Appendix

Table 1: Historical trends in total health expenditure and out-of-pocket shares in OECD countries, 1960-2010

	Annual growth (%) 1960 to 2010		Out-of-pocket share of total health expenditure (%)		
Country	Nominal	Real	1960	1980	2010
Australia	7.8	4.1	35.8	15.3	19.3
Austria	8.5	4.7	..	..	17.0
Belgium	..	..	..	..	20.7
Canada	7.4	3.7	..	..	14.4
Chile	..	..	..	..	36.5
Czech Republic	..	..	..	..	14.9
Denmark	..	..	..	11.4	13.2
Estonia	..	..	..	..	18.6
Finland	8.2	4.5	43.6	18.4	19.2
France	8.5	4.7	30.3	12.8	7.4
Germany	..	..	..	10.3	13.1
Greece	..	..	..	..	28.8
Hungary	..	..	..	..	26.3
Iceland	8.5	4.7	..	..	18.2
Ireland	9.4	5.6	..	..	18.2
Israel	..	..	..	..	26.0
Italy	..	..	..	..	17.5
Japan	9.8	6.0	..	..	14.4
Korea	..	..	..	73.4	34.2
Luxembourg	..	..	..	7.2	10.0
Mexico	..	..	..	..	49.0
Netherlands	..	..	..	..	5.7
New Zealand	..	..	..	10.4	10.5
Norway	9.8	6.1	..	..	..
Poland	..	..	..	..	22.1
Portugal	..	..	..	..	25.8
Slovak Republic	..	..	..	..	25.9
Slovenia	..	..	..	..	12.2
Spain	11.1	7.2	..	..	19.8
Sweden	..	..	..	..	16.3
Switzerland	7.2	3.5	..	..	25.1
Turkey	..	..	..	..	..
United Kingdom	7.7	4.0	..	8.6	9.2
United States	8.4	4.6	48.9	23.3	11.7
Average	8.6	4.9	39.6	19.1	19.4

Source: OECD Health Data 2013 (http://www.oecd.org/health/healthdata), accessed September 20, 2013.
Notes: The OECD has health expenditure data back to 1960 for 13 countries but only has out-of-pocket spending data in the 1960s for Australia, Finland, France and the United States. The OECD reports data in current US purchasing power parity dollars. To indicate real trends in spending, the authors have corrected the series with the US GDP deflator. As a result, the figures are only an approximation of the true real spending trends.

Table 2. Countries included and excluded from 1995-2009 sample

Countries excluded due to missing data in one or more years:
Afghanistan
Bahrain
Cyprus
Guinea
Guyana
Kuwait
Malawi
Malta
Montenegro
New Zealand
Oman
Qatar
Romania
Sierra Leone
South Korea
Suriname
Countries excluded due to with inaccurate data for variables of interest:
Democratic Republic of Congo
Zimbabwe
Countries included:
Albania
Algeria
Argentina
Armenia
Australia
Austria
Azerbaijan
Bangladesh
Belarus
Belgium
Benin
Bhutan
Bolivia
Botswana
Brazil
Bulgaria
Burkina Faso
Cambodia
Canada
Cape Verde
Central African Repub
Chad
Chile
China
Colombia
Congo, Rep.
Costa Rica
Cote d'Ivoire
Croatia

Czech Republic
Denmark
Djibouti
Dominican Republic
Ecuador
Egypt, Arab Rep.
El Salvador
Estonia
Ethiopia
Fiji
Finland
France
Gambia, The
Georgia
Germany
Ghana
Greece
Guatemala
Haiti
Honduras
Hungary
India
Indonesia
Iran, Islamic Rep.
Ireland
Israel
Italy
Jamaica
Japan
Jordan
Kazakhstan
Kenya
Kyrgyz Republic
Lao PDR
Latvia
Lebanon
Lesotho
Lithuania
Luxembourg
Macedonia, FYR
Madagascar
Malaysia
Mali
Mauritius
Mexico
Moldova
Mongolia
Morocco
Mozambique
Namibia
Nepal
Netherlands

Nicaragua
Niger
Norway
Pakistan
Panama
Papua New Guinea
Paraguay
Peru
Philippines
Poland
Portugal
Russian Federation
Rwanda
Saudi Arabia
Senegal
Singapore
Slovak Republic
Slovenia
Solomon Islands
South Africa
Spain
Sri Lanka
Sudan
Swaziland
Sweden
Switzerland
Syrian Arab Republic
Tajikistan
Tanzania
Thailand
Togo
Trinidad and Tobago
Tunisia
Turkey
Uganda
Ukraine
United Arab Emirates
United Kingdom
United States
Uruguay
Uzbekistan
Venezuela, RB
Vietnam
Yemen, Rep.
Zambia

Table 3. First-differenced results without year dummies

	LN TOTAL HEALTH EXPENDITURE PER CAPITA		LN GOV'T HEALTH EXPENDITURE PER CAPITA		LN OOP HEALTH EXPENDITURE PER CAPITA		OOP SHARE OF TOTAL HEALTH EXPENDITURE	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(GDPpc)	0.691***	0.760***	0.691***	0.760***	0.705***	0.695***	-0.011	-0.0208
	(0.150)	(0.161)	(0.150)	(0.161)	(0.065)	(0.0605)	(0.015)	(0.0170)
Proportion GE/GDP	1.401***	1.460***	1.401***	1.460***	0.159*	0.157*	-0.190***	-0.193***
	(0.246)	(0.270)	(0.246)	(0.270)	(0.094)	(0.0944)	(0.035)	(0.0344)
Proportion age 60+	4.263*	2.070	4.263*	2.070	2.705*	3.561**	0.558	0.544
	(2.233)	(2.067)	(2.233)	(2.067)	(1.488)	(1.555)	(0.363)	(0.386)
Constant	0.009	0.0142*	0.009	0.0142*	0.012***	0.00535	-0.002*	-0.002*
	(0.008)	(0.00823)	(0.008)	(0.00823)	(0.004)	(0.00404)	(0.001)	(0.001)
R^2 Within	0.064	0.110	0.064	0.110	0.081	0.117	0.033	0.084
F-statistic	18.72	9.82	18.72	9.82	39.34	27.49	10.12	13.85
Year FEs	No	No	No	No	No	No	No	No
Reg'l dum		Yes		Yes		Yes		Yes
CD	Yes	No	Yes	No	Yes	No	No	No

[b] Robust standard errors clustered by country are in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%.

Box. Cross-sectional dependence, autocorrelation, and unit roots

The challenge

One obstacle to deriving accurate conclusions from our model is that the estimates may be biased if the panel data exhibits cross-sectional dependence. In our dataset, cross-sectional dependence would occur, for example, if health spending in a particular country and year were systematically influenced by similar changes in a neighbouring country. Because our dataset has a relatively small number of years (T) and large number of countries (N), we use the Pesaran (2004) test for cross-sectional dependence (CD).

Our estimates from level regressions will also be biased if the dependent variable exhibits a unit root process. Therefore, we test for unit root processes in the levels of each dependent variable using the augmented Dickey-Fuller (ADF) regression. We also conduct panel unit root tests which both assume cross-sectional independence (i.e. Im, Pesaran, & Shin (IPS), 2003; Maddala & Wu, 1999). If we fail to reject the null (i.e., the presence of a unit root), then regressions in levels may be spurious. In contrast, as noted earlier, the error term is serially uncorrelated in regressions in first-differences where the variable has a unit root (integrated at order one).

Empirical findings

Estimates in the level fixed-effects models may be biased if the error terms in the panel data are autocorrelated. In fact, in all six level fixed-effect models, the assumption that there is no autocorrelation is rejected. Fortunately, the first-differences models presented in columns 6 and 8 of Table III are consistent with the assumption of no autocorrelation in the error terms and the first-differences model (column 8) is therefore, in our judgment, the preferred specification.

The estimates are also subject to bias when the data has cross-sectional dependence, as indicated by the CD test in the first six columns of Table III. Following Baltagi and Moscone (2010), we address this problem by including regional averages of the dependent variable and independent variables for both the level fixed-effects and first-differences models (columns 7 and 8, Table III). With this addition to the specification, the regressions no longer exhibit cross-sectional dependence. Therefore, the first-differences model in the final column of Table III remains our preferred specification. The income elasticity increases slightly when including cross-sectional averaged variables, and is notably higher than the estimate of 0.446 which Baltagi and Moscone (2010) obtain for OECD countries in a regression with cross-sectional average of the dependent and independent variables. Our tests for unit roots such that non-stationary is not a serious problem

The first-differences model that addresses cross-sectional dependence (column 8 in Table III) yields an estimate of 0.723 for the income elasticity of total health expenditure, which is significantly lower than the estimates in the level fixed-effects model and slightly higher than in the first-differences model that ignores cross-sectional dependence. The estimate of 0.723 is similar to the income elasticity of 0.674 presented by Baltagi and Moscone (2010) for a regression that addresses cross-sectional dependence without covariates. However, our estimate is higher than the estimate of 0.446 which they derive when covariates are included.

In terms of unit roots, both the IPS and the Fisher-type ADF tests suggest that at least one of the country series is stationary for the main dependent variables of interest (see appendix). The IPS test rejects the null hypothesis of a unit root when including a trend and fails to

reject the null when lags are included. The Fisher-type ADF tests also reject the null in most cases, even when lags are included. These results would suggest that non-stationarity is not a serious problem except for the fact that these unit root tests are not robust in the presence of cross-sectional dependence. Nevertheless, our tests for unit roots are similar to those found in Baltagi and Moscone (2010) who additionally apply a novel test that accounts for cross-sectional dependence, leading them to reject the hypothesis of a unit root when variables are analysed in first-differences (see Appendix Table 4 below). This confirms our preference for the results of the first-differences models.

References

Baltagi, B. H., & Moscone F. (2010). Health care expenditure and income in the OECD reconsidered: Evidence from panel data. *Economic Modelling*, *27*, 804-11.

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 Statistics*, *61*, 631-652.

Pesaran, M. H. (2004). General diagnostic tests for cross section dependence in panels. University of Cambridge, Faculty of Economics, Cambridge Working Papers in Economics No. 0435.

Table for Box. Cross-sectional dependence, autocorrelation, and unit roots
Unit root tests

	IPS unit root tests				
Variable	*L0*	*L0 Trend*	*L1 Trend*	*L2 Trend*	*L3 Trend*
GDPpc	13.20	1.86	2.78	1.64	4.14
THEpc	8.79	-6.14***	-0.86	-0.23	-1.71**
GHEpc	8.13	-6.01***	0.80	0.51	-0.77
OOPpc	8.46	-6.59***	-1.23	0.56	0.27
OOP/THE	0.29	-5.93***	0.14	-1.58*	1.67
	Fisher-type ADF				
Variable	*L0*	*L0 Trend*	*L1 Trend*	*L2 Trend*	*L3 Trend*
GDPpc	167.9	224.8	310.5***	268.9	248.6
THEpc	286.1	383.2***	377.6***	319.3***	401.6***
GHEpc	241.6	329.9***	331.9***	343.3***	436.7***
OOPpc	415.2***	345.9***	394.6***	345.3***	304.8***
OOP/THE	347.9***	280.1	373.5***	403.8***	268.6

Notes: *** significant at 1%; ** significant at 5%; * significant at 10%. For IPS unit root tests, t-tilde-bar values are presented. For Fisher-type ADF tests, inverse chi-squared statistics are presented. Natural logs of per capita variables were used. 'L' refers to the number of lags included and 'trend' indicates a test for trend-stationarity.