

U.S. State-Level Carbon Dioxide Emissions: Does It Affect Health Care Expenditure?

Nicholas Apergis^{*}, Rangan Gupta^{**}, Chi Keung Marco Lau^{***} and Zinnia Mukherjee^{****}

^{*} Corresponding author. Department of Banking and Financial Management, University of Piraeus, Piraeus, Greece. Email: napergis@unipi.gr.

^{**} Department of Economics, University of Pretoria, Pretoria, South Africa. Email: rangan.gupta@up.ac.za

^{***} Department of Accountancy, Finance, and Economics, University of Huddersfield, UK. Email: c.lau@hud.ac.uk.

^{****} Department of Economics, Simmons College, Boston, USA. Email: zinnia.mukherjee@simmons.edu.

U.S. State-Level Carbon Dioxide Emissions: Does It Affects Health Care Expenditure?

ABSTRACT

This paper is the first to provide an empirical analysis of the short run and long run effects of carbon dioxide (CO₂) emissions on health care spending across U.S. states. Accounting for the possibility of non-linearity in the data and the relationship among the variables, the analysis estimated various statistical models to demonstrate that CO₂ emissions led to increases in health care expenditures across U.S states between 1966 and 2009. Using quantile regressions, the analysis displayed that the effect of CO₂ emissions was stronger at the upper-end of the conditional distribution of health care expenditures. Results indicate the effect of CO₂ emissions on health care was relatively stronger for states that spend higher amounts in health care expenditures. The primary policy message of the paper is that there can be tangible health related benefits associated with policies that aim to reduce carbon emissions across U.S. states.

Keywords: health care expenditure, carbon dioxide emissions, panel cointegration, panel quantile regression

JEL Classification: I18, C31, C33

1. Introduction

The relationship between environmental quality and healthcare has long been an area of interest among scholars. Studies from medical science provide evidence that air pollution affects all types of mortality. An early study, Ostro and Rothschild (1989) use Health Interview Surveys to find that the association with small particulate matter can lead to work loss and even bed disability in adults. Schwartz and Dockery (1992) use data over the period 1973–1980 for Philadelphia air pollutants, such as total suspended particulate (TSP) and sulfur dioxide increased daily mortality rates. Spix and Wichmann (1996) show that in Koln, sulfur dioxide leads to 3-4% increase in mortality and particulates to a 2% increase in mortality. Wordley et al. (1997) find that in the U.K., ambient outdoor concentrations of particulate matter (PM10) significantly affect numerous health indicators. While there has been several scholarly attempts to find scientific evidence on the relationship between environmental quality and health care, identifying both the short term and long-run effects of pollutants on health is often very challenging for various reasons. For example, given the lack of effective monitoring systems, the levels of exposure to pollutants are often unknown. Also, the length of exposure to air pollutants, multiple exposures to different pollutants, and the cumulative effects of exposures all pose difficulties in fully understanding the impact of each pollutant on human health (Briggs, 2003).

From an economic perspective, a key issue of both academic and policy interest is the potential spatial and temporal effects of different environmental quality indicators on healthcare expenditures. Economists have long been interested in identifying the factors that determine an individual's healthcare expenditures. Early studies, such as Abel-Smith (1967), show that income is a key driver of healthcare spending. Murthy and Ukpolo (1995) find evidence that U.S. per capita health expenditure is cointegrated with its determinants. Using data from 1960–1987, they identify key explanatory factors such as per capita income, health

services and Medicare prices, age, and practicing physicians that explain the variation in health care spending across the U.S. population. Focusing on Canada, Matteo and Matteo (1998) find that both income and age have a positive effect on per capita provincial healthcare expenditure.

While a relatively large body of literature exists on the determinants of healthcare expenditure, the empirical literature on the relationship between environmental quality indicators and health care expenditure is still limited in spite of the associated economic and social implications. For example, the externalities generated by air pollution can have potentially negative consequences for labor productivity, which has direct implications for industrial performances and economic growth. The nature and extent of these effects essentially depend on the pollutant. For example, using data from Oslo, Hansen and Selte (2000) found that an increase in small particulate matter (PM) leads to a rise in sick leaves, which negatively affects output and trade in the city of Oslo. However, they also find that the effects of nitrogen dioxide and sulfur dioxide on sick leaves of employees are rather ambiguous. Jerrett et al. (2003) make use of data for 49 counties in Ontario and a sequential two-stage regression model to find that counties with higher pollution tend to experience higher health expenses, while counties that spend more on protecting environmental quality have lower expenditures on health care. Jacobson (2008) is the first paper that specifically isolated the human health effect of carbon dioxide from other global warming agents. The paper used a high-resolution model to show that chemical and meteorological changes due to carbon dioxide may increase U.S. annual air pollution deaths anywhere between 350 – 1800 as increases in fossil-fuel CO₂ increases can lead to increase in U.S. surface ozone, carcinogens, and particulate matter, which in turn can increase hospitalization and death rates. Allen et al (2015) simulated indoor environmental quality conditions in “green” and

“conventional” buildings. They found office workers who worked in environments with lower exposure to carbon dioxide had significantly higher cognitive functioning score.

In a 2008 publication, Narayan and Narayan examined the role of environmental quality in explaining per capita health expenditure for a sample of OECD countries. They adopted a panel cointegration approach to obtain estimates of short-run and long-run effects of environmental quality on health care expenditure for eight OECD countries. Their results indicate that carbon monoxide emissions, sulfur oxide emissions, per capita income and per capita health expenditure are panel cointegrated. Interestingly, they find that in the short-run, both carbon monoxide emissions and income and carbon monoxide emissions have a statistically significant positive effect on health expenditure. Income has an elastic and positive effect in the long run while carbon monoxide and sulfur oxide have an inelastic and positive impact on health expenditure. Assadzahed et al. (2014) focus specifically on the linkage between carbon dioxide emissions and health care expenses. They use a panel dataset for eight oil exporting countries that cover the years 2000 through 2010. Their findings reveal that short-run elasticities for carbon dioxide and income are positive and statistically significant, while the impact of life expectancy on health expenditures turns out to be negative.

Given the serious implications of rise in atmospheric greenhouse gas (GHG) concentrations on life on earth, an important area of research involves identifying the linkages between greenhouse gases emissions and expenditures on health care. Carbon dioxide (CO₂), water vapor, methane (CH₄), nitrous oxide (N₂O), hydrofluorocarbons and chlorofluorocarbons are among the most abundantly found greenhouse gases in the earth’s atmosphere (NOAA website, *retrieved October 10, 2017*). Larger emissions of greenhouse gases lead to higher concentrations of pollutants in the atmosphere, which lead to excess heat trapped on the earth’s surface that would have otherwise radiated back to the atmosphere. In

this paper, we focus on the relationship between carbon dioxide and health care expenditures. Carbon dioxide emissions that stem from various anthropogenic causes such as fossil fuel combustion, soil erosion, and deforestation is likely to play a critical role in defining current and long-term global environmental quality. Scientific evidence shows that carbon dioxide can stay trapped in the earth atmosphere for a very long time. The negative externalities that stem from carbon dioxide emissions have long last adverse effects on both economic growth and human welfare. This makes it critically important to develop our understanding of the role of carbon dioxide emissions in human health over time. Therefore, this paper contributes to the existing environmental and health economics literature by providing an empirical analysis of the impact of per capita CO₂ emissions on real per capita health care expenditure across all the 50 U.S. states. We control for a measure of output, *i.e.*, real per capita personal disposable income, given the widespread evidence of the latter being a strong predictor of health care expenditures (Freeman, 2003, 2012; Caporale et al., 2015).

The United States is the second largest emitter of carbon dioxide in the world behind China and ahead of the European Union (EU) and India (Union of Concerned Scientist website, *retrieved October 10, 2017*). Within the U.S., there is a considerable variation in CO₂ emissions across states. For example, in 2013, aggregate CO₂ emissions in Texas for all five sectors, *i.e.* commercial, industrial, residential, transportation, electric power, was 712.86 million metric tons, whereas for Vermont was 5.97 million metric tons (U.S. EPA website, *retrieved October 10, 2017*). There is also some variation in per capita health care spending across these states. For example, in 2009, the per capita health care spending in the District of Columbia (D.C.), Alaska, and Massachusetts were \$10348.85, \$9127.63, and \$9277.89, respectively, indicating the highest spending per capita across all U.S. states. In comparison, per capita health care spending for Utah, Georgia and Idaho were \$5030.94, \$5467.46, and

\$5657.99, respectively, three states with the lowest per capita spending in the country (Kaiser Family Foundation website, *retrieved October 10, 2017*).

The novelties of this paper are twofold. First, it is the first paper to provide an empirical analysis of the short- and long-run effects of CO₂ emissions of healthcare spending across U.S. states using a panel dataset. The results can be useful in the context of designing and evaluating U.S. health care and environmental policies, particularly, those that account for cross-state variation. Second, the paper makes a methodological contribution as well. To account for the possibility of non-linearity in the data of the individual variables as well as in the relationship amongst the variables, we estimate various conditional mean-based statistical models. We also conduct quantile regressions to account for the variability of the results across the US states, conditioned on their level of health care expenditures.

The remainder of the paper is organized as follows: In Section 2, the data set and the econometrics methodologies are detailed. Section 3 discusses the empirical findings. Finally, Section 4 provides discusses some policy implications and provides a few concluding remarks.

2. Data and Method

For the empirical analysis, the study makes use of annual data on healthcare expenditures for all 50 US states for the period 1966 to 2009. We obtained the data on health care spending by state of residence from the Center for Medicare and Medicaid Services Health Expenditure. The data on nominal personal disposable income were obtained from the regional database of the U.S. Bureau of Economic Analysis (BEA). Per capita values were obtained for both variables by dividing the data with the population data, which was also obtained from the BEA regional database. We obtained the data on per capita carbon dioxide from the Carbon Dioxide Information Analysis Center. The data was obtained from sectoral accounting, which

reflects human/economic activity during a certain period of time. The CO₂ data are measured in thousand metric tons of carbon.

Because the state level consumer price index (CPI) levels are not available for the sample period, we converted the per capita nominal personal disposable income and the nominal per capita health care expenditure data into their real values by deflating with the aggregate US CPI. Freeman (2012) is our source for the data on real per capita health expenditures (H) and personal disposable income (INCOME) for the 50 US states. The data on income inequality (Gini coefficient) is obtained from Mark W. Frank's website.¹ The paper provides a complete description of the dataset. For the analysis, the data are transformed into their natural logarithmic values.²

As is standard practice in panel data econometrics (N=50) with a long time series component (T=44), we start off by conducting unit root testing on the data. Given the evidence of non-linearity in the three variables of interest (i.e. CO₂ emissions per capita, health expenditure per capita, and per capita gross domestic product (GDP)), as mentioned in the research of Li et al. (2014), Caporale et al. (2015) and Zerihun et al. (2017), the analysis uses non-linear unit root tests along with standard linear versions. In case that unit root tests support non-stationarity, the analysis moves ahead with cointegration tests to analyze the relationship between health care expenditures and CO₂ emissions, after controlling for income.

2.1. Non-linear panel unit roots

Following Cerrato et al. (2011, 2013), the variable of interest y_{it} , is modeled through the following Exponential Smooth Transition Autoregressive model:

¹ Source: http://www.shsu.edu/eco_mwf/inequality.html

² The dataset is available upon request.

$$y_{it} = \xi_i y_{i,t-1} + \xi_i^* y_{i,t-1} Z(\theta_i; y_{i,t-d}) + \mu_{it} \quad t=1, \dots, T \quad i=1, \dots, N, \quad (1)$$

Equation (1) in its first-difference form yields:

$$\Delta y_{i,t} = \alpha_i + \xi_i y_{i,t-1} + \sum_{h=1}^{h-1} \delta_{ijh} \Delta y_{ij,t-h} + (\alpha_i^* + \xi_i^* y_{i,t-1} + \sum_{h=1}^{h-1} \delta_{ih}^* \Delta y_{i,t-h}) * Z(\theta_i; y_{i,t-d}) + \gamma_i f_t + \varepsilon_{it} \quad (2)$$

Cerrato *et al.* (2011) use a first-order Taylor series approximation methodology that re-parameterizes Equation (2) and the auxiliary regression yields:

$$\Delta y_{i,t} = a_i + \delta y_{i,t-1}^3 + \gamma_i f_t + \varepsilon_{i,t} \quad (3)$$

If the errors are serially correlated, then equation (3) can be extended as follows,

$$\Delta y_{i,t} = a_i + \delta y_{i,t-1}^3 + \sum_{h=1}^{h-1} \rho_{ih} \Delta y_{i,t-h} + \gamma_i f_t + \varepsilon_{i,t} \quad (4)$$

The non-linear cross-sectionally augmented DF (NCADF) regression in its final form can be written as:

$$\Delta \bar{y}_{i,t} = a_i + b_i \bar{y}_{i,t-1}^3 + c_i \Delta \bar{y}_t + d_i \Delta \bar{y}_{t-1} + \varepsilon_{i,t} \quad (5)$$

t-statistics could be derived from \hat{b}_i , which are denoted by:

$$t_{iNL}(N, T) = \frac{\hat{b}_i}{s.e.(\hat{b}_i)} \quad (6)$$

where \hat{b}_i is the OLS estimate of b_i , and $s.e.(\hat{b}_i)$ is its associated standard error. The t -statistic in Equation (10) can be used to construct a panel unit root test by averaging the individual test statistics:

$$\bar{t}_{iNL}(N, T) = \frac{1}{N} \sum_{i=1}^N t_{iNL}(N, T) \quad (7)$$

This is a non-linear cross-sectionally augmented version of the IPS test (NCIPS). The Pesaran (2007) test (CIPS) takes into account the cross sectional dependence among panel members.

Note our model is a model version of the Environmental Kuznets Curve, which captures the relationship between environmental degradation and the income level of an economy. The following figure 1 shows the relationship illustrated by the EKC, which shows that during early stages of development environmental degradation rises (i.e., environmental quality falls) up to a certain level of income and then starts rising with further increases in income.

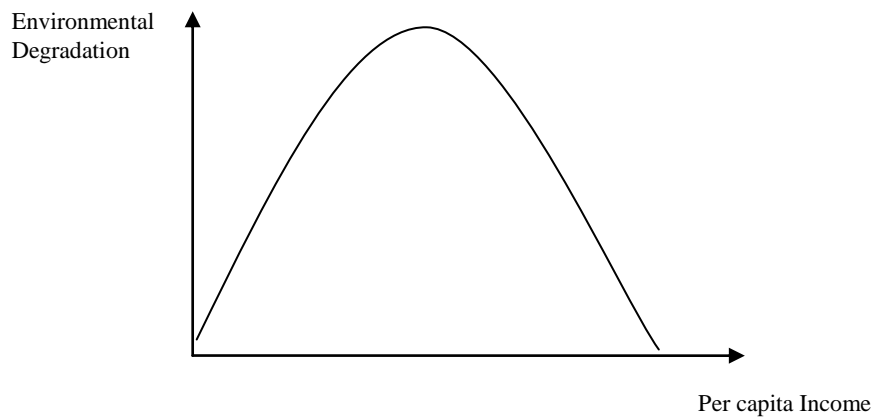


Fig 1: The Environmental Kuznets Curve

The panel cointegration results indicate the presence of a long-run relationship between health care expenditure, personal disposable income per capita, and CO₂ emissions. Thus, we estimate the long-run coefficients through the following model:

$$H_{it} = \alpha_{i+} \beta_i INCOME_{it} + \gamma_i CO2_{it} + e_{it} \quad (8)$$

$$e_{it} = \lambda_i' F_t + \mu_{it}, \quad (9)$$

where, the subscript 'it' denotes the observation on the *i*'th state at time *t*, for $i=1,2,\dots,N$ and $t=1,2,\dots,T$. The dependent variable H_{it} denotes real health care expenditures per capita, while $INCOME_{it}$ is real personal disposable income per capita; $CO2_{it}$ is per capita CO₂ emissions. All the variables are represented in their natural logarithmic forms. F_t is the $m \times 1$

vector of unobserved common factors. Three conventional estimation methods are used to estimate the long-run relationship, namely, the Mean Group (MG) estimation method discussed in Pesaran and Smith (1995), the Group Mean Fully Modified OLS (GM-FMOLS) estimation method (Pedroni, 2000, 2001) and the Group Mean Dynamic OLS (GMDOLS) estimation technique (Pedroni, 2001).

2.2. Autoregressive Distributed Lag (ARDL) model

This sub-section uses the autoregressive distributed lag (ARDL) model, proposed by Pesaran et al. (2001), to examine the relationship between real income per capita, real health expenditures, and per capita CO₂ across the 50 US states. It is worth noting that we pursue this approach to check out for the robustness of the above results. While the unit root tests indicate that the entire panel of the three variables are non-stationary, there are cases (i.e., states) where the variables are found to be stationary (similarly to the results in Freeman, 2012; Caporale et al., 2015). The advantage of the ARDL methodological approach is that it does not require pre-testing of unit roots, and hence, is a more general approach, while it accommodates for any possible issues of endogeneity that could be present. The ARDL representation of the effects of real income per capita and per capita CO₂ on real health expenditures can be described as follows:

$$\Delta H_{it} = \lambda_0 + \sum_{i=1}^m \Delta \lambda_{1i} \Delta H_{it-1} + \sum_{i=1}^m \Delta \lambda_{2i} \Delta INCOME_{it-1} + \sum_{i=1}^m \Delta \lambda_{3i} \Delta CO2_{it-1} + \lambda_4 H_{it-1} + \lambda_5 INCOME_{it-1} + \lambda_6 CO2_{it-1} + v_{it} \quad (10)$$

where m denotes the lag order. The error term v_t is assumed to be an independent and identically distributed i.i.d. process with a finite second moment. We can transform equation (10) can be transformed into the following Error Correction model:

$$\Delta H_{it} = \sigma_{0+} + \sum_{i=1}^k \Delta \sigma_{1i} \Delta H_{it-1} + \sum_{i=1}^k \Delta \sigma_{2i} \Delta INCOME_{it-1} + \sum_{i=1}^k \Delta \sigma_{3i} \Delta CO2_{it-1} + \xi(H_{it-1} - \beta_1 INCOME_{it-1} + \beta_2 CO2_{it-1}) + \mu_{it} \quad (11)$$

where, ξ is the speed of the adjustment parameter, and β_1, β_2 are the long-run coefficients for real income and CO₂ per capita, respectively. The short-run parameters are represented by σ_{1i}, σ_{2i} , and σ_{3i} . The ARDL (p, q, k) model is given by:

$$\Delta H_{it} = \sigma_{0+} + \sum_{i=1}^p \Delta \sigma_{1i} \Delta H_{it-1} + \sum_{i=1}^q \Delta \sigma_{2i} \Delta INCOME_{it-1} + \sum_{i=1}^k \Delta \sigma_{3i} \Delta CO2_{it-1} + \xi(H_{it-1} - \beta_1 INCOME_{it-1} + \beta_2 CO2_{it-1}) + \mu_{it} \quad (12)$$

2.3. Panel quantile regressions (PDQ)

In the relevant literature, OLS estimations have been used extensively to consider the determinants of health expenditures (Freeman, 2012). However, there are reasons to believe that the influence of income and per capita CO₂ emissions is likely to differ across states. We conduct the BDS test, with the statistics rejecting linearity in the majority of states across all three variables, indicating that some type of hidden structure is contained in the series. This is reinforced by the linearity test recommended by Tsay (1996). We have not reported the results here to save space, but available upon request. Therefore, the panel quantile regression (PQR) methodology, in relevance to Equation (16), is pursued; it accounts for the likelihood of heterogeneity as it estimates the parameters of the model at different points on the (conditional) per capita health expenditure distribution. The non-linear nature of PQR has a distinct advantage in terms of the parameters that can be estimated. First, we can obtain estimates of the different parameters of the logarithms of income and CO₂ for under-expenditures, which reflect the lower end regions of the conditional distribution of per capita health expenditure. Also, the parameters of the over-expenditures (regions at the upper end) can be estimated. Further advantages of the PQR methodology include the non-sensitivity of

estimated coefficients relating to observations of the dependent variable of health expenditures that are outliers. Also, the estimators are more efficient than those provided by OLS when the error term is not normally distributed.

Since the mean regression methodologies fail to take into account the effects that can be potentially heterogeneous in nature, we specify the τ -th quantile ($0 < \tau < 1$) of the conditional distribution of the dependent variable (i.e., the log of per capita health expenditures), given a set of independent variables X_{it} , as follows:

$$Q_{\tau} \left(\frac{LnH_{it}}{X_{it}} \right) = \alpha_{\tau} + \beta_{\tau} X_{it} + \alpha_{\tau} \mu_{it} \quad (13)$$

where LnH_{it} is per capita health expenditure in a log form of state i at time t , and X_{it} represents a vector of two independent variables, i.e. income in a log form (i.e., $\ln INCOME_{it}$) and per capita CO_2 emissions, also in its log form (i.e., $\ln CO_{2it}$). u_{it} denotes unobservable factors, such as institutional and socio-demographic factors (e.g., unmeasured disease severity or other health limitations). To estimate the parameters in equation (13), the following objective function is used to minimize the absolute value of the residual:

$$\begin{aligned} Q_{\tau}(\beta_{\tau}) &= \min_{\beta} \sum_{i=1}^n [|LnH_{it} - \beta_{\tau} X_{it}|] \\ &= \min_{\beta} [\sum_{i: LnH_{it} \geq \beta X_{it}} \tau |LnH_{it} - \beta_{\tau} X_{it}| + \sum_{i: LnH_{it} < \beta X_{it}} (1 - \tau) |LnH_{it} - \beta_{\tau} X_{it}|], \quad (14) \end{aligned}$$

A limitation of the use of QR methods is revealed when the model includes a large number of fixed effects. A number of papers highlight the problem of capturing the unobserved factors with a fixed effects quantile model (Koenker, 2004; Canay 2011). When the estimation includes a large number of cross-sectional units along with a relatively small number of observations for each cross-sectional unit, the estimated parameters of the fixed effects are not consistent. Koenker (2004) proposes a class of penalized quantile regression estimators

(i.e., the shrinkage methodology) to address the above problem by estimating directly a vector of individual effects. However, Canay (2011) finds that the Koenker's methodology is computationally intensive and he subsequently introduces a two-step methodology of estimating panel quantile regression models with fixed effects. In the first stage, the conditional mean of u_{it} is estimated and then the analysis employs the estimated parameters to obtain the individual fixed effect $\hat{\alpha}_i = \frac{\sum_{t=1}^T (LnH_{it} - X'_{it}\hat{\beta}_\mu)}{T}$, where $\hat{\beta}_\mu$ denotes the estimated versions of the parameters from the conditional mean regression.

In the second stage, the analysis subtracts the estimated individual effect from the dependent variable, $\widehat{LnH}_{it} = LnH_{it} - \hat{\alpha}_i$ and then the standard estimation of the quantile regression is used. For parameters inference, Canay (2011) proposes a bootstrap procedure for estimating the variance-covariance matrix for this estimator. The bootstrap methodology is implemented by randomly drawn samples with the replacement of a sample of size NT from the original data and computing the two-step estimator, as described above for B times, resulting in a total of B different estimates. The estimated bootstrapped variance-covariance matrix at quantile τ is constructed as:

$$\frac{1}{B} \sum_{j=1}^B (\hat{\beta}_j^*(\tau) - \bar{\beta}^*(\tau))(\hat{\beta}_j^*(\tau) - \bar{\beta}^*(\tau))'$$

where $\hat{\beta}_j^*(\tau)$ are the estimated parameters from the j th bootstrap and the τ th quantile, whereas

$$\bar{\beta}^*(\tau) = \frac{1}{B} \sum_{j=1}^B \hat{\beta}_j^*(\tau).$$

3. Empirical Results

The results of NCIPS statistics are reported in Table 1 and they support the presence of a unit root across all three variables, while similar conclusions are reached through the CPIS

test (Panel A, Table 2). Because we find that all variables are integrated of order 1, we test the null hypothesis of no cointegration. Panel B in Table 2 reports the panel cointegration results. The first three cointegration tests are ADF-based and PP-based cointegration tests (Pedroni, 2000; Pedroni, 2004) and Kao (1999) ADF-based tests, which assume cross sectional independence, indicative of no cointegration. These tests are ADF-based and PP-based cointegration tests (Pedroni, 2000; Pedroni, 2004) and Kao (1999) ADF-based tests. All three tests suggest the rejection of the null of no cointegration. The tests suggest the rejection of the null of no cointegration. It is worth noting that the evidence of cointegration is stronger for the periods 1995 to 2009 and 1985 to 2009. An influential event could be the passing of the Consolidated Omnibus Budget Reconciliation Act (COBRA) of 1985, which amended the 1974 Employee Retirement Income Security Act (ERISA). The amendment enabled some employees to continue their health insurance coverage after leaving their employment. The health reform of the 'Health Security Express' started at the end of July 1994. It is shown to be influential as well.

[Insert Tables 1 and 2 about here]

We also use the error-correction-based panel cointegration tests incorporating cross sectional dependence (Westerlund, 2007) to conduct robustness checks. Westerlund develops four normally distributed tests, namely G_t , G_a , P_a , and P_t . The first two tests are mean-group tests, since they only assume unit-specific error correction parameters, while the rejection of the null hypothesis can be taken as evidence of cointegration of at least one of the cross-sectional units. The latter two test statistics pool information overall the cross-sectional units and, hence, the rejection of the null should be taken as evidence of cointegration for the panel as a whole.

As another check for robustness, we make use of two tests (i.e., τN and ΦN) for the null hypothesis of no cointegration, proposed by Westerlund and Edgerton (2008); they take into account the presence of structural breaks within the heterogeneous panel. τN and ΦN test are derived from the popular Lagrange multiplier (LM)-based unit-root tests (for details please refer to page 675 of Westerlund and Edgerton; 2008). The tests are taking into account of heteroskedastic and serially correlated errors, unit-specific time trends, cross-sectional dependence and unknown structural breaks in both the intercept and slope of the cointegrated regression, which are the common factual characteristics in panel data.

Panel A in Table 3 reports the break date for both level break and regime shift. The cointegration results in panel B document that the null of no cointegration is rejected at the 1% significance level in the no break model for both τN and ΦN tests.

[Insert Table 3 about here]

Moreover, the CD test of Pesaran (2004) confirms the presence of cross section dependence in the residuals for those three mean group methodologies (MG, GM-FMOLS, and GM-DOLS); the test results violate the assumption of cross-section independence (Table 4); therefore, we employ a novel general methodology which allows cross-sectional dependence and cross section-specific slope coefficients (i.e., CCE-MG methodology) proposed by Pesaran (2006) and Kapetanios et al. (2011). The CCE-MG estimator suggests that the coefficients are statistically significant at the 5% significant level.

[Insert Table 4 about here]

Table 5 presents the results for Equation (12); they illustrate that the error-correction coefficient is negative and statistically significant at the 1% significance level. Importantly, the long-run coefficients from the cointegrating equation are reported; a 10% increase in per

capita income results in a long-run increase of 8.61% in per capita health expenditures, while a 10% increase in per capita carbon emissions results in a long-run increase of 1.57% in per capita health expenditures. The ECM coefficient is -0.319 , implying that the adjustment speed is about 32%. We also apply the panel granger causality test of Dumitrescu and Hurlin (2012) confirms that CO_2 causes health care expenditure. The null hypothesis of CO_2 (in log form) does not causes health expenditure (in log form) is rejected with the probability of 0.0134, where the optimal lag length is 3 as chosen from the SC criteria. The null hypothesis of CO_2 growth does not causes health expenditure growth is also rejected with the probability of 0.0154, where the optimal lag length is 1 as chosen from the SC criteria.

[Insert Table 5 about here]

Table 6 reports the results from the quantile regression approach by Canay (2011). The findings display the coefficients for the per capita health expenditures variable between the 10th and 90th quantiles that was obtained from estimating equation (13). When using panel quantile regression coefficients on per capita CO_2 emissions, the estimates tend to be relatively high at higher quantiles (i.e., for over-expenditure regions; those per capita health expenditures are high, given the values of the explanatory variables). The influence of 1% increase in per capita CO_2 emissions on per capita health expenditures is only 0.13% at the 10th percentile, in comparison to 0.16% at the 90th percentile. However, the coefficients for the income variable are slightly smaller for the higher quantiles; the influence of 1% increase in income on per capita health expenditure is 0.61% at the 10th percentile in comparison to 0.59% at the 90th percentile.

[Insert Table 6 about here]

As a robustness check, we also ran the quantile regressions including a measure of income distribution as a covariate but it does not change the conclusion at all. The results are presented in table 7 in the appendix.

4. Conclusion and Discussion

The identification of the determinants of health care expenditures has important scholarly and policy implications. An issue that remains relatively less explored in this area of study is the potential relationship between environmental quality and health care spending. In this paper, we present estimates of the causal effect of carbon dioxide emissions on per capita health care expenditure across all U.S. states. Carbon dioxide is a primary greenhouse gas and is known to stay trapped for decades in the earth's atmosphere, which necessitates the need to understand its impact on human health. The U.S is the second largest emitter of CO₂ in the world. It exhibits a considerable variation in both CO₂ emissions and per capita health care expenditure across states. This paper contributed to the literature by providing for the first time a rigorous empirical analysis of the short run and long run effects of CO₂ emissions on health care spending across U.S. states. Recognizing both the non-linearity in the data and the relationships among the individual variables, we estimated different statistical models to obtain evidence that indicates carbon emissions over the sample period led to an increase in per capita health care expenses.

We found that the effect of CO₂ emissions on health care expenses was relatively stronger for states that spend higher amounts in health care expenditures. In general, the results from the panel quantile regressions show that the influence of 1% increase in per capita CO₂ emissions on per capita health expenditures is only 0.13% at the 10th percentile, in comparison to 0.16% at the 90th percentile. As an example, we observed that for the time period 1966 to 2009, the average health care expenditure for Alaska is \$1932 and the per capita CO₂ emissions is 14.83 metric tons. For Idaho, the figure for health care expenditure is

\$1363 while the per capita CO₂ emissions is 3.63 metric tons between 1996 and 2009. Note this does not imply that the CO₂ emissions must be reduced necessarily in states with high per capita health care expenditures at the risk of reducing health benefits. For example, in Alaska, where the temperature drops to critical levels, fossil fuel combustion might be critically important in generating heat that produces a comfortable living environment in such a cold climate. Hence, any significant reduction in CO₂ emissions by reducing fossil fuel combustion there might pose health risks for the state's population. Our results also do not indicate the CO₂ emissions must be reduced uniformly across all U.S. states. In contrast, the results of this paper emphasises two important issues – first, per capita CO₂ emissions lead to an increase in per capita health care expenses and the relationship varies across considerably across U.S. states, which implies the net benefit associated with reducing a unit of carbon dioxide emissions can vary significantly across the states.

Our findings have economic and social implications for individuals, households and society at large. The carbon dioxide content in an individual's indoor and outdoor environmental quality have impact on per capita health expenses. More aware individuals are about the effects of air quality on human health and health expenses, more enabled individuals will be able to take precaution to limit exposure to low quality ambient air. The primary policy message of the paper is that there can be tangible and varied levels of health related benefits associated with policies that aim to reduce carbon emissions across U.S. states. For example, the health care-associated savings, mostly from avoided hospital visits and reduced spending on pollution-related illnesses, which can stem from a carbon-reducing policy can be effective in capturing efficiency gains across the economy. However, to achieve significant reductions at the lowest cost to the economy will require strong, coordinated, political will and economy-wide actions.

Acknowledgement:

We would like to thank Donald G. Freeman, Sam Houston State University, for providing the dataset.

References

- Abel-Smith, B. 1967. *An International Study of Health Expenditure and its Relevance for Health Planning*, World Health Organization, Geneva.
- Allen, J.G., MacNaughton, P., Satish U., Santanam S., Vallarino J., Spengler J.D., 2016. Associations of cognitive function scores with carbon dioxide, ventilation, and volatile organic compound exposures in office workers: a controlled exposure study of green and conventional office environments. *Environmental Health Perspectives*, 124, 805-812.
- Assadzadeh, A., Bastan, F., Shahverdi, A., *The Impact of Environmental Quality and Pollution on Health Expenditures: A Case Study of Petroleum Exporting Countries*, Proceedings of 29th International Business Research Conference 24 - 25 November 2014, Novotel Hotel Sydney Central, Sydney, Australia, ISBN: 978-1-922069-64-1
- Briggs, D. 2003. Environmental pollution and the global burden of disease, *British Medical Bulletin*, 68 (1), 1-24.
- Canay, I. A. (2011). A simple approach to quantile regression for panel data. *The Econometrics Journal*, 14(3), 368-386.
- Caporale, G.M., Cunado, J., Gil-Alana, L.A, Gupta, R., (2015). The relationship between healthcare expenditure and disposable personal income in the US states: A fractional integration and cointegration analysis. CESIFO Working Paper No. 5407.
- Cerrato, M., de Peretti, C., Larsson, R., Sarantis, N., (2011). A nonlinear panel unit root test. Working Paper, Department of Economics, University of Glasgow.
- Cerrato, M., De Peretti, C., & Stewart, C. (2013). Is the consumption-income ratio stationary? Evidence from linear and non-linear panel unit root tests for OECD and non-OECD countries. *Manchester School*, 81(1), 102-120.
- Dumitrescu, E. I., & Hurlin, C. (2012). Testing for Granger non-causality in heterogeneous panels. *Economic Modelling*, 29(4), 1450-1460.
- Freeman, D.G., (2003). Is health care a necessity or a luxury? Pooled estimates of income elasticity from US state-level data, *Applied Economics*, 35:5, 495-502.
- Freeman, D.G., (2012). Is health care a necessity or a luxury? New evidence from a panel of US state-level data. Sam Houston State University, Economics and International Business Working Paper No. 12-03.

- Jacobson, M.Z. 2015. On the causal link between carbon dioxide and air pollution mortality. *Geophysical Research Letters*, 35(3), L03809, doi:10.1029/2007GL031101.
- Jerrett, M., Eyles, J., Dufournaud, C., Birch, S., 2003. Environmental influences on health care expenditures: an exploratory analysis from Ontario, Canada. *Journal of Epidemiology and Community Health* 57, 334–338.
- Hansen, A.C., Selte, H.K., 2000. Air pollution and sick-leaves: a case study using air pollution data from Oslo. *Environmental and Resource Economics* 16, 31–50.
- Kaiser Family Foundation website: <http://kff.org/other/state-indicator/health-spending-per-capita/#>
- Kao, C.D., (1999). Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics* 90, 1-44.
- Kapetanios, G., Pesaran M.H., Yagamata T., (2011) Panels with non-stationary multifactor error structures. *Journal of Econometrics* 160, 326-348.
- Karatzas, G., 2000. On the determinants of the US aggregate health care expenditure? *Applied Economics* 32, 1085–1099.
- Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91(1), 74-89.
- Li, X-L, Tang, D.P., and Chang, T. (2014). CO₂ emissions converge in the 50 U.S. states — Sequential panel selection method. *Economic Modelling*, 40(C), 320-333.
- Matteo, L.D., Matteo, R.D., 1998. Evidence on the determinants of Canadian provincial government health expenditures. *Journal of Health Economics* 17, 211–228.
- Murthy, N.R.V., Ukpolo, V., 1995. Aggregate health care expenditure in the United States: new results. *Applied Economics Letters* 2, 419–421.
- NOAA website: <https://www.ncdc.noaa.gov/monitoring-references/faq/greenhouse-gases.php>
- Ostro, B., Rothchild, S., 1989. Air pollution and acute respirator morbidity: an observational study of multiple pollutants. *Environmental Research*. vol. 50, 238–247.
- Pedroni, P., (2000). Fully modified OLS for heterogeneous cointegrated panels. In Baltagi, B. H. (Ed.) *Nonstationary Panels, Panel Cointegration and Dynamic Panels*, 15, Amsterdam: Elsevier, 93-130.
- Pedroni, P., (2001). Purchasing power parity tests in cointegrated panels. *The Review of Economics and Statistics* 83, 727-731.

- 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.
- Pesaran, M.H., Smith, R., (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics* 68, 79-113.
- Pesaran, M. H., Shin, Y. and Smith, R. J., (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, 16, 289–326.
- Pesaran, M.H. (2004). General Diagnostic Tests for Cross Section Dependence in Panels", CESifo Working Paper Series No. 1229; IZA Discussion Paper No. 1240. Available at SSRN: <http://ssrn.com/abstract=572504>
- Pesaran, MH., (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74, 967-1012.
- Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, 22(2), 265-312.
- Schwartz, J., Dockery, D.W., 1992a. Increased mortality in Philadelphia associated with daily air pollution concentrations. *American Review of Respirator Disease*. vol. 145, 600–604.
- Spix, C., Wichmann, H.E., 1996. Daily mortality and air pollutants: findings from Koln, Germany. *Journal of Epidemiology and Community Health*. vol. 50, 52–58.
- Tsay, R. S. (1986). Nonlinearity tests for time series. *Biometrika*, 73(2), 461-466.
- Union of Concerned Scientists website:
http://www.ucsusa.org/global_warming/science_and_impacts/science/each-countrys-share-of-co2.html#.VtNOMfkrLIU
- U.S. EPA - https://www3.epa.gov/statelocalclimate/resources/state_energyco2inv.html
- Westerlund, J. (2007). Testing for error correction in panel data*. *Oxford Bulletin of Economics and statistics*, 69(6), 709-748.
- Westerlund, J., Edgerton, D. L. (2008). A simple test for cointegration in dependent panels with structural breaks. *Oxford Bulletin of Economics and Statistics*, 70(5), 665-704.

- Wordly, J., Walters, S., Ayres, J.G., 1997. Short term variations in hospital admissions and mortality and particulate air pollution. *Occupational and Environmental Medicine*. vol. 54, 108–116.
- Zerihun, M. F., Cunado, J., & Gupta, R. (2017). Are Health Care Expenditures and Personal Disposable Income Characterised by Asymmetric Behaviour? Evidence from US State-Level Data. *Social Indicators Research*, 131(2), 527-542.

Appendix:**Table 1. Nonlinear panel unit root test results (NCIPS)**

States	lnH	lnCO2	lnINCOME
Alabama	0.2281	-0.6296	-0.4173
Alaska	-1.2728	-0.9185	-1.9742**
Arizona	-2.2031	-1.0420	-1.6987
Arkansas	-0.5635	-3.0873***	-0.0142
California	-1.5961	-1.4477	-1.0819
Colorado	-2.1183***	-1.3036	-0.6376
Connecticut	-0.9418	-1.9567**	-1.0490
Delaware	-0.6795	-0.6418	-1.6152
Florida	-1.6404	-1.0436	-1.8622
Georgia	-0.5410	-1.5568	-1.7974
Hawaii	-1.8919	-2.5422***	-2.0403***
Idaho	-0.6640	-0.6576	-1.4348
Illinois	-0.6708	-1.1427	-0.3145
Indiana	-0.3513	-2.7062***	-0.9864
Iowa	-1.2861	-1.6210	-2.7583***
Kansas	-1.4415	-0.5033	-0.5024
Kentucky	-0.7987	-1.3769	-1.2238
Louisiana	-0.7154	-1.8805	-1.5103
Maine	-0.1249	-2.3277***	-1.3601
Maryland	-1.3286	-2.0348**	-2.0571***
Massachusetts	-1.8555	-2.7840***	-2.9790***
Michigan	-0.8053	-1.2608	-0.6291
Minnesota	-2.1724***	-0.8490	-1.7785
Mississippi	-0.1562	-1.8917	-2.6489***
Missouri	-2.0276	-1.0366	-2.1741***
Montana	-1.8244	-1.5028	-1.4937

Nebraska	-1.5302	-2.5234***	-3.3648***
Nevada	-1.9666**	-0.8284	-1.1182
New Hampshire	-0.9716	-2.5431***	-1.3798
New Jersey	0.1557	-2.9379***	-1.5312
New Mexico	-1.4389	-2.4254***	-0.5335
New York	-1.4054	-0.4861	-1.8177
North Carolina	-0.1883	-0.6752	-1.3827
North Dakota	-2.0016**	-0.9644	-1.4556
Ohio	-0.6865	0.5596	-0.6247
Oklahoma	-1.5786	-0.9153	-1.6319
Oregon	-2.3474***	-1.3173	-0.4397
Pennsylvania	-0.9866	-1.7410	-1.1888
Rhode Island	-1.3137	-1.5499	-2.2142***
South Carolina	-0.0655	-1.0599	-1.9088*
South Dakota	-1.7900	-2.0728**	-1.9273*
Tennessee	-0.7978	-1.1999	-0.0643
Texas	-2.2012***	-0.4324	-1.3550
Utah	-1.4523	-1.7071	-1.2776
Vermont	-0.8074	-2.3887***	-2.1440***
Virginia	-1.3209	-1.8505	-2.2531***
Washington	-2.0600***	-2.2630***	-0.8430
West Virginia	-0.3119	-0.2882	-1.2910
Wisconsin	-1.3015	-2.4041***	0.3709
Wyoming	-1.3757	-1.1364	-1.0535
AVERAGE	-1.0307	-1.0828	-1.0817

Critical values of Panel NCADF Distribution (N = 50, T = 44):

1%	-3.68
5%	-3.04
10%	-2.76

Critical Values of Individual NCADF Distribution (N = 50, T = 44):

1%	-2.05
5%	-1.96
10%	-1.91

Note: Critical values are from table 13. and table 14. of Cerrato et al., (2011).***, **, & * denote 1%, 5%, & 10% critical values respectively.

Table 2. Panel unit root test results

	lnH	lnCO ₂	lnINCOME	
Panel A: Panel Unit Root Test				
CIPS	-1.715	-2.447	-1.981	
NCIPS	-1.184	-1.498	-1.409	
Panel B: Panel Cointegration				
	(1966-2009)	(1990-2009)	(1995-2009)	(1985-2009)
Pedroni-ADF	-3.517***	-7.837***		
with trend	-3.273***	-6.701***		
Pedroni-PP	-1.474*	-6.981***		
with trend	0.1749	-5.790***		
Kao-ADF	-6.041***	-5.658***		
Gt	-1.435	-1.941	-3.957	-2.533***
Ga	-9.662**	-7.423**	-0.563**	-7.711***
Pt	-6.907	-10.968	-5.685***	-15.603***
Pa	-4.919	-5.919*	-0.338***	-6.383***

Note: CIPS denotes the Pesaran (2007) panel unit root test. NCIPS is the nonlinear version of CIPS from Cerrato et al. (2011). Pedroni-ADF, Pedroni-PP, Kao-ADF, stand for Pedroni (2000; 2004) ADF-based and PP-based, and Kao (1999) ADF-based cointegration tests respectively. Ga, Gt, Pa, and Pt stand for the cointegration tests of Westerlund (2007). The tests proposed by Westerlund (2007) account for cross sectional dependence and was calculated through the calculation of robust standard errors by bootstrapping with 1000 replications. The null hypothesis of the reported cointegration tests is "no cointegration". The Pa and Pt test statistics pool information overall the cross-sectional units. Rejection of H₀ should therefore be taken as evidence of cointegration for the panel as a whole. Since cross sectional units are suspected to be correlated, robust critical values can be obtained through bootstrapping.

Table 3. Panel unit root test with structural changes

Panel A: Break date for level shift and regime shift		
States	Break Date (level break)	Break Date (regime shift)
Alabama	1998	1971
Alaska	1980	1980
Arizona	1995	1970
Arkansas	1991	1991
California	1980	1980
Colorado	1993	1980
Connecticut	1980	1980
Delaware	1979	1979
Florida	1982	1976
Georgia	1986	1986
Hawaii	1997	1979
Idaho	1991	1991
Illinois	1991	1991
Indiana	1980	1980
Iowa	1991	1991
Kansas	1991	1991
Kentucky	1991	1991
Louisiana	2000	1982
Maine	1980	1980
Maryland	1998	1972
Massachusetts	1980	1980
Michigan	1979	1976
Minnesota	1991	1991

Mississippi	1991	1991
Missouri	1991	1982
Montana	1991	1991
Nebraska	1980	1980
Nevada	1980	1980
New Hampshire	1980	1980
New Jersey	1980	1980
New Mexico	1991	1979
New York	1983	1971
North Carolina	1980	1980
North Dakota	1980	1980
Ohio	1982	1981
Oklahoma	1991	1991
Oregon	1980	1980
Pennsylvania	1979	1979
Rhode Island	1980	1980
South Carolina	1991	1991
South Dakota	1980	1980
Tennessee	2006	2006
Texas	1994	1973
Utah	1980	1980
Vermont	1986	2002
Virginia	1992	1992
Washington	2002	1979
West Virginia	1991	1991
Wisconsin	1986	1986
Wyoming	1991	1991
AVERAGE	1987	1983

Panel B: Panel cointegration with structural break

(1966-2009)

τ_N

Φ_N

Model	Value	p-value	Value	p-value
No break	-0.098	0.461	1.826	0.966
Level break	-0.965	0.167	0.884	0.812
Regime shift	-0.008	0.497	1.831	0.966
(1995-2009)				
No break	-2.74	0.003	0.009	0.008
Level break	1.549	0.939	0.009	0.504
Regime shift	-0.699	0.242	0.398	0.655

Notes: The test is implemented using the Campbell and Perron (1991) automatic procedure to select the lag length. We use three breaks, which are determined by grid search at the minimum of the sum of squared residuals. The p-values are for a one-sided test based on the normal distribution.

Table 4. Mean group estimations and residual tests

		α	β	γ
MG		1.264	0.202	0.2
		[0.001]	[0.017]	[0.00]
CD test	86.85			
	[0.007]			
MG-FMOLS			0.224	0.233
			[0.002]	[0.00]
CD test	32.22			
	[0.00]			
MG-DOLS			0.332	0.331
			[0.004]	[0.00]
CD test	7.504			
	[0.00]			
CCE-MG	-0.59	0.093	0.242	0.041
	[0.557]	[0.836]	[0.001]	[0.047]

Notes: Figures in square brackets denote p-values. Equation (12) was estimated with Fully Modified OLS (Pedroni, 2000, 2001) and Group Mean Dynamic OLS (Pedroni, 2001). MG, GM-FMOLS, and GM-DOLS stand for standard Mean Group (Pesaran and Smith, 1995), Group Mean Fully Modified OLS (Pedroni, 2000, 2001) and Group Mean Dynamic OLS (Pedroni, 2001). MG, GM-FMOLS and GM-DOLS assume cross section independence. CCE-MG refers to the Common Correlated Effects Mean Group estimation and inference method (Pesaran, 2006) and allows for cross sectional dependence. CD-test refers to the Pesaran's (2004) test of cross sectional dependence.

Table 5. ARDL estimation results

Selected Model: ARDL(2, 4, 4)

Dependent Variable: D(LNH)

Variable	Coefficient	Std. Error	t-Statistic	Prob
Long Run Equation				
LNINCOME	0.8611	0.0543	15.8539	0
LNCO ₂	0.1576	0.0263	5.9812	0
Short Run Equation				
COINTEQ01	-0.3119	0.0321	-9.7176	0
D(LNH(-1))	0.2021	0.0316	6.3977	0
D(LNINCOME)	-0.2851	0.0587	-4.8605	0
D(LNINCOME(-1))	-0.3068	0.0619	-4.9566	0
D(LNINCOME(-2))	-0.0510	0.0634	-0.8039	0.4218
D(LNINCOME(-3))	0.0513	0.0490	1.0464	0.2957
D(LNCO ₂)	-0.0377	0.0188	-2.0071	0.0452
D(LNCO ₂ (-1))	-0.0915	0.0214	-4.2812	0
D(LNCO ₂ (-2))	-0.0076	0.0209	-0.3635	0.7164
D(LNCO ₂ (-3))	-0.0122	0.0216	-0.5670	0.5709
C	-0.5685	0.0658	-8.6367	0
Time Trend	0.0049	0.0007	7.3414	0

Table 6. Quantile estimation of the relationship between personal health care expenditures, personal disposable income, and CO2 emissions in per capita form

		Fixed Effect quantile regression								
Variables		10%	20%	30%	40%	50%	60%	70%	80%	90%
Per capita CO2 in log		0.1319*** (0.0123)	0.1299*** (0.0098)	0.1363*** (0.0085)	0.1413*** (0.0064)	0.1376* (0.0064)	0.143*** (0.0062)	0.1442*** (0.0063)	0.1531*** (0.0070)	0.1632*** (0.0099)
Per capita GDP in log		0.6079*** (0.054)	0.6333*** (0.0375)	0.6124*** (0.0324)	0.6174*** (0.02785)	0.6059*** (0.0242)	0.5913*** (0.0233)	0.5912*** (0.0268)	0.5917*** (0.0329)	0.6076*** (0.0248)

Note: Bootstrap SEs in parentheses with 2000 replications. * significant at 10% and *** at 1%. OLS SEs are robust. The number of observations is 2200.

Table 7: Quantile estimation of the relationship between personal health care expenditures, GDP, inequality, and CO2 emissions in per capita form

Variables	Fixed Effect quantile regression								
	10%	20%	30%	40%	50%	60%	70%	80%	90%
Per capita C02 in log	0.1319*** (0.0123)	0.1299*** (0.0098)	0.1363*** (0.0085)	0.1413*** (0.0064)	0.1376* (0.0064)	0.143*** (0.0062)	0.1442*** (0.0063)	0.1531*** (0.0070)	0.1632*** (0.0099)
Per capita GDP in log	0.6079*** (0.054)	0.6333*** (0.0375)	0.6124*** (0.0324)	0.6174*** (0.02785)	0.6059*** (0.0242)	0.5913*** (0.0233)	0.5912*** (0.0268)	0.5917*** (0.0329)	0.6076*** (0.0248)
Inequality	-0.0923 (0.184)	-0.0719 (0.1529)	-0.0911 (0.1339)	-0.0345 (0.1326)	-0.0074 (0.1227)	0.0745 (0.1335)	0.1051 (0.1281)	0.1425 (0.1456)	0.1065 (0.1952)

Note: Bootstrap SEs in parentheses with 2000 replications. * significant at 10% and *** at 1%. OLS SEs are robust. The number of observations is 2200.