



Health expenditure, longevity, and child mortality: dynamic panel data approach with global data

Devdatta Ray¹ · Mikael Linden¹

Received: 12 December 2017 / Accepted: 29 August 2019 / Published online: 6 September 2019
© The Author(s) 2019

Abstract

In this study, effects of public and private health expenditures on life expectancy at birth and infant mortality are analysed on a global scale with 195 countries in the years 1995–2014. The global data set is divided into country categories according to growth in life expectancy, decrease in infant mortality rate, and level of gross national income per capita. Some new dynamic panel model estimators, argued to be more efficient with high persistence series and predetermination compared to popular but complex *GMM* estimators, show that public health expenditures are generally more health-promoting than private expenditures. However, the health effects are not as great as primary education effects. Although the new estimators provide some new and valuable information on health expenditure effects on life expectancy and infant mortality on a global scale, they do not show desired robustness.

Keywords Health expenditures · Low and high incomes · Life expectancy · Dynamic panel methods

JEL Classification I15 · I18 · H51 · C30

Introduction

Globally, health spending is highly unequal. It is even more unequally distributed than the national income of countries (Deaton 2013). Countries that spend little on health also have poorer health conditions. Although OECD countries have less than 20% of the world's population, they accounted for over 80% of world spending on health at the turn of the new century, whereas the poorest three quarters of the world's population accounted for only 7% of the world's health expenditures (*HE*). Looking across regions, Africa contains about 10% of the world's population, yet it accounted for 3% of the world's health spending. Asia and the Pacific (including China), with almost 30% of the world's population, accounted for only 4% of the world's health spending (OECD 2014; Poullier et al. 2002; WHO 2015).

✉ Devdatta Ray
leehalonen@gmail.com

¹ Department of Health and Social Management (Health Economics), University of Eastern Finland (UEF Kuopio Campus), P.O. Box 162, 70211 Kuopio, Finland

Although health spending should affect health conditions, it is important to note that the efficiency of transforming spending into better health outcomes varies significantly. Even at low levels of *HE* per capita spending, some countries achieve better health than others. This means that there may be an opportunity for public policy to also make a difference. On the contrary, among high-spending countries, sometimes additional spending bears little relationship to improvements in life expectancy. This is one reason behind the concern in wealthier countries for cost containment.

During most of the second half and especially the last decades of the twentieth century, public *HE* grew at a faster rate than the national income in developed countries. Empirical studies show that demographic factors, such as population ageing, have a positive effect on public *HE* growth, but rather of a second order, when compared with other drivers, such as income level, technology, relative prices, and institutional settings (Medeiros and Schwierz 2013; Oliviera Martins and de la Maisonneuve 2006).

The focus of this paper is on the following question: Can public and private *HE* explain the health status variable like life expectancy at birth and infant mortality across different countries? This question is answered on a global scale with 195 countries for the years 1995–2014. Our analysis concerns the health impacts of public and private *HE*. We expect them to have different effects, as their provision and demand are not distributed uniformly within the population in each country. In addition, the nature and the quality of public health services differ significantly from those of private ones, especially in the poorest countries. Thus, the second question of interest is as follows: Do increased public and private expenditures in non-high-income countries bring more improvements in health status compared to high-income countries? This is the well-known income gradient hypothesis.

This paper is divided into five sections. In the following section, we provide a short review of material relevant to our questions and of some seminal papers. The third section elucidates the data used and describes the models applied and the methods involved. The fourth section presents the results. Subsequently, the paper ends with “**Discussion and conclusions**” section, which includes discussions and conclusions on the results.

Health expenditures and health status

Background

A major share of *HE* is publicly financed (i.e. through taxes or compulsory social insurance contributions), at least in OECD countries. While some believe that this and general income level may raise public *HE* because of additional demand resulting from a decrease in the net price of care, others suggest that the public financing of *HE* serves as a restraining factor. Research has shown that income increases are not the only primary drivers of the increase in the health share of the *GDP* (Gerdtham and Johnsson 2000). An interesting possibility is that institutional factors such as the spread of insurance coverage have not only directly encouraged spending but also have induced the adoption and diffusion of new medical technologies (Acemoglu and Finkelstein 2008). However, in high-income countries, macro-economic pressures on public budgets spill over to health budgets.

One approach to reducing the fraction of public financing is to increase out-of-pocket payments or private insurance. There are, however, major problems with this substitution. First, there is a limit as to how much out-of-pocket payments can be increased if the goal of equity is concerned. Second, private insurance as a means of financing poses a problem,

because those with the highest potential expenditure also quite often have the lowest incomes. The public fraction of HE is highest in rich countries, which also often have the highest total expenditure. Private financing dominates in low-income countries, where direct out-of-pocket payments are more important than private insurance (Gerdtham and Johnsson 2000).

Focusing on the three major sources of HE —tax-financed spending, social security spending, and private spending—one sees clear differences in the structures of health care systems. South Asia is the region with the largest private sector share and virtually no reliance on social security systems. Africa and the Middle East rely heavily on private financing, but appear to have larger public tax-based (or externally supported) sources. In East Asia and the Pacific region, private spending is also high, but the public share has a significant portion in social security, driven almost exclusively by China (Hu and Ljungwall 2013; OECD 2014; Poullier et al. 2002; WHO 2014, 2015). The Americas also rely heavily on private financing, but somewhat less than other regions. It is only in Europe and partly in the OECD that health systems depend less on private financing and rely instead on significant shares of both social security and tax-based funding.

The estimates of public expenditure on health (HE_{PUB}) range from as low as 7% to almost 100% of all health spending. The wealthier and healthier countries tend to rely more heavily on public sources of funds as a share of total spending. Generally, HE_{PUB} as a share of total health spending is poorly correlated with per capita GDP , even if the correlation is statistically different from zero (Ke et al. 2011; Poullier et al. 2002). However, once again, countries and regions vary significantly along this dimension. One way of looking at this wide range of public commitment to HE is to compare countries in groups with similar per capita income levels or health outcomes.

Although private health spending (HE_{PRIV}) is overwhelmingly paid out-of-pocket, the share of private health insurance in total HE is insignificant in most countries. Prepaid private insurance accounts for more than 5% of HE_{PRIV} in only about one-third of the world's countries. In those countries where private insurance has some significance, this averages only 26% of private spending, while private spending as a whole accounts only for an average of 10% of all health spending (Liang and Mirelman 2014; Poullier et al. 2002; WHO 2015). The bulk of private spending is paid out-of-pocket at the time of service. This out-of-pocket spending accounts for a much greater share of HE in relatively poor countries than in richer ones. The high level of out-of-pocket spending or very low public spending in some low- and middle-income countries stands out as one of the most troubling areas for public health policy.

There is a growing call for bilateral and multilateral agencies to increase their financial support to the health systems in low-income and high-disease-burden countries (e.g. Ebola epidemic in Western Africa in 2014–2015). There is also a growing concern among countries that provide grants and loans that their funds are targeted effectively to the populations with most need. However, most external funding goes to countries with large populations. In other cases, countries seem to be picked out for special assistance because they are recovering from war or dealing with severe hunger. Political attachments between particular countries also play a role in this context, but general, non-targeted aid seems to be unsuccessful (Deaton 2013).

Expenditure on health naturally depends on the number of people in need of health care. This is determined by factors such as population size and age composition. Expenditure is perceived to increase considerably at older ages, as older people often require costly medical treatment due to multi-morbidities and chronic illnesses. Improvements in life expectancy may therefore lead to increases in HE if not accompanied by improvements in general health

status. The relation between life expectancy and *HE* is complicated by the fact that it is also influenced by the proximity to death, at least in the high-income countries.

Econometric studies with *HE* effects on health outcomes

Improvements in human and health capital are critical catalysts to economic growth and development in the macroeconomic literature (López-Casasnovas et al. 2005). Good health improves not only individuals' consumption and production in the short run, but also returns from investments in productive activities and education in the long run. At the macro level, investment in the health workforce and infrastructure is expected to improve health conditions and hence the health capital of the population. However, in many developing regions where resources are relatively scarce, *HE* has received less attention in government budgets. At the African regional level, Anyanwu and Erhijakpor (2007) found with a panel data fixed effect model that total *HE* was a significant contributor to health outcomes, with a 10% increase in total health care expenditure per capita resulting in a 21% and 22% decrease in under-five and infant mortality rates, respectively. Akinkugbe and Afeikhena (2006) also provided evidence that the effect of health care expenditure as a ratio of *GDP* on life expectancy, under-five mortality, and infant mortality was positive and significant in developing regions.

Filmer and Pritchett (1997) provided evidence that while health care spending impacted on child mortality, it was *not* the dominant driver of this health outcome. Factors such as education, technological change, and social capabilities have been identified by some researchers as major drivers of health outcomes rather than health care spending (Caldwell 1986; Easterly and Levine 1996; Lleras-Muney and Sherry 2008; Musgrove 1996). To understand why public spending on health has failed to have a strong effect on reducing mortality, Filmer and Pritchett (1999) stressed in their study that they were not suggesting that medical services are not (potentially) effective. They were not arguing that penicillin, immunizations, or oral rehydration therapy are ineffective as health interventions. But the impact of HE_{PUB} on health is much more complicated than the effectiveness of particular services purchased. For public spending to improve health cheaply (i.e. if money mattered), three things need to happen. Firstly, public spending is required to create effective health services. Secondly, the existence of those new public services is expected to change the total amount of effective health services consumed by the population. Thirdly, the additional services consumed must be cost-effective in improving health. If any one of these conditions is not met, the actual cost of services becomes high (Deaton 2013, chapter 9; Filmer and Pritchett 1999).

Partly contrary to these results, some studies claim to have found a clear significant impact of HE_{PUB} on health (Anand and Ravallion 1993; Baldacci et al. 2002; Bidani and Ravallion 1997; Jamison et al. 1996). For example, Bidani and Ravallion (1997) used a particular functional form to separate out the impacts of various variables on the poor and the non-poor. They found health status effects of HE_{PUB} spending for the poor but not for the non-poor. Their findings highlighted the importance of considering the incidence of the health benefits (i.e. some benefits helped best only the poor, so cuts without reallocation also fall on the poor). Note that Filmer and Pritchett (1999) also found that the impact on the poor versus the non-poor depended on the composition and efficacy of public spending.

One of the purposes of the study by Novignon et al. (2012) was to investigate the impact of total health care spending on various health outcomes after controlling for country-specific demographic structures and economic conditions. A differential analysis of public and private health care spending was also performed. The hypothesis was that there was no significant relationship between health spending and health outcomes in Sub-Saharan Africa. A further

hypothesis was that there was no significant difference in the effects of public (HE_{PUB}) and HE_{PRIV} on health outcomes. Results from the fixed and random effects models showed that a 1% increase in total HE (as % of GDP) was more likely to increase life expectancy at birth by approximately 0.6–0.7 years. Disintegrating the effect of total HE showed that a 1% increase in both HE_{PUB} and HE_{PRIV} significantly increased life expectancy at birth by about one and 0.4–0.5 years, respectively. Similar results were obtained for death rate (per 1000 people) and infant mortality rate (per 1000 live births).

For OECD countries, the HE effects on health are not directly comparable to the above low-income country results (see e.g. Barthold et al. 2014; Heijink et al. 2013; Jaba et al. 2014; Nixon and Ulmann 2006; van Baal et al. 2013). Note that private vs. public expenditure effects are analysed in very few papers. Only the papers by Cremieux et al. (2005), Lichtenberg (2000), and Or (2000) provide some information on this question. Cremieux et al. (2005) use data on Canadian provinces over the period 1975–1998. They focus on public and private spending on drugs with many additional variables with panel fixed effects regression methods. Results show that for life expectancy at birth, increases in drug spending and private spending effect are somewhat larger than public expenditure effect. Lichtenberg (2000) provides time series evidence from the US over the period 1960–1997 for life expectancy at birth in dynamic models where public and private expenditures predict life expectancy together with GDP and new molecular drug approvals. Public expenditure's short- and long-term effects are statistically significant, but private effects are not precise, especially when lagged GDP is added in the model. Or (2000) uses similar methods as Cremieux et al. (2005) but explains premature death in 21 OECD countries in 1970–1992 with total HE and with share of public HE of total HE augmented with public health, environmental factors, and GDP . Note that Heijink et al. (2013) control for vast number variables and time trends in their analysis, but HE remains a significant determinant of avoidable mortality.

Models, data, and methods

Setup

The argument that income level—either personal or GDP per capita level—determines the health conditions of individuals and the population is profound in the health economics literature. However, the heterogeneity in health status between individuals or nations, even at the same income levels, requires a more detailed relationship between health conditions and specific expenditures targeted to promote health and care provision. The distinction between public and private expenditures here is important, since the former is likely a policy variable determined by the level of $GDPc$ and the political agenda by the state and local public authorities, and the latter reflects the individual-level resources devoted to health care. Thus, both are endogenous variables in the long run. However, past findings suggest that the exogenous direct and delayed effects of HE_{PUB} on life expectancy and on infant mortality are positive and significant. Taking HE_{PRIV} as an exogenous variable is less warranted, as it is a form of derived demand (i.e. sickness and illness force people with short-run income constraints to put their money on HE).

Next, we propose dynamic panel data models to determine the levels of life expectancy at birth (LE) and infant mortality rate (IM), depending on HE_{PRIV} and HE_{PUB} . We prefer the logarithmic forms of the variables in the following dynamic panel fixed effects (FE) model.

$$\begin{aligned}
 \ln LE_{it} &= \alpha_0 + \alpha_i + \alpha_1 \ln LE_{i,t-1} + \alpha_2 \ln HE_{PRIV,it} \\
 &\quad + \alpha_3 \ln HE_{PUB,it} + \alpha_4 \ln PCR_{it} + \alpha_5 \ln RDE_{it} + \varepsilon_{i,1t} \\
 \ln IM_{it} &= \beta_0 + \beta_i + \beta_1 \ln IM_{i,t-1} + \beta_2 \ln HE_{PRIV,it} \\
 &\quad + \beta_3 \ln HE_{PUB,it} + \beta_4 \ln PCR_{it} + \beta_5 \ln FS_{it} + \varepsilon_{i,2t}
 \end{aligned} \tag{1}$$

In the first model, life expectancy is determined by private and public expenditures, and by exogenous variables of primary education rate (*PCR*) and level of R&D expenditures per capita (*RDE*). In the model for infant mortality, we replace *RDE* for food supply (*FS*). We stress that *HE* are direct means and resources to achieve good health and care among the population in the country, not the income level of the country as such. Thus, in the above models, the level of education, level of technology, and *FS* per capita refer generally (among many other similar variables) to the country's development level that sustains life expectancy and lowers infant mortality.

One-period lagged health variables, LE_{-1} and IM_{-1} , in models reflect the dynamics of health status (i.e. past health status determines the current one). Note, however, that both equations can be recursively solved for current and past values of other variables in the models and for the *starting values* of life expectancy and infant mortality (i.e. $LE_{i,0}$ and $IM_{i,0}$). These and other variables' effects on current values of LE_{it} and IM_{it} are determined by the sizes of adjustment parameters (α_1 and β_1). If they are close to but below one, the past variable values can still have large effects on current-level health status (see Eq. 3 below).

In general terms, the model captures more directly the income-driven health part of the bi-directional income–health relationship (Weil 2009, chapter 6). Income per capita, and other indicators of the living standard determine the health status of a country's population. For example, if primary schooling is missing and the *FS* per capita is low, the income level of the country is typically low and the average health status is also low. Evidently, the so-called growth process has not started or it has halted because of missing factors that are important to sustain income generation. Although the needs for health care and medication are most urgent, the resources for them are sparse, even missing, or used elsewhere.

Our main argument is that, at least for poor countries, the resources devoted to public health provision—the HE_{PUB} —are more important for the population's health status than the private expenditures. The reason for this stems from the large (income) inequalities prevailing in most poor countries supporting high incomes and HE_{PRIV} only for a small fraction of the population. The large population share of the poor can only get health benefits of public health care that is not exclusive.

Data

Annual data from 195 countries in the years 1995–2014 were collected from different sources. The life expectancy variable (*LE*) (life expectancy at birth in total years) was collected from world development indicators (World Bank 2015). Additional data for life expectancy for some countries were taken from Gapminder (2015). The infant mortality rate (*IM*) (number of deaths of less than 1-year-old infants/1000 live births) was provided by the World Bank and WHO. Total, public, and private *HE* as % shares of *GDP* (i.e. HE^S , HE_{PUB}^S , and HE_{PRIV}^S) were taken from the Global Health Expenditure Database of the WHO (2015). Further health spending data for OECD countries were taken from OECD (2015) and derived for non-OECD countries with data on *GNI* per capita (formerly *GNP* per capita). *GNIc* is the gross national income, converted to U.S. dollars using the World Bank Atlas method, divided by the mid-year population (World Bank 2016). HE_{PRIV} and HE_{PUB} were derived as

Table 1 Clusters in average growth rate of life expectancy

	Cluster	
	1	2
Cluster mean	0.00872	0.00257
Number of countries	42	153

Table 2 Clusters in average growth rate of infant mortality

	Cluster		
	1	2	3
Cluster mean	- 0.0125	- 0.0349	- 0.0615
Number of countries	56	100	39

fractions of *GNIc* with HE_{PUB}^S and HE_{PRIV}^S . The variable *PRC* (total of primary education completion rate as a % of the relevant age group) is the % of students completing the last year of primary school (World Bank 2015). R&D expenditures (*RDE*) were derived from R&D expenditure as a % of *GDPc* (World Bank 2015). Finally, we obtained the variable *FS* (kilocalories per person per day) from the World Bank (2016).

Country clusters and groups

In order to analyse effectively the public and private expenditure effects on life expectancy and infant mortality, we used the following country grouping strategy. We need different country clusters and groups to identify how public and private expenditure determines life expectancy and infant mortality that are quite heterogeneous across the sample countries and sample years. The *K*-means cluster method identified two clusters of countries with average country-specific growth rates of life expectancy in the sample period 1995–2014 (i.e. $\frac{1}{T-1} \sum_{t=2}^T \Delta \ln LE_{it}$). Table 1 reports the cluster mean values and number of cluster countries. We observe that in cluster 1 the mean growth rate of life expectancy was 3.5 times larger than in cluster 2. Typically, cluster 1 includes some of the poorest countries that have experienced significant health benefits from their care systems started in recent years.

Next, the *K*-means method was also applied to growth rates of infant mortality rates (see Table 2). Due to the heterogeneous growth rates of infant mortality, the method proposed three clusters for average growth rates of infant mortality rates ($\frac{1}{T-1} \sum_{t=2}^T \Delta \ln IM_{it}$). Here cluster 3 comprises countries that belong to both *LE* growth clusters, that is, countries whose development process started before the sample period and their rapid progress in health status can also be seen in fast-declining infant mortality rates (e.g. China, Turkey, Brasil). Cluster 1 contains some of the poorest countries but also some developed countries that have already obtained a low level of infant mortality that is not declining anymore. Thus, most cluster countries here belong to life expectancy growth cluster 2. In infant mortality cluster 2, a typical country is a rich country (i.e. European country) with relatively low growth in life expectancy, but also some non-rich countries with a rapidly rising life expectancy (e.g. India).

Finally, we divided the countries into two groups based on their average level of *GNIc* during the sample period. If a country's average *GNIc* level was below 2440 US\$ during the sample years, it belonged to group 1 (77 countries, 39.5% of countries); countries with a

Table 3 Summary statistics for life expectancy growth clusters

	$\Delta \ln LE$	LE	HE_{priv}	HE_{pub}
<i>CLUSTER 1</i>				
Mean	0.0081	60.638	39.718	60.906
SE of mean	0.0012	0.286	2.174	5.076
Standard deviation	0.035	8.287	63.016	147.119
CV	4.402	0.137	1.586	2.415
Median	0.0076	60.101	17.275	12.387
Sample size	798	840	840	840
<i>CLUSTER 2</i>				
Mean	0.0027	71.523	271.344	636.829
SE of mean	0.0002	0.141	8.285	23.621
Standard deviation	0.0145	7.815	458.309	1119.407
CV	5.297	0.109	1.689	1.757
Median	0.0027	73.601	98.058	173.977
Sample size	2907	3060	3060	3060

level higher than 2440 US\$ formed group 2 (118 countries, 60.5%). Note that in the sample, the mean income is 10085 US\$ and the median is 3298 US\$. Thus, 2440 US\$ is close to 75% of the global median income in the years 1995–2014. This means that group 1 countries are globally the poorest countries.

Summary statistics

Tables 3, 4, 5 provide detailed summary statistics in different clusters and income groups. In Table 3, clusters based on average life expectancy growth across the sample countries show that between low- and high-growth countries the difference between life expectancy is 11 years. Thus, during the sample period (1995–2014) the high level of life expectancy means less growth in life expectancy than at a lower level of life expectancy. The level of private and public HE per capita is 7–10 times larger in cluster 2 than in cluster 1.

The distributions of expenditures are skewed towards low values corresponding to $GNIc$ distributions among the global countries. Note also the large standard deviation (and CV) values showing large heterogeneity, especially in $\Delta \ln LE$ and expenditure observations.

Similar remarks are valid for the infant mortality growth clusters, but now we also observe (see Table 4), that the levels of HE across the clusters are not as large as in the above life expectancy clusters. Particularly, the level of HE_{PRIV} does not vary across the $\Delta \ln IM$ clusters and levels of IM . Thus, infant mortality seems to be independent of HE_{PRIV} . However, the level of IM clearly determines the speed of its decline (i.e. the lower the level of infant mortality rate is, the larger the decrease is). Note also that IM distributions are skewed towards low values.

Table 5 provides summary statistics for GNI per capita level groups. The most interesting result is that the rate of decrease of infant mortality ($\Delta \ln IM$) is almost the same in both $GNIc$ groups, albeit there is a huge difference between the levels of infant mortality (IM). The difference between life expectancy (LE) is 12 years, but in poor countries the growth rate of life expectancy is almost two times greater than in non-poor countries. However, a huge difference remains between the country groups in levels of HE .

Table 4 Summary statistics for infant mortality clusters

	$\Delta \ln IM$	IM	HE_{priv}	HE_{pub}
<i>CLUSTER 1</i>				
Mean	- 0.0146	40.583	234.764	420.675
SE of mean	0.0006	0.954	18.083	26.263
Standard deviation	0.0247	31.924	605.189	882.296
CV	- 1.482	0.786	2.577	2.048
Median	- 0.0161	31.551	47.969	118.399
Sample size	1064	1120	1120	1120
<i>CLUSTER 2</i>				
Mean	- 0.0341	33.754	219.487	592.067
SE of mean	0.0006	0.618	7.410	25.985
Standard deviation	0.0247	30.491	331.392	1162.122
CV	- 0.721	0.903	1.510	1.963
Median	- 0.0331	22.610	57.637	83.022
Sample size	1900	2000	2000	2000
<i>CLUSTER 3</i>				
Mean	- 0.0592	26.900	207.348	427.396
SE of mean	0.0011	1.108	9.205	22.888
Standard deviation	0.0287	28.449	257.096	799.483
CV	- 0.485	1.057	1.239	1.823
Median	- 0.0579	15.301	116.807	160.625
Sample size	741	780	780	780

Table 5 Summary statistics for *GNI per capita* level groups

	$\Delta \ln LE$	LE	$\Delta \ln IM$	IM	HE_{priv}	HE_{pub}
<i>GROUP 1</i>						
Mean	0.0052	61.846	- 0.0333	58.871	27.097	29.181
SE of mean	0.0008	0.197	0.0007	0.749	0.836	1.376
Standard deviation	0.0311	7.736	0.0277	29.412	32.827	53.021
CV	5.979	0.125	- 0.031	0.499	1.211	1.851
Median	0.0045	62.010	- 0.0315	56.851	15.741	14.050
Sample size	1463	1540	1463	1540	1540	1540
<i>GROUP 2</i>						
Mean	0.0030	73.964	- 0.0339	18.340	348.282	828.355
SE of mean	0.0020	0.129	0.0006	0.392	10.237	24.966
Standard deviation	0.0095	6.299	0.0300	19.033	497.315	1212.87
CV	3.162	0.085	- 0.887	1.037	1.428	1.464
Median	0.0027	75.001	- 0.0309	12.701	169.356	309.718
Sample size	2242	2360	22420	2360	2360	2360

Generally, these findings with respect to our dynamic panel data models mean that we do not expect much success in infant mortality modelling, as the variable seems to be insensitive to the level of private expenditure. However, the large within heterogeneity in clusters and income groups masked by the above location statistics need a country-level fixed effect (*FE*) modelling approach that can provide some valuable results across the cluster countries.

Dynamic panel data models

Consider the following dynamic fixed (*FE*) or random (*RE*) effect model (for more details, see Pesaran 2015, chapters 26–27):

$$y_{it} = \alpha_i + \lambda y_{i,t-1} + \beta' x_{it} + \mu_{it}, \quad i = 1, 2, \dots, N \text{ and } t = 1, 2, \dots, T. \quad (2)$$

Typically, regressors, x_{it} , are assumed to be *strictly exogenous* (i.e. $E[\mu_{it}|x_{it}] = 0$) for all i and t . However, the assumption of strict exogeneity is not valid by construction for lagged dependent variable $y_{i,t-1}$, since even if we assume that $E[\mu_{it}\alpha_i] = E[\mu_{it}y_{i0}] = 0$, the *FE/RE* demeaning term $E[\mu_{it}\bar{y}_{i,-1}] \neq 0$ will not vanish for short panels. In the process without regressors x_{it} , this will cause bias for the *FE* or *RE* estimators of λ with its size depending on the true value of $|\lambda| < 1$ and the length of panels (Nickell 1981; Pesaran 2015, p. 679)

$$plim_{N \rightarrow \infty}(\hat{\lambda}_{FE/RE} - \lambda) = -\frac{(1 + \lambda)}{T} + O(T^{-2}).$$

The bias is order of $1/T$ and vanishes when $T \rightarrow \infty$. For example, when λ is close to 1 (the non-stationary case) and $T=20$, the bias is close to -0.1 . Note that if regressors x_{it} are included in the model, the size of bias for λ and β depends on the correlation between $y_{i,t-1}$ and x_{it} . If regressors x_{it} are only weakly exogenous (i.e. allowing for feedbacks from $\mu_{i,t-1}$) or if they are endogenous variables, the *FE/RE* bias for β is still present, even if no lagged dependent variable is found in the model.

The generic problems of the above dynamic panel model can be seen when we solve for y_{it} recursively from the initial values of y_{i0}

$$y_{it} = \lambda^t y_{i0} + \sum_{j=0}^{t-1} \lambda^j \beta' x_{i,t-j} + \frac{1 - \lambda^t}{1 - \lambda} \alpha_i + \sum_{j=0}^{t-1} \lambda^j \mu_{i,t-j}, \quad t = 1, 2, \dots, T \quad (3)$$

When λ is close to one, initial values y_{i0} and unit-specific effects α_i have large and permanent effects on the y_{it} observations determining the properties of dynamic panel data model estimators. As the process for $y_{i,t-1}$ has a similar presentation, we obtain, abstracting from terms for regressors and errors:

$$y_{it} - y_{i,t-1} = (\lambda^t - \lambda^{t-1})y_{i0} + \lambda^{t-2}\alpha_i.$$

This shows that initial effects, but not necessarily the unit-specific α_i effects, have a small role in determining the one-period differenced values of y_{it} . Subsequently, the following difference model has also been popular to eliminate the unit-specific effects on λ and β estimates:

$$\Delta y_{it} = \lambda \Delta y_{i,t-1} + \beta' \Delta x_{it} + \Delta \mu_{it}. \quad (4)$$

However, this will not solve the (OLS) estimation problems for the model parameters, since

$$E[\Delta y_{i,t-1} \Delta \mu_{it}] = E[\lambda \Delta \mu_{i,t-1} \Delta \mu_{it}] \neq 0.$$

$$\text{Because of } E[\Delta\mu_{i,t-s}\Delta\mu_{it}] = \begin{cases} 2\sigma_u^2, & \text{for } s = 0 \\ -\sigma_u^2, & \text{for } s = 1 \\ 0, & \text{for } s > 1 \end{cases}$$

we need at least two-period lagged values of $y_{i,t-j}$ and $\Delta y_{i,t-j}$ ($j \geq 2$) that do not correlate with $\Delta\mu_{it}$ (but correlate with $\Delta y_{i,t-1}$). We can use them as instruments for $\Delta y_{i,t-1}$ as long as $\lambda < 1$, but as $\lambda \rightarrow 1$, we face the weak instrument problem for $y_{i,t-2}$ because $E[y_{i,t-2}, \Delta y_{i,t-1}]$ depends on the size of λ (for more details, see Pesaran 2015, p. 682).

The short panel bias of *FE/RE* and the efficiency problem of the *IV* approach for the first difference model started the search for *IV/GMM*-type estimators, leading to consistent and more efficient estimators like *GMM* estimators by Arellano and Bond (1991), Ahn and Schmidt (1995), and Blundell and Bond (1998). These surprisingly popular methods are extremely complex estimators, which are unbiased and efficient only when no residual serial correlation is found, the dynamic lag order of the model is correctly specified, we have strictly exogenous regressors, no correlation is found between explanatory variables and unit-specific effects a_i , errors are homoscedastic, the sample length is small (i.e. $T/n \rightarrow 0$ convergence), low autocorrelation is present in endogenous series, and the problem of weak or too many instruments is not present (see e.g. Dang et al. 2015; Gouriéroux et al. 2010; Hahn et al. 2007; Kiviet et al. 2017).

Empirical drawbacks of the *IV/GMM* agenda have led to a large group of alternative estimators that have tried in several different ways to correct for $1/T$ time series bias. Chudik and Pesaran (2015) divide this literature into the following broad categories: (i) analytical corrections based on an asymptotic bias formula (Bruno 2005; Bun 2003; Bun and Carree 2005, 2006; Bun and Kiviet 2003; Hahn and Kuersteiner 2002; Hahn and Moon 2006; Kiviet 1995, 1999), (ii) bootstrap and simulation-based bias corrections (Everaert and Ponzi 2007; Phillips and Sul 2003, 2007), and (iii) other methods, including jackknife bias corrections (Dhaene and Jochmans 2012) and the recursive mean adjustment correction procedures (So and Shin 1999). In addition, some methods have been proposed on long differences (Hahn et al. 2007; Han and Phillips 2013; Han et al. 2014), on forward filtering (Keane and Runkle 1992; Keane and Neal 2016; Pesaran 2015, chapter 27.2), and on the transformed likelihood method (Hayakawa and Pesaran 2015; Pesaran 2015, chapter 27.6).

In the following, we take methods that are planned to be robust enough against near unit-root case and avoid strict exogeneity assumption. This means that we use a long difference *IV* method, LD_{IV} , proposed by Hahn et al. (2007) as well as the Keane–Runkle estimator (1992), which allows for predetermined variables as instruments.

The LD_{IV} technique uses long differencing, i.e. $\Delta_k y_t = y_t - y_{t-k}$ with $k = 2, 3, 4, \dots$, instead of first differencing and iterated two-stage least square (2SLS) in estimating persistent dynamic models with a short time dimension. The LD_{IV} estimator uses lagged levels of the regressors (including $y_{i,t-k-1}$) and the residuals as instruments. The setup for the model is (Hahn et al. 2007, pp. 586–587; Huang and Ritter 2009, p. 269):

$$\Delta_k y_{it} = \lambda \Delta_k y_{it-1} + \beta' \Delta_k x_{it} + \Delta_k \mu_{it} \tag{5}$$

where we can use $y_{i,t-k-1}, x_{i,t-k}$ (if strictly exogenous or predetermined) as instrument variables. After obtaining 2SLS estimates for model 5 we calculate the residuals

$$y_{i,t-1} - \hat{\lambda} y_{i,t-2} - \hat{\beta}' x_{i,t-1}, \dots, \text{ and } y_{i,t-k} - \hat{\lambda} y_{i,t-k-1} - \hat{\beta}' x_{i,t-k}.$$

Next we use these as additional instrumental variables with $y_{i,t-k-1}, x_{i,t-k}$ to estimate (5) once again. This is the first iteration. Next new 2SLS estimations are then further iterated via the new results. Typically, less than five iterations are sufficient for convergence.

The Keane and Runkle (1992) estimator (KR_{PRE}) uses the idea of forward filtering or decomposition from the time-series literature to improve the efficiency of the estimates when the error contains some form of serial correlation. Under Cholesky transformation, the orthogonality conditions implied by predetermination are maintained (Keane and Neal 2016). In practice, a key feature of the approach is to use only one or two lags of the predetermined variables as instruments rather than all available lags back to the first period like in complex GMM estimation. Keane and Runkle assume that x_{it} are predetermined, in the sense that $E[x_{is}\mu_{it}] = 0$, for $t \geq s$. This is a natural approach in this context where public and private HE in current and previous periods drive life expectancy but not necessarily vice versa. However, life expectancy targets or its unobserved determinants in coming periods $t+i$ ($i=1,2,\dots$) will affect public and private HE in the future (i.e. $E[x_{it}\mu_{it}] \neq 0$, for $r>t$). Note that in the first difference model x_{it} is correlated with $\mu_{i,t-1}$ because x_{it} is predetermined but not strictly exogenous. However, $y_{i,t-2}$ and $x_{i,t-1}$ are now valid instruments.

In the Keane–Runkle method, the model 2 has a general covariance specification for $v_{it} = \alpha_i + \mu_{it}$. That is, $E[vv'] = I_N \otimes \Sigma$, where v is a stacked $NT \times 1$ vector of $v_i = (v_{i1}, v_{i2}, \dots, v_{iT})'$ and $\Sigma = E[v_i.v_i']$. To implement the KR_{PRE} estimator, we need an estimate for Σ . It is obtained from consistent preliminary 2SLS/IV estimation of model 2 using the instruments Z to obtain the 2SLS/IV residuals \hat{v}_i , and $\hat{\Sigma} = \frac{1}{N} \sum_{i=1}^N \hat{v}_i.\hat{v}_i'$. Note that a similar two-step procedure can be applied also for difference model 4.

Estimation results

Life expectancy model

Tables 6 and 7 provide the dynamic panel data model estimation results for $\ln LE$ based on standard FE , weighted $FE_{W/TR}$ with trend, LD_{IV} , and KR_{PRE} estimation methods. These methods have validity in this context, since in preliminary data analysis the panel data non-stationarity tests (not reported) supported the stationarity alternative. The FE methods assume that all explanatory variables are strictly exogenous, but in LD_{IV} and KR_{PRE} this erroneous assumption is noticed (i.e. lagged $\ln LE$ is not exogenous), and we need IV methods to overcome this problem. In addition, in KR_{PRE} we use a less restricted approach on variables $\ln LE_{-1}$, $\ln HE_{PRIV}$, and $\ln HE_{PUB}$ when we assume that they are predetermined with periods $t \geq s$. Thus, the FE approach is expected to be biased in short panels with $1/T$ rate. However, using a sample length of $T=19$, its role is less important than the evident sample heterogeneity and trending behavior of life expectancy panels. To obtain greater robustness in estimation, we use trend variable and cross-section weights in $FE_{W/TR}$ estimation (i.e. we estimate the model with a feasible GLS specification, assuming the presence of cross-section heteroscedasticity). This has a large effect on the point estimates of model parameters, but their qualitative effects remain the same as in standard FE estimation. In the following tables below 10% level p values with coefficient estimate values are marked with bold font. In the high life expectancy growth rate cluster (cluster 1), HE_{PRIV} have imprecise, statistically non-significant effects on life expectancy, but HE_{PUB} and primary education rate ($\ln PCR$) predict it with correct signs with 10% level significance. In the low life expectancy growth rate cluster (cluster 2), private expenditures also have some significance along with R&D expenditures. Note that education effects on life expectancy are generally larger than HE or R&D effects. Note that DW values refer to a Durbin–Watson type test statistics calculated on the stacked set of pooled model

Table 6 *lnLE* models with $\Delta lnLE$ clusters (*p* values in parentheses)

Cluster 1 N = 42, T = 19	FE ¹	FE ¹ _{WTR}	LD ^{2,3} _{IV}	KR ⁴ _{PRE}
<i>C</i>	1.499 (0.053)	0.539 (0.001)	–	2.800 (0.000)
<i>lnLE</i> ₋₁	0.597 (0.004)	0.862 (0.000)	0.597 (0.029)	0.194 (0.001)
<i>lnHE</i> _{priv}	0.006 (0.283)	– 0.001 (0.248)	0.003 (0.477)	– 0.049 (0.012)
<i>lnHE</i> _{pub}	0.005 (0.030)	0.001 (0.073)	0.005 (0.250)	0.061 (0.000)
<i>lnPCR</i>	0.028 (0.094)	0.005 (0.001)	0.021 (0.205)	0.092 (0.001)
<i>lnRDE</i>	0.009 (0.077)	0.0006 (0.194)	0.011 (0.147)	– 0.002 (0.867)
<i>DW</i> value	2.44	2.36	–	–
Cluster 2 N = 153, T = 19	FE ¹	FE ¹ _{WTR}	LD ^{2,3}	KR ⁵ _{PRE}
<i>C</i>	1.089 (0.024)	0.512 (0.000)	–	3.381 (0.000)
<i>lnLE</i> ₋₁	0.727 (0.000)	0.875 (0.000)	0.709 (0.001)	0.122 (0.036)
<i>lnHE</i> _{priv}	0.002 (0.111)	0.0007 (0.000)	0.0024 (0.103)	0.012 (0.000)
<i>lnHE</i> _{pub}	0.0004 (0.005)	0.0004 (0.006)	0.0025 (0.002)	0.009 (0.001)
<i>lnPCR</i>	0.008 (0.002)	0.003 (0.000)	0.009 (0.158)	0.052 (0.000)
<i>lnRDE</i>	0.002 (0.007)	0.0004 (0.000)	0.018 (0.132)	0.004 (0.029)
<i>DW</i> value	1.63	1.73	–	–

¹SEs calculated with White’s cross-section method

²SEs adjusted for cross-section clusters

³Instruments: *lnLE*(– 6), *lnHE*_{priv}(– 5), *lnHE*_{pub}(– 5), *lnPCR*(– 5), *lnRDE*(– 5), *res*(– 1 to – 4)

⁴Instruments: $\Delta lnLE$ (– 1), $\Delta lnHE$ _{priv}(– 1 to – 2), $\Delta lnHE$ _{pub}(– 1 to – 2), *lnPCR*, *lnRDE*, constant

⁵Instruments: $\Delta lnLE$ (– 1), *lnHE*_{priv}(– 1 to – 2), *lnHE*_{pub}(– 1 to – 2), *lnPCR*, *lnRDE*, constant

residuals. Thus the reported *DW* values are indicative, i.e. when the values are close to 2 we take it as a sign of appropriate model lag choice.

Results with the *LD*_{IV} method are not promising. A search over a suitable difference length provided a five-period difference approach. Models with a trend variable were also considered, but they provided non-significant and partly wrongly signed estimates. Results with the *KR*_{PRE} method were quite different from the above-mentioned ones. This method seems to scale down the adjustment coefficient but provides larger point estimates for other model parameters. Note that *lnHE*_{PRIV} gets a negative coefficient estimate for cluster 1. *SE*s of *KR*_{PRE} estimates are not corrected for possible heteroscedasticity. Also, an instrument validity test is needed for *LD*_{IV} and *KR*_{PRE} estimations before their full validity can be evaluated. We do not report any model diagnostic values like *R*² or *F*-test and residual diagnostics except *DW* values for *FE* models, as the different model estimators are based on different methods and provide statistics that are not comparable.

Table 7 provides the *lnLE* model estimates in *GNIc*-level groups. Results with *FE* methods are comparable with the results above. *HE*_{PUB} effects are positive and significant on life

Table 7 *lnLE* models with *GNIPC* groups (*p* values in parentheses)

Group 1 N = 77, T = 19	FE ¹	FE ¹ _{WTR}	LD ^{2,3} _{IV}	KR ⁵ _{PRE}
<i>C</i>	1.309 (0.022)	0.503 (0.000)	–	0.013 (0.122)
<i>lnLE</i> _{–1}	0.655 (0.000)	0.872 (0.000)	0.617 (0.001)	0.025 (0.532)
<i>lnHE</i> _{priv}	0.004 (0.029)	– 0.0001 (0.177)	0.003 (0.212)	0.043 (0.000)
<i>lnHE</i> _{pub}	0.004 (0.014)	0.001 (0.009)	0.004 (0.260)	0.014 (0.065)
<i>lnPCR</i>	0.023 (0.019)	0.006 (0.000)	0.021 (0.060)	0.104 (0.001)
<i>lnRDE</i>	0.006 (0.028)	0.001 (0.000)	0.006 (0.068)	0.012 (0.044)
<i>DW</i> value	2.42	1.84	–	–
Group 2 N = 118, T = 19	FE ¹	FE ¹ _{WTR}	LD ^{2,4} _{IV}	KR ⁶ _{PRE}
<i>C</i>	0.796 (0.000)	0.426 (0.000)	–	– 0.021 (0.032)
<i>lnLE</i> _{–1}	0.806 (0.000)	0.899 (0.000)	0.781 (0.000)	0.746 (0.000)
<i>lnHE</i> _{priv}	0.001 (0.152)	0.0005 (0.004)	0.021 (0.351)	0.012 (0.033)
<i>lnHE</i> _{pub}	0.005 (0.01)	0.0006 (0.003)	0.005 (0.026)	0.063 (0.061)
<i>lnPCR</i>	– 0.000 (0.970)	0.000 (0.671)	– 0.002 (0.897)	– 0.026 (0.152)
<i>lnRDE</i>	0.002 (0.000)	0.0003 (0.001)	0.0014 (0.001)	– 0.024 (0.069)
<i>DW</i> value	1.84	1.69	–	–

¹SEs calculated with White’s cross-section method

²SEs adjusted for cross-section clusters

³Instruments: *lnLE*(– 6), *lnHE*_{priv}(– 5), *lnHE*_{pub}(– 5), *lnPCR*(– 5), *lnRDE*(– 5), *res*(– 1 to – 4)

⁴Instruments: *lnLE*(– 16), *lnHE*_{priv}(– 15), *lnHE*_{pub}(– 15), *lnPCR*(– 15), *lnRDE*(– 15), *res*(– 1 to – 14)

⁵Instruments: Δ *lnLE*(– 1), *lnHE*_{priv}(– 1 to – 2), *lnHE*_{pub}(– 1 to – 2), *lnPCR*, *lnRDE*, constant

⁶Instruments: Δ *lnLE*(– 1), Δ *lnHE*_{priv}(– 1), Δ *lnHE*_{pub}(– 1), *lnPCR*, *lnRDE*, constant

expectancy, but the size of the effects is less than for primary education. Surprisingly, the education effects are non-significant for non-poor countries. Results with the *LD*_{IV} method are close to standard *FE* results, but non-significant *HE* effects are obtained. *KR*_{PRE} estimation results are once again in their own category in income level group 1. This depends partly on the demeaning of variables that provided the best results. All but adjustment coefficients are significant, and the rest are comparable with *FE* and *LD* coefficients when these are solved for long-run presentation. However, for income group 2, although estimated with demeaned data, results are quite different. In general, these results—and many others not represented here—show that the *KR*_{PRE} method has some stability and robustness problems.

Infant mortality model

These problems are clearly evident for infant mortality model estimation, as the country-level time series have *AR*(2) presentation more likely than *AR*(1) models. Although the *KR*_{PRE}

method needs consistent IV estimation in the first stage, which runs easily into problems when the adjustment coefficient sum is close to one, the second-stage estimation will not correct these but also runs into problems. For these reasons, we did not provide results on $lnIM$ models with $\Delta lnIM$ clusters and take a critical stance on KR_{PRE} estimation in $\Delta lnIM$ clusters.

Table 8 provides FE and LD_{IV} results for $lnIM$ models with $\Delta lnIM$ clusters. We observe that LD_{IV} based on a long difference works with $AR(1)$ presentation. Estimation results show that increases in HE_{PUB} reduce infant mortality, but the FS has an infant-reducing effect only in cluster 3, where the decline in infant mortality is largest. Surprisingly, HE_{PRIV} have an increasing infant mortality effect in this cluster. In other clusters, the effects are imprecise.

Table 9 provides results on $lnIM$ models with $GNIC$ groups and also includes some KR_{PRE} estimation results. Generally, results in Table 9 are not favorable for HE variables. Public expenditures reduce infant mortality in non-poor countries, but similar effects are not found for private expenditures in either income group. Education effects reducing infant mortality are clearly present in Table 9 compared to Table 8. FS effects are mortality-reducing in poor countries.

Generally, we observe that different methods provide quite varying results on point estimates, but clear qualitative accordance is found in the signs of coefficient estimates across the estimation methods. The LD_{IV} method provides results that are comparable to FE results, but KR_{PRE} produced results that are less clear. Nevertheless, we calculate long-run estimates or elasticities for HE variables from estimation results at 10% or below the significance level in the above tables. Note that we did not calculate SEs of the long-run estimates. That can be done with a delta method. Table 10 provides the sum-up. It shows that across the different data configurations and model estimations, 22 HE_{PUB} coefficient estimates from 36 possible ones were significant and their mean of long-run elasticities is 0.0364 for life expectancy and -0.223 for infant mortality. Thus, if public expenditures increase by 10%, they will increase life expectancy by 0.36% and reduce infant mortality by 2.31%. For private expenditures, we find 0.23% and -0.83% mean elasticities for a much smaller number of 10%-level significant estimates.

These estimates and related ones in the literature imply that HE_{PUB} are more effective to improve health status than private expenditures for a large number of countries. However, we cannot say that this is also true for the poorest countries in our sample (see Table 6: cluster 1, and Tables 7 and 9: $GNIC$ group 1), since we have too few estimates to prove this. Some results with a larger data set (34 OECD countries in the years 1970–2012) and different methods have shown that public expenditure effects in non-poor countries dominate the private ones (see e.g. Linden and Ray 2017). At the moment, partly because first-class data on life expectancy and infant mortality from the poorest countries are not available, the question with poor countries is not fully answerable.

Discussion and conclusions

The HE effects on life expectancy at birth and infant mortality were analyzed with dynamic panel data models for 195 countries in the years 1995–2014. The target questions of ‘Can public and private HE explain the health status variables like life expectancy at birth and infant mortality across different countries?’ and ‘Do increased public and private expenditures in non-high-income countries bring more improvements in health status compared to high-

Table 8 *lnIM* models with $\Delta lnIM$ clusters (*p* values in parentheses)

Cluster 1 N = 56, T = 18	FE ¹	FE ¹ _{W/TR}	LD ^{2,3} _{IV}
<i>C</i>	– 0.106 (0.475)	0.060 (0.047)	–
<i>lnIM</i> _{–1}	1.501 (0.000)	1.698 (0.000)	0.977 (0.001)
<i>lnIM</i> _{–2}	– 0.542 (0.000)	– 0.739 (0.000)	–
<i>lnHE</i> _{pub}	– 0.003 (0.0002)	– 0.0005 (0.254)	0.002 (0.301)
<i>lnHE</i> _{priv}	– 0.005 (0.000)	– 0.0007 (0.017)	– 0.004 (0.018)
<i>lnPCR</i>	– 0.006 (0.094)	– 0.0007 (0.336)	0.004 (0.567)
<i>lnFS</i>	0.036 (0.032)	0.011 (0.003)	0.046 (0.028)
<i>DW</i> value	2.35	2.44	–
Cluster 2 N = 100, T = 18	FE ¹	FE ¹ _{W/TR}	LD ^{2,4} _{IV}
<i>C</i>	0.210 (0.021)	0.126 (0.000)	–
<i>lnIM</i> _{–1}	1.079 (0.000)	1.491 (0.000)	0.975 (0.001)
<i>lnIM</i> _{–2}	– 0.119 (0.492)	– 0.532 (0.000)	–
<i>lnHE</i> _{priv}	– 0.003 (0.299)	– 0.0008 (0.101)	– 0.004 (0.345)
<i>lnHE</i> _{pub}	– 0.006 (0.001)	– 0.0007 (0.046)	– 0.004 (0.291)
<i>lnPCR</i>	– 0.020 (0.005)	– 0.004 (0.005)	– 0.018 (0.031)
<i>lnFS</i>	0.0009 (0.941)	0.002 (0.313)	0.003 (0.108)
<i>DW</i> value	2.20	2.47	–
Cluster 3 N = 39, T = 18	FE ¹	FE ¹ _{W/TR}	LD ^{2,5} _{IV}
<i>C</i>	0.444 (0.019)	0.186 (0.000)	–
<i>lnIM</i> _{–1}	1.215 (0.000)	1.657 (0.061)	1.012 (0.000)
<i>lnIM</i> _{–2}	– 0.225 (0.088)	– 0.677 (0.000)	–
<i>lnHE</i> _{priv}	0.005 (0.081)	0.002 (0.078)	0.016 (0.020)

Table 8 continued

Cluster 3 N = 39, T = 18	FE ¹	FE ¹ _{W/TR}	LD ^{2,5} _{IV}
<i>lnHE_pub</i>	– 0.007 (0.005)	– 0.002 (0.070)	– 0.006 (0.292)
<i>lnPCR</i>	– 0.034 (0.000)	– 0.007 (0.126)	– 0.050 (0.000)
<i>lnFS</i>	– 0.038 (0.075)	– 0.014 (0.0006)	– 0.038 (0.266)
DW value	2.30	2.49	–

¹SEs calculated with White's cross-section method

²SEs adjusted for cross-section clusters

³Instruments: $\ln IM(-4)$, $\ln HE_priv(-3)$, $\ln HE_pub(-3)$, $\ln PCR(-3)$, $\ln FS(-3)$, $res(-1 \text{ to } -2)$. Model includes a trend

⁴Instruments: $\ln IM(-16)$, $\ln HE_priv(-15)$, $\ln HE_pub(-15)$, $\ln PCR(-15)$, $\ln FS(-15)$, $res(-1 \text{ to } -14)$

⁵Instruments: $\ln IM(-6)$, $\ln HE_priv(-5)$, $\ln HE_pub(-5)$, $\ln PCR(-5)$, $\ln FS(-5)$, $res(-1 \text{ to } -4)$

income countries?' were analyzed with country clusters sensitive to life expectancy growth rates and rates of decrease in infant mortality, and with low and high *GNI* per capita levels.

The first question gets an affirmative answer, and obtained results show that HE_{PUB} are generally more health-promoting than private expenditures. However, either of the HE effects is not as large as primary education effects. We were not able to give a positive answer to the second question. In country clusters and groups identifying poor countries, positive public expenditure effects are found but private expenditure effects are either non-significant or of comparable size to public effects. This outcome is partly sensitive to the methods used. The new dynamic panel model estimators introduced in this context, LD_{IV} and KR_{PRE} estimators, are not robust enough to provide an answer to the second question. We can observe like Kiviet et al. (2017, pp. 46–48) on *GMM* estimators that 'However, not too much is known yet about the actual accuracy in practical situations on the abundance of different not always asymptotically equivalent implementations of estimators and test procedures', and 'Our results demonstrate that, especially under particular unfavorable settings, there is great urge for developing more refined inference procedures for structural dynamic panel data models' to also be valid on LD_{IV} and KR_{PRE} estimators used here. For the LD_{IV} , we need some guidance on how to determine the long difference length. KR_{PRE} needs some more robust developments under error heterogeneity and longer *AR* settings than *AR*(1).

The assumptions of variable strict exogeneity and predetermination made in the analysis are not harmless. Both the LD_{IV} and KR_{PRE} methods allow for settings where HE are endogenous (i.e. life expectancy and infant mortality determine the levels of private and public expenditures). This is not only the starting value or initial-level problem, but reasonable arguments can support the idea that the current levels of health status also determine HE . Typically, these arguments rest on the population's age structure and on the 'healthier are wealthier' type of argument where a specific distinction is not made between private and public expenditures. On the contrary, our approach stressed the fact the HE_{PUB} are a policy-driven variable in the short run (i.e. some sort of exogeneity is a natural starting point in analysis). Some model estimations were conducted with endogenous HE , but they did not provide any results beyond what we have presented already. However, this important question must be addressed in detail in future research.

Table 9 *lnIM* models with *GNIC* groups (*p* values in parentheses)

Group 1 N = 77, T = 18	FE ¹	FE ¹ _{W/TR}	LD ^{2,3} _{IV}	KR ⁵ _{PRE}
<i>C</i>	0.458 (0.004)	0.098 (0.001)	–	– 0.004 (0.049)
<i>lnIM</i> _{–1}	0.970 (0.000)	1.610 (0.000)	0.980 (0.000)	0.819 (0.000)
<i>lnIM</i> _{–2}	0.0134 (0.946)	– 0.612 (0.000)	–	–
<i>lnHE</i> _{priv}	– 0.002 (0.591)	– 0.0005 (0.247)	0.007 (0.014)	– 0.026 (0.247)
<i>lnHE</i> _{pub}	– 0.001 (0.401)	0.0004 (0.090)	– 0.001 (0.767)	– 0.007 (0.415)
<i>lnPCR</i>	– 0.018 (0.0016)	– 0.0019 (0.087)	v0.088 (0.092)	– 0.003 (0.415)
<i>lnFS</i>	– 0.045 (0.009)	– 0.004 (0.040)	– 0.043 (0.069)	– 0.431 (0.122)
<i>DW</i> value	2.14	2.53	–	–
Group 2 N = 118, T = 18	FE ¹	FE ¹ _{W/TR}	LD ^{2,4} _{IV}	KR ⁶ _{PRE}
<i>C</i>	– 0.044 (0.487)	0.013 (0.584)	–	– 0.002 (0.034)
<i>lnIM</i> _{–1}	1.397 (0.000)	1.615 (0.000)	0.986 (0.000)	0.889 (0.000)
<i>lnIM</i> _{–2}	– 0.417 (0.000)	– 0.635 (0.000)	–	–
<i>lnHE</i> _{priv}	0.0001 (0.927)	– 0.0008 (0.105)	0.001 (0.878)	– 0.002 (0.834)
<i>lnHE</i> _{pub}	– 0.006 (0.0005)	– 0.0007 (0.352)	– 0.009 (0.063)	– 0.019 (0.045)
<i>lnPCR</i>	– 0.010 (0.073)	0.0014 (0.381)	– 0.035 (0.081)	0.001 (0.885)
<i>lnFS</i>	0.019 (0.018)	0.004 (0.201)	0.059 (0.035)	0.026 (0.055)
<i>DW</i> value	2.42	2.43	–	–

¹SEs calculated with White’s cross-section method

²SEs adjusted for cross-section clusters

³Instruments: *lnIM*(– 6), *lnHE*_{priv}(– 5), *lnHE*_{pub}(– 5), *lnPCR*(– 5), *lnFS*(– 5), *res*(– 1 to – 4). Model includes a trend

⁴Instruments: *lnIM*(– 16), *lnHE*_{priv}(– 15), *lnHE*_{pub}(– 15), *lnPCR*(– 15), *lnFS*(– 15), *res*(– 1 to – 14). Model includes a trend

⁵Instruments: *lnIM*(– 2), *lnHE*_{priv}(– 1 to – 2), *lnHE*_{pub}(– 1 to – 2), *lnPCR*, *lnRDE*, constant, (difference model)

⁶Instruments: *lnIM*(– 2), *lnHE*_{priv}(– 1 to – 2), *lnHE*_{pub}(– 1 to – 2), *lnPCR*, *lnRDE*, constant, (difference model)

Table 10 Long-run elasticities

Variables	Number of significant estimates	LR-elasticity estimate range	Mean LR-elasticity
<i>lnHE</i> _{priv} → <i>lnLE</i>	7	[– 0.0607, 0.0560]	0.0229
<i>lnHE</i> _{pub} → <i>lnLE</i>	11	[0.00144, 0.248]	0.0364
<i>lnHE</i> _{priv} → <i>lnIM</i>	5	[– 1.333, 0.50]	– 0.0833
<i>lnHE</i> _{pub} → <i>lnIM</i>	11	[– 0.6435, 0.0759]	– 0.2231

From a health policy perspective, the obtained results are interesting. In many country types analysed here there is still room for health improvements with larger HE . Globally, private expenditures seem less health-productive than HE_{PUB} . Thus, the policy option would be a shift of resources from the private to the public sector. However, in practice this would be difficult by harming the consumer's sovereignty with taxes and regulation. Also, first we must know in detail to what extent private and public health care services are complements or substitutes to each other. Note here that all three components of care—price, volume, and quality—have their impact on both the care provisions and their utilization. Typically, at least in non-developed countries, the price and quality components dominate in the private sector, as much of care provision here is based on imported medical skills and goods. As in many developed countries, care differentiation refers to the sector financing but not to the provision of health care, and the distinction between the two is not evident. In our data, high-income countries' private and public expenditures correlate positively with each other much more strongly than in low-income countries (0.707 and 0.277, respectively). As the expenditure leakage is smaller in poor countries, correctly targeted HE_{PUB} and their marginal increases matter greatly in low-income countries. Thus, a general policy option is to subsidise the use of health care services so that the poorest can have access to and resources for them.

Acknowledgements Open access funding provided by University of Eastern Finland (UEF) including Kuopio University Hospital. We thank two anonymous referees for helpful comments.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Acemoglu, D., & Finkelstein, A. (2008). Input and technology choices in regulated industries: evidence from the health care sector. *Journal of Political Economy*, 116(5), 837–880.
- Ahn, S. C., & Schmidt, P. (1995). Efficient estimation of models for dynamic panel data. *Journal of Econometrics*, 68, 5–27.
- Akinkugbe, O., & Afeikhen, J. (2006). Public health care spending as a determinant of health status: a panel data analysis for SSA and MENA. In A. Adenikinju & O. Olaniyan (Eds.), *Applied macroeconomics and economic development*. Nigeria: Ibadan University Press.
- Anand, S., & Ravallion, M. (1993). Human development in poor countries: on the role of private incomes and public services. *Journal of Economic Perspectives*, 7(1), 133–150.
- Anyanwu, C. J., & Erhijakpor, E. O. A. (2007). Health expenditures and health outcomes in Africa. African Development Bank Economic Research Working Paper No. 91.
- Arellano, M., & Bond, S. R. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58, 227–297.
- Baldacci, E., Guin-Siu, M. T., & de Mello, L. (2002). More on the effectiveness of public spending on health care and education: A covariance structure model. *Journal of International Development*, 15, 709–725.
- Barthold, D., Nandi, A., Mendoza-Rodriguez, J. M., & Heymann, J. (2014). Analyzing whether countries are equally efficient at improving longevity for men and women. *American Journal of Public Health*, 104(11), 2163–2169.
- Bidani, B., & Ravallion, M. (1997). Decomposing social indicators using distributional data. *Journal of Econometrics*, 77, 125–139.
- Blundell, R. W., & Bond, S. R. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87, 115–143.
- Bruno, S. F. (2005). Approximating the bias of the LSDVC estimator for dynamic unbalanced panel data model. *Economics Letters*, 87, 361–366.
- Bun, M. J. G. (2003). Bias correction in the dynamic panel data model with a nonscalar disturbance covariance matrix. *Econometric Reviews*, 22, 29–58.

- Bun, M. J. G., & Carree, M. A. (2005). Bias-corrected estimation in dynamic panel data models. *Journal of Business and Economic Statistics*, 23, 200–210.
- Bun, M. J. G., & Carree, M. A. (2006). Bias-corrected estimation in dynamic panel data models with heteroscedasticity. *Economic Letters*, 92, 220–227.
- Bun, M. J. G., & Kiviet, J. F. (2003). On the diminishing returns of higher order terms in asymptotic expansions of bias. *Economics Letters*, 79, 145–152.
- Caldwell, J. C. (1986). Routes to low mortality in poor countries. *Population and Development Review*, 12(2), 171–220.
- Chudik, A., & Pesaran, M. H. (2015). Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of Econometrics*, 188(2), 393–420.
- Cremieux, P.-Y., Meillur, M.-C., Quéllette, P., Petit, P., Zelderm, M., & Potvin, K. (2005). Public and private pharmaceutical spending as determinants of health outcomes in Canada. *Health Economics*, 14, 107–116.
- Dang, V. A., Kim, M., & Shin, Y. (2015). In search of robust methods for dynamic panel data models in empirical corporate finance. *Journal of Banking & Finance*, 53, 84–98.
- Deaton, A. (2013). *The great escape: health, wealth, and the origins of inequality*. Princeton: Princeton University Press.
- Dhaene, G., & Jochmans, K. (2012). Split-panel jackknife estimation of fixed-effect models. Université catholique de Louvain, Center for Operations Research and Econometrics Discussion Paper No. 2010003, revised 21 July 2012.
- Easterly, W., & Levine, R. (1996). Africa's growth tragedy: policies and ethnic divisions. Institute for International Development/Harvard, Papers 536
- Everaert, G., & Ponzi, L. (2007). Bootstrap-based bias correction for dynamic panels. *Journal of Economic Dynamics and Control*, 31, 1160–1184.
- Filmer, D., & Pritchett, L. (1997). Child mortality and public spending on health: how much does money matter. World Bank Policy Research Working Paper No. 1864.
- Filmer, D., & Pritchett, L. (1999). The impact of public spending on health: Does money matter? *Social Science and Medicine*, 49, 1309–1323.
- Gapminder (2015). <http://www.gapminder.org/data>
- Gerdtham, U. G., & Johnsson, B. (2000). International comparisons of health expenditure: theory, data and econometric analysis. In A. J. Culyer & J. P. Newhouse (Eds.), *Handbook of health economics* (Vol. 1A, pp. 11–53). The Netherlands: Elsevier.
- Gouriéroux, C., Phillips, P. C. B., & Yu, J. (2010). Indirect inference for dynamic panel models. *Journal of Econometrics*, 157, 68–77.
- Hahn, J., Hausman, J., & Kuersteiner, G. (2007). Long difference instrumental variables estimation for dynamic panel models with fixed effects. *Journal of Econometrics*, 140, 574–617.
- Hahn, J., & Kuersteiner, G. (2002). Asymptotically unbiased inference for a dynamic panel model with fixed effects when both N and T are large. *Econometrica*, 70, 1639–1657.
- Hahn, J., & Moon, H. (2006). Reducing bias of MLE in a dynamic panel model. *Econometric Theory*, 22, 499–512.
- Han, C., & Phillips, P. C. B. (2013). First difference maximum likelihood and dynamic panel estimation. *Journal of Econometrics*, 175, 35–45.
- Han, C., Phillips, P. C. B., & Sul, D. (2014). X-Differencing and dynamic panel model estimation. *Econometric Theory*, 30(1), 201–251.
- Hayakawa, K., & Pesaran, M. H. (2015). Robust standard errors in transformed likelihood estimation of dynamic panel data models with cross-sectional heteroscedasticity. *Journal of Econometrics*, 188(1), 111–134.
- Heijink, R., Koolman, X., & Westert, G. P. (2013). Spending more money, saving more lives? The relationship between avoidable mortality and healthcare spending in 14 countries. *European Journal of Health Economics*, 14, 527–538.
- Hu, S., & Ljungwall, C. (2013). China's healthcare system: Overview and quality improvements. Swedish Agency for Growth Policy Analysis (pp. 1–40). Östersund, Sweden.
- Huang, R., & Ritter, J. R. (2009). Testing theories of capital structure and estimating the speed of adjustment. *Journal of Financial and Quantitative Analysis*, 44, 237–271.
- Jaba, E., Baln, C. B., & Robu, I.-B. (2014). The relationship between life expectancy at birth and health expenditures estimated by a cross-country and time-series analysis. *Procedia Economics and Finance*, 15, 108–114.
- Jamison, D. T., Wang, J., Hill, K., & Londofio, J.-L. (1996). Income mortality and fertility control in Latin America: country-level performance 1960–1990. *Analisis Economico*, 11, 219–261.
- Ke, X., Saksena, P., & Holly, A. (2011). The determinants of health expenditure: a country-level panel data analysis. Working paper of the Results for Development Institute, WHO.

- Keane, M., & Neal, T. (2016). The Keane and Runkle estimator for panel-data models with serial correlation and instruments that are not strictly exogenous. *The Stata Journal*, 16, 523–549.
- Keane, M., & Runkle, D. (1992). On the estimation of panel-data models with serial-correlation when instruments are not strictly exogenous. *Journal of Business & Economic Statistics*, 10(1), 1–9.
- Kiviet, J. F. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics*, 68, 53–78.
- Kiviet, J. F. (1999). Expectations of expansions for estimators in a dynamic panel data model; some results for weakly-exogenous regressors. In C. Hsiao, K. Lahiri, L.-F. Lee, & M. H. Pesaran (Eds.), *Analysis of panel data and limited dependent variables* (pp. 199–225). Cambridge, MA: Cambridge University Press.
- Kiviet, J. F., Pleus, M., & Poldermans, R. (2017). Accuracy and efficiency of various GMM inference techniques in dynamic micro panel data models. *Econometrics Journal*, 5(1), 14–57.
- Liang, L. L., & Mirelman, A. J. (2014). Why do some countries spend more for health? An assessment of sociopolitical determinants and international aid for government health expenditures. *Social Science and Medicine*, 114, 161–168.
- Lichtenberg, F. (2000). Sources of U.S. longevity increase, 1960–1997. CES-info wp-405.
- Linden, M., & Ray, D. (2017). Life expectancy effects of public and private health expenditures in OECD countries 1970–2012: Panel time series approach. *Economic Analysis & Policy*, 56, 101–113.
- Lleras-Muney, A., & Sherry, G. (2008). Health inequality, education and medical innovation. *Demography*, 45(3), 741–761.
- López-Casasnovas, G., Rivera, B., & Currais, L. (2005). *Health and economic growth: Findings and policy implications*. Cambridge: MIT Press.
- Medeiros, J., & Schwierz, C. (2013). Estimating the drivers and projecting long-term public health expenditure in the European Union: Baumol's cost disease revisited. European Commission, Economic Papers 507. Brussels: European Economic Commission.
- Musgrove, P. (1996). Public and private roles in health theory and financing patterns. World Bank Discussion Paper No. 339. Washington D.C.: The World Bank.
- Nickell, S. J. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6), 1417–1426.
- Nixon, J., & Ulmann, P. (2006). The relationship between health care expenditure and health outcomes: Evidence and caveats for a causal link. *European Journal of Health Economics*, 7(1), 7–18.
- Novignon, J., Olakojo, S. A., & Nonvignon, J. (2012). The effects of public and private health care expenditure on health status in sub-Saharan Africa: New evidence from panel data analysis. *Health Economics Review*, 2(1), 1–8.
- OECD. (2014). *Health at a glance: Europe 2014*. Paris: OECD Publishing.
- OECD (2015). [Data.oecd.org/healthres/health-spending](http://data.oecd.org/healthres/health-spending). Paris: OECD.
- Oliviera Martins, J., & de la Maisonneuve, C. (2006). The drivers of public expenditure on health and long-term care: An integrated approach. *OECD Economic Studies*, 43, 116–154.
- Or, Z. (2000). Exploring the effects of health care on mortality across OECD countries. Labour Market and Social Policy Occasional-Papers No 46. Paris: Organization for Economic Cooperation and Development.
- Pesaran, M. H. (2015). *Time series and panel econometrics*. Oxford: OUP.
- Phillips, P. C. B., & Sul, D. (2003). Dynamic panel estimation and homogeneity testing under cross section dependence. *Econometrics Journal*, 6, 217–259.
- Phillips, P. C. B., & Sul, D. (2007). Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence. *Journal of Econometrics*, 137, 162–188.
- Poullier, J. P., Hernandez, P., Kawabata, K., & Savedoff, D. W. (2002). Patterns of global health expenditures: results for 191 countries. WHO Discussion Paper No 51.
- So, B. S., & Shin, D. W. (1999). Recursive mean adjustment in time series inferences. *Statistics and Probability Letters*, 43, 65–73.
- Van Baal, P., Oqbulqasim, P., Brouwer, W., Nselder, W., & Mackenbach, J. (2013). The influence of health care spending on the life expectancy. NETSPAR panel paper 35.
- Weil, D. N. (2009). *Economic growth* (2nd ed.). London: Pearson International.
- WHO. (2015). *World health statistics 2015*. Geneva: WHO.
- World Bank. (2015). *Population statistics*. Washington: WB.
- World Bank. (2016). *Economic statistics*. Washington: WB.