Tithe an Oireachtais
Houses of the Oireachtas

Working Paper Series

No. 1 of 2026

Benchmarking Public Health Spending Efficiency in Europe: A DEA Approach

Jacopo Bedogni

February 2026

An Oifig Buiséid Pharlaiminteach
Parliamentary Budget Office

Disclaimer

Benchmarking Public Health Spending Efficiency in Europe: A DEA Approach

Jacopo Bedogni[1]

24 February 2026

Abstract

Public spending efficiency is a growing policy concern as fiscal space tightens and service demands rise. This paper demonstrates how Data Envelopment Analysis (DEA) can benchmark public health spending across 25 European OECD countries using public health expenditure per capita in Purchasing Power Parity (PPP)–adjusted US dollars as input and a multi-output health bundle (life expectancy, infant mortality, avoidable mortality). Robustness is tested using jackknife resampling and order-m partial frontiers, and correlates of inefficiency are examined via a Simar–Wilson double-bootstrap truncated regression. Results consistently characterise Ireland as outcome-strong but input-intensive: the input-oriented model implies a ~45% frontier distance (a benchmarking metric, not short-run savings). Output-oriented conclusions depend on the outcome bundle, headroom is limited in the baseline, life-expectancy-inclusive specification, but is materially larger when focusing on mortality outcomes (infant and avoidable mortality).

Keywords: Data Envelopment Analysis (DEA); Health spending; Public spending efficiency; Benchmarking

JEL Classification: H51, D24, C61

---

[1] Jacopo Bedogni, PhD, is Head of the Economic Modelling and Policy Costing Unit, Parliamentary Budget Office (PBO), Houses of the Oireachtas. The author thanks colleagues at the PBO for helpful comments on earlier drafts. Correspondence: Jacopo.Bedogni@Oireachtas.ie.

1. Introduction

Ensuring value for money in public expenditure has moved to the centre of policy discussion. Fiscal capacity is under growing pressure from demographic change, climate-related investment needs, macroeconomic volatility, and geopolitical uncertainty. At the same time, citizens continue to expect improvements in service quality and outcomes. A core policy challenge therefore arises: how can governments ensure that existing spending delivers the strongest possible results?

In public finance, efficiency should not be confused with indiscriminate cost cutting. An efficient system is one that achieves strong outcomes for a given level of resources or achieves a given set of outcomes using fewer resources than peers. Inefficiency carries a double cost: it absorbs scarce public funds without commensurate benefits, and it delays improvements that could otherwise be achieved within existing budgets. For example, if two countries spend broadly similar amounts on health, but one achieves better outcomes, the other may have scope to improve performance without increasing spending, provided that differences in context and population risks are taken into account.

Measuring efficiency in public services, however, is not straightforward. Simple indicators such as spending per capita, spending as a share of GDP, or "cost per outcome" ratios provide only a partial picture. For example, public health systems pursue multiple objectives at the same time, improving longevity, reducing avoidable deaths, and supporting quality of life, while also ensuring access and equity. Cross-country comparisons are further complicated by differences in demographics, disease burden, and delivery models. As a result, raw comparisons can mistakenly attribute poor performance to inefficiency when it partly reflects structural conditions.

To address these challenges, this paper uses Data Envelopment Analysis (DEA) to benchmark health spending efficiency across countries. DEA compares each country to a best-practice benchmark observed in the data. Importantly, the benchmark does not have to be a single country: it can be a weighted blend of similar high-performing peers. For example, Ireland might be compared to a mix of Spain and Italy if that combination provides the closest best-practice reference for Ireland's spending level and outcome profile. The method can be applied to public services because it can assess performance using multiple outcome indicators at once and produces results that are intuitive for policy, such as the implied scope to reduce inputs while maintaining outcomes, or to improve outcomes at current spending.

This paper makes four contributions. First, it provides a transparent, multi-outcome benchmark of public health spending efficiency across a European OECD peer group. Second, it reports results from both a cost-containment perspective (how much spending could theoretically be reduced for given outcomes) and an improvement perspective (how much outcomes could rise for given spending). Third, it tests whether results are sensitive to the choice of peer countries using standard robustness checks. Fourth, it explores how contextual factors such as income and health risks relate to measured efficiency gaps, helping to interpret results.

Empirically, the framework is applied to OECD data on health expenditure and health outcomes, focusing on a European sample and on Ireland's position relative to peer systems. The purpose is not to create a simple league table.

Rather, it is to quantify the size of efficiency gaps, identify plausible peer benchmarks, and clarify whether gaps appear more consistent with operational factors or broader structural constraints.

The paper proceeds as follows. Section 2 reviews the relevant literature. Section 3 sets out the methodology. Section 4 describes the data and variables. Sections 5–7 present results, robustness checks, and correlates of inefficiency. Section 8 concludes, and Section 9 outlines limitations and directions for future research.

2. DEA Applications to Public Spending Efficiency

Since the seminal work of Charnes, Cooper, and Rhodes (1978) introduced Data Envelopment Analysis (DEA) as a non-parametric method for measuring efficiency, the approach has been widely adopted to benchmark public expenditure against outcomes. Applications now span health, education, infrastructure, and public administration. In cross-country OECD studies, DEA is used to construct efficiency frontiers and estimate potential spending reductions (input-oriented models) or performance improvements (output-oriented models) for countries operating below best practice.

A representative example is Dutu and Sicari (2020), who apply DEA to welfare spending across OECD members, covering health, secondary education, and general public services. Using composite indicators to control for socio-economic context, their results demonstrate that DEA is particularly useful for positioning countries relative to peer frontiers rather than producing simple rankings. Systematic reviews reinforce this breadth. For example, Rostamzadeh et al. (2021) survey DEA benchmarking between 2003 and 2020, documenting its extensive use in domains where market prices are unavailable. They conclude that DEA's value lies in its ability to flag weak functions and quantify potential improvement targets, even if it cannot prescribe specific operational changes.

Health is one of the most extensively studied areas in the DEA literature. Jung et al. (2023) synthesise 123 studies from 2017–2022, revealing a surge in research driven by the COVID-19 pandemic and the need to optimise resource allocation under demographic pressure. Their review highlights a maturation in the field: while traditional Charnes-Cooper-Rhodes (CCR) and Banker-Charnes-Cooper

(BCC) models remain common, there is a clear shift toward advanced techniques such as Network DEA and Slack-Based Measures (SBM).

Crucially, modern health efficiency research increasingly integrates second-stage statistical tools. Jung et al. (2023) note the widespread use of Tobit regression to identify determinants of inefficiency, while complementary reviews (e.g., Cantor & Poh, 2018) illustrate how combining DEA with econometric analysis provides a richer picture of the drivers of efficiency. This evolution is evident in longitudinal OECD studies. For instance, Aydın (2022) applies a DEA-Malmquist productivity index to OECD health systems (2000–2015), distinguishing between gains made by using existing resources better versus gains driven by technological innovation. Similarly, Gavurová et al. (2021) utilize dynamic network DEA to trace efficiency developments, distinguishing between public health (prevention) and medical care (treatment) sub-systems.

In the Irish context, recent assessments point to a "productivity paradox", wherein substantial resource inputs have not translated into commensurate gains in health outcomes. Sicari & Sutherland (2023) note that despite Ireland's comparatively high per-capita health spending, the system continues to suffer from structural inefficiencies, most notably a two-tier access model and a hospital-centric pattern of care delivery that drives up costs. National spending reviews echo these concerns. Meaney, Oyewole, and Bedogni (2018), benchmarking Ireland against EU counterparts, highlight a missed demographic dividend: although Ireland has one of the youngest populations in Europe, it nonetheless spends more than many older EU countries while achieving comparable or, in some cases, weaker outcomes.

This study complements the existing operational literature by utilizing the Simar–Wilson framework. While advanced bootstrapping techniques have appeared in cross-country studies (e.g., Dutu & Sicari, 2020), they have rarely been applied to rigorously test the determinants of Irish public spending efficiency. By adopting this approach, we can econometrically isolate environmental drivers, specifically testing the hypothesis that national wealth (Baumol's cost disease) and population morbidity exert non-linear pressures that define the efficiency frontier in high-income jurisdictions.

3. Methodology

This paper benchmarks how effectively countries translate health-system resources into health outcomes using Data Envelopment Analysis (DEA). DEA is a frontier method: instead of estimating the average relationship between spending and outcomes, it identifies the best observed performance in the data and measures how far each country is from that best-practice frontier. In practical terms, it answers a policy-relevant question: given what peers already achieve, what does best practice look like at this level of spending (or what level of spending is consistent with this level of outcomes)? Since DEA can handle multiple outcomes simultaneously, it is particularly suitable for public services where performance is multidimensional.

Each country is treated as a decision-making unit. Let $x_j \geq 0$ denote the input (public health expenditure per capita, PPP-adjusted, USD) for country *j*, and let $y_j \geq 0$ denote a vector of health outcomes. DEA constructs a benchmark for each country using peer weights $\lambda_j \geq 0$, which form a feasible "best-practice comparison point" as a weighted blend of observed countries. The peer benchmark does not need to be a single country; it can be a combination of several relevant peers. For example, a country may be benchmarked against a blend of Spain and Italy if that mix provides the closest best-practice reference for its spending and outcome profile.

DEA can be estimated from two complementary perspectives that map directly to policy questions. The input-oriented model asks: holding outcomes fixed, by how much could spending be reduced if the country operated as efficiently as best practice? For a given country k, the input-oriented model under Variable Returns to Scale (VRS) solves:

$$\min_{\theta,\lambda} \theta$$

subject to:

$$\sum_{j=1}^{n} \lambda_j x_j \leq \theta x_k, \quad \sum_{j=1}^{n} \lambda_j y_j \geq y_k, \quad \sum_{j=1}^{n} \lambda_j = 1, \quad \lambda_j \geq 0$$

The scalar $\theta \in (0,1]$ is the efficiency score. A value such as $\theta=0.55$ implies that, relative to best practice in the sample, the country's spending could theoretically be reduced proportionally to 55% of its current level while maintaining the same measured outcomes, i.e. an indicative 45% benchmarking distance $(1-\theta)$. This is a relative frontier distance, not a short-run savings plan: it does not account for transition costs, institutional constraints, or feasibility in the near term.

The output-oriented model asks the mirror question: holding spending fixed, by how much could outcomes improve under best practice? For country k, the output-oriented VRS model solves:

$$\max_{\phi,\lambda} \ \phi$$

subject to:

$$\sum_{j=1}^{n} \lambda_j \mathrm{x}_j \leq x_k, \qquad \sum_{j=1}^{n} \lambda_j \mathrm{y}_j \geq \phi y_k\ , \qquad \sum_{j=1}^{n} \lambda_j = 1, \quad \lambda_j \geq 0$$

Here $\phi \geq 1$ is the proportional outcome expansion factor; for example ϕ=1.03 corresponds to an indicative 3% proportional improvement in the measured outcomes at current spending (ϕ−1).

A key practical issue in cross-country benchmarking is scale. Countries may differ not only in operational performance but also in whether they operate at an efficient scale. We therefore estimate DEA under both VRS and Constant Returns to Scale (CRS). CRS is obtained by dropping the convexity constraint $\sum_{j=1}^{n} \lambda_j = 1$. Let $\theta_k^{VRS}$and θ_k^{CRS} denote the input-oriented efficiency scores for country *k* under VRS and CRS respectively. Scale efficiency is then:

$$SE_k = \frac{\theta_k^{CRS}}{\theta_k^{VRS}}$$

This decomposition helps distinguish operational inefficiency (captured under VRS) from inefficiency related to operating at a non-optimal scale (captured by the VRS–CRS gap).

Some health outcomes are "bads" (e.g. mortality rates), where lower values indicate better performance. Because DEA treats outputs as "goods" (more is better), we transform mortality indicators using a monotonic mapping that preserves rankings. In the baseline we use:

$$y^* = \frac{1}{m + \varepsilon}$$

where m is the mortality rate and ε is a small constant to avoid division by zero. This transformation ensures that lower mortality corresponds to higher output.

DEA also produces peer benchmarks and targets. For country k, the projected frontier point is:

$$\hat{x_k} = \sum_{j=1}^{n} \lambda_j x_j, \quad \hat{y_k} = \sum_{j=1}^{n} \lambda_j y_j$$

These projections help translate scores into an interpretable benchmark: the λ-weights indicate which peers matter most in defining best practice for a given country.

Since DEA is deterministic and frontier estimates can be influenced by extreme observations, we assess the stability of results using complementary sensitivity approaches, including leave-one-out analysis and partial frontier methods. These tests examine whether conclusions depend on any single peer or on a small set of extreme performers.

Finally, to examine correlates of inefficiency while accounting for the statistical properties of DEA scores, we use the Simar–Wilson (2007) double-bootstrap procedure. DEA efficiency scores are not independent because the frontier is estimated from the same data used to compute the scores; naïve second-stage regressions can therefore yield invalid inference. The double bootstrap addresses this by (i) bias-correcting DEA efficiency estimates and (ii) providing valid confidence intervals and p-values in the second stage. We transform bias-corrected efficiency scores θ_k^{BC} into an inefficiency distance δ_k:

$$\delta_k = \frac{1}{\theta_k^{BC}} - 1, \quad \delta_k \geq 0,$$

and relate δ_k to contextual factors z_k (e.g., obesity prevalence and GDP per capita, including a quadratic term) in a truncated regression:

$$\delta_k = z_k \beta + \varepsilon_k$$

where the error term ε_k is distributed as a normal distribution $N(0,\sigma^2)$ with left-truncation, ensuring that inefficiency never falls below zero.

Second-stage results are interpreted as associations that help contextualise benchmarking gaps, not as causal estimates of policy effects.

4. Data and sample selection

The dataset is constructed from the OECD Health Statistics 2025 database, which provides the most comprehensive set of internationally comparable indicators on health expenditure, resources, and outcomes across OECD member countries. The analysis focuses on a cross-section of 25 European OECD countries, selected to improve comparability by limiting heterogeneity in health-system governance, epidemiological profiles, and data-reporting standards.[2]

To enable cross-country comparability, all financial variables are expressed in US Dollars at Purchasing Power Parity (PPP), adjusting for differences in price levels and purchasing power between jurisdictions. The primary input is Public Health Expenditure per Capita (Current Prices, PPP). This covers expenditure by government and compulsory contributory health financing schemes, excluding voluntary private payments.

The analysis uses three indicators of health outcomes to capture different aspects of system performance:

- Life expectancy at birth (years): a broad measure of population health.
- Infant mortality rate (per 1,000 live births): an indicator responsive to perinatal care quality and the effectiveness of acute care for infants.
- Avoidable mortality (per 100,000 population): deaths considered preventable and/or treatable with effective public health interventions and timely, high-quality medical care. This indicator is generally more directly responsive to health care system performance than life expectancy, which is also influenced by non-health care factors such as lifestyle and environmental conditions.

[2] Norway is excluded because of substantial lags in key outcome data, most notably avoidable mortality, for which the latest observation is 2016, which would weaken the validity of cross-country comparisons in a predominantly 2023 (or near-2023) dataset. The European OECD sample includes: Austria, Belgium, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, United Kingdom; other non-European OECD members are excluded to improve comparability.

Two environmental variables are included to account for structural differences across countries: GDP per capita (PPP) reflects national economic capacity, while the obesity rate (% of the population) serves as a proxy for the underlying morbidity burden and risk-factor prevalence.

OECD indicators are not always available for the same year across countries, so the dataset follows a latest-available-year strategy. Typically, 2023 values are used; otherwise, the most recent valid observation is retained (particularly for avoidable mortality, which may be reported for 2022 or 2021). More flexibility is allowed for obesity rates, since they rely on self-reported surveys collected at irregular intervals. This approach maximises sample coverage while remaining close to a common reference period.

Table 1 summarises the characteristics of the final sample (n = 25). A clear pattern emerges: financial inputs vary widely, while life expectancy is comparatively clustered.

Table 1. Descriptive Statistics of the European Sample (n=25)

Variable	**Mean**	**Std. Dev.**	**Min**	**Max**
Inputs				
Public Health Expenditure (US$ PPP, per capita)	4,357	1,682	1,612	7,308
Outputs				
Life Expectancy at Birth (years)	81.2	2.4	75.6	84.3
Infant Mortality (per 1,000 live births)	3.0	0.9	1.7	5.6
Avoidable Mortality (per 100,000 population)	213.4	85.3	114	412

Environmental				
Obesity Rate (%)	18.2	4.0	11.8	29.0
GDP per Capita (US$ PPP)	67,936	26,603	42,045	152,293

Source: Author's analysis based on OECD Health Statistics 2025. Values in US dollars (PPP). Sample excludes Norway due to outcome data lags. Data correspond to the latest available year (primarily 2023; otherwise typically 2022/2021). Obesity rates are self-reported. Avoidable mortality covers preventable and treatable deaths.

Public health expenditure shows substantial dispersion, with an average of $4,357 per capita and a standard deviation of $1,682. The highest-spending country allocates $7,308 per person, more than four times the level of the lowest spender ($1,612).

By contrast, life expectancy varies much less. The average is 81.2 years (with a standard deviation of 2.4), and the gap between the highest (84.3 years) and lowest (75.6 years) values is 8.7 years, modest relative to the dispersion in spending. This pattern is consistent with the possibility of diminishing marginal gains in longevity at higher levels of resourcing, although life expectancy is also shaped by broader socio-economic and behavioural factors.

Other outcomes exhibit greater divergence. Infant mortality averages 3.0 deaths per 1,000 live births but ranges from 1.7 to 5.6, reflecting larger cross-country differences in early-life outcomes and care, including perinatal services. Avoidable mortality averages 213.4 deaths per 100,000 population and ranges from 114 to 412, an almost fourfold difference, suggesting marked cross-country differences in prevention and the timeliness and effectiveness of care for conditions amenable to healthcare. Taken together, these patterns imply that longevity outcomes are relatively concentrated across Europe, while dispersion is greater for indicators more responsive to prevention and timely, effective care.

Environmental indicators also vary considerably. GDP per capita averages $67,936 (PPP) with a standard deviation of $26,603, reflecting Europe's economic diversity. The obesity rate averages 18.2%, reaching 29% in the highest-obesity country. This variation is important, as higher morbidity burdens typically raise the resources required to achieve comparable health outcomes.

5. Results and analysis

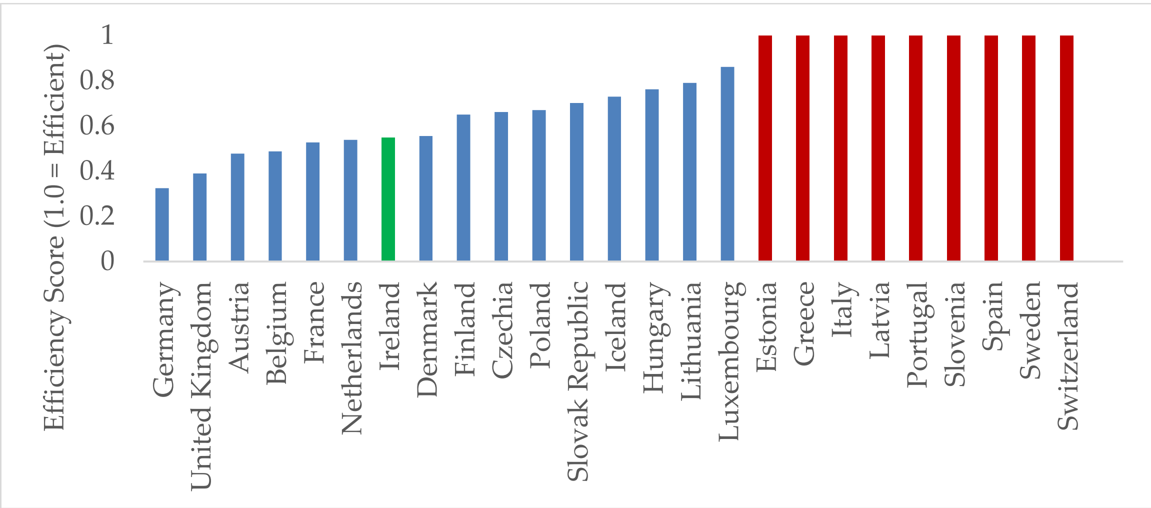
5.1. Spending Efficiency (Input-Oriented DEA): How Far Countries Are from Best Practice

Applying the input-oriented VRS DEA model shows that countries reach (or define) the efficiency frontier through several distinct routes (Figure 1). Because the input is expenditure per capita (rather than physical inputs such as staff or beds), these scores should be read primarily as value-for-money (spending) efficiency, i.e., how much spending is required to achieve a given bundle of outcomes, rather than standard technical (operational) efficiency.

First, a low-spending route is evident. In the current sample, Latvia (public health expenditure $1,612 per capita) lies on the efficiency frontier despite comparatively weak outcomes (life expectancy 75.6 years; avoidable mortality 412 per 100,000). In DEA terms, Latvia represents a "corner solution": exceptionally low inputs can place a country on the frontier even when outcomes are weaker, because the frontier is constructed from the observed combinations of inputs and outputs.[3]

However, the frontier is not defined by low spending alone. Greece reaches the frontier with low public spending ($2,095) while maintaining relatively strong longevity (81.8 years) and mid-range avoidable mortality (213). Estonia is also on the frontier at $2,555, combining moderate longevity (79.1 years) with very low infant mortality (1.7 per 1,000).

Figure 1. Input-Oriented Efficiency Score (VRS), Higher Score = More Efficient Spending.



Source: Author's analysis. **Note**: Results are from the full multi-output DEA specification (outputs: life expectancy, infant mortality, avoidable mortality; input: public health expenditure per capita). The input-oriented efficiency score (θ) should therefore be interpreted primarily as an expenditure/value-for-money measure (the proportional spending contraction implied for a given outcome bundle), rather than as pure technical efficiency based on physical inputs (e.g., staff or beds).

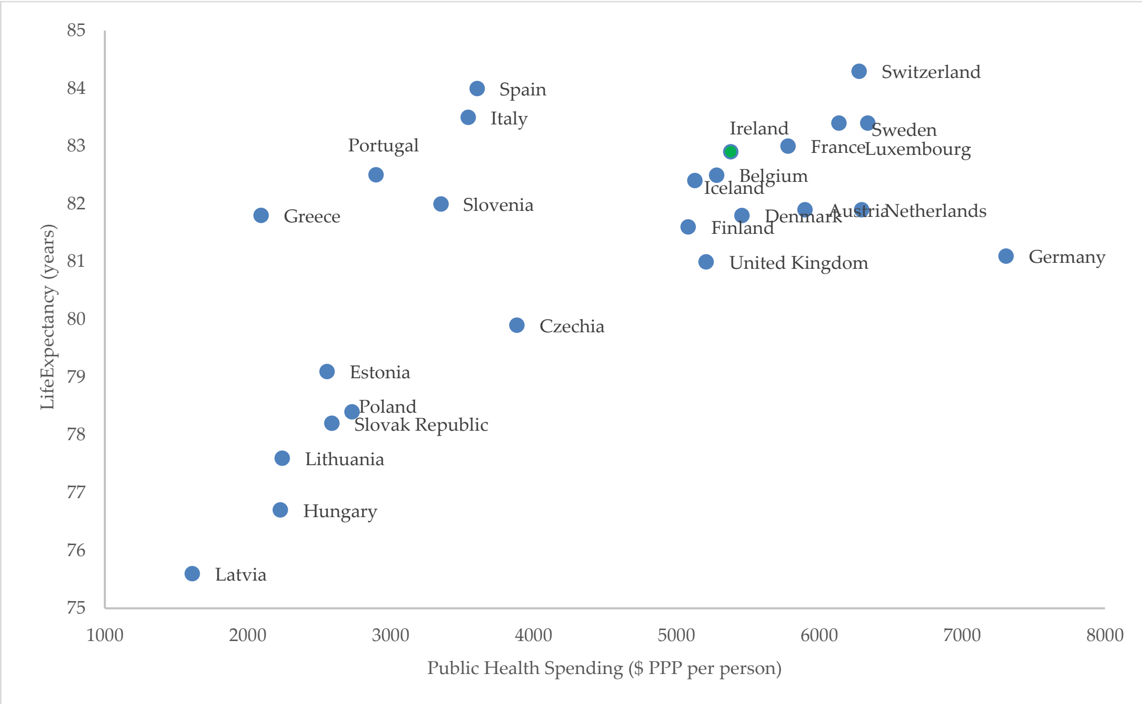
[3] To assess whether Ireland's estimated distance to the frontier is driven by low-spending "corner" observations, we inspected the VRS reference set weights (λ). Ireland's input-oriented benchmark is a convex combination of Spain ($\lambda \approx 0.57$) and Greece ($\lambda \approx 0.43$), while the output-oriented benchmark is a convex combination of Switzerland ($\lambda \approx 0.57$) and Spain ($\lambda \approx 0.43$). Latvia receives $\lambda = 0$ in both cases, indicating that Ireland's results are not mechanically anchored by the lowest-spending corner point.

Second, an outcomes-strong, moderate-spending route provides a more policy-relevant benchmark for higher-income systems. Portugal, Italy, Slovenia, and Spain are on the efficiency frontier while combining comparatively strong outcomes with spending at or below the sample average (Portugal $2,899; Italy $3,542; Slovenia $3,352; Spain $3,606). Among these, Spain is particularly salient: it achieves very high life expectancy (84.0 years) with public spending ($3,606) below the sample mean ($4,357).

Third, the frontier also includes high-spending, high-performance systems such as Sweden ($6,339) and Switzerland ($6,279). These countries combine higher spending with exceptionally strong outcomes, especially very low avoidable mortality (133 in Sweden; 114 in Switzerland), which can also place them on the frontier in a multi-output setting.

Figure 2 provides an intuitive visualisation using life expectancy as one outcome dimension (noting that the DEA frontier is estimated in multi-output space). The association between spending and life expectancy is steep at lower spending levels, but becomes flatter around the $3,500–$3,700 range, close to Italy ($3,542) and Spain ($3,606), consistent with diminishing marginal gains in longevity at higher spending levels in this cross-section.

Figure 2. Illustrative relationship between spending and life expectancy.



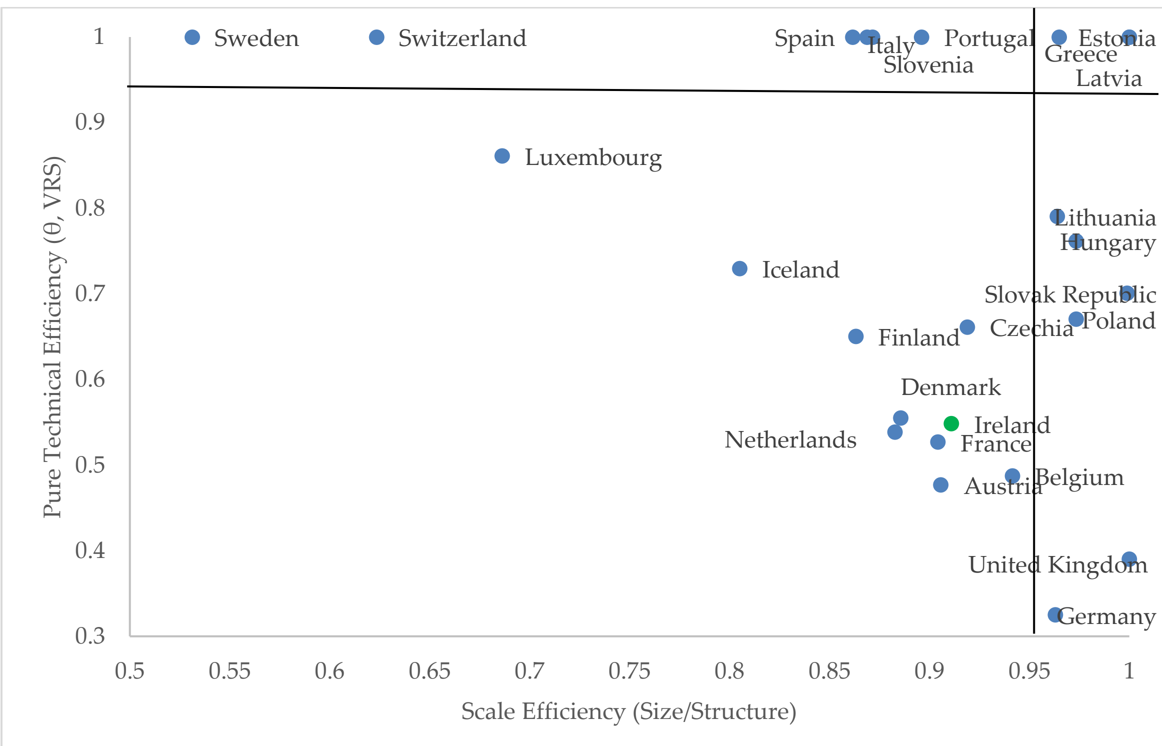
Source: Author's analysis. **Notes:** Illustrative relationship between spending and life expectancy; the DEA frontier is estimated in multi-output space.

The country comparisons underscore this pattern. Germany spends $7,308 per person, approximately twice Spain's level, yet reports life expectancy 2.9 years lower (81.1 versus 84.0). France ($5,781) and Ireland ($5,382) likewise spend substantially more than Spain while achieving only marginally lower longevity (83.0 and 82.9, respectively).

Ireland's input-oriented VRS efficiency score is $\theta = 0.55$. Mechanically, this corresponds to an indicative frontier gap of about 45% (i.e., the proportional input reduction required to reach the frontier while holding the outcome bundle constant in the model). This should be interpreted as a theoretical benchmarking distance rather than a short-run savings target, given fixed costs, workforce constraints, and the high-wage platform characteristic of the Irish economy.

Figure 3 breaks Ireland's efficiency gap into two parts: system size (scale) and how well the system runs at its current size. Ireland's scale efficiency is 0.911, which suggests the system is operating close to an appropriate overall scale in this cross-country comparison. The remaining gap is therefore more likely to reflect how resources are used within the existing system, rather than the system being fundamentally "too big" or "too small". This points to greater potential in cost control and service organisation, for example, care pathways, procurement, workforce mix, and governance, and/or differences in input prices and case mix that are not fully captured by PPP adjustments.

Figure 3. Technical vs. Scale Efficiency



Source: Author's analysis. **Notes**: The input-oriented efficiency score (θ) (y-axis) should be interpreted primarily as an expenditure/value-for-money measure (the proportional spending contraction implied for a given outcome bundle), rather than as pure technical efficiency based on physical inputs (e.g., staff or beds). Scale efficiency $SE_k = \frac{\theta_k^{CRS}}{\theta_k^{VRS}}$ (x-axis) measures inefficiency related to operating at a non-optimal scale.

To ensure that this conclusion is not driven by the exclusive use of public spending, we also re-estimated the multi-output, input-oriented VRS DEA model using total health expenditure per capita (OECD Health Statistics) as the single input, keeping the same outcome bundle and transformations.

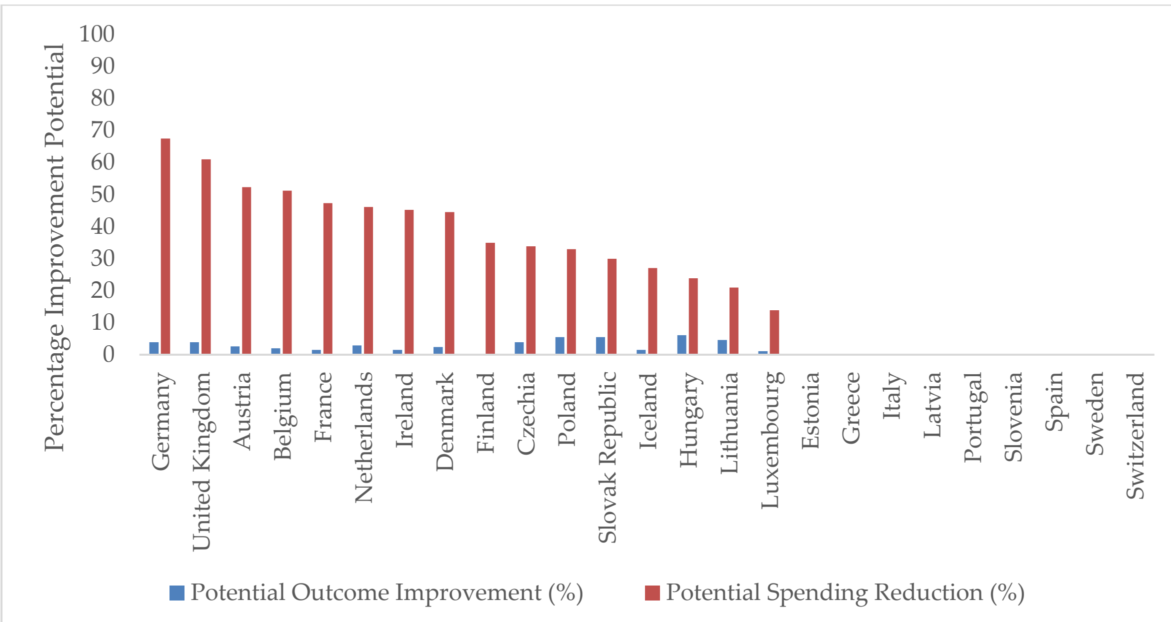
Results are highly consistent across the two input definitions: country efficiency scores remain closely aligned, and Ireland continues to be characterised as outcome-effective but input-intensive. Under the total-spending specification, Ireland's estimated efficiency rises modestly (θ from 0.548 to approximately 0.609), implying a smaller, though still substantial, indicative input gap (from roughly 45% to about 39%). Overall, the spending definition affects the magnitude of the estimated gap, but not the substantive conclusion on scope to improve value for money at broadly similar outcomes.

5.2. Output-Oriented Efficiency: Potential Outcome Gains at Current Spending

The output-oriented model provides a complementary perspective. Instead of asking how much spending could be reduced while holding outcomes constant (input orientation), it asks how much outcomes would need to improve, proportionally and simultaneously, to match best practice at current spending levels. The results show that the implied scope for proportional outcome gains among high-spending countries is small. Germany's output-oriented score is ϕ = 1.038, implying an indicative outcome expansion (ϕ -1) of about four percent; France's is ϕ = 1.015 (about one and a half percent); and Ireland's is ϕ = 1.015 (also about one and a half percent).

This is very small relative to the input-oriented gaps estimated for the same countries. For example, Germany's input-oriented score is θ = 0.325 (an indicative input gap (1- θ) of roughly two-thirds), while France (θ = 0.527) and Ireland (θ = 0.548) imply input gaps of roughly one-half. In other words, the model suggests large potential savings distances on the input side, but only modest proportional gains on the output side.

Figure 4. Asymmetry of Inefficiency: Savings vs Gains



Source: Author's analysis. **Notes**: Multi-output DEA (life expectancy, infant mortality, avoidable mortality). Output "gain" is calculated as (ϕ − 1); input "gap/savings distance" is calculated as (1 − θ).

A plausible interpretation is that measured outcomes, especially longevity, are already close to practical ceilings in many high-spending systems, so further proportional improvements become progressively harder to achieve absent a major technological breakthrough. However, this 'limited headroom' result is sensitive to the inclusion of life expectancy, which is clustered in Europe. In a robustness check excluding life expectancy and focusing on mortality outcomes (infant and avoidable mortality), Ireland's input-oriented score is unchanged (θ≈0.55), but the output-oriented score rises to ϕ≈1.25 (transformed space; see footnote for mapping back to mortality units), indicating greater scope for mortality improvement at existing spending.[4]

Overall, for the baseline longevity-inclusive bundle, the frontier comparison suggests that the main opportunity is less about generating large additional proportional gains in these particular outcomes at current spending, and more about reducing the cost of producing existing outcomes—through cost control, service organisation, and productivity-related reforms.

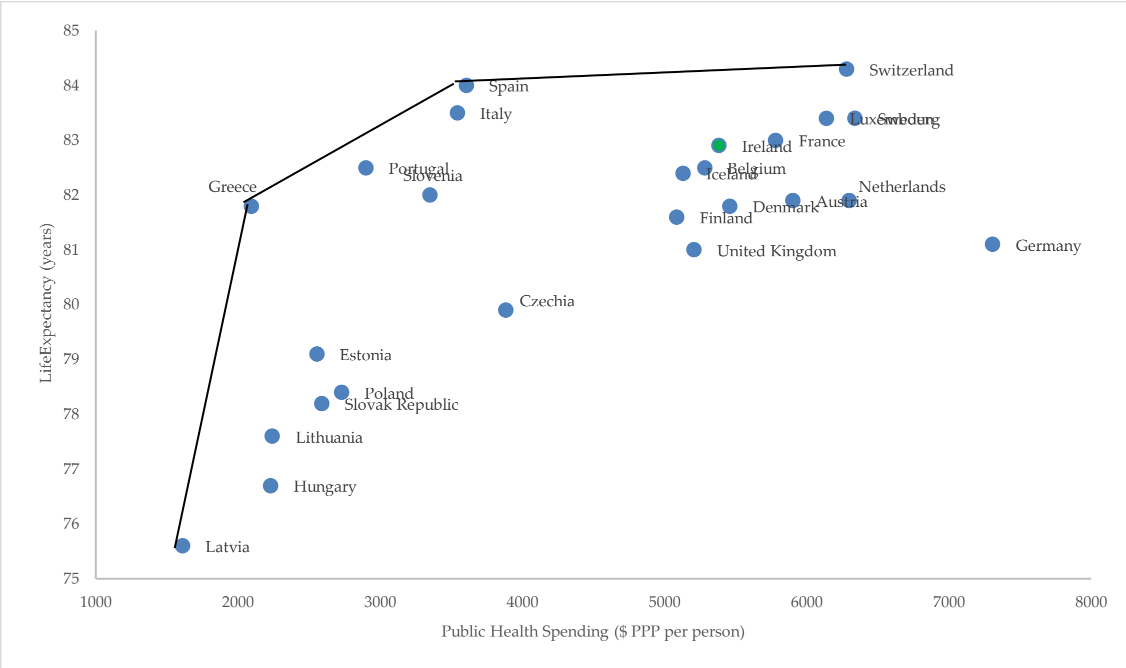
5.3. Disaggregated Efficiency Analysis

The multi-output DEA results summarise overall value for money across several outcomes at once. To understand which outcome dimensions matter most for a country's distance from the frontier, we also estimated three separate input-oriented VRS models, each using the same spending input but focusing on a single outcome: life expectancy, infant mortality, and avoidable mortality. The purpose of these single-outcome models is not to replace the multi-output results, but to identify where the implied efficiency gap is largest.

---

[4] Under the inverse transformation, a proportional increase in $y*$ by ϕ corresponds approximately to scaling mortality levels by 1/ϕ. For Ireland, that is roughly a 20% reduction in mortality indicators. In this mortality-focused output-oriented model, Ireland's benchmark is a composite of Sweden (λ≈0.50), Spain (λ≈0.35), and Switzerland (λ≈0.15).

Figure 5. Frontier 1: Life Expectancy (Efficiency Frontier)



Source: Author's analysis. **Notes**: Latvia, Greece, Spain, and Switzerland are countries that lie on the efficiency frontier for life expectancy, based on an input-oriented VRS model.

Figure 5 shows the frontier for life expectancy. Greece, Latvia, Spain and Switzerland lie on the efficiency frontier in the single-outcome life expectancy model.

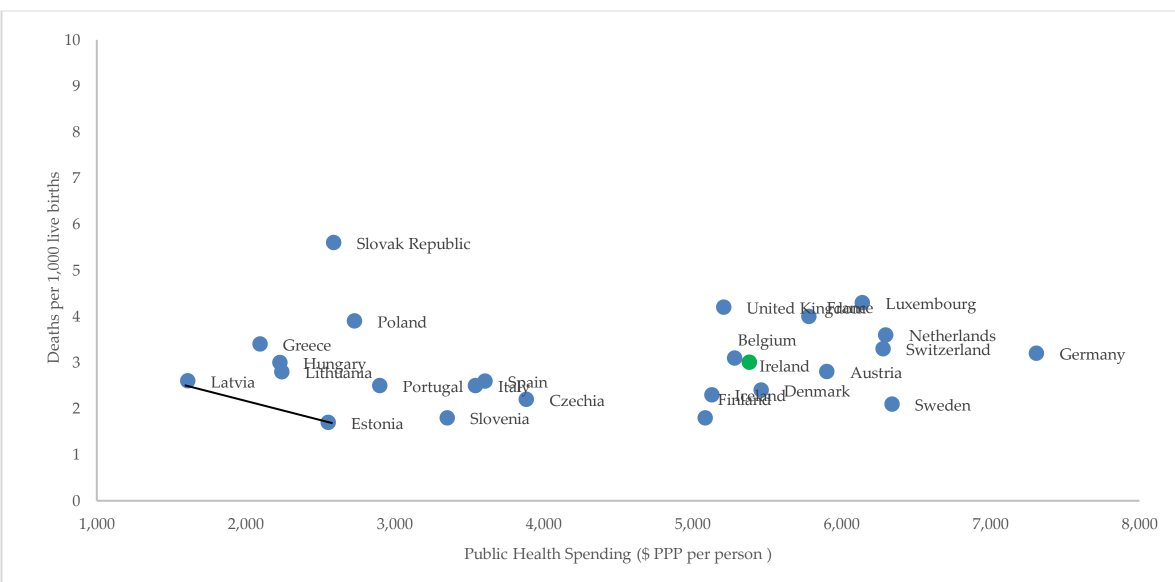
Ireland's life expectancy is very high (82.9 years), but the single-outcome efficiency score for life expectancy is 0.530, which indicates that, relative to the best-practice benchmark set, Ireland achieves its longevity outcomes at a comparatively high spending level. Put simply: Ireland's longevity is excellent, but expensive to produce in this cross-country benchmarking exercise.

Infant mortality is often treated as an indicator of early-life and perinatal outcomes. The outcome leaders include Estonia (1.7), Finland (1.8) and Slovenia (1.8) deaths per 1,000 live births. In the single-output efficiency model, Estonia and Latvia sit on the efficiency frontier. Estonia is particularly notable because it combines the lowest infant mortality rate (1.7) with comparatively low spending ($2,555). Latvia again represents a low-spending "corner" point that can define the frontier even if outcomes are not the very best in absolute terms.

Ireland's infant mortality rate is 3.0. While this is good by international standards, it is materially higher than the best performers in the sample. Importantly, Ireland's single-outcome efficiency score for infant mortality is 0.300, which is substantially lower than its life expectancy score, suggesting that, in this benchmarking framework, infant mortality contributes the largest share of Ireland's distance from best practice.

A simple comparison illustrates the point: Ireland spends slightly more than Finland ($5,382 vs $5,084) yet records an infant mortality rate that is about two-thirds higher (3.0 vs 1.8).

Figure 6. Frontier 2: Infant Mortality (Lower is Better)

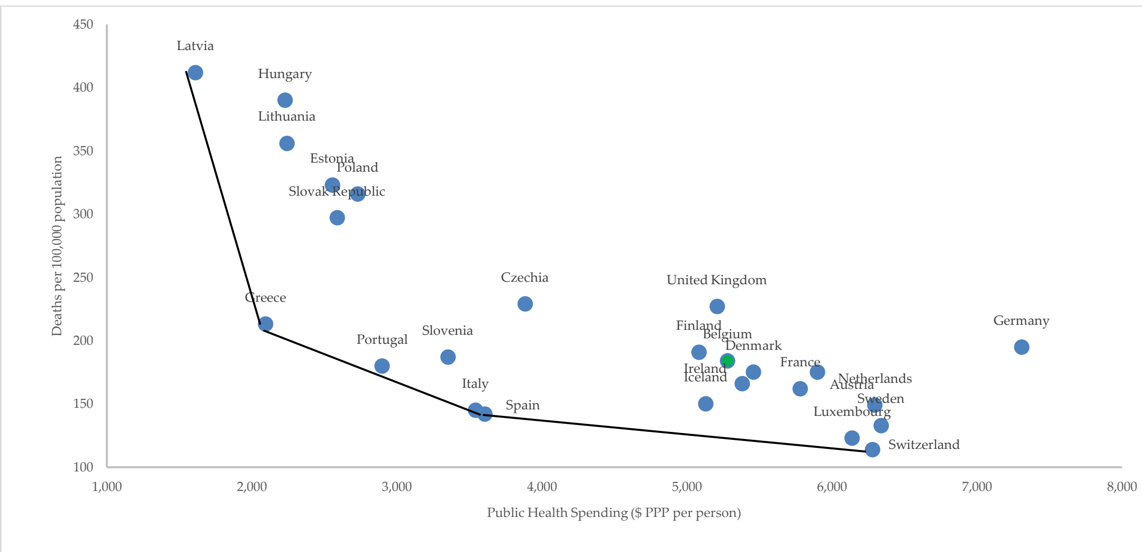


Source: Author's analysis. **Notes**: Latvia and Estonia are countries that lie on the efficiency frontier for infant mortality, based on an input-oriented VRS model.

Avoidable mortality captures deaths that should be reducible through effective prevention and timely healthcare. The lowest avoidable mortality rate is in Switzerland (114), followed by Luxembourg (123) and Sweden (133), with Spain (142) and Italy (145) also performing strongly.

In the single-output efficiency model, the frontier countries are Greece, Latvia, Spain and Switzerland. As with life expectancy, this set reflects both: high performance (Switzerland, Spain), and low-spending corner points (Latvia, and Greece).

Figure 7. Frontier 3: Avoidable Mortality (Lower is Better)



Source: Author's analysis. **Notes**: Latvia, Greece, Spain and Switzerland are countries that lie on the efficiency frontier for avoidable mortality, based on an input-oriented VRS model. In an input-oriented DEA, a country can lie on the frontier by achieving a given outcome level at exceptionally low spending, even if its outcomes are not high in absolute terms. Such 'corner' observations define the low-input boundary of feasible performance.

Ireland's avoidable mortality rate is 166 per 100,000, and its single-outcome avoidable mortality efficiency score is 0.548. This places Ireland closer to its overall multi-output result than infant mortality does, but still indicates a meaningful "value for money" gap relative to the frontier. In simple terms: Ireland's avoidable mortality is better than many peers, but not as strong as the best performers given its spending level.

Taken together, the single-outcome models reinforce a consistent message:

- Ireland's performance is strongest on life expectancy (excellent longevity, but achieved at relatively high cost).
- The largest efficiency shortfall (in this benchmarking sense) appears in infant mortality, where Ireland's single-outcome efficiency score is much lower (0.300) than for life expectancy (0.530) or avoidable mortality (0.548).
- Avoidable mortality sits in between: Ireland performs well, but the frontier suggests some peers achieve similar or better outcomes with less spending.

This pattern supports the interpretation that Ireland's biggest opportunities are less about buying large additional health gains through higher spending, and more about improving how effectively resources translate into outcomes.

5.4. Ireland's spending composition in an EU context

The DEA scores summarise Ireland's distance to the multi-output efficiency frontier, but they do not show how Ireland's spending mix differs from comparator systems. To provide interpretive context, this subsection uses Eurostat's System of Health Accounts (SHA) breakdown of current healthcare expenditure (CHE) to compare Ireland with the EU average across providers (who delivers care), functions (what care is for), and financing (who pays). The comparison is descriptive rather than causal: different spending shares do not automatically signal inefficiency, but they can indicate structural features that distinguish Ireland's health system from the European benchmark.

On the provider side, Ireland's profile is noticeably more institutional than the EU average. In 2023, hospitals accounted for 39.91% of Ireland's CHE versus 36.80% for the EU (a +3.11 percentage-point difference). The more striking divergence is in residential long-term care (LTC) facilities, which account for 17.08% in Ireland compared with 10.36% in the EU (+6.72pp). Offsetting these higher institutional shares, Ireland is lower than the EU in both ambulatory providers (21.03% vs 25.62%, −4.59pp) and retailers/other medical goods providers (11.57% vs 16.24%, −4.67pp). In simple terms, Ireland allocates more of its current health spending through hospitals and residential long-term care settings, and less through ambulatory and retail channels than the EU average.

The same broad pattern appears when spending is decomposed by function. Ireland allocates 58.40% of CHE to curative and rehabilitative care compared with an EU average of 52.69% (+5.71pp), and 21.60% to long-term care (health) compared with 17.10% (+4.50pp). These higher service shares are matched by materially lower spending on medical goods (11.76% vs 17.77%, −6.01pp) and lower shares for ancillary services (2.57% vs 5.04%, −2.47pp). Ireland is also below the EU average in preventive care (2.64% vs 3.65%, −1.01pp) and governance/health system administration (2.23% vs 3.55%, −1.32pp). Overall, Ireland's functional mix is more heavily concentrated in curative and long-term care services, with comparatively smaller allocations to medical goods, ancillary services, and prevention.

Finally, the financing split underscores that Ireland sits in a different institutional funding family than the EU average. Ireland's CHE is predominantly financed by government schemes (76.0%), with compulsory contributory schemes at 0.6%, voluntary health insurance at 10.58%, and household out-of-pocket payments at 11.3% (with small additional shares in other categories). The EU average shows a much larger role for compulsory schemes (52.31%) and a smaller government share (28.24%), with out-of-pocket payments at 14.88%. Taken together, Ireland resembles a predominantly tax-funded system with a non-trivial voluntary insurance layer, whereas the EU average reflects insurance-based financing as the dominant pillar.

5.5. Frontier archetypes: how "efficient" mixes vary across frontier countries

The baseline DEA frontier countries, Estonia, Greece, Italy, Latvia, Portugal, Slovenia, Spain, Sweden, and Switzerland do not share a single spending template. Instead, Eurostat's SHA data suggest several distinct "frontier archetypes", meaning coherent clusters of provider and functional mixes that are compatible with frontier status. This matters for interpretation: if frontier countries achieve benchmark outcomes with different spending structures, then Ireland's distance from the frontier is less about matching any one model and more about understanding how Ireland's particular combination of shares differs from the configurations observed among efficient peers.

A first archetype can be described as "hospital-heavy, LTC-light". Spain, Italy, and Portugal all allocate a substantially larger share of spending through hospitals than the EU average: Spain's hospital share is 45.44% (+8.64pp vs EU), Italy's 43.47% (+6.67pp), and Portugal's 43.61% (+6.81pp). Yet these systems simultaneously record markedly lower SHA-defined long-term care (health) than the EU: Spain 10.19% (−6.91pp), Italy 10.24% (−6.86pp), and Portugal 5.16% (−11.94pp). What defines this archetype is not low hospital spending, hospitals take a relatively large share, but rather a comparatively small share allocated to SHA-classified long-term care (health), so the overall mix can still align with frontier efficiency.

A second archetype is “hospital-heavy with a strong medical-goods/retail channel and very low LTC”, illustrated most clearly by Greece. Greece’s hospital share is 44.96% (+8.16pp vs EU), but what distinguishes it is the scale of medical goods/retail: retailers and other medical-goods providers account for 30.06% (+13.82pp vs EU), and the functional share for medical goods is 29.18% (+11.41pp vs EU). At the same time, Greece records an exceptionally low LTC(health) share of 1.91% (−15.19pp vs EU). In this configuration, frontier status is consistent with a system that is hospital-weighted but also channels a much larger proportion of spending into medical goods, alongside minimal SHA-recorded long-term care.

A third archetype can be labelled “LTC-heavy frontier”, exemplified by Sweden and Switzerland. Sweden’s LTC (health) share is 27.51% (+10.41pp vs EU) and Switzerland’s is 20.94% (+3.84pp), and both are also above the EU average in residential long-term care facility spending (Sweden 18.92%, +8.56pp; Switzerland 16.12%, +5.76pp). Their retail/medical-goods provider shares are notably below the EU average (Sweden 10.49%, −5.75pp; Switzerland 9.39%, −6.85pp). This archetype shows that high long-term care shares are compatible with frontier status too, what varies is the balancing across other components of the spending mix.

Seen against these archetypes, Ireland looks like a hybrid that does not align neatly with any single frontier pattern. Ireland is hospital-heavy relative to the EU (+3.11pp), which is common among frontier peers, but it is also LTC-heavy on both the provider and functional definitions (residential LTC +6.72pp; LTC(health) +4.50pp), echoing elements of the Sweden/Switzerland pattern. Where Ireland differs most sharply from many frontier configurations is that it combines these two “heavy” components with below-EU ambulatory provision (−4.59pp) and below-EU retail/medical-goods provision (−4.67pp), alongside a materially lower medical goods functional share (−6.01pp).

Descriptively, the frontier set suggests several “routes” to frontier performance, hospital-heavy with low LTC, hospital-heavy with high medical goods/retail, or LTC-heavy with compensating patterns elsewhere, whereas Ireland combines higher hospitals and higher LTC with comparatively smaller ambulatory and retail/goods channels. This helps frame Ireland’s DEA result as “outcome-effective but input-intensive” without implying that inefficiency can be pinned to any single spending line; rather, Ireland’s mix appears structurally distinct from the frontier configurations observed in the SHA data.

6. Sensitivity and robustness analysis

A recognised limitation of standard DEA is its deterministic nature: the efficiency frontier is constructed from the most extreme observations in the sample. As a result, a single outlier or measurement error can shift the frontier and potentially distort the scores of all countries. To ensure that the conclusions drawn for Ireland and its peers are not driven by such effects, the baseline VRS model was subjected to a number of robustness checks.

6.1. Jackknife Resampling

To assess the stability of the efficiency scores and identify potential influential observations, we employed a leave-one-out Jackknife resampling procedure following the approach of Wilson (1995). In this analysis, the DEA model is re-estimated *n* times; in each iteration, one country is removed from the reference set, and the efficiency scores of all other nations are recalculated against this reduced frontier. This tests whether the inefficiency of a country like Ireland is driven by a specific, high-performing peer whose exclusion would radically alter the benchmark.

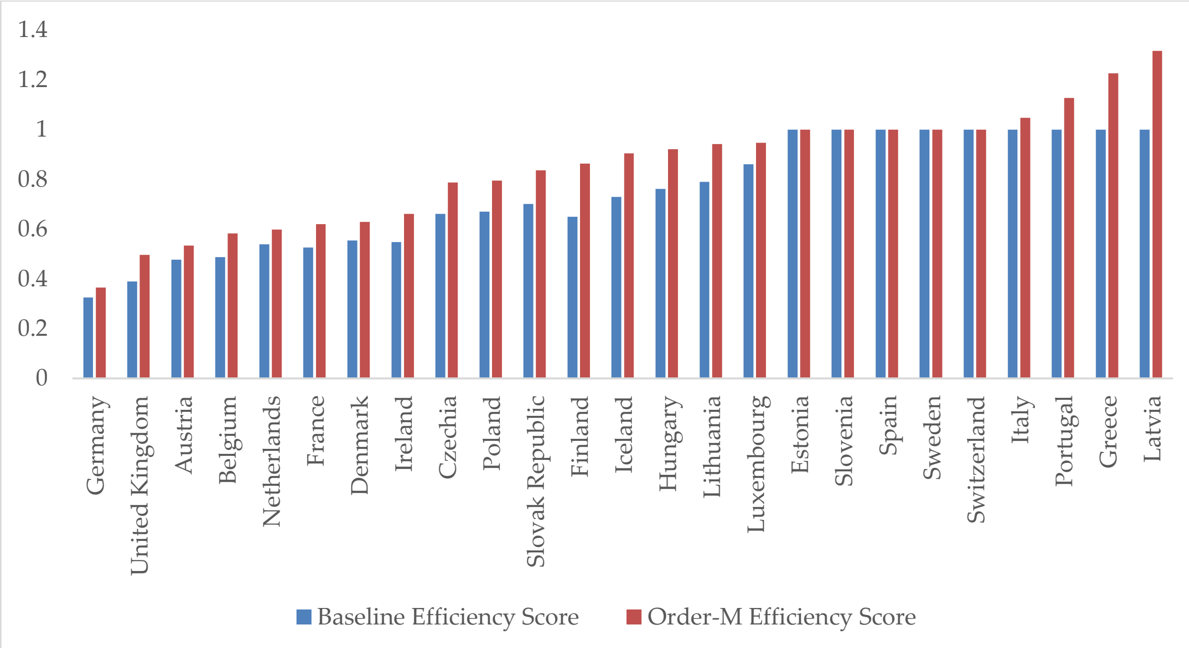
The results indicate a very high degree of stability. The correlation between the baseline VRS scores and the mean jackknife scores is approximately 1.00, implying that country rankings are essentially unchanged under leave-one-out perturbations. Importantly, Ireland's score varies only modestly across iterations: the jackknife mean is 0.550 (very close to the baseline 0.548), with a small standard deviation (0.007) and a narrow range from 0.548 to 0.580. These findings suggest that Ireland's distance from the frontier is not an artefact of a single outlier observation. Rather, it reflects a systematic gap relative to the best-practice envelope defined by the European sample as a whole.

6.2. Order-m Partial Frontiers

To mitigate the sensitivity of standard DEA to outliers and measurement noise, we estimated Order-m efficiency scores. Instead of benchmarking each country against the best observed performers in the full sample, Order-m constructs a "partial" frontier by comparing each unit to the expected performance of a randomly drawn subset of *m* peers (here, m = 15). By design, this reduces the influence of extreme observations and yields efficiency estimates that are less susceptible to atypical data points.

Following the computational procedure described by Tauchmann (2012), we computed Order-m scores using a Monte Carlo approximation. For each country, we drew B = 200 random subsamples of size *m* from the reference set, with replacement, and calculated the corresponding pseudo-efficiency score in each replication. The final Order-m score is then defined as the arithmetic mean across replications. Sampling with replacement is important in this context because it allows the method to approximate the underlying distribution of health-system performance, rather than treating the observed set of 25 countries as a fixed and exhaustive population.

Figure 8. Order-m efficiency scores vs baseline VRS scores



Source: Author's analysis. **Notes**: Order-m Efficiency Score is the arithmetic mean across replications. In some countries, efficiency scores exceed 1, indicating they are 'super-efficient.' This means these countries outperform what the typical m-peer frontier can replicate, making them 'super-efficient' relative to that benchmark.

As expected (see Figure 8), Order-m estimates are generally higher than the baseline VRS scores, reflecting the less stringent nature of a probabilistic benchmark. Nevertheless, the substantive conclusion for Ireland remains unchanged. Ireland's Order-m score is 0.661, higher than its baseline VRS score (0.548) but still well below the benchmark implied by the partial frontier.

Taken together, the consistency across the baseline model, jackknife resampling, and Order-m benchmarking strengthens confidence in the core empirical conclusion: Ireland's measured shortfall persists across alternative frontier constructions and is not driven primarily by peer selection, extreme observations, or sampling variation.

7. Determinants of inefficiency: second-stage analysis

DEA quantifies the size of each country's efficiency gap, but it does not explain why that gap arises. To investigate potential drivers, we conducted a second-stage analysis using the Simar–Wilson (2007) double-bootstrap truncated regression, which is specifically designed for DEA-based efficiency measures and corrects for serial correlation and bias in the first-stage estimates. The dependent variable is the bias-corrected inefficiency score, defined as:

$$\delta_k = \frac{1}{\theta_k^{BC}} - 1,$$

where higher values indicate greater inefficiency. The second-stage estimates are associational and should not be interpreted as causal effects. They describe how measured inefficiency co-varies with country characteristics in this cross-section, not how inefficiency would change if a covariate were manipulated.

Table 2 reports the estimated coefficients. In addition to the obesity rate, the model includes both a linear and quadratic term for GDP per capita (PPP, in $1,000s) to test whether the relationship between national income and measured inefficiency is non-linear. Given the limited cross-country sample to reduce heterogeneity (n = 25), the specification is intentionally parsimonious to reduce the risk of overfitting and unstable estimates; adding many additional covariates would quickly exhaust degrees of freedom and weaken inference in a small-sample setting.[5]

[5] As an additional environmental control, we augment the Simar–Wilson (2007) double-bootstrap truncated regression with the population share aged 65+ (percentage points). The estimated association is positive and statistically significant ($\beta = 0.218$, bootstrap SE = 0.104; 95% percentile CI [0.066, 0.467]; $p = 0.003$), indicating that, conditional on obesity prevalence and the non-linear GDP terms, countries with older populations tend to exhibit higher measured inefficiency in this cross-section. This result is interpreted as an association rather than a causal effect.

Table 2: Determinants of Health System Inefficiency (Simar-Wilson Double Bootstrap)

Independent Variable	Coefficient	Std. Error	CI_2.5	CI_97.5	z-statistic	Bootstrap p-value
(Intercept)	-24.98***	11.27	-49.66	-9.57	-2.22	<0.001
Obesity Rate	0.25***	0.11	0.11	0.51	2.25	<0.001
GDP per Capita ('000s)	0.43***	0.19	0.17	0.85	2.30	<0.001
GDP per Capita Sq ('000s)	-0.002***	0.001	-0.004	-0.001	-2.25	<0.001

Source: Author's analysis. **Notes**: The dependent variable is the bias-corrected inefficiency score (δ), where higher values indicate greater inefficiency. Coefficients for GDP per capita are expressed in units of $1,000 USD (PPP); thus, the estimates reflect the marginal effect of a $1,000 increase in GDP per capita. Results are associational, not causal; specification is parsimonious given the small sample (n = 25). Statistical significance is indicated by ∗p<0.1; ∗∗p<0.05; ∗∗∗p<0.01. P-values are derived from the empirical distribution of 2,000 bootstrap replications. Bootstrap p-values are computed as twice the smaller tail probability of the coefficient's sign across bootstrap replications; p=0 indicates no sign changes in 2,000 draws and should be read as p<0.001. σ (error SD) = 1.11 (SE 0.30)

Two findings stand out. First, obesity is strongly and positively associated with inefficiency. Second, the relationship between GDP per capita and inefficiency is non-linear and follows an inverted-U pattern.

Obesity is positively and statistically significantly associated with inefficiency (β = 0.254). Holding income terms constant, a one-percentage-point increase in the obesity rate is associated with a 0.254-unit increase in δ. This pattern is consistent with the possibility that a higher chronic-disease burden increases the volume and complexity of care required to achieve a given level of population health outcomes, which may worsen measured efficiency when outcome gains (e.g., longevity improvements) are incremental.

GDP per capita is positive in the linear term (β = 0.433) and negative in the quadratic term (β = −0.002), implying an inverted-U relationship between income and inefficiency. For most countries in the sample, higher income is associated with higher inefficiency, consistent with a cost-pressure channel: as economies become richer, health-sector input prices (especially wages, but also administrative and technology-related costs) tend to rise. At very high levels of income, however, the negative quadratic term suggests that further increases in wealth begin to offset these pressures, potentially through superior capacity for innovation, diffusion of advanced technologies, and institutional capability.

The implied turning point can be approximated at $GDP^{*} \approx \frac{-\beta_1}{2\beta_2} \approx 96.8$ (i.e., about $97,000 PPP per capita). This indicates that inefficiency increases with income up to roughly this level, after which the marginal association reverses. Because many European countries are below this threshold, the estimated relationship implies that much of the sample lies on the upward-sloping portion of the curve.

These results help refine the interpretation of Ireland's first-stage DEA performance in two ways. First, they underscore the role of the broader cost environment. Ireland's GDP per capita is exceptionally high (about $134k PPP in the dataset), placing it above the estimated turning point. This is consistent with strong input-price pressures (particularly wages) that can make frontier-level efficiency harder to achieve when incremental outcome improvements are limited. At the same time, because Ireland lies beyond the turning point, the fitted relationship suggests that income effects may not continue to worsen measured inefficiency indefinitely at very high-income levels, reinforcing the need for caution in attributing Ireland's gap to income alone.[6]

Second, obesity represents an additional headwind that is independent of income. Even after controlling for GDP per capita (and its non-linearity), higher obesity prevalence is associated with worse measured efficiency. For Ireland (obesity rate 21.3% in the dataset), this suggests that population health risk factors may compound cost pressures by increasing care needs and complexity, potentially widening the observed efficiency gap unless prevention and chronic-disease management improve in parallel.

8. Conclusion

This study compares European health systems using a benchmarking method (DEA) that looks at how much health outcome is achieved for a given level of public spending. We use public health expenditure per person as the input, and three outcomes together, life expectancy, infant mortality, and avoidable mortality, to assess spending efficiency.

Across the models, Ireland comes out as high performing on outcomes but relatively expensive compared with the best-performing peers in the sample. In the main (input-oriented, VRS) model, Ireland's score is $\theta = 0.548$, which corresponds to an indicative gap to the benchmark frontier of about 45%. It is

[6] Ireland's GDP per capita is upward-distorted by globalisation and multinational-related effects, so it may overstate the income level most relevant for domestic cost conditions; accordingly, the Central Statistics Office publishes modified GNI (GNI*) to net out major globalisation distortions (including depreciation on relocated IP and leased aircraft, and redomiciled PLC income). In 2023, GNI* per capita was €55,084 (CSO, 2025), well below GDP per capita, implying that using GNI*-based income would materially reduce Ireland's measured per-capita income and, depending on the PPP conversion used in the second stage, would likely move Ireland closer to or below the estimated turning point at which higher income is associated with rising inefficiency.

important to stress what this number means: it is a model-based distance to the best observed practice in the sample, not a realistic "savings target".

When we split the overall gap into "system size" (scale) versus "how efficiently the system runs at its current size," Ireland appears close to the right scale. Ireland's scale efficiency is 0.911, suggesting the larger part of the gap relates to within-system performance and costs, rather than the system being fundamentally too large or too small.

The output-oriented model asks a different question: how much could outcomes improve at current spending? For the baseline, life-expectancy-inclusive outcome bundle, Ireland's output-oriented score is $\phi = 1.015$, which implies a proportional improvement of around 1.5% in the combined outcome bundle in the model's terms. The output-oriented results therefore indicate limited proportional headroom for this longevity-inclusive specification. However, when focusing on mortality outcomes alone, the frontier implies materially greater scope for improvement at existing spending.

The results for Ireland are stable under alternative frontier checks. The jackknife exercise shows baseline scores and jackknife means are almost perfectly aligned ($r \approx 1.00$), and Ireland's score varies only modestly across iterations (0.548 to 0.580). Using a less extreme "partial frontier" (Order-m), Ireland's score increases to 0.661, but it still remains clearly below the benchmark.

A second-stage Simar–Wilson regression suggests two patterns: higher obesity prevalence is associated with higher measured inefficiency, and the association between GDP per capita and inefficiency is non-linear (inverted-U), with an implied turning point around $97,000 PPP per capita. Because the sample is small (n = 25), the model is intentionally parsimonious, and these relationships should be read as correlations, not as causal effects.

In summary, Ireland looks “outcome-strong but input-intensive” in this cross-country benchmarking framework, and this finding is robust to alternative frontier checks. While output-oriented headroom is limited for the baseline longevity-inclusive outcome bundle, a mortality-focused specification indicates greater potential for improvement, and the remaining explanations require more detailed, system-level investigation beyond DEA.

9. Limitations and future directions

While the framework applied in this paper yields useful insights into the comparative performance of European health systems, several methodological limitations should be noted to avoid over-interpretation.

First, DEA provides a measure of relative, not absolute, efficiency, and the estimates here are based on a cross-sectional snapshot of 2023 data. Scores are defined by the best performers within the sample at that point in time: a score of 1 indicates that no other country in this dataset achieves better outcomes for a given level of spending on the observed dimensions, not that the system is “perfect” or operating at a global optimum. If the entire comparison set were collectively below international best practice, DEA would not identify that shared shortfall. Moreover, as a static approach, the analysis may not capture the effects of very recent reforms, time-lagged impacts of policy changes, or longer-run cyclical dynamics in health system performance.

Second, the DEA models use public health expenditure per capita as the single input. As a result, the estimated efficiency scores should be interpreted as expenditure efficiency (or cost efficiency) rather than "pure" technical efficiency in physical units. Differences in unit costs, particularly labour costs, may therefore contribute to measured inefficiency alongside differences in service volumes and operational performance. This is not a weakness for the policy question addressed here (value-for-money in public spending), but it does mean that the frontier captures a composite of price and volume effects. In this sense, higher input prices are treated as part of the cost of producing outcomes and therefore remain relevant to a value-for-money assessment from a public finance perspective. Although PPP conversion and the second-stage controls for GDP per capita help to address cross-country cost differences, the first-stage frontier cannot fully disentangle price effects from volume effects. Future research could incorporate physical input indicators (e.g., staffing ratios, bed supply, or activity volumes) as additional inputs in a multi-input DEA specification, allowing a clearer separation between price-driven and volume-driven sources of measured inefficiency.

Third, while the Simar–Wilson double-bootstrap procedure improves statistical validity relative to naïve second-stage regressions, the estimated relationships remain associative rather than causal. For example, the positive association between obesity prevalence and inefficiency should not be interpreted as proof that obesity causes organisational inefficiency. Unobserved heterogeneity, such as health behaviours, cultural norms, the structure of private provision, or underlying socio-economic determinants, may influence both risk-factor prevalence and measured system performance.

Finally, the estimated frontier is necessarily sensitive to the choice of outputs. This study prioritises “hard” mortality outcomes (life expectancy, infant mortality, and avoidable mortality) because they are broadly comparable across countries and less prone to measurement inconsistency. However, this definition of performance is incomplete. It does not capture process and experience measures, such as waiting times, access equity, continuity of care, or patient-reported outcomes, dimensions that are central to how citizens experience the health system. A system could therefore appear efficient in preventing mortality while still performing poorly in timely access or service quality. As data availability improves, future work should incorporate patient-reported outcome measures, waiting-time indicators, and distributional measures of access to ensure that measured efficiency gains are not achieved in theory at the expense of timeliness, equity, or patient experience.

More broadly, the methodological approach demonstrated here can be extended to other major areas of public expenditure, such as education, social protection, long-term care, or climate adaptation, where governments face similar pressures to deliver better outcomes within constrained fiscal space. Applying comparable frontier methods across policy domains could help identify where public resources generate the highest marginal returns across the wider welfare state, supporting a more coherent and evidence-based approach to fiscal efficiency beyond the health sector.

References

Aydın, A. (2022) 'Benchmarking health care systems of OECD countries: A DEA–based Malmquist Productivity Index Approach', Alphanumeric Journal, 10(1), pp. 25–40. doi:10.17093/alphanumeric.1057559.

Cantor, V.J.M. and Poh, K.L. (2018) 'Integrated analysis of health care efficiency: A systematic review', Journal of Medical Systems, 42, Article 8. doi:10.1007/s10916-017-0848-7.

Central Statistics Office (CSO) (2025) 'Economy', Measuring Ireland's Progress 2023. CSO Statistical Release, 4 April. Available at: https://www.cso.ie/en/releasesandpublications/ep/p-mip/measuringirelandsprogress2023/economy/ (Accessed: 24 February 2026).

Charnes, A., Cooper, W.W. and Rhodes, E. (1978) 'Measuring the efficiency of decision making units', European Journal of Operational Research, 2(6), pp. 429–444. doi:10.1016/0377-2217(78)90138-8.

Dutu, R. and Sicari, P. (2020) 'Public spending efficiency in the OECD: Benchmarking health care, education, and general administration', Review of Economic Perspectives, 20(3), pp. 253–280. doi:10.2478/revecp-2020-0013.

Eurostat (2025) Healthcare expenditure statistics by function, provider and financing scheme (Statistics Explained). European Commission. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php/Healthcare_expenditure_statistics_by_function,_provider_and_financing_scheme (Accessed: 3 February 2026).

Gavurova, B., Kocisova, K. and Sopko, J. (2021) 'Health system efficiency in OECD countries: Dynamic network DEA approach', Health Economics Review, 11(1), Article 40, pp. 1–25. doi:10.1186/s13561-021-00337-9.

Jung, S., Son, J., Kim, C. and Chung, K. (2023) 'Efficiency measurement using data envelopment analysis (DEA) in public health care: Research trends from 2017 to 2022', Processes, 11(3), Article 811. doi:10.3390/pr11030811.

Meaney, K., Oyewole, V. and Bedogni, J. (2018) Comparative levels and efficiency of Irish public spending. Available at: https://assets.gov.ie/7322/262f606010344ca3adeba2c88087add6.pdf (Accessed: 23 December 2025).

OECD (2025) OECD Health Statistics 2025. Available at: https://www.oecd.org/en/data/datasets/oecd-health-statistics.html (Accessed: 22 December 2025).

Rostamzadeh, R., Akbarian, O., Banaitis, A. and Soltani, Z. (2021) 'Application of DEA in benchmarking: A systematic literature review from 2003–2020', Technological and Economic Development of Economy, 27(1), pp. 175–222. doi:10.3846/tede.2021.13406.

Simar, L. and Wilson, P.W. (2007) 'Estimation and inference in two-stage, semi-parametric models of production processes', Journal of Econometrics, 136(1), pp. 31–64. doi:10.1016/j.jeconom.2005.07.009.

Sicari, P. and Sutherland, D. (2023) Health sector performance and efficiency in Ireland. OECD Economics Department Working Papers, No. 1750. Paris: OECD Publishing. Available at: https://doi.org/10.1787/6a000bf1-en.

Tauchmann, H. (2012) 'Partial frontier efficiency analysis', The Stata Journal, 12(3), pp. 461–478. doi:10.1177/1536867X1201200309.

Wilson, P.W. (1995) 'Detecting influential observations in data envelopment analysis', Journal of Productivity Analysis, 6(1), pp. 27–45. doi:10.1007/BF01073493.