

OECD Health Working Papers No. 95

Future trends in health care expenditure: A modelling framework for cross-country forecasts

Alberto Marino,
David Morgan,
Luca Lorenzoni,
Chris James

https://dx.doi.org/10.1787/247995bb-en

OECD

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

21-Jun-2017

English text only

DIRECTORATE FOR EMPLOYMENT, LABOUR AND SOCIAL AFFAIRS
HEALTH COMMITTEE

DELSA/HEA/WD/HWP(2017)5
Unclassified

Cancels & replaces the same document of 20 June 2017

Health Working Papers

OECD Health Working Paper No. 95

FUTURE TRENDS IN HEALTH CARE EXPENDITURE: A MODELLING FRAMEWORK FOR CROSS-COUNTRY FORECASTS

Marino A.*, James C.*, Morgan D.* and Lorenzoni L.*

JEL classification: C53, H51, I18, J11

Authorized for publication by Stefano Scarpetta, Director, Directorate for Employment, Labour and Social Affairs

(*) OECD, Directorate for Employment, Labour and Social Affairs, Health Division.

All health Working Papers are now available through the OECD's website at:
http://www.oecd.org/health/health-working-papers.htm

JT03416536

Complete document available on OLIS in its original format
This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

English text only

DIRECTORATE FOR EMPLOYMENT, LABOUR AND SOCIAL AFFAIRS

www.oecd.org/els

OECD HEALTH WORKING PAPERS

http://www.oecd.org/health/health-working-papers.htm

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the author(s).

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcomed, and may be sent to health.contact@oecd.org.

This series is designed to make available to a wider readership selected health studies prepared for use within the OECD. Authorship is usually collective, but principal writers are named. The papers are generally available only in their original language – English or French – with a summary in the other.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

© OECD 2017

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for commercial use and translation rights should be submitted to rights@oecd.org.

ACKNOWLEDGEMENTS

The authors would like to thank Christine de la Maisonneuve and Peter Hoeller (OECD Economics Department), and Joaquim Oliveira Martins (OECD Centre for Entrepreneurship, SMEs, Local Development and Tourism) for their comments and contributions to the methodological discussion for this paper. The authors would also like to thank Francesca Colombo (OECD Health Division) for comments on an early version of this paper.

ABSTRACT

Across the OECD, healthcare spending has typically outpaced economic growth in recent decades. While such spending has improved health outcomes, there are concerns about the financial sustainability of this upward trend, particularly as healthcare systems are predominantly funded from public resources in most OECD countries. To better explore this financial sustainability challenge, many countries and international institutions have developed forecasting models to project growth in future healthcare expenditure.

Despite methodological differences between forecasting approaches, a common set of healthcare spending drivers can be identified. Demographic factors, rising incomes, technological progress, productivity in the healthcare sector compared to the general economy (Baumol's cost disease) and associated healthcare policies have all been shown to be key determinants of healthcare spending.

- For **demographics**, death-related costs has been shown to be the main factor behind increasing healthcare costs, with costs for people in their last year before death between 2 and 15 times higher than for people who survive. The impact of ageing on increased health expenditure, then, is predominantly in terms of the share of a country's population being close to death (non-survivors).

- For **income**, most cross-country econometric studies have found income elasticity (the relative share of GDP allocated to the health sector as a country becomes richer) in high-income countries to be less than one, after other spending drivers are accounted for. More specifically, the average elasticity estimate is 0.75. At the same time, there is evidence that low- and middle-income countries show higher elasticities.

- Low productivity in the health sector – commonly referred to as **Baumol's cost disease** – has been widely documented in high-income country settings. On average, the literature points to over half of productivity gains in the overall economy being translated to wage increases in the healthcare sector.

- **Technology** has also been shown to have, on aggregate, a positive impact on health spending. Estimates of its exact effect, and the methods used to derive such estimates, vary widely.

Based on this review, this paper sets out a theoretical framework for forecasting health expenditure trends that reflects the relative strengths of each of these drivers through panel regressions, while ensuring that modelling assumptions are transparent and internally consistent. Results from the literature offer plausible ranges for key model variables. The forecasting framework allows for both total and public health expenditures to be projected, with an indirect estimation of private spending. A distinction is also made between spending increases which can offer quality or access improvements, and those that solely increase costs. Such a distinction can inform policy discussions around financial sustainability and the extent to which any increases in health spending can provide real added value.

RÉSUMÉ

Au sein des pays de l'OCDE, l'augmentation des dépenses de santé a généralement dépassé la croissance économique au cours de ces dernières décennies. Bien que ces dépenses aient permis d'améliorer les conditions de santé des populations, la question de leur pérennité financière reste entière, en particulier parce que le financement de ces systèmes de santé repose principalement sur des fonds publics. Afin de mieux mesurer ce futur défi financier, de nombreux pays et organisations internationales ont développé des modèles de prévisions de la croissance des dépenses de santé.

Malgré des différences méthodologiques entre ces modèles de prévision, des facteurs communs de la croissance des dépenses de santé peuvent être identifiés. La démographie, l'augmentation des salaires, le progrès technologique, la productivité du secteur de la santé relativement au reste de l'économie (*Loi de Baumol*) et autres politiques liées à la santé se révèlent être les principaux déterminants des dépenses de santé.

- Pour la **démographie**, les coûts entraînés par les décès sont le principal facteur expliquant l'augmentation des dépenses de santé, avec des coûts 2 à 15 fois supérieurs pour les patients mourant dans l'année du traitement que ceux qui survivent. L'impact du vieillissement de la population correspond donc principalement à la part de la population dans la dernière année de vie (*les non-survivants)*.

- Pour les **revenus**, la plupart des études économétriques comparatives concluent sur une élasticité-revenu (la sensibilité des dépenses de santé à l'augmentation du PIB) inférieure à 1 dans les pays à hauts revenus. Plus précisément, l'élasticité moyenne se situe autour de 0,75. Cependant, d'autres études montrent que celle-ci est bien plus élevée dans les pays à moyens et bas revenus.

- La faible productivité dans le secteur médical - ou ***Loi de Baumol*** - est très largement documentée dans les pays à hauts revenus. En moyenne, la littérature conclut que plus de la moitié des gains de productivité de l'économie se retrouvent dans des augmentations de salaire dans le secteur de la santé.

- Il a été prouvé que le **progrès technologique** a un impact positif, au niveau agrégé, sur les dépenses de santé. Les estimations et méthodes d'estimation utilisées pour approcher cet effet sont très diverses.

A partir de ce constat, cette étude propose un cadre théorique pour prévoir les prochaines évolutions des dépenses de santé prenant en compte, au travers de régressions de panel, l'importance relative de chacun de ces déterminants, tout en s'assurant de la transparence et cohérence des hypothèses de modélisation. Les résultats de la littérature offrent un éventail de valeurs plausibles pour les variables clées du modèle. Le modèle de projection permet de prévoir à la fois les dépenses publiques de santé et les dépenses totales, avec une estimation indirecte des dépenses privées. Le modèle distingue aussi les augmentations de dépenses qui se traduisent par des améliorations de qualité ou d'accès aux soins, de celles qui impliquent uniquement une hausse des coûts. Cette distinction peut alimenter les discussions autour de la soutenabilité financière des dépenses de santé et dans quelle mesure des augmentations de dépenses peuvent engendrer une réelle valeur ajoutée.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS 3
ABSTRACT 4
RÉSUMÉ 5
1. INTRODUCTION 8
2. MAIN APPROACHES TO FORECASTING HEALTHCARE EXPENDITURE 9
2.1 Categories of forecasting models 9
2.2 The System of Health Accounts framework and different components of healthcare expenditure 11
2.3 Forecasting spending by healthcare function, and the case of long-term care 13
2.4 Forecasting spending by financing scheme: potential differences between public and private healthcare 14
3. DRIVERS OF HEALTH CARE EXPENDITURE 15
3.1 Main drivers of health care expenditure 15
3.2 Demographics 17
3.2.1 Ageing: share of the elderly 17
3.2.2 Time-to-death models 18
3.3 Income elasticity 22
3.3.1 Levels of aggregation 23
3.3.2 Country income groups 24
3.3.3 Simultaneity and reverse causality 25
3.4 Baumol's cost disease 26
3.4.1 Baumol's theory 26
3.4.2 Medical prices as a proxy for Baumol's effect 26
3.4.3 Deriving and incorporating the Baumol's variable 27
3.5 Technological advancements 28
3.5.1 The residual approach 29
3.5.2 The proxy approach 29
3.6 Policy and variable interactions 30
3.6.1 Institutional characteristics of health care systems and policy measures 30
3.6.2 Interactions between ageing, technology and policy 31
4. FORECASTING FRAMEWORK 32
4.1 Scope of the analysis 32
4.2 Regression model for health care expenditure growth 33
4.2.1 Data specifications 33
4.2.2 Econometric specification 33
4.2.3 Death-related costs dynamic model for demographic effects 34
4.2.4 Use of income elasticity and country income groups 35
4.2.5 Derivation of Baumol's variable 35
4.2.6 Model integration of technological advancements 36
4.3 Scenario analysis and plausibility ranges 36
5. CONCLUSIONS 39
REFERENCES 40
ANNEX 1 45

CROSS-SECTIONAL STUDIES 45
NATIONAL LEVEL STUDIES 46
AGGREGATE PANEL STUDIES 47
BAUMOL'S COST DISEASE STUDIES 49
OECD HEALTH WORKING PAPERS 51

Boxes

Highlights: main approaches to forecasting HCE and this paper's forecasting framework 9
Box 1. Advantages and limitations of different categories of forecasting models 10
Highlights: main drivers of health care expenditure 15
Highlights: demographics 17
Highlights: income elasticity 22
Box 2. Unit roots, stationarity and co-integration 24
Highlights: Baumol's cost disease 26
Highlights: technological advancements 28
Highlights: policy and variable interactions 30

1. INTRODUCTION

1. Across the OECD, there have been concerns about the financial sustainability of the upward trend in healthcare spending in recent decades, particularly as funding is largely from public sources in most OECD countries. For example, previous OECD analysis shows that without reforms to contain health-related costs, public spending on healthcare for OECD countries will increase on average from 6% of GDP in 2010, to 9% in 2030, and to 14% in 2060 (de la Maisonneuve and Oliveira Martins 2013).

2. To better understand this financial sustainability challenge, many countries and international institutions have developed **forecasting models** to project growth in future healthcare expenditure (Astolfi *et al.* 2012). These models use a variety of different approaches and underlying assumptions, making their results context-specific and not always straightforward to compare.

3. In this context, the first objective of this paper is to understand the main **approaches used to forecast healthcare expenditure growth**. Forecasting models from national and international institutions can be grouped according to the unit of analysis – that is, whether individuals, population sub-groups or the whole population is used as the modelling basis. Forecasting models also differ in terms of the components of healthcare expenditure analysed, particularly whether forecasts are split by financing source (total, public or private spending) or disaggregated by healthcare function (e.g. long-term care, medical goods, inpatient and outpatient care). These different forecasting approaches are discussed in Section 2.

4. Despite methodological differences between approaches, a common set of **healthcare spending drivers** can be identified. Demographic factors, rising incomes, technological progress, productivity in the healthcare sector compared to the general economy (Baumol's cost disease) and associated healthcare policies have all been shown to be key determinants of healthcare spending. The second objective of this paper is to review the literature, focusing on the estimation of the relative contribution of these drivers to healthcare spending growth. An analysis of these drivers is presented in Section 3.

5. Building on these findings, the last objective of this paper is to establish a **forecasting framework** that: *(a)* reflects the main drivers of spending growth and is methodologically robust; and *(b)* allows for comparative analysis across a wide range of countries. This forecasting framework is set out in Section 4. As well as reflecting key drivers of healthcare spending, it sets out technical details of the model, particularly the derivation and construction of variables to be used in regression models; the rationale for modelling healthcare expenditure by component (financing scheme or type of expenditure); and the use and standardisation of data. It also includes the potential use of scenario analysis to assess how different policies may impact the trajectory of healthcare expenditure over the coming decades, reflecting both commonalities and differences across countries.

2. MAIN APPROACHES TO FORECASTING HEALTHCARE EXPENDITURE

Highlights: main approaches to forecasting HCE and this paper's forecasting framework

- Forecasting models can be categorized under three groups, depending on the unit of analysis:
  - Microsimulation models produce detailed results that can incorporate behaviours at the individual level, but are data intensive and do not include the broader economic environment.
  - Macro-level models produce broad forecasts using time series with key economic variables. They are not data intensive, but the series need to be long and without breaks, and take into account cyclicality and broader economic trends.
  - Component-based models are between microsimulation and macro-level models in terms of data needs and complexity. They usually use population sub-groups and break down expenditure by sub-components.
- The forecasting framework developed in this paper uses a component-based approach. It allows for a greater level of detail than macro-level models, while the data required is light enough to allow comparable cross-country forecasts across OECD and non-OECD countries. It also allows for the integration of System of Health Accounts sub-categorisations of health expenditure.
  - A split by financing scheme (total, public, private) will be implemented so that total and public expenditure can be estimated (using the same set of drivers), while private expenditure will be derived.
  - A split by function of care (long-term care, pharmaceutical, etc.) can provide new understandings of the mechanisms of expenditure, but it will not be implemented at this stage, since further research is required to understand the drivers behind these functions.

2.1 Categories of forecasting models

6. Forecasting models are typically grouped according to the unit of analysis employed, which in turn affects both the level of sub-categorization of healthcare expenditure analysed and the amount of data and simulation required. Three broad categories can be identified in the form of **microsimulation** models (individual level), **component-based** models (cohorts of individuals and categories of expenditure) and **macro-level** models (population as a whole). These are briefly presented below – a more detailed discussion can be found in Astolfi *et al.* 2012.

7. **Microsimulation models** rely on individuals as the main unit of analysis. Examples in this family of models are the Canadian Population Health Model (Statistics Canada 2011) and the Swedish Long-term Demand for Welfare Services (Klevmarken 2011). These models are usually highly intensive in terms of data requirements, since they aim to simulate the characteristics and behaviours of individuals that populate the model.

8. **Component-based models**, of which cohort-based and actuarial (or cell-based) models are important sub-classes, usually rely on larger sub-groupings of the population as the unit of analysis, making use of important splits of expenditure (by function or provider). Examples of such models include the European Union/Ageing Working group model (European Commission 2011) and the Health Expenditure Growth Projection Method (de la Maisonneuve and Oliveira Martins 2013) developed by the OECD Economics Department. These models are less demanding in terms of data requirements, since they are more flexible regarding the main stratification component for the population (usually age), but they

retain a good amount of specificity and complexity for the resulting projections. For this reason, they are the dominant group in terms of popularity of use.

9. **Macro-level models** use the population as a whole as the unit of analysis. Examples include the Australian Government Productivity Commission Model (2005) and the US Center for Medicare and Medicaid Services Dynamic Computable GE Model (Borger *et al.* 2008; Friedman 2010). These models are the least demanding in terms of data requirements, since they do not necessarily need to stratify the population within more specific sub-groups. Instead, they rely on panel analysis and time trends to project (total) health expenditure in the future.

10. Within the past decade, several institutions have used component-based models to project healthcare expenditure. As an example, the model developed by de la Maisonneuve and Oliveira Martins (2013) uses ageing, income, technology and Baumol's cost disease to project public health and long-term care expenditure to 2030 and 2060 under different scenarios. The model uses an accounting framework (simple sum analysis) and residual approach (the residual factor from the sum corresponds to the unexplained part of the model) to model unobserved components (technology and policies). A similar residual approach has been used by other institutions, such as the WHO (Holly *et al.* 2011) and the IMF (IMF Fiscal Affairs Department 2010).

11. The theoretical framework presented in this paper is firmly rooted within the component-based approach, both in terms of drivers and stratification of the population (e.g. age groups for demography, country income groups for income elasticity and Baumol's cost disease). It also aims to take the best of macro-level models by introducing total health expenditure as a key dependent variable, and by using regression techniques similar to those most commonly found in the literature to estimate the parameters for the variables included in the model.

Box 1. Advantages and limitations of different categories of forecasting models

Microsimulation models require large amounts of data to assemble a sample that adequately represents the whole population of interest and includes all the characteristics of interest. Data is often gathered from a variety of sources, and sophisticated statistical techniques are often required to standardize the various databases so that they can be used to populate all of the desired attributes of individuals included in the sample. Datasets used include cancer and hospital registries, data from the pharmaceutical industry, population health surveys and clinical trials. For dynamic microsimulation, a second data-related issue concerns the design of realistic behaviours for all individuals over their life course. Estimating relative risks or hazards of transition from one state to another require analysis of longitudinal data or a review of the health economic literature where relative risks or hazards of transitions have been reported. Degrees of response that individuals may have to changes in external factors (elasticity) can be estimated through econometric regressions based on individuals' past experiences and choices or taken from the literature. Major limitations of this category of models include the difficulty to include in the model health system characteristics and policies that determine the supply and provision of healthcare services and may modify diagnostic and therapeutic pathways; the difficulty to include components of expenditure such as administration and research/investments; and the fact that these models are not conceived to explicitly include the broader economic environment in which "virtual" individuals live.

Component-based models are typically more data demanding than macro-level models but less demanding than microsimulation models, and this partially explains their popularity. Basic versions of component-based models typically break down health expenditure into major spending categories and age classes and employ actuarial projections as the main driver of future health spending. The development of more sophisticated versions of these models requires additional information, such as health spending broken down by gender and disease categories. The implementation and maintenance of component-based

models tends to be relatively simple and inexpensive, and they can be integrated into a broader framework that projects other spending, such as social security expenditure. In these models, the impact of policy changes can be assessed by simply modifying the value of the parameters of the drivers that were included. Major limitations of this category of models include the difficulty to explore distributional impacts of different policies to control healthcare expenditure growth; the limited ability to test "what if" scenarios about the impact of new policies to control expenditure growth; the fact that these models will only be able to test assumptions about the future burden of disease/health status of the population; and the difficulty to estimate the population's engagement in healthy lifestyles and its impact on healthcare expenditures.

Macro-level models restrict analysis to aggregate healthcare expenditure. They depend on the presence of clear and undisturbed trends and the absence of structural breaks. Thus, they are most appropriate for short-term projections, as the extrapolation methods benefit from the inertia in the healthcare system. Macro-level models are typically the least demanding in terms of data requirements. This is particularly the case for pure extrapolation methods which use a single time series, and for regression-based models which very often include just a few explanatory variables. However, time series need to be relatively long and consistent. The presence of breaks in the series, due to either changes in the methodology, the implementation of specific policies, or to limitations in the availability of data points may harm the results. In addition, the identification of relevant explanatory variables in regression-based models is particularly complex. In fact, swings from expansion to contraction in health spending (turning points) can be predicted only if explanatory variables are able to anticipate peaks and troughs. The computational and data requirements for dynamic computable general equilibrium models (CGE), on the other hand, are generally much higher and depend on the specification of the equations included in the model. CGE models depend on assumptions of equilibrium that may not account for observed trends and rely on strong simplifying assumptions about the behaviour of individuals, firms and governments.

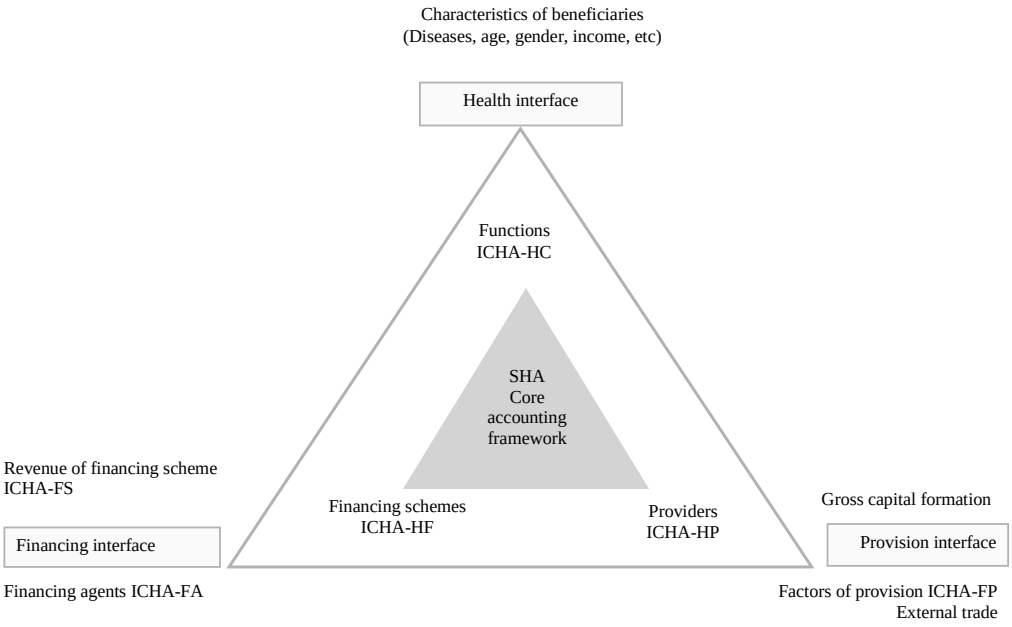
2.2 The System of Health Accounts framework and different components of healthcare expenditure

12. The forecasting framework suggested in this paper is grounded in the System of Health Accounts (SHA) standard (OECD, Eurostat, WHO, 2011). SHA provides a comprehensive framework for reporting internationally comparable data on healthcare expenditure and is organised around a tri-axial system of classifications: the different healthcare **functions** representing the types of healthcare services and goods consumed, the **providers** of these services and goods and the **financing schemes** paying for them (Figure 1). SHA comprises a set of dimensions, and corresponding classifications, that enable the systematic tracking of the flow of resources in a country's health system. SHA 2011 takes into account both public and private sector activities in health and provides key inputs into the design, implementation, and evaluation of health policies.

13. From a forecasting perspective, the SHA framework is useful for defining specific **components** of healthcare expenditure that could be separately forecast. For example, forecasts could be disaggregated by healthcare function. In particular, drivers for long-term care may well be different to drivers for other healthcare spending, such as inpatient and outpatient care, or the significance and effect of demographics. Forecasts could also be disaggregated by type of financing, such as government and other compulsory financing schemes as compared with direct payments by households and voluntary private health insurance. These disaggregation approaches are further discussed in Section 2.3 (healthcare function) and Section 2.4 (health financing scheme).

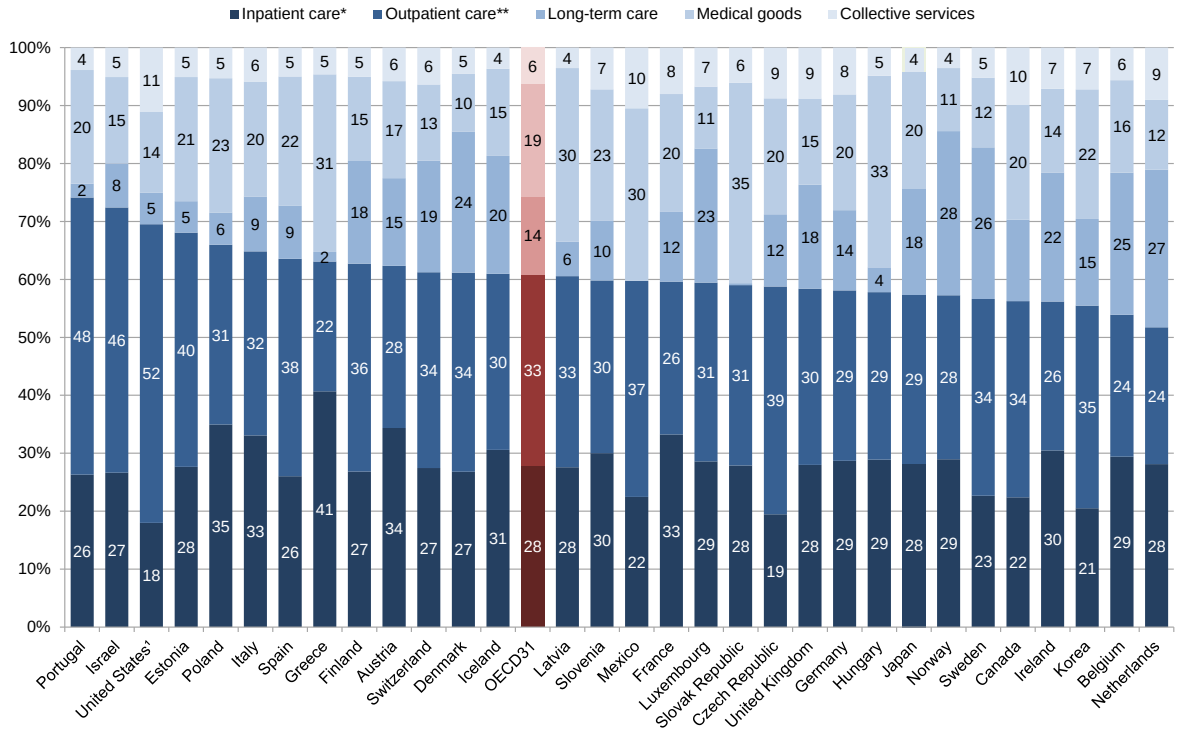
14. The SHA framework is also useful for determining the **boundaries** of healthcare expenditures, and therefore the boundaries of related forecasting models. For example, long-term nursing care and personal care with a health purpose are classified by countries as healthcare expenditures, whereas other assistance services such as household support and instrumental activities of daily living (that is, help with cleaning, cooking, etc.) are typically classified as social care (non-health) expenditures.

Figure 1: The core and extended accounting framework of SHA 2011



Source: OECD, Eurostat and WHO (2011).

Figure 2: Health expenditure by function of healthcare, 2014 (or nearest year)

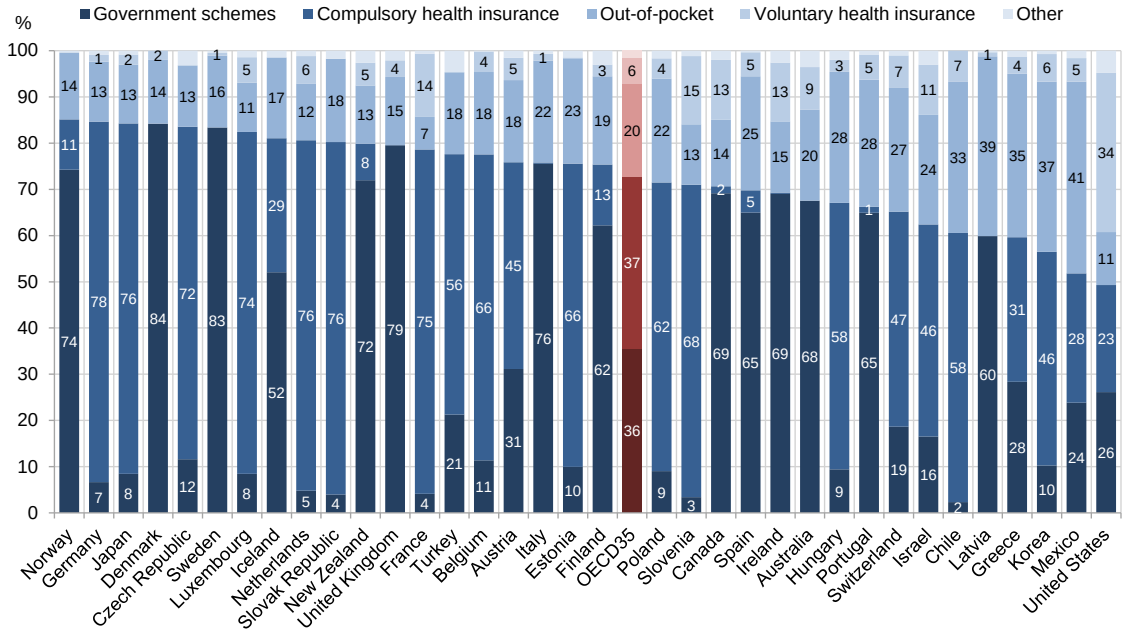


Note: Countries are ranked by curative-rehabilitative care as a share of current expenditure on health.
*Refers to curative-rehabilitative care in inpatient and day care settings. **Includes home-care and ancillary services.
[1] Inpatient services provided by independent billing physicians are included in outpatient care for the United States.
Source: OECD Health Statistics 2016.

15. Figures 2 and 3 show the distribution of health spending according to the main functions and financing schemes respectively. As can be seen, the majority of health spending is allocated to individual healthcare services aiming at the cure and rehabilitation of patients across the different institutional settings of the healthcare system. Medical goods, predominantly retail pharmaceuticals, also account for an important share of spending; while the relevance of the LTC sector can vary significantly. The remainder is spent of collective services – namely prevention and public health, and the overall administration of the healthcare system.

16. Public, or more precisely government resident-based schemes or compulsory health insurance, account for the majority of health spending in almost all OECD countries – typically accounting for three-quarters of the spending on healthcare goods and services. On the private side, households' own spending is the most important source, while voluntary health insurance plays a varying role in a large group of countries. The rest of the private financing comes from a mix of non-governmental organisations (*e.g.* charities) and private businesses.

Figure 3: Health expenditure by type of financing, 2014 (or nearest year)



Source: OECD Health Statistics 2016.

2.3 Forecasting spending by healthcare function, and the case of long-term care

17. It is likely that the drivers of spending for different healthcare functions differ in significance and effect. Perhaps most notably, the main determinants of **long-term care** (LTC) spending may be different from those driving other functions of healthcare such as inpatient and outpatient care. For example, there may be a stronger relationship between LTC spending and demographic change, given that a high share of LTC patients are elderly.[1] LTC spending growth is also likely to be more dependent on the extent of informal care from family members than other types of healthcare spending. But other functional differences could also be important. Spending growth in **pharmaceuticals** and medical goods, for example, may depend more critically on technological developments (such as in the increasing importance of effective but expensive personalised medicine), than other functions of healthcare.

[1] The relationship between spending and the so-called "pure ageing effect", measured by the share of the elderly (65+) is found to be highly significant for LTC, but not for health care as a whole. This will be discussed further in section 3.2 on the demographic component.

18. A few forecasting models have attempted to separately model specific healthcare functions, typically focusing on LTC. The model previously developed at the OECD (2013), for example, had a special module for LTC. This included dependency ratios as a key cost driver of LTC (since patients in long-term care are dependents) but not for healthcare, and modelled demographic effects differently. It also had a measure for Baumol's cost disease, which was not included in the general model for healthcare. Although new evidence suggests that Baumol's cost disease applies to healthcare as a whole, the intuition behind it being a major driver of LTC stems from the fact that the sector is, in theory, more labour intensive than healthcare, and therefore more subject to Baumol's wage increases in excess of general productivity growth. Therefore, it was hypothesized that Baumol's coefficient would be higher as a share of total determinants of LTC. Finally, the OECD 2013 model also included a variable to proxy informal care (participation rate of women in employment).

2.4 Forecasting spending by financing scheme: potential differences between public and private healthcare

19. Another possible split of expenditure is by type of financing – primarily the distinction between public and private components. The rationale behind this categorisation is similar to the one by function: the drivers to forecast public and private components might differ. First, spending by households on different healthcare services might be determined by very different factors compared to government expenditure; second, even assuming similar drivers for both components, it is still likely that the importance of these drivers would differ. For example, households' willingness to shift their budget towards healthcare might impact income elasticity very differently, depending also on the type of payment system used in a country, the relevant insurance schemes, whether care is subsidized and to what extent (co-payments, reimbursement, etc.), and so on. Therefore, there might also be a degree of simultaneity when estimating public and private expenditure separately, which will be discussed later in this section.

20. Component-based models, especially those developed by international institutions, have traditionally modelled public healthcare expenditure only, rather than splitting healthcare spending into public and private, or modelling total expenditure in its entirety. This approach is motivated by the fact that public spending on healthcare is the frame of reference for public finances reviews.

21. However, component-based models are structured in such a way that the central estimates for the main determinants and income elasticity in particular, are extrapolated from studies which use total healthcare expenditure as the dependent variable rather than just public spending. Therefore, using such estimates is equivalent to making a strong, untested assumption that the income elasticity for total healthcare expenditure is the same as that for public healthcare expenditure.

22. If, therefore, income elasticity takes different values for public, private and total health expenditure, the use of estimates for public expenditure which were regressed on total will result in a functional form misspecification[2]. Moreover, the Baumol variable, which is introduced later in the paper, cannot be used as an explanatory variable for public healthcare expenditure since it takes into account general productivity of the economy as a whole (Hartwig 2008; Colombier 2012). In summary, modelling both total and public healthcare expenditure to control for estimation differences between the two – the approach suggested by Holly et al., 2011 – can control for both the functional form misspecification and a possible reporting bias mentioned in the literature.[3]

---

[2] A functional form misspecification occurs when the regression model is functionally incorrect, i.e. by including variables that do not relate to each other.

[3] There are major cross-country differences in how tax expenditures are treated and reported, and therefore the split between public and private might introduce large reporting biases into the projections (Moscone *et*

3. DRIVERS OF HEALTH CARE EXPENDITURE

3.1 Main drivers of health care expenditure

Highlights: main drivers of health care expenditure

- A number of studies have analysed trends in healthcare expenditure over the past forty years. Based on this literature, five main drivers can be identified:
  - Demographics. Evidence shows that death-related costs are the main factor behind increased costs. The impact of ageing on increased health expenditure is therefore predominantly in terms of the share of a country's population being close to death.
  - Income. Rising incomes increase expectations on the quality and scope of healthcare, thereby increasing healthcare expenditures.
  - Productivity (Baumol's cost disease). Low productivity in the health sector has been widely documented in high-income country settings, increasing healthcare costs.
  - Technology. New technologies can extend the scope, range and quality of healthcare services, but can also be very expensive.
  - Policy measures. A range of policies can influence healthcare spending growth, with provider payment reform and pharmaceutical policies two notable cost containment reform areas.
- A distinction can be made between drivers of increased spending that also offer possible quality or access improvements (income, technology and policy measures), as compared with spending increases that only reflect higher costs (Baumol's cost disease and demographics). At the same time, it is important to recognise potential interactions between these drivers.

23. Academic and institutional terminology commonly groups the *spectrum* of variables that have been used over the past 40 years to explain trends in healthcare expenditure under two categories — demographic and non-demographic drivers. This terminology was first employed in the late 70s, and it mainly referred to **ageing**, mostly expressed as the share of the elderly aged 65 and over (demographic) and **income elasticity**, in the form of GDP per capita (non-demographic) (Newhouse 1977; Kleiman 1974).

24. Notwithstanding the fact that income is, understandably, a major driver of healthcare expenditure, recent research has shown that income and ageing do not constitute the totality of determinants, since they fail to explain a sizable share of observed expenditure growth[4]. Other drivers have been proposed as candidates, the most common being **Baumol's cost disease** (e.g. wages in excess of

al. 2010; de la Maisonneuve and Oliveira Martins 2013). Using total health expenditure allows us to overcome this bias altogether.

4 Regressions in the first set of cross-sectional studies that use only ageing and income as main determinants were, on average, able to explain around 50% of health care expenditure. See Annex 1 for a review of these studies.

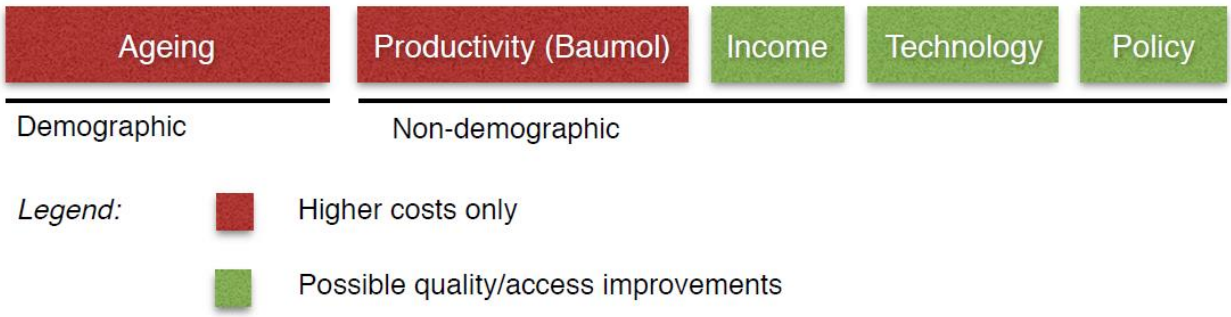
productivity growth), **technological advancements** (e.g. new drugs, treatments, medical technologies and diagnostic tests) and **policy measures**.

25. These drivers affect health spending in ways that are often difficult to assess. It would be an error to categorize all increases in spending as inherently bad and unwanted. Some of the increases in healthcare spending come, in fact, with higher quality or improved access to care. This share is affected by some of the drivers such as technological advancements, policy measures, and changes in income elasticity (a country getting richer and devoting more resources to healthcare). However, part of the increase in spending comes solely with increases in costs, with no possible improvement in quality or access. For example, the higher costs associated with Baumol's cost disease (rising wages in the medical sector that are not on par with the productivity increases) or the costs associated with the pure ageing effect of the population (the more the population ages, the more care will be needed). It is important that such effects are broken down by driver, so that policy options can be evaluated also in the light of those drivers whose effect should be mitigated (higher costs only), and those that require further analysis in terms of their cost-effectiveness (Figure 4).

26. These determinants interact with each other, blurring our understanding of their effect, but also reflect the fact that the demographic/non-demographic terminology in place is a false dichotomy — *i.e.* when, for example, life expectancy or mortality measures are used as a proxy for medical technology, we are effectively using a demographic determinant to measure the effect of a non-demographic one.

27. Moreover, if a true measure for medical technology was available, it would still cause an endogeneity problem[5], because of its interaction with the ageing component (measured as the share of the elderly). In fact, the impact of technology on the ageing cost curve distribution is well documented, and reflected in the forecasting literature through the implementation of assumptions of expansion or compression of morbidity, or the so-called **healthy ageing** hypothesis (de la Maisonneuve and Oliveira Martins 2013; James *et al.* 2015). Other interactions with policy effects are discussed in Section 3.6.2.

Figure 4: Drivers of health care expenditure



28. A key variable, which several recent studies have found to be highly significant in explaining health care spending growth, is the quantification of Baumol's cost disease – a theory positing that the health care sector suffers from lower productivity growth with respect to the overall economy, with wages having to stay on par with the wages of more progressive sectors. This leads to productivity gains being translated (up to a 1 to 1 ratio) to wages in non-progressive sectors of the economy, such as health care or education, where productivity improvements are structurally harder to achieve.

5 Endogeneity can happen when there is a correlation among two or more variables in the same regression function.

29. Another potential explanatory factor for variability in health care expenditure is the effect of policy measures undertaken by individual countries or internationally. This encompasses factors such as the provider payment system adopted by countries, regulations regarding subsidies for drugs, the share of public and private insurance schemes, and more generally any action undertaken by the government to regulate health care systems or improve effectiveness and safety of care.

30. Changes in policy are, however, largely unexplored, at least within the macroeconomic and component-based approach. This is due to the nature of the driver itself, which is best suited for a quasi-natural or experimental approach at the national level. Additional considerations will be made with regard to the proxies used for medical technology, which only partly capture prevention policies on addictive behaviours (e.g. smoking and alcohol consumption). Table 1 below summarises the drivers of health care expenditure and presents an overview of the variables used in regression models from the literature.

Table 1: Empirical variables for drivers of health care expenditure

Driver of Health care Expenditure	Variables used in models
Demographics	Share of the elderly (65+), Share of the young (15-), Death-related costs and age-specific cost curves
Baumol's cost disease	Medical prices, Wages in excess of general productivity
Income	GDP per capita
Technology	Index of hospital-country characteristics, Infant mortality, Life expectancy, Share of R&D, Share of patents, Residual calculation*
Policies	Index of policies and institutional characteristics of countries

* Note: Residual calculation refers to an approach to derive technology through a simple accounting sum (see Section 4).

3.2 Demographics

Highlights: demographics

- Demographics has been modelled through the share of the elderly, although recent evidence shows that time to death is a better predictor of health expenditure. The impact of ageing on increased health spending is, then, predominantly in terms of the share of a country's population being close to death.
- Death-related costs are modelled dynamically by splitting expenditure curves by age groups and looking at costs of those who are in their final years of life.
- These models are able to estimate a parameter that shows how much larger expenditure is for each population depending on how many people are close to death.

3.2.1 Ageing: share of the elderly

31. The demographic (or ageing) component in models of health care expenditure growth has traditionally referred to the inherent demographic shift of the population in any country across time. In the form of the share of the population older than 65 (sometimes coupled with the share of the population younger than 15), it aims to capture the increasing health care expenditure associated with a rising elderly share of the population. The hypothesis of increasing health care expenditure for the elderly reflects a higher prevalence of comorbidities and chronic illnesses in older people.

32. Although the share of the elderly is a relatively easy variable to obtain and to integrate in both macro and component based models, its effect on health care expenditure has traditionally been largely

insignificant. Literature results reported in Appendix 1 show the size and significance of coefficients for ageing measured as share of the elderly. These results show that the share of the elderly fails to take into account the nuances of how expenditure is associated to age groups, and to account for the fact that costs associated with deaths at a young age are much higher. A recent and growing body of literature suggests, in fact, that the share of the elderly might only be appropriate for modelling long-term care expenditure rather than health expenditure as a whole. This provides evidence for the alternative hypothesis that increasing expenditure is correlated with time-to-death, rather than with ageing *per se*, after splitting the relevant population cohort into survivors and non-survivors (see below). (Lis *et al.* 2016; Zweifel *et al.* 1999).

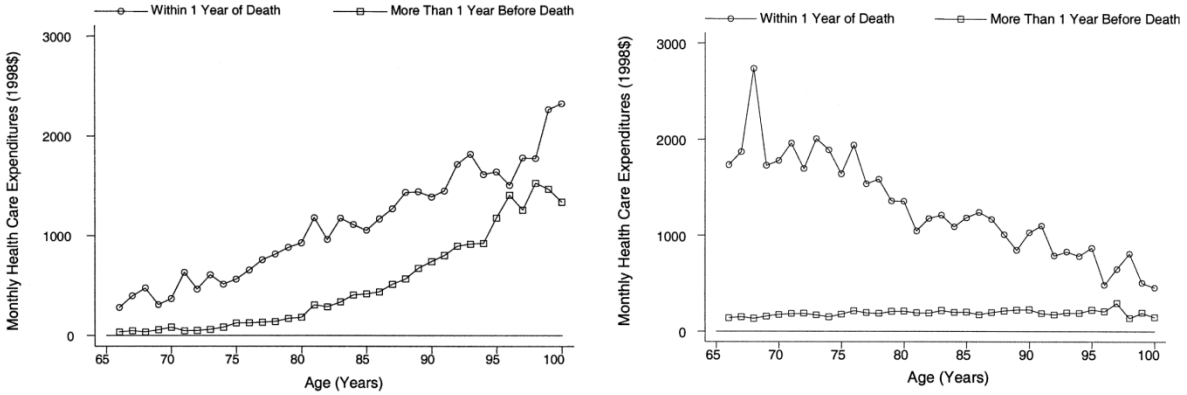
33. The ageing and technology components have been traditionally modelled separately. However, as discussed in the previous section, technological advances might intuitively increase life expectancy and – consequently – change the age distribution of the population. This potential endogeneity issue should be kept in mind when evaluating proxies for technology such as life expectancy and infant mortality, which incorporate both demographic and technological effects in one variable.

3.2.2 Time-to-death models

34. Zweifel et al. (1999) were among the first to question the relationship between ageing and health care expenditure, claiming that changes in the population structure due to lower fertility rates and higher life expectancy are unlikely to affect health care expenditure in the long run. They propose that the weak link between ageing and health care expenditure is driven by the high costs associated to dying, which is modelled by how close any individual is to death, with expenditure rising significantly in the two years prior to death, and very significantly in the final six months. Therefore, the observed relationship is in line with the hypothesis that, as people get older, they are more likely to be in their final years of life, which in turn affects health care expenditure.

35. They find that once remaining lifetime is controlled for in the regressions, ageing has no impact on health care expenditure for non-survivors, suggesting that ageing is only a significant factor among survivors with chronic illnesses. This finding is echoed in recent papers by Aprile (2007), Breyer and Felder (2006) and Yang *et al.* (2003). The latter test the hypothesis that age expenditure curves split by inpatient and nursing care differ significantly, with positive results – expenditure increases linearly with age at a very flat rate for inpatient expenditure (while declining rapidly for older decedents), but increases steeply for nursing care, making the case for pure ageing (as share of the elderly) being considered a major driver in long-term care expenditure calculations, but not for health care. Figure 5 below shows this graphically.

Figure 5: (left) Inpatient expenditure by age and time to death; (right) Nursing expenditure by age and time to death.



Source: Yang et al. (2003).

36. Several other studies in recent years show similar results supporting the time-to-death hypothesis. A study by Breyer and Felder (2006) lays out a set of competing hypotheses regarding healthy ageing and the behaviour of life expectancy in projecting future health expenditures, of which the time-to-death hypothesis has the most significant empirical support.

37. The point of reference is the **status quo** hypothesis, stating that age-specific expenditures remain stable once the impact of technological change is controlled for, and therefore life expectancy impact can be estimated by applying current age-expenditure profiles to future age distributions. This hypothesis does not, therefore, impact future expenditures through additional gains in life expectancy.

38. At the negative end of the spectrum is the **expansion of morbidity** hypothesis, which states that technology only prolongs the life of the very sick, therefore health status for the population decreases as the population gets older, increasing expenditure through additional use of health care goods and services. Therefore, the expansion of morbidity hypothesis states that a simulation based on status quo expenditure profiles will necessarily underestimate expenditure growth.

39. At the positive end of the spectrum is **time-to-death**, defined as the conjecture that "the observed difference in the health care expenditure between young and old […] are not primarily due to calendar age, but are caused by the differences of time-to-death" (Breyer and Felder 2006). In this case, they posit, an exogenous increase in life expectancy (because of new medical technologies, or simply healthier lifestyles) would cause less and less people in each age group to be in their last years of life, therefore lowering age-specific death rates. Therefore, a simulation based on status quo expenditure profiles will, in this case, overestimate expenditure growth.

40. An even stronger hypothesis in positive terms is referred to as **compression of morbidity**, or **healthy ageing**, which posits that as technology improves, more people reach the limits of longevity, and morbidity gets compressed in shorter periods before death. Intuitively, this means that as life expectancy increases, people are pushed towards their natural limits of death, and time-to-death becomes less relevant as people's inpatient expenditures in their final years of life will go down in favour of more natural, and less costly, deaths. This hypothesis claims that increases in life expectancy will result in a reduction of health expenditure growth in the future, since mortality rates will decrease.

41. While the status quo hypothesis has weak empirical evidence, there is no support in the literature for the expansion of morbidity hypothesis. The time-to-death hypothesis, however, has compelling

empirical proof, explored further below. The compression of morbidity hypothesis can also be observed in some behaviours observed in treatments for old people (Breyer and Felders, 2006).

42. Studies by Yang *et al.* (2003), Jacobzone (2003), Lis *et al.* (2016), Seshamani and Gray (2004) on health care expenditure stratify the study sample by age cohorts and survivors. Such studies employ expenditure curves by age and sex for survivors and non-survivors to test whether there is an increasing trend in expenditure for older age groups. They then use the non-survivors curves to differentiate between people who are within six months and five years before death, and people outside of that limit, to test whether people who are closer to death show higher health care expenditures. They finally use the values found to interpolate a curve for death-related costs where both age and time to death are taken into account.

43. When using average expenditure curves, the population for every country is split into age and sex groups, and the average health care expenditure for every age group is linked to the population to construct a curve that better reflects demographic shifts in the population. This curve is the basis for additional assumptions and scenarios within the model, such as the healthy ageing hypothesis, where the expenditure curves are assumed to shift to the right in time (life expectancy goes up without any changes in morbidity). From there, death-related costs are calculated – they reflect the higher expenditure associated to the remaining time-to-death for the average individual in a country.

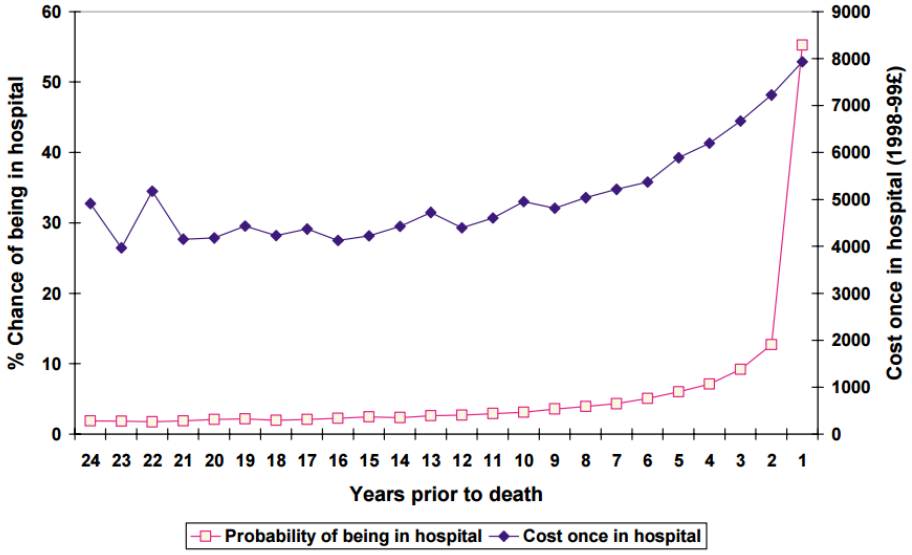
44. Results show that the *cost of dying* is several times higher than standard health care expenditures for both survivors and non-survivors that are not within the limits of time to death. This cost is also even higher for younger non-survivors rather than older ones.

45. The latest study by Lis *et al. (*2016), for example, finds that the age structure has a very modest impact on health spending, and the impact is best modelled by the share of the young[6]. However, when the breakdown on expenditures is based on time to death, the results are clear and significant. They find that expenditures rise 2-3 times faster for people within three years of death, and that these differences fade the further people are from dying, disappearing completely at the 10 years threshold. They also find that the age profile steepens in the last year of life, supporting the result that younger people in bad health drive spending up.

46. A study by Seshamani and Gray (2004) finds, instead, that while costs increase linearly starting at 16 years before death, the probability of being hospitalised has an exponential growth during those 16 years, with a quadrupling effect going from the penultimate to the last year of life (see Figure 6 below). Average costs (given by the probability of being hospitalized times the cost once in hospital), therefore, showed an exponential increase from 15 years before death on, with a seven-fold increase from year 3 to the year prior to death. They also show that, when splitting curves by age groups, the oldest age groups show more gradual increases in expenditure as death approaches, with a lower surge at the end of life. This effect was only calculated for ages 65+, with costs quintupling for a 65 year old but tripling for a 95 year old.

[6] Share of the young acts in similar ways, as a proxy, to the share of the elderly – but with a reversed sign.

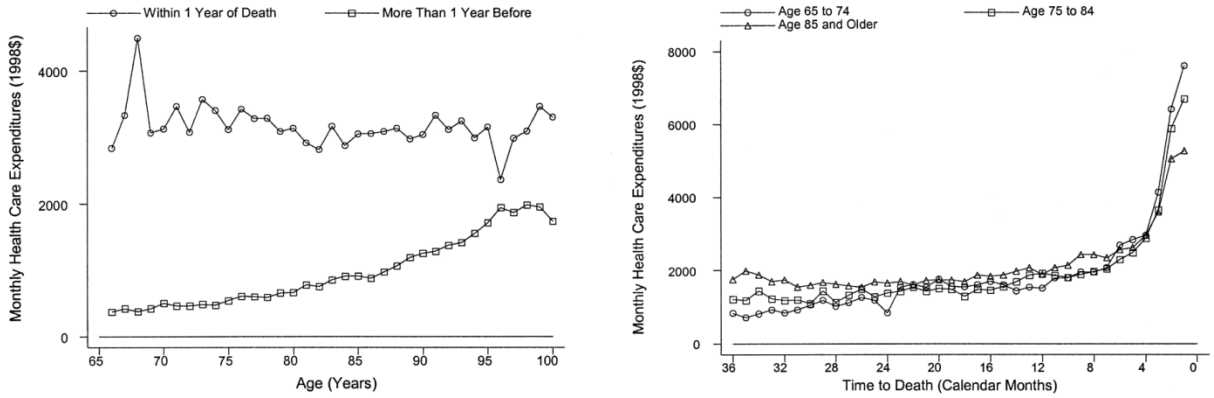
Figure 6: Time-to-death cost curve



Source: Seshamani and Gray (2004).

47. A similar approach by Yang *et al.* (2003) shows the same time to death chart, coupled with a simpler way of visualising survivors against non-survivors. Using a sample from the United States, their results are consistent with the rest of the literature, with expenditure for non-survivors significantly higher than for survivors, and independent of age. Expenditure for survivors is much smaller and increasing with age (when inpatient and nursing care are not separated) (Figure 7).

Figure 7: (left) Health care expenditure by age and time to death; (right) Health care expenditures by time to death and age



Source: Yang et al. (2003).

48. Lastly, a study by Jacobzone (2003) summarises some of the earlier studies that split expenditures by survivors and non-survivors. Table 2 below summarises some of the estimates for these studies, either described as expenditure in the last months of life or non-survivors expenditure (equivalent). These estimates differ by several magnitudes depending on the sample used (mostly national samples).

Table 2: Summary of studies on time-to-death and survivor curve split

Study	Findings
Mustard et al. (1998)	Non-survivors expenditure 15 times higher than average public expenditure
Madsen et al. (2000)	Expenditure in the last three months of life corresponds to 62% of total expenditure, with younger groups having more pronounced increases than the very old (over 80).
Busse (1996, 1999)	Non-survivors expenditure between 4 and 12 times higher than survivors.
Lagergren and Batljan (2000)	Non-survivors expenditure 12 times higher than survivors for inpatient care.
Zweifel (1999)	Expenditure in the last quarter of life 2-3 times higher than reference period.
Riley and Lubitz (1993)	Non-survivors expenditure 6.5-7 times higher than survivors.

Source: Jacobzone (2003).

49. The derivation of a coefficient for death-related costs which can be integrated in a regression framework will be possible through the use of survivors and non-survivors curves, split by age and sex. These curves will be used to estimate the ratio of costs between survivors and non-survivors for each country, which will be subsequently plugged into the regression. Section 4 details the dynamic regression specification used to estimate death-related costs.

3.3 Income elasticity

Highlights: income elasticity

- Income elasticity is one of the main drivers of HCE. It represents the proportional share of GDP that is allocated to health as income rises in a country.
- While health care was considered a luxury in the past, recent evidence – that accounts for other spending drivers – shows consistently that income elasticity is positive but below 1, making it a necessity good. That is, as income rises, health expenditure rises but at a lower rate.
- Income elasticity may vary depending on a country's income level, with some evidence that low- and middle-income countries show higher elasticities. Therefore, the framework will split the forecasts by country income groups.

50. Rising incomes increase expectations on the quality and scope of health care. Decades of research have consistently shown a positive relationship between national income and aggregate health care expenditures (see, for example, Farag *et al.* 2012). What is more debatable is the size of this effect, namely the income elasticity of health care expenditure. Income elasticity has been debated for the past 40 years — since Kleiman[7] (1974), a wide range of methods have been proposed to analyse its distribution and behaviour. The concept of income elasticity, in its simplest form, can be defined as the responsiveness of demand for health care (at different levels of aggregation) to changes in income of those demanding that good, all else being equal. Therefore, it captures people's behavioural preferences regarding how much of their additional budget (or the national budget, at higher levels of aggregation) they are willing to allocate to health care. This frames the forecasting of health care expenditure as a classic supply and demand

[7] Even though most attribute the inception of the topic of income elasticity to Newhouse's 1977 seminal paper, Kleiman preceded him using 1968 data while at the same time considering, ahead of his time, that elasticities of private and public outlay might differ substantially.

model: as a country's budget – measured by GDP per capita - increases, income elasticity determines, in large part, the share of additional funds that should be allocated to health care goods and services.

51. Economic definition tells us that if elasticity is above (below) 1, health care is a luxury (necessity) good. This means that, if elasticity is above 1, the share of expenditure devoted to health care would increase as income increases. The opposite is true if income elasticity is below 1.

52. This is important in two regards: first, it makes the estimation and sensitivity analysis of the parameter crucial, since its size has a larger impact *the more it deviates from 1* on the final forecast in either reducing or increasing health care expenditure as a share of GDP; second, it tells us whether public involvement in the market is needed (or indeed useful) to reduce expenditure (Farag *et al.* 2012). Recent literature is largely settled on health care being a necessity, even at the highest level of aggregation (international). This is reflected in the OECD model (de la Maisonneuve and Oliveira Martins 2013), which assumes an income elasticity of 0.8 as a central estimate, ranging from 0.6 to 1 (two standard deviations) in the sensitivity analysis. Similar component-based projection models, however, adopt extremely different values, with the IMF using 0.3, the Congressional Budget Office using 1 and the European Commission using 1.1 [8].

3.3.1 Levels of aggregation

53. The first studies of income elasticity, after Kleiman and Newhouse, and up to the beginning of the 1990s, found it to be consistently above one, making it a luxury good. These studies were based on a cross-sectional specification (a single year) from OECD countries, without the use of additional controls. Central estimates ranged from ~1.2 to 1.5, with some exceptions, such as Parkin *et al.* (1987), finding values of 0.9 when using a purchasing power parity (PPP) deflator for health care.

54. Cross-sectional studies were later found to be biased in two ways: firstly, they did not take into account the large omitted variable bias in excluding demand and supply-side variables within the analysis. This was evident in the exclusion from regression analysis of important factors beyond GDP and, later, ageing (which was later found to be largely insignificant for health care). Secondly, they were ignoring country and time fixed effects that could only be explored with panel data and models, therefore ignoring important cyclical effects of time and country differences in driver estimates. The omission of both types of controls resulted in a large upward bias, since growth was indeed increasing and income was the only regressor considered.

55. Later studies started implementing panel techniques to account for time trends and variations between countries. Some of these studies, even recent ones, restrict their analytical focus on a single country (national aggregation level). The reason why these studies are considered separately from studies of international panels was first theorised by Getzen (2000): there is a trend for income elasticity to become a luxury good at higher levels of aggregation.

56. While, in many developed countries, individuals may have near zero elasticities because of insurance and payment schemes that dilute or remove out-of-pocket expenditures (such as in the United Kingdom or France), aggregated expenditure does not necessarily follow the same risk behaviour pattern. Even though recent evidence suggests that health care might still be a necessity, even at the highest level of aggregation, this pattern of higher elasticities at higher levels of aggregation is observed both when shifting from national to international panels, and when comparing international panels of a few OECD countries against the only two studies (Holly *et al.* 2011; Farag *et al.* 2012) with over 130 countries.

[8] The European Commission's Ageing Working Group also assumes that income elasticity changes over time according to changes in consumers' preferences.

57. The early 2000s marked the completion of the transition from cross-sectional (and later pooled OLS) to panel studies. While a wide range of estimation techniques have been employed through the years, the benchmark still remains fixed effects estimation[9], with country and year trends and demand and supply-side controls. Estimations are often corroborated by tests on unit roots, co-integration and stationarity of the panel that became common practice in the late 1990s. Box 1 below briefly explains the issue.

Box 2. Unit roots, stationarity and co-integration

Stationarity of a panel is defined as the property of time series data of having mean or variance that do not change when shifted in time. Several co-integration studies in the late 1990s have found first and sometimes second level unit roots in many countries that are commonly analysed in the literature — unit roots different from zero cause violation of stationarity that must be corrected (some macroeconomic variables are found to have stochastic trends). The use of differencing and logarithmic transformation can help mitigate the non-stationarity of the panels, and these techniques are now common practice across recent studies.

Hansen and King (1996) were among the first to find non-stationarity in panels on health care expenditure. Among others, studies by Blomqvist and Carter (1997), McCoskey and Selden (1998), Gerdtham and Löthgren (2002), Herwartz and Theilen (2003), focused on the issue of co-integration.

58. The studies reviewed use a characterisation of demand and supply side controls similar to the one used for the drivers of the OECD model: demography (ageing measured as share of the population under 15 and/or over 65), technological progress (proxied by infant mortality, life expectancy, alcohol use, disease patterns, etc.), health prices and Baumol effect (physicians per 1000 population, public share of expenditure, substitution effects), and policy (system characteristics, voice and accountability, federal cash transfers).

3.3.2 Country income groups

59. One of the most relevant findings in recent studies (Holly *et al.* 2011; Farag *et al.* 2012; Di Matteo 2003) is that the distribution of income elasticity is not linear across countries, but follows a bell shape where low and high-income countries have lower income elasticities than middle-income ones. In the light of this hypothesis, even early studies like Newhouse's are found to be consistent with this. In fact, the simple inclusion of Greece (which had a much lower GDP than the other countries in his analysis) brought Newhouse's estimates up from 1.18 to 1.31.

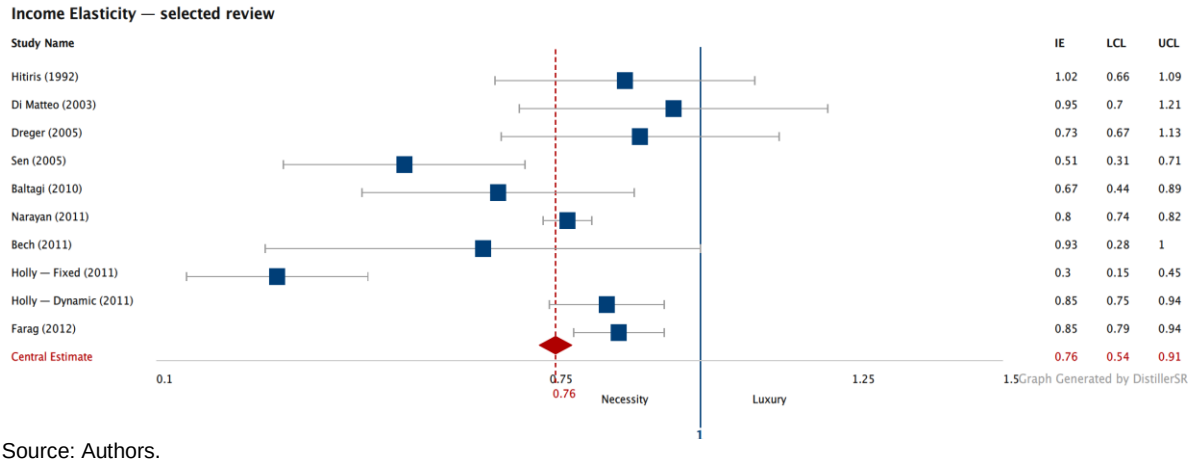
60. Figure 8 at the end of this section provides estimate ranges for income elasticity. The table of studies in Annex 1 also provides estimates for different income levels across countries for the studies that report on them. Even though the literature on this issue is still occasional, these studies are among the most robust ones as they employ a variety of controls, fairly large panels and large data sets.

61. The question of whether country-specific elasticities should be used in a forecasting model is not a trivial one: such elasticities have the advantage of being much more accurate in terms of both the true mean of the parameter and the used to estimate it. Estimating country-specific elasticities would prove much stronger in terms of robustness, since differences across countries in measurement and health systems characteristics can be taken into account. However, collecting data and estimating individual elasticities for every country in the OECD model would be challenging and time consuming.

[9] Although Hausman specification tests are regularly performed for random effects, and in some cases dynamic forms used, such as in Holly *et al.* (2011).

62. A plausible compromise would be to approximate elasticities according to country income groups (using – as an example – the World Bank categories: low, low-middle, upper-middle, high-income countries). This would increase the accuracy of the predictions by reflecting the finding that low and high-income countries seem to have lower income elasticities than middle-income ones, while retaining the international level of aggregation that ensures a degree of comparability and generalisability of the estimates. The assumption that income elasticity changes according to income groups implies also a change in preferences over time as real income is not constant.

Figure 8: Forest plot of income elasticity estimates



Source: Authors.

3.3.3 Simultaneity and reverse causality

63. Related to the Baumol effect is the issue of reverse causality brought up by Holly *et al.* (2011). They note that, while most of the literature has focused on the effect of GDP on health care expenditures, very few studies have considered that the causality could run in the opposite direction as well. Simultaneity could happen for two reasons — first, if we consider health a component of human capital such as education, we know that investment in human capital is related to GDP growth; second, a better health status increases labour supply and productivity, which are also related to economic growth.

64. A study by Erdil and Yetkiner (2009) finds that there is a significant degree of simultaneity between GDP and health care through the use of a Granger-causality test with lags. They find that 46 countries show simultaneity. Whenever they find one-way causality - in low and middle-income countries - it seems to run as expected from GDP to health care, while in high-income countries the reverse holds (Holly *et al.* 2011). However, testing with Granger-causality is not easy, and perhaps the inclusion of productivity data for the general economy for Baumol's cost disease would mitigate some of the bias by reducing the endogeneity of the simultaneity.

3.4 Baumol's cost disease

Highlights: Baumol's cost disease

- Baumol's theory predicts that health care costs will increase simply due to their inability to have significant productivity increases with respect to the rest of the economy.
- This effect is modelled in the framework through the derivation of a 'Baumol variable' meant to capture the size of wages in excess of productivity growth in the general economy.
- The Baumol effect is highly dependent on the relative state of the economy of a country, so the framework will also benefit from a differentiation between country income groups (like for income).
- The Baumol effect may apply with varying degrees of strength depending on the component of health care expenditure under analysis.
- The Baumol effect captures part of the health price effects.

3.4.1 Baumol's theory

65. Baumol was among the first to predict, with a theoretically grounded macroeconomic argument, that the health care sector would outgrow the general economy. He proposed that health care would be plagued, in the long run, by what we know now as "Baumol's cost disease" — the idea that health care (together with education and other smaller luxury sectors of the economy) is non-progressive, meaning that technological advancements, capital investments and economies of scale do not make for a cumulative rise in output that is on par with progressive sectors of the economy (Baumol 1967). The consequence of this argument is that productivity in the health care sector will be outgrown by general productivity of the economy, and therefore an ever increasing share of spending will have to be allocated to wages and salaries in the health sector to offset growth in the general economy and prevent a major shift in the labour force (Hartwig 2008, 2011).

66. Health care is one of the sectors supposedly suffering from this disease, since automation and better technology generally do not allow for large productivity increases. A health professional is difficult to substitute, in particular by using new technologies, which may actually also bring an increase in volume (e.g. faster diagnostic tests). Increases in volume likely brought about by new technology will also drive up expenditure, since new health professionals will have to be hired to treat everyone. Moreover, new technologies require more specialised training for say doctors, driving wages up further since more years of experience are required.

67. Baumol theorised a model whereby productivity increases in progressive sectors of the economy would eventually drive up labour and wages (and therefore expenditure) in non-progressive sectors. The technicalities of Baumol's cost disease, which is among the few theory-backed macroeconomic factors that could explain health care expenditure trends, are beyond the scope of this paper, and can be found in Hartwig (2008; 2011) and Colombier (2012).

3.4.2 Medical prices as a proxy for Baumol's effect

68. Baumol's variable is deeply rooted in his theory and revolves around our acceptance or rejection of its assumptions regarding the behaviour of progressive and non-progressive sectors of the economy. A series of papers by Hartwig (2008; 2011) set the tone for the discussion, by proposing that the historical

rejection of Baumol's theory to explain part of health care expenditure growth might be misguided in two respects. First, he argues, the rejection of Baumol on the premise that "the production of health care services might be as capital-intensive as manufacturing activities — or even more so" (Hartwig 2008) is missing the point — new medical technologies do not displace labour, for the reasons mentioned earlier, at the same rate (or at all) as in progressive sectors of the economy. Therefore, the assumption that Baumol theory only applies to sectors of health care that are less likely to be influenced by the use of new technologies (such as long-term care) might not hold.

69. Second, the lack of attention to Baumol's theory might be because of the assumption that **medical prices** were, in fact, able to proxy Baumol's cost disease (and consequent findings of non-significant coefficients). Several reasons are put forward as to why medical prices might *not* be a suitable proxy. First, there is the widespread belief that they carry a substantial upward bias because of omitted quality adjustments. Second, even when a GDP deflator is used to take into account country differences, international trends are extremely divergent across countries — due to differences in how the index is calculated and how government policies influence medical prices differently.

70. Governments have the possibility to influence health care expenditure in different ways, and "in an environment of heavily regulated prices, there are ways to shift resources into this sector other than raising prices" (Hartwig, 2008). With the concept of quality competition between hospitals becoming increasingly appealing, and the use of prospective payment systems where prices are set nationally, it could be argued that medical prices is an inherently endogenous variable, for which efforts to disentangle the effect of external factors which are not necessarily captured by the common country fixed effects might be futile.

71. Nevertheless, Hartwig (2011) proposes to test the Baumol effect using the best possible version of medical prices as a proxy, testing for all the criteria that emerged in the literature recently - the use of growth rates because of unit roots, deflators for medical prices, total current expenditure as the dependent variable. As is common for models that test for Baumol, the Hausman test suggests using a random effects specification possibly because the OECD samples are usually smaller than the one used in studies on income elasticity and therefore might be random in how they are selected. Controlling for income, Hartwig finds a positive coefficient for medical prices in all specifications, ranging from 0.38 to 0.60 in robustness analysis, and confirming that a positive and significant Baumol effect exists in health care, even when an imperfect proxy is used.

3.4.3 Deriving and incorporating the Baumol's variable

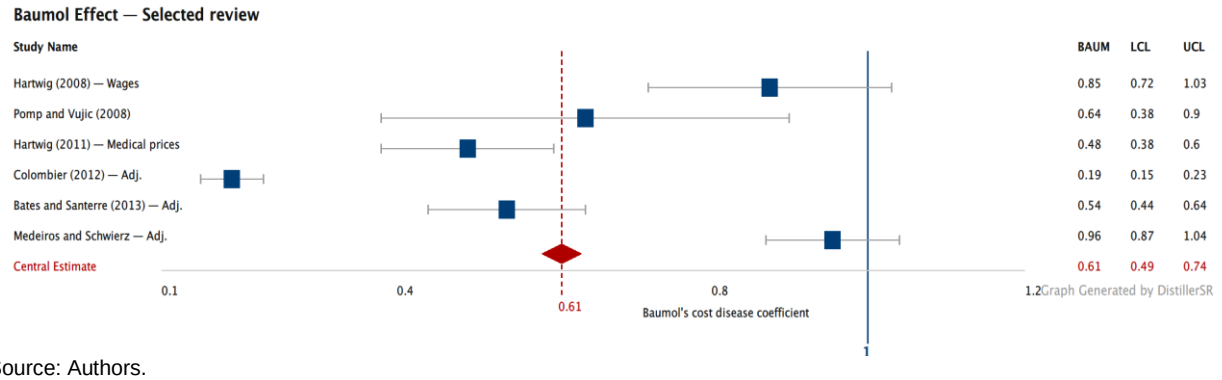
72. Hartwig's motivation for constructing a new way of testing Baumol's theory is therefore clear — medical prices are a misleading proxy which suffers from endogeneity in several respects. Baumol states that productivity growth brought about by quality-adjustment of service prices differs fundamentally from regular productivity growth. Therefore, in his 2008 paper, Hartwig defined the so-called **Baumol variable** by manipulating the equations in the original model of unbalanced growth by Baumol (1963). This variable measures the differential wage increases in excess of productivity growth. The derivation can be found in Section 4.2.5.

73. The advantage of using such a measure is that data are readily available from OECD, and that it does not have to cope with quality adjustment issues like for medical prices. This variable has been largely welcomed in the literature, with several papers using it to test for Baumol's theory in place of medical prices. Colombier (2012) later proposed an adjustment to the variable by weighting it by the share of the Baumol-affected sectors with respect to the overall economy, which was followed by others (Medeiros and Schwierz 2013; Ho and Zhou 2014; Bates and Santerre 2013).

74. The intuitive interpretation of the coefficient of Baumol's variable is as a ratio of wages in the Baumol-affected sector (health care) compared to productivity gains in the general economy. Therefore, a coefficient of 1 (the coefficient posited by Baumol in 1967) means that productivity gains in the general economy, much alike an inflationary factor, cause a **directly proportional** increase in wages in the medical sector, even though productivity in the sector is more stagnant than the rest of the economy. If wages in health sector of the economy were not kept at a rate similar to the general productivity of the country, individuals would eventually find that they are too low compared to the rest of the economy to justify the investment required to become a health professional.

75. Most of the econometric specifications that employ Baumol's variable control for at least two of the main determinants (income and ageing), with significant effects in their coefficients. The specifications are almost always in the form of OLS, fixed and random effects, with some using first differences or log differences instead of growth rates. Figure 9 below shows a forest plot of the central estimates for the studies reviewed, which find an average Baumol effect of 0.6, meaning that wages in the medical sector rise at a rate of 60% of the general economy productivity gains.

Figure 9: Forest plot of Baumol's variable estimates



Source: Authors.

3.5 Technological advancements

Highlights: technological advancements

- The introduction of new technologies is likely to increase health spending. However, its impacts are often difficult to model, because of unclear function of certain variables, computational complexity, or lack of data.
- Several variables have been used as a proxy for technological advancements, including life expectancy, infant mortality, share of the elderly, indexes of medical technologies, hospital research, coverage and general research and development.
- All the above variables have been found, at different times, to be significant at different levels and effect size. However, research and development expenditure has been consistently used over time, data is readily available for several countries, and is less likely to correlate with other variables in the framework.

76. New health technologies can extend the scope, range and quality of health care services. While some technologies can be cost-saving, others can increase costs, by offering better but more expensive care for complex illnesses, including those that may not have been previously treatable. Technological change in the health sector, though, represents the most complex and endogenous of the drivers of health care expenditure to model. Several considerations have already been made regarding its endogeneity, especially

taken together with the ageing component, where variables such as life expectancy interact with both the demographic and technological coefficients in ways that are not yet fully understood. Technological change has long been identified as the variable that is continuously and most rapidly changing in health care expenditure forecasting models (Chernew and Newhouse 2012).

77. The direction of the effect of technological change on health expenditure trends is not straightforward. As Chernew and Newhouse (2012) explain in their comprehensive analysis of definitions and methods used in the literature, technological advances come in different forms, such as **product**, **knowledge or process innovation**. The relationship between costs and quality of such innovations is ambiguous: some technologies might come with similar costs and higher quality (e.g. a new drug is introduced in the market, with better health outcomes and at a similar price than what is currently available). However, a large number of technologies come with higher quality, but also higher costs – or even higher costs for similar quality outcomes. Some technologies might reduce waiting times because they allow for increased volumes (e.g. a new diagnostic test that is much faster but just as precise as the next best alternative), while some others might increase them (e.g. a new diagnostic test with much better precision, but slower and/or more expensive than the next best alternative).

78. While the evidence suggests that, on balance, technology has a positive effect on expenditure, it should be kept in mind that this effect is composed of a positive and negative vector, depending on the definition of technology that is being used for the analysis. Moreover, depending on the method used to measure technological change (exogenous or endogenous technology), the estimate of its effect might incorporate policy effects or other confounding variables that are not specified in the model. Two main approaches have been commonly used in the literature: the most common one in component-based models, following the macroeconomic insights of Solow (1957), is the residual approach; while the second one, mostly preferred in macro-level literature, is what Chernew and Newhouse refer to as the affirmative approach, or put more simply, the proxy approach.

3.5.1 The residual approach

79. The residual approach, most commonly found in recent component-based health care expenditure forecasting models (and some macro-economic studies), uses regression methods to estimate the impact of time-varying drivers on health care expenditure growth, excluding any measure of technological change from the regression. The residual coefficient derived from the regression is then interpreted as the estimate that captures the impact of technology. This method is relatively straightforward, and is the standard when better measures of technological change are not available, or determined to be too endogenous to be feasibly implemented (Holly *et al.* 2011; de la Maisonneuve and Oliveira Martins 2013). In fact, this method is particularly sensible to endogeneity issues, arising from interactions among determinants that are discussed further in Section 3.7.

3.5.2 The proxy approach

80. The alternative approach uses a suitable proxy to be included directly in the regression model. Such proxies are varied, and they most often include mortality rates, life expectancy, research and development expenditure, patents registered, or other national indexes based on medical devices or hospital characteristics. Annex 1 includes a list of proxies used in macro-level studies on the determinants of health care expenditure.

81. Similarly, Chernew and Newhouse suggest capturing technology through the identification of which specific diseases and technologies have contributed to the increases in health care expenditure. However, the shortcoming of this approach is that it is not possible to fully capture the extent of it, let alone for a panel of countries.

82. The model developed at the OECD in 2013 assumes that there is an OECD-wide technology frontier for health (captured by the relative number of patents), and that the access to this frontier is mediated by the amount of spending in research and development (R&D). However, these proxies might not capture the effect that we are interested in as we know that, in today's globalised world, knowledge spreads quickly, and new medicines, diagnostic tests and therapies might be patented in one country but still used elsewhere. This is especially true for OECD countries (and Europe in particular), where pharmaceutical companies usually concentrate their research in a single country, but then export their products elsewhere in the world. Therefore, the effect of patents and R&D might be biased downwards.

83. A recent paper by You and Okunade (2017) explores most of the aforementioned variables as proxies for technological change, including the residual approach, in a case study for Australia: research and development expenditure, hospital research expenditure, hospital treatment coverage, share of the elderly (used in a similar way to life expectancy in other studies), infant mortality rate, and weighted and unweighted indexes of medical device technology. The rationale behind using the share of the elderly is that "changes in fertility rates and life expectancy (and consequently the share of the aged) might be influenced by technologies for improving longevity or delaying fertility".

84. After controlling for income and ageing used in different specifications as a proxy for demographics and technology, they find significant coefficients for all proxies. Predictably, the variable for R&D has sometimes a lower coefficient than a proxy like the medical devices index[10]. This is because new medical devices can induce patient demand with a much shorter lag as compared to changes in health care expenditure due to changes in the expenditure on research and development. Therefore, this suggests that the R&D variable could be introduced with a lag in the model to take into account the hypothesis that current health care expenditure is influenced by past increases of spending in R&D that may result in new technologies introduced only after several years.

3.6 Policy and variable interactions

Highlights: policy and variable interactions

- Policy measures for cost containment or wider reforms of the health system also contribute to health spending trends, but quantification is inherently complex and better suited for other types of studies.
- A residual approach (used as an alternative to calculate technology) might also partly capture unexplained policy effects.

3.6.1 Institutional characteristics of health care systems and policy measures

85. Institutional characteristics of health care systems are also relevant to health care spending, with a wide range of policies shown to be effective in containing the costs of health care. A recent review highlighted the health care policies that have been most effective in cost containment. Notable among these are provider payment reforms and pharmaceutical policies encouraging increased generics uptake. However, transforming such insights into variables that can be used for forecasting is not straightforward. One of the few papers that attempted to do this is by de la Maisonneuve *et al.* (2016). This study quantified the effect of structural influences – i.e. policies and institutional determinants of health spending

[10] The results for this study vary between 0.14 and 0.35 for the proxies used, which is in line with the upper bound of 0.35 found through a residual analysis in the same paper. However, these coefficients are not generalizable because they are greatly dependent on the specification used and are not bound by theoretical notions (like income elasticity and Baumol's cost disease).

– finding that these explain around a quarter of cross-country differences in health spending in a panel of OECD countries from 2000-2011.

86. Furthermore, results of econometric analyses (Lorenzoni *et al.* 2017) show that cyclical economic effects have further suppressed health care expenditure growth in the aftermath of the 2008 economic and financial crisis, while they have contributed positively to the health care expenditure growth when the economy has been expanding. These results lend support to the hypothesis that cyclical effects play a role in fluctuations in health care expenditure growth, however, health care spending growth still trends downward after controlling for cyclical effects, which suggests that structural influences also play a role.

87. There is little empirical evidence on these regulatory factors linked to policy decisions upon which models may be developed. Thus the impact of changes in those factors on health care expenditure trends will be a residual from the regression model.

3.6.2 Interactions between ageing, technology and policy

88. The literature is not conclusive as to the impact of technology proxies, since studies usually do not focus on the technological component but rather on income, ageing and (more recently) on Baumol's cost disease. One of the most comprehensive studies on technology by Chernew and Newhouse (2012) puts forward the argument that technology might be inherently endogenous, and therefore quite challenging to model without a suitable instrument. Regarding ageing, instead, most studies find largely insignificant effect when they model it using the share of the elderly or the share of the young (sometimes both).

89. The main issue with the proxies used in the literature is that they might pick up a different part of the same demographic effect that the ageing component tries to capture. If we take into account life expectancy, for example, it is intuitively clear how a variable that captures the increasing share of the elderly might interact with a variable that shows the monotonic trend of people consistently living to older ages.

90. Similarly, infant mortality might capture the effect that the population expands faster since births are monotonically more likely to result in a healthy life. Moreover, these effects are likely to capture the tendency of individual countries to act on the behaviour of individuals with prevention policies regarding their lifestyles. It might also be argued that, if not necessarily a proxy for policy, it could be seen as a proxy for culture itself. Some societies might view (for social or religious reasons) drinking, smoking, or obesity in a different light than others, which will in turn impact both infant mortality and life expectancy as a proxy for technology.

4. FORECASTING FRAMEWORK

91. In this section, we outline possible future directions for work on the construction of the forecasting framework for health care expenditure, by (1) describing the scope of the analysis with regard to the unit of analysis and the potential splits by financing scheme and function of care; (2) outlining the regression model for the forecasts, with mathematical derivation for constructed variables (demographics and Baumol's variable) and details about data specifications; (3) summarising the evidence discussed in the paper in the form of plausibility ranges for the drivers and scenario analysis.

4.1 Scope of the analysis

92. The forecasting framework will follow a component-based approach, which allows the forecasts to be comparable at the international level and to be split into sub-categories (or components) to allow for more precise estimates of the drivers, while also being only mildly data intensive. It will make use of regression methods to estimate the parameters of the model, residual and trends analysis for additional robustness, and scenario analysis to answer relevant policy questions and reflect different plausible situations in the mid-long term regarding the evolution of parameters such as life expectancy, demographic shifts, and technological advances. This approach is preferred to micro or macro simulation models because of its compromise between the data requirements and computational complexity that makes it detailed, yet still internationally comparable.

93. As discussed throughout the paper, while most previous models focused on public health expenditure, the forecast estimates would benefit from additional levels of detail found in the (partial) disaggregation of expenditure by either financing type or function of care. This level of detail should provide more accurate predictions on the relative effect of drivers on expenditure. However, we acknowledge that the drivers for disaggregated expenditure types, such as private health expenditure or long-term care, might differ substantially from those of health care as a whole.

94. Theory, unfortunately, is scarce, at the international level of aggregation, regarding common sets of drivers for functions of health care expenditure such as pharmaceutical or long-term care. In addition, there are issues regarding data availability both in measuring private health expenditure drivers, where national models that have modelled it in the past have access to data with a degree of detail that a cross-country model can't feasibly achieve; and in measuring long-term care drivers such as dependency ratios and making sure that countries reporting on LTC refer to the same common basket of care (social vs. health).

95. Therefore, the framework will adopt a partial split by financing type, with two forecasts being produced for public health expenditure and total health expenditure, using the same drivers and coefficients for both. This will allow for an indirect measure of private health expenditure by simple subtraction of the two final forecasts. However, a split by function of care is unfeasible at present time because of lacking theory at the international level on the drivers of such subcomponents of care. Future iterations of the framework should investigate further possible regression functions for this split.

96. In summary, the model will run under the following specification:

$$\boldsymbol{Health\ Care\ Expenditure}\ (Total, Public) \xrightarrow[country,time]{} f(Demographics, Income, Wages\ and\ \ Productivity, Technology)$$

$$\boldsymbol{Demographics}\ (Age, Sex) \xrightarrow[country,time]{} f(Probability\ of\ dying, Time\ to\ death, Standardized\ cost\ ratio)$$

$$\boldsymbol{Technology} \xrightarrow[country,time]{} \begin{cases} f(Share\ of\ research\ and\ development, Time\ lags) \\ or \\ f(Residual\ estimation\ for\ HCE) \rightarrow includes\ policy\ impact \end{cases}$$

$$\boldsymbol{Policies} \xrightarrow[country,time]{} f(Residual\ estimation\ for\ HCE) \rightarrow includes\ technology\ impact$$

4.2 Regression model for health care expenditure growth

4.2.1 Data specifications

97. Having identified all the main determinants of health care expenditure growth, the regression model to estimate the parameters for the forecast is simple to construct. We follow OECD recommendations in terms of how variables such as health care expenditure and GDP are adjusted: health care expenditures should not be used in the PPP-adjusted form, since the model specification employs growth rates as the unit of analysis (Hartwig, 2008; de la Maisonneuve and Oliveira Martins 2013). Instead, exchange rate and inflation adjustments should be made (by using constant international US$ transformed variables[11]) to allow for cross-country comparisons in a medium-long panel, since the OECD recommends using nominal data for very short panels only (Lorenzoni *et al.* 2015).

98. Whenever individual observations for isolated years are missing (for R&D and total employment), the data is projected linearly when a visible increasing (or decreasing) linear trend is present, to ensure the correct functioning of the first differencing adjustment required for the specification (Bates and Santerre 2013).

99. The measure for health care expenditure is derived by interacting GDP per capita with the share of total, public and private spending on health care as a percentage of GDP. Productivity Y is measured as real GDP divided by total employment. All other variables, including wages per worker, are already available in their model form in OECD.Stat. A table in Annex 2 summarizes possible data sources for the data requirements.

4.2.2 Econometric specification

100. The econometric specification of the model follows the mathematical notation of most macro-level studies to harmonise with the largest portion of the literary base used in this paper. The first difference of the natural log of all the variables (growth rate form) is used to avoid the issue of unit roots of level 1, as mentioned in Section 3.1, 3.3 (Box 1). The general equation is as follows:

[11] The wage component of the Baumol variable should be expressed, following Hartwig's preferred specification, in nominal terms. However, Hartwig also adjusts for wages in real terms in later model specifications, therefore the use of real wages is deemed appropriate.

$$\Delta \ln HCE_{c,t,f} = \alpha + \beta_1 \Delta \ln Demo_{c,t} + \beta_2 \Delta \ln GDP_{c,t} + \beta_3 \left(\Delta \ln W_{c,t} - \Delta \ln Y_{c,t}\right) * \delta_1 Income_{c,t} + \beta_4 \Delta \ln Tech_{c,t} + \eta_c + \tau_t + \varepsilon_{c,t} \quad (1)$$

$$Demo_{c,t} = \Psi^{surv}_{c,t,s,x}[(1 - \pi) + \pi \mathrm{k}_{c,t,s,x}] \quad (2)$$

101. The dependent variable HCE is health care expenditure per capita in country c for year t and function f (total or public); $Demo$ refers to the separately estimated demographic component in equation (2), in the form of the coefficient for the death-related costs by country; on β_2 we measure income elasticity as the first differenced log of GDP per capita, which will subsequently be integrated with OECD-produced GDP growth forecasts in order to obtain the share of HCE growth by country; on $\beta_3$we calculate the Baumol variable [12], measured as wages in the overall economy W in excess of productivity per worker Y— both GDP and Baumol's variable are interacted with a country income group dummy on δ_1; on on β_4, we estimate the technology component, either by the use of a proxy (in the form of national rate of research and development) or by a separate estimation, using a residual approach to simulate the joint impact of technology and any other potential driver explaining expenditure growth (mainly policy measures); lastly, ; η_c and τ_t are country and time effects, depending on the estimation methods (OLS, fixed and random effects), while $\varepsilon_{c,t}$ is the residual component of the regression.

102. Another possible way of modelling the regression for the forecast is to apply the subdivision for country income groups prior to the regression, producing a total of 8 regression equations (if all country income groups are present in the sample). This method can be used as a control to check for robustness of the estimates, since the country groups will be interacted with all variables in the regression rather than Baumol and income only.

4.2.3 Death-related costs dynamic model for demographic effects

103. The idea behind the death-related costs model is to project the estimation of health care expenditure by decomposing it in two components; a first one for *non-survivors* that takes into account time to death (one year for our model), and a second one for "normal" health costs of *survivors* (all other people). This division, further split into age and sex groups, makes it so that differences in the state of health are taken into account in the derivation of health care expenditure, so that a multiplier for costs of non-survivors can be estimated from the regressions.

104. A general equation for the model is found in Aprile (2007), as follows:

$$\Psi_{t,s,x} = \Psi^{surv}_{t,s,x}(1-\pi) + \Psi^{death}_{t,s,x}\,\pi = \Psi^{surv}_{t,s,x}[(1 - \pi) + \mathrm{k}_{t,s,x}\,\pi] \quad (3)$$

105. Where $\Psi^{surv}_{t,s,x}$ represents expenditure for survivors over time, age and sex; $\Psi^{death}_{t,s,x}$ is expenditure for non-survivors; π is the probability of death by age and sex; and $\mathrm{k}_{t,s,x}$ represents the expenditure ratio $\Psi^{death}_{t,s,x}/\Psi^{surv}_{t,s,x}$.

106. The above equation takes into account the evolution of age consumption profiles due to the dynamics of the probabilities of death underlying the demographic scenario (April 2007). Several assumptions can be made about the behaviour of parameter $\mathrm{k}_{t,s,x}$: if kept at a constant rate of 1, we would be assuming a pure age scenario, meaning that costs for survivors are translated linearly with costs for non-survivors (equivalent to a healthy ageing assumption where increases in life expectancy are translated to years in good health, since costs do not increase with age); if the upper limit bound is adopted (infinity),

[12] This variable captures part of the health price effects too.

we would be assuming that only the number of deaths matters in projecting expenditure, meaning that the cost of death would be infinitely higher than survivors'.

4.2.4 Use of income elasticity and country income groups

107. Income elasticity is expressed mathematically as the first difference of the logged GDP per capita by country. This is the traditional way of modelling it, since it avoids the issue of unit roots of level 1. The resulting elasticity estimate will be country-specific, allowing for comparison with the bell-shape hypothesis found in the literature. Moreover, the variable will be interacted with a dummy variable for country income groups, following the same classification as World Bank income groups (low income, lower middle income, upper middle income, high income). This will produce a set of four dummy interacted coefficients (one of which will be used as a baseline to avoid multicollinearity) for income elasticities.

4.2.5 Derivation of Baumol's variable

108. Baumol's model of unbalanced growth can be expressed mathematically as:

$$(4) \quad Y^{NP}(t) = aL^{NP}(t)$$

$$(5) \quad Y^{P}(t) = bL^{P}(t)e^{rt}$$
$$(6) \quad W(t) = W_0(e^{rt})$$

109. Where only one type of input, labour L, generates output in both the non-progressive (Y^{NP}) and progressive (Y^{P}) sectors. He assumes that, while the non-progressive sector in equation (4) stagnates, with output proportional to the amount of labour introduced, the progressive sector in (5) has output rising at a constant growth of r over time. He further assumes that wages in the two sectors are equal at each point in time, and that they rise over time with productivity improvements in the progressive sector, as per (6). It follows, therefore:

$$(7) \quad C^{NP}(t) = \frac{W_0(e^{rt})}{a}$$

$$(8) \quad C^{P}(t) = \frac{W_0(e^{rt})}{be^{rt}} = \frac{W_0}{b}$$

110. Equation (7) implies that unit costs rise continually over time in the non-progressive sector, depending on productivity improvements in the progressive one, while costs in (8) remain constant over time. This hypothesis can be tested by examining whether unit cost changes in the stagnant sector are directly proportional to the excess of wage rate growth less labour productivity growth in the overall economy, as in:

$$(9) \quad \Delta \log(C^{NP}) = \lambda[\Delta \log(W) - \Delta \log(Y)]$$

111. The expression in brackets is the empirical formulation of the Baumol variable, with W representing overall wages and Y representing output per worker (productivity in the general economy). A positive coefficient for λ provides support for Baumol's cost disease, with $\lambda = 1$ representing a directly

proportional relationship between general productivity in the economy and wages in the non-progressive sectors of the economy (in our case, health).[13]

4.2.6 Model integration of technological advancements

112. The technology variable, as discussed earlier, will be proxied through the use of the share of research and development of GDP in its first differenced log form (growth rate). This proxy is consistently used in the literature with significant and positive coefficients, although few studies have employed the use of lags for the variable. It is intuitively important to take time lags into account when using such a variable – if the aim is to model the impact of budget changes to research and development in health care expenditure, we are assuming that such changes are somehow modelling the development and implementation of new technological advancements in the medical field. However, increases in budget for R&D will have their full effect realized with a significant time lag, since medical technologies need several years of research, clinical tests and analyses before being released in the market. Therefore, as suggested by You and Okunade (2017), the implementation of 3 or 5 year lags will be crucial in modelling such a proxy.

113. An additional check for robustness can be performed for technological change: following past forecasting models, a residual approach can be employed to measure how significantly different the estimated proxy measure is from the accounting residual. This is because there is no conclusive evidence of the validity of R&D as a proxy for technology, and the unavailability of better proxies (such as medical technology country-specific indexes) makes the calculation of a residual the only alternative approach at the present time.

4.3 Scenario analysis and plausibility ranges

114. Forecasting models can be predictive or policy-oriented (Kopec *et al*. 2010; Weinstein *et al*. 2003). For example, weather forecasting models are essentially predictive models in that no policy can be implemented to influence the weather in the very near future. In contrast, policy-oriented health expenditure forecasting models aim to inform policy makers about when and how to implement reforms and what effects those reforms are likely to produce.

115. Transparent policy-oriented models should provide policy makers with a measure of the uncertainty associated with a projected value. Three tools are theoretically available to that end: the prediction interval, scenarios, and sensitivity analysis (Astolfi *et al*. 2012).

116. A prediction interval consists of an upper and a lower bound between which the future value is expected to lie with a prescribed probability. Knowledge of the prediction interval enables users to assess future uncertainty, plan different strategies for the range of possible outcomes indicated by the prediction interval, compare forecasts from different methods more thoroughly and explore different scenarios based on different assumptions more carefully.

117. Scenario analysis provides information about the range of possible outcomes of the model. Starting from the most likely outcome, the baseline, scenarios analysis sets values for a series of variables which lead to identifying the most favourable outcome (best-case scenario) and the most unfavourable outcome (worst-case scenario). Although possible, the best- and the worst-case scenarios are not necessarily associated with the highest probability to occur. Instead, they fix the extremes of the range of all possible outcomes.

[13] Derivation taken from Hartwig (2008), Bates and Santerre (2013).

118. A different valuation approach is given by sensitivity analysis, the process of assessing the impact of modifying the value of a single variable on the final outcome of the model. Sensitivity analysis is carried out to test policy scenarios such as a cost-containment international (or national) policy, an exogenous shock in the number of deaths and/or time to death (such as a pandemic), an exogenous shock in technological advancements (such as a major breakthrough for the cure of a large cause of mortality), and similar. Sensitivity analysis modifies one variable at the time, whereas scenarios analysis modifies a series of variables at once. Sensitivity analysis can be seen as a subset of scenario analysis.

119. Probably the most important use of the model that produce forecasts is to consider variable scenarios around the central forecast, based on alternative assumptions about policy or key conditioning variables. Those scenarios are often published alongside a commentary on the central forecast as a means of illustrating uncertainty and risks surrounding the central forecast (Turner 2016).

120. For some of the drivers, there is a clear suggested range within which the estimated parameter should lie (either because of mathematical constraints, like the Baumol variable, or because of overwhelming evidence, such as income elasticity). For others, such as the death-related cost ratio of time-to-death and normal expenditure, the evidence is scarce and largely dependent on the sample used, and therefore suggesting a plausibility range is less obvious. Table 3 below summarizes the plausibility ranges for the drivers of health care expenditure.

Table 3: Plausibility ranges for health care expenditure drivers

Driver of Health care Expenditure	Plausibility ranges
Demographics	Estimation approach with non-survivor costs between 2 and 15 times higher than survivors, with the parameter k taking similar values if assumed.
Baumol's cost disease	Coefficient hovering around 0.6 in health care.
Income	Coefficient hovering around 0.75 aggregated, following a bell shape with middle income countries showing higher elasticities.
Technology	Coefficient to be estimated econometrically, lagged at 3 or 5 years, with estimates hovering around 0.25 (approximately one half to one third of other proxies'effect); residual estimation.
Policies	Residual estimation for technology also takes policy effects into account. No other plausibility ranges available.

121. Regarding income elasticity, evidence is overwhelmingly in favour of the idea that health care is a necessity (below 1), with an average estimate of 0.75 for highly developed countries. The evidence regarding the shape of the distribution of income elasticity for country income groups is not yet settled – however, there is a growing body of findings in favour of the "bell shape" hypothesis, claiming that low and high income countries have lower elasticities than middle income countries. However, one of the most recent studies on the issue finds mixed evidence for this hypothesis (Baltagi *et al.* 2017). A regression approach with individual income elasticities will, therefore, provide specific estimates by country or by income group.

122. Baumol's variable, on the other hand, is generally constrained by its mathematical derivation, whereby a coefficient of 1 indicates a full, direct transfer of productivity gains from the general sector to wages in health care. The evidence finds that Baumol's cost disease is almost never realized fully in the real world, with coefficients hovering around 0.6.

123. The demographic component does not have either a mathematical constraint or a largely consistent body of literature behind its estimation. While this means there is no clear recommendation on its plausibility range, it is also the parameter that can be most easily interpreted in terms of scenarios and

sensitivity analysis. The parameter k can either be estimated or assumed – in the case of k = 1, we assume a pure ageing effect; for k = $\propto$, we are in a full death-costs scenario where only demographic mortality affects health care costs. Similarly, the parameter π can be either derived from real world data (demographic mortality by sex and age group) or tweaked in conjunction with assumptions regarding life expectancy and the impact of technology (expansion or compression of morbidity).

124. Lastly, the technology component has also no mathematical constraints or consistent evidence from the literature. The preferred proxy used in our model is share of research and development, which does not interact with other components (like infant mortality and life expectancy do) and is readily available on our database for all countries (unlike indexes of medical technology). Since coefficients for R&D are generally lower than other coefficients (an average of 0.25 compared to 0.5-0.75 for other proxies), as discussed in section 3.5 and 4.2.6, the proxy could be lagged at 3 or 5 years, to take into account potential time effects on health care expenditure.

5. CONCLUSIONS

125. Rapid health care spending growth is a concern for policymakers, particularly in countries with tightened fiscal constraints. Indeed, the majority of health care systems are publicly financed, implying an important fiscal sustainability challenge from rising health care expenditures. Forecasting scenarios for future health care expenditure can quantify the likely extent of this sustainability challenge, thereby helping policymakers act in a timely manner.

126. To produce plausible cross-country forecasts, this working paper first reviewed the literature on the main spending drivers for health care. Demographics, income effects, low productivity and technological advances have all been shown to increase health care spending. However, the nature and strength of these effects varies. This depends on the countries studied and related policy effects, as well as differences in empirical methodologies.

127. For demographics, time-to-death has been shown to be the main factor behind increasing health care costs, with expenditure for non-survivor ranging between 2 and 15 times higher than survivors. The impact of ageing on increased health expenditure, then, comes predominantly in terms of the share of a country's population being close to death (non-survivors). For income, most studies (using appropriately specified cross-country econometric studies) have found income elasticity in high-income countries to be less than one, with an average elasticity estimate of 0.75. At the same time, there is evidence that middle-income countries show higher elasticities. Low productivity in the health sector – commonly referred to as Baumol's cost disease – has been widely documented in high-income country settings. On average, the literature points to over half of productivity gains in the overall economy being translated to wages in the health care sector. Technology has also been shown to have, on aggregate, a positive impact on health spending. Estimates of its exact effect, though, vary widely.

128. Based on this review, a health care expenditure forecasting framework has been developed that can reflect the relative strengths of each of these drivers, while ensuring that modelling assumptions are transparent and internally consistent. Results from the literature offer plausible ranges for key model variables. The forecasting framework allows for both total and public health expenditures to be forecast, with indirect estimates of private spending derived (since drivers of private spending likely differ from those that underlie public and total). A distinction can also be made between spending increases which can offer quality or access improvements, and those that solely increase costs. Such a distinction can not only inform discussions around financial sustainability, but also on the extent to which any increases in health spending can provide real added value.

REFERENCES

Aprile R (2007). How to Take Account of Death-Related Costs in Projecting Health Care Expenditure – Updated Version. Ragioneria Generale Dello Stato, Italy.

Astolfi R, Lorenzoni L and Oderkirk J (2012). A Comparative Analysis of Health Forecasting Methods. OECD Health Working Papers No 59. OECD Publishing Paris http://dx.doi.org/10.1787/5k912j389bf0-en.

Australian Productivity Commission (2005). Impacts of Advances in Medical Technology in Australia. Research Report, Melbourne.

Baltagi BH, Lagravinese R, Moscone F and Tosetti E (2017). Health Care Expenditure and Income: A Global Perspective. Health Economics 26 (7): 863-874.

Baltagi BA (2010). Healthcare Expenditure and Income in the OECD Reconsidered: Evidence from Panel Data. Bonn: Institute for the Study of Labour.

Bates LJ and Santerre RE (2013). Does the U.S. health care sector suffer from Baumol's cost disease? Evidence from the 50 states. Journal of Health Economics 32(2): 386-391.

Baumol WJ (1967). Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis. The American Economic Review 57(3): 415-426.

Baumol WJ (1993). Social wants and dismal science: the curious case of the climbing costs of health and teaching. Proceedings of the American Philosophical Society 137(4): 612-637.

Borger C, Rutherford TF and Won GY (2008). Projecting Long-term Medical Spending Growth. Journal of Health Economics 27(1): 69-88.

Breyer F and Felder F (2006). Life expectancy and health care expenditure: a new calculation for Germany using the costs of dying. Health Policy 75 (2): 178-186.

Chernew ME and Newhouse JP (2012). Health Care Spending Growth. In: Handbook of Health Economics, Volume 2. Elsevier B.V.

Colombier C (2012). Drivers of health care expenditure: does Baumol's cost disease loom large? FiFo Discussion Papers 12(5).

de la Maisonneuve C, Moreno-Serra R, Murtin F and Oliveira Martins J (2016). The drivers of public health spending: integrating policies and institutions. OECD Economics Department Working Papers, No. 1283, OECD Publishing, Paris. http://dx.doi.org/10.1787/5jm2f76rnhkj-en.

de la Maisonneuve C and Oliveira Martins J (2013). A Projection Method for Public Health and Long-Term Care Expenditures. OECD Economics Department Working Papers, No. 1048, OECD Publishing. http://dx.doi.org/10.1787/5k44v53w5w47-en.

Di Matteo L (2003). The Income Elasticity of Health Care Spending: A Comparison of Parametric and Nonparametric Approaches. European Journal of Health Economics 4(1): 20-29.

Dieleman JL, Templin T, Sadat N, Reidy P, Chapin A, Foreman K, Haakenstad A, Evans T, Murray C J and Kurowski C (2016). National spending on health by source for 184 countries between 2013 and 2040. The Lancet 387: 2521-2535.

Dreger C and Reimers H-E (2005). Health Care Expenditures in OECD Countries: A Panel Unit Root and Cointegration Analysis , s.l.: IZA Discussion Paper No. 1469.

Erdil E and Yetkiner IH (2009). The Granger-causality between health care expenditure and output: a panel data approach. Journal of Applied Economics 41(4): 511-518.

European Commission (2011). The 2012 Ageing report. Underlying Assumptions and Projection Methodologies.

Farag M, NandaKumar AK, Wallack S, Hodgkin D, Gaumer G and Erbil C (2012). The income elasticity of health care spending in developing and developed countries. International Journal of Health Care, Finance and Economics 12: 145-162.

Fiscal Affairs Department – International Monetary Fund (2010). Macro-Fiscal Implications of Health Care Reform in Advanced and Emerging Economies. Fiscal Affairs Department.

Freeman DG (2003). Is healthcare a necessity or a luxury? Pooled estimates of income elasticity from US state- level data. Applied Economics 35 (5): 495-502.

Friedman JN (2010). Predicting Medicare Cost Growth. In Improving Health Care Cost Projections for the Medicare Population, Summary of a Workshop (G. Wunderlich ed.). Washington D.C.: The National Academies Press, pages 83-106.

Gerdtham U-G and Jonsson, B (2000). International comparisons of health expenditure: theory, data and econometric analysis. Handbook of Health Economics, Volume 1, Elsevier Science.

Gerdtham U-G and Lothgren M (2002). New panel results on cointegration of international health expenditure and GDP. Applied Economics 34 (13): 1679-1686.

Gerdtham U-G, Sogaard J, Andersson F and Jonsson B (1992). An econometric analysis of health care expenditure: A cross-section study of the OECD countries. Journal of Health Economics (11): 63-84.

Getzen T E (2000). Health care is an individual necessity and a national luxury: applying multilevel decision models to the analysis of health care expenditures. Journal of Health Economics 19: 259-270.

Hansen P and King A (1995). The determinants of healthcare expenditure: A cointegration approach. Journal of Health Economics 15: 127-137.

Hartwig J (2008). What drives health care expenditure? — Baumol's model of 'unbalanced growth' revisited. Journal of Health Economics 27(3): 603-623.

Hartwig J (2011). Can Baumol's model of unbalanced growth contribute to explaining the secular rise in health care expenditure? An alternative test. Journal of Applied Economics 43(2): 173-184.

Herwartz H and Theilen B (2003). The determinants of healthcare expenditure: testing pooling restrictions in small samples. Journal of Health Economics 12: 113-124.

Hitiris T and Posnett J (1992). The determinants and effects of health expenditure in developed countries. Journal of Health Economics 11: 173-181.

Ho C-Y and Zhou W (2014). Unbalanced growth and health care expenditure. Economics of Transition 22(4): 739-758.

Holly A, Xu K and Saksena P (2011). The determinants of health expenditure: A Country-Level Panel Data Analysis. WHO Working Papers.

Jacobzone S (2003). Ageing and the Challenges of New Technologies: Can OECD Social and Health Care Systems Provide for the Future? The Geneva Papers on Risk and Insurance 28 (2): 254-74.

James C, Blecher M, de Lagasnerie G, Kumar A and Vammalle C (2015). Fiscal Sustainability of health systems — Why is it an issue, what can be done? In: Fiscal Sustainability of Health Systems. Paris: OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264233386-en.

Johannson A, Guillemette Y, Murtin F, Turner D, Nicoletti G, de la Maisonneuve C, Bagnoli P, Bousquet G and Spinelli F (2013). Long-term growth scenarios. OECD Economics Department Working Papers, No. 1000, OECD Publishing, Paris. http://dx.doi.org/10.1787/5k4ddxpr2fmr-en.

Kleiman E (1974). The Determinants of National Outlay on Health. In: The Economics of Health and Medical Care. UK: Palgrave Macmillan, pages 66-88.

Klevmarken A (2011). Microsimulation for Public Policy: Experiences from the Swedish Model SESIM. ESRI Discussion Paper Series, Economic and Social Research Institute, Tokyo, Japan, No. 242.

Kopec JA, Finès P, Manuel DG et al. (2010). Validation of population-based disease simulation models: a review of concepts and methods. BMC Public Health 10:710.

Lis M (2016). Age or time-to-death — what drives health care expenditures? Panel data evidence from the OECD countries. IBS Working Papers 04/2016.

Liu S and, Chollet D (2006). Price and Income Elasticity of the Demand for Health Insurance and Healthcare Services: A Critical Review of the Literature. Mathematica Policy Research Inc., Washington, DC.

Lorenzoni L, Morgan D, Murakami Y and James C (2015). Public Expenditure Projections for Health and Long-Term Care for China until 2030. OECD Health Working Paper No. 84, OECD Publishing, Paris. http://dx.doi.org/10.1787/5jrs3c274vq3-en.

Lorenzoni L, Millar J, Sassi F and Sutherland D (2017). Cyclical vs structural effects on health care expenditure trends in OECD countries. OECD Health Working Paper No 92, OECD Publishing, Paris. http://dx.doi.org/10.1787/be74e5c5-en.

Medeiros J and Schwierz C (2013). Estimating the drivers and projecting long-term public health expenditure in the European Union: Baumol's «cost disease» revisited, Belgium: Economic Papers 507.

Moscone F and Tosetti E (2010). Health expenditure and income in the United States. Health Economics 19: 1385-1403.

Musgrove P, Zeramdini R and Carrin G (2002). Basic patterns in national health expenditure. Bulletin of the World Health Organization, Geneva.

Narayan PK, Narayan S and Smyth R (2011). Is healthcare really a luxury in OECD countries? Evidence from alternative price deflators. Applied Economics 43 (15): 3631-3643.

Newhouse J P (1977). Medical-Care Expenditure: A Cross-National Survey. The Journal of Human Resources 12(1): 115-125.

Nordhaus WD (2006). Baumol's disease: a macroeconomic perspective. National Bureau of Economic Research, no. 12118.

OECD, Eurostat, WHO (2011). System of Health Accounts. OECD, Eurostat, WHO, European Union.

Pammolli F, Riccaboni M and Magazzini L (2011). The sustainability of European health care systems: beyond income and aging. European Journal of Health Economics 13: 623-634.

Parkin D, McGuire A and Yule B (1987). Aggregate health care expenditures and national income. Is health care a luxury good? Journal of Health Economics 6(2): 109-127.

Pomp M and Vujic S (2008). Rising health spending, new medical technology and the Baumol effect. CPB Discussion Papers, no. 115.

Przywara B (2010). Projecting future health care expenditure at European level: drivers, methodology and main results, Brussels: Economic Papers 417.

Sen A (2005). Is Health Care a Luxury? New Evidence from OECD Data. International Journal of Health Care Finance and Economics 5: 147-164.

Seshamani M and Gray A (2004). Ageing and health care expenditure: the red herring argument revisited. Health Economics 13 (4): 303-314.

Solow RM (1957). Technical Change and the Aggregate Production Function. The Review of Economics and Statistics 39 (3): 312-320.

Statistics Canada (2011). Health Models. www.statcan.gc.ca/microsimulation/health-sante/healthsante-eng.htm.

Turner D (2016). The Use of Models in Producing OECD Macroeconomic Forecasts. OECD Economics Department Working Paper 1336, OECD Economics Department Working Papers, No. 1336, OECD Publishing, Paris. http://dx.doi.org/10.1787/5jlnb59tmdls-en.

Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C and Luce BR (2003). Principles of Good Practice for Decision Analytic Modeling in Health-Care Evaluation: Report of the ISPOR Task Force on Good Research Practices—Modeling Studies. Value in Health 6(1): 9-17.

White C (2007). Healthcare Spending Growth: How Different Is The United States From The Rest Of The OECD? Health Affairs 26 (1): 154-161.

Yang Z, Norton EC and Stearns SC (2003). Longevity and health care expenditures the real reasons older people spend more. Journals of Gerontology Series B: Psychological Sciences and Social Sciences 58: 2–10.

Yavuz NC, Yilanci V and Ozturk ZA (2013). Is health care a luxury or a necessity or both? Evidence from Turkey. European Journal of Health Economics (14): 5-10.

You X and Okunade AA (2017). Income and Technology as Drivers of Australian Health Care Expenditure. Health Economics 26 (7): 853-862.

Zortuk M and Ceken S (2015). The relationship between health care expenditures and income in the selected transition economies: a panel smooth transition regression analysis. Romanian Journal of Economic Forecasting 18(2): 105-118.

Zweifel P, Felder S and Meiers M (1999). Ageing of population and health care expenditure: a red herring? Health Economics 8: 485–496.

ANNEX 1

CROSS-SECTIONAL STUDIES

Study	Elasticity				Specification	Dependent	Year	Controls	Comments
	Low income	*Middle Income*	*High Income*	***Total***					
Kleiman (1974)*	—	—	—	1.2	Cross-section	Total	1968	—	16 countries.
Newhouse (1977)	—	1.26–1.31	1.13–1.18	1.26	Cross-section	Total	1972	None	13 developed countries. Even with the use of a cross-section, the non-linear effect is found for different levels of income.
Maxwell (1981)*	—	—	—	1.4	Cross-section	—	1975	—	10 countries.
Parkin (1987)*	—	—	—	1.18	Cross-section	Total	1980	—	18 countries, HPPP-deflated spend.
Gerdtham (1992)	—	—	1.29–1.44	1.44 (0.08)	Cross-section	Total	1987	Ageing, health prices, urbanization, substitution, public provision/financing	19 OECD countries. PPP/HPPP-deflated yield same results. Notes how GDPs across OECD countries vary greatly.

Table 1: Summary of elasticities for cross-sectional specifications. Total represents the value of the authors' preferred model specification (usually the one with the most controls applied). Standard errors in brackets, when reported by the authors. * indicates the paper for the study is unavailable for consultation and values are extrapolated from other sources.

NATIONAL LEVEL STUDIES

Study	Elasticity				Specification	Dependent	Year	Controls	Comments
	Low income	*Middle Income*	*High Income*	***Total***					
Di Matteo (1998)	—	—	0.77	0.77	Pooled OLS	Public	1965–1991	Ageing (over 65), federal transfers, EPF	Canada provincial expenditures only.
Freeman (2003)	—	—	0.81–0.84	0.84	DOLS	Total	1966–1998	Time and country fixed effects	United States data. Unit roots and cointegration analysis performed.
Moscone (2010)	—	—	0.36–0.96	0.43 (0.05)	Panel (fixed)	Total (individual)	1980–2004	None	49 US states. Extensive analysis of unit roots. Estimates with fixed effects and CCE (common correlated effects) yield very different results.
Yavuz (2013)	—	0.75	—	0.75	Panel (ARDL)	Total	1975–2007	Ageing (over 65), infant mortality, physicians per 1000	Turkey data. Use of lags for long-term projections show that elasticity will gradually go up to >1. Good summary table of income elasticities in the literature.
Acemoglu (2013)	—	—	0.55–1.13	0.72 (0.21)	Time series variation	Total	1970–1990	Hospital beds and expenditures, population, year fixed effects	United States data. Use of oil prices time series to extrapolate health care income elasticity.

Table 2: Summary of elasticities for national level specifications. Total represents the value of the authors' preferred model specification (usually the one with the most controls applied). Standard errors in brackets, when reported by the authors.

AGGREGATE PANEL STUDIES

Study	Elasticity				Specification	Dependent	Year	Controls	Comments
	Low income	*Middle Income*	*High Income*	***Total***					
Hitiris (1992)	—	—	0.66–1.09	1.02 (0.02)	Pooled OLS	Total	1960–1987	Ageing (over 65), share of public spend, country fixed effects	20 OECD countries. Sen notes they do not account for year effects (fixed) and therefore biased.
Di Matteo (2003)	—	—	0.70–1.21	0.70–1.21	Pooled OLS (non-parametric)	Total	1980–1997 1965–2000 1960–1997	Ageing (over 65), federal cash transfers, country fixed effects	Three panels (US, Canada, 20 OECD countries respectively). Use of non-parametric methods that do not impose linearity. Like Farag, finds that middle-income countries have higher elasticities. Like Getzen, finds that aggregating generates higher elasticities.
Dreger (2005)	—	—	0.67–1.13	0.73 (0.14)	Panel (fixed)	Total (private deflator)	1975–2001	Ageing (over 65), life expectancy, infant mortality	21 OECD countries. Unit roots and cointegration tests taken. Three proxies for medical progress used.
Sen (2005)	—	—	0.51	0.51 (0.20)	Panel (fixed)	Total	1990–1998	Ageing (over 65), life expectancy, infant mortality, physicians per 1000, avg length of stay, country/year fixed effects, linear trends	13 OECD countries. Sen uses various specification, he also finds that OLS is a luxury and all the other models are not (adding fixed effects and controls).

Baltagi (2010)	—	—	0.44–0.89	0.67–0.86 (0.05)	Panel (fixed)	Total	1971–2004	Ageing (dependency rates), public share	20 OECD countries. Very detailed with in-depth econometric analysis. Tests for unit roots, non-stationarity, cross-section dependence are applied and discussed.
Narayan (2011)	—	—	0.74–0.82	0.80	Panel (fixed)	Total	1972–2004	Ageing (over 65), time trends	G18 from OECD (G13 and G5 specifications also used). Unit roots and non-stationarity applied and discussed.
Bech (2011)	—	—	0.28– 1	0.93	Panel (fixed)	Total	1980–2003	Ageing (over 65/75), mortality, alcohol consume, public share, health system dummies	EU-15. Co-integration panel approach that tries to estimate short and long run elasticities.
Holly (2011)	0.39 (0.06)–0.92 (0.04)	0.46 (0.08)–0.75 (0.06)	0.15 (0.06)–0.94 (0.07)	0.15–0.46 0.75–0.94	Panel (fixed, dynamic)	Public, OOP, Total	1995–2008	Ageing (over 60), public share, disease pattern, system characteristics, time trends, substitution	143 countries (pop over 300k). Use of different specifications for different expenditures which might interact differently with drivers. Most cohesive study. Impossible to test for reverse causality in FE. Only totexp estimates reported (in line with other studies).
Farag (2012)	0.51 (0.10)	0.87 (0.07)	0.64 (0.10)	0.83–0.9 (0.04)	Panel (fixed)	Total	1995–2006	Ageing, voice and accountability, GINI, life expectancy, physicians per 1000, newborns, FE, proxy	174 countries. Instrumental variable approach for endogeneity of controls (governance and ageing) and proxies for demand and supply. First OLS model finds luxury, every other specification finds necessity. Strong evidence.
Zortuk (2015)	—	0.42–0.68	—	0.42	Panel (transition)	Total	1995–2011	Ageing, year fixed effects	11 transition countries (Eastern Europe). Elasticities are calculated individually for each economy and various controls are employed.

Table 3: Summary of income elasticities found in the literature. Total represents the value of the authors' preferred model specification (usually the one with the most significant controls and tests applied). Reported standard errors in brackets.

BAUMOL'S COST DISEASE STUDIES

Study	Estimates			Specification	Dependent	Year	Controls	Proxy and countries
	Ageing	*GDP*	*Baumol*					
Hartwig (2008)	—	0.50**	0.85**	Panel (random, fixed), OLS	Total current spending	1960–2004	Income	Baumol = Growth rate of nominal wages per worker (per capita wage growth + real GDP growth) – productivity growth. 19 OECD countries. Reported estimates from RE model controlling for income.
Pomp and Vujic (2008)	0.33 (0.22)	0.86*** (0.27)	0.64** (0.26)	Panel (fixed), FD, LD	Real health spending	1970–2004	Income, ageing (65+)	Baumol = Economy-wide productivity growth (GDP per hour worked). 24 OECD countries. Reported estimates are based on LD in 5-year splines for a smaller sample of 11 countries.
Hartwig (2011)	0.59*	0.93*	0.48*	Panel (random, fixed), OLS	Total current spending	1971–2003	Income, ageing (65+)	Baumol = Relative medical care prices. 9 OECD countries. Reported estimates from time FE model with all controls applied.
Colombier and Carsten (2012)	0.17* (0.10)	0.65*** (0.09)	0.19*** (0.04)	Panel (fixed, random)	Total current spending	1965–2007	Income, ageing (65+), morbidity, technology	Baumol = Growth rate of nominal wages per worker – productivity growth, adjusted by 1/(share of Baumol-related employment). 20 OECD countries. Reported estimates are based on FE model with all controls applied.
Bates and Santerre (2013)	0.35***	0.45***	0.54***	Panel (fixed), FD, 2SLS	Nominal health spending*	1980–2009	Income, ageing (65+), employment, poverty, union coverage	Baumol = Nominal wage per worker – economy-wide output per worker. 50 US states. Reported estimates are based on 2SL2 results (growth rates) divided by the health care labor share. *Spending for private ambulatory care, hospital and nursing and residential care industries only (78% of total).

Medeiros and Schwierz (2013)	—	0.83***	0.96***	Panel (fixed)	Total current spending	1960–2011	Income	Baumol = Colombier specification. 22 OECD countries. Part of a larger model of determinants of health care which does not include Baumol.
Ho et al. (2014)	-0.45** (0.21)	0.25*** (0.08)	0.001** (0.00)	Panel (random)	Public and private spending per capita*	2002–2010	Income, ageing, government debt, pollution, health quality	Baumol = Colombier specification. Chinese provinces. *Spending is about 85% of the total and doesn't include social spending, similar to Bates & Santerre. Health quality is measured by a set of hospital characteristics.

Table 4: Selected studies on Baumol's disease. Standard errors in brackets, when reported by the authors. *** indicates significance at the 1% level, ** at the 5% level, * at the 10% level.

OECD HEALTH WORKING PAPERS

A full list of the papers in this series can be found on the OECD website: http://www.oecd.org/els/health-systems/health-working-papers.htm

No. 94 UNDERSTANDING VARIATIONS IN HOSPITAL LENGTH OF STAY AND COST: RESULTS OF A PILOT PROJECT (2017)
Luca Lorenzoni and Alberto Marino

No. 93 MEASURING SOCIAL PROTECTION FOR LONG-TERM CARE (2017)
Tim Muir

No. 92 CYCLICAL VS STRUCTURAL EFFECTS ON HEALTH CARE EXPENDITURE TRENDS IN OECD COUNTRIES (2017)
Luca Lorenzoni, Jonathan Millar, Franco Sassi, Douglas Sutherland

No. 91 AN OECD ANALYSIS OF HEALTH SPENDING IN NORWAY (2017)
David Morgan, Michael Gmeinder, Jens Wilkens

No. 90 HOW OECD HEALTH SYSTEMS DEFINE THE RANGE OF GOOD AND SERVICES TO BE FINANCED COLLECTIVELY (2016)
Ane Auraaen1, Rie Fujisawa, Grégoire de Lagasnerie, Valérie Paris

No. 89 *THE ORGANISATION OF OUT-OF-HOURS PRIMARY CARE IN OECD COUNTRIES* (2016)
Caroline Berchet, Carol Nader

No. 88 *HEALTH CARE COVERAGE IN OECD COUNTRIES IN 2012* (2016)
Valérie Paris, Emily Hewlett, Ane Auraaen, Jan Alexa, Lisa Simon

No. 87 *PHARMACEUTICAL EXPENDITURE AND POLICIES: PAST TRENDS AND FUTURE CHALLENGES* (2016) Annalisa Belloni, David Morgan, Valérie Paris

No. 86 *THE LABOUR MARKET IMPACTS OF OBESITY, SMOKING, ALCOHOL USE AND RELATED CHRONIC* (2016) Marion Devaux, Franco Sassi

No. 85 *INTERNATIONAL COMPARISON OF SOUTH AFRICAN PRIVATE HOSPITALS PRICE LEVELS* (2015) Luca Lorenzoni, Tomas Roubal

No. 84 *PUBLIC EXPENDITURE PROJECTIONS FOR HEALTH AND LONG-TERM CARE FOR CHINA UNTIL 2030* (2015) Luca Lorenzoni, David Morgan, Yuki Murakami, Chris James

No. 83 *EMERGENCY CARE SERVICES: TRENDS, DRIVERS AND INTERVENTIONS TO MANAGE THE DEMAND* (2015) Caroline Berchet

No. 82 *MENTAL HEALTH ANALYSIS PROFILES* (MhAPs) *Sweden* (2015) Pauliina Patana

No. 81 *MENTAL HEALTH ANALYSIS PROFILES* (MhAPs) *England* (2015) Emily Hewlett and Kierran Horner

No. 80 *ASSESSING THE IMPACTS OF ALCOHOL POLICIES* (2015)
Michele Cecchini, Marion Devaux, Franco Sassi

RECENT RELATED OECD PUBLICATIONS

OBESITY UPDATE (2017)
http://www.oecd.org/health/obesity-update.htm *(Electronic Format Only)*

PRIMARY CARE REVIEW OF DENMARK (2017)

NEW HEALTH TECHNOLOGIES - MANAGING ACCESS, VALUE AND SUSTAINABILITY (2017)

TACKLING WASTEFUL SPENDING ON HEALTH (2017)

OECD REVIEWS OF HEALTH SYSTEMS: LATVIA (2016)

HEALTH AT A GLANCE: EUROPE (2016)

HEALTH AT A GLANCE: ASIA/PACIFIC (2016)

BETTER WAYS TO PAY FOR HEALTH CARE (2016)

HEALTH WORKFORCE POLICIES IN OECD COUNTRIES: RIGHT JOBS, RIGHT SKILLS, RIGHT PLACES (2016)

OECD REVIEWS OF HEALTH CARE QUALITY: UNITED KINGDOM (2016)

OECD REVIEWS OF HEALTH SYSTEMS: MEXICO (2016)

OECD REVIEWS OF HEALTH SYSTEMS: COLOMBIA (2015)

HEALTH AT A GLANCE (2015)

HEALTH DATA GOVERNANCE: PRIVACY, MONITORING AND RESEARCH (2015)

FISCAL SUSTAINIBILITY OF HEALTH SYSTEMS: BRIDGING HEALTH AND FINANCE PERSPECTIVES (2015)

OECD HEALTH STATISTICS (2016) (database available from http://www.oecd.org/health/healthdata)

ADDRESSING DEMENTIA – THE OECD RESPONSE (2015)

CARDIOVASCULAR DISEASE AND DIABETES – POLICIES FOR BETTER HEALTH AND QUALITY OF CARE (2015)

DEMENTIA RESEARCH AND CARE: CAN BIG DATA HELP? (2015)

FIT MIND, FIT JOB – FROM EVIDENCE TO PRACTICE IN MENTAL HEALTH AND WORK (2015)

TACKLING HARMFUL ALCOHOL USE – ECONOMICS AND PUBLIC HEALTH POLICY (2015)

MAKING MENTAL HEALTH COUNT: THE SOCIAL AND ECONOMIC COSTS OF NEGLECTING MENTAL HEALTH CARE (2014)

GEOGRAPHIC VARIATIONS IN HEALTH CARE: WHAT DO WE KNOW AND WHAT CAN BE DONE TO IMPROVE HEALTH SYSTEM PERFORMANCE? (2014)

PAYING FOR PERFORMANCE IN HEALTH CARE: IMPLICATIONS FOR HEALTH SYSTEM PERFORMANCE AND ACCOUNTABILITY (2014)

OECD REVIEWS OF HEALTH CARE QUALITY –AUSTRALIA (2015)

OECD REVIEWS OF HEALTH CARE QUALITY – JAPAN (2015)

OECD REVIEWS OF HEALTH CARE QUALITY – ITALY (2015)

OECD REVIEWS OF HEALTH CARE QUALITY – PORTUGAL (2015)

OECD REVIEWS OF HEALTH CARE QUALITY – CZECH REPUBLIC (2014)

OECD REVIEWS OF HEALTH CARE QUALITY – NORWAY (2014)

For a full list, consult the OECD health web page at http://www.oecd.org/health/