Available online at www.sciencedirect.com
ScienceDirect
Journal of Health Economics 26 (2007) 682–698
JOURNAL OF HEALTH ECONOMICS
www.elsevier.com/locate/econbase

Explaining China's regional health expenditures using LM-type unit root tests

Win Lin Chou *

Department of Economics, Chinese University of Hong Kong, Shatin, Hong Kong

Received 12 October 2004; received in revised form 30 November 2006; accepted 6 December 2006
Available online 12 January 2007

Abstract

This paper investigates the relationship between health care expenditure, income, and other factors that are not related to income for China with pooled cross-section and time series data. To study the stationarity property of these variables, we use panel Lagrange Multiplier (LM) unit root tests that allow for structural changes. To perform the LM unit root tests, we employ finite-sample critical values derived through the bootstrap method, instead of relying on the critical values from the asymptotic normal distribution. An important finding based on the estimated panel cointegrated regressions is that the government budget deficits have a significant long-run impact on China's health care expenditure. This provides supportive evidence on the differences between rich and poor areas in China's health care financing policy, and the substantial disparities in health service coverage in China.

JEL classification: C12; C23; I10

Keywords: Lagrange multiplier test; Panel unit root tests; Bootstrapped critical values; Panel cointegration tests

1. Introduction

The growth of health care expenditure and of its share in GDP is a subject that has been constantly discussed among academics, administrators, and politicians in many countries. One approach to this issue used cross-country comparisons of health care expenditure.[1] To date, all the studies on international comparisons of health expenditure used the OECD data. A common result

* Tel.: +852 2609 8001; fax: +852 2603 5805.
E-mail address: winlinchou@cuhk.edu.hk.

[1] The main reason for examining health expenditure rather than the quantity of care that is demanded is that there are no generally accepted quantity measures which could be used to combine heterogeneous health care services into a single quantity index.

doi:10.1016/j.jhealeco.2006.12.002

of these studies (see for example, Newhouse, 1977; Leu, 1986; Gerdtham et al., 1992; Hitiris and Posnett, 1992) is that income (proxied by GDP) is the most significant factor explaining variations in health care expenditure, while the impact of factors such as age structure, relative prices and amount of health care funded by the public sector is found to be small. These studies have enhanced our understanding of the relations between economic and non-economic variables with respect to health care expenditure.

Perhaps due to the lack of data, research on China's health expenditure and its determinants is quite scarce. Previous research that dealt with the health sector in China tended to employ the survey research approach with simple descriptive statistics (see for example, World Bank, 1984, 1992, 1997), and econometric methods were rarely used. With the increased availability of time-series and cross-section data for health care expenditure in China, it has now become possible to investigate the relationship between health care expenditure and its main determinants in China using econometric methods.

The main objective of this paper is to investigate the relationship between health care expenditure, income, and other variables that are not related to income by using pooled cross-section and time series data for China. This study differs from the research that uses the pooled OECD data in two aspects. First, the paper uses the Lagrange Multiplier (LM) type panel unit root tests to study the stationarity property of the time series variables, while much of the previous research had employed the Dickey-Fuller (DF)-type tests. The use of the LM-type tests is motivated by the findings of Amsler and Lee (1995) and Im et al. (2005). In conducting the panel unit root tests, it is necessary to make available distributions of the individual test statistics, their means, and variances. To this end, a variety of simulations have been implemented to establish the asymptotic distributions of different unit root test statistics. Using the Monte Carlo method, Amsler and Lee (1995) demonstrated that allowing or not allowing for a structural break in the intercept does not affect the asymptotic distribution of the LM unit root test statistics under the null hypothesis. This finding suggests that the LM tests are not affected by misplacement of structural break or by allowing for a break when there is no break or vice versa. This invariant result is in sharp contrast with the DF-type tests such as that of Perron's (1989) in which the structural break affects the asymptotic distributions of the test statistics. Im et al. (2005) further showed that this invariance property is carried over to the panel LM tests. They found that when implementing the panel LM unit root tests, the same means and variances of the asymptotic distributions of individual LM test statistics can be employed whether or not the structural break is allowed for in the model. Another attractive feature of the LM-type test is that unlike the previous studies on endogenous break tests, such as Zivot and Andrews (1992), which assumed no breaks under the null, the LM-type test allows for breaks under both the null and alternative hypotheses. The advantage of the latter is that rejection of the unit root null is unambiguously an evidence of trend stationarity, whereas assuming no break under the null in the endogenous break tests can imply two possibilities, that is, trend stationarity with breaks and a unit root with breaks, thus suggesting test results are ambiguous.

Second, unlike Jewell et al. (2003) who conducted the panel LM unit root tests using the OECD data with critical values that are taken from the asymptotic normal distribution, we derive finite-sample critical values by the bootstrap method, as asymptotic critical values often differ from their finite-sample counterparts in practice. Even if the distribution of the test statistic adopted is asymptotically invariant to break points, such invariance may not hold in finite samples. The motivation that drives the use of bootstrap method is that bootstrap has been found to be effective in correcting size distortions of panel tests in the presence of cross-sectional dependency. Panel unit root tests have been constructed under the assumption of independence across members of the panel; however, the possibility of cross-sectional dependency cannot be excluded in the panel. Consequences of ignoring cross-sectional dependency such as severe size distortions have been evidenced in Maddala and Wu (1999). Many different approaches have recently been developed to deal with the issue of cross-sectional dependency in panel unit root tests. Chang (2004) is a noticeable example. Simulations in Chang (2004) show that the bootstrap panel unit root tests perform well in finite samples relative to the IPS test statistic of Im et al. (2003).

Justification for employing the pooled cross-section (i.e. by province) and time series data of China is that the substantial disparities are observed in health service coverage and people's health status between coastal and inland provinces, and between urban and rural areas.[2] The aggregate time series statistics reported in the World Development Indicators (WDI) are unable to reflect the observed regional disparities.

When considering the main determinants for China, we basically follow Gerdtham et al. (1992), except the number of practicing physicians and the relative price index, due to the unavailability of data. In addition, we consider a new variable on government budget deficits. The importance of this variable has been given in Gerdtham and Jonsson (2000), who argued that this "new" variable is likely to be a strong constraint on public health expenditure. As government budget deficits are likely to become a constraint for many regions in China, we will investigate their impact on China's health expenditures.

The paper is organized as follows: In Section 2, we discuss the model that is used to derive the panel LM unit root test statistics allowing structural breaks under the null hypothesis. Section 3 reports the results of panel unit root and panel cointegration tests, and the estimated parameters of panel cointegrated regressions. Finally, Section 4 gives the conclusion.

2. Panel LM unit root test statistics with breaks and data

2.1. Panel LM unit root test statistics

The panel LM unit root test statistics that allow for structural breaks are based on the univariate LM unit root test statistics of Schmidt and Phillips (1992, SP hereafter). Consider the following data generating process (DGP) for y_t:

$$y_t = \delta_1 + Z_t\delta + x_t, \tag{1}$$

$$x_t - \alpha x_{t-1} = \varepsilon_t. \tag{2}$$

where Z_t contains exogenous variables, and $\varepsilon_t \sim iidN(0, \sigma^2)$. We are interested in deriving the LM statistic to test the null hypothesis of the unit roots $\alpha = 1$. When the DGP for y_t allows for a structural break in the level, Z_t takes the form

$$Z_t = (t, DU_t) \tag{3}$$

where $DU_t = 1$ for $t \geq TB + 1$, and 0 otherwise, and where TB stands for the time period when the break occurs. We denote this model as Model 1. When a break is present in both intercept and

[2] The Chinese statistics indicate that starting from the early 1990s richer provinces in China such as Beijing, Shanghai and Shangdong, which have low infant mortality rates received greater budgetary resources for health, whereas poor provinces like Guizhou, Henan and Yunnan where infant mortality rates are higher than the national average, received lower health budgetary resources for health.

slope, Z_t is specified as

$$Z_t = (DU_t, t, DT_t) \tag{4}$$

where $DT_t = t$ for $t \geq TB + 1$, and 0 otherwise. This model is denoted as Model 2. Model 3 allows for a break in the slope and Z_t is specified as

$$Z_t = (t, DT_t). \tag{5}$$

The LM unit root test statistic is estimated by regression according to the LM principal as follows. Under the null hypothesis $\alpha = 1$, the maximum likelihood estimators of δ in (1), which are denoted by $\tilde{\delta}$, are obtained by the estimation of the following equation by the ordinary least squares (OLS) method:

$$\Delta y_t = \Delta Z_t \delta + \mathrm{e}_t, \tag{6}$$

where ΔZ_t is the first difference of the regressors Z_t. We write the OLS residuals $\tilde{S}_t$ as

$$\tilde{S}_t = y_t - \tilde{\delta}_1^* - Z_t \tilde{\delta}, \tag{7}$$

where $\tilde{\delta}_1^* = y_1 - Z_1 \tilde{\delta}$, in which y_1 and Z_1 denoting the first observations of y_t and Z_t, respectively.[3]

The LM-type unit root test statistic can be estimated from the test regression:

$$\Delta \tilde{S}_t = \Delta Z_t \delta + \phi \tilde{S}_{t-1} + \mathrm{e}_t \tag{8}$$

To allow for the existence of autocorrelation in errors, we can estimate an augmented regression that is specified as:

$$\Delta \tilde{S}_t = \Delta Z_t \delta + \phi \tilde{S}_{t-1} + \sum\nolimits_1^k c_j \Delta \tilde{S}_{t-j} + \mathrm{e}_t, \tag{9}$$

where k can be chosen based on the alternative selection procedures, as in the case of the ADF-type tests. The LM test statistic is given by the t-statistic for the testing of the null of $\phi = 0$, and is defined as $\tau_\phi^{TB}(m)$, where m denotes Model 1, 2 and 3 that are described in (3)–(5), respectively, and TB is the break date.

2.2. *Break selection methods*

When a break is allowed under the null hypothesis, the LM unit root tests require the estimation of the break date (TB). Several break date selection methods have been proposed in the literature. We consider two endogenous break selection methods. In the first case, we choose the break date over a range of possible break points when its corresponding t-statistic testing the null of a unit root is minimized in Eq. (9).[4] Let the break date that is selected in this way be denoted as $\widehat{TB}$, and its corresponding t-statistic is the minimum LM test statistic of Lee and Strazicich (2003, 2004), and is denoted as $t_\phi^{\widehat{TB}}(m)$ or simply *MinLM*. As an alternative, we consider the break date selection methods proposed by Nunes (2004), who suggests the selection of the break date by the maximization of the absolute value of the t-statistic for ΔDU_t in regression (9) for Model 1, and by the maximization of the F-statistic for ΔDU_t and ΔDT_t in (9) for Model 2. The estimated break

3 See Schmidt and Phillips (1992) for details.

4 The number of k augmented terms is determined jointly with the selected break date. The maximum value of k is set at 5 due to the relatively short time series data for China.

date is denoted as $\widehat{tb}$ and the estimated t-statistics are written as $t_{\phi}^{\widehat{tb}}(m)$ or $Max|tb|$. Both Lee and Strazicich (2003, 2004) and Nunes (2004) find that the LM-type test statistics are asymptotically invariant to a change in intercept.

2.3. Panel LM test

We now turn to the panel LM unit root tests. We define the univariate LM unit root test statistic for the ith cross-section unit obtained from the augmented regression (9) as LM_i^{τ}. The panel LM unit root test statistic is computed by averaging the univariate LM statistics for each cross-section units as

$$L\bar{M} = \frac{1}{N}\sum_{i=1}^{N} LM_i^{\tau},$$

where LM_i^{τ} represents either $MinLM$ or $Max|tb|$ of the ith cross-section unit, and N stands for the number of cross-section units. A standardized panel LM statistic is defined as

$$\text{Panel LM test statistic} = \frac{\sqrt{N}[L\bar{M} - E(LM_T(k)]}{\sqrt{Var(LM_T(k))}}, \tag{10}$$

where $E(LM_T(k))$ and $Var(LM_T(k))$ are the mean and variance of the univariate LM test statistic under the null hypothesis, and are computed for various combinations of T and k. The invariance property of the panel LM test statistic enables the same mean and variance values to be used with or without a structural change in the level.[5] The panel LM test statistic converges to a standard normal distribution asymptotically. But considering that the asymptotic critical values often differ from their finite-sample counterparts in practice, we therefore derive finite-sample critical values by the bootstrap method. As mentioned in the previous section the bootstrap method has been found to be effective in correcting size distortions of panel tests in the presence of cross-sectional dependency. For example, Maddala and Wu (1999) have documented the consequences of ignoring cross-sectional dependency such as severe size distortions. Simulations conducted by Chang (2004) show that the bootstrap panel unit root tests perform well in finite samples relative to the IPS test statistic. In view of the relatively short ($T=27$) time series data we are using in this study, we only consider the panel LM test with one break in the level.[6]

2.4. Data

The panel LM unit root tests are performed using a sample of panel data on 28 provinces in China covering the period 1978–2004. We have taken out two provinces, namely, Hainan and Chongqing for which time series for health care expenditure are only available for selected years. In addition, Tibet is excluded from the sample because of the lack of price series which prevents us from deriving the health care expenditure in real terms. The health care expenditure series are expressed in per capita terms (denoted as HE) and logarithmic form. Income is measured by

[5] When there is a change in the slope, such asymptotic invariance property does not hold. Therefore, Im et al. (2005) did not consider Models 2 and 3. The simulated means and variances are available on request.

[6] Nunes (2004) shows that the null distributions of the test statistics under the two break date selection methods are asymptotically invariant to a change in the level but not to a change in the slope.

provincial GDP. It is expressed in per capita and logarithmic form. Factors that are not related to income include the dependency ratio of the old aged population (NI1), the proportion of the population aged 65 and over (NI2), the ratio of the health expenditure funded by public sector (NI3), and government budget deficits (NI4). Variables NI1, NI2, and NI3 are expressed in logarithmic form, whereas NI4 is measured in 1990 RMB, and is adjusted for inflation rates. The major sources for HE, GDP, and non-income variables are various issues of *Yearbook of Public Health in Peoples' Republic of China, China Population Statistics Yearbook, China Statistical Yearbook*, provincial Statistical Yearbooks, and Almanacs by province.

3. Empirical test results

3.1. Panel LM unit root tests

To obtain panel LM test statistics, we first calculate univariate LM test statistics by applying the two test statistics described in the previous section to variables HE, GDP, NI1, NI2, NI3 and NI4. The univariate test results of applying the test statistics *MinLM* and $Max|tb|$ to HE and GDP for each of the 28 provinces of China are shown in Tables 1 and 2. The corresponding panel LM test results are given in Table 3. However, for space consideration, we do not show all the details of the univariate test results for non-income variables NI1–NI4, and only report their panel LM test results in Table 3.[7]

Our tests of the unit root null hypothesis are based on the bootstrapped critical values which are obtained by simulating the fitted Eq. (9) using the bootstrap procedure with 5000 replications.[8] For variables HE and GDP, the bootstrapped critical values are presented in the last three columns of Tables 1 and 2, respectively. With the bootstrapped critical values, empirical sizes are computed for each variable. For space consideration, only the sizes for HE and GDP are presented under the column of empirical size in Tables A1 and A2 of Appendix A.[9]

Comparing the empirical sizes and the frequency of true break selection under the null hypothesis between the two LM-type test statistics, the LM tests proposed by Nunes (2004), namely, $Max|tb|$, are found to outperform the minimum LM test statistics, *MinLM*. The test results of $Max|tb|$ indicate that allowing for one structural break all variables are non-stationary except the proportion of the population aged 65 and over (NI2), which is found to be stationary. For the health expenditure variable HE, structural breaks are found to exist in majority of the provinces in China. Most of the breaks occurred during the period 1994–2002 when China undertook more significant economic and health policy changes.

3.2. Panel cointegration tests

Having completed the test for the order of integration of the variables in the panel, we now turn to the cointegration test. Conventional tests for cointegration include the residual-based test of Engle and Granger (1987) and the system-based method of Johansen (1988, 1991). The problem with these conventional cointegration tests is low-test power, which is a major reason for pooling data into a panel. Recently developed panel cointegration tests, such as those of Kao (1999)

[7] Detailed results for non-income variables NI1 to NI4 are available upon request.
[8] The bootstrap procedure is available upon request.
[9] The empirical sizes and other finite sample properties for variables NI1 to NI4 are available upon request.

Table 1
LM unit root tests on per capita health expenditure by region, 1978–2004

Series	Break selection scheme	LM test statistic	Break year	k	Bootstrapped critical values		
					1%	5%	10%
Beijing	*MinLM*[a]	−2.53	2002	3	−5.15	−4.27	−3.83
	Max|tb|[b]	−2.53	2002	3	−4.64	−3.85	−3.43
Tianjin	*MinLM*	−2.84	1994	5	−5.04	−4.25	−3.78
	Max|tb|	−1.53	2002	5	−4.67	−3.75	−3.37
Hebei	*MinLM*	−2.78	1987	4	−5.06	−4.17	−3.79
	Max|tb|	−1.77	2002	4	−4.46	−3.61	−3.22
Shanxi	*MinLM*	−1.90	–	5	−5.26	−4.28	−3.87
	Max|tb|	−2.02	1987	2	−4.83	−3.85	−3.41
Liaoning	*MinLM*	−2.25	1987	5	−5.26	−4.33	−3.93
	Max|tb|	−1.90	2002	1	−4.64	−3.76	−3.34
Jilin	*MinLM*	−1.77	–	5	−5.36	−4.41	−3.96
	Max|tb|	−1.73	1996	1	−4.84	−3.91	−3.44
Heilongjiang	*MinLM*	−3.97*	2002	5	−5.20	−4.37	−3.94
	Max|tb|	−3.97*	2002	5	−4.98	−4.05	−3.59
Shanghai	*MinLM*	−2.05	1997	5	−5.05	−4.24	−3.84
	Max|tb|	−2.05	1997	4	−4.66	−3.82	−3.39
Jiang su	*MinLM*	−1.36	1986	3	−5.29	−4.31	−3.87
	Max|tb|	−1.35	2002	3	−4.52	−3.68	−3.26
Zhejiang	*MinLM*	−1.81	1985	5	−5.21	−4.27	−3.89
	Max|tb|	−1.81	1985	0	−4.92	−3.96	−3.52
Anhui	*MinLM*	−0.96	–	0	−5.34	−4.34	−3.85
	Max|tb|	−0.51	1992	0	−4.72	−3.88	−3.38
Fujian	*MinLM*	−1.31	–	2	−4.97	−4.19	−3.78
	Max|tb|	−1.63	1986	5	−4.55	−3.63	−3.21
Jiangxi	*MinLM*	−2.54	–	5	−5.23	−4.30	−3.88
	Max|tb|	−2.78	1986	4	−4.74	−3.89	−3.41
Shangdong	*MinLM*	−2.39	1993	4	−5.09	−4.27	−3.87
	Max|tb|	−2.39	1996	4	−4.69	−3.87	−3.44
Henan	*MinLM*	−1.24	–	5	−5.08	−4.17	−3.75
	Max|tb|	−1.36	1995	5	−4.46	−3.64	−3.22
Hubei	*MinLM*	−1.52	–	1	−4.92	−4.00	−3.58
	Max|tb|	−1.61	1997	0	−4.29	−3.50	−3.11
Hunan	*MinLM*	−1.28	–	2	−5.21	−4.25	−3.86
	Max|tb|	−1.65	1996	2	−4.69	−3.69	−3.32
Guangdong	*MinLM*	−1.80	–	0	−5.11	−4.34	−3.91
	Max|tb|	−2.01	1988	0	−4.82	−3.90	−3.47
InnerMonglia	*MinLM*	−2.86	2002	5	−4.86	−3.94	−3.55
	Max|tb|	−2.86	2002	5	−4.45	−3.55	−3.10
Guangxi	*MinLM*	−1.57	–	2	−5.31	−4.37	−3.92
	Max|tb|	−1.56	1992	3	−4.84	−3.98	−3.50

Table 1 (*Continued*)

Series	Break selection scheme	LM test statistic	Break year	*k*	Bootstrapped critical values		
					1%	5%	10%
Sichuan	*MinLM*	−2.35	2002	2	−5.15	−4.30	−3.92
	Max|*tb*|	−2.35	2002	2	−4.55	−3.75	−3.34
Guizhou	*MinLM*	−1.19	–	4	−5.07	−4.24	−3.83
	Max|*tb*|	−1.71	1987	4	−4.56	−3.74	−3.30
Yunnan	*MinLM*	−3.28	1987	2	−5.11	−4.20	−3.83
	Max|*tb*|	−2.22	1994	1	−4.68	−3.87	−3.44
Shaanxi	*MinLM*	−1.48	1998	5	−5.07	−4.17	−3.78
	Max|*tb*|	−1.48	1998	0	−4.52	−3.68	−3.24
Gansu	*MinLM*	−2.24	–	5	−4.98	−4.23	−3.81
	Max|*tb*|	−1.04	1997	0	−4.42	−3.67	−3.24
Qinghai	*MinLM*	−0.57	–	4	−5.55	−4.29	−3.63
	Max|*tb*|	−2.11	2000	1	−4.30	−3.10	−2.69
Ningxia	*MinLM*	−2.10	1995	4	−5.25	−4.33	−3.89
	Max|*tb*|	−1.85	1994	2	−4.63	−3.80	−3.35
Xingjiang	*MinLM*	−1.80	–	5	−5.24	−4.26	−3.87
	Max|*tb*|	−1.92	1998	5	−4.74	−3.88	−3.43

Notes: (–): Refers to the case of no breaks. (*k*): lag order. Bootstrapped critical values are from Table A1 of Appendix A.

[a] *MinLM* = Minimum LM test of Lee and Strazicich (2003, 2004).

[b] *Max*|*tb*| = test statistics of Nunes (2004).

[*] Denotes significance at the 10% level.

and Pedroni (1999, 2004), that take advantage of combining information from the time-series observations with that from the cross-section data are shown to be more powerful compared with the conventional tests that are purely time-series based.

Both the Pedroni and Kao tests are similar in essence to the Engle and Granger (1987) analysis, since they involve the use of residuals derived from the panel. The Pedroni's test takes "no cointegration" as the null hypothesis, whereas the Kao test took 'cointegration' as the null. As we have taken "no unit roots" as the null hypothesis in the previous section on panel unit root tests, we follow the same by applying Pedroni's method to test the null of "no cointegration" between HE and the related variables. Moreover, the Pedroni's test has other attractive features such as allowing for heterogeneity among the individual members of the panel and endogeneity of the regressors. In particular, the ability to capture heterogeneity due to the member specific characteristics is important in the analysis of panel data. Pedroni proposes *seven* panel cointegration statistics, four based on pooling along the within-dimension and three based on pooling along the between-dimension. Within the first category, the panel variance (ν), panel *rho*, and panel PP, involve the use of non-parametric corrections, and the fourth is a parametric test called the panel ADF test. In the second category, the group *rho* and group PP use non-parametric corrections while the group ADF test is an ADF-based test. The details of these statistics can be found in Pedroni (1999).

The estimated Pedroni's test statistics are given in Table 4. Since only non-stationary series are considered as regressors, the proportion of the population aged 65 and over (NI2) that is found to be stationary is therefore excluded from the analysis. The results presented in Table 4 show that the null of no cointegration is rejected by four of the seven test statistics at the 5% significance level when testing the cointegration between per capita health care expenditure (HE), the dependency

Table 2
LM unit root tests on per capita income by region, 1978–2004

Series	Break selection scheme	LM test statistic	Break year	k	Bootstrapped critical values		
					1%	5%	10%
Beijing	*MinLM*[a]	−2.74	1988	1	−7.28	−5.58	−4.94
	Max|*tb*|[b]	−2.74	1988	1	−5.75	−4.47	−3.94
Tianjin	*MinLM*	−2.14	–	1	−5.34	−4.26	−3.81
	Max|*tb*|	−1.82	1988	1	−5.84	−4.56	−4.00
Hebei	*MinLM*	−2.5	1985	1	−7.08	−5.68	−5.04
	Max|*tb*|	−1.81	1989	5	−5.57	−4.40	−3.90
Shanxi	*MinLM*	−5.48	–	5	−5.14	−4.23	−3.80
	Max|*tb*|	−6.39	1995	5	−5.62	−4.48	−3.99
Liaoning	*MinLM*	−6.34**	1996	5	−7.16	−5.69	−5.13
	Max|*tb*|	−4.58**	1989	1	−5.37	−4.35	−3.90
Jilin	*MinLM*	−4.94*	1988	4	−6.68	−5.45	−4.91
	Max|*tb*|	−4.94**	1988	4	−5.60	−4.51	−3.96
Heilongjiang	*MinLM*	−2.27	–	4	−5.38	−4.32	−3.85
	Max|*tb*|	−2.27	–	4	−5.38	−4.32	−3.85
Shanghai	*MinLM*	−2.02	1985	1	−7.42	−5.90	−5.23
	Max|*tb*|	−1.41	1988	3	−5.76	−4.53	−4.02
Jiang su	*MinLM*	−3.31	1991	1	−6.72	−5.46	−4.93
	Max|*tb*|	−3.31	1991	1	−5.54	−4.48	−3.96
Zhejiang	*MinLM*	−3.62	–	3	−5.41	−4.33	−3.81
	Max|*tb*|	−2.57	1988	1	−5.58	−4.53	−3.98
Anhui	*MinLM*	−4.11	1990	3	−7.17	−5.54	−4.90
	Max|*tb*|	−4.11*	1990	3	−5.74	−4.38	−3.91
Fujian	*MinLM*	−3.46	1992	4	−7.07	−5.68	−5.10
	Max|*tb*|	−3.46	1992	4	−5.72	−4.55	−4.02
Jiangxi	*MinLM*	−3.11	–	5	−5.45	−4.30	−3.83
	Max|*tb*|	−4.09*	1992	1	−5.88	−4.50	−3.97
Shangdong	*MinLM*	−2.43	–	1	−5.37	−4.28	−3.83
	Max|*tb*|	−2.43	–	1	−5.37	−4.28	−3.83
Henan	*MinLM*	−5.87**	1986	5	−7.18	−5.71	−5.13
	Max|*tb*|	−4.34*	1989	5	−5.66	−4.42	−3.93
Hubei	*MinLM*	−2.69	–	1	−5.40	−4.30	−3.80
	Max|*tb*|	−1.87	1989	0	−5.72	−4.50	−3.98
Hunan	*MinLM*	−2.99	–	1	−5.27	−4.28	−3.77
	Max|*tb*|	−2.19	1988	1	−5.53	−4.47	−3.99
Guangdong	*MinLM*	−2.70	1988	4	−7.18	−5.79	−5.16
	Max|*tb*|	−1.22	1992	0	−5.67	−4.54	−4.01
InnerMonglia	*MinLM*	−5.38	–	3	−5.42	−4.20	−3.73
	Max|*tb*|	−5.38	–	3	−5.42	−4.20	−3.73
Guangxi	*MinLM*	−1.79	–	5	−5.32	−4.30	−3.83
	Max|*tb*|	−1.79	–	5	−5.32	−4.30	−3.83

Table 2 (*Continued*)

Series	Break selection scheme	LM test statistic	Break year	*k*	Bootstrapped critical values		
					1%	5%	10%
Sichuan	*MinLM*	−2.76	–	5	−5.23	−4.28	−3.78
	Max|tb|	−3.15	1988	5	−5.60	−4.41	−3.91
Guizhou	*MinLM*	−3.52	1990	5	−7.26	−5.65	−5.01
	Max|tb|	−2.75	–	5	−5.44	−4.19	−3.71
Yunnan	*MinLM*	−4.80	1987	3	−7.23	−5.81	−5.15
	Max|tb|	−4.80**	1987	3	−5.87	−4.54	−3.96
Shaanxi	*MinLM*	−9.02**	1987	5	−7.29	−5.55	−4.92
	Max|tb|	−9.02**	1987	5	−5.54	−4.38	−3.92
Gansu	*MinLM*	−4.98	1985	3	−7.39	−5.73	−5.12
	Max|tb|	−3.59	1989	5	−5.76	−4.49	−3.95
Qinghai	*MinLM*	−2.96	1989	4	−7.20	−5.38	−4.74
	Max|tb|	−2.96	1989	4	−5.69	−4.36	−3.79
Ningxia	*MinLM*	−3.45	–	2	−5.11	−4.11	−3.66
	Max|tb|	−1.96	1988	4	−5.91	−4.47	−3.89
Xingjiang	*MinLM*	−3.10	1999	4	−7.28	−5.76	−5.09
	Max|tb|	−1.54	1988	1	−5.73	−4.54	−3.97

Notes: (–): Refers to the case of no breaks. (*k*): lag order. Bootstrapped critical values are from Table A1 of Appendix A.
[a] *MinLM* = Minimum LM test of Lee and Strazicich (2003, 2004).
[b] *Max|tb|* = test statistics of Nunes (2004).
[*] Denotes significance at the 10% level.
[**] Denotes significance at the 5% level.

Table 3
Panel LM unit root tests on HE, GDP and four non-income series, 1978–2004

Series	Break selection scheme	LM statistic	Bootstrapped critical values		
			1%	5%	10%
HE	*MinLM*	0.80	−6.86	−5.92	−5.46
	Max|tb|	1.52	−3.62	−2.76	−2.37
GDP	*MinLM*	−3.66	−6.93	−5.99	−5.55
	Max|tb|	−0.15	−3.91	−3.09	−2.61
Nil	*MinLM*	−10.56***	−6.82	−5.89	−5.43
	Max|tb|	−0.84	−3.46	−2.59	−2.08
NI2	*MinLM*	−11.57***	−6.75	−5.91	−5.46
	Max|tb|	−3.15**	−3.36	−2.54	−2.12
NI3	*MinLM*	−2.94	−6.99	−5.98	−5.50
	Max|tb|	0.77	−3.81	−2.93	−2.45
NI4	*MinLM*	−0.75	−6.58	−5.66	−5.26
	Max|tb|	2.51	−3.27	−2.50	−2.09

Notes: HE = real per capita health expenditure; GDP = real per capita income; NI1 = the dependency ratio of old aged population (NI1); NI2 = the proportion of population aged 65 and over; NI3 = the ratio of the health expenditure funded by public sector; and NI4 = real government budget deficits.
[**] Denotes significance at the 5% level.
[***] Denotes significance at the 1% level.

Table 4
Results from Pedroni panel cointegration tests

Variables included in cointegrating regression	Panel statistics				Group statistics		
	ν	*rho*	PP	adf	*Rho*	PP	adf
GDP, NI1, NI3 & NI4	−0.90	2.87	1.22	1.02	3.94	0.69	0.21
NI1 & NI4	0.27	−0.86	−3.01**	−2.94**	1.20	−2.35**	−2.30**

Note: $N=28$, $T=27$. The 5% and 10% critical values are −1.65 and −1.28, respectively. All Pedroni statistics use the left tail of the normal distribution to reject the null, except the panel ν test which variance ratio statistic and takes positive values.
** Denotes significance at the 5% level.

ratio of the old aged population (NI1) and the government budget deficits (NI4). These results suggested that the dependency ratio of the old age population and the government budget deficits are helpful in explaining the behavior of China's health care expenditure in the long run.

3.3. Modelling panel cointegrated regressions

On the basis of the existence of the cointegrated relationship found above, we now proceed to estimate the cointegrated regressions, that is

$$\mathrm{HE}_{it} = \alpha_0 + \alpha_1 \mathrm{NI1}_{i,t} + \alpha_2 \mathrm{NI4}_{it} + \varepsilon_{i,t}, \tag{11}$$

where NI1 and NI4 are, respectively, the dependency ratio of the old aged population, and government budget deficits, and ε_t is an error term. The estimated coefficient $\hat{\alpha}_1$ is the long-run elasticity on NI1. One issue that concerns us is the bias caused by the two regressors of (11) which are likely to be endogenous. In their study of panel cointegrated regressions, Kao and Chiang (2000) found that the OLS estimator is biased due to the endogeneity problem of the regressors, and suggested the bias-correction methods such as the dynamic OLS (DOLS) and fully modified (FM) OLS using pooled cross-section and time-series data. Kao and Chiang (2000) also showed that the DOLS estimator outperforms the FMOLS and OLS estimators in the estimation of cointegrated panel regressions. These findings have motivated our use of the DOLS and FMOLS in dealing with the endogeneity bias of the regressor.

Using the conventional OLS and the two bias-correction methods DOLS and FMOLS methods, we estimate Eq. (11). Table 5 shows the estimated long-run elasticity of the variable NI1 based on the conventional OLS, DOLS, and FMOLS methods along with their corresponding t-values. It is noteworthy that when employing the DOLS method, for bias-correction purposes, we have included one lag and 4 leads of the first differences of the right-hand side variables of Eq. (11),

Table 5
Parameter estimates of cointegrated panel regressions using OLS, DOLS and FMOLS methods (dependent variable: HE)

	OLS	DOLS	FMOLS
NI1	−0.055 (−0.100)	1.851 (2.692)***	1.317 (2.371)***
NI4	−0.102 (−5.019)***	−0.263 (−10.075)***	−0.061 (−2.910)***
R^2	0.679	0.732	0.701

Notes: Equations are based on the pooled data from 1978 to 2004 for 28 provinces. Figures in parentheses are t-values.
*** Denotes significance at the 1% level.

and estimate the following augmented cointegrated regression:

$$\mathrm{HE}_{it} = \alpha_0 + \alpha_1 \mathrm{NI1}_{it} + \alpha_2 \mathrm{NI4}_{it} + \beta_1 \Delta\mathrm{NI1}_{it} + \beta_2 \Delta\mathrm{NI4}_{it} + \gamma_{11} \Delta\mathrm{NI1}_{i,t-1} + \gamma_{21} \Delta\mathrm{NI4}_{i,t-1} + \sum_{j=1}^{4} \delta_{1j} \Delta\mathrm{NI1}_{i,t+j} + \sum_{j=1}^{4} \delta_{2k} \Delta\mathrm{NI4}_{t+k} + \varepsilon_{i,t} \quad (12)$$

Table 5 shows the estimated long-run elasticities of the variable NI1 based on the conventional OLS, DOLS, and FMOLS methods along with their corresponding *t*-values. As such we have the following observations. First, the OLS estimate of the old aged dependency ratio variable (NI1) has an unexpected negative sign. Since the OLS estimators are biased, the t-statistics of the OLS estimators presented in Table 5 do not have the conventional t-distributions. Therefore, we should not place too much emphasis on the values of the t-statistics for the OLS estimators. Both the DOLS and FM estimates of NI1 have values greater than one, suggesting that China's health care expenditure is elastic to the old aged dependency ratio. An increase in the old aged dependency ratio by 1%, will increase health care expenditure by 1.31–1.85%, ceteris paribus, in China in the long run. This is consistent with the situation in China in which most Chinese pay directly for health services when they receive them; the share of out-of-pocket health spending in China has risen steadily since the late 1970s.[10] With this development, the increase in the old aged population is therefore associated with much larger increases in health care expenditure. Second, the two bias-correction estimates of DOLS and FMOLS indicate that the budget deficit variable (NI4) has significant impact on health expenditure. As Kao and Chiang (2000) have pointed out that the FM estimator could be inferior to the conventional OLS in some cases, we can make our conclusion based on the DOLS estimates. Comparing the OLS estimate of the coefficient of the budget deficit variable with that of the DOLS, we see that the OLS method tends to understate the impact of this variable. The DOLS estimate associated with the budget deficits shows that with for every 10 million increase in budget deficits, the health expenditure will shrink by about 26.3%, ceteris paribus, in the long run. This finding suggests the government budget deficits have imposed a significant constraint on public health expenditures. It is consistent with the existing situation in many parts of China that with eroding government revenues,[11] people receive less from publicly funded health care, and their out-of-pocket health care spending increases.

4. Conclusion

In this study, from a pooled cross-section and time series data set of China, we provide the evidence on the stationarity property of the income and non-income variables by using panel LM unit root tests that allow for structural breaks under both the null and alternative hypotheses. We consider two endogenous break selection methods. To implement the LM unit root tests, the bootstrapped critical values are obtained instead of relying on the asymptotic normal distribution. We find that the reliance on asymptotic critical values tends to reject the null hypothesis of unit root too frequently. This is probably the reason that Jewell et al. (2003) find that the health expenditure and GDP are stationary using the OECD data.

[10] World Bank report (1997), p. 25.

[11] Statistics show that local government revenues as a share of GDP have declined since 1978. To keep the budget deficit in check, local government expenditures as a share of GDP reduced as well.

The panel unit root test supports the hypothesis of non-stationarity of health care expenditure, income, and the non-income variables on the dependency ratio of the old aged population, the proportion of health expenditures that is publicly funded, and the government budget deficits. The only variable that is found to be stationary is the proportion of the population aged 65 and above. Variables such as the relative price of health care, and the number of physicians are not considered due to the lack of data.

One important finding that emerges from this study is the government budget deficits are found to have significant long-run impact on regional health care expenditures. It is now well known that the economic development in China is geographically skewed, with coastal provinces growing much faster than the inland ones. This skewed pattern of economic development has resulted in more developed coastal provinces such as Jiangsu and Zhejiang which enjoy government budget surplus, whereas inland provinces such as Yunnan, and Sichuan suffer from severe government budget deficit problems. The government budget deficits had imposed a stronger constraint on public health expenditures in the less developed inland provinces compared to the richer ones. Our estimate shows that, on the average, for every 10 million RMB increase in budget deficits, the health expenditure will be reduced by about 26.3% in the long run, ceteris paribus. This finding therefore provides an explanation as to why there are substantial disparities in health service coverage, and in China's health care financing policy between rich and poor provinces. Although the income variable GDP is not directly used in the cointegration estimation, we may say that the use of the government budget deficits to some extent has incorporated the effect of GDP in China.

Another finding is that the old aged dependency ratio has significant impact on China's health care expenditure. The high long-run elasticity associated with the old aged dependency ratio suggests that finding funds to increase public spending on health care for old age population is needed, in particular, the rural poor who pay directly for health services when they receive them.

Appendix A

Tables A1 and A2

Table A1
Bootstrapped critical values, empirical sizes and frequency of true break for HE, 1978–2004

Series	Break selection scheme	Regression (9)						Empirical size[a] (%)	Frequency of true break[b] selection under H_0 (%)
		DGP: 0 breaks			DGP: 1 break				
		1%	5%	10%	1%	5%	10%		
Beijing	*MinLM*	−4.38	−3.61	−3.22	−5.15	−4.27	−3.83	4.72	19.08
	Max|tb|	−4.27	−3.48	−3.11	−4.64	−3.85	−3.43	5.14	50.08
Tianjin	*MinLM*	−4.26	−3.52	−3.13	−5.04	−4.25	3.78	4.36	7.78
	Max|tb|	−4.46	−3.60	−3.21	−4.67	−3.75	−3.37	5.82	37.06
Hebei	*MinLM*	−4.24	−3.48	−3.14	−5.06	−4.17	−3.79	6.24	16.76
	Max|tb|	−4.42	−3.58	−3.20	−4.46	−3.61	−3.22	5.54	92.04
Shanxi	*MinLM*	−4.40	−3.62	−3.22	−5.26	−4.28	−3.87	5.22	6.12
	Max|tb|	−4.21	−3.53	−3.15	−4.83	−3.85	−3.41	5.20	36.88
Liaoning	*MinLM*	−4.28	−3.48	−3.10	−5.26	−4.33	−3.93	4.98	15.88
	Max|tb|	−4.44	−3.58	−3.18	−4.64	−3.76	−3.34	4.94	65.1

Table A1 (*Continued*)

Series	Break selection scheme	Regression (9)						Empirical size[a] (%)	Frequency of true break[b] selection under H_0 (%)
		DGP: 0 breaks			DGP: 1 break				
		1%	5%	10%	1%	5%	10%		
Jilin	*MinLM*	−4.34	−3.49	−3.16	−5.36	−4.41	−3.96	4.36	6.32
	Max|tb|	−4.17	−3.42	−3.11	−4.84	−3.91	−3.44	4.96	48.28
Heilongjiang	*MinLM*	−4.39	−3.52	−3.16	−5.20	−4.37	−3.94	4.84	11.14
	Max|tb|	−4.39	−3.51	−3.19	−4.98	−4.05	−3.59	4.00	26.48
Shanghai	*MinLM*	−4.21	−3.43	−3.07	−5.05	−4.24	−3.84	4.60	19.56
	Max|tb|	−4.38	−3.61	−3.22	−4.66	−3.82	−3.39	−5.00	57.22
Jiangsu	*MinLM*	−4.27	−3.48	−3.11	−5.29	−4.31	−3.8	−4.66	18.86
	Max|tb|	−4.26	−3.52	−3.13	−4.52	−3.68	−3.26	4.76	85.5
Zhejiang	*MinLM*	−4.46	−3.60	−3.21	−5.21	−4.27	−3.89	5.86	11.62
	Max|tb|	−4.24	−3.48	−3.14	−4.92	−3.96	−3.52	5.16	26.72
Anhui	*MinLM*	−4.42	−3.58	−3.20	−5.34	−4.34	−3.85	4.68	5.18
	Max|tb|	−4.40	−3.62	−3.22	−4.72	−3.88	−3.38	5.02	40.54
Fujian	*MinLM*	−4.21	−3.53	−3.15	−4.97	−4.19	−3.78	5.00	5.94
	Max|tb|	−4.28	−3.48	−3.10	−4.55	−3.63	−3.21	5.20	97.98
Jiangxi	*MinLM*	−4.44	−3.58	−3.18	−5.23	−4.30	−3.88	5.22	5.4
	Max|tb|	−4.34	−3.49	−3.16	−4.74	−3.89	−3.41	4.82	31.52
Shangdong	*MinLM*	−4.17	−3.42	−3.11	−5.09	−4.27	−3.87	4.84	23.36
	Max|tb|	−4.39	−3.52	−3.16	−4.69	−3.87	−3.44	4.82	63.84
Henan	*MinLM*	−4.21	−3.51	−3.19	−5.08	−4.17	−3.75	5.30	5.88
	Max|tb|	−4.21	−3.43	−3.07	−4.46	−3.64	−3.22	4.42	96.86
Hubei	*MinLM*	−3.94	−3.23	−2.90	−4.92	−4.00	−3.58	4.60	5.36
	Max|tb|	−4.40	−3.54	−3.17	−4.29	−3.50	−3.11	4.68	99.34
Hunan	*MinLM*	−4.44	−3.58	−3.21	−5.21	−4.25	−3.86	5.56	4.72
	Max|tb|	−4.06	−3.26	−2.92	−4.69	−3.69	−3.32	5.12	29.54
Guangdong	*MinLM*	−4.46	−3.63	−3.21	−5.11	−4.34	−3.91	4.82	8.14
	Max|tb|	−4.36	−3.57	−3.15	−4.82	−3.90	−3.47	5.00	41.14
InnerMonglia	*MinLM*	−4.39	−3.50	−3.11	−4.86	−3.94	−3.55	5.22	22.14
	Max|tb|	−4.32	−3.52	−3.14	−4.45	−3.55	−3.10	4.62	43.4
Guangxi	*MinLM*	−4.27	−3.52	−3.12	−5.31	−4.37	−3.92	4.74	4.68
	Max|tb|	−4.34	−3.48	−3.13	−4.84	−3.98	−3.50	5.08	36.6
Sichuan	*MinLM*	−4.34	−3.45	−3.06	−5.15	−4.30	−3.92	5.22	21.94
	Max|tb|	−4.36	−3.53	−3.14	−4.55	−3.75	−3.34	5.56	55.62
Guizhou	*MinLM*	−4.30	−3.52	−3.17	−5.07	−4.24	−3.83	6.18	6.24
	Max|tb|	−3.94	−3.23	−2.90	−4.56	−3.74	−3.30	5.38	75.7
Yunnan	*MinLM*	−4.40	−3.54	−3.17	−5.11	−4.20	−3.83	5.88	15.68
	Max|tb|	−4.44	−3.58	−3.21	−4.68	−3.87	−3.44	5.04	52.16
Shaanxi	*MinLM*	−4.06	−3.26	−2.92	−5.07	−4.17	−3.78	4.58	18.88
	Max|tb|	−4.46	−3.63	−3.21	−4.52	−3.68	−3.24	5.18	43.1
Gansu	*MinLM*	−4.36	−3.57	−3.15	−4.98	−4.23	−3.81	5.04	5.5
	Max|tb|	−4.39	−3.50	−3.11	−4.42	−3.67	−3.24	5.18	90.04

Table A1 (*Continued*)

Series	Break selection scheme	Regression (9)						Empirical size[a] (%)	Frequency of true break[b] selection under H_0 (%)
		DGP: 0 breaks			DGP: 1 break				
		1%	5%	10%	1%	5%	10%		
Qinghai	*MinLM*	−4.32	−3.52	−3.14	−5.55	−4.29	−3.63	5.70	5.68
	Max|*tb*|	−4.27	−3.52	−3.12	−4.30	−3.10	−2.69	5.06	94.7
Ningxia	*MinLM*	−4.34	−3.48	−3.13	−5.25	−4.33	−3.89	4.54	6.48
	Max|*tb*|	−4.34	−3.45	−3.06	−4.63	−3.80	−3.35	4.76	70.4
Xingjiang	*MinLM*	−4.36	−3.53	−3.14	−5.24	−4.26	−3.87	5.40	6.98
	Max|*tb*|	−4.30	−3.52	−3.17	−4.74	−3.88	−3.43	4.36	34.8
Panel LM	*MinLM*	−3.22	−2.46	−2.03	−6.86	−5.92	−5.46	5.10	
	Max|*tb*|	−3.22	−2.46	−2.03	−3.62	−2.76	−2.37	6.00	

[a] Empirical size is computed by allowing one break under the null but the break is neglected.
[b] True break dates are assumed to be the same as those estimated in Table 1.

Table A2
Bootstrapped critical values, empirical sizes and frequency of true break for GDP

Series	Break selection scheme	Regression (9)						Empirical size[a] (%)	Frequency of true break[b] selection under H_0 (%)
		DGP: 0 breaks			DGP: 1 break				
		1%	5%	10%	1%	5%	10%		
Beijing	*MinLM*	−4.27	−3.43	−3.08	−5.04	−4.30	−3.87	4.30	22.74
	Max|*tb*|	−4.27	−3.43	−3.08	−4.58	−3.63	−3.24	4.82	33.32
Tianjin	*MinLM*	−4.34	−3.55	−3.18	−5.02	−4.28	−3.89	5.50	17.10
	Max|*tb*|	−4.34	−3.55	−3.18	−4.93	−4.02	−3.59	4.58	5.36
Hebei	*MinLM*	−4.20	−3.50	−3.13	−5.05	−4.20	−3.80	4.92	9.62
	Max|*tb*|	−4.20	−3.50	−3.13	−4.68	−3.82	−3.38	4.58	18.02
Shanxi	*MinLM*	−4.30	−3.58	−3.24	−5.18	−4.35	−3.93	4.96	11.06
	Max|*tb*|	−4.30	−3.58	−3.24	−4.91	−3.96	−3.51	4.58	8.18
Liaoning	*MinLM*	−4.28	−3.46	−3.11	−5.06	−4.19	−3.82	4.98	16.72
	Max|*tb*|	−4.28	−3.46	−3.11	−4.70	−3.86	−3.39	5.08	8.42
Jilin	*MinLM*	−4.23	−3.54	−3.17	−5.03	−4.28	−3.85	4.58	26.56
	Max|*tb*|	−4.23	−3.54	−3.17	−4.59	−3.69	−3.27	5.46	77.34
Heilongjiang	*MinLM*	−4.33	−3.54	−3.16	−5.12	−4.15	−3.76	4.82	21.46
	Max|*tb*|	−4.33	−3.54	−3.16	−4.90	−3.96	−3.51	4.94	5.28
Shanghai	*MinLM*	−4.41	−3.63	−3.24	−5.06	−4.22	−3.81	5.42	22.44
	Max|*tb*|	−4.41	−3.63	−3.24	−4.82	−3.96	−3.52	4.74	33.08
Jiangsu	*MinLM*	−4.28	−3.52	−3.14	−5.08	−4.22	−3.83	5.14	21.32
	Max|*tb*|	−4.28	−3.52	−3.14	−4.67	−3.79	−3.38	4.62	53.36
Zhejiang	*MinLM*	−4.47	−3.55	−3.15	−5.22	−4.29	−3.89	4.82	7.54
	Max|*tb*|	−4.47	−3.55	−3.15	−4.59	−3.73	−3.31	4.98	80.38
Anhui	*MinLM*	−4.34	−3.48	−3.10	−5.12	−4.24	−3.85	4.66	18.02
	Max|*tb*|	−4.34	−3.48	−3.10	−4.55	−3.72	−3.35	5.34	30.36

Table A2 (*Continued*)

Series	Break selection scheme	Regression (9)						Empirical size[a] (%)	Frequency of true break[b] selection under H_0 (%)
		DGP: 0 breaks			DGP: 1 break				
		1%	5%	10%	1%	5%	10%		
Fujian	*MinLM*	−4.36	−3.60	−3.20	−5.19	−4.36	−3.93	4.62	4.84
	Max|tb|	−4.36	−3.60	−3.20	−4.82	−3.98	−3.50	5.00	24.10
Jiangxi	*MinLM*	−4.31	−3.56	−3.18	−5.12	−4.30	−3.86	4.78	16.86
	Max|tb|	−4.31	−3.56	−3.18	−4.87	−3.96	−3.50	4.72	40.92
Shangdong	*MinLM*	−4.29	−3.53	−3.14	−5.03	−4.25	−3.85	5.90	21.10
	Max|tb|	−4.29	−3.53	−3.14	−4.62	−3.80	−3.41	4.64	48.14
Henan	*MinLM*	−4.37	−3.53	−3.15	−5.27	−4.29	−3.84	5.64	7.46
	Max|tb|	−4.37	−3.53	−3.15	−4.60	−3.73	−3.34	5.48	43.18
Hubei	*MinLM*	−4.42	−3.59	−3.23	−5.19	−4.34	−3.93	4.88	15.32
	Max|tb|	−4.42	−3.59	−3.23	−4.74	−3.90	−3.43	5.20	38.44
Hunen	*MinLM*	−4.18	−3.52	−3.15	−5.03	−4.29	−3.86	4.34	25.36
	Max|tb|	−4.18	−3.52	−3.15	−4.63	−3.75	−3.36	4.88	61.10
Guangdong	*MinLM*	−4.34	−3.57	−3.19	−5.24	−4.30	−3.89	4.64	18.34
	Max|tb|	−4.34	−3.57	−3.19	−5.00	−3.91	−3.48	4.82	47.80
InnerMonglia	*MinLM*	−4.39	−3.56	−3.16	−5.16	−4.25	−3.85	4.96	15.00
	Max|tb|	−4.39	−3.56	−3.16	−4.74	−3.82	−3.40	5.64	49.46
Guangxi	*MinLM*	−4.32	−3.57	−3.17	−5.09	−4.23	−3.84	5.32	21.70
	Max|tb|	−4.32	−3.57	−3.17	−4.71	−3.85	−3.42	4.92	49.72
Sichuan	*MinLM*	−4.23	−3.55	−3.17	−5.02	−4.18	−3.78	5.50	19.96
	Max|tb|	−4.23	−3.55	−3.17	−4.47	−3.70	−3.31	4.88	71.26
Guizhou	*MinLM*	−4.15	−3.45	−3.09	−5.12	−4.23	−3.79	4.90	5.52
	Max|tb|	−4.15	−3.45	−3.09	−4.76	−3.80	−3.31	4.52	7.22
Yunnan	*MinLM*	−4.47	−3.55	−3.20	−5.19	−4.31	−3.87	4.36	15.16
	Max|tb|	−4.47	−3.55	−3.20	−4.81	−3.87	−3.46	4.66	38.86
Shaanxi	*MinLM*	−4.34	−3.55	−3.17	−5.23	−4.31	−3.93	5.18	14.56
	Max|tb|	−4.34	−3.55	−3.17	−4.62	−3.76	−3.33	5.24	59.30
Gansu	*MinLM*	−4.28	−3.55	−3.16	−5.18	−4.27	−3.88	5.14	5.64
	Max|tb|	−4.28	−3.55	−3.16	−4.90	−3.81	−3.36	4.66	56.60
Qinghai	*MinLM*	−4.11	−3.25	−2.93	−5.11	−4.10	−3.69	5.08	17.46
	Max|tb|	−4.11	−3.25	−2.93	−4.58	−3.58	−3.15	5.24	14.44
Ningxia	*MinLM*	−4.19	−3.38	−3.02	−4.89	−4.09	−3.69	5.76	14.72
	Max|tb|	−4.19	−3.38	−3.02	−4.51	−3.61	−3.18	5.32	18.34
Xingjiang	*MinLM*	−4.28	−3.49	−3.12	−5.21	−4.29	−3.87	5.28	11.30
	Max|tb|	−4.28	−3.49	−3.12	−4.77	−3.87	−3.42	4.94	22.52
Panel LM	*MinLM*	−3.31	−2.48	−2.04	−6.93	−5.99	−5.55	5.90	
	Max|tb|	−3.31	−2.48	−2.04	−3.91	−3.09	−2.61	4.80	

[a] Empirical size is computed by allowing one break under the null but the break is neglected.
[b] True break dates are assumed to be the same as those estimated in Table 2.

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.

Chang, Y., 2004. Bootstrap unit root tests in panels with cross-sectional dependency. Journal of Econometrics 120, 263–293.

Engle, R., Granger, C.W.J., 1987. Cointegration and error-correction: representation, estimation, and testing. Econometrica 55, 251–276.

Gerdtham, U.-G., Jonsson, B., 2000. International cmparisons of health expenditure: theory, data and econometric analysis. In: Culyer, A.J., Newhouse, J.P. (Eds.), Handbook of Health Economics, vol. 1A. Elsevier, Amsterdam.

Gerdtham, U.-G., Sogaard, J., Jonsson, B., Andersson, F., 1992. A pooled cross-section analysis of the health care expenditures of the OECD countries. In: Zweifel, P., Frech III, H.E. (Eds.), Health Economics Worldwide. Kluwer Academic Publishers, Dordrecht.

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., Tieslau, M., 2005. Panel LM unit root tests with level shifts. Oxford Bulletin of Economics and Statistics 67, 393–419.

Im, K.S., Pesaran, M.H., Shin, Y., 2003. Testing for unit roots in heterogeneous panels. Journal of Econometrics 115, 53–74.

Jewell, T., Lee, J., Tieslau, Strazicich, M.C., 2003. Stationarity of health expenditures and GDP: evidence from panel unit root tests with heterogeneous structural breaks. Journal of Health Economics 22, 313–323.

Johansen, S., 1988. Statistical analysis of cointegration vectors. Journal of Economic Dynamics and Control 12, 154–231.

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 59, 1551–1580.

Kao, C., 1999. Spurious regression and residual-based tests for cointegration in panel data. Journal of Econometrics 90, 1–44.

Kao, C., Chiang, M., 2000. On the estimation and inference of a cointegrated regression in panel data. In: Badi, H., Baltagi (Eds.), Nonstationary Panels, Panel Cointegration, and Dynamic Panels. Advances in Econometrics, vol. 15. Elsevier Science, JAI, Amsterdam, pp. 179–222.

Lee, J., Strazicich, M.C., 2003. Minimum LM unit root test with two structural breaks. Review of Economics and Statistics 85, 1082–1089.

Lee, J., Strazicich, M.C. (2004). Minimum LM unit root test with one structural break. Working Papers 04-17, Department of Economics, Appalachian State University.

Leu, R.E., 1986. The public-private mix and international health care costs. In: Culyer, A.J., Jonsson, B. (Eds.), Public and Private Health Services: Complementarities and Conflicts. Basil Blackwell Ltd., Oxford, pp. 41–63.

Maddala, G.S., Wu, S., 1999. A comparative study of unit root tests with panel data and a new simple test. Oxford Bulletin of Economics and Sstatistics, Special Issue, 365–386.

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

Nunes, L.C., (2004). LM-type tests for a unit root allowing for a break in trend. Working Paper, Universidade Nova de Lisboa.

Pedroni, P., 1999. Critical values for cointegration tests in heterogeneous panels with multiple regressors. Oxford Bulletin of Economics and Statistics 61, 653–670.

Pedroni, P., 2004. Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. Econometric Theory 20, 597–625.

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

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

World Bank, 1984. China: The Health Sector. A World Bank Country Study, Washington, DC.

World Bank, 1992. China: Long-Term Issues and Options in the Health Transition. A World Bank Country Study, Washington, DC.

World Bank, 1997. Financing Health Care: Issues and Options for China. A World Bank Country Study, Washington, DC.

Zivot, E., Andrews, D.W.K., 1992. Further evidence of the great crash, the oil price shock and the unit root hypothesis. Journal of Business and Economic Statistics 10, 251–270.