NBER WORKING PAPER SERIES

UNDERSTANDING CROSS-COUNTRY DIFFERENCES IN HEALTH STATUS AND EXPENDITURES

Raquel Fonseca
François Langot
Pierre-Carl Michaud
Thepthida Sopraseuth

Working Paper 26876
http://www.nber.org/papers/w26876

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2020, Revised February 2023

Corresponding Author: François Langot, flangot@univ-lemans.fr. We thank Mariacristina De Nardi, Eric French, Olivia Mitchell, Josep Pijoan-Mas, Jose Victor Rios Rull and Pascal St-Amour for helpful comments as well as seminar and conference participants at ASFE annual conference, GATE workshop on "Heterogeneity and Health", CEMFI, SCSE, LACEA, CIRANO, UQAM, Cergy University, and Evry University. This research was supported by the National Institute on Aging, under grant R01AG030824, PANORisk Regional grant (Pays de la Loire, France), and Institut Universitaire de France. This research is also part of the program of the Research Chair in Intergenerational Economics. Fonseca and Michaud are immensely indebted to the late James P. Smith, who left us in 2022, for stimulating their interest in the understanding of the origins of health differences across time, space and socioeconomic status. Errors are our own. Edited by Harald Uhlig. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Understanding Cross-country Differences in Health Status and Expenditures
Raquel Fonseca, François Langot, Pierre-Carl Michaud, and Thepthida Sopraseuth
NBER Working Paper No. 26876
March 2020, Revised February 2023
JEL No. E21,H51,I10

ABSTRACT

Using a general equilibrium heterogeneous agent model featuring health production, we quantify the contribution of health price in explaining cross-country differences in health expenditures and health status. Considering other country-specific explanatory factors, U.S. health prices are estimated to be 33% higher than those of European countries. This price differential explains more than 60% of the difference in health expenditures and more than half of the difference in health status between Europe and the U.S. Despite its large impact at the aggregate level, these price differences increase the lifetime cost of living of Americans by two percentage points.

Raquel Fonseca
ESG UQAM
320, rue Sainte-Catherine Est
Montréal (Québec) H2X 3X2
Canada
and CIRANO
rfonseca.benito@gmail.com

François Langot
Le Mans University
Paris School of Economics
Avenue Olivier Messiaen
72085 Le Mans cedex 9
France
Francois.Langot@univ-lemans.fr

Pierre-Carl Michaud
HEC Montréal
3000 Côte-Sainte-Catherine Road
Montréal QC H3T 2A7
and NBER
pierre-carl.michaud@hec.ca

Thepthida Sopraseuth
CY Cergy-Paris University, Thema
33 boulevard du Port
95011 Cergy-Pontoise
France
thepthida.sopraseuth@u-cergy.fr

1 Introduction

A considerable variation in health expenditures is observed among different countries. In 2019, the United States (U.S.) spent 16.8% of its GDP on health while Germany spent 11.7% and Italy 8.7% (Health Statistics, OECD).[1] Yet, the intensity of health expenditure does not appear to be strongly correlated with health outcomes at the aggregate level. For example, Americans have been repeatedly found to be in worse health along with having shorter life spans than Europeans (Banks et al., 2006, Institute of Medicine and National Research Council, 2013). It is tempting to use this low correlation as evidence to show that the marginal productivity of health expenditures is low. However, many distortions blur this picture.

There is a vast body of literature relating cross-country differences in health and health expenditures to differences in income (Gerdtham and Jonsson, 2000) and underlying health risk factors (Thorpe et al., 2007). More fundamentally, health expenditures mask price and quantity variation in health care and there is compelling evidence of cross-country price differences for the same health services (Anderson et al., 2003, Cutler and Ly, 2011). This study aims to measure the size, impact, and welfare costs of price differences.

We construct a parsimonious general equilibrium heterogeneous agents model *à la* Aiyagari (1994), augmented with health production as in Grossman (1972) where health provides a direct utility benefit. We parameterize many of the relevant differences across countries in the economic environment along with distortions in the provision of health care, which leads to variation in the price of health services, and differences in the production technology of goods for each country, in particular Total Factor Productivity (TFP). In the model, agents face country-specific risks (health and income) and benefit from health insurance that partially insures against health expenditure risk. The price of health services is determined by a physician-provider-payer model with information frictions, administrative costs, and imperfect competition, leading to price distortions.

Although the model imposes constraints on the richness of institutional differences across countries, we show that it is suitable for estimation using existing micro and macro data. Our framework allows identifying and estimating the extent of differences in the price of health services at the ag-

[1] From https://www.oecd.org/health/health-statistics.htm. We refer to health expenditure as the final consumption of health care goods and services (i.e., current health expenditure). This includes personal health care (curative care, rehabilitative care, long-term care, ancillary services, and medical goods) and collective services (prevention and public health services as well as health administration), but excludes spending on other investments.

gregate level. The relative price of health services, country-specific health risks and TFP differences are directly estimated along with other parameters of the model. We estimate the welfare costs of these differences using a Hicksian compensating variation measure adapted to our dynamic setting. We are able to estimate the effect of these differences on the income-health gradient which also varies substantially across countries because of the cross-sectional dimension of the model (Avendano et al., 2009, Smith, 1999). We show that the income-health gradient is useful for the identification of parameters in the model.

We find that the United States is characterized by the highest (relative) price of health services while one of the highest efficiency for producing goods.[2] We estimate that the price of health services in the United States is 33% greater than the average price level in European countries. Using a counterfactual experiment, we estimate that, if the U.S. had price levels similar to those found in Europe, the U.S.-Europe gap in health expenditures would have been reduced by 68% and the gap in fraction of healthy people by 58%. Estimated price levels are lower in countries with more competition, higher incentives for quality of care, more stringent price regulation, and lower administrative costs. Differences in country-specific health risks, which are correlated with health behaviors (i.e., obesity, smoking, physical inactivity, and drugs), also account for a substantial share of the differences in health spending and health outcomes. In terms of welfare, the additional cost of living that Americans bear because of higher prices is equivalent to 2 percentage points of lifetime consumption expenditures.

The paper is structured as follows. In section 2, we document substantial differences in the price of health services across countries and discuss other sources of variation that can explain differences in health expenditure and health services across countries. In section 3, we present the model used to capture these differences and quantify the welfare effects. In section 4, we present the data, estimation method and report the estimates derived from the model. In section 5, we perform several counterfactual experiments aimed at understanding the relative strength of various factors that explain cross-country differences in health and health expenditures. We then estimate the welfare implications of these differences (section 6). Finally, section 7 concludes the paper.

[2] We consider eight countries in the estimation: Denmark, France, Germany, Italy, The Netherlands, Spain, Sweden, and the United States. We have limited the set of countries because of data availability and to ensure the comparison among countries with similar levels of economic development. In the subsequent sections, we refer to Europe as the set of European countries considered for the study. We argue that they are fairly representative of the heterogeneity across the continent (Western Europe).

2 Price Differences Across Countries and the Growth of Health Expenditures

2.1 Growth Accounting

Differences in economic growth and population aging partially explain the cross-country differences in health expenditures (Gerdtham and Jonsson, 2000). How much do health expenditures vary across countries after accounting for these known factors? We estimate using OECD data over the period 1970-2007 a regression of log health expenditures on the log of GDP per capita and the log of the age 65+ population share, capturing population aging. The regression includes a common time trend (linear splines per decade) and country fixed effects.[3] We obtain an estimate of the elasticity of health expenditures to GDP per capita of 0.9, which is well within the range found in the literature. The elasticity to the 65+ population share is 0.7. For each country, we can decompose growth into a component due to economic growth, population aging, and a residual. Table 1 reports the implied long-term annual average growth rates using this decomposition. First, we observe that economic growth translates into roughly 1.5 to 2 percentage points long-term annual growth in health expenditures for each country. Second, the contribution of population aging is much smaller and the contribution is similar across countries, if not higher in Europe. Third, the residual component is by far the component that varies the most across countries and for which the U.S. has the highest growth (1.9% compared to 1.1% for the European average). This difference in the rate of growth of the unexplained component doubles the real health expenditures per capita in the U.S. over the period while it implies a 50% growth for the European countries selected for the study. Hence, we focus on explaining this residual to understand what drives much higher health expenditures in the U.S.

Insert Table 1 here

Interestingly, the gap in the residual between the U.S. and other European countries widened considerably in the 1980s. Residual growth, in deviation to the common trend during that decade,

[3] We use both GDP and total health expenditures in National Currency Units (NCU) at 2015 constant prices. From the countries we consider in our analysis, we have to exclude Italy from this exercise because data on health expenditures is only available after 1990. For France, the data is reported every 5 years. We interpolate linearly missing years.

averaged 2.2% in the U.S. while it was virtually zero in Europe. Horenstein and Santos (2018) provide evidence of the emergence of higher markups in the U.S. health care sector in that same decade. These higher markups may lead to higher prices which, in turn, may lead to higher health expenditures, provided the (negative) price elasticity of the demand for health services is not too large.

2.2 Direct Evidence on Prices

We can also look at prices directly. Each country produces a Price index for health services. Such indices are problematic over a long horizon as the basket of goods and services covered and the quality (effectiveness) of these services change over time (Berndt et al., 2001). Hence, we focus on a relatively short period, from 1996 to 2007, to demonstrate how price series diverge considerably between the U.S. and Europe. EuroStat in Europe has produced a comparable Health Price index since 1996 while the Bureau of Labor Statistics (BLS) produces a comparable index for the U.S.[4] First, we adjust health expenditures in real terms using the overall Consumer price index (CPI) in each country and define the share of GDP devoted to health as $s_t = p_t \frac{m_t}{y_t}$, where p_t is the relative price of health services (relative to other goods), m_t is real health expenditures and y_t is real GDP. In some starting year, here 1996, p_{1996} will differ across countries. However, this relative price approximately evolves according to $p_t = (1 + i_{m,t} - i_t)p_{t-1}$ where $i_{m,t}$ is the inflation rate for health services while i_t is the overall inflation rate using the CPI. We use these rates to construct a counterfactual where we purge from the evolution of s_t the excess inflation for health goods and services (imposing $i_{m,t} = i_t$ for all years) over the 1996-2007 period. We denote this counterfactual GDP share of health expenditures as $\widehat{s}_t = p_{1996} \frac{m_t}{y_t}$. In the first panel of Table 2, we see that overall annual inflation in the U.S., measured by the CPI (i_τ) was on average 2.33% compared to 3.62% for health products and services ($i_{m,\tau}$) over this period. The excess inflation was therefore on average 1.20% per year in the U.S., compared to 0.38% per year in our group of European countries. While the U.S. spent 15.04% of its GDP on health in 2007, European countries spent on average 9.55%. If we control for excess inflation over this relatively short time period, we find counterfactual shares

[4] Prices indices for health products and services from EuroStat and BLS are fairly comparable. They are both Laspeyres Indices. One notable difference is that the price index for health products and services of BLS includes insurance premiums while the EuroStat index excludes them. We find that excluding health insurance premiums from the BLS index, or including them in the EuroStat index, has no significant impact on results presented in Table 2. Results are available upon request and we provide details on these price indices in Appendix C.1.

of 12.42% and 9.16% for the U.S. and Europe respectively. Hence, the GDP share in the U.S. is 21.09% higher because of excess inflation while it is on average 4.26% higher in Europe. This suggests that prices played an important role in explaining the higher expenditure on health by the U.S. as compared with European countries.

Insert Table 2 here

Evidence that prices are different for the same procedures or drugs abound (Anderson et al., 2003). In Table 3, we report various price estimates collected from studies seeking to compare prices across countries. First, we can look at particular components of health expenditures, such as the cost of hospital stays or drug prices. Based on OECD Statistics Health data, we find that hospital spending per discharge is lower in Europe compared to the U.S. (ranging from 28% to 74% of U.S. spending). Canada Patented Medicine Prices Review Board (2016) construct a price index for patented drugs in OECD countries. The price index reveals that prices are substantially lower in European countries relative to the U.S. We can also look at specific treatments or conditions. The price of an angiogram and bypass surgery is much higher in the U.S. than in Europe (IFHP, 2013). Yet, 30-day mortality rates for patients admitted with a heart attack (Acute Myocardial Infarction, AMI) are quite similar across countries. This rate is even lower in Germany as compared to the U.S.[5] This suggests that price differences are large relative to outcomes differences, at least for heart attacks. In fact, Garber and Skinner (2008) review evidence on productive efficiency and conclude that other countries have achieved similar health outcomes as the U.S. at a lower cost.

Insert Table 3 here

Cutler and Ly (2011) reviewed many explanations to conclude that a considerable portion of these cost differences come from the administrative burden of managing a complex reimbursement system. In addition, the relationship between physicians, providers (hospitals), and payers (insurers) may lead to important rents due to asymmetric information. Finally, imperfect competition among providers may give rise to significant markups (Horenstein and Santos, 2018). Direct price comparisons have limitations. For instance, higher prices or cost per patient could also reflect higher

[5]Fatality rate after an AMI within 30 days after hospital admission provides a robust indicator for health quality as AMI *(i)* is a common, fatal and costly disease, *(ii)* is not related to the patients' choice of providers, and *(iii)* is an event similarly defined across all countries.

efficiency (quality) and level of service in the U.S., as opposed to Europe where supply restrictions and price controls exist. However, it is hard to find supportive evidence since the difference in health outcomes, such as life expectancy, do not suggest an American advantage. Comparing the residual growth in health expenditures presented in Table 1 and the increase in life expectancy (at 50), the U.S. witnessed an increase in 4.68 years over the 1970-2005 period compared to 4.79 years in Europe. As shown in Table 3, heart attack mortality rates are more similar across countries than what price differences would suggest.

2.3 Health Differences

In fact, we can compare various health outcomes across countries. Table 4 presents comparable health measures for our set of countries. Using data from the Survey of Health, Ageing and Retirement (SHARE, Europe) and the Health and Retirement Study (HRS, U.S.) on the population of age 50 to 75, we observe a much higher prevalence of five common health conditions (hypertension, diabetes, lung disease, heart disease and stroke) in the U.S. In relative terms, the prevalence of these conditions is 40.1 to 76.5% higher in the U.S. than in Europe. Some of these differences may be due to differences in diagnostic thresholds. However, Banks et al. (2006) show that large differences exist even if one accounts for differences in diagnostics. In terms of limitations with activities of daily living (eating, bathing, getting out of bed, etc.), the gap is 57%. The fraction without ADL, a measure of health we will use in this study, is 4.7% higher in Europe than the U.S. Ultimately, this gap matches well the gap in terms of mortality. The U.S. has 4.1% lower expected number of years of life remaining at age 50 (1.31 year deficit). Another important feature of the data across countries is the extent to which health outcomes are different by socio-economic status. Taking ADLs as a measure of health, the fraction in good health is 27.6% higher in the top quartile of the income distribution in the U.S. compared to an average of 6.1% in Europe.[6] There is evidence that this income-health gradient is larger in the U.S. across a large number of health conditions (Avendano et al., 2009, Banks et al., 2009). Hence, not only is health lower but it is also more unequally distributed in the U.S. than in Europe.

Insert Table 4 here

[6] We explain in section 4 how we construct these measures.

Garber and Skinner (2008) reviewed several outcomes and argued that the U.S. does not achieve unambiguous better outcomes given the additional spending. In fact, based on data from the Dartmouth Atlas, Skinner et al. (2005) estimate that 20% of health expenditures in Medicare have insignificant value in terms of health. Chandra and Skinner (2012) argue that rapid growth in U.S. health expenditures is partially driven by the adoption of less cost-efficient treatments in the U.S. In fact, some cost-effective innovations, such as Beta Blockers were implemented in Europe before being implemented in the U.S. (Chandra and Skinner, 2012). In Appendix A, we also show that there is no clear evidence that supply restrictions lead to less access and use of health care in Europe (Table A.1). In fact, the frequency of visits to doctors is higher among Europeans than Americans. Horenstein and Santos (2018) show that excess growth in health expenditures in the U.S. is not driven by factor intensity (capital and labor).

2.4 Putting the Pieces Together

Explaining the underlying sources of differences in health expenditures and health status requires a framework that fits both the cross-country as well as the time series dimension of the data. Such a framework can help isolate the magnitude of potential price differences needed to explain a higher share of resources devoted to health and lower health status at an aggregate level in the U.S. compared to Europe as well as the excess growth witnessed over time. It also allows for normative analysis. We consider a simple static framework which we expand to a complete dynamic general equilibrium model in Section 3.

To illustrate, consider two "countries," $g \in \{US, EU\}$. Within-country heterogeneity is summarized by the presence of agents with different income levels, $y_{g,i}$ (4 quartiles, $i = 1, \ldots, 4$). Income $y_{g,i}$ can be used to buy c units of consumption goods (non-medical) at a price normalized to unity and m units of medical goods (or services) at relative price p_g. Each agent faces health risk. With some probability $\overline{\pi}_g(m)$, an agent is in good health ($h = 1$) and with probability $1 - \overline{\pi}_g(m)$ he is in poor health ($h = 0$). Assume $\overline{\pi}'_g(m) > 0$ and $\overline{\pi}''_g(m) < 0$, where $\overline{\pi}'_g$ is the first derivative of the production (probability) function. Let preferences be represented by an additive utility function, $\frac{c^{1-\sigma}}{1-\sigma} + \phi h$, where ϕ captures the utility benefit of being in good health. Each agent maximizes expected utility $U_{g,i} = \frac{c_{g,i}^{1-\sigma}}{1-\sigma} + \phi\overline{\pi}_g(m_{g,i})$ subject to the constraint $y_{g,i} = c_{g,i} + p_g m_{g,i}$. The interior solution ($m_{g,i} \in [0; y_{g,i}/p_g]$) satisfies

$$p_g \left(y_{g,i} - p_g m_{g,i}\right)^{-\sigma} \quad = \quad \phi \overline{\pi}_g'(m_{g,i}) \quad \forall i. \tag{1}$$

The solution yields a demand function for medical goods as a function of income and the price of medical goods, $m_{g,i}(y_{g,i}, p_g)$. It is instructive to investigate, given differences in income and preferences, what (relative) price levels are consistent with observed health and health expenditures differences. This needs to account for the fact that $m_{g,i}(y_{g,i}, p_g)$ depends on income and prices (among other factors). If we take the iso-elastic specification for consumption as given, we need a functional form for the production function to make progress. In Appendix A.3.1 we show that a two-parameter exponential form, $\overline{\pi}_g(m_{g,i}) = 1 - \exp(-\alpha_{1,g} - \alpha_0 m_{g,i})$, fits well the data on health and health expenditures by income for the U.S. With this function, the parameter α_0 governs the productivity of m while $\alpha_{1,g}$ captures other factors impacting health. As a first approximation, we assume that α_0 is the same across countries. This would be the case if technology diffuses relative fast across countries.[7] In this case, we have that $\overline{\pi}'_{g,i} = \alpha_0(1 - \overline{\pi}_{g,i})$ for each country.

Given these functional forms, and several facts (moments) from the data, we can back out all parameters of this model for each country, including p_{EU}.[8] For the U.S., using as moments the GDP share of health expenditures, the fraction in good health and the health income gradient, we obtain $\{\phi, \alpha_0, \alpha_{1,US}\} = \{3.00, 7.517, 1.466\}$ using equation (1) and the normalizations $y_{US} = p_{US} = 1$. For Europe, the GDP share of health expenditures and the fraction in good health lead to $\{p_{EU}, \alpha_{1,EU}\} = \{0.567, 1.890\}$ using (1). Hence, a lower (relative) price in Europe and better health ($\alpha_{1,EU} > \alpha_{1,US}$) is consistent with the data and this simple framework.[9] Consistent with the data, the model predicts a lower health-income gradient in Europe than in the U.S. ($\frac{\overline{\pi}_{4,EU}}{\overline{\pi}_{1,EU}} = 1.131 < \frac{\overline{\pi}_{4,US}}{\overline{\pi}_{1,US}} = 1.276$). Identification of preference and production parameters is ensured by a relatively small set of moments which can be computed from the data. The extension to a dynamic model involves a similar set of moments and a similar identification scheme.

This simple model also tracks the time series of health expenditure growth in the U.S. if we allow

[7] We discuss the plausibility of this assumption in section 4.

[8] See Appendix A.3.2 for more details on the calibration.

[9] We use limitations in activities of daily living for this exercise. If we instead use other health measures in Table 4, we obtain price estimates ranging from 0.21 to 0.54. Hence, the general conclusion does not depend on the health measure used and the price index for Europe is always much lower than one.

for excess health care inflation. Using the calibration, the income elasticity of health expenditures is 1.22 in the U.S., within the confidence interval from the estimate of 0.9 (standard error = 0.115) uncovered in the growth accounting exercise. It is broadly consistent with the evidence from the literature (Gerdtham and Jonsson, 2000). Hence, the GDP share of health expenditures is not impacted much when income rises, at least over a short horizon. In fact, as Table 1 shows, higher growth in income in the U.S. is not sufficient to explain the faster increase in the GDP share of health expenditures in the U.S. over the long-run. The estimates from the model are consistent with this decomposition. The model also has implications for future growth, suggesting that the share of GDP devoted to health expenditures would eventually stop growing and even decline (slightly) if income was the sole driver. This is due to the fact that the income elasticity declines with the level of income.[10] With income twice as high, the income elasticity would decline from 1.22 to 0.91. Prices on the other hand have the potential of explaining part of the difference in growth over recent decades and sustain higher growth in the future. The price elasticity of health expenditures is below one (around -0.71) and remains below one even at higher levels of income or prices. The direct implication is that the GDP share of health expenditures increases in prices. Consistent with the decomposition in Table 3, this admittedly simple model predicts that an excess inflation for health services in the U.S. of 1.1pp is needed to explain the rise in the GDP share of health expenditures in the U.S. compared to Europe since the early 1980s (See Appendix A.3.3).

Up to this point, we have left out health insurance from discussion. Fonseca et al. (2021) show that the increased generosity of health insurance, in particular with the introduction of Medicare in the late 1960s may help explain the rapid rise of health expenditures in the U.S. However, a commonly held belief is that the U.S. is, in aggregate, less generous that Europe in terms of health insurance. But if we measure generosity as the share of total health expenditures paid by consumers (out-of-pocket), OECD data reveals that this share is 13% in the U.S. compared to 15% in Europe. According to that measure, some European countries are even less generous than the U.S. This is due for the most part to the exclusion of certain type of medical services, in particular drugs and supplementary health services, from public coverage in some countries. If we introduce in our model a co-insurance rate, μ_g, and a contribution rate to finance public health expenditures, τ_g,

[10] Recall that the changes in the GDP share of health expenditures ($s = \frac{pm}{y}$) are given by $\frac{ds}{s} = (1 + \varepsilon_{m|p})\frac{dp}{p} + (\varepsilon_{m|y} - 1)\frac{dy}{y}$ where $\varepsilon_{m|p}$ and $\varepsilon_{m|y}$ are respectively the price and income elasticities of health expenditures.

we obtain a new budget constraint $(1-\tau_g)y_{g,i} = c_{g,i} + \mu_g p_g m_{g,i}$. Redoing the calibration exercise with a contribution rate that finances total public health spending (see Appendix A.3.4), we obtain $\{\phi, \alpha_0, \alpha_{1,US}\} = \{0.40, 7.55, 1.47\}$ and $\{p_{EU}, \alpha_{1,EU}\} = \{0.46, 1.70\}$. Hence, the estimated (relative) price difference is slightly lower due to the presence of health insurance. Due to redistribution, the predicted income-health gradient is also flatter when insurance is introduced. Interestingly, the elasticity of health expenditures to a change in μ, keeping income and the contribution rate constant, equals -0.66 with this static model. Although price sensitivity is larger than gold standard estimates produced by the RAND Health Insurance Experiment (reported elasticities between -0.2 and -0.3), this simple model yields a demand for health services which is inelastic to price and roughly unit elastic to income, consistent with the literature (Manning et al., 1987, Newhouse, 1992).

This simple framework illustrates a path that can be followed to recover information on relative price levels and health risks, through the lens of an economic model. Ultimately, this strategy accounts for behavioral responses to income and price differences. In other words, the relative price differences consistent with observed data and preferences account for the fact that medical consumption is increasing in income and decreasing in price. Therefore, estimating preferences and the health production function is crucial for gauging price differences consistent with observed data. Also, the preference specification has implications for health variations across individuals with different levels of income. To be plausible, these specifications should lead to price and income elasticities for medical consumption that are consistent with evidence from the empirical literature.

In a dynamic setting, differences in health insurance, income risk, imperfect capital markets and total factor productivity (production efficiency), need to be taken into account to properly assess the role of price and health risks differences. Furthermore, general equilibrium effects may be significant. In the subsequent section, we present a parsimonious estimable dynamic general equilibrium model that allows filtering price differences from health risks, income and other drivers of health expenditures and health status. Such a model can be used to perform counterfactuals and compute welfare effects.

3 General Equilibrium Model

We build a tractable model that can be used for estimation utilizing the information available across countries. We consider a stationary economy with infinitely lived agents and abstract from population aging because aging does not drive the cross-country differences (see Table 1). On the demand side, the model features agents who value health and therefore have a demand for health services.[11] Agents value health in the model because it provides a direct utility flow. Since health is persistent, health is modeled as a stock with transition probabilities taken as a function of the consumption of health services. On the supply side, we adopt a simplified framework that delivers price differences which are incorporated in the model and summarized by the (relative) price level of health services. Price differences in the model are explained by differences in administrative costs, distortions created by imperfect information in the physician-provider relationship, and imperfect competition among providers, leading to markups. Government regulations also impact each of these distortions. This model leads to a relative price of health services which is estimated from the data. After estimating this relative price level, we compute correlations with indicators of underlying distortions across countries. Finally, the modeling of the demand and supply of health services is incorporated in a heterogeneous agent general equilibrium with incomplete capital markets and uncertainty in wages (Aiyagari, 1994) as well as contributions (taxes or premiums) to fund health insurance. Uncertainty in wages is important as it drives precautionary savings and can also impact the desire of agents to invest in healthcare to avoid financial risks associated with health.

3.1 Households

Agents are heterogeneous with respect to their productivity level $e \in \mathcal{E}$, discrete health status $h \in \mathcal{H}$ and asset holdings $a \in \mathcal{A}$. Agents are either in good health ($h = 1$) or in bad health ($h = 0$). The probability $\pi(h' = 1|h, m)$ of being in good health ($h' = 1$) next period, given the current health

[11] In the spirit of Grossman (1972), there is a literature proposing dynamic life-cycle models of health and other economic decisions. For example, Hall and Jones (2007) and Fonseca et al. (2021) build models with endogenous health to account for the growth in health expenditures and health in the U.S. De Nardi et al. (2010) and Hugonnier et al. (2013) investigate wealth and health dynamics at older ages using models with endogenous health. In a cross-country setting, data by age is limited, both in terms of behavior and outcomes as well as in terms of institutional differences over the life-cycle. Hence, we do not include life-cycle dimensions in the model. Papanicolas et al. (2020) show that the gap in health expenditures between the U.S. and a comparison group, including some countries we consider, is independent of age. Koijen et al. (2016) do not account for age in a model with endogenous health investment where they study the impact of policy uncertainty on health expenditures and medical R&D.

status h, depends on the choice of health services m. It can be interpreted as a health production function and probabilities are endogenous to the choice of m. The probability of being in bad health is given by $\pi(h' = 0|h, m) = 1 - \pi(h' = 1|h, m)$, $\forall h, m$.

Preferences. Agents value both their consumption and their health status. Households' preferences are represented by expected discounted utility

$$\sum_{t=0}^{\infty} \beta^t \sum_{e^t} \sum_{h^t} \widetilde{\pi}_t(e^t)\pi_t(h^t, m_t)u(c_t, h_t), \tag{2}$$

where $0 < \beta < 1$ is the time discount factor, and $c \geq 0$ is consumption of goods.[12] We assume that the instantaneous utility is additive in consumption c and health h:

$$u(c, h) = \frac{c^{1-\sigma}}{1-\sigma} + \phi h, \tag{3}$$

with $\phi > 0$ the utility benefit of good health, and σ is a risk aversion parameter. Our specific formulation is similar to Hall and Jones (2007) adapted to discrete health.[13]

In the model, agents live infinitely. This leads to two simplifications. First, the benefit of investing in health to reduce mortality is not explicitly modeled. However, the parameter ϕ captures both the quality of life and life extension benefits of investing in health since infinitely-lived agents invest in health to increase utility each year as well as the number of years they stay healthy (since health is persistent). Hence, we argue that the quantity of life years trade-off is partially captured by our specification.[14] The second potential limitation is that we do not model the possibility of large expenditure at older ages, and in particular at the end-of-life and its difference across countries. However, recent evidence suggests that end-of-life spending is much lower than previously thought and the share of total health expenditures devoted to the last years of life is relatively constant across countries (French et al., 2017).

[12] Denote by e^t and h^t the histories of productivity levels and health status up to and including time t. A Markov process $\{e \in \mathcal{E}, \widetilde{\pi}(e'|e), \widetilde{\pi}_0(e_0)\}$, where $\widetilde{\pi}(e'|e)$ is a transition matrix and $\widetilde{\pi}_0(e_0)$ its initial value, induces distributions $\widetilde{\pi}_t(e^t)$ over time-t histories e^t Another Markov process $\{h \in \mathcal{H}, \pi(h'|h, m), \pi_0(h_0, m_0)\}$ induces distributions $\pi_t(h^t, m_t)$ over time-t histories h^t for an optimal choice for health service, m_t.

[13] Hall and Jones (2007) use $u(c, h) = b + \frac{c^{1-\sigma_c}}{1-\sigma_c} + \phi\frac{h^{1-\sigma_h}}{1-\sigma_h}$ where h is continuous. Since we do not model life extension benefits, b is zero in our framework.

[14] Given the additive nature of instantaneous utility, expected utility in (2) is composed of two terms: the first involving consumption and the second a discounted sum of the probability of being in good health times ϕ.

Health Production. Each agent can spend his resources on consumption c and health services m. Health services m improve the probability of being in good health next period. Next period's variables are denoted with a prime. In addition, we assume that the production function that maps health services in health status is

$$\pi(h' = 1|h, m) = 1 - \exp(-(\alpha_0 m + \alpha_{1h})). \tag{4}$$

Parameters α_{10} and α_{11} are exogenous and govern both the level and persistence of health, conditional on m, while α_0 captures the productivity of m. One can interpret α_{11} as being inversely related to health behaviors such as smoking and obesity (the ability to stay in good health). We interpret α_{10} as the individual ability to recover from poor health, irrespective of the extent of health services consumed. In section 4.5, we show that our estimates of these parameters are correlated with observable indicators of health behaviors, such as obesity, smoking, and years lost due to health behaviors.

Resource Constraint. Labor income is a function of an idiosyncratic stochastic process e that determines the value of efficient labor. For tractability and identification, we assume that health does not affect labor income, hence shutting down one of the mechanisms present in Grossman (1972) that generates investment in health (the investment motive). We discuss an extension of the model that allows for this mechanism in Section 4.7.[15] The term e is an AR(1) persistent shock, $e_{t'} = \rho_e e_t + \nu_t$ where the variance of the shock ν_t is σ_e^2. Financial market incompleteness prevents agents from insuring against this risk. In addition to labor income, agents collect capital income from asset holding a, with interest rate r. The next period's asset a' is then

$$a' = (1 + r)a + we(1 - \tau) - c - \mu pm, \quad a' \geq 0. \tag{5}$$

[15] The evidence regarding the effect of health on labor income, conditional on participation, is mixed according to the survey of Currie and Madrian (1997). This is partially due to variation across studies in handling measurement error in health and the potential simultaneity between health and labor income. Evidence of a positive effect of good health is more conclusive for hours worked than for wages. Given that the model does not allow for an intensive labor supply margin, we choose to shut down this channel in the model. A second motivation for not including such a pathway is that we already allow labor income to affect health, through the consumption of health services which increases with income. We think that the reverse effect, from health to labor income, would be hard to identify across countries given the data available.

Disposable labor income is net of contributions, at a rate τ, that finances the health insurance system in equilibrium.[16] Disposable income and assets are allocated between consumption c, health services m and saving for the next period. The relative price of health services with respect to consumption is denoted p while the co-insurance rate is denoted μ (the fraction of out-of-pocket expenditures in total health expenditures). By using a constant μ, we make two simplifying assumptions. First, we assume that the insurance system is well captured by a proportional health insurance reimbursement schedule which does not vary according to the type of care received. In practice, the generosity of health insurance may differ depending on the type of care or health insurance agents have. For example, Americans may face different co-insurance rates before and after 65, when eligible to Medicare. Firstly, Fonseca et al. (2021) report that co-insurance rates are similar for employer-provided health insurance and Medicare. However, the U.S. also stands out with a higher fraction of workers without health insurance. Second, co-insurance rates could be non-linear in health expenditures. Recognizing that co-insurance rates may be different across the income distribution, either because of coverage, generosity, or non-linear co-insurance rates, we perform in Section 4.7 a robustness exercise where we allow for μ to depend on e for the U.S. (where data allows us to estimate μ as a function of e).

Demand for Health Services and Savings. For the agent, the state variables are the realizations of the stock of wealth, a, health status h, and the household-specific shock, e. The dynamic program solved by an agent in state (a, h, e) is

$$V(a,h,e) \quad = \quad \max_{m,c} \left\{ \frac{c^{1-\sigma}}{1-\sigma} + \phi h + \beta \sum_{e'} \widetilde{\pi}(e'|e) \left[\begin{array}{l} \pi(h'=1|h,m)V(a',h'=1,e') \\ +(1-\pi(h'=1|h,m))V(a',h'=0,e') \end{array} \right] \right\} \tag{6}$$

subject to equation (5). V denotes the value function. The solution of this problem is a set of decision rules that map the individual state into choices for consumption and health services. We denote these rules by $\{c(a,h,e), m(a,h,e)\}$.

[16] The health sector in the model encompasses all health services in the economy, including public and private health services. τ, therefore, captures the aggregate cost of health services to households, such as private insurance premiums and social security taxes to the health care system.

3.2 The Supply of Health Services

Given the complexity and variety of health care systems, it is difficult to encapsulate the richness of country-specific interactions between providers, physicians, and insurers in the model which may lead to higher prices in some countries compared to others. We discussed in section 2 potential explanations for price differences including productivity, administrative costs, information rents, and imperfect competition.

Our empirical strategy allows us to retrieve price differences in the estimation directly. Therefore, we assume that the sources of these price differences do not have a direct effect on the general equilibrium. Two assumptions are required for this to hold: 1) constant returns to scale of the health services production function, 2) market imperfections or regulation independent of the demand for health services. The first assumption yields constant marginal cost which could differ across countries because of productivity differences in providing health services or administrative costs. The second assumption states that other market imperfections, such as information rents, imperfect competition, or direct price regulation, do not depend directly on demand conditions.[17]

In section 4.4, we present a framework on the supply side of the health services market which allows interpreting price difference estimates across countries. We correlate our estimates with observed indicators of the sources of price differences. This simple model of price determination is consistent with the general equilibrium model we estimate. It leads to a health price that is independent of other aggregates and is different across countries due to information frictions, administrative costs, and market concentration.

3.3 Good-Producing Firm

Production Y is characterized by constant returns to scale using aggregate capital K and labor N as inputs : $Y = AK^{\alpha}N^{1-\alpha}$. A captures technological factor productivity (TFP) and $0 < \alpha < 1$ the capital share in GDP. The firm operates under perfect competition such that profit maximization

[17] In a general equilibrium framework, the use of the profits coming from health market distortions must be specified. For simplicity, we assume that they are used to finance fixed costs and entry costs in this market. Hence the value of total healthcare demand is the sum of goods (input and fixed costs) used to produce it.

leads to:

$$
\begin{aligned}
r &= \alpha A \left(\frac{N}{K}\right)^{1-\alpha} - \delta_k && (7)\\
w &= (1-\alpha) A \left(\frac{K}{N}\right)^{\alpha}, && (8)
\end{aligned}
$$

with w the wage rate, r the interest rate, and δ_k the depreciation rate.

3.4 Health Insurance System

Health insurance reimburses medical expenditures using a fixed contribution rate (tax or premiums) on labor income:

$$
\tau w N = (1-\mu) p \sum_e \sum_h \sum_a m(a,h,e)\lambda(a,h,e) \tag{9}
$$

where $\lambda(a,h,e)$ is the stationary distribution of individuals across individual states (a,h,e). Given the co-insurance rate μ, the contribution rate τ must finance expenditures. Using equation (8), we note that the contribution rate is proportional to the GDP share of health expenditures.

3.5 Definition of Equilibrium

A steady-state equilibrium for this economy is a household value function, $V(a,h,e)$; household optimal policy functions, $\{c(a,h,e), m(a,h,e)\}$; a health insurance system, τ; a stationary probability measure of households, λ; factor prices, $\{r,w\}$; and macroeconomic aggregates, $\{K,N\}$, such that the following conditions hold:

(a.) Factor inputs, contributions to health insurance, and transfers are obtained aggregating over households: $K = \sum_e \sum_h \sum_a a\lambda(a,h,e)$ and $N = \sum_j e_j N_j$.

(b). Given $\{K,N\}$, factor prices $\{r,w\}$ equal their marginal productivities (Equations (7) and (8)).

(c.) Given $\{r,w,\tau\}$, the household's optimal policy functions solves the household's problem (6).[18]

(d.) The contribution rate τ adjusts such that health expenditures supported by insurance are paid, Equation (9) is satisfied.

[18] We model closed economies and do not allow for capital flows among the set of countries we model.

(e.) Goods and health markets clear, $Y = \sum_e \sum_h \sum_a [c(a,h,e) + pm(a,h,e)]\lambda(a,h,e) + \delta_k K$.

(f.) The measure of households $\lambda(a,h,e)$ is stationary.[19]

[19] We create a discrete state-space and solve the household problem using value function iteration. We then compute the stationary distribution and compute equilibrium conditions. We iterate until an equilibrium that meets all the conditions above is obtained. Appendix B contains more information on the solution method.

4 Data and Estimation

We aim to estimate the price of health services along with other parameters of the model for countries $g = 1, ..., G$. We first estimate and calibrate several country-specific parameters, which we label as *auxiliary* parameters. We then estimate *structural* parameters for preferences, health production, the price of health services, and TFP using a method of simulated moments (MSM).

4.1 Auxiliary Parameters

We allow for cross-country heterogeneity in income risk $\{\rho_e, \sigma_e\}$ and the goods-producing technology $\{\alpha, \delta\}$, as well as the generosity of health insurance across countries, μ.[20]

Income Risk. The level of labor income is determined in equilibrium by w. However, income risk, e, needs to be estimated. As the model does not account for social insurance, we account for the insurance effects when estimating income risk by considering a measure of after-tax household income to estimate the process for e. We use data from the Panel Study of Income Dynamics in the U.S. and the European Community Household Panel to estimate the persistence and volatility of income. For each country, we first partial out the effect of age from income by regressing (log) after-tax household income on a set of age dummies and obtain residuals which we use to model shocks, e. We assume the following process: $\eta_t = e_t + u_t$ with $e_t = \rho_e e_{t-1} + \nu_t$, where ν_t is the innovation to the persistent component distributed $N(0, \sigma_e^2)$, whereas the transitory component u_t is distributed $N(0, \sigma_u^2)$. We account for this component to partially control for measurement error as is often done in the literature (French, 2005). Table 5 shows the estimates of the income shock process using a minimum distance estimator.[21] We find considerable persistence in income, with estimates of the auto-correlation coefficient (ρ_e) ranging from 0.912 (Denmark) to 0.972 (Spain). If we compare across countries, the variance of the transitory component, σ_u^2, varies less across countries (from 0.077 to 0.263) than the variance of the persistent component, which, assuming stationarity, is given by $\frac{\sigma_e^2}{1-\rho_e^2}$. We find much more variation for this component with variance estimates ranging from 0.095 (Denmark) to 0.486 (U.S.). The contribution of the persistent component to the total variance of shocks is also larger for the U.S. compared to European countries. Therefore, the main

[20] More details on data sources used to estimate auxiliary and structural parameters can be found in Appendix C.
[21] See Appendix C.2.1 for details on the estimation procedure.

source of difference in this measure of income risk is the scale of the innovation to the persistent income shock. The variance of the persistent shock is approximately twice as large in the U.S. compared to Europe.[22,23]

Insert Table 5 here

Co-insurance Rates. We use OECD Health Data over the period 1995-2015 to compute a measure of how much patients support in terms of health expenditures. The co-insurance rate, μ is defined as private out-of-pocket household expenditures as a percentage of total health expenditures.[24] Table 6 shows estimates of μ across countries. Spain and Italy have larger shares of out-of-pocket health expenditures, while France and the Netherlands have the smallest shares. The U.S. ranks in the middle with a co-insurance rate of 13.6%. However, the U.S. health care system effectively provides substantial coverage against health expenditures. After the age of 65, this coverage is provided by Medicare while it is provided by employer insurance as well as Medicaid (for the poor) before age 65. Overall, Fonseca et al. (2021) show that the median co-insurance rate is similar for employer-provided health insurance and Medicare but it is lower for Medicaid. This is despite around 15% of workers in 2005 without any health insurance. In Section 4.7, we discuss an extension that allows for co-insurance rates to depend on income.[25]

Insert Table 6 here

[22] This is consistent with De Nardi et al. (2021), who show that social insurance reduces substantially the variance of household income in the Netherlands compared to the U.S.

[23] Following French (2005), we assume that the transitory component captures in large part measurement error (we use survey data) and set it to zero in the model. As the variance of the transitory component is relatively comparable across countries, this assumption does not have a sizable effect on explaining differences in health and health expenditures.

[24] Out-of-pocket payments are defined by the OECD as expenditures borne directly by a patient after taking into account public and private insurance coverage. Private health insurance premiums are not included in out-of-pocket payments. Hence, they are captured by the contribution rate τ.

[25] Some European countries, like Spain and Italy, have relatively large co-insurance rates for drugs as well as for some services such as dental and vision care. The U.S. has a significant fraction of uninsured workers with a marginal co-insurance rate close to one. However, uninsured workers use very little health services, which may explain why the overall co-insurance rate, measured as a ratio of out-of-pocket health expenditures as a ratio of total health expenditures, is relatively small for the U.S. In an ideal world, one would want to consider the non-linearity in insurance contracts and consider both the average co-insurance rate (for the budget constraint and setting the contribution) and the marginal co-insurance rate (for the decision to consume health services at the margin). We are not aware of estimates of the marginal co-insurance rate, or non-linear insurance schedule, across countries which would allow us to make a distinction between the two.

Technology of the Good-producing Firms. We use Penn World Table (Feenstra et al., 2015) to estimate the country-specific shares of capital (α) and the depreciation rates (δ_k). Table 6 reports estimates for the period 1995-2015. The share of capital in production (α) is lowest in Denmark (0.36) and largest for Italy (0.47), with an intermediate value for the US (0.384). In the case of the depreciation rate (δ_k), estimates range between 0.038 (Spain) to 0.049 (US).

4.2 Structural Parameters

We follow a two-step method of simulated moments approach to estimate structural parameters. First, a set of common parameters $\{\sigma, \phi, \alpha_0\}$ and U.S. specific parameters $\{\alpha_{10}, \alpha_{11}\}$ are estimated using U.S. data. Identification of β from σ in a heterogeneous agent model is notoriously difficult. Hence, we fix β at 0.97.[26] Second, we estimate the (relative) price of health services, p_g, and TFP (relative to the U.S.), A_g, using the set of common parameters, for a group of seven European countries: Denmark, France, Germany, Italy, the Netherlands, Spain and Sweden.[27] In addition to these country-specific parameters, we allow for variation in parameters $\{\alpha_{10}, \alpha_{11}\}$ across countries to capture unobserved differences in health status (obesity, smoking, etc.). Therefore, the vector of structural parameters we estimate is given by

$$\Theta = \left\{\sigma, \phi, \alpha_0, \{\alpha_{g,10}\}_{g=1}^{G}, \{\alpha_{g,11}\}_{g=1}^{G}, \{A_g\}_{g=1}^{G}, \{p_g\}_{g=1}^{G}\right\}.$$

The assumption of common preferences is commonly made in macro models estimated across countries (Chari et al., 2007, Ohanian et al., 2008). The assumption that α_0 is common to all countries implies that the ability of any m to produce h is the same across countries. This is not as restrictive as it may seem. The marginal cost of producing good health, given by $\frac{p}{\pi'_m(h'|h,m)}$, is country-specific, despite a common α_0. The marginal cost of producing a certain level of health is heterogeneous across countries as the distortions are reflected in the equilibrium price. This is consistent with the view that the U.S. has witnessed widespread adoption of more costly medical treatments without necessarily any significant additional health benefits relative to existing treatments (Chandra and

[26]The exact value of β we use does not have much effect on other estimates and qualitative conclusions from the counterfactuals we perform. The fixed value of β impacts mostly the estimated value of ϕ, the benefit of health in utility, and to some extent the value of σ. Our estimation procedure has no problem fitting the moments for various levels of β suggesting that its value is not distinctly identified from other parameters.

[27]As we explain when discussing identification, we impose the normalization $p_{US} = 1$ as well as $A_{US} = 1$ so that price and TFP estimates are relative to the U.S.

Skinner, 2012, Garber and Skinner, 2008). Hence, we allow for imperfect diffusion of innovations of this type. However, we assume rapid diffusion of medical innovations which are cost-effective across countries. Evidence from Table 3 suggests that price (and cost) dispersion is much larger than dispersion in outcomes (for example measured by AMI case-fatality rates), which is consistent with the assumption of a common α_0 combined with a country-specific p which could reflect the use of more expensive treatments to reach similar health outcomes.

Estimator. We denote the set of country specific auxiliary parameters estimated earlier χ_g and $\chi = \{\chi_1, ..., \chi_G\}$. For each country, consider a set of Ψ_g simulated moments denoted:

$$\psi_g(\Theta_g, \chi_g) = \left\{\psi_{g,1}(\Theta_g, \chi_g), ..., \psi_{g,\Psi_g}(\Theta_g, \chi_g)\right\},$$

while moments from the data are $\psi_{g,N}$. Our objective is therefore to pick parameters Θ_g so that moments from the data and model are as close as possible.[28]

We could stack moments of each country and estimate parameters jointly, but this procedure is numerically difficult as it requires solving the general equilibrium in each country for each trial value of the parameter vector. Hence, it does not exploit the fact that many parameters are country specific and therefore only impact moments for a given country. We first estimate common parameters $\{\sigma, \phi, \alpha_0\}$, and U.S.-specific parameters $\{\alpha_{US,10}, \alpha_{US,11}\}$ using a set of U.S. moments $\psi_{US}(\Theta_{US}, \chi_{US})$.[29] We then estimate country specific parameters $\Theta_{g\neq US} = \{\alpha_{g,10}, \alpha_{g,11}, A_g, p_g\}$ given parameter estimates Θ_{US} using a second set of moments $\psi_g(\Theta_g, \chi_g)$.[30,31]

Moments. We combine a set of aggregate moments and moments derived from micro data to estimate parameters. Identification follows closely the calibration exercise of the simple model in

[28] We use a weighting matrix $W_{g,N}$, a positive definite matrix, which is based on the data. We choose a diagonal matrix with elements equal to the inverse of the variance of each moment. For moments involving micro data, we use the bootstrap method to find their variances. For other moments, we use time-series variations around 2005 to compute their variances.

[29] Formally, we use the following estimator $\Theta_{US} = \arg\min\,[\psi_{US}(\Theta_{US}, \chi_{US}) - \psi_{US,N}]'\, W_{US,N}\,[\psi_{US}(\Theta_{US}, \chi_{US}) - \psi_{US,N}]$ where $\Theta_{US} = \{\sigma, \phi, \alpha_0, \alpha_{US,10}, \alpha_{US,11}\}$. We use the NEWUOA algorithm proposed by Powell (2006).

[30] More formally, we consider the estimator $\Theta_g = \arg\min\,[\psi_g(\Theta_g, \chi_g) - \psi_{g,N}]'\, W_{g,N}\,[\psi_g(\Theta_g, \chi_g) - \psi_{g,N}]$, $\forall g \neq US$ where $\Theta_{g\neq US} = \{\alpha_{g,10}, \alpha_{g,11}, A_g, p_g\}$.

[31] We obtain standard errors as follows. Denote by $D_{g,N}$ the matrix of derivatives of the moment vector relative to parameters for country g. We obtain this matrix numerically at the estimated value of the parameters. When using as weighting matrix $W_{g,N}$, the inverse of the covariance matrix of the data, the variance of estimates collapses to $V_{g,N} = (D'_{g,N} W_{g,N} D_{g,N})^{-1}$ (see Cameron and Trivedi (2005), page 174).

Section 2. The vector of moments for each country g is given by:

$$\begin{aligned} \psi_{US} &= \left\{C/Y, s, \tilde{\pi}_{1|0}, \tilde{\pi}_{1|1}, \overline{\pi}_2, \overline{\pi}_3, \overline{\pi}_4\right\} \\ \psi_{g \neq US} &= \left\{\tilde{Y}_g, s_g, \tilde{\pi}_{1|0,g}, \tilde{\pi}_{1|1,g}, \overline{\pi}_{2,g}, \overline{\pi}_{3,g}, \overline{\pi}_{4,g}\right\}, \end{aligned} \tag{10}$$

where C/Y is the ratio of consumption to GDP; $\tilde{Y}_g$ the GDP per capita relative to the U.S.; s_g the share of health expenditures as a fraction of GDP; $\tilde{\pi}_{1|0,g}$ and $\tilde{\pi}_{1|1,g}$ the transition rates from bad to good and good to good health status; $\overline{\pi}_{i,g}$ the relative probability of being in good health within income quartiles $i = 2, 3, 4$, using the first quartile as a reference group (income gradient).[32]

First, we estimate the five parameters using six moments based on the U.S. data. The parameter σ is pinned down by targeting C/Y. Transition rates by health status, $\tilde{\pi}_{1|0,US}$ and $\tilde{\pi}_{1|1,US}$, inform $\alpha_{US,10}$ and $\alpha_{US,11}$. Parameters $\{\phi, \alpha_0\}$ are determined by the share of health expenditures in GDP and the relative probability of being in good health within income quartiles (or income gradient of health status) $\overline{\pi}_{q,g}$, for $q = 2, 3, 4$.

In a second step, we look at European countries and estimate the country-specific price of health services and productivity (TFP), relative to the US. The health transition rates also allow the identification of $\{\alpha_{10,g}, \alpha_{10,g}\}$ in each country g. GDP per capita relative to U.S. ($\tilde{Y}_g$) pins down A_g. To pin down p_g, we use GDP shares of health expenditures, the transition rates, and the income gradient of health status.

Data. We use two types of data to compute moments: national statistics and micro data. For moments involving national statistics, our approach is to target a period around 2005. To account for variation in moments around that period, we typically select the period 1995-2015 when data is available. We use the ratio of consumption to GDP (C_g/Y_g) and GDP per capita relative to U.S. ($\tilde{Y}_g = Y_g/Y_{US}$) from the Penn World Tables (Feenstra et al., 2015) over the years 1995 to 2015.[33] We use information from OECD Health Data for 1995 to 2015 for the GDP share of health

[32] Given that we target transition rates in the estimation, the markovian assumption on the health process implies that, if we fit transition rates, we also fit the fraction in good health (assuming stationarity). Hence, we do not target the fraction in good health as an additional moment. In counterfactual simulations (Section 5), we show that the fraction in good health is well matched by the model.

[33] Only for the U.S., we use real consumption and real GDP per capita at 2011 level National prices (in millions of 2011 PPP-adjusted U.S. dollars) to compute C/Y over the same period. We compute consumption as a share of GDP consistent with the model. We subtract government consumption from Penn World Table consumption but add health spending which is public, Medicare and Medicaid, i.e. 33% of the share of health expenditures as a fraction of GDP.

expenditures ($s \equiv \frac{pm}{Y}$).

For moments involving micro data, we use two longitudinal surveys to estimate health state transitions. For the U.S., we use the Health and Retirement Study (HRS, waves 2004 and 2006), while for Europe we use the Survey of Health, Ageing, and Retirement in Europe (SHARE, waves 2004 and 2006). We selected respondents aged 50 to 75.[34] We select limitations with activities of daily living (ADL) as a measure of health in the present study. These limitations include whether someone has difficulty with dressing, bathing, getting in and out of bed, eating, and walking across a room.[35] We estimate the probability of having no ADL ($h = 1$) at time $t + 1$ given state j at time t as $\tilde{\pi}_{1|j,g}$. We compute these transition rates for states at t, no ADL ($j = 1$) and with ADL ($j = 0$) in each country g.[36] To compute the health gradient, we use the distribution of after-tax household income in 2005 PPP adjusted U.S. dollars. We use the quartiles of the within-country distribution. We compute the fraction without ADL within each quartile, $\tilde{\pi}_{q,g}$ for $q = 1, 2, 3, 4$. We use as moments the fraction in good health relative to the first quartile: $\overline{\pi}_{q,g} = \overline{\pi}_{q,g}/\overline{\pi}_{1,g}$ for $q = 2, 3, 4$.[37,38]

Insert Figure 1 here

Estimated moments. Figure 1 reports moments from the data. GDP per capita is generally 10 to 35% lower in European countries relative to the U.S. ($\tilde{Y}_g$). The U.S. spends 14.7% of GDP on health (s_g) while only two countries lie above 10% in Europe (France and Germany). In terms of

[34] Two reasons motivate this choice for the age range: 1) Both surveys only include respondents above the age of 50. 2) At older ages, both surveys are less representative of the overall population since the institutionalized population is partially covered in HRS. However, they are not included in the first wave of SHARE and were only partially covered in some countries in 2006. We do not use additional waves of SHARE since the third wave was a retrospective survey and additional waves are a bit far from our observation window (2011).

[35] While one might be interested in additional dimensions of health, the computational burden of doing so prohibits this possibility. Limitations with activities of daily living is a reliable overall health measure, with a strong predictive power on mortality and use of physician services. It is likely to be less affected by reporting scale bias than self-reported health (reported from poor to excellent); see Kapteyn et al. (2007).

[36] Given that surveys measure health every two years, we recompute annual transition rates by solving $\tilde{\Pi}_2 = \tilde{\Pi}_1^2$ for $\tilde{\Pi}_1$ where Π_q is the Markov transition matrix for q year transitions.

[37] We use ratios. Indeed, the health transition rates already pin down the aggregate fraction in good health, as we assume a stationary process. Using ratios allows focusing on the information contained in the gradient for identification.

[38] Using SHARE and HRS, we have tested whether the income-health gradient is sensitive to controls for risk factors such as obesity and smoking. We estimated a logit model for whether the respondent was in good health as a function of quartiles of income fixed effects (capturing the income-health gradient) and controls for obesity and smoking. Including risk factors did not impact the estimated differences in health by income (see Table D.1 in Online Appendix). Hence, the assumption is reasonable. Banks et al. (2009) reach the same conclusion that risk factors do not impact substantially the health-income gradient in the U.S. and England.

transition rates into good health, the U.S. ranks last in terms of transition rates to good health irrespective of the initial state (good or bad). Finally, the health gradient by income quartile is much steeper in the U.S. than in any European country.

4.3 Estimation Results

Estimation results for common parameters of the general equilibrium model are reported in Table 7. Three parameters are common to all countries $\{\sigma, \phi, \alpha_0\}$. Other parameters, prices, TFP, and exogenous health risks are country-specific.

Insert Table 7 here

The coefficient of relative risk aversion, σ, is estimated at 2.09. In the literature, this parameter varies across a wide range depending on the model and focus. Our estimate is very close to the value assumed by Hall and Jones (2007) even though our model is quite different in terms of how health is modeled.[39] Below we investigate the implications of this estimate for the income and price elasticity of health expenditures which provides a more informative test of the plausibility of our preference estimates. The marginal utility of being in good health (ϕ) is estimated to be 0.304. Using steady-state values for consumption and health, this implies that individuals are on average willing to accept a loss of roughly 1% of their permanent consumption to increase their probability of being in good health marginally by one percentage point.[40] This valuation of the welfare benefits of being healthy is slightly smaller than the estimate computed by De Nardi et al. (2018), showing that individuals are willing to pay about 3% of average income to increase the probability of being healthy by one percentage point.[41] The parameter governing the marginal productivity of health investment α_0 equals 0.169. This implies an elasticity of health transition from bad to bad health

[39]The standard errors on this parameter and others are quite small. One of the reasons is that moments based on aggregate statistics are computed over a short period around year 2005. As a result, some parameters are mostly identified from very precisely aggregate moments from the time-series. Hence, these moments will be given a lot of weight when computing standard errors, giving the illusion that we get a very sharp estimate of these parameters, given the uncertainty about model specification, choice of moments, etc.

[40]With a steady-state approximation of the welfare given by $\frac{1}{1-\beta}\left(\frac{c^{1-\sigma}}{1-\sigma} + \Pr(h=1)\phi\right)$, we deduce that consumption reduction compensated by a rise in probability of being in good health is $\delta_c = 1 - \left(1 - \Delta\Pr(h=1)\phi / \left(\frac{c^{1-\sigma}}{1-\sigma}\right)\right)^{1/(1-\sigma)}$. Given the estimated parameters $\{\sigma, \phi, \alpha_0\}$ and $c = 2.329$ and $\Pr(h=1) = 0.899$ in the U.S., we deduce that $\delta_c = -1\%$ with $\Delta\Pr(h=1) = 1\%$.

[41]The estimate of 3% comes from noting that, in De Nardi et al. (2018), 60% of the welfare gain (total of 5%) comes from health benefits and 40% from increased lifetime income. The last effect is not considered in our model.

to medical expenditures of -0.5 at values in steady-state. It is difficult to find comparable studies to benchmark this estimate. For particular diseases (pneumonia, heart attacks, and congestive health failure), Romley and Sood (2013) report elasticities of 30-day mortality rates to medical care intensity ranging from -0.31 to -0.71.

As estimation of preferences is model-specific, we gauge the plausibility of these parameter estimates by computing the elasticity of health expenditures pm to the co-insurance μ generated by the model. For the U.S. this elasticity is -0.379 in partial equilibrium (wage, interest rate, and taxation are kept constant). This estimate is very close to the elasticity found in the RAND Health Insurance Experiment which is in the range -0.2 to -0.3 (Manning et al., 1987). The income elasticity estimate of health expenditures pm is 0.52. Acemoglu et al. (2013) estimate an elasticity ranging from 0.55 (standard error = 0.230) to 0.8 (standard error = 0.155). Our estimate is however lower than macro studies which tend to estimate an elasticity closer to one but higher than micro studies which tend to find estimates below 0.25 (see Gerdtham and Jonsson (2000) for a review of estimates).[42] Hence, our estimates do not suggest that health is a luxury good: higher income cannot lead to a higher GDP share of health expenditures. We conclude that our preference specification does not lead to abnormal price or income responses when compared to reduced-form empirical estimates. Table 8 reports estimates of country-specific parameters: prices, health risks and TFP differences.

Insert Table 8 here

4.4 Price Differences and their Origin

We report price estimates, relative to the U.S. (where it is normalized to one) in Table 8. All European countries have much lower health prices than the U.S. For example, France (0.615), Spain (0.65), the Netherlands (0.698), and Italy (0.713) have prices which are more than 30% lower than in the U.S. while Germany has prices which are 15.3% lower. The price of health services in the U.S. is 33% higher than the average European price which is consistent with the evidence from the static calibration exercises in Sections 2 and 4.2. While the estimated price gap between the U.S. and Europe was larger in the static model, it is worth nothing that static computations did not

[42] We compute this income elasticity for one percent change in the equilibrium wage.

incorporate dynamic and general equilibrium effects. Yet, all these different computations point to a substantial price gap which appears to be robust to other considerations. Also consistent with our estimate for the U.S. relative to Europe, Skinner et al. (2005) estimate that 20% of health expenditures in Medicare appear to have little value based on data from the Dartmouth Atlas of health care.

The relation of estimated price differences to potential drivers of price differences presented in the literature is worth examining. We first present a simple model and then correlate our estimates to proxies for drivers of price difference.

Consider a health care system consisting of three actors: physicians (or other health care professionals), providers (hospitals or health agencies), and payers (insurers). This set of actors is flexible enough to account for the heterogeneity of institutional settings across countries. In particular, it allows us to consider three potential price distortions, for which we have indicators across countries: information rents, imperfect competition and administrative costs. Finally, governments can influence each of these distortions using, for example, price controls.

First, the physician-provider relationship is potentially affected by information frictions which lead to informational rents for physicians.[43] Some countries, such as the U.S., may be particularly prone to having more information frictions because of the highly decentralized nature of the interactions between physicians and providers. It is also because most of the physicians are paid on a fee-for-service basis and potentially have more freedom in choosing treatments for patients. While countries have access to the same monitoring technologies, their use may vary both due to the complexity of the interactions they need to monitor as well as the administrative capacity of governments. Consider a principal-agent model with potential shirking on quality of care. In such a simple model, the incentive-compatible contract that influences the physician to provide high-quality health services pays a price that is decreasing in the shirking detection probability, with a limit equal to marginal cost when the detection probability equals one (quality of care is perfectly observed). This leads to the prediction that countries with more information frictions pay physicians more. Formally, let ζ be the detection probability and z the marginal productivity of physicians in transforming goods (and time) into health services. The optimal payment given a linear production

[43] Information frictions in the provision of health care and induced distortions were first discussed in Arrow (1963).

function is $\frac{1}{\zeta z}$.[44] An alternative, or complement, to a model with information rents is a model with (Nash) bargaining between physicians and providers. Centralized providers, such as national health agencies, often found in European countries, may negotiate lower prices for health services.

Second, consider the decentralized relationship between payers (insurers) and providers (say hospitals). The role of providers is to act as an intermediary between physicians and insurers. When insurers are distinct from providers, there is the potential for considerable administrative burden in managing the payments to physicians and patients with the insurers. The U.S. is a case where providers and payers are often distinct entities while in many countries, payers and providers are integrated (for example in a single-payer setting). Cutler and Ly (2011) document that administrative costs of managing their reimbursement system are much higher in the U.S. compared to other countries. The U.S. has also many more workers in health care administration than countries such as Germany or the Netherlands. Given that much more freedom is given to physicians to choose treatments, the verification of health insurance claims is potentially more costly in the U.S. than in Europe. We can introduce these administrative costs as proportional to the level of health services, ι, such that the price of health services is inflated by an additional markup $1/(1-\iota)$.[45]

A third distortion comes from imperfect competition among providers and the market power it confers to them on the provider-payer side of the market. Assume there are P providers. If these P providers compete in an oligopoly *à la Cournot*, and payers operate in a competitive market with a constant returns to scale technology, the equilibrium price is multiplied by a factor $1/(1-1/P)$ where $1/P$ can be interpreted loosely as market concentration. The U.S. has seen an increase in market concentration, in particular due to mergers among hospitals (Gaynor and Town, 2012).[46]

[44] To get this result, consider that the physician transforms b_h units of the consumption good into health services through the production function $b = zb_h$, where z is the productivity of the physician (or health care professional). The physician's compensation is $p_p b$. For simplicity, assume that the output of the physician can be of high or low quality: $q \in \{0; 1\}$. When quality is high, the provider supplies the adequate service to a patient and collects profit $\Pi_b^h = p_p b - b_h$. When the quality is low ($q = 0$), the physician does not provide the requisite service (shirks), and thus, does not buy any input and collects profit $\Pi_b^s = p_p b$. The provider can detect shirking behavior with probability $\zeta \in [0, 1]$. To maximize profits, the provider will propose an incentive contract such that $p_p = \frac{1}{\zeta z}$.

[45] To get this result, we adapt the physician-provider setting to include administrative costs $\iota p_p b$, with a fraction $\iota > 0$ proportional to firm revenue $p_p b$, where p_p is the physician's compensation. This yields the incentive-compatible price $p_p = \frac{1}{\zeta(1-\iota)z}$.

[46] Using the provider-physician equilibrium contract, which ensures $q = 1$ at the equilibrium, we get the input quantity of the i-type provider $m_r(i) = q(p_p)b = b$, and its total revenue $p_r m_r(i)$, where p_r is the price paid by payers. There are P providers ($i \in [1, P]$) producing each a single health service. The aggregate demand $\mathcal{D}$ is unit elastic ($\mathcal{D} = p_r \sum_{i=1}^{P} m_r(i)$) and the cost function is $\mathcal{C}_i = p_p b = p_p(m_r(i) + F)$ with F a fixed cost. The P provider

Bringing all three distortions together in this illustrative model, the equilibrium price of health services is given by

$$p = \underbrace{\frac{1}{\zeta}}_{\text{information rents}} \times \underbrace{\frac{1}{1-1/P}}_{\text{imperfect competition}} \times \underbrace{\frac{1}{(1-\iota)}}_{\text{admin. costs}} \times \underbrace{\frac{1}{z}}_{\text{marginal cost}}. \tag{11}$$

An alternative interpretation of the variation in the price of health services across countries is that it is the result of direct regulation (for example price or cost controls). One example is the price of drugs which is regulated in several European countries. Pharmaceutical innovation has been shown to be quite sensitive to expected market size (revenue) (Acemoglu and Linn, 2004). Hence, European price controls could depress (global) innovation which ultimately negatively impacts patients while price distortions in the U.S. provide an incentive for innovation (Lakdawalla et al., 2008).[47] This has implications for computing the welfare loss associated with such distortions. Evidence from Table A.2 (appendix) does not suggest that R&D in the U.S., as measured by the number of patents, is necessarily more tilted towards medical innovation than in Europe. In fact, there is evidence that the link between price distortions, such as imperfect competition, and innovation is non-linear (Aghion et al., 2005).

Insert Figure 2 here

Using (11), we can test whether estimates of price differences correlate with proxies for the origin of price differences. For example, we can use indices from OECD of competition (degree of patient choice among practitioners), incentives for quality (a measure of how tight surveillance of physicians is), regulation (State policies controlling quantities and prices of healthcare suppliers), and finally a measure of administrative costs.[48] Panel (a) of Figure 2 shows that stronger competition between health providers is correlated with lower health prices in Europe. These results on the negative correlation between the price of health services and competition are supported by the literature on market concentration and prices in the U.S. as reviewed in Gaynor and Town (2012). Panel (b) of Figure 2 shows that an index of price regulation is strongly correlated with lower prices, even

play an oligopoly game à la Cournot leading to the equilibrium price $p_r = \frac{1}{1-1/P} p_p$. Finally, if payers act on a competitive market with a linear production function, then the consumer price for the health basket goods is equal to $p = p_r$.

[47] For a model with endogenous innovation driven by markups in health care, see Koijen et al. (2016).

[48] Data for the U.S. is not available from the OECD, except for administrative costs. See De la Maisonneuve et al. (2016) and Appendix C.3.3 for more details.

within Europe. Panel (c) of Figure 2 displays a weaker negative relationship observed for an index of incentives for quality.

Finally, panel (d) of Figure 2 shows that there is a positive link between estimates of health prices and the administrative costs in hospitals, such as those measured by Himmelstein et al. (2014).[49] Hence, the estimated relative prices of health services correlate well with institutional differences across countries which were discussed in Section 2.

4.5 Health Risk Differences and their Origin

Table 8 reports estimates of parameters in the health production function capturing health risks. In all countries, there is a strong state-dependence in health transition probabilities: for a given level of health expenditures, the probability of being in good health next period is much larger in case of good than in bad health ($\alpha_{11} > 0 > \alpha_{10}$). However, there are large differences across countries in this state-dependence in health transition probabilities. We examine whether these differences in α_{1h} across countries are related to health risk differences across countries.

Insert Figure 3 here

Panel (a) of Figure 3 shows that the α_{11} are closely correlated with an index of the prevalence of risky behaviors (obesity, smoking, drug consumption and absence of physical activity).[50] Panel (b) of Figure 3 shows that the estimated α_{10} are consistent with the information on disease persistence: a low α_{10} in the transition from bad to good health occurs in countries with an observed large number of years of life lost due to premature death and of healthy years lost due to illness (DALYs), both caused by these risky behaviors.[51]

4.6 Other Estimates and Fit

We report in Table 8 estimates of differences in productivity (TFP). According to these estimates, Denmark is the only country with a statistically more productive technology than the U.S.

[49] Administrative costs in hospitals are available for a limited set of countries.

[50] See the Appendix C.3.2 for the construction of this index.

[51] The DALYs give the number of years of life lost due to each risky behavior (obesity, smoking, drug abuse, and absence of physical activity). This number of years of life lost is the sum of the number of years lost due to premature death and the number of healthy years of life lost due to illness. Hence, the DALYs indicate the persistence of an unhealthy state.

($A_{DK} = 1.255 A_{US}$), while Germany and the Netherlands appear as productive. The other European countries have lower estimated productivity. The Spearman correlation between our TFP estimates with the other estimates ranges from 0.74 (Bergeaud et al., 2016) to 0.83 (Groningen Growth and Development Center).

The model fits well the moments used in estimation. In particular, we get a close match of the share of health expenditures in GDP. In terms of health, the model does well, in particular for the transition from good to good health and for the gradient. For example, it is able to replicate a larger gradient in the U.S. and a much lower gradient in European countries. The transition from bad to good health is not as well matched within Europe, although the difference between European countries and the U.S. is well captured (the fraction of healthy people is lower in the U.S. than in Europe on average). Appendix D provides a detailed comparison of the fit of the moments. In addition, we show in the same Appendix that the model matches the income gradient of health expenditures in the cross-section for the U.S. (the only country for which we have data).

4.7 Extensions

We consider two extensions of the baseline specification. Results from the two extensions are found in Appendix E. The first allows for the wage to depend on health. Although we fit well the observed income-health gradient, a model with wages that depend on health may yield different predictions from a model that forces the gradient to emerge only from the demand for health. Given the moments we use, it is not possible to identify the parameter governing the link between health and productivity (wage) from other parameters in the model. We consider a common wage penalty of 20% for those in bad health.[52] We re-estimate all parameters. We find that parameters are quite insensitive to this wage penalty. In particular, estimates of the price of health services are overall similar and the ranking of countries does not change.[53]

The second extension allows co-insurance rates to depend on income since our characterization of the health insurance system is quite crude, in particular for the U.S. In the case of the U.S., we use the Medical Expenditure Panel Study (MEPS) to estimate the gradient in co-insurance rates

[52] Currie and Madrian (1997) reviewed estimates of the effect of health on earnings. Estimates of the effect vary because of the definition of health and econometric methods used but an effect of 20% is roughly in the middle of the range of estimates reported.

[53] The wage penalty may be different across countries. However, we lack comparable data to consistently estimate these differences.

by income. We find that co-insurance rates increase with income. Co-insurance rates increase by 50% from the lowest decile of income to the highest one. This yields co-insurance rates in the model that range from 10.88% to 16.32%. We re-estimate common parameters for the U.S. to check how sensitive our estimates are to these differences. We find that this results in minor differences in parameter estimates. This is consistent with the low elasticity of health expenditures with respect to the price that we reported earlier. Therefore, if we use the estimates of common parameters to estimate country-specific parameters, we still find qualitatively similar price differences between the U.S. and Europe. We have also examined the sensitivity of our price estimates to the co-insurance rates we use. The effective price paid by agents is $p \times \mu$. Hence, a badly measured μ may bias our price estimates. We examine this sensitivity in the context of Italy, which has a very large μ because of excluded services from health insurance. If instead we assume that Italy has the same μ as the U.S., the price estimate we obtain increases from 0.71 to 0.78. We conclude that the price estimates we obtain may be somewhat contaminated by mismeasured μ but that this mismeasurement does not reverse the conclusion that prices are lower in Europe.

5 Explaining Variation in Health Expenditures and Health Across Countries

The price of health services is 33% (1/0.751-1) larger in the U.S. than in Europe (average), while overall productivity (TFP) is 8.76% higher in the U.S. than in Europe.[54] Our estimation results also reveal that heterogeneity in exogenous health risks is important. We focus on these three sources of heterogeneity and investigate their relative importance to explain observed differences.

Insert Table 9 here

To quantify the effect of these differences, we focus on the share of health expenditures in GDP (s), the fraction of individuals in good health ($\pi(h=1)$), and health inequalities (income-health gradient) measured by the relative fraction of individuals in good health within the fourth income quartile ($\overline{\pi}_4$). We simulate counterfactual general equilibrium scenarios where we neutralize each source of heterogeneity. Table 9 reports results. We consider four scenarios: *i)* a baseline scenario where all country-specific heterogeneity is accounted for; *ii)* a scenario where we remove price differences, setting the price of health services equal to the European average; *iii)* a scenario where TFP heterogeneity is removed, setting the TFP equal to the European average; and *iv)* a scenario where exogenous health risks heterogeneity (α_{11}, α_{10}) is removed, setting exogenous health risks equal to the European average.[55]

The total simulated differences (Δ) between the U.S. and the E.U countries are 5.428 percentage points (pp) for the GDP share of health expenditures s and -7.695 percentage points for the fraction in good health $\pi(h=1)$. The U.S. spends more but has lower health on average. As for the health-income gradient, it is larger in the U.S. than in Europe: The ratio of the fraction in good health for those within the top 75% of the income distribution to those within the bottom 25% of income ($\overline{\pi}_4$) equals 1.310 in the U.S., versus 1.066 in Europe, which represents a 24.396 percentage point gap.

[54] We obtain these average differences by using individual estimates for each country. The average relative price is 0.751 in Europe which translates into $1/0.751 - 1 = 0.332$. For TFP, the average relative TFP is 0.919 in Europe which translates into a $1/0.919 - 1 = 0.0876$ higher TFP in the U.S.

[55] Given that the characteristics of the production function of U.S. goods $\{\alpha, \delta\}$, as well as the co-insurance rate (μ), are close to the average of their European counterparts, they do not play any significant role in explaining differences. In addition, the experiment in which we remove heterogeneity in the income process $\{\rho_e, \sigma_e\}$ yields results that are similar to the removal of heterogeneity in TFP. We do not report results from these experiments but they are available upon request.

When heterogeneity in the price of health services is removed ("Price" scenario), the Europe-US gap in expenditures is reduced by 68.6%, going from 5.428pp to 1.706pp. The gap in the fraction of individuals in good health is reduced by 58%, going from -7.695pp to -3.231pp. Hence, the price of health services has a sizable effect on health expenditures and health status differences between the U.S. and Europe. This wedge has also a sizeable impact on health inequalities. In the U.S. baseline, individuals in the top quartile are 31% more likely to be in good health than their counterparts in the first quartile. The gradient is reduced to 19.3% when the U.S. price of health services is set at the European level. By removing price differences, the gap in the gradient between the U.S. and Europe further reduces by 50%.

When TFP heterogeneity is removed ("TFP" scenario), the GDP share of health expenditures increases marginally in the U.S. by 0.21pp, but it also increases marginally in European countries by 0.04pp.[56] This result is in line with our estimate of the income elasticity of health expenditures which is below one. The GDP share of health spending is declining in TFP. The U.S. in this study, as in others (e.g. Ohanian et al. (2008)), is found to have higher TFP (except for Denmark). Therefore, technological efficiency cannot explain why the U.S. has a higher GDP share of health expenditures in our model.

Worse health status in the U.S., for example, due to higher prevalence of risky behaviors (obesity, smoking, drugs, and physical inactivity), could also explain differences in expenditures and health status (Thorpe et al., 2004, 2007). For example, the rapid growth of obesity in the U.S. relative to other countries could play a role (Cutler et al., 2003). As shown in section 4.5, these risky behaviors are captured in our model by exogenous health risks (α_{10}, α_{11}). When the heterogeneity in exogenous health risks is removed ("Health risks" scenario), differences across countries in health expenditures, health status, and health inequalities decrease sharply. The gaps in expenditures virtually disappear, going from 5.428 to 0.498 percentage points. At the same time, the gap in the fraction of individuals in good health is reduced by 71.1%, from -7.695 to -2.224 percentage points, and the gap in income-health gradient falls by 55.6%, from 24.396 to 10.839. The fact that adding up the effects of price and unobserved health differences sums to more than 100% of the gap implies that there is a negative complementarity between these forces; that is, price effects would be lower

[56] This result is driven by the large decline in TFP for Denmark when we use European average TFP. If we omit Denmark, the gap between the U.S. and a group of countries characterized by a low-TFP (all the E.U. countries except Denmark) unambiguously increases.

if Americans had better health behaviors and vice versa.

This decomposition of differences between the U.S. and the European countries with respect to the GDP share of health expenditures, fraction of individuals in good health, and income-health gradient suggest that both the price of health services and exogenous health risks largely explain the gap between the U.S. and Europe while TFP differences do not play any significant role.

6 Welfare Consequences of the Higher Price of Health Services

We find sizeable differences in the price of health services between the U.S. and European countries. In this section, we investigate the welfare implications of these differences. We perform a counterfactual exercise in which we assume that Americans have access to health services at the same prices as Europeans. We then ask the question: What is the compensation required by Americans for higher prices? This is a Hicksian compensation exercise that we adapt to our dynamic general equilibrium model.

Precisely, we compute the compensation required for each agent living in an economy where the health price is the same as in Europe and then increases to the level observed in the U.S. Using the model, we compute the welfare of each (a, e, h)-type agent in the U.S. economy, $V(a, h, e|p_{US}, \Omega^X_{US})$, and $V(a, h, e|p_{EU}, \Omega^X_{US})$ with lower prices estimated in Europe. These welfare values depend on $\Omega = \{\Omega^X_{US}\}_{X=US,EU}$, a set of two vectors regrouping *(i)* all US-specific characteristics (income risk, risky health behaviors, co-insurance rate) and *(ii)* equilibrium factor prices $(r(p_X),\ w(p_X))$ and tax rate $(\tau(p_X))$. The index X is used to emphasize partial and general equilibrium effects. With $X = US$, the compensation is evaluated in partial equilibrium (PE), because general equilibrium prices (wages, interest rate, and tax rate) are held constant at their equilibrium values in the U.S. economy (with U.S. prices). When $X = EU$, the calculation takes into account general equilibrium (GE) adjustments of the interest rate, wage rate, and contribution rate since we use the counterfactual equilibrium prices (wage, interest rate, and tax rate) that would be obtained in a U.S. economy with European health prices. The compensation $\mathcal{P}^X(a, h, e)$ that makes an agent indifferent between the two price levels, p_{US} and p_{EU}, is given by

$$V\left(a + \mathcal{P}^X(a, h, e), h, e|p_{US}, \Omega^{US}_{US}\right) = V\left(a, h, e|p_{EU}, \Omega^X_{US}\right) \quad X = US, EU. \tag{12}$$

The compensation $\mathcal{P}^X(a, h, e)$ can be spent on consumption goods and health services and therefore we allow for re-allocation of resources. In Table 10, we report estimates of the compensation as a fraction of consumption in the steady-state for the baseline scenario. We do this for those in both poor ($h = 0$) and good ($h = 1$) health and for three levels of income (the lowest level e_0, e_4 middle and e_9 the top-level). Finally, we compute compensation in partial equilibrium ($X = US$) and general equilibrium ($X = EU$).

Insert Table 10 here

Overall, the compensation required represents 2.542% of baseline consumption when accounting for general equilibrium effects. General equilibrium effects are important. Without those, the compensation required would represent less than two thirds of a percentage point (0.603%). A significant portion of the general equilibrium effect is due to the increase in the contribution rate for health insurance. In terms of distribution, it is perhaps not surprising that those in poor health require a larger compensation as they consume more health services. The compensation is lower for those at the bottom of the income distribution since they consume fewer health services and are affected negatively by increased contributions (taxes).

Instead of looking at compensations directly as a fraction of consumption, it is interesting to reformulate these effects in terms of price or cost-of-living indices. Indeed, we can construct a cost-of-living index from these compensation measures. For an agent with status (a, h, e), moving from European to U.S. prices induces a change in lifetime expenditures which can be represented by an index (100 represents no change) equal to

$$I_{LT}^{X}(a,h,e) = 100 \times \frac{a + \mathcal{G}_X(e) + \mathcal{P}^X(a,h,e)}{a + \mathcal{G}_X(e)}, \tag{13}$$

where $\mathcal{G}_X(e)$ is the lifetime present value of income, at equilibrium wage and rent with prices in region $X = US, EU$. The numerator measures the resources needed to reach a certain level of welfare in an economy where $p = p_{US}$ and the denominator measures the resources needed to reach the same welfare but in an economy where $p = p_{EU}$.[57] When $I_{LT}^{X} > 100$, the lifetime cost of living is higher in the economy where $p = p_{US}$ than in another where p_{EU}. Similar to the Hicksian measures of the cost of living first proposed by Konüs (1924), we can aggregate the lifetime cost of living index by using the agents' distribution obtained in general equilibrium for the baseline economy with p_{US}. Therefore, the average cost-of-living index for differences in the price of health services

[57]See Appendix F.1 for more details on the computation method.

is given by:

$$I_{LT}^{X} = \sum_{a}\sum_{e}\sum_{h} \lambda(a, e, h | p_{US}, \Omega_{US}^{US}) I_{LT}^{X}(a, e, h), \tag{14}$$

where $\lambda(a, e, h | p_{US}, \Omega_{US}^{US})$ is the stationary distribution of agents in the baseline scenario. Table 11 reports lifetime cost of living indices in the U.S. induced by the higher price of health services.[58] These indices are calculated both under partial (PE) and general equilibrium (GE), and are reported for agents in bad and good health as well as for three different levels of income (e_0 lowest, e_4 middle and e_9 highest levels). The overall lifetime cost of living index is also reported.

Insert Table 11 here

Our estimates of the average lifetime cost of living index are respectively 102.234 with general equilibrium (GE) adjustments and 100.577 in partial equilibrium (PE). The cost of living index using GE effects is larger than a traditional Laspeyres index using steady-states values of consumption and health expenditures. In the case of our experiment, one would define a Layspeyres Index as $I_L = \frac{c_{US} + \mu p_{US} m_{US}}{c_{US} + \mu p_{EU} m_{US}}$, where c_{US} and m_{US} are the stationary values of consumption and health expenditures. Using steady-state values, it yields a value of 100.87 which is a bit less than half of the cost-of-living increase measured in GE. Berndt et al. (2001) discuss these shortcomings in the context of constructing a price index for medical services. They point out that a theoretically grounded cost of living index would account for the production of health (health market services, health insurance, and ability of individuals to use care for being in good health) and consumption of health services (preferences and budget).[59]

[58] Appendix F.2 shows the cost-of-living index as a function of various dimensions in the space (a, h, e).

[59] It is well known that a Laysperes index suffers from several limitations: *i)* it is valid only in a static environment, *ii)* it does not allow for substitution and hence does not keep utility constant, *iii)* it assumes a representative agent and *iv)* it assumes an economy without uncertainty. Moreover, one also needs to account for general equilibrium adjustments: a change in health price induces changes of other equilibrium prices (wages, interest rate) and the contribution rate (tax).

7 Conclusion

Through the lens of a general equilibrium heterogeneous agent model, the present study estimates the differences in prices of health services across countries, and then measures their contributions, as well the ones of other factors, in explaining cross-country differences in health expenditures and health status. The model is estimated using the method of simulated moments on macro and microdata from the U.S. and seven European countries.

The model fits the data reasonably well and matches evidence on the income elasticity of health expenditures (Acemoglu et al., 2013) and the price elasticity of health expenditures from the RAND Health Insurance Experiment (Manning et al., 1987). Estimated prices are negatively correlated with a measure of competition, the intensity of price regulations, and incentives for quality of care. These factors, along with administrative costs lead to important distortions and higher prices in the U.S. (Cutler and Ly, 2011). Estimates of unobserved health risks are also strongly correlated with an index of risky behavior grouping obesity, smoking, drug abuse, and physical inactivity prevalence which provides some confidence that these estimates do not pick up other unspecified differences across countries.

The estimates show that the U.S. is the country where the price of health services is the highest, roughly 33% larger than in Europe. This is consistent with evidence from the Dartmouth Atlas of health care which suggests that more than 20% of health expenditures in Medicare could be eliminated with little impact on health outcomes (Skinner et al., 2005). According to our estimates, the gap in health expenditures between the U.S. and the set of European countries considered could be reduced by 68% and the gap in the fraction in good health by 58% if the price of health services in the U.S. was the same as in Europe. By removing the price differences of health services, the gap in income-health gradient would be reduced by 50%. We find that differences in TFP cannot account for cross-country differences in health expenditures while we find that unobserved health risks have a large impact on health and health expenditures, which is consistent with the evidence presented by Thorpe et al. (2004, 2007).

Our welfare calculations suggest that the impact of a higher price for health services on the average lifetime cost of living is approximately two percentage points. This computation of the welfare loss abstracts from various dynamic efficiency considerations when innovation is taken into account. For example, we have found that the price of health services is lower when there is more competition in the health care sector. In a dynamic setting, markups for innovators could be impacted negatively by lower prices which could slow innovation (medical progress). Hence, lower prices from more competition could lead to a welfare loss. This conjecture assumes that innovation increases when competition is reduced. However, there is some evidence that there exists a U-inverted relationship between market concentration and innovation suggesting that more competition may encourage R&D (Aghion et al., 2005). Therefore, even in this dynamic perspective, stronger competition could also increase welfare. Yet, the innovation process could also be impacted negatively by aggregate uncertainty induced by changes in government policies (Koijen et al., 2016) and deliberate price controls (Lakdawalla et al., 2008) which we have not taken into account. Hence, one interpretation of our result is that our estimate of the welfare effect is an upper bound on the welfare effect of price distortions.

One interesting channel to explore is the extent to which technological interdependence leads to an asymmetric equilibrium where two types of countries coexist: the first with high-powered incentives for medical innovation (the U.S.) and the second, the followers (Europe), benefiting from progress due to the rapid dissemination of knowledge, without paying the costs associated to innovation with the implementation of price controls (Acemoglu et al., 2017). Consistent with our results, low prices in Europe are related to the better screening of innovations with high costs relative to their effectiveness. This may lead to higher prices in the U.S. relative to Europe without any significant advantage in terms of health. Within the U.S., the widespread adoption of costly and not necessarily cost-effective technologies has been noted as one explanation for the rapid growth in health expenditures (Chandra and Skinner, 2012, Garber and Skinner, 2008). This is certainly consistent with the negative relationship we find between prices and health across countries.

Data Availability

Code and Data replicating the tables and figures in this article can be found in Fonseca, Langot, Michaud, and Sopraseuth (Fonseca et al.) in the Harvard Dataverse, .

References

Acemoglu, D., A. Finkelstein, and M. Notowidigdo (2013). Income and Health Spending: Evidence from Oil Price Shocks. *Review of Economics and Statistics 95*(4), 1079–1095.

Acemoglu, D. and J. Linn (2004). Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry. *The Quarterly Journal of Economics 119*(3), 1049–1090.

Acemoglu, D., J. A. Robinson, and T. Verdier (2017). Asymmetric Growth and Institutions in an Interdependent World. *Journal of Political Economy 125*(5), 1245–1305.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005). Competition and Innovation: an Inverted-U Relationship. *The Quarterly Journal of Economics 120*(2), 701–728.

Aiyagari, S. (1994). Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics 109*(3), 659–84.

Anderson, G., U. E. Reinhardt, P. Hussey, and V. Petrosyan (2003). It's The Prices, Stupid: Why The United States Is So Different From Other Countries. *Health Affairs 22*(3), 89–105.

Arrow, K. (1963). Uncertainty and the Welfare Economics of Medical Care. *The American Economic Review 53*(5), 941–973.

Avendano, M., M. Glymour, J. Banks, and J. Mackenbach (2009). Health Disadvantage in US Adults Aged 50 to 74 Years: A Comparison of the Health of Rich and Poor Americans With That of Europeans. *American Journal of Public Health 99*(3), 540–548.

Banks, J., M. Marmot, Z. Oldfield, and J. Smith (2006). Disease and Disadvantage in the United States and in England. *Journal of the American Medical Association 295*(17), 2037–2045.

Banks, J., M. Marmot, Z. Oldfield, and J. Smith (2009). The SES Health Gradient on Both Sides of the Atlantic. In D. Wise (Ed.), *Developments in the Economics of Aging*, pp. 259–406. University of Chicago Press.

Bergeaud, A., G. Cette, and R. Lecat (2016). Productivity Trends in Advanced Countries between 1890 and 2012. *Review of Income and Wealth 62*, 420–444.

Berndt, E., D. Cutler, R. Frank, Z. Griliches, J. Newhouse, and J. Triplett (2001). Price Indexes for Medical Care Goods and Services – An Overview of Measurement Issues. In D. Cutler and E. Berndt (Eds.), *Medical Care Output and Productivity*, Chapter 4, pp. 141–200. University of Chicago Press.

Cameron, A. and P. Trivedi (2005). *Microeconometrics*. Cambridge University Press.

Canada Patented Medicine Prices Review Board (2016). Annual report. Technical report, Health Division.

Chandra, A. and J. Skinner (2012). Technology Growth and Expenditure Growth in Health Care. *Journal of Economic Literature 50*(3), 645–80.

Chari, V., P. Kehoe, and E. McGrattan (2007). Business Cycle Accounting. *Econometrica 75*(3), 781–836.

Currie, J. and B. Madrian (1997). Health, Health Insurance and the Labor Market. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3, Chapter 50, pp. 3309–3416. Elsevier.

Cutler, D., E. Glaeser, and J. Shapiro (2003). Why Have Americans Become More Obese? *The Journal of Economic Perspectives 17*(3), 93–118.

Cutler, D. and D. Ly (2011). The (Paper) Work of Medicine: Understanding International Medical Costs. *Journal of Economic Perspectives 30*(6), 1174–1187.

De la Maisonneuve, C., R. Moreno-Serra, F. Murtin, and J. Martins (2016). The Drivers of Public Health Spending: Integrating Policies and Institutions. Technical Report 1283, OECD Economics Department Working Papers.

De Nardi, M., G. Fella, M. Knoef, G. Paz-Pardo, and R. Van Ooijen (2021). Family and Government Insurance: Wage, Earnings, and Income Risks in the Netherlands and the U.S. *Journal of Public Economics 193*, 104327.

De Nardi, M., E. French, and J. Jones (2010). Why Do the Elderly Save? The Role of Medical Expenses. *Journal of Political Economy 118*(1), 39–75.

De Nardi, M., S. Pashchenko, and P. Porapakkarm (2018). The Lifetime Costs of Bad Health. NBER Working Papers 23963, National Bureau of Economic Research.

Feenstra, R., R. Inklaar, and M. Timmer (2015). The Next Generation of the Penn World Table. *America Economic Review 105*(19), 3150–3182.

Fonseca, R., F. Langot, P.-C. Michaud, and T. Sopraseuth. Replication Data and Code for: "Understanding Cross-country Differences in Health Status and Expenditures: Health Price Matters". *Journal of Political Economy*. https://doi.org/10.7910/DVN/PV7LVY.

Fonseca, R., P.-C. Michaud, T. Galama, and A. Kapteyn (2021). Accounting for the Rise of Health Spending and Longevity. *Journal of European Economic Association 19*(1), 536–579.

French, E. (2005). The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour. *The Review of Economic Studies 72*(2), 395–427.

French, E., J. McCauley, M. Aragon, P. Bakx, M. Chalkley, S. Chen, B. Christensen, H. Chuang, A. Cote-Sergent, M. De Nardi, E. Fan, D. Echevin, P.-Y. Geoffard, C. Gastaldi-Menager, M. Gortz, Y. Ibuka, J. Jones, M. Kallestrup-Lamb, M. Karlsson, T. Klein, G. de Lagasnerie, P.-C. Michaud, O. O'Donnell, N. Rice, J. Skinner, E. van Doorslaer, N. Ziebarth, and E. Kelly (2017). End-Of-Life Medical Spending In Last Twelve Months Of Life Is Lower Than Previously Reported. *Health Affairs 36*, 1211–1217.

Garber, A. M. and J. Skinner (2008). Is American Health Care Uniquely Inefficient? *Journal of Economic Perspectives 22*(4), 27–50.

Gaynor, M. and R. Town (2012). Competition in Health Care Markets. In A. Culyer and J. Newhouse (Eds.), *Handbook of Health Economics*, Volume 2, pp. 499–637. Elsevier.

Gerdtham, U.-G. and B. Jonsson (2000). International Comparisons of Health Expenditure: Theory, Data and Econometric Analysis. In A. Culyer and J. Newhouse (Eds.), *Handbook of Health Economics*, Volume 1, Part A, pp. 11–53. Elsevier.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy 80*(2), 223–255.

Hall, R. and C. Jones (2007). The Value of Life and the Rise in Health Spending. *The Quarterly Journal of Economics 122*(1), 39–72.

Himmelstein, D., M. Jun, R. Busse, K. Chevreul, A. Geissler, P. Jeurissen, S. Thomson, M. Vinet, and S. Woolhandler (2014). A Comparison Of Hospital Administrative Costs In Eight Nations: US Costs Exceed All Others By Far. *Health Affairs 33*(9), 1586–1594.

Horenstein, A. and M. Santos (2018). Understanding Growth Patterns in US Health Care Expenditures. *Journal of the European Economic Association 17*(1), 284–326.

Hugonnier, J., F. Pelgrin, and P. St-Amour (2013). Health and (Other) Asset Holdings. *Review Economic Studies 80*(2), 663–710.

IFHP (2013). Comparative Price Report: Variation in Medical and Hospital Prices by Country. Technical report, International Federation of Health Plans.

Institute of Medicine and National Research Council (2013). *U.S. Health in International Perspective: Shorter Lives, Poorer Health*. Washington, DC: The National Academies Press.

Kapteyn, A., J. Smith, and A. van Soest (2007). Vignettes and Self-Reports of Work Disability in the United States and the Netherlands. *American Economic Review 97*(1), 461–473.

Koijen, R. S. J., T. J. Philipson, and H. Uhlig (2016). Financial Health Economics. *Econometrica 84*(1), 195–242.

Konüs, A. A. (1924). The Problem of the True Index of the Cost of Living. *Bulletin of the Institute of Economic Conjuncture* (9-10), 64–71. Reprinted in Econometrica, January, 1939, 10-29.

Lakdawalla, D. N., D. P. Goldman, P.-C. Michaud, N. Sood, R. Lempert, Z. Cong, H. de Vries, and I. Gutierrez (2008). U.S. Pharmaceutical Policy In A Global Marketplace. *Health Affairs 27*(Supplement 1), w138–150.

Manning, W., J. Newhouse, N. Duan, E. Keeler, and A. Leibowitz (1987). Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment. *The American Economic Review 77*(3), 251–277.

Newhouse, J. (1992). Medical Care Costs: How Much Welfare Loss? *The Journal of Economic Perspectives 6*(3), 3–21.

Ohanian, L., A. Raffo, and R. Rogerson (2008). Long-term Changes in Labor Supply and Taxes: Evidence from OECD Countries, 1956-2004. *Journal of Monetary Economics 55*(8), 1353–1362.

Papanicolas, I., A. Marino, L. Lorenzoni, and A. Jha (2020). Comparison of Health Care Spending by Age in 8 High-Income Countries. *JAMA Network Open 3*(8), 2574–3805.

Powell, M. J. D. (2006). The NEWUOA Software for Unconstrained Optimization Without Derivatives. In *Large-Scale Nonlinear Optimization*, pp. 255–297. Springer, Boston, MA.

Romley, J. A. and N. Sood (2013). Identifying the Health Production Function: The Case of Hospitals. NBER Working Papers 19490, National Bureau of Economic Research, Inc.

Skinner, J. S., E. S. Fisher, and J. Wennberg (2005). The Efficiency of Medicare. In D. Wise (Ed.), *Analyses in the Economics of Aging*, pp. 129–160. University of Chicago Press.

Smith, J. (1999). Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status. *Journal of Economic Perspectives 13*(2), 145–166.

Thorpe, K., C. Florence, D. Howard, and P. Joski (2004). Trends: The Impact Of Obesity On Rising Medical Spending. *Health Affairs 22*(2), W4–480–486.

Thorpe, K., D. Howard, and K. Galactionova (2007). Differences In Disease Prevalence As A Source Of The U.S.-European Health Care Spending Gap. *Health Affairs 26*(6), w678–686.

LT growth rates	DE	DK	FR	NL	SE	SP	US	Europe Av.
GDP growth	0.011	0.016	0.017	0.017	0.016	0.020	0.018	0.015
Pop. Aging	0.007	0.004	0.005	0.006	0.004	0.010	0.004	0.007
Unexplained	0.010	0.001	0.014	0.005	0.002	0.016	0.019	0.011

Table 1: **Decomposition of Growth in Health Expenditures 1970-2007**: We use real health expenditures per capita $pm_{g,t}$, real GDP per capita $y_{g,t}$ and the 65+ population share $p65^+_{g,t}$, where each country is denoted g, and year t. We first estimate the regression $\log(pm_{g,t}) = a\log(y_{g,t}) + b\Delta\log(p65^+_{g,t}) + c_g + \eta_t + e_{g,t}$. We include a spline function of time with nodes at each decade and country fixed effects. We obtain estimates $\widehat{a} = 0.901$ ($se = 0.115$) and $\widehat{b} = 0.704$ ($se = 0.110$). The first row reports the effect of GDP growth on long-term (LT) annual growth rates in health expenditures. The second row reports the effect of population aging. Finally, the last row reports the portion that is left unexplained (including the common trend). For the last column, we use population shares in Europe to compute a weighted average.

	(%)	DE	DK	FR	IT	NL	SE	SP	US	Europe Av.
Inflation rates	health $\bar{i}_m$	2.58	1.29	1.79	2.41	3.14	2.51	1.48	3.62	2.19
	overall $\bar{i}$	1.34	1.78	1.54	2.06	2.12	1.43	2.66	2.33	1.81
	$\bar{i}_m - \bar{i}$	1.24	-0.49	0.25	0.35	1.01	1.08	-1.18	1.20	0.38
GDP share	$s = p\frac{m}{y}$	10.63	9.34	10.62	8.26	9.25	9.03	8.12	15.04	9.55
	$\widehat{s} = p_{1996}\frac{m}{y}$	9.37	10.03	10.43	7.82	8.13	8.25	9.24	12.42	9.16
	$\frac{s-\widehat{s}}{\widehat{s}}$	13.42	-6.85	1.86	5.57	13.69	9.48	-12.10	21.09	4.26

Table 2: **Prices and the Share of GDP Devoted to Health Expenditures.** We use consumer price indices (CPI) from EuroStat and the Bureau of Labor Statistics from 1996 to 2007. The first panel shows average inflation rates computed for health products and services and all goods along with the difference between average rates for each country over the period. The second panel provides the observed share of health expenditures in GDP (s), the counterfactual share with 1996 prices ($\widehat{s}$) and the relative difference between the two measures. The last column reports a population weighted average of these numbers for the selected group of European countries.

	US	DE	DK	FR	IT	NL	SE	SP
Prices								
Angiogram	914			264				125
relative to US	1			0.288				0.136
Bypass surgery	73420			22344		14061		17437
relative to US	1			0.304		0.191		0.275
Drug price index	1	0.34		0.268	0.285	0.272	0.306	0.275
relative to US								
Hospital spending per discharge	18788	5251	11468	5348		13909	9953	
relative to US	1	0.28	0.61	0.28		0.74	0.53	
Efficiency								
Case-fatality rate after AMI (%)	6.66	5.5	9.85	7.05	7.25	9.5	6.3	10.3
relative to US	1	0.82	1.47	1.05	1.08	1.42	0.94	1.54

Table 3: **Price and Efficiency Differences Across Countries (2013 dollars)**: Price information for angiogram and bypass surgery from International Federation of Health Plans (IFHP (2013)) while the drug price index is taken from Canada Patented Medicine Prices Review Board (2016). Hospital spending per discharge for 2009 from OECD Health Data. Case-fatality rate (% Admission-based) in adults aged 45 and over within 30 days after admission for Acute Myocardial Infarction (AMI), OECD Health Data (2001-2014).

	hypertension	diabetes	lung	heart	stroke	total cond.	ADLs	Life exp (50)
DE	0.338	0.105	0.042	0.090	0.034	0.610	0.067	31.14
DK	0.277	0.066	0.059	0.063	0.038	0.504	0.069	30.15
FR	0.256	0.086	0.048	0.103	0.023	0.516	0.072	32.56
IT	0.350	0.107	0.060	0.082	0.021	0.619	0.075	32.66
NL	0.239	0.074	0.062	0.091	0.035	0.501	0.052	31.22
SE	0.261	0.079	0.025	0.109	0.028	0.502	0.054	32.11
SP	0.289	0.125	0.045	0.077	0.018	0.554	0.069	32.26
Europe (EU)	0.310	0.102	0.049	0.090	0.027	0.573	0.069	31.96
US	0.445	0.147	0.069	0.159	0.045	0.865	0.108	30.65
Δ (US-EU)	0.136	0.045	0.020	0.069	0.018	0.293	0.039	-1.31
Δ(%)	0.439	0.446	0.401	0.765	0.689	0.511	0.570	-0.04

Table 4: **Health Status across Countries**: We report the prevalence of various health conditions (hypertension, diabetes, lung disease, heart disease and stroke), the average number of conditions and the prevalence of limitations in activities of daily living (ADLs) in those age 50 to 75. The prevalence is computed from the 2004 Survey of Health, Ageing and Retirement in Europe (SHARE) and the 2004 Health and Retirement Study (HRS, U.S.). Statistics are weighted using sampling weights in the respective surveys. The last column reports the remaining life expectancy at age 50 based on the Human Mortality Database. The bottom two lines of the Table report the difference (and relative difference) between the U.S. and the (population-weighted) set of European countries.

	DE	DK	FR	IT	NL	SE	SP	US
ρ_e	0.921	0.913	0.967	0.937	0.943	0.913	0.972	0.959
σ_e^2	0.039	0.016	0.014	0.029	0.019	0.016	0.015	0.039
σ_u^2	0.077	0.071	0.133	0.157	0.119	0.071	0.263	0.097
$\frac{\sigma_e^2}{1-\rho_e^2}$	0.256	0.095	0.215	0.237	0.170	0.095	0.270	0.486

Table 5: **Estimates of Income Shocks Process**: Parameter estimates by minimum distance as outlined in Appendix C. ρ_e refers to the auto-correlation of the persistent shocks, σ_e^2 the variance of innovation of the persistent shocks and σ_u^2 the variance of transitory shocks.

	DE	DK	FR	IT	NL	SE	SP	US
μ	0.128	0.149	0.086	0.238	0.098	0.163	0.229	0.136
α	0.373	0.360	0.379	0.470	0.393	0.461	0.374	0.384
δ_k	0.039	0.043	0.040	0.040	0.041	0.047	0.038	0.049

Table 6: **Calibration of Auxiliary Parameters**: μ refers to the co-insurance rate of health insurance (out-of-pocket payments as percentage of total health expenditures), α refers to the expenditure share of capital while δ_k refers to the depreciation rate on capital.

σ	ϕ	α_0
2.097	0.304	0.169
(0.010)	(0.136)	(0.019)

Table 7: **Common Parameters**: Estimated by MSM. Standard errors in parenthesis.

	US	DE	DK	FR	IT	NL	SE	SP
α_{10}	-0.967	-1.296	-1.601	-1.094	-0.716	-1.265	-1.525	-0.006
	(0.161)	(0.098)	(0.01)	(0.013)	(0.032)	(0.03)	(0.011)	(0.008)
α_{11}	3.487	3.954	4.273	3.698	3.868	3.998	4.311	3.394
	(0.051)	(0.031)	(0.015)	(0.023)	(0.029)	(0.141)	(0.019)	(0.02)
$\frac{p}{p_{US}}$	1	0.847	0.888	0.615	0.713	0.698	0.844	0.65
	-	(0.012)	(0.015)	(0.005)	(0.180)	(0.021)	(0.026)	(0.03)
$\frac{A}{A_{US}}$	1	1.009	1.255	0.924	0.639	1.01	0.788	0.811
	-	(0.099)	(0.008)	(0.026)	(0.007)	(0.078)	(0.003)	(0.034)

Table 8: **Country-Specific Parameters**: Estimated by MSM. Standard errors in parenthesis.

	GDP share of health expenditures s			Fraction good health $\pi(h=1)$			Income-Health gradient $\overline{\pi}_4$		
	U.S.	Europe	Δ	U.S.	Europe	Δ	U.S.	Europe	Δ
Baseline	14.333	8.910	5.428	87.772	95.417	-7.695	1.310	1.066	24.396
Price	10.648	8.941	1.706	92.097	95.328	-3.231	1.193	1.072	12.018
TFP	14.541	8.890	5.651	87.774	95.864	-8.090	1.338	1.051	28.704
Health risks	9.999	9.499	0.498	92.100	94.325	-2.224	1.198	1.089	10.839

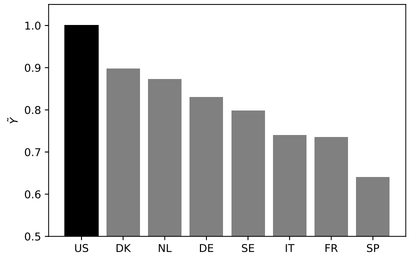
Table 9: **Decomposition of the Differences between U.S. and Europe**: s is the GDP share of health expenditures (%), $\pi(h=1)$ is the fraction of individuals in good health (%) and $\overline{\pi}_4$ is the relative probability to be in good health in the fourth income quartile (Income-health gradient). For each scenario, Δ measures the percentage difference between the U.S. and European average.

		e_0	e_4	e_9	Aggregate
GE	Poor health	2.242	3.121	2.546	2.542
	Good health	2.168	2.601	2.347	
PE	Poor health	0.311	0.896	0.824	0.603
	Good health	0.452	0.576	0.677	

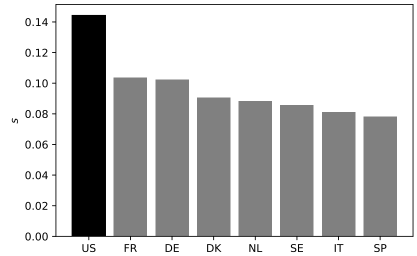
Table 10: **Compensation for Higher Prices in the U.S.** We compute compensation required to maintain welfare fixed at higher prices. We express compensations as a percentage of steady-state consumption in the baseline. We report compensations in partial equilibrium (PE) and accounting for general equilibrium effects (GE) for individuals in poor and good health as well as for three levels of income (lowest e_0, middle e_4, and e_9 highest level).

		e_0	e_4	e_9	Aggregate
GE	Bad health	100.479	101.975	103.531	102.234
	Good health	100.999	102.117	103.313	
PE	Bad health	100.075	100.632	101.241	100.577
	Good health	100.236	100.522	101.032	

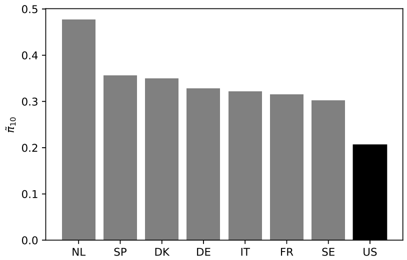
Table 11: **Lifetime Cost of living in the U.S. Induced by Price Differences**: We compute the lifetime cost of living index (multiplied by 100) for a change in price of health services (p) from the European to U.S. level. We report indices in partial equilibrium (PE) and accounting for general equilibrium effects (GE) for individuals in bad and good health as well as for three levels of income (lowest e_0, middle e_4, and e_9 highest).



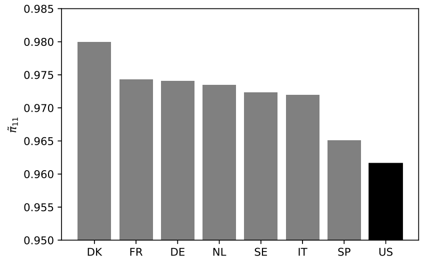
(a) Relative GDP per capita



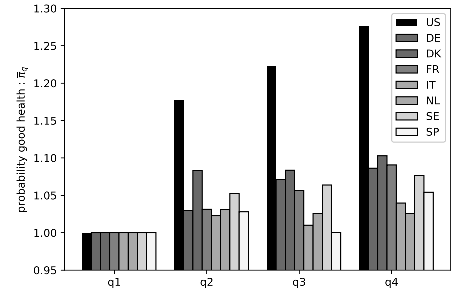
(b) GDP share of health expenditures



(c) Transition from Bad to Good Health

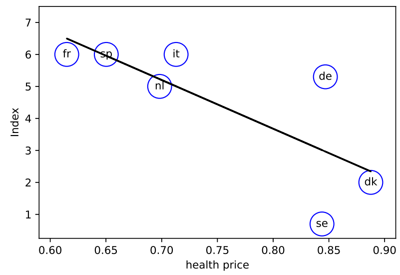


(d) Transition from Good to Good Health

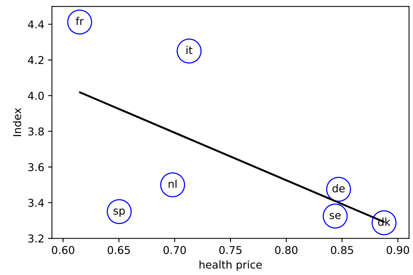


(e) Health-Income Gradient

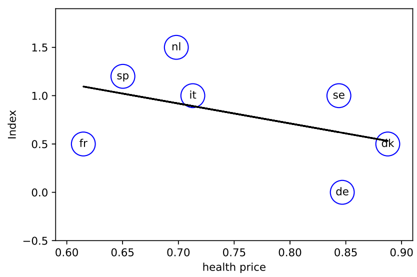
Figure 1: **Moments used in Estimation**: See text for description of how each was constructed.



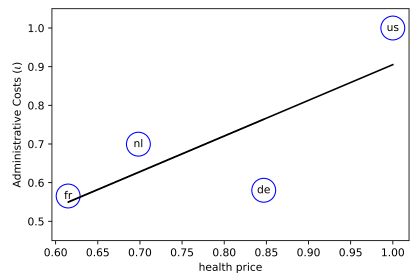
(a) Competition



(b) Regulation

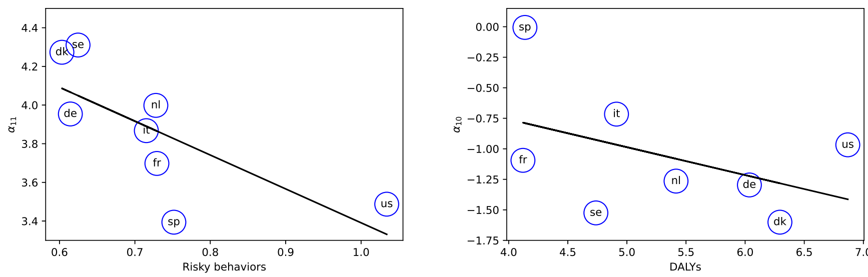


(c) Incentive for quality



(d) Administrative costs

Figure 2: **Correlation between Price Estimates and Proxies for Distortions.** Indices based on De la Maisonneuve et al. (2016). See Appendix C.3.3 for more details on data.



(a) From good to good health

(b) From bad to good health

Figure 3: **Health Behaviors and Country-Specific Effects on Transitions to Good Health.** (a) Health Behaviors: sum of the prevalence of individuals with at least one of the risky behavior (obesity, smoking, absence of physical activity and drug consumption). (b) Disability Adjusted Life Years (DALYs) measure the sum of years of potential life lost due to premature mortality and the years of productive life lost due to disability induced by selected risky behaviors. See Appendix C.3.2 for more details on the data used and the construction of the indices.