THE CAUSALITY BETWEEN HEALTH EXPENDITURE AND ECONOMIC GROWTH IN IRAN

Mohsen Mehrara, Faculty of Economics, University of Tehran, Tehran, Iran
Maysam Musai, Faculty of Social Sciences, University of Tehran

Abstract
This paper examines causal relationships between health expenditure and GDP for Iran using annual data over the period 1970-2008. The Gregory-Hansen (1996) cointegration technique, allowing for the presence of potential structural breaks in data, is applied to empirically examine the long-run co-movement between health spending and output. The results suggest that there is a long-run relationship between these variables and the income elasticity for health care spending is greater than one during the period after Islamic Revolution (1979-2008). The Granger Causality test indicates a strong unidirectional effect from GDP to health expenditure, although providing no support to the view that health expenditure promotes long-term economic growth. The evidence for Iran clearly supports the ‘Income View’ over the ‘Health View’.

Keywords: non-linear Causality, Government Expenditure, Revenues, Iran Economy

1. Introduction
Theoretical literature suggests that there could be a two way relationship between health and income. The effect of economic growth on health is well known. By definition, health expenditure is a function of income or resources available both in private and public sectors. Higher income implies that there is more money to spend on health. A large body of research within health economics indicate that variations in per capita health care expenditure could be mostly explained by variations in per capita GDP (Gerdtham and Jonsson, 2000). On the other hand, a reverse causation from health expenditure to income has also a theoretical basis. Health is a capital and hence investment on health is an important source for economic growth. The report of the WHO’s Commission on Macroeconomics and Health (2001), states that “extending the coverage of crucial health services…to the world’s poor could save millions of lives each year, reduce poverty, spur economic development and promote global security.” (World Health Organization, 2001). Theoretically, health is a determinant of human capital, and labor productivity. So, regarding health expenditure as an investment in human capital and accordingly the engine of growth, an increase in health expenditure is expected to lead to higher income. In addition, rises in health expenditure possibly increase labor supply and productivity, which eventually must lead to a higher income (e.g., Muyssken, Yetkiner, and Ziesemer, 2003). Finally, there may be some intermediate variable, which causes both better health and higher income. For example, more education increases health and income for households.

The relationship between health and economic growth has been empirically investigated intensely, although, the evidence is mixed. Moreover, most of empirical studies have focused on developed countries by using a panel data analysis (e.g., Roberts, 1999; Freeman, 2003; Gerdtham and Lothgren, 2000; Sen, 2005; Wang and Rettenmaier, 2007). Therefore, a country-specific study on developing countries such as Iran is relatively scarce. Roberts (1999) identifies three weaknesses in the estimation based on cross countries or panel data which have been used to explore the relationship between income and health. First, the use of cross sectional data imposes homogeneity on sample countries, whereas there are considerable differences between the way health care is financed and organized in different economies. The second weakness is the failure to take into account the dynamics in the relationship though the use of an appropriate lag structure. The third weakness is the difficulty of dealing with non-stationary variables. Moreover, none of the studies allow for endogenous structural break in data.
So, we account for all these weaknesses with modeling the relationship between two variables based on the time series data from an individual country, Iran.

The focus of the paper is, therefore, to examine the relationship between health expenditure and income in the case of Iran country for the period from 1975-2008. The direction of causality between these two variables is examined by utilizing a cointegration and error correction modeling framework, which provides a more comprehensive test of causality than the standard Granger causality test. We apply the Zivot and Andrews (1992) unit root tests and the Gregory and Hansen (1996) cointegration tests, allowing for the presence of potential structural breaks in data during the sample period. The paper is organized in four sections. Section 2 discusses the methodology, data and results. Section 3 concludes.

2. Methodology and Empirical Results

In this section we use the Granger causality to study the causal relationship between health expenditure and GDP in Iran. The macroeconomic variables used in the model are (logarithm of) real health expenditure (HE) and real GDP per capita (GDP). The data series are obtained from Central Bank of Iran (CBI). The data are annual from 1970-2008, reflecting data availability. Considering the short sample period, a bivariate model is used to empirically examine the long-run co-movement and the causal relationship between health expenditure and real GDP.

2.1. Zivot and Andrews Unit Root Test

Conventional tests for identifying the existence of unit roots in a data series include that of the Augmented Dickey Fuller (ADF) (1979, 1981) or Phillips-Perron (1988). So in the first step of the empirical analysis, the Phillips-Perron unit-root tests have been carried out for the both variables: real health expenditure per capita, and GDP per capita, both in logarithm. The results reported in Table 1, indicate that both of the variables are nonstationary. However, recent contributions to the literature suggest that such tests may incorrectly indicate the existence of a unit root, when in actual fact the series is stationary around a one-time structural break (Zivot and Andrews, 1992; Pahlavani, et al, 2006). Zivot and Andrews (ZA) (1992) argue that the results of the conventional unit root tests may be reversed by endogenously determining the time of structural breaks. The null hypothesis in the Zivot and Andrews test is a unit root without any exogenous structural change. The alternative hypothesis is a stationary process that allows for a one-time unknown break in intercept and/or slope. Following Zivot and Andrews, we test for a unit root against the alternative of trend stationary process with a structural break both in slope and intercept. Table 1 provides the results. As in the Phillips-Perron case, the estimation results fail to reject the null hypothesis of a unit root for both variables. The same unit root tests have been applied to the first difference of the variables and in all cases we rejected the null hypothesis of unit root. Hence, we maintain the null hypothesis that each variable is integrated of order one or I(1).

Table 1: Unit-root tests of Phillips-Perron(PP) and Zivot and Andrews (ZA)

<table>
<thead>
<tr>
<th>Health Expenditure per capita(HE)</th>
<th>GDP per capita</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP</td>
<td>ZA</td>
</tr>
<tr>
<td>-1.43</td>
<td>-2.16(1979)</td>
</tr>
</tbody>
</table>

Note: The break point in ZA unit root test is presented in brackets. Empirical results indicate that the null
hypothesis of unit-root is rejected in all cases. The lag lengths for the ZA and PP tests are chosen by using SC’s information criterion and Newey and West (1987) method respectively. Critical values for ZA tests were obtained from Zivot and Andrews (1992). Break points are reported in ( )

2.2 The Gregory-Hansen Cointegration Analysis
As noted by Perron(1989), ignoring the issue of potential structural breaks can render invalid the statistical results not only of unit root tests but also of cointegration tests. Kunitomo (1996) argues that in the presence of a structural change, traditional cointegration tests, which do not allow for this, may produce spurious cointegration. Therefore one has to be aware of the potential effects of structural effects on the results a cointegration test, as they usually occur because of major policy changes or external shocks in the economy.

The Gregory-Hansen approach (1996) (hereafter, GH) addressed the problem of estimating cointegration relationships in the presence of a potential structural break by introducing a residual-based technique so as to test the null hypothesis (no cointegration) against the alternative of cointegration in the presence of the break (such as a regime shift). In this approach the break point is unknown, and is determined by finding the minimum values for the ADF t-statistic.

By taking into account the existence of a potential unknown and endogenously determined one-time break in the system, GH introduced three alternative models. The first model includes an intercept or constant (C) and a level shift dummy. The second alternative model (C/T) contains an intercept and trend with a level shift dummy. The third model is the full break model (C/S), which includes two dummy variables, one for the intercept and one for the slope, without including trend in model. This model allows for change in both the intercept and slope.

These tests detect the stability of cointegrating vectors over time in the presence of structural breaks in the form of level shift, level shift with trend, and regime shift. Table 2 reports all cases. when dependent variable is health expenditure, the null hypothesis of no cointegration relationships is rejected in favor of the existence of one cointegrating relationship, allowing for a one time structural break (although not rejected when GDP is dependent variable). The results show that the variables under examination do not drift apart for Iran. The estimated long run relationship using the C/S is of the form:

\[ HE = -2.25 + 1.12 GDP - 1.3D + 0.81D(GDP) \]

\[ t \quad (3.56) \quad (3.25) \quad (2.51) \quad (4.12) \]

where dummy \( D = 0 \) if \( t \leq 1979 \) and \( DU, = 1 \) if \( t > 1979 \). Both the intercept and the intercept at the time of regime shift (Islamic Revolution in Iran) are significant. Both the income elasticity before the regime shift and at the time of regime shift are significant. The income elasticity before the regime shift is 1.12 that is closer to unity. The change of income elasticity at the time of regime shift is 0.81. The slope of the function is interpreted as the change in the growth of the series. Therefore, we can see that income elasticity increased after regime shift and took a different path, implying more attention paid to health expenditure after the revolution. So, the elasticity of demand for health care in Iran, at least after the revolution amounting to 1.93 is greater than 1, implying that people and government spend progressively more health when income increase.
Table 2: Gregory-Hansen cointegration tests

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Model</th>
<th>Test Statistic</th>
<th>Break Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>HE</td>
<td>C</td>
<td>-6.30*</td>
<td>1978</td>
</tr>
<tr>
<td></td>
<td>C/T</td>
<td>-5.89*</td>
<td>1979</td>
</tr>
<tr>
<td></td>
<td>C/S</td>
<td>-6.92*</td>
<td>1979</td>
</tr>
<tr>
<td>GDP</td>
<td>C</td>
<td>-2.11</td>
<td>1991</td>
</tr>
<tr>
<td></td>
<td>C/T</td>
<td>-1.70</td>
<td>1974</td>
</tr>
<tr>
<td></td>
<td>C/S</td>
<td>-2.08</td>
<td>1991</td>
</tr>
</tbody>
</table>

Notes: C denotes level shift, C/T denotes level shift with trend, and C/S denotes regime shift. The lag length is chosen based on minimum SC. * denotes significant at the 5% level. Critical values were obtained from Gregory and Hansen (1996).

2.3. Granger Causality Tests

The existence of cointegrating relationship between HE and GDP for Iran suggests that there must be long run Granger causality in at least one direction (Hatanaka, 1996). In this section, we test for Granger Causality between log of real health expenditure per capita (HE) and log of real GDG per capita(GDP). Cointegration implies that causality exists between the two series but it does not indicate the direction of the causal relationship. The dynamic Granger causality can be captured from the vector error correction model (VECM) derived from the long-run cointegrating relationship (Granger 1988). Engle and Granger (1987) showed that if the two series are cointegrated, the vector-error correction model for the HE and GDP series can be written as follows:

\[
\Delta HE = \alpha_y + \beta_R ECT_{t-1} + \sum_{i=1}^{n} \gamma_{Ri} \Delta GDP_{t-i} + \sum_{i=1}^{n} \delta_{Ri} \Delta HE_{t-i} + \epsilon_{Rt} \]  
\[
\Delta GDP = \alpha_e + \beta_E ECT_{t-1} + \sum_{i=1}^{n} \gamma_{Ei} \Delta GDP_{t-i} + \sum_{i=1}^{n} \delta_{Ei} \Delta HE_{t-i} + \epsilon_{Ei} \]  
\[
ECT = HE + 2.25 -1.12GDG + 1.3D10.81D(GDP) \]  

where \( \Delta \) is a difference operator; ECT is the lagged error-correction term derived from the long-run cointegrating relationship; The \( \beta_i \) \((i = y, e)\) are adjustment coefficients; a is long run coefficient or elasticity and the \( \epsilon_{y,e} \) are disturbance terms assumed to be uncorrelated and random with mean zero.

Sources of causation can be identified by testing for significance of the coefficients on the lagged variables in Eqs. (1) and (2). First, by testing \( H_0 : \gamma_{Ri} = 0 \) for all i in Eq. (1) or \( H_0 : \delta_{Ei} = 0 \) for all i in Eq. (2), we evaluate Granger weak causality. This can be implemented using a standard F-test. Masih and Masih (1996) and Asafu-Adjaye (2000) interpreted the weak Granger causality
as ‘short run’ causality in the sense that the dependent variable responds only to short-term shocks to the stochastic environment.

Another possible source of causation is the ECT in Eqs. (1) and (2). In other words, through the ECT, an error correction model offers an alternative test of causality (or weak exogeneity of the dependent variable). The coefficients on the ECTs represent how fast deviations from the long run equilibrium are eliminated following changes in each variable. If, for example, $\beta_R$ is zero, then HE does not respond to a deviation from the long run equilibrium in the previous period. Indeed $\beta_R = 0$ or $\beta_E = 0$ is equivalent to both the Granger non-causality in the long run and the weak exogeneity (Hatanaka, 1996). This can be tested using a simple t-test.

It is also desirable to check whether the two sources of causation are jointly significant, in order to test Granger causality. This can be done by testing the joint hypotheses $H_0 : \beta_R = 0$ and $\gamma_{Ri} = 0$ for all i in Eq. (1) or $H_0 : \beta_E = 0$ and $\delta_{Ei} = 0$ for all i in Eq.(2). This is referred to as a strong Granger causality test. The joint test indicates which variable(s) bear the burden of short run adjustment to re-establish long run equilibrium, following a shock to the system (Asafu-Adjaye, 2000). A test of these restrictions can be done using F-tests.

Another concept related to Granger-causality is that of instantaneous causality. Roughly speaking, a variable HE is said to be instantaneously causal for another time series variable GDP if knowing the value of HE in the forecast period helps to improve the forecasts of GDP. It turns out, however, that, in a bivariate VAR process, this concept reduces to a property of the model residuals. More precisely, let $\varepsilon_I = (\varepsilon_{RI}, \varepsilon_{EI})$ be the residual vector of $y_I = (\Delta HE, \Delta GDP)$; then, $\Delta HE$ is not instantaneously causal for $\Delta GDP$ if and only if $\varepsilon_{EI}$ and $u \varepsilon_{RI}$ are uncorrelated. In turn, $\Delta HE$ is instantaneously causal for $\Delta GDP$ if and only if $\varepsilon_{RI}$ and $\varepsilon_{EI}$ are correlated. Consequently, the concept is fully symmetric. If $\Delta GDP$ is instantaneously causal for $\Delta HE$, then $\Delta HE$ is also instantaneously causal for $\Delta GDP$. Hence, the concept as such does not specify a causal direction. The causal direction must be known from other sources. Still, if it is known from other sources that there can only be a causal link between two variables in one direction, it may be useful to check this possibility by considering the correlation between the residuals (Lutkepohl, 2004).

The results of the tests on causality are presented in Table 3. It can be seen from the HE equation that GDP strongly Granger-causes HE. The coefficient of the ECT and lagged explanatory variables are significant in the health expenditure equation (HE) which indicates that long-run as well as short run causalities run from GDP to health expenditure. The adjustment coefficient in health equation (2) is estimated about -0.64, implying that, health expenditure adjusts at a reasonable speed to the long-run equilibrium, where almost two-third of the disequilibrium is corrected in the first period. Moreover, the interaction term (ECT and GDP) in the spending equation is significant at 5% level. The results for the other equation suggest that HE has no effect on GDP in short- and long-run. Therefore, there is unidirectional Granger causality running from GDP to health expenditure.
Table 3: Result of causality tests

<table>
<thead>
<tr>
<th>Source of causation</th>
<th>Short-run</th>
<th>Long-run</th>
<th>Joint(short-run/long-run)</th>
</tr>
</thead>
<tbody>
<tr>
<td>∆HE</td>
<td>∆GDP</td>
<td>ECT(-1)</td>
<td>ECT(-1)</td>
</tr>
<tr>
<td>HE does not cause GDP</td>
<td>1.50</td>
<td>-</td>
<td>1.74</td>
</tr>
<tr>
<td>p-value</td>
<td>(0.31)</td>
<td>(0.52)</td>
<td>(0.41)</td>
</tr>
<tr>
<td>GDP does not cause HE</td>
<td>-</td>
<td>4.12</td>
<td>-</td>
</tr>
<tr>
<td>p-value</td>
<td>(0.04)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
</tbody>
</table>

Notes: the lag length has been chosen based on minimum SC. ∆ denotes series in first difference.

Testing for instantaneous causality can be done by determining the absence of instantaneous residual correlation. Because only one correlation coefficient is tested to be zero, the number of degrees of freedom of the approximating chi-square distribution is one. Clearly, it is sufficient to report the test result for only one instantaneous causal direction because the test value for the other direction is identical given that it tests the very same correlation coefficient. The test statistics based on the residuals of the VECM is 9.56, being highly significant.

These results imply that, there is instantaneous as well as unidirectional Granger causality running from GDP to health expenditure, while health has a insignificant effect on GDP in both the short- and long-run. In other words, HE is strongly exogenous and whenever a shock occurs in the system, health expenditure must be reduced to maintain the long run relationship.

3-conclusion

This paper investigates the causal relationship between health expenditure and GDP in Iran, using the Gregory-Hansen (1996) cointegration and error correction modeling techniques based on annual data from 1970 to 2008. Our results support the existence of a long run relationship between GDP and health expenditure and the hypothesis that over the period after the revolution, health expenditure in Iran rose at a faster rate than GDP. The value of the long run income elasticity of health spending is much more than unity (1.93) during this period.

We also find strong support for the exogeneity of GDP. The main results in this paper confirm that there is an instantaneous and unidirectional causal link running from GDP to health spending. Yet, health spending does not Granger-cause per-capita GDP growth with a positive sign. The lack of strong link from health to economic growth is not necessarily a reason to reallocate health investment away from the health sector. The vital problem is to make health spending more effective in improving health outcomes, through improvements in accountability and incentives. The improvements in health status will be worth the effort even if they turn out to have little effect on growth.

Acknowledgements

This paper was funded by grant from the University of Tehran submitted to the author.
References