkin, D., McGuire, A., Yule, B., 1987. Aggregate health care expenditures and national incomes: is health care a luxury good? Journal of Health Economics 6, 109–127.

Perron, P., 1989. The great crash, the oil price shock, and the unit root hypothesis. Econometrica 57, 1361–1401.

Phillips, P. C. B., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75, 407–436.

Schmidt, P., P. C. B. Phillips, 1992. LM tests for a unit root in the presence of deterministic trends. Oxford Bulletin of Economics and Statistics 54, 257–287.

Weisbrod, B. A. 1991. The health care quadrilemma: an essay on technological change, insurance, quality of care, and cost containment. Journal of Economic Literature 29, 523–552.

Table 1
LM unit root tests of health care expenditures 1960-1997 with optimal number of breaks and without adjustment for time fixed effects

Country	Univariate LM unit root test statistic	Optimal # of breaks	Optimal lag length k	Break location
Australia	–2.052	0	3	-
Austria	–1.373	0	1	-
Belgium	–2.324	1	7	1981
Canada	–1.073	1	1	1974
Denmark	–1.796	1	6	1979
Finland	–1.977	1	1	1974
France	0.299	0	6	-
Germany	–3.207	1	8	1974
Greece	–2.744	0	8	-
Iceland	–1.224	0	4	-
Ireland	–3.741**	1	5	1975
Italy	–2.558	1	2	1978
Japan	–2.161	0	4	-
Netherlands	–2.466	0	3	-
Norway	–3.775**	1	8	1986
Spain	–1.346	1	5	1992
Sweden	–2.287	1	7	1972
Switzerland	–1.817	0	4	-
United Kingdom	–3.044	1	7	1982
United States	–3.110	1	2	1975
LM panel test statistic	–2.334***			

In no case was more than one break significant.
All regressions include an intercept and time trend.
Critical values for the LM unit root test with no break are –3.63, –3.06, and –2.77 at the 1%, 5%, and 10% significance levels, respectively, and come from Schmidt and Phillips (1992).
Critical values for the minimum LM test with one break are –4.239, –3.566, and –3.211 at the 1%, 5%, and 10% levels, respectively, and come from Lee and Strazicich (1999a).
Critical values for the LM panel unit root test (with or without breaks) are distributed asymptotic standard normal and are –2.326, –1.645, and –1.282 at the 1%, 5%, and 10% levels, respectively.
*** Significant at the 1% level; ** significant at the 5% level; * significant at the 10% level.

Table 2
LM unit root tests of health care expenditures 1960-1997 with optimal number of breaks and with adjustment for time fixed effects

Country	Univariate LM unit root test statistic	Optimal # of breaks	Optimal lag length k	Break location
Australia	–1.656	0	0	-
Austria	–3.590**	1	4	1970
Belgium	–2.672	1	6	1970
Canada	–3.964**	1	8	1981
Denmark	–2.426	0	0	-
Finland	–4.334***	1	8	1992
France	–5.045***	1	6	1973
Germany	–3.778***	1	6	1980
Greece	–3.359*	1	8	1986
Iceland	–2.482	1	4	1974
Ireland	–4.647***	1	5	1975
Italy	–3.845**	1	2	1979
Japan	1.276	0	6	-
Netherlands	–3.418**	0	8	-
Norway	–4.269***	1	8	1986
Spain	–1.139	0	5	-
Sweden	–2.152	1	5	1976
Switzerland	–5.823***	1	8	1970
United Kingdom	–1.701	1	7	1979
United States	–2.375	0	5	-
LM panel test statistic	–8.695***			

In no case was more than one break significant.
All regressions include an intercept and time trend.
Critical values for the LM unit root test with no break are –3.63, –3.06, and –2.77 at the 1%, 5%, and 10% significance levels, respectively, and come from Schmidt and Phillips (1992).
Critical values for the test with one break are –4.239, –3.566, and –3.211 at the 1%, 5%, and 10% levels, respectively, and come from Lee and Strazicich (1999a).
Critical values for the LM panel unit root test (with or without breaks) are distributed asymptotic standard normal and are –2.326, –1.645, and –1.282 at the 1%, 5%, and 10% levels, respectively.
*** Significant at the 1% level; ** significant at the 5% level; * significant at the 10% level.

UNIVERSITY OF NORTH TEXAS

Margie A. Tieslau
Associate Professor
Department of Economics
phone: 940.565.3442

P.O. Box 311457
Denton, TX 76203-1457
tieslau@unt.edu
fax: 940.565.4426

Professor J.P. Newhouse
Division of Health Policy Research and Education
Harvard University
180 Longwood Avenue
Boston, MA 02115-5899

September 20, 2001

Dear Professor Newhouse:

Please consider the enclosed manuscript, "Stationarity of health expenditures: a re-examination using panel unit root tests with heterogeneous structural breaks," co-authored with Todd Jewell, Junsoo Lee, and Mark Strazicich, for publication in the *Journal of Health Economics*. I have enclosed four copies of the manuscript, as per the instructions in the most recent issue of the journal. In addition, since our paper draws heavily on the paper by Im and Lee ("LM Unit Root Test with Panel Data: A Test Robust to Structural Changes," 2001) and this paper is not yet in print, I have included a copy of this paper in order to assist the referees. For the convenience of the referee(s), all computer programs can be found at: www.junsoo.lee. Working papers by Lee and Strazicich can be found at: www.junsoo.lee2 .

If you require any additional information, please do not hesitate to contact me. Thank you for your consideration. I look forward to hearing from you on this matter in the near future.

Sincerely,

Margie A. Tieslau
Associate Professor

Enclosures