APPENDIX: Health Care Spending Growth Has Slowed: Will the Bend in the Curve Continue?

Sheila D. Smith, Joseph P. Newhouse, Gigi Cuckler

---

1. ADJUSTMENTS TO OECD DATA

Health and economic data used in this analysis were obtained from the Organization for Economic Co-operation and Development (OECD) OECD Health Statistics 2021 online database ("Health expenditure and financing" category) (OECD 2021). The measure of health spending was the "Total current expenditure" series (OECD definition HC.1-HC.9). This series includes "Individual and collective health care" but excludes investment spending—a distinction that reflects the primary focus of the analysis on the allocation of health care consumption out of total income. Gross Domestic Product (GDP) by country was used as a measure of income rather than other measures such as disposable personal income in order to maximize the availability of the data across countries and time.

We converted all these series to a per capita basis and then into U.S. dollar terms using purchasing power parities (PPP). We then deflated these nominal time-series by the U.S. GDP deflator to convert the series to real (constant dollar) terms (Bureau of Economic Analysis 2020).[1]

The health spending data from the OECD Health Statistics 2021 presented two key challenges: (1) the time-series for health spending contained a number of omitted values; and (2) there were conceptual changes in the definition over time in several countries that were detailed in the OECD documentation.

Further contributing to omitted values, beginning with the OECD Health Statistics 2015 database, health statistics availability was limited to the years from 1970 forward, such that the years 1960-69 could no longer be downloaded from the web.[2] In the current OECD Health Statistics 2021 database, there were also notably fewer countries with GDP data populated from

[1] Table 1.1.4. Price Indexes for Gross Domestic Product (line 1)

[2] Through email communication, OECD staff indicated that this change was made due to limitations in country coverage, breaks in time series due to methodological changes, and limited data availability for deflators and other macro-economic indicators before 1970.

1960-69. Archived data from the 2015 database were obtained through communications with OECD. For GDP data, OECD economic data were supplemented with data from the World Bank to extrapolate economic data back through 1960 when possible (World Bank 2021).

For instances with missing annual values of OECD health expenditures because of periodic surveys that were less frequent than annual, constant growth was assumed between years of available data. In cases in which a break was evident in the data series (such as a break in the series definition or change in methodology identified by OECD or a break identified through visual inspection), we employed a smoothing methodology. We treated the archived health statistics data for 1960-69 as a definitional change in the series and as such, the same smoothing methodology was applied as with other definitional changes in the data series. This process followed the methodology used in historical analysis of international health spending data published by OECD (OECD 2014). Essentially this method revises the growth rate in the break year by replacing it with the average growth rate in the preceding 5 years (to approximate trend growth). The levels before the break year were then revised by retropolation using actual growth rates (those of the unadjusted series).

The OECD 2021 health database source documentation identified key areas with definitional modifications due to changes in reporting requirements or other discontinuities in the health data series (available directly from the OECD database (OECD 2021) via linked documentation files by country, called “Data, Sources, and Methods” and under the sub-heading, “Notes on Comparability”). We investigated these country-specific data idiosyncrasies, related to definitional or methodological changes for the OECD health accounts, and smoothed where appropriate. However, we investigated major policy, economic, and political changes by country to ensure that any adjustments that we made in the data did not override an important historical trend. A few obvious cases of trend breaks related to economic, political, or policy changes include major health reform in Australia in 1975; the revolutions of 1989 in the former Communist states of Central and Eastern Europe; entry into the European Union by Ireland and that country’s ensuing economic fluctuation; and, more recently, the 2008-2009 financial crisis, which is evident in the data across several OECD countries.

The relevant definitional and methodological breaks for the OECD health spending accounts data used in modeling include the following:

Country	**Year(s) Adjusted**
All countries with archived data	1970
Australia	1971, 1998
Austria	1975, 1981, 1990
Belgium	1995, 2003
Canada	1975
Denmark	1980, 2003, 2010
Finland	1993, 2015
France	1995, 2003, 2006
Germany	1992
Iceland	1980, 2003

Japan	1995, 2011
Netherlands	1998
Norway	1990, 1993, 1997, 2001
Portugal	1995, 2000
Spain	1999, 2003
Sweden	1993, 2011
Switzerland	1995
United Kingdom	1997

As stated earlier, historical data were not available for Gross Domestic Product in the 2021 OECD database for several countries from 1960-69. Where possible we obtained economic data from the World Bank to backfill GDP data (World Bank 2021). Specifically, we used the data series "GDP per capita (current US$)" (series code NY.GDP.PCAP.CD) from the World Bank DataBank of World Development Indicators for all countries where data were available. To fill in the historical levels, growth rates from World Bank data were applied to the levels from the 2021 OECD data for GDP per capita in US dollars to avoid any potential shifts in the level. In cases where World Bank data were unavailable for the historical years, but archived GDP data from prior OECD 2014 data were available, growth rates from archived OECD data were applied to the OECD 2021 levels to backfill for omitted historical values. In addition, in three instances, we adjusted Gross Domestic Product data to account for apparent breaks in the data based on visual inspection of charted trends, again using the same smoothing method applied to the health expenditure data. Below is a list by country of the historical data source and also of the three smoothing adjustments.

Adjustments to OECD time series: data source and relevant period

Country	**GDP Historical Source Data (for 1960s)**	**Adjusted Year**
Australia	OECD 2021 Data (1960-)	
Austria	World Bank (1960-69)	
Belgium	World Bank (1960-69)	
Canada	World Bank (1960), OECD 2021 Data (1961-)	
Chile	World Bank (1960-85)	
Colombia	World Bank (1960-74)	
Costa Rica	World Bank (1960-90)	1981
Czech Republic	None Available	
Denmark	OECD 2014 Data (1960-65)	
Estonia	None Available	
Finland	World Bank (1960-69)	
France	OECD 2021 Data (1960-)	
Germany	OECD 2014 Data (1960-69)	1991
Greece	OECD 2021 Data (1960-)	
Hungary	None Available	
Iceland	OECD 2014 Data (1960-69)	
Ireland	World Bank (1960-69)	

Israel	None Available	
Italy	World Bank (1960-69)	
Japan	World Bank (1960-69)	
Korea	World Bank (1960-69)	
Lithuania	None Available	
Latvia	None Available	
Luxembourg	World Bank (1960-69)	
Mexico	World Bank (1960-69)	
Netherlands	World Bank (1960-69)	
New Zealand	World Bank (1960-69)	
Norway	World Bank (1960-69)	
Poland	None Available	
Portugal	World Bank (1960-69)	
Slovak Republic	None Available	
Slovenia	None Available	
Spain	World Bank (1960-69)	
Sweden	OECD 2021 Data (1960-)	
Switzerland	OECD 2014 Data (1960-69)	
Turkey	World Bank (1960-69)	1998
United Kingdom	OECD 2021 Data (1960-)	
United States	World Bank (1960-69)	

Sample Selection

The selection of the sample of countries for this analysis was largely determined by the availability of the country-level data in the OECD database. Given that one focus of the analysis was the long-term relationship between health spending and income, one would like a sample with the longest possible and most complete history – that is, a balanced sample with no missing intermittent observations. Against that, one would also like a sufficient number of countries in the cross-sectional sample to ensure that estimated coefficients were stable and robust to changes in the sample time period.

Twelve countries have health spending data available from 1960. They are as follows: Australia, Austria, Canada, Finland, France, Iceland, Ireland, Japan, Spain, Switzerland, United Kingdom, United States. If this list is expanded to include countries with health spending data back from 1970, the sample rises to the following 21 countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Iceland, Ireland, Japan, Korea, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, United States.[3] We used the sample available from 1970 forward, excluding Ireland from the modeling sample because it was a significant outlier in terms of health and economic trends. We ended the sample in 2019

[3] Note that the Netherlands (data available from 1972) and Denmark (data available from 1971) were included in this sample. For these two countries, growth from the first available year was extrapolated back through 1970 to allow for these countries to be included in the 1970 country sample.

because, in addition to the temporary distortions from the pandemic, 2020 data were unavailable for several countries when we began the study.

2. Estimation of demographic indexes for OECD countries

Demographic effects on health care spending result from shifts in the composition of population across demographic cohorts with varying intensity of health care utilization. For the 20-country sample of OECD countries in our decomposition of health care spending growth we estimate the effect of changes across age cohorts, capturing the effects of variation in the rate of population aging on growth in health expenditures.

We assume relative spending on medical care per capita for a given age cohort remains constant over time at its base year value. The impact of demographic change on health care spending is then measured by the change in the distribution of population across age cohorts between year *t* and *t+1*, weighted by constant relative health care spending per capita for each age cohort. The index of spending associated with the mix of the population across age cohorts under this method is D_t as defined in Equation A1 below: an aggregate of population shares in each age cell weighted by real spending per capita corresponding to those cells in the base year distribution. The contribution of changing demographic mix on spending equals the growth rate of D_t.

$$(A1) \qquad D_t = \sum_{a=1}^{N} (n_{a,t}/n_t) * h_a$$

D_t = Index of spending for mix of population across age
$n_{g,a,t}$ = population in age cohort a, time t
n_t = total population across all cohorts, time t
h_a = base year spending per person by age cohort a, relative to mean
N = number of age cohorts

Calculating the D_t index requires two inputs. First, we need time series for population by age for each of our 20 OECD countries, which are readily available (OECD 2021). Second, we need estimated variation in spending per person for each age cohort for a base year. For this purpose, we rely on estimates for an eight-country sample for the year 2015 (Papanicolas et al 2020, estimates shown in Figure A1).

While the eight countries included in the Papanicolas et al (2020) sample overlap with our 20-country OECD sample, we lack data for spending by age cohort for the 12 remaining countries in the sample. However, we found that there was relatively little variation in the distribution of health care spending by age across the eight countries for which data was available.

We estimated the demographic index D_t based on two methods: 1) own-country data for base-year spending by age, versus 2) a weighted average of spending by age for the eight-country sample. Figure A2 shows the contribution to growth from demographic change for 1970-2019 based on these two methods for each of the eight countries; the sensitivity of the estimated contribution to growth to own versus 8-country average health spending by age is small.

We then generated demographic indexes based on equation A1 for all 20 countries in our pooled OECD country sample based on the 8-country mean from (Papanicolas 2020).

3. IMPUTATION OF US MEDICAL PRICE INFLATION, 1970-1995, AND DECOMPOSITION OF RELATIVE MEDICAL PRICE INFLATION.

Relative medical price inflation for 1970 to 2019 is estimated based on two different methods for pre and post-1996. The change in method is conditional on the availability of data for Producer Price Indexes (PPI) for all major health sectors (hospitals, physician offices, nursing homes, home health agencies, prescription drugs, and medical durables) beginning in 1996. Before 1996, relative medical price inflation is imputed based on underlying drivers of growth (factor price inflation and total factor productivity growth) and relies on data and methodology from the Bureau of Labor Statistics Office of Productivity and Technology (OPT). From 1996 forward, medical price inflation is based on the CMS chain-weighted price deflator for personal health care from the National Health Expenditure Accounts – largely composed of PPI's for the individual health sectors – as a measure of price inflation for medical care (Centers for Medicare & Medicaid Services 2020).

Imputation of medical price inflation for the pre-1996 period

Prior to the availability of the PPI, price data for health sectors were limited to the Consumer Price Index (CPI), which is seriously flawed in several ways as a measure of long-term medical price inflation (Berndt et al 2000, 2001). Major concerns include the use of list prices rather than transactions prices (over a period when the divergence between these two metrics was increasing in magnitude) as well as an item structure within the CPI that subsumes variation in the composition of treatment over time as well as price change. As an alternative to the CPI, we impute price inflation before 1996 based on its underlying determinants. Specifically, medical price inflation is equal to the sum of growth in input costs of production, total factor productivity (TFP), and producer profits. The contribution from input costs is based on estimates from the BLS Office of Productivity and Technology (described below) for 1987 forward, and extrapolated back to 1970 based on estimates prepared for CMS by Dale Jorgenson and Mun Ho for 1970 to 1987 (Ho and Samuels, 2006).

Data from the BLS Office of Productivity and Technology (OPT)

Estimated input price inflation for health care is based on estimates from the BLS Office of Productivity and Technology, using variables that are defined in the process of estimation of growth in total factor productivity. TFP is estimated as a residual, and thus requires estimates of real growth in factor inputs for (capital (K), labor (L), energy (E), materials (M), and business services (S), referred to as KLEMS). The estimation of factor input prices is a necessary step in the estimation of real spending. The OPT also generates a composite index of input prices ('Combined inputs price deflator') for all KLEMS inputs for each major and detailed industry sector. The BLS price deflator shown in Figure 2 and Table 3 is based on the 'Sectoral output price deflator' for each NAICS code as shown below. All indexes are available from 1987 through 2020, however data for 2020 was excluded from the sample due to distortions from the effects of Covid-19. Industry sectors used in our analysis draw from files for both major industry sectors, and detailed manufacturing sectors, as shown below. Data files were accessed from https://www.bls.gov/productivity/tables/ in November 2022.

Industry sector	NAICS	Type*	File name
Ambulatory health care services	621	M	major-industry-total-factor-productivity-klems.xlsx
Hospitals and nursing and residential care	622, 623	M	major-industry-total-factor-productivity-klems.xlsx
Pharmaceuticals and medicine	3254	D	total-factor-productivity-manufacturing-and-transportation-detailed-industries.xlsx
Medical equipment and supplies	3391	D	total-factor-productivity-manufacturing-and-transportation-detailed-industries.xlsx

*Major sector (M), or detailed manufacturing sector (D)

The concepts available for major and detailed industry sectors differ slightly in detail; major sectors break out each factor input within KLEMS. Detailed industry sectors consolidate all intermediate inputs into a single category (energy, materials, and business services).

<table>
<tr><th></th><th colspan="2">Price deflator</th><th colspan="2">Quantity (real)</th><th colspan="2">Factor share</th></tr>
<tr><td></td><td>Major</td><td>Detailed</td><td>Major</td><td>Detailed</td><td>Major</td><td>Detailed</td></tr>
<tr><td>Output</td><td>X</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Input price (KLEMS)</td><td>X</td><td>X</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Capital(K)</td><td>X</td><td>X</td><td>X</td><td>X</td><td>X</td><td>X</td></tr>
<tr><td>Labor (L)</td><td>X</td><td>X</td><td>X</td><td>X</td><td>X</td><td>X</td></tr>
<tr><td>Energy (E)</td><td>X</td><td rowspan="3">X</td><td>X</td><td rowspan="3">X</td><td>X</td><td rowspan="3">X</td></tr>
<tr><td>Materials (M)</td><td>X</td><td>X</td><td>X</td></tr>
<tr><td>Business Services (S)</td><td>X</td><td>X</td><td>X</td></tr>
</table>

Price indexes (output price, combined inputs price, and factor prices by KLEMS) for are chain-weighted across all sectors by NAICS to generate aggregate indexes of input price inflation that match the scope of total personal health care expenditures within the National Health Expenditure Accounts published by the Centers for Medicare & Medicaid Services (CMS 2021).

Extrapolation of input price inflation for 1970 through 1987

Input price inflation for 1970 to 2000 is estimated based on a database compiled for the Centers for Medicare & Medicaid Services for the estimation of health care productivity (Ho and Samuels 2006). These estimates rely on a KLEM-based model, closely paralleling the framework employed by the US Bureau of Labor Statistics. Differences include the use of a single category for inputs of materials and business services, and industry definitions based on the Standard Industrial Classification (SIC) rather than the North American Industrial Classification System (NAICS) definitions. In addition, there are some differences in underlying source data, particularly for the estimate of unit costs of labor inputs. When aggregated to the level of personal health care expenditures, the implied growth in input price inflation for the overlapping period from 1987 to 2000 tracks reasonably well for input price indexes based on the two sources. We used the resulting input price index for personal health care based on the CMS estimates to extrapolate the BLS-based input price input price index for personal health care backwards from 1987 to 1970 to generate an index that covered the entire period from 1970 to 2019.

Total factor productivity assumption

The issues with the CPI also imply concerns about the use of estimates for TFP for the pre-1996 era; estimating productivity growth necessarily relies on accurate measures of output price in order to obtain measures of growth in quantity of output. Estimates of growth in TFP that rely upon CPI price data are consistently negative, implying persistent and lasting reductions in productivity. This unlikely result probably stems from an upward bias in the CPI measures of output price inflation. For this reason, we base the contribution to relative medical price inflation from TFP prior to 1996 on an assumed range relative to economy-wide growth in TFP – from a lower bound of zero to an upper bound of 75 percent of estimated economy-wide growth in TFP.

The assumption of slower growth in TFP for the production of health care than for economy-wide TFP is based on a key assumption in the Baumol 'cost disease' model (Baumol 1967, 2012). The Baumol model predicts that productivity growth in service industries will be systematically slower than for the economy as a whole, particularly in areas like health care services which require a high degree of customization of services. The estimated contribution to health care spending growth from relative medical price inflation is not highly sensitive to the assumed range for relative productivity growth (the range for relative TFP corresponds to a range in contribution to growth from 6.4% to 9.7% over 1970 to 2019). Economy-wide growth in TFP is based on estimates from the US Bureau of Labor Statistics for total private business (https://www.bls.gov/productivity/tables/, file: total-factor-productivity-major-sectors.xlsx), accessed September 2021. To obtain an economy-wide estimate, we assumed zero productivity growth for public sector output.

Decomposition of Relative Medical Price Inflation

The decomposition of growth in relative medical price inflation in Table 3 is based on the same BLS OPT database described above, utilizing indexes of output price inflation, input price inflation, factor shares by KLEMS, factor input prices by KLEMS, and estimated TFP described above for the imputation of relative medical price. For the decomposition of relative output price inflation, we use data from 1996 through 2019 (again based on PPI availability).

4. Estimation of lag distribution for income variable for OECD countries

We estimated the lag distribution for the income variable in Equations 1.1 and 1.2 after constraining the dependent variable to exclude effects from non-income variables consistent with those equations. We estimated the equation in first differences (growth rate) in order to fit the short-term cyclical variation in *h* and *y* that is the critical to estimation of the the lag in the relationship. The OECD exUS sample estimation is based on a time-series for the equal-weighted aggregate of the 20-country sample. The estimation interval is 1970 to 2019. Our objective was to find the best fit on average across all countries in the sample; aggregation across countries tempers time-series volatility in individual country data and allows for an approximation of the lag structure that best fits the aggregate relationship on average across all countries.

We estimated the model with a polynomial-distributed-lag (PDL) for the coefficients on growth in real per capita GDP, constraining the coefficients on lagged values to fit to a second-degree polynomial (Equation A2). This allows for impacts from lagged income effects that vary over time in a systematic pattern. The number of periods specified for the estimation was chosen to maximize the explanatory power of the regression based on adjusted R-squared.

$$\text{(A2)} \qquad \mathrm{d}\left(\ln\left(H_{20,t}\right) - \beta_I \ln\left(I_{20,t}\right) - \ln\left(D_{20,t}\right)\right) = \sum_{s=0}^{5} \left(\gamma_0 + \gamma_1 s + \gamma_2 s^2\right) \mathrm{d}\left(\ln\left(Y_{20,t}\right)\right) + \mu_t$$

$H_{20,t}$ = real per capita spending, health consumption expenditures, 20-country aggregate
$Y_{20,t}$ = real per capita spending, health consumption expenditures, 20-country aggregate
$I_{20,t}$ = out-of-pocket share of health consumption expenditures, 20-country aggregate
$D_{20,t}$ = demographic index, 20-country aggregate

Based on estimating equation A2 with varying lag lengths of up to eight years, a lag from 0 to 5 years on growth in real per capita GDP produced the best fit (Figure A3). Coefficients for the first and second degrees of the quadratic lag function were insignificant in this equation, so we assumed the lag distribution was best represented as a simple moving average (equal weights on all lags). This specification of the lag distribution is assumed in the specification of equations 2.1 and 2.2.

5. Estimation of standard error for income-technology interaction effect

The standard error (SE) of the interaction effect between income and technology is the SE for the difference between the estimated coefficients on the income variable in equations 2.1 and 2.2 (β_y' - β_y). The standard error of difference in coefficients will equal the square root of the variance (Var(β_y' - β_y) = Var(β_y') + Var(β_y) - 2*Cov(β_y',β_y)). This can be approximated by assuming that the covariance is zero (SE(β_y' - β_y) = sqrt(Var(β_y') + Var(β_y)). This method implies a standard error of 0.0361.

We also estimated SE(β_y' - β_y) more precisely by estimating a two-stage model as described by Tofighi et al (2011). The resulting estimate for the standard error based on this approach is close to the approximation based on the assumption of zero covariance.

6. BAUMOL VARIABLE AND ASSOCIATED MODEL EQUATIONS

The 'Baumol variable'(Hartwig, 2008) can be defined for a large subsample of the OECD pooled sample used for the estimation of income and expenditures elasticities. Data for this estimation are drawn from the OECD Compendium of Productivity Indicators, 2022. Labor productivity is defined as GDP (value added) per employed person. As an indicator of economy-wide wages, we use compensation per employed person, defined as the sum of compensation for employees and self-employed workers, including wages, in-kind compensation, and contributions to social insurance by employers, divided by total employment. Labor income received by self-employed is recorded in national accounts as mixed income, which bundles both their labor and capital income. Labor compensation for the self-employed is imputed based on the assumptions that hourly compensation received by self-employed workers is equal to that for employed workers by industry.

While data for GDP per employee are broadly available, the labor compensation data have gaps for countries in our OECD sample. This limits our sample to 14 countries for 1990 to 2019, with gaps in the data for some countries in the sample before 1995.[4]

The Baumol variable is defined as the differential between growth in economy-wide real wages (w) and real labor productivity (y).

(A3) $$\ln(c_h) = \lambda \ln(w/y)$$

c_h = unit costs of production, health care sector
w = real compensation per employed person (economy-wide)
y = real GDP per employed person (economy-wide)

We add this variable to the model specification in equations A3.1 and A3.2 to produce the revised specification shown in equations A3.1 and A3.2 below.

(A3.1) $$\ln(H_t) - \beta_I \ln(I_t) - D_t = \alpha + \beta_y^b \mathrm{MA}\,(\ln(Y_t), 5) + \lambda \ln\left(\frac{\mathrm{w}}{\mathrm{y}}\right) + \sum_{c=0}^{I} c_i + \sum_{t=0}^{T} yr_t + \mu_{it}$$

(A3.2) $$\ln(H_t) - \beta_I \ln(I_t) - D_t = \alpha + \beta_y^{b\prime} \mathrm{MA}\,(\ln(Y_t), 5) + \lambda' \ln\left(\frac{\mathrm{w}}{\mathrm{y}}\right) + \sum_{c=0}^{I} c_i + \mu_{it}$$

H_t	= real per capita spending on health consumption, time t
I_t	= out-of-pocket share of health care spending, time t
D_t	= demographic index, time t
Y_t	= real per capita GDP, time t
c_i	= fixed effects for each country i
yr_t	= fixed effects for each year t
MA(ln(Y_t),5)	= moving average of ln(Y_t), 0 to 5 years

[4] Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Netherlands, Norway, New Zealand, Portugal, Spain, UK, US

The Baumol model predicts that λ will be positive and less than or equal to 1.0. The results from estimating equations A3.1 and A3.2 are consistent with this prediction; the coefficients on these variables are significant in both equations. The coefficient for the model with 2-way fixed effects is 0.67, and for the model with country-fixed effects is 0.39.

The estimated results of the model including and excluding the Baumol variable confirm a positive interaction effect between income and relative medical price effect; the coefficient on real per capita GDP is significantly lower for the model with 2-way fixed effects (equation A3.1). The reduction for the model with country fixed effects (equation A3.2) with the inclusion of the Baumol variable is consistent with a predicted positive interaction effect, however the implied interaction between income and relative medical price is small for this specification. Estimated coefficients on the Baumol variable and on the income variable in equations A3.1 and A3.2 are shown in Table 1.

7. Test for structural break in residual

The residual is estimated based on Equation 1b (the simplified residual) for the United States and the OECD exUS to allow for definitional consistency. We test for structural change in the mean contribution to growth from the residual over the 1970-2019 period for both the United States and the OECD exUS (countries are equal-weighted). Given the volatility of the time-series for the residual contribution to growth, we estimate the test for structural breaks based on a smoothed version of the series ($MA(\epsilon_{i,t}(1), 3)$ = centered three-year moving average of the estimated contribution to growth from the residual).

The regression specification is shown in Equation A4 below.

$$\text{(A4)} \qquad d(MA(\epsilon_{i,t}(1), 3)) = c + \mu_{it}$$

i = United States, OECD ex US

Results of tests for structural breaks based on the Quandt-Andrews test for structural change are shown below. In the context of Equation A4, we are testing for a structural break in the mean contribution to growth in health spending from the residual. The null hypothesis of no structural change is rejected at the 1% level for both the United States and the OECD ex US (Hansen 1997). The most probable structural break is identified as 2005 for the United States, and 2004 for the OECD ex US.

United States

Quandt-Andrews unknown breakpoint test
Null Hypothesis: No breakpoints within 15% trimmed data
Equation Sample: 1972 2018
Test Sample: 1980 2011
Number of breaks compared: 32

Statistic	Value	Prob.
Maximum LR F-statistic (2005)	20.58030	0.0002
Maximum Wald F-statistic (2005)	20.58030	0.0002
Exp LR F-statistic	8.038679	0.0000
Exp Wald F-statistic	8.038679	0.0000
Ave LR F-statistic	10.89939	0.0000
Ave Wald F-statistic	10.89939	0.0000

Note: probabilities calculated using Hansen's (1997) method

OECD excluding United States

Quandt-Andrews unknown breakpoint test
Null Hypothesis: No breakpoints within 15% trimmed data
Varying regressors: All equation variables
Equation Sample: 1972 2018
Test Sample: 1980 2011
Number of breaks compared: 32

Statistic	Value	Prob.
Maximum LR F-statistic (2004)	29.58810	0.0000
Maximum Wald F-statistic (2004)	29.58810	0.0000
Exp LR F-statistic	11.37808	0.0000
Exp Wald F-statistic	11.37808	0.0000
Ave LR F-statistic	6.657487	0.0010
Ave Wald F-statistic	6.657487	0.0010

Note: probabilities calculated using Hansen's (1997) method

References

Bates, Laurie J. and Santerre, Rexford E. 2013. "Does the U.S. health care sector suffer from Baumol's cost disease? Evidence from the 50 states." *Journal of Health Economics* (Elsevier) 32 (2): 286-291. doi:https://doi.org/10.1016/j.jhealco.2012.12.003.

Baumol, William J. 2012. *The Cost Disease: Why Computers Get Cheaper and Health Care Doesn't.* New Haven: Yale University Press.

Baumol, William J. 1967. "The Economics of Unbalanced Growth: The Anatomy of Urban Crisis." *The American Economic Review* 57 (3): 415-426.

Berndt, Ernst R., David M.Cutler, Richard G Frank, Zvi Griliches, Joseph P.Newhouse, and Jack E.Triplett. 2001. "Price Indexes for Medicare Care Goods and Services: An Overview of Measurement Issues." Chap. 4 in *Medical Care Output and Productivity*, edited by Ernst R.Berndt David M.Cutler, 141-200. Chicago: NBER Studies in Income and Wealth.

Berndt, Ernst, David M.Cutler, Richard G Frank, Zvi Griliches, Jospeh P.Newhouse, and Jack E. Triplett. 2000. *Medical Care Prices and Output.* Vol. 1a, chap. 3 in *Handbook of Health Economics*, by Anthony J.Culyer and Joseph P.Newhouse, 119-180. Amsterdam: Elsevier.

Bureau of Economic Analysis. 2020. *https://www.bea.gov.* November 25. Accessed May 24, 2022. https://apps.bea.gov/iTable/iTable.cfm?reqid=19&step=2&isuri=1&1921=survey#reqid=19&step=2&isuri=1&1921=survey.

Bureau of Labor Statistics Multifactor Productivity Program. 2007. *Technical Information about the Multifactor Productivity Measures.* Bureau of Labor Statistics. https:/www.bls.gov/mfp/mprtech.pdf.

Bureau of Labor Statistics, Office of Productivity and Technology (OPT). n.d. *https://www.bls.gov/productivity/tables/.* Accessed September 2021.

Centers for Medicare & Medicaid Services. 2020. "National Health Expenditures Accounts Methodology Paper, 2020." Baltimore. https://www.cms.gov/files/document/definitions-sources-and-methods.pdf.

Centers for Medicare and Medicaid Services, Office of the Actuary, National Health Statistics Group. 2021. *National Health Expenditures Data - Historical.* December 21. Accessed February 1, 2022. https://www.cms.gov/files/zip/national-health-expenditures-type-service-and-source-funds-cy-1960-2020.zip.

Hansen, Bruce E. 1997. "Approximate Asymptotic P-Values for Structural Change Tests." *Journal of Business and Economic Statistics* 15 (1): 60-67. doi:10.2307/1392074.

Hartwig, Jochen. 2008. "What drives health care expenditures? -- Baumol's model of 'unbalanced growth' revisited." *Journal of Health Economics* (Elsevier) 27 (3): 603-623.

Ho, Mun S., Jon D. Samuels. 2006. *An Intertemporal General Equilibrium Model of U.S. Health Expenditures.* Cambridge: Dale Jorgenson Associates, 181-205.

Jorgenson, Dale, and Kevin Stiroh. 2000. "Raising the Speed Limit: U.S. Economic Growth in the Information Age." *Brookings Papers on Economic Activity* 157.

Luca Lorenzoni, Alberto Marino, David Morgan, Chris James. 2019. *Health Projections to 2030: New results based on a revised OECD methodology.* OECD Health Working Papers, OECD, Paris: OECD Publishing, 46. doi:https://doi.org/10.1787/5667f23d-en.

OECD. 2014. *Health Care Systems: Efficiency and Policy Settinsg.* Paris: OECD Publishing. Accessed 26 2014, September. doi:10.1787/978926409901-en.

—. 2021. *Health Expenditure and Financing.* Accessed August 30, 2021. https://stats.oecd.org/Index.aspx?/DataSetCode=SHA#.

—. 2021. *Health Expenditure and Financing.* Accessed July 15, 2021. https://stats.oecd.org/Index.aspx?/DataSetCode=SHA#.

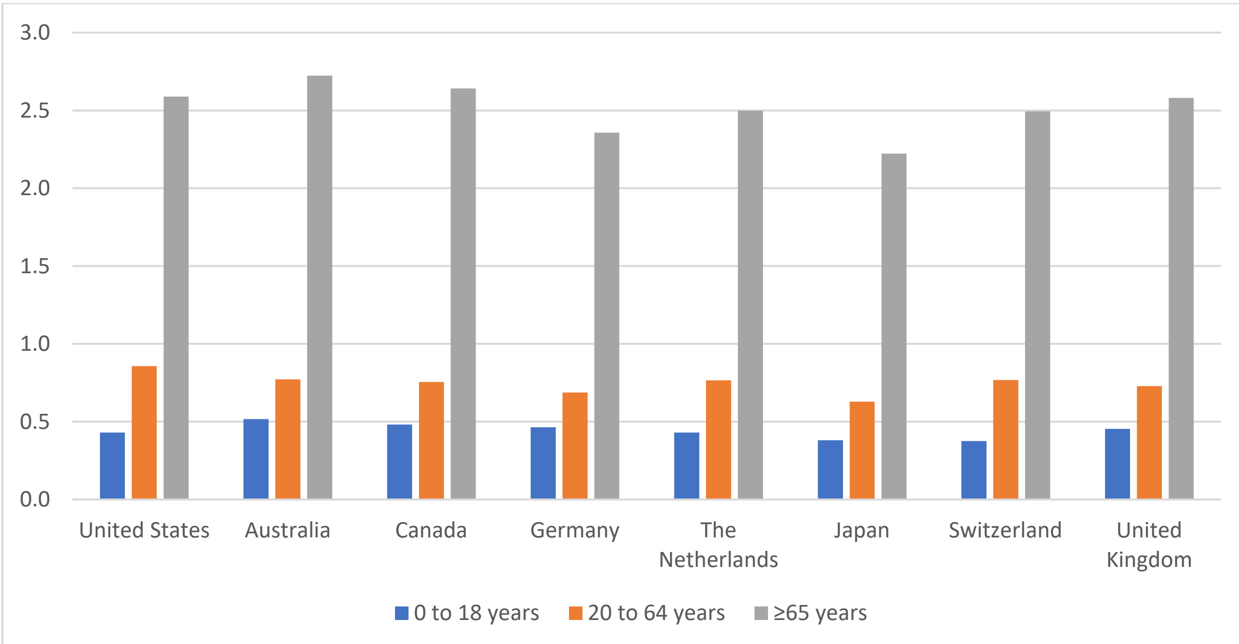
OECD. 2021. *OECD Compendium of Productivity Indicators, 2021.* Paris: OECD Publishing. Accessed September 2022. doi:https://doi.org/10.1787/f25cdb25-en.

Papanicolas, Irene, Alberto Marino, Luca Lorenzoni, Ashish Jha. 2020. "Comparison of Health Care Spending by Age in 8 High-Income Countries." *JAMA Network Open* 8: 3. doi:10.1001/jamanetworkopen.2020.14688.

The Boards of Trustees of the Federal Hospital Insurance and Federal Supplementary Medical Insurance Trust Funds. 2021. *www.cms.gov.* August 31. https://www.cms.gov/files/document/2021-medicare-trustees-report.pdf.

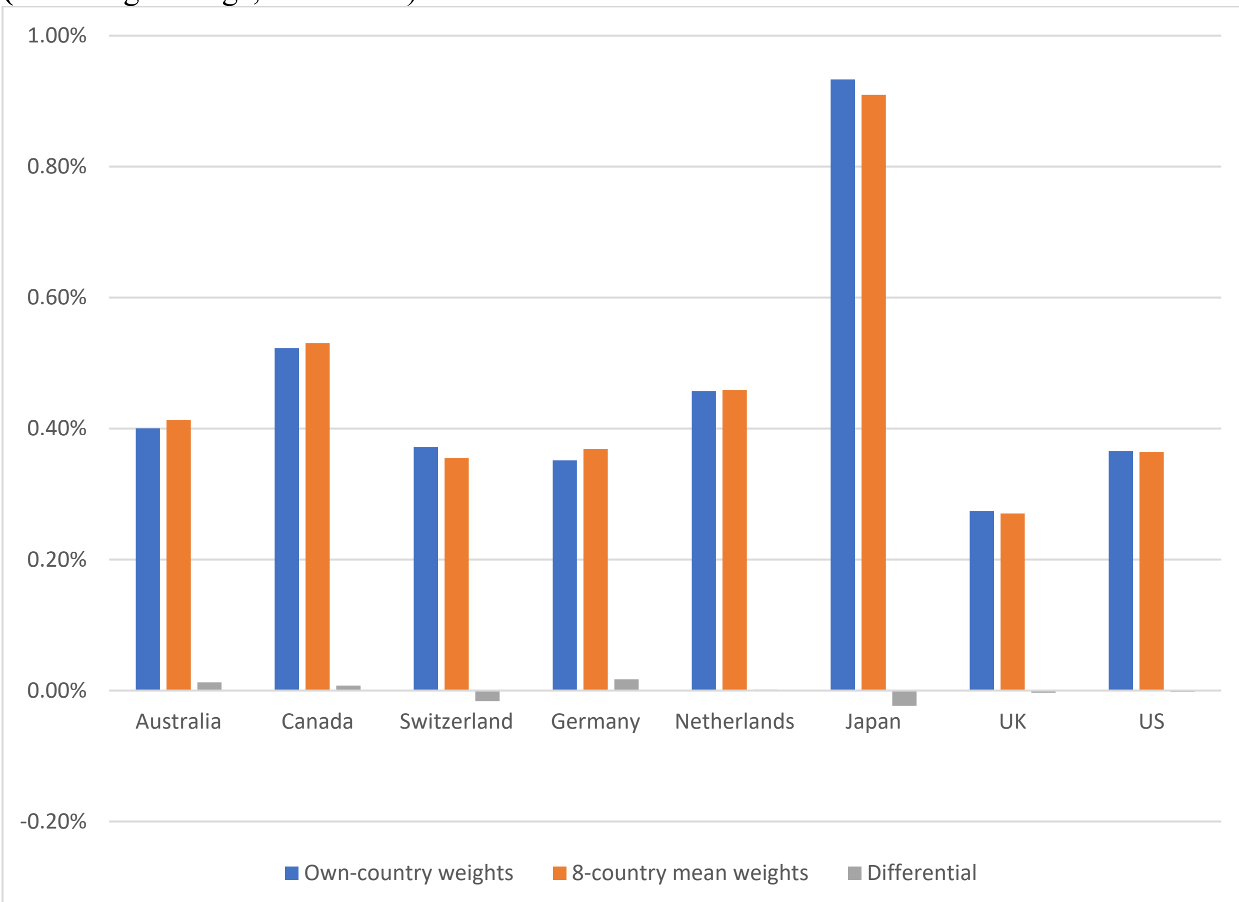
Tofighi, Davood, David P.MacKinnon, Myeongsun Yoon. 2011. "Covariances between regression coefficient estimates in a single mediator model." *British Journal of Mathematical and Statistical Psychology* 312-320. doi:10.1348/000711008X331024.
World Bank. 2021. *World Development Indicators.* Accessed August 18, 2021. https://databank.worldbank.org/indicator/NY.GDP.PCAP.CD/1ff4a498/Popular-Indicators.

Figure A1: Relative health care expenditures per capita by age cohort, shown as a ratio to mean spending



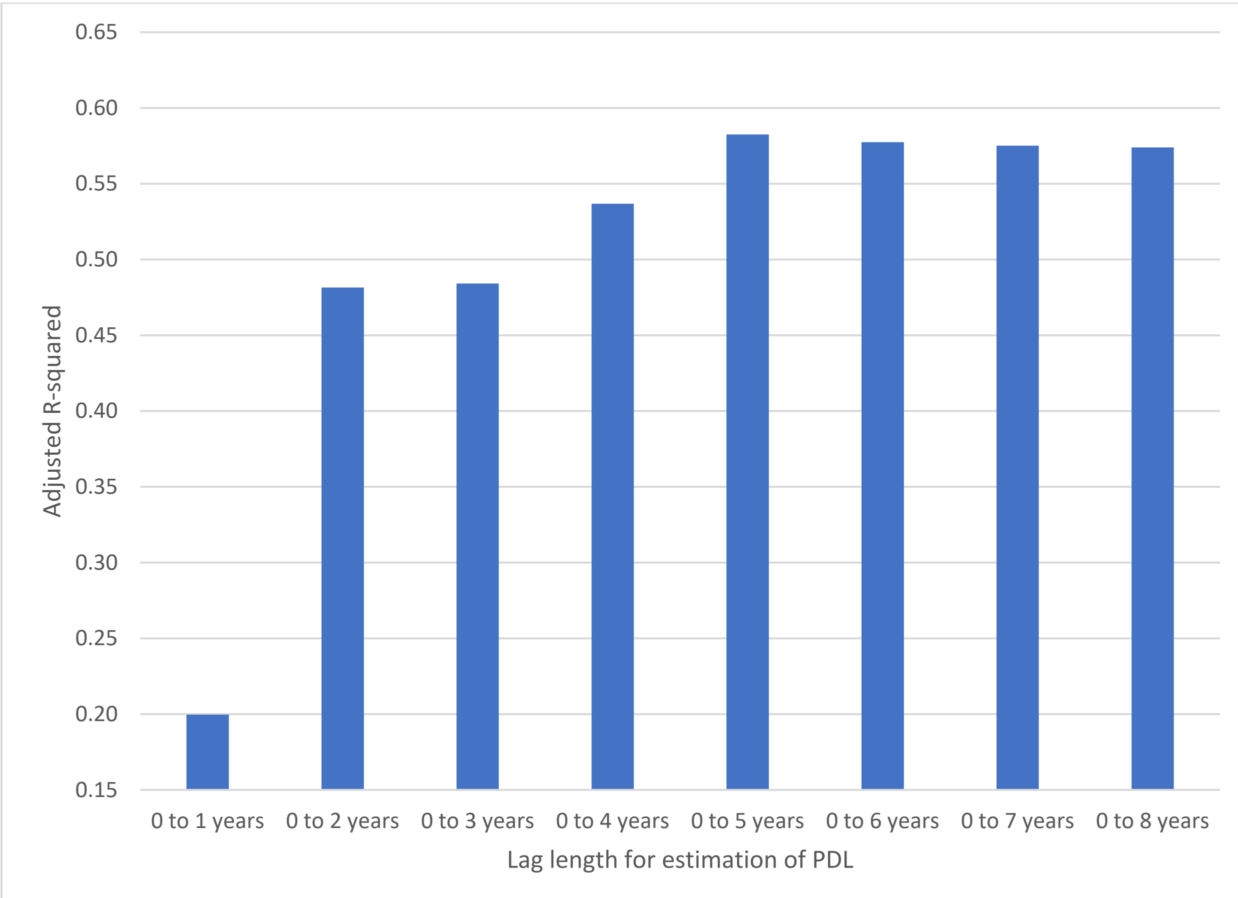
Source: Papanicolas et al (2020)

Figure A2: Contribution to health care spending growth from demographic change: own-country health spending by age versus 8-country mean health spending by age
(Percentage change, 1970-2019)



Source: Authors' calculations.

Figure A3: Variation in Adjusted R-Squared with variation in lag length of the polynomial-distributed-lag (PDL) in the estimation of Equation A2



Source: Authors’ calculations.