Does Marginal Healthcare Funding Reduce Avoidable Mortality and Improve Wellbeing? Evidence from the English NHS

Wing Lin

Department of Economics, University of Warwick

February 2026

Abstract

Estimating the short-run effects of public healthcare funding on health outcomes and wellbeing is particularly relevant for policymaking under budget cycles and rigid capacity. However, an empirical challenge arises: funding is endogenous due to the continuous targeting of resources toward evolving health needs. I employ a simulated-instruments approach to address this challenge in panel settings, estimating the short-run causal effects of funding allocations in the English National Health Service (NHS). I construct the instrument, "simulated allocations", a counterfactual set of allocations to local areas under fixed local characteristics, isolating variation driven solely by national rule changes from local endogenous responses. I find that an £1,000/head increase in allocation (mean = £1,270/head) reduces preventable mortality by 101.2 per 100,000 population, with no significant effect on treatable mortality. Beyond survival, I find suggestive evidence that funding improves subjective wellbeing through increases in happiness, life satisfaction, and reduction in anxiety. This allows for a wellbeing cost-benefit analysis, which yields a marginal cost of approximately £1,282/WELLBY (one-unit increase in life satisfaction per person-year). This estimate is significantly below the UK government valuation benchmark of £13,000/WELLBY, indicating that marginal expansions in healthcare funding are likely welfare-enhancing.

JEL codes: H51, H75, I10, I18, I31.

Acknowledgements: I am grateful to Ludovica Gazze for their guidance and support. I thank Sascha O. Becker, James Fenske, Thijs van Rens, Lazaros Andronis, Nathan Canen, Matthew Ridley for helpful comments and suggestions. All errors are my own.

Correspondence: Department of Economics, University of Warwick, Coventry, CV4 7AL, UK. Email: wing.lin@warwick.ac.uk

1. Introduction

Healthcare spending accounts for over 10% of GDP across OECD countries, a share projected to rise as populations age and costs of medical technology increase (Lloyd-Sherlock, 2000; de Meijer et al., 2013; World Health Organization, 2025). Policymakers now face increasing difficulty in allocating finite resources for healthcare under rising needs and intensifying fiscal pressures. The National Health Service (NHS) in the United Kingdom exemplifies such challenges as a public, tax-funded, centralised healthcare system free at point of delivery. As politicians have to justify healthcare funding decisions within budgetary and electoral cycles, estimating the marginal returns to healthcare resources on health outcomes and welfare is crucial to determining the optimal funding level. Despite extensive literature studying the long-run structural equilibrium spending returns[1], there is limited evidence on the short-run causal impacts of healthcare funding. These short-run estimates align more closely with actual decision-making, as policymakers manage healthcare funding to maximise welfare within the constraints of annual budgets and rigid organisational capacities.

An empirical challenge arises when estimating short-run effects, which requires panel settings to isolate "within-area" variation[2]. Funding is not only confounded by health differences across areas, but also further confounded by time-varying health needs within areas. This confounding leads to limitations in existing approaches in panel settings. Instrumental variable strategies often suffer from dynamic endogeneity concerns, weak instrument bias, or lack of clear interpretation (Nakamura et al., 2016; Roodman, 2009). Quasi-experimental studies offer internal validity (Doyle, 2011; Finkelstein et al., 2012) but rely on one-off events or shocks that can be difficult to generalise.

This paper addresses such challenges and provides new evidence on the short-run causal effects of healthcare funding on health outcomes and wellbeing by adapting a simulated-instruments approach, as in other fields of economic analysis (Gruber and Saez, 2002; Biasi, 2023). I study NHS England's funding allocation rules to Clinical Commissioning Groups (CCGs) between 2013/14 and 2019/20, using funding allocation per head as the treatment variable. This represents an intent-to-treat (ITT) framework on healthcare policy, as it evaluates the impact of the funding determined by central planners rather than the demand-led local spending (Angrist and Pischke, 2009). To overcome funding endogeneity, I construct the simulated instrument, "simulated allocations", a

[1] For example, existing English NHS studies estimate marginal productivity primarily in cross-sectional settings (Martin et al., 2008a,b,c; Claxton et al., 2015; Andrews et al., 2017; Lomas et al., 2019; Martin et al., 2021a,b, 2022, 2023). Cross-sectional estimators rely on "between" variation to capture the long run structural returns. They assume areas are in long-run equilibrium, where areas have adjusted their capital stock and infrastructure to their respective funding levels (Baltagi and Griffin, 1984; Kuh, 1959).

[2] Panel estimators focus on "within" variation to capture short-run response to changes. It holds long-run structural factors constant with fixed effects or first differences. In contrast,

counterfactual set of allocations that CCGs would have received if their demographics and needs had remained fixed at base year[3] levels. The simulated allocation varies over time only through national allocation rule changes and total budget variations, which are determined exogenous to local health needs or outcomes. The instrument isolates exogenous variation in funding from local endogenous responses, allowing for causal identification of the local average treatment effect (LATE). I estimate effects on avoidable mortality (preventable and treatable conditions) and subjective wellbeing measures to provide both health-specific and welfare-based evaluations of marginal healthcare funding returns.

My results show that marginal increases in funding allocations have a short-run effect[4] on reducing overall avoidable mortality. While OLS estimates suggest increased allocations reduces both preventable and treatable mortality, simulated-IV results reveal a distinct difference in effects across the types: effects on preventable mortality are significantly larger than treatable mortality. The estimated causal effect is that a £1,000/head marginal increase in allocation (mean = £1,270/head) reduces preventable mortality by 101.2 per 100,000 population on 3-year rolling averages, while having no statistically significant effect on treatable mortality. This distinction links to: (1) the mechanical prioritisation of preventable over treatable conditions in the allocation rules, (2) fundamental differences in marginal returns between preventive and curative care, (3) shifting policy focus toward prevention, and (4) asymmetric payment mechanisms channelling funding growth toward preventative services.

Beyond survival, I find suggestive evidence for wellbeing improvements. A £1,000/head allocation reduces anxiety contemporaneously by 1.61 and generates lagged improvements in happiness by 0.93, life satisfaction by 0.79, and overall wellbeing by 0.85. Although individual measures show limited precision at conventional significance levels, consistency across specifications supports a positive causal relationship. IV estimates exceed the OLS magnitudes as predicted, with wider intervals reflecting the efficiency-consistency trade-off (Wooldridge, 2020) and the reduction in sample size from instrument construction. High IV first-stage F-statistics rule out weak identification, indicating results suggest healthcare funding improves wellbeing from lagged increases in happiness and life satisfaction, alongside with immediate anxiety reduction.

I go beyond traditional health valuation frameworks like quality-Adjusted life year (QALY) (Luce et al., 2006; Lomas et al., 2019), and conduct a cost-benefit analysis for healthcare funding in

[3]This can be considered baseline levels of demographics and needs. The word choice of "base year" prevents confusion with "baselines" in allocation formulae, which represents the previous year's allocation levels.

[4]I characterise this as "short run" estimates under panel settings with fixed effects or first differences, as structural determinants are absorbed, my estimates are consistent with a short-run interpretation. While the 3-year rolling averages introduces a temporal structure longer than a single fiscal year, it remains distinct from long-run equilibrium estimates derived from cross-sectional variation where full structural adjustment is assumed.

wellbeing terms. Wellbeing analysis is an experience-based valuation of health state, which is more suitable compared to existing preference-based methods, such as QALY (Dolan, 2008). I estimate the marginal cost per WELLBY of the NHS, which measures the funding required to generate a one-unit increase in life satisfaction (on a 0–10 scale) for one person for one year (Frijters and Krekel, 2021; Frijters et al., 2024). The estimated marginal cost per WELLBY at £1,282/WELLBY lies significantly below societal valuation benchmark of £13,000/WELLBY proposed by the HM Treasury (2021). This indicates that additional NHS funding can potentially generate substantial social welfare benefits.

This paper complements a broader literature of evaluations on the impact of healthcare spending on outcomes and marginal productivity of healthcare funding. Specifically, this paper contributes to the literature on healthcare economics in four ways.

First, I adapt the simulated instrument methodology to a healthcare funding setting, particularly for short-run causal estimation for publicly funded healthcare systems under panel settings. Previous cross-sectional studies on the marginal productivity and cost effectiveness of the NHS rely on external instruments (Martin et al., 2008a,b,c; Claxton et al., 2015), which might raise concerns on weak instruments or instrument endogeneity. Under the innovation of instrumenting expenditure with allocation rule components (Andrews et al., 2017), these cross-sectional studies have shifted to using "allocation rule" instruments (Lomas et al., 2019; Martin et al., 2021a,b, 2022, 2023). However, these instruments become endogenous in panel settings, as they track evolving local needs. In this paper, I address this endogeneity by using simulated instruments, constructing counterfactual allocations based on fixed baseline demographics and needs. The instrument isolates variation driven solely by national policy changes, which is independent of local shocks. My empirical strategy accounts for endogeneity, addresses existing limitations in cross-sectional studies, and allows for short-run causal inference for marginal returns in rule-based healthcare funding under panel settings.

Second, I provide new estimates of the short-run returns on healthcare funding and descriptively compare them with existing long-run spending estimates in the literature. Existing UK evidence, relying on cross-sectional variation, estimates that the long-run NHS spending elasticity of population mortality ranges from -0.5 to -1.1 (Martin et al., 2008a; Claxton et al., 2015; Martin et al., 2021a,b, 2023). By converting my estimates into elasticities for population mortality at sample mean, my lower bound estimate[5] for short-run funding elasticity (-0.13) is significantly smaller than their long-run spending counterparts. The divergence in results relates to treatment variable choice and time horizons. On one hand, ITT elasticity is naturally smaller than spending elastic-

[5]This lower bound estimate (unrealistically) assumes the causal impact of funding on total deaths is driven entirely by avoidable deaths, which yields a likely lower bound estimate. This will be discussed in section 6.1.

ity in absolute terms, since allocations must filter through local decision-making before reaching actual health services. On the other hand, the distinction between short-run response (panel) and long-run equilibrium returns (cross-sectional) reflects the short-run constraints on fixed factors in the NHS, such as hospital infrastructure and its workforce. Such differences in results imply that using long-run structural elasticities for annual performance targets risks overstating returns within budget cycles, which demonstrates the relevance of using short-run funding responses when devising or interpreting annual targets.

Third, I add evidence for higher marginal returns to preventative than curative care. For comparison, I conduct a descriptive meta-analysis on existing literature by mapping existing Programme Budget Category (PBC) elasticities into avoidable mortality types. Within the study period (2013/14–2019/20), my results align with other cross-sectional NHS studies in observing such differences in returns (Martin et al., 2022, 2023). As both long-run and short-run estimates reflects such differences, they collectively suggest that the fundamental marginal productivity of preventative care (e.g. community health services or mental health services) may be greater than curative care (e.g. elective hospital care or urgent care). This is consistent with the "flat of the curve" hypothesis for curative care (Fisher et al., 2000, 2003; Thornton, 2002, 2011; Hensher et al., 2024; Finkelstein and Gentzkow, 2026) and the higher marginal returns to preventive care documented in the literature (Mays and Smith, 2011; Bradley et al., 2016; Masters et al., 2017).

Finally, I expand the measurement of healthcare return beyond survival to consider subjective wellbeing, a more holistic but less studied approach. Subjective wellbeing can capture broader healthcare benefits beyond mortality, such as health state and care quality (Dolan, 2008; Lee et al., 2013; LaVela and Gallan, 2014; Brazier and Tsuchiya, 2015). Wellbeing estimates also allow cost-benefit analysis in wellbeing terms, aligning with recent guidance from Frijters and Krekel (2021) and the HM Treasury (2021, 2022). My results verify Rutten and Reed's (2009) computable general equilibrium model predictions that additional NHS resources would yield a net societal welfare gain, which could justify further investment in the NHS.

The remainder of the paper is organised as follows. Section 2 describes the institutional background of NHS England, CCG commissioning, and allocation rules. Section 3 details the data used. Section 4 presents the conceptual framework for healthcare spending and outcomes. Section 5 explains empirical strategy, endogeneity issues, instrument construction and presents the results. Section 6 interprets and discusses the implications of my results with existing literature. Section 7 concludes.

2. Institutional Background

2.1. National Health Service (NHS) and Commissioning

Established in 1948, the National Health Service (NHS) in the United Kingdom is one of the largest publicly funded universal healthcare systems in the world. The NHS founding principles are to provide healthcare services based on needs, not the ability to pay, and free at point of delivery for all UK citizens (Clement, 2023). To provide such services, the NHS continuously plans, agrees, and monitors its services through NHS commissioning, a process that funds and ensures NHS services are provided effectively to meet the needs of the population. The commissioning process also supports the delivery of policy objectives in the NHS, such as improving quality of services, value for money, and health outcomes.

The funding of healthcare services is a core function of commissioning, consisting of three main stages since 2012. First, the Department of Health and Social Care receives a budget from the UK government to provide health services in England. Second, the majority of the funding is transferred to NHS England (NHSE) to distribute locally. NHSE is a strategic leadership and commissioning authority, established to direct and oversee commissioning of NHS services on a regional or national level[6]. It uses allocation rules (see section 2.3) to allocate the funding to Clinical Commissioning Groups (CCGs)[7]. CCGs are regional statutory bodies with the duty to plan and commission a majority of NHS services for people in their defined geographic area (Harker, 2019). Third, CCGs commission providers, such as NHS Foundation Trusts, NHS Trusts, and private providers, to provide these services. Commissioners and providers use NHS standard contracts to set the price of the services provided under the national tariff payment system, block contracts, or an approved payment approach (NHS England, 2020).

2.2. CCG Commissioning Duties

Clinical Commissioning Groups (CCGs) are responsible for commissioning healthcare services to meet the reasonable needs of patients who are registered with their member practices or live within their defined geographic area. CCGs are led by a Governing Body made up of GPs and other clinicians to implement clinical-led decision making (National Audit Office, 2018). The health services that CCGs commission are listed in Table 1 (Department of Health, 2012).

[6]Established in 2013 as part of the Health and Social Care Act (2012)

[7]Clinical Commissioning Groups (CCGs) were established in April 2013, replacing Primary Care Trusts (PCTs) as the statutory bodies responsible for commissioning. Subsequently, CCGs were replaced by Integrated Care Boards (ICBs) in July 2022. As this paper focuses on the period 2014-15 to 2019-20, when CCGs were the active statutory bodies, the later transition to ICBs is excluded from this analysis.

Table 1: List of Health Services Commissioned by CCGs (CCG services)

	Health Service
1	Community health services
2	Maternity services
3	Elective hospital care
4	Rehabilitation services
5	Urgent and emergency care including A&E, ambulance and out-of-hours services
6	Older people's healthcare services
7	Healthcare services for children
8	Healthcare services for people with mental health conditions
9	Healthcare services for people with learning disabilities
10	Continuing healthcare
11	Abortion services
12	Infertility services
13	Wheelchair services
14	Home oxygen services
15	Treatment of infectious diseases

Source: Department of Health (2012)

In this paper, NHS funding only refers to CCG's funding, which are based on their allocations for commissioning CCG responsible services (Table 1) and administration. It excludes funding on primary care or specialised services, which is a limitation arising from ensuring instrument and data consistency and stability (as discussed in Appendix B).

2.3. CCG Allocation Rules

NHSE allocates funding to every CCG with a four-step rule-based funding rule (formula), established on the principle of equal opportunity of access for equal need (NHS England, 2025)[8].

1. Baselines are carried forward from previous year's allocations with adjustments.
2. Target allocations are determined based on relative need and unavoidable costs using the weighted capitation formulae. For all CCG service types (e.g. mental health and maternity), service-specific weighted population are calculated and aggregated. After applying adjustments, the overall weighted population is obtained to reflect CCG's population, needs for healthcare services related to age, needs exceeding those predicted solely by age, unmet

[8]NHSE publishes actual funding allocations publicly for the financial year to CCG on its website. In each major allocation round, NHSE also publishes a technical document that explains how actual allocations were set, under a set of rules and calculations (NHS England, 2025).

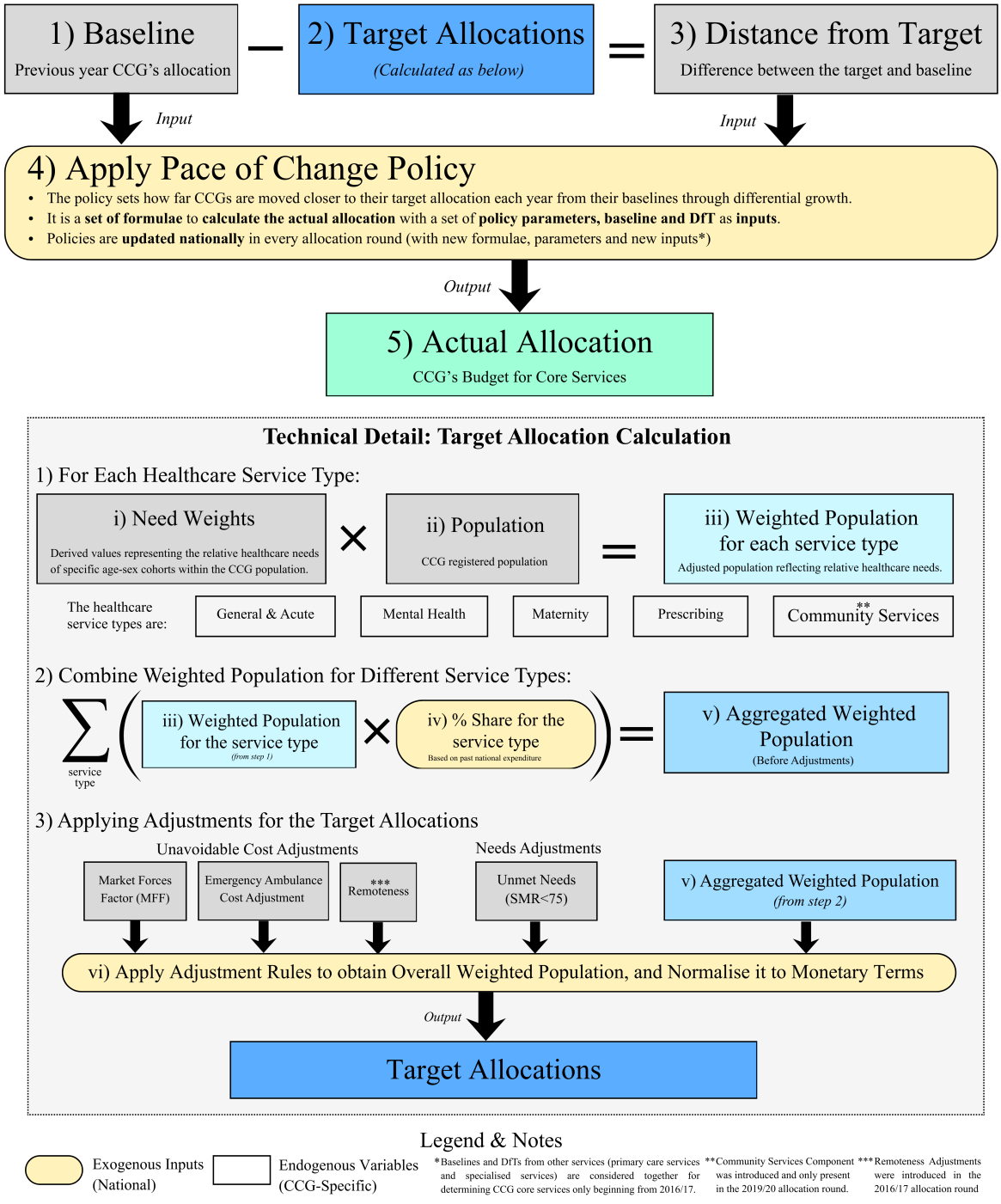


Figure 1: Overview of the CCG Allocation Process

needs, health inequalities, unavoidable higher costs for that service. To calculate each CCG's monetary target allocations, CCG's share of the overall weighted population for England is multiplied by the national budget.

3. Distance from target is calculated by the difference between baseline and the target allocation.
4. Pace of change policy is applied to determine the actual allocation. Pace of change policy sets how far CCGs are moved closer to their target allocation each year from their baselines through differential growth, based on its opening distance from target. It aims to balance the need for funding stability and the objective of moving furthest under-target CCGs closer to their target.

Figure 1 and Equation 1 summarise how actual allocations for CCG in period t are calculated based on allocation rules R_t. Appendix A presents a detailed breakdown of the allocation rules and their changes over time.

$$Allocation_{CCG,t} = R_t\big(NationalBudget_t, Baseline_{CCG,t}, Population_{CCG,t}, Needs_{CCG,t}, Adjustments_{CCG,t}\big) \tag{1}$$

Over time t, there are several changes in the allocation rules R_t. Formulae, weights and models for different needs and costs adjustments have been refreshed or introduced over time. Pace of change policies are updated in main allocation rounds. Estimation for CCG's population, unmet need, and health inequalities adjustments are refined. These allocation rule changes provide a source of exogenous variation in the funding received by CCGs, allowing for the construction of the simulated instrument.

3. Data

3.1. Allocations to CCG as the Intent-to-Treat Policy Variable

To estimate the effect of increased healthcare funding, I use CCG funding allocations for the treatment variable to represent the policy lever, instead of using actual expenditure. This is because the UK government and NHSE control the assignment of the allocations, but not the precise expenditure. The overall NHS budget and allocations to CCGs are determined ex ante by the UK government and NHSE respectively, based on annual budgets, spending reviews, weighted capitation formulae and historical baselines. In constrast, actual spending is demand-led, since the NHS is free at point of delivery (Department of Health and Social Care, 2012). Driven by the obligation to meet concurrent patient needs, local expenditure can diverge from the predetermined budget, creating a "soft budget constraint" (Kornai, 2009; National Audit Office, 2020). CCGs adjust expenditure in real-time response to unobserved health shocks, local demand spikes, or deficits (Gainsbury, 2016; National Audit Office, 2020). Therefore, I evaluate the policy lever

directly manipulable by central planners (i.e. estimating the return on allocating a marginal pound to a local health system), which is often more relevant than the structural return on realised spending (Angrist and Pischke, 2009).

This paper focuses on funding to CCGs from 2013/14 to 2019/20. This period ensures data consistency by excluding major structural re-organisations and the structural break introduced by the COVID-19 pandemic. To standardise cross-area comparisons, I normalise allocations to CCG by its population size to derive allocation per head, $AllocationPH_{CCGt}$.

3.2. Avoidable Mortality

Avoidable mortality is an indicator that measures premature deaths from certain conditions that should not occur in the presence of timely and effective healthcare. The indicator is widely used as a high-level outcome measure for the general performance, quality, and effectiveness of health systems in terms of prevention and intervention (Nolte and McKee, 2004; OECD, 2023). This paper uses the avoidable mortality panel dataset from the UK's Office for National Statistics (2025a,d). It consists of 3-calendar-year rolling averages of age-standardised avoidable deaths rates per 100,000 population at the local authority level. Such deaths are defined as at age between 0 to 74 and meet any conditions that the International Classification of Diseases considers preventable and treatable (Appendix E).

Avoidable mortality has four characteristics that make it suitable for this analysis: consistency, comparability, comprehensiveness, and variability. A consistent definition and full coverage of death registrations[9] provide reliable data quality for this paper. Calculation of death rates and age-standardisation enable comparability across local areas. Avoidable mortality focuses on healthcare system outcomes specifically, which balances the need for sufficient variation and comprehensive measurement of healthcare system performance. It is more suitable than narrower measures, such as infant, child or maternal mortality, which focuses specifically on children healthcare or maternity services. It also provides larger year-to-year variation compared against broader measures, such has life expectancy, which is heavily influenced by longer-term trends in income and living standards.

Avoidable mortality is further classified into two types, preventable mortality and treatable mortality. Preventable mortality refers to causes of death that can be mainly avoided through effective public health and primary prevention interventions. It considers deaths with conditions that are preventable preceding the incidence of diseases or injury, such as influenza, HIV, lung cancer, and alcohol poisoning. Treatable mortality refers to causes of death that can be mainly

[9]Records of death certificates for avoidable mortality are assigned to the deceased's usual residence and aggregated at a local level. They have 100% coverage, since it is a legal requirement that all deaths are registered (Office for National Statistics, 2025d).

Table 2: Questions for the Four ONS Wellbeing Measures in Annual Population Surveys

Measure	Question
Happiness	Overall, how happy did you feel yesterday, where 0 is "not at all happy" and 10 is "completely happy"?
Worthwhile	Overall, to what extent do you feel that the things you do in your life are worthwhile, where 0 is "not at all worthwhile" and 10 is "completely worthwhile"?
Life satisfaction	Overall, how satisfied are you with your life nowadays, where 0 is "not at all satisfied" and 10 is "completely satisfied"?
Anxiety	On a scale where 0 is "not at all anxious" and 10 is "completely anxious", overall, how anxious did you feel yesterday?

Source: Office for National Statistics (2025i)

avoided through timely and effective healthcare interventions, including secondary prevention and treatment. It considers deaths with conditions that fatality can be avoided after the incident of disease with treatment, such as sepsis, renal failure, epilepsy, and medical misadventures. Avoidable mortality rate for local authority a for every 3-year-period t is the sum of the two types, where $AvoidableMortality_{at} = PreventableMortality_{at} + TreatableMortality_{at}$.

3.3. Subjective Wellbeing

I employ subjective wellbeing as an outcome measure to capture the broader benefits of healthcare beyond mortality (Brazier and Tsuchiya, 2015). Dolan (2008) argues that subjective wellbeing offers a better valuation of health states compared to existing preference-based methods, such as Quality-Adjusted Life Years (QALYs). This is because subjective wellbeing represents an experience-based valuation rather than hypothetical preference. Lee et al. (2013) and LaVela and Gallan (2014) propose that subjective wellbeing is effective for evaluating patient experience and overall care quality. Furthermore, these estimates allow for the incorporation of wellbeing into formal cost-benefit analysis, aligning with recent guidance from Frijters and Krekel (2021) and the HM Treasury (2021, 2022). Estimating healthcare returns with subjective wellbeing provides a more holistic but less studied approach, complementing existing valuations.

To assess individuals' subjective evaluations of their wellbeing and quality of life that they experienced themselves, Office for National Statistics (2025i) collects data on personal wellbeing measures from ONS Annual Population Survey and the Opinions and Lifestyles Survey. They assess wellbeing through the four measures: "Happiness", "Life satisfaction", "Feeling the things done in life are worthwhile", and "Anxiety". Survey participants are asked to respond to the four questions with a score from 0 to 10, where 0 is "not at all" and 10 is "completely" (Table 2). This

Table 3: Data Definition and Sources

Measure	Description
Crime Rate	Recorded offense rate per person by Community Safety Partnership area.
Disposable Income	Log of gross disposable household income per person (£1000/head) at current basic prices by local authorities.
Emissions	Log of total greenhouse gas emissions (tCO2e) per capita by local authorities.
First Mortgage	Number of first-time buyer mortgage sales per 10 dwelling stock by local authorities.
Pupil Absence Rate	Proportion of pupils in state-funded schools that were absent for 10% of academic year by county and unitary authorities.
Unemployment Rate	Proportion of economically active people aged 16 years and over without a job and actively seeking work by local authorities.

Sources: Office for National Statistics (2025e), Office for National Statistics (2025k), Department for Energy Security and Net Zero (2025), Office for National Statistics (2025f), Office for National Statistics (2025h), and Office for National Statistics (2024)

paper uses Office for National Statistics (2020) annual personal wellbeing dataset, which reports the mean scores for each wellbeing measure in financial year t at the local authority level a.

I aggregate the four ONS measures of wellbeing with equal weights to calculate an overall mean score of wellbeing[10]. This follows the findings of O'Donnell and Oswald (2015) findings, which the sampled individuals give all four measures a considerable share of weights for an overall estimate of personal wellbeing. I have used a signed average of the four measures as in Equation 2, such that 0 represents minimum wellbeing and 10 represents the maximum wellbeing across all measures.

$$Wellbeing_{at} = \frac{Happiness_{at} + Satisfaction_{at} + Worthwhile_{at} + (10 - Anxiety_{at})}{4} \tag{2}$$

3.4. Controls

Table 3 lists the set of control variables for time-varying CCG specific confounders of outcomes or wellbeing. I rescale the variables for better representation of results. I apply log transformation to variables that show strong evidence of right-skewness, while I keep other variables as levels (see Appendix F). Dummy variables for missing controls are added to preserve sample size.

[10]The construction of a composite index requires the assumption of cardinality. While I acknowledge the theoretical constraints of treating ordinal Likert data as cardinal (Bond and Lang, 2019), I follow the standard convention in wellbeing economics (Frijters and Krekel, 2021). Previous research suggests that assuming cardinality for 0–10

Table 4: Descriptive Summary Statistics of Key Variables

	Observations	Mean	SD	Min	Max
Wellbeing *(0–10 score)*					
Happiness	2,264	7.519	0.227	6.700	8.350
Feeling worthwhile	2,264	7.875	0.197	7.200	8.700
Life satisfaction	2,265	7.681	0.216	6.750	8.590
Anxiety	2,260	2.873	0.326	1.750	3.950
Overall mean	2,260	7.550	0.194	6.845	8.405
Avoidable Mortality *(Deaths per 100,000, 3-year rolling average)*					
Preventable mortality	1,470	142.97	34.66	78.700	272.000
Treatable mortality	1,470	83.08	15.69	47.100	137.000
Overall avoidable mortality	1,470	226.05	49.36	137.500	403.000
CCG Allocation per head *(£1,000/head)*					
For CCGs	1,438	1.295	0.164	0.912	2.090
Distributed to LADs	2,265	1.270	0.152	0.912	2.090
Controls					
Crime Rate	2,352	0.068	0.029	0.015	0.381
Disposable Income	2,368	20.807	11.608	11.474	211.115
Emissions	2,368	7.320	8.670	2.360	202.060
First Buyer Mortgage	2,360	0.106	0.029	0.024	0.296
Pupil Absence	2,360	0.109	0.017	0.028	0.195
Unemployment Rate	2,302	0.048	0.019	0.017	0.145

Note: The data used in this study is a combined panel dataset, spanning over 7 financial years (financial year 2014-15 to financial year 2019-2020). SD refers to the standard deviation.

3.5. Geographical Mapping

Given that the outcome variables are recorded in local authority districts (LAD), I perform my analysis at the LAD level to ensure boundary consistency and measurement accuracy. To merge the datasets into a panel, CCG allocations are mapped to LAD level[11]. I use Lower-layer Super Output Areas (LSOAs) as a bridge for precise re-aggregation, because LSOAs do not cross any LAD or CCG boundaries. I assign CCG allocations per head to constituent LSOAs, and calculate the total allocation for each LSOA with its population estimates[12]. Then, LSOAs' total allocations are aggregated to LAD level. The allocation per head for LAD a, $AllocationPH_{at}$, is obtained by dividing the aggregated total allocation by its population. Ultimately, this approach yields $AllocationPH_{at}$ based on a population-weighted average of the geographic exposure of their

scales introduce little bias compared to treating them as strictly ordinal (Riedl and Geishecker, 2014).

[11] Mapping accounts for boundary updates (e.g., LAD18 vs. LAD24) and historical CCG mergers. Controls are all mapped or available at LAD24 geographic boundaries, and then mapped to LAD18 if needed.

[12] This ensures no cross-cutting of boundaries. LSOA-level populations are from Office for National Statistics (2025g).

constituent CCGs.

To align datasets across time, I convert calendar-year populations and control variables to financial-year equivalent (for wellbeing) or three-year-rolling averages (for avoidable mortality) with weighted averages[13]. This maximises accuracy by preventing extensive manipulation of the outcome data. Table 4 presents descriptive statistics for the final panel, and Figure 2 illustrates the variation across geography.

4. Conceptual Framework

I present the conceptual framework for how increased funding can lead to improved outcomes and wellbeing. To begin, suppose a CCG receives a funding allocation per head of $AllocationPH_{CCGt}$ for financial year t. Allocation per head is determined by the function for allocation rules for the financial year $R_t(.)$, which is increasing in its inputs, national budget b_t, CCG's relative need n_{CCGt}. Allocations increase with needs, because the weighted capitation formulae allow CCGs to commission similar levels of health services for populations with similar levels of need, and appropriately higher levels of health services for populations with higher levels of need[14].

$$AllocationPH_{CCGt} = R_t(b_t, n_{CCGt})$$

Based on the geographical mapping, I assume the per head funding for different LSOAs are the same across a CCG, since it does not consider variation in needs at a granular level. After aggregating to LAD level, allocation per head for an LAD a at financial year t, $AllocationPH_{at}$, represents the total per head funding by their parent CCG(s).

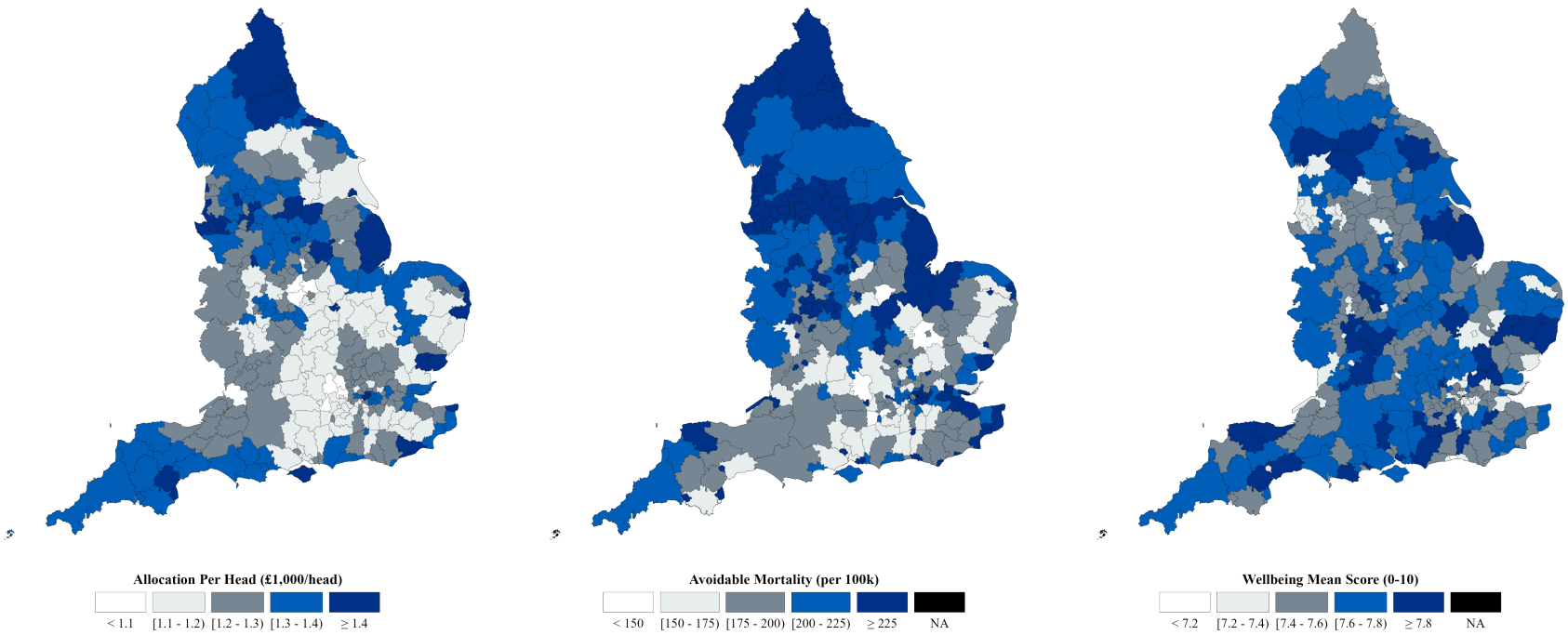
$$AllocationPH_{at} = AllocationPH_{CCGt}$$

I model actual spending as a function $S(.)$ of funding allocation $AllocationPH_{at}$ and idiosyncratic changes μ_{at} in spending due to unanticipated changes in healthcare needs or demand. For simplicity, this framework abstracts from the distribution of spending between different services, treating healthcare spending per head on different services as one variable, $SpendingPH_{at}$.

$$SpendingPH_{at} = S(AllocationPH_{at},\ \mu_{at})$$

[13] For avoidable mortality, 3-year rolling averages were matched directly to the corresponding 3-year average of allocations (e.g., 2014–16 mortality linked to 2014/15–2016/17 allocations). This minimises mismatch error while preserving the lag structure.

[14] This is outlined in the technical guides to allocation, which aligns with NHSE's aim to provide equal opportunity of access for equal need and contribute to the reduction of health inequalities through allocations.



(a) NHS Allocations (b) Avoidable Mortality Rates (c) Wellbeing Scores

Figure 2: Geographical Differences in Healthcare Variables

A CCG spends to commission h_{at} healthcare services for the LAD area per person, where h_{at} reflects both availability and quality of healthcare.

$$h_{at} = H\left(SpendingPH_{at}\right)$$

I assume that increased spending improves the availability or quality of healthcare services, such that $H(.)$ is an increasing function. For example, given an increase in spending, a CCG can commission greater service volume in musculoskeletal therapy services, or an improvement the quality of that service by offering a more comprehensive but costly treatment plan.

Followingly, health outcomes o_{at} (measured with avoidable mortality) for an average individual in the LAD is expressed in a function $O(.)$ of healthcare services h_{at}, fundamental healthcare needs n_{at}, and other non-health factors X_{at} in the LAD area. Non-health factors X_{at} determine outcomes or wellbeing, for example income, education, environment factors. Hence, X_{at} is a vector of control variables for time-varying CCG-specific outcome confounders.

$$o_{at} = O(h_{at},\ n_{at},\ X_{at})$$

I hypothesise health outcomes o_{at} improves when healthcare services h_{at} improves in availability and/or quality, given a certain need n_{at}. Also, given a certain quantity h_{at} and quality q_{at} of healthcare services, higher need n_{at} is expected to result in worsening health outcomes n_{at}.

Finally, I model subjective wellbeing or quality of life w_{at} for an average individual in the LAD as a function $W(.)$ of health outcomes o_{at}, healthcare services h_{at}, and other non-health factors X_{at}.

$$w_{at} = W(o_{at},\ h_{at}, X_{at})$$

Similarly, I hypothesise when health outcomes o_{at} or quantity h_{at} or quality q_{at} of healthcare services improves or increases, subjective wellbeing w_{at} improves. Health outcomes and wellbeing can then be expressed the following form. $F(.)$ and $G(.)$ are functions that turn funding allocations and spending shock into health outcomes and wellbeing respectively.

$$o_{at} = F\left(AllocationPH_{at},\ X_{at},\ \mu_{at}\right) = F\left(R_t\left(b_t, n_{at}\right),\ X_{at}, \mu_{at}\right)$$

$$w_{at} = G\left(AllocationPH_{at},\ X_{at},\ \mu_{at}\right) = G\left(R_t\left(b_t, n_{at}\right),\ X_{at},\ \mu_{at}\right)$$

Under these assumptions and hypotheses, the first-order partial derivative of outcomes and wellbeing will be positive with respect to the allocation per head for each CCG $AllocationPH_{at}$,

holding needs n_{at} and other factors X_{at} constant.

$$\left.\frac{\partial o_{at}}{\partial AllocationPH_{at}}\right|_{n_{at},X_{at},\mu_{at}} > 0 \quad \text{and} \quad \left.\frac{\partial w_{at}}{\partial AllocationPH_{at}}\right|_{n_{at},X_{at},\mu_{at}} > 0$$

This reflects our research question that I am inquiring: whether exogenous increase in healthcare funding allocations leads to improvements in health outcomes and wellbeing at the local authority level. To evaluate with an estimation equation, I apply linear approximations to the functions $F(.)$ and $G(.)$, as follows:

$$\begin{aligned}
o_{at} \approx\ & F\left(\overline{AllocationPH}_{at},\ \overline{\mu}_{at}\right) \\
& + F_1\left(\overline{AllocationPH}_{at},\ \overline{\mu}_{at}\right)\left(AllocationPH_{at} - \overline{AllocationPH}_{at}\right) \\
& + F_2\left(\overline{AllocationPH}_{at},\ \overline{\mu}_{at}\right)(x_{at} - \overline{x}_{at}) \\
& + F_3\left(\overline{AllocationPH}_{at}, \overline{x}_{at},\ \overline{\mu}_{at}\right)(\mu_{at} - \overline{\mu}_{at})
\end{aligned}$$

$$\begin{aligned}
w_{at} \approx\ & G\left(\overline{AllocationPH}_{at},\ \overline{x}_{at},\ \overline{\mu}_{at}\right) \\
& + G_1\left(\overline{AllocationPH}_{at}, \overline{x}_{at},\ \overline{\mu}_{at}\right)\left(AllocationPH_{at} - \overline{AllocationPH}_{at}\right) \\
& + G_2\left(\overline{AllocationPH}_{at}, \overline{x}_{at},\ \overline{\mu}_{at}\right)(x_{at} - \overline{x}_{at}) \\
& + G_3\left(\overline{AllocationPH}_{at}, \overline{x}_{at},\ \overline{\mu}_{at}\right)(\mu_{at} - \overline{\mu}_{at})
\end{aligned}$$

where F_k or G_k is the partial derivative of the functions of health outcomes and wellbeing with respect to the kth input of the function. The bars on top of the variables refers to the given level of allocation, non-health shocks, and idiosyncratic spending changes that I am approximating around. Since the bar-values are unspecified, in the following estimation, I proxy them with the observed values from the previous year, i.e. $t-1$.

5. Empirical Strategy and Results

5.1. OLS Estimation (Panel A)

To evaluate the exogenous effect of allocations on outcomes and wellbeing, I arrive at the first difference estimation model from above linear approximations and adding a time fixed effect term

to control for secular trends in outcomes or wellbeing[15]. (equation 3: FD model):

$$\Delta y_{at} = \beta_0 + \beta_1 \Delta AllocationPH_{at} + \beta_X \Delta X_{at} + \varphi_t + v_{at} \tag{3}$$

where $\Delta y_{at} = y_{at} - y_{at-1}$, refers to changes in health outcomes o_{at} or wellbeing w_{at} for local authority a between financial year t and year $t-1$. $\Delta AllocationPH_{at} = AllocationPH_{at} - AllocationPH_{at-1}$, referring to changes in per head allocation for local authority a between financial year t and year $t-1$. ΔX_{at} reflects the changes in the control variables (non-health factors) such that $\Delta X_{at} = X_{at} - X_{at-1}$. β_X is a vector for the coefficients for ΔX_{at}. Time fixed effects φ_t is the time fixed effects and v_{at} is the error term.

However, differencing highly persistent allocation data reduces the variation for identification, increasing attenuation bias and noise-to-signal ratio (Griliches and Hausman, 1986). To preserve more meaningful variation, I also arrive at the following two-way fixed effects estimation model (equation 4: FE model):

$$y_{at} = \beta_0 + \beta_1 AllocationPH_{at} + \beta_X X_{at} + \delta_a + \varphi_t + \varepsilon_{at} \tag{4}$$

where y_{at} is the dependent variable, which can health outcomes o_{at} or wellbeing w_{at} for local authority a in financial year t. $AllocationPH_{at}$ is the NHSE allocation per head for local authority a in financial year t. β_X is a vector for the coefficients of X_{at}. Local authority (LAD) fixed effects δ_a control for LAD-specific time-invariant determinants of health outcomes or wellbeing. Time fixed effects φ_t control for secular trends in outcomes or wellbeing. The variable ε_{at} is the error term, where is related to FD's error term in $v_{at} = \Delta\varepsilon_{at} = \varepsilon_{at} - \varepsilon_{at-1}$,

The coefficient of interest is β_1 in both models. It captures the short-run effect of a £1000 increase in allocation per head (mean = £1,270) for a local authority on avoidable mortality rate or wellbeing scores. Tables 5 and 6 present the OLS estimates from FD and FE models for the entire study period, 2013-14 to 2019-20, respectively for avoidable mortality and wellbeing measures.

Regarding the calculation of standard errors, I implement CCG-intersection clustering, following the design-based framework of Abadie et al. (2023). Since healthcare allocations are assigned at the CCG level, clustering should reflect this assignment mechanism. However, my analysis uses LAD-level data because outcomes are measured in this geography, and approximately 22% of LADs span multiple CCG boundaries for at least one period. The representative resident in these crosscutting LADs (e.g. CCG A & B) receives a mix of multiple CCGs' healthcare services, creating treatment heterogeneity within what would otherwise be a single CCG cluster.

[15]There is no need to add local authority fixed effects, since for LAD-specific determinants of outcomes are cancelled out by the first difference.

Table 5: Effect of NHS Allocations on Avoidable Mortality: OLS Estimates (2013/14-2019/20)

	Avoidable Mortality		Preventable Mortality		Treatable Mortality	
	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
Allocation per head	-54.002**	-36.431*	-31.451*	-15.107	-22.376*	-21.143*
	(22.618)	(21.099)	(17.419)	(15.602)	(11.753)	(11.142)
Controls						
Crime	15.211	-16.801	3.524	-15.958	11.462	-1.130
	(74.188)	(62.375)	(48.713)	(42.597)	(39.883)	(34.015)
Disposable Income	-82.164**	-78.624***	-45.858**	-45.643**	-36.211**	-33.147**
	(32.812)	(28.071)	(23.114)	(19.216)	(15.917)	(13.724)
Emissions	-5.452	-4.111	-4.866	-2.138	-0.575	-1.937
	(7.378)	(7.527)	(5.803)	(5.854)	(3.001)	(2.907)
First Mortgage	88.330	70.853	26.476	23.821	61.849**	47.088**
	(62.802)	(47.995)	(39.887)	(31.473)	(30.531)	(22.917)
Pupil Absence Rate	-141.357	-90.867	-59.088	-59.290	-81.631	-31.159
	(91.532)	(83.131)	(76.699)	(64.508)	(56.704)	(45.567)
Unemployment Rate	89.838	-10.058	79.257	1.112	10.826	-11.274
	(93.559)	(66.946)	(61.116)	(46.084)	(52.397)	(38.684)
Local area FE	Yes	-	Yes	-	Yes	-
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.980	0.028	0.978	0.011	0.946	0.024
Within R^2	0.060	0.022	0.033	0.011	0.057	0.021
Observations	1,470	1,176	1,470	1,176	1,470	1,176
Serial Correlation Tests						
Wooldridge test	447.709		309.483		497.430	
	(0.000)		(0.000)		(0.000)	
Breusch-Godfrey test	203.079	5.605	194.410	0.525	212.824	3.955
	(0.000)	(0.061)	(0.000)	(0.769)	(0.000)	(0.138)

Note: The table shows OLS FE and FD estimates of the impact of NHS allocations on mortality measures with controls. Dummy variables are created for missing controls. Columns 1, 3, and 5 (FE) show estimates based on the two-way fixed effects OLS model, while columns 2, 4, and 6 (FD) show estimates based on the first difference OLS model. The dependent variable is given by the column heading. Standard errors are clustered at the CCG-intersection level and given in parentheses. The Wooldridge F-statistic is a test for first order serial correlation. Breusch-Godfrey χ^2_2-statistic is a test for second order serial correlation. P-values for all the serial correlation tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Standard CCG clustering would misrepresent this treatment exposure by assigning crosscutting LADs to a single cluster based on their primary CCG. This implicitly assumes they experience the same healthcare environment as LADs fully within that CCG. In reality, residents of an LAD x spanning CCGs A and B face a different healthcare environment than residents of a nearby LAD y entirely within CCG A. This is because LAD x residents receive services from providers in both CCGs and are subject to both allocation of both CCGs' allocation.

Table 6: Effect of NHS Allocations on Wellbeing: OLS Estimates (2013/14-2019/20)

	Wellbeing		Happiness		Worthwhile		Life Satisfaction		Anxiety	
	FE	FD	FE	FD	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: OLS										
Allocation per head	0.362*	-0.009	0.335	-0.246	0.262	0.045	0.230	-0.137	-0.619*	-0.343
	(0.188)	(0.189)	(0.205)	(0.247)	(0.187)	(0.259)	(0.178)	(0.240)	(0.363)	(0.396)
Controls										
Crime	-0.661	-0.268	-0.649	-0.306	-0.539	-0.218	-0.697	0.080	0.833	0.676
	(0.484)	(0.590)	(0.575)	(0.827)	(0.511)	(0.545)	(0.552)	(0.538)	(0.838)	(1.174)
Disposable Income	-0.243	0.131	0.071	0.328	-0.013	0.098	0.077	0.444	1.065**	0.315
	(0.238)	(0.319)	(0.254)	(0.390)	(0.266)	(0.382)	(0.264)	(0.426)	(0.465)	(0.579)
Emissions	-0.018	-0.011	-0.064	-0.041	0.053	0.051	-0.068	-0.032	-0.011	0.028
	(0.072)	(0.089)	(0.094)	(0.115)	(0.076)	(0.106)	(0.098)	(0.104)	(0.113)	(0.158)
First Mortgage	0.552	1.093**	0.407	1.173*	0.910**	0.675	1.372***	1.418**	0.522	-0.833
	(0.441)	(0.537)	(0.519)	(0.684)	(0.418)	(0.584)	(0.408)	(0.633)	(0.906)	(0.960)
Pupil Absence Rate	0.440	0.154	0.319	-0.712	-0.583	-0.162	-0.238	0.441	-2.052	-0.619
	(0.926)	(1.036)	(1.028)	(1.237)	(0.971)	(1.078)	(0.934)	(1.088)	(1.615)	(1.786)
Unemployment Rate	-2.519***	-1.399*	-1.790**	0.263	-2.183***	-2.184***	-2.311***	-1.517*	3.842***	2.257*
	(0.610)	(0.767)	(0.762)	(0.958)	(0.634)	(0.805)	(0.636)	(0.825)	(1.113)	(1.348)
Local area FE	Yes	-	Yes	-	Yes	-	Yes	-	Yes	-
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.418	0.062	0.321	0.031	0.399	0.021	0.455	0.046	0.304	0.051
Within R^2	0.015	0.003	0.006	0.002	0.012	0.004	0.016	0.005	0.014	0.002
Observations	2,260	1,933	2,264	1,938	2,264	1,938	2,265	1,940	2,260	1,933
Serial Correlation Tests										
Wooldridge test	27.007		4.495		16.841		20.052		33.135	
	(0.000)		(0.034)		(0.000)		(0.000)		(0.000)	
Breusch-Godfrey test	45.916	291.159	60.409	371.777	37.795	301.496	36.283	308.520	55.719	275.777
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note: The table shows OLS FE and FD estimates of the impact of NHS allocations on wellbeing measures with controls. Dummy variables are created for missing controls. Columns 1, 3, 5, 7, and 9 (FE) show estimates based on the two-way fixed effects OLS model, while columns 2, 4, 6, 8, and 10 (FD) show estimates based on the first difference OLS model. The dependent variable is given by the column heading. Standard errors are clustered at the CCG-intersection level and given in parentheses. The Wooldridge F-statistic is a test for first order serial correlation. Breusch-Godfrey χ^2_2-statistic is a test for second order serial correlation. P-values for all the serial correlation tests are given in parentheses. $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

CCG-intersection clustering addresses this by grouping LADs according to their unique combination of CCG exposures. LADs fully within one CCG cluster by that CCG, while LADs spanning multiple boundaries form distinct clusters (e.g. a LAD intersecting CCGs A and B forms cluster "AB"). This yields 178 pure-CCG clusters and 76 intersection clusters, appropriately reflecting both the CCG-level assignment mechanism and the mixing of treatment exposure from different CCGs created by administrative boundaries. The results are robust to alternative clustering strategies, including standard CCG clustering and LAD clustering.

Results using CCG-intersection clustering are presented as the main specification throughout the paper. For comparison, I report standard errors under two alternative approaches in Appendix C: primary CCG clustering (assigning each LAD to its primary parent CCG) and LAD-level clustering. The choice of clustering approach affects standard errors modestly, but does not alter the substantive conclusions.

For avoidable mortality measures in Table 5, the FE results indicate that a £1,000 increase in allocation per head for a LAD (mean = £1,270) is associated with a reduction of 31.451 preventable and 22.376 treatable mortality per 100,000 population on a 3-year average (p-values 0.072 and 0.058 respectively). The results of the FD model are roughly consistent with the FE model with smaller coefficients.

For wellbeing measures in Table 6, FE results show that an increase of £1,000 in allocation per head for a local area is associated with a 0.362 increase in wellbeing scores, a 0.335 increase in happiness scores and 0.619 reduction of anxiety scores (p-values 0.054, 0.103 and 0.090). Other measures in the FE model and all measures in the FD model show no significant effect from increased funding.

Both models offer insights into our research question. The FD model (equation 3) aligns better with the theoretical model and the goal of evaluating the effect of marginal changes. However, the FE model (equation 4) is generally superior to the FD model with more reliable estimates. Although both models have similar degrees of freedom, the FD model suffers from attenuation bias and higher standard errors due to the differencing of highly persistent allocation data (Griliches and Hausman, 1986). Therefore, I mainly report on the results of the FE model, with the FD model as additional evidence.

5.2. Endogeneity problem of OLS

Under the assumption that the treatment variable ($AllocationPH_{at}$) is exogenous, OLS estimates of β_1 captures the causal effect of increased allocations on outcomes and wellbeing. However, my framework demonstrates that the NHSE allocation rules is a main source of an endogeneity problem: endogenous targeting of resources. On one hand, allocations are a year-specific function of national budget and needs of the CCG, $AllocationPH_{CCGt} = R_t(b_t, n_{CCGt})$. This propagates to a LAD level, where allocations $AllocationPH_{at}$ depend on the relative need of the area. On the other hand, a determinant of health outcomes is also the needs n_{at}. An endogeneity problem arises because the variable n_{at} appears simultaneously in both equations. Allocations are confounded by both area-specific and time-varying healthcare needs. Standard fixed effects cannot resolve this since n_{at} varies over time within areas. Therefore, in a standard OLS regression, the error term ε_{at} contains unobserved components of need that dictate funding, where $corr\,(AllocationPH_{at}, \varepsilon_{at}) \neq 0$. OLS estimates of β_1 are biased and only returns associations between allocations and outcomes.

The direction of the bias is likely downwards in absolute magnitude when estimating returns to funding. This is because health needs n_{at} are expected to negatively correlated with outcomes o_{at} and wellbeing w_{at} (i.e. higher needs imply worse health and wellbeing), while positively correlated

with allocations $AllocationPH_{at}$ (i.e. funding follows need). Allocations and the error term are negatively correlated, $corr\ (AllocationPH_{at}, \varepsilon_{at}) < 0$, leading to underestimated effects.

To disentangle the source of endogeneity, I find five key sources of endogeneity within NHSE allocation policy rules that likely drives outcomes simultaneously.

1. Demographics and population changes drive total allocation changes for the CCG. A population growing older with increasing needs receiving more funding is endogenous.
2. NHSE uses standardised mortality rates for those under 75 years of age (SMR<75) as a measure for unmet needs. Allocations are then adjusted according to the unmet needs, where poor outcomes translate directly to more allocations.
3. The needs weights for each age-CCG matrix are updated in the main allocation rounds to reflect relative need. Relative need is estimated from past patterns of utilisation of health services. Since utilisation is likely linked to health outcomes, such changes present another channel of endogeneity.
4. Supply-side adjustments, such as unavoidable costs, might reflect changes in price levels and costs in an area. This might in turn affect living costs and living standards within that area, which could affect outcomes and cause potential endogeneity.
5. Pace of change policy also considers CCG-specific time-varying variables. Endogeneity is therefore present when the policy's variables are applied to estimate the actual allocations for the CCG given the target and baseline allocations.

Followingly, my geographical mapping methodology translates the above sources of endogeneity to a LAD level.

5.3. Accounting for Endogeneity

A major challenge in estimating returns to healthcare is isolating exogenous variation. Standard OLS estimators fail to account for funding or spending endogeneity, yielding biased estimates (Hitiris and Posnett, 1992; Crémieux et al., 1999; Karaman et al., 2020). The literature has commonly relied on instrumental variable (IV) strategies to address this bias. For validity, instruments must satisfy the exclusion restriction (independence from the error term) and the relevance condition (strong correlation with the treatment) (Staiger and Stock, 1997). However, finding exogenous instruments with high power remains difficult within the context of healthcare spending or funding.

Cross-country studies mostly employ external instruments, such as national fiscal deficit, emissions, or military expenditures of the neighbouring countries (Filmer and Pritchett, 1999; Bokhari

et al., 2007; Farahani et al., 2010; Moreno-Serra and Smith, 2015). However, the validity of these instruments remains contested, as they rarely satisfy both the exclusion and relevance conditions simultaneously. Findings in this literature are highly sensitive to instrument choice. Many previously accepted instruments suffer from weak instrument bias or fail over-identification tests when applied to extended datasets, leading to unstable estimates of the returns to spending (Nakamura et al., 2016).

More recent studies utilise internal instruments from lagged values or heteroscedasticity with Generalized Method of Moments (Rahman et al., 2018; Ray and Linden, 2020; Anwar et al., 2023). However, such strategies often lack economic interpretation and rely on strong assumptions regarding the absence of serial correlation in errors (Roodman, 2009). This assumption is likely violated in health data where outcomes are highly persistent.

For within-country studies, endogeneity is often addressed via quasi-experimental variation rather than traditional IVs[16]. For example, Currie and Gruber (1996), Card et al. (2008), Finkelstein et al. (2012), Sohn (2017) and Goodman-Bacon (2018) leverage exogenous variation from Medicaid eligibility to estimate its effect on health outcomes. Similarly, Doyle (2011) removes local area confounders by utilising random far-from-home emergency healthcare demand to identify the returns to hospital care. While internally valid, these methodologies rely heavily on unique, idiosyncratic events or shocks, making implementation and generalisation difficult for national healthcare systems.

Given the NHS's structural stability, the main identification strategy in the UK context remains as instrumental variables. A innovation in instrument choice is instrumenting expenditure with exogenous components in the NHS resource allocation rules for cross-sectional analysis (Andrews et al., 2017). The exogenous "allocation rule" instruments are the age index, input price index, and distance from target index in Primacy Care Trust (PCT) allocation rules. These instruments are empirically shown superior to other socioeconomic factors in isolate funding variation unrelated to immediate health shocks. As a result, cross-sectional studies on NHS marginal productivity and cost effectiveness have shifted from using socioeconomic instruments[17] (Martin et al., 2008a,b,c; Claxton et al., 2015) to these "allocation rule" instruments (Lomas et al., 2019; Martin et al., 2021a,b, 2022, 2023).

However, the "allocation rule" instruments lose power and become endogenous in the panel

[16]Within-country studies utilising tradition IV includes Thornton (2002), Thornton (2011), Mays and Smith (2011) and Brown et al. (2014)

[17]Instruments used in existing literature include variables such as the proportion of the population providing unpaid care, the proportion of lone pensioner households, socioeconomic deprivation, and the availability of informal care. There are endogeneity concerns with these instruments since these factors likely influence health outcomes directly, rather than solely through the channel of healthcare spending (violating the exclusion restriction).

setting employed by this paper. Firstly, because these instruments are stable over time, CCG fixed effects absorb most of the variation, greatly reducing their power. Secondly, the remaining variation is likely endogenous, as identified in section 5.2. Changes in the age index reflect demographic shifts that directly correlated with health outcomes, the input price index captures local cost-of-living dynamics, which can be linked to poverty and stress, and the "distance from target" index introduces dynamic endogeneity, as past health needs directly influences current allocations. In sum, any variation in components of allocation rules driven by local-specific time-varying factors is endogenous.

Only remaining source of exogeneity that I can leverage is the variations in allocation rules itself. Specifically, it is national allocation rules changes that drive heterogeneous, exogenous variations in allocations across CCGs and time.

5.4. Simulated Instrument Construction

To address above challenges when accounting for the endogeneity issue, I employ a simulated instrumental variable approach, which has been used in other fields of economic analysis (Gruber and Saez, 2002; Biasi, 2023). It constructs a counterfactual measure of treatment that varies solely due to changes in national policy rules, holding local responses constant. This addresses challenges of finding suitable instruments and limited econometric modelling alternatives in continuous treatment for all units.

Specifically, I construct the instrument "simulated allocation per head", $SAllocation_{at}$. Simulated allocation is the counterfactual set of allocations a CCG would have received if its demographics and needs had remained fixed at their base year[18] levels. Hence, the instrument isolates variation in allocations driven by national allocation rule changes from local endogenous responses. To do so, I interact the fixed characteristics with time-varying national allocation rules (formulae). The fixed inputs are:

1. Demographics and population changes
2. Needs weights for each age-CCG matrix
3. Unmet needs adjustments from SMR<75
4. Supply-side adjustments
5. Any CCG-specific time-varying variables used in the pace of change policy are held constant at the base year.

[18]This can be considered baseline levels of demographics and needs. The word choice of "base year" prevents confusion with "baselines" in allocation formulae, which represents the previous allocation levels.

I then allow the instrument to evolve over time only through changes in allocation rules and the total budget. Exogenous changes include:

1. Changes to the aggregate national budget for CCG allocations
2. Updates to national proportions that combine weighted populations for different areas of services (e.g. the split between maternity and prescription services based on past national expenditure)
3. The pace of change policy (i.e. the convergence of baseline to target allocations over time)
4. The introduction of new national equity adjustments

Consequently, the simulated instrument $SAllocation_{at}$ represents the counterfactual set of allocations a CCG would have received over time, if its demographics and needs had remained fixed at their base year[19] levels. Equation 5 summarises the construction of the simulated instrument $SAllocation_{CCG,t}$:

$$SAllocation_{CCG,t} = R_t\big(NationalBudget_t, SBaseline_{CCG}, \overline{Population_{CCG,t}}, \overline{Needs_{CCG,t}}, \overline{Adjustments_{CCG,t}}\big) \tag{5}$$

where the barred variables refer to variables fixed at the base year levels ($\overline{Population_{CCG,t}}$, $\overline{Needs_{CCG,t}}$, $\overline{Adjustments_{CCG,t}}$). For $SBaseline_{CCG}$, it refers to the simulated baseline, i.e. the simuluated allocation of the previous year, $SAllocation_{CCG,t-1}$.

I set 2015 as the base year, which has weighted population levels used to calculate 2016/17 allocations, where section 5.7 explains the base year choice. I divided it by CCG's fixed population and mapped the simulated allocation per head to the local authority level accordingly. Appendix B details the process of constructing the instrument.

By strictly holding local inputs constant, this construction mechanically isolates variation driven solely by national allocation rule updates, offering identification advantages.

5.5. Causal Validity of Simulated Instruments

The causal validity of the simulated-instrument approach relies on five key assumptions required to identify the local average treatment effect (LATE). I address each assumption in turn.

First, my simulated instrument satisfies the relevance of the instrument assumption. Unlike the standard approach of using observables as instruments, which often discard useful variation, this

[19] This can be considered baseline levels of demographics and needs. The word choice of "base year" prevents confusion with "baselines" in allocation formulae, which represents the previous allocation levels.

simulated approach retains all variation driven by allocation rules or national funding changes. This strategy preserves statistical power, enhances the first-stage F-statistics ($F > 50$), and improves the efficiency of causal estimation, compared to standard IV approaches.

Second, I argue that the simulated instrument satisfies the instrument exogeneity assumption after addressing potential threats. While the construction of the instrument mechanically eliminates reverse causality from local feedback loops, two threats remain: dynamic endogeneity and policy endogeneity (local health factors influencing policy rules changes).

Regarding dynamic endogeneity, I prevent it by calculating simulated instruments based on the fixed base year (Appendix B). I use simulated baselines, $SBaseline_{CCGt}$, not updating the instrument with actual baselines or allocations. The instrument solely captures the role of rule changes in the evolving allocations, without introducing endogenous change in allocations.

Regarding policy endogeneity, I argue that the institutional structure of the NHS allocations ensures that the rule changes are exogenous variation at the local level, while any aggregate compensatory changes ensure conservativeness.

All allocation rule changes in this study period are "formula-focused", rather than "needs-based", as detailed in Appendix A. The changes are mostly technical updates on how weighted populations are combined or how pace of change policy is set across all CCGs, rather than reacting to specific health needs or financial deficits (which are held constant with the instrument). Furthermore, allocation rule changes are recommended by the Advisory Committee on Resource Allocation (ACRA), an independent technical body. These rules are set centrally based on national constraints, spending reviews and broad political priorities, rather than specific local effects or influence. This institutional separation means that individual CCGs or LADs cannot unilaterally affect national formula decisions to offset its own health shocks. From the perspective of a single local area, rule changes represent exogenous shocks in funding.

However, even if the strict exogeneity is relaxed, the general compensatory design of resource allocation ensures no exaggerated estimates. While individual local areas cannot dictate formula changes, ACRA or NHSE might change the allocation rules in a way that it favours poorer areas, since they are legally mandated by the Health and Social Care Act (2012) to reduce health inequalities. This implies that even if national rules do systematically respond to specific local health patterns, they will compensate and target resources toward areas with worsening outcomes[20]. Any endogeneity induced negatively correlates the instrument and the error term, where worsening

[20]This structural mechanism can be illustrated by the 2016/17 pace of change policy, where NHS England introduced a national funding floor ensuring no CCG remained more than 5% below its target. Such changes redistribute funds to areas furthest below their target allocations (often those with higher relative needs). This confirms that the policy reaction function is compensatory: funding follows deprivation. This structure ensures that any endogeneity biases the estimated health effects toward zero, rather than inflating them.

health triggers positive funding shocks. Consequently, this specific form of reverse causality generates a downward bias and incorporates conservativeness, where the true causal effect of funding on health outcomes can potentially be larger than my estimates.

Third, my estimation models satisfy the exclusion restriction assumption, which requires that the simulated instrument affects outcomes only through actual allocations, not through correlated trends or omitted factors. The simulated instrument has no direct effect on outcomes beyond its impact on funding. This holds because rule changes are technical formula adjustments, not policy interventions that CCGs or residents observe directly. I ensure conditional parallel trends by including area and time fixed effects, plus time-varying LAD-specific controls for potential confounders (see Table 3). These controls ensure that, conditional on observed characteristics, any remaining differential trends in outcomes are attributable to the funding variation captured by the instrument.

Fourth, the definition of CCG's responsibility ensures stable unit treatment value assumption (SUTVA). CCGs have a statutory duty to provide care for the population registered in their specific geographic area, instead of the location of the healthcare consumption. Because funding is tied to a specific registered list of patients, the "treatment" is largely contained within the administrative boundary. This eliminates interference where funding in one CCG improves outcomes in another. When CCG allocations are mapped to the LAD level, the assumption remains satisfied. For LADs located entirely within one CCG, they receive identical treatment. For LADs crosscutting CCG boundaries, I employ population-weighted averages to construct allocation values reflecting the specific CCG exposure mix. Combined with CCG-intersection clustering, this ensures that treatment is appropriately defined and spillovers are contained within exposure clusters.

Fifth, monotonicity holds because CCGs cannot refuse or modify their formula-determined allocations. Rule changes that increase simulated funding mechanically increase actual funding for all affected CCGs (i.e. no defiers).

Under these five assumptions, the IV estimator identifies the LATE of additional funding on avoidable mortality and wellbeing (Angrist and Pischke, 2009). Despite the simulated instrument providing a counterfactual for every local authority, IV estimates only capture the "local" effect: the effect on CCGs whose allocations are shifted significantly in direct response to the national rule changes. For CCGs without significant rule-driven allocation changes, their funding variation contributes little to identification. Interpreting these results as an average treatment effect (ATE) would require homogeneous treatment effects across all CCGs, which is unlikely true given the diverse differences across England.

5.6. Instrumental Variables Estimation (Panel B)

Now, I now estimate the models with instrumental variables (IV), using simulated allocations $AllocationPH_{at}$ as an instrument. For the first difference (FD) model, the IV estimation equations are as follows:

$$1^{\text{st}} \text{ stage:} \quad \Delta AllocationPH_{at} = \gamma_0 + \gamma_1 \Delta SAllocationPH_{at} + \gamma_X \Delta X_{at} + \varphi_t + v_{1,at} \tag{4.1}$$

$$2^{\text{nd}} \text{ stage:} \quad \Delta y_{at} = \beta_0 + \beta_1 \Delta \widehat{AllocationPH}_{at} + \gamma_X \Delta X_{at} + \varphi_t + v_{2,at} \tag{4.2}$$

For the two-way fixed-effects (FE) model, the IV estimation equations are as follows:

$$1^{\text{st}} \text{ stage:} \quad AllocationPH_{at} = \gamma_0 + \gamma_1 SAllocationPH_{at} + \gamma_X X_{at} + \delta_a + \varphi_t + \varepsilon_{1,at} \tag{5.1}$$

$$2^{\text{nd}} \text{ stage:} \quad y_{at} = \beta_0 + \beta_1 \widehat{AllocationPH}_{at} + \gamma_X X_{at} + \delta_a + \varphi_t + \varepsilon_{2,at} \tag{5.2}$$

The variable y_{at} or Δy_{at} are the dependent variables, which can be health outcomes o_{at} or wellbeing w_{at} for local authority a in financial year t. The variable $SAllocationPH_{at}$ is the instrument, simulated allocation per head for local authority a in financial year t. The hats refer to the fitted values of the allocation per head from the first stage of instrument variable estimation. Local authority fixed effects δ_a are used in FE model to control for LAD-specific time-invariant determinants of health outcomes or wellbeing. Time fixed effects φ_t control for secular trends in outcomes or wellbeing for both models. The variables v_{at} and ϵ_{at} are the error terms.

The study period is trimmed to be between 2014/15 and 2019/20. This is necessary because the first year's allocation in this study (2013/14) has to be used as a starting baseline for actual allocation in 2014/15. Simulated allocation per head for 2013/14 is therefore not possible to be constructed. OLS and IV estimates for both FE and FD models for the trimmed period are shown in tables 7 and 9 for avoidable mortality and wellbeing measures respectively. My instruments are all strong with large first-stage F-statistics (F > 50 across FE specifications, compared with standard rule-of-thumb at F > 10).

Furthermore, to test for lagged effects, I estimate the effect of previous year's allocation on outcomes, using a 1-period lagged model. I replaced $AllocationPH_{at}$ with $AllocationPH_{at-1}$ in all the empirical models outlined above. OLS and IV estimates for both lagged FE and FD models are shown in tables 8 and 10[21]. Attempting to study the dynamic effects, I also estimated

[21]For avoidable mortality estimates, a 1-period lag refers to a 1-year lag in 3-year averages (i.e. 1-period lag for 3-average from 2015/16-2017/18 is 2014/15-2016/17).

a two-period model with results in Appendix D. However, there is strong collinearity between current and lagged allocation. The strength of the instruments in the first stage reduces, leading to highly imprecise second stage coefficient estimates with large standard errors. Therefore, isolating the current and lagged effects of allocation is not possible in this study.

Avoidable Mortality Results

For avoidable mortality measures (tables 7 & 8), OLS and IV estimates indicate a significant negative effect of allocation on avoidable mortality in the short run[3]. Lagged models have lower coefficient estimates and larger standard error, which suggest that the effect of increase in allocation on 3-year rolling averages is mostly realised within the 3-year period.

There is significant discrepancy between OLS and IV estimates. OLS estimates suggest increased allocation benefits for both preventable and treatable mortality. In contrast, the IV models reveal that the benefits are entirely driven by preventable mortality. IV coefficients for preventable mortality are large and significant. FE-IV results indicate that a £1,000 increase in allocation per head for a LAD has a causal effect in reducing 101.23 preventable mortality per 100,000 population on a 3-year rolling averages ($p = 0.039$). In contrast, IV coefficients for treatable mortality are statistically indistinguishable from zero. I leave the discussion on the distinct difference in marginal returns in section 6, while in the following, I focus on the direction of endogeneity in treatable mortality opposite to expected and being insignificant, where higher treatable mortality potentially receives less allocation.

This counter-intuitive direction of endogeneity can be explained by the construction of the allocation rules. NHSE allocation formulae utilise standardised mortality ratio for those under 75 (SMR<75) to adjust for unmet needs. Since SMR<75 captures all causes of deaths, it includes avoidable deaths. Because preventable mortality makes up more (64.1%) of avoidable mortality than treatable mortality (35.9%)[22], SMR<75 is more responsive to preventable than treatable mortality in percentage terms. Hence, allocation formulae respond to proportional changes in preventable mortality more sensitively than treatable mortality. Since resources are split from a fixed national budget, allocations are mechanically favour for reducing preventable mortality over treatable mortality. This explains when accounting for spending endogeneity through IV models, a larger reduction for preventable but no reduction for treatable mortality is observed.

This pattern is consistent with the Wu-Hausman tests results[23], which indicate significant

[22] Figures calculated using "avoidable mortality by cause" data for England, excluding deaths attributed to COVID-19 (Office for National Statistics, 2025b).

[23] The Durbin-Wu-Hausman test (referred to as the Wu-Hausman test) evaluates the consistency of an estimator when compared to an alternative, less efficient estimator that is already known to be consistent (Wu, 1973; Hausman, 1978).

Table 7: Effect of NHS Allocations on Avoidable Mortality: OLS and IV Estimates (2014/15-2019/20)

	Avoidable Mortality		Preventable Mortality		Treatable Mortality	
	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
Allocation per head	-48.543*	-33.055	-24.178	-12.993	-24.190*	-20.047*
	(29.442)	(24.602)	(24.398)	(19.219)	(13.994)	(11.797)
Panel B: IV						
2nd Stage						
Fitted allocation per head	-103.681	-111.769*	-101.232**	-101.563**	-2.346	-10.309
	(65.306)	(57.650)	(48.648)	(42.063)	(28.011)	(24.848)
1st Stage						
Simulated allocation per head	0.634***	0.596***	0.634***	0.596***	0.634***	0.596***
	(0.087)	(0.073)	(0.087)	(0.073)	(0.087)	(0.073)
Local area FE	Yes	-	Yes	-	Yes	-
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.983	0.010	0.980	-0.029	0.957	0.029
Within R^2	0.055	0.012	-0.004	-0.021	0.059	0.029
Observations	1,176	882	1,176	882	1,176	882
Diagnostic Tests						
F-statistic (1st stage)	53.721	66.443	53.721	66.443	53.721	66.443
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic	0.895	2.279	3.352	5.604	0.810	0.198
	(0.344)	(0.131)	(0.067)	(0.018)	(0.368)	(0.656)
Wooldridge test	263.881		215.061		259.024	
	(0.000)		(0.000)		(0.000)	
Breusch-Godfrey test	145.161	0.062	124.520	1.471	153.284	0.330
	(0.000)	(0.970)	(0.000)	(0.479)	(0.000)	(0.848)

Note: The table shows OLS and IV estimates of the impact of NHS allocations on mortality measures with controls. Panel A reports OLS results; Panel B reports IV results where allocation per head is instrumented by the simulated allocation per head. Dummy variables are created for missing controls. Columns 1, 3, and 5 (FE) show estimates based on the two-way fixed effects models, while columns 2, 4, and 6 (FD) show estimates based on the first difference models. Standard errors are clustered at the CCG-intersection level and given in parentheses. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under clustered standard errors. The Wu-Hausman statistic tests for endogeneity. The Wooldridge test and the Breusch-Godfrey test checks for first-order and second-order serial correlation respectively. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

endogeneity for preventable but not for treatable mortality. Null hypothesis of exogeneity rejected for preventable but not rejected for treatable mortality. The difference in test statistics shows the endogeneity effect is stronger in preventable than in treatable mortality. This aligns with the

Specifically, it tests for the significance of endogeneity by comparing Ordinary Least Squares (OLS) estimates to Instrumental Variables (IV) estimates. The test is valid under the assumption of the instrument is strong (F > 10). A significant test statistic rejects the null hypothesis that the regressors are exogenous, indicating that OLS is inconsistent and IV estimation is required.

Table 8: Lagged Effect of NHS Allocations on Avoidable Mortality: OLS and IV Estimates (2014/15-2019/20)

	Avoidable Mortality		Preventable Mortality		Treatable Mortality	
	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
Lagged allocation per head	-43.210	-33.714	-29.546	-24.482	-13.522	-8.982
	(34.412)	(27.862)	(27.958)	(22.598)	(17.550)	(14.141)
Panel B: IV						
2nd Stage						
Lagged fitted allocation per head	-76.826	-77.634	-77.679	-81.457**	1.327	4.370
	(59.872)	(50.073)	(48.531)	(40.741)	(29.895)	(24.329)
1st Stage						
Lagged simulated allocation per head	0.806***	0.795***	0.806***	0.795***	0.806***	0.795***
	(0.098)	(0.082)	(0.098)	(0.082)	(0.098)	(0.082)
Local area FE	Yes	-	Yes	-	Yes	-
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.986	0.030	0.984	-0.002	0.966	0.029
Within R^2	0.057	0.032	0.018	0.008	0.056	0.032
Observations	882	588	882	588	882	588
Diagnostic Tests						
F-statistic (1st stage)	67.667	93.463	67.667	93.463	67.667	93.463
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic	0.471	1.114	1.472	2.825	0.376	0.455
	(0.493)	(0.291)	(0.225)	(0.093)	(0.539)	(0.500)
Wooldridge test	96.786		111.680		73.540	
	(0.000)		(0.000)		(0.000)	
Breusch-Godfrey test	150.423	2.079	152.863	1.127	136.222	4.300
	(0.000)	(0.354)	(0.000)	(0.569)	(0.000)	(0.116)

Note: The table shows OLS and IV estimates of the impact of lagged NHS allocations (3-year lag) on mortality measures with controls. Panel A reports OLS results; Panel B reports IV results where allocation per head is instrumented by the simulated allocation per head. Dummy variables are created for missing controls. Columns 1, 3, and 5 (FE) show estimates based on the two-way fixed effects models, while columns 2, 4, and 6 (FD) show estimates based on the first difference models. Standard errors are clustered at the CCG-intersection level and given in parentheses. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under clustered standard errors. The Wu-Hausman statistic tests for endogeneity. The Wooldridge test and the Breusch-Godfrey test checks for first-order and second-order serial correlation respectively. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

mechanical prioritisation effect explained above.

However, the failure to reject the null in the Hausman test for treatable mortality may reflect low test power effect rather than true exogeneity. Hence, it is not optimal to rely solely on this result to select the OLS specification for treatable mortality, especially given endogeneity is likely a concern across all mortality types as shown in section 5.2. To maintain consistency, I present IV estimates as the primary specification for both types of avoidable mortality, avoiding ad-hoc aggregation of different estimator types. Therefore, I conclude that the causal benefits on avoidable

mortality are primarily driven by preventable mortality, with limited evidence for benefits from treatable mortality.

Table 9: Effect of NHS Allocations on Wellbeing: OLS and IV Estimates (2014/15-2019/20)

	Wellbeing		Happiness		Worthwhile		Life Satisfaction		Anxiety	
	FE	FD	FE	FD	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: OLS										
Allocation per head	0.186	-0.009	0.103	-0.251	0.189	0.006	0.126	-0.129	-0.325	-0.385
	(0.204)	(0.193)	(0.227)	(0.245)	(0.214)	(0.266)	(0.205)	(0.246)	(0.394)	(0.410)
Panel B: IV										
2nd Stage										
Fitted allocation per head	0.759	0.715	0.275	-0.679	0.468	-0.124	0.536	2.044	-1.606*	-0.883
	(0.519)	(1.776)	(0.572)	(2.135)	(0.512)	(1.542)	(0.559)	(2.156)	(0.907)	(3.151)
1st Stage										
Simulated allocation per head	0.531***	0.220***	0.531***	0.220***	0.531***	0.220***	0.531***	0.220***	0.531***	0.220***
	(0.064)	(0.046)	(0.064)	(0.046)	(0.064)	(0.046)	(0.064)	(0.046)	(0.064)	(0.046)
Local area FE	Yes	-	Yes	-	Yes	-	Yes	-	Yes	-
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.387	0.042	0.300	0.025	0.370	0.007	0.394	-0.007	0.303	0.047
Within R^2	0.008	0.000	0.006	0.000	0.012	0.005	0.013	-0.025	0.001	0.002
Observations	1,936	1,610	1,940	1,614	1,940	1,614	1,941	1,616	1,936	1,610
Diagnostic Tests										
F-statistic (1st stage)	68.532	22.924	68.532	22.924	68.532	22.924	68.532	22.924	68.532	22.924
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic	1.441	0.168	0.107	0.041	0.360	0.007	0.622	1.030	2.465	0.025
	(0.230)	(0.682)	(0.743)	(0.840)	(0.548)	(0.932)	(0.430)	(0.310)	(0.116)	(0.873)
Wooldridge test	16.458		3.202		9.896		11.400		17.444	
	(0.000)		(0.074)		(0.002)		(0.001)		(0.000)	
Breusch-Godfrey test	60.656	231.100	87.177	275.107	61.458	204.766	67.953	224.689	68.497	203.053
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note: The table shows OLS and IV estimates of the impact of NHS allocations on wellbeing measures with controls. Panel A reports OLS results; Panel B reports IV results where allocation per head is instrumented by the simulated allocation per head. Dummy variables are created for missing controls. Columns 1, 3, 5, 7, and 9 (FE) show estimates based on the two-way fixed effects models, while columns 2, 4, 6, 8, and 10 (FD) show estimates based on the first difference models. Standard errors are clustered at the CCG-intersection level and given in parentheses. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under clustered standard errors. The Wu-Hausman statistic tests for endogeneity. The Wooldridge test and the Breusch-Godfrey test checks for first-order and second-order serial correlation respectively. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Wellbeing Results

For wellbeing measures (table 9 & 10), wellbeing estimates shows some causal relationship with allocations but uncertainty increases. For the contemporaneous models, FE-IV results estimate a £1,000 increase in allocation per head for a LAD results in a reduction of 1.606 in anxiety scores and increase in overall wellbeing score of 0.759 (p-values 0.078 and 0.145). For the lagged models, FE-IV results indicate that a £1,000 increase in allocation per head for a LAD results in an improvement of 0.933 in happiness scores, 0.792 in life satisfaction scores and a reduction of 1.357 in anxiety scores (p-values 0.157, 0.131, 0.152). Such increase in allocation per head show stronger evidence for a lagged causal effect on increasing overall wellbeing scores by 0.849 (p-value 0.083). Moreover, for these measures, FD-IV models estimate a larger effect despite having greater

uncertainty.

Table 10: Lagged Effect of NHS Allocations on Wellbeing: OLS and IV Estimates (2014/15-2019/20)

	Wellbeing		Happiness		Worthwhile		Life Satisfaction		Anxiety	
	FE	FD	FE	FD	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: OLS										
Lagged allocation per head	0.023	-0.034	0.061	0.085	0.021	0.019	0.062	0.212	0.064	0.446
	(0.247)	(0.292)	(0.293)	(0.384)	(0.255)	(0.299)	(0.289)	(0.337)	(0.441)	(0.492)
Panel B: IV										
2nd Stage										
Lagged fitted allocation per head	0.849*	1.955*	0.933	3.606**	0.262	0.191	0.792	1.684	-1.357	-2.330
	(0.488)	(1.174)	(0.657)	(1.778)	(0.463)	(1.186)	(0.523)	(1.234)	(0.945)	(1.966)
1st Stage										
Lagged simulated allocation per head	0.746***	0.415***	0.746***	0.415***	0.746***	0.415***	0.746***	0.415***	0.746***	0.415***
	(0.080)	(0.055)	(0.080)	(0.055)	(0.080)	(0.055)	(0.080)	(0.055)	(0.080)	(0.055)
Local area FE	Yes	-	Yes	-	Yes	-	Yes	-	Yes	-
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.387	0.030	0.297	-0.010	0.374	0.007	0.391	0.015	0.309	0.045
Within R^2	0.003	-0.019	-0.001	-0.043	0.014	0.005	0.005	-0.005	-0.003	-0.012
Observations	1,612	1,287	1,615	1,289	1,615	1,289	1,616	1,291	1,612	1,287
Diagnostic Tests										
F-statistic (1st stage)	86.448	56.833	86.448	56.833	86.448	56.833	86.448	56.833	86.448	56.833
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic	3.841	3.056	2.204	4.115	0.390	0.022	2.806	1.539	2.893	2.127
	(0.050)	(0.080)	(0.138)	(0.043)	(0.532)	(0.881)	(0.094)	(0.215)	(0.089)	(0.145)
Wooldridge test	10.365		1.384		3.083		3.594		13.528	
	(0.001)		(0.240)		(0.079)		(0.058)		(0.000)	
Breusch-Godfrey test	88.058	127.822	132.083	155.883	112.919	152.477	109.959	141.494	78.912	109.257
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note: The table shows OLS and IV estimates of the impact of lagged NHS allocations (1-year lag) on wellbeing measures with controls. Panel A reports OLS results; Panel B reports IV results where allocation per head is instrumented by the simulated allocation per head. Dummy variables are created for missing controls. Columns 1, 3, 5, 7, and 9 (FE) show estimates based on the two-way fixed effects models, while columns 2, 4, 6, 8, and 10 (FD) show estimates based on the first difference models. Standard errors are clustered at the CCG-intersection level and given in parentheses. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under clustered standard errors. The Wu-Hausman statistic tests for endogeneity. The Wooldridge test and the Breusch-Godfrey test checks for first-order and second-order serial correlation respectively. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Despite the models showing limited significance at the 5% threshold, the consistency in the direction and magnitude of coefficients across models provides suggestive evidence for a positive causal relationship between healthcare funding and wellbeing. OLS estimates (Table 6) establish a robust baseline by utilising the full seven-year dataset, showing a significant positive association with overall wellbeing around the 5% significance threshold ($\beta = 0.362$, $p = 0.054$). When correcting for endogeneity with IV models, the estimated magnitude of the causal effect increases substantially, as predicted in my conceptual framework. The IV coefficients across specifications are consistently larger in magnitude than the OLS coefficients for all wellbeing measures (except worthwhile), as shown in following robustness checks. Although the significance of IV-FE estimates for individual wellbeing measures falls to around 15%, the lagged effect on overall wellbeing remains significant at a 10% level ($p = 0.083$). For IV-FD estimates, overall wellbeing and happiness indicate potential for a much larger lagged positive effect. When comparing OLS and

IV estimates, the wider intervals estimated by IV reflect the trade-off between efficiency and consistency under sufficient instrument strength (high first-stage F-statistics). IV estimators are consistent but naturally less efficient than OLS (Wooldridge, 2020). This precision issue is further exacerbated by the one-year sample reduction required for instrument simulation. The lack of significance at 5% thresholds in IV models likely reflects limited precision, rather than an absence of effect. In sum, my results collectively support for a positive effect on overall wellbeing from increased healthcare funding, which is primarily driven by lagged increases in happiness and life satisfaction, and an immediate fall in anxiety.

5.7. Robustness Checks

I perform two robustness checks on my findings and report the results with figures 3 and 4[24]. Firstly, I estimate the models without the use of any controls, X_{at}. Secondly, I estimate the results for different base years used to construct the simulated instrument. There are three available base years for refreshed CCG weighted population data: 2013 (for allocations 2014/15), 2015 (for allocations 2016/17), and 2018 (for allocations 2019/20). I use 2015 as my main base year specification because it optimally balances instrument relevance with data availability. It is close enough to the actual baseline in 2013/14 to maintain relevance, while having new data for the new allocation rule components. Base year 2018 has lower instrument strength since it is the furthest away from the starting baseline, despite having the most complete data for new components. Base year 2013 suffers from missing new allocation components and the greatest distance from later-year allocations, resulting in the weakest first-stage relationship. My results verify that base year 2015 is the strongest instrument, followed by base year 2018. Results from using the 2013 base year are deemed unreliable and grayed out in the figures due to larger uncertainty based on the above issues. Appendix B details the differences in constructing the instrument under different base years, then compares and explains the difference in instrument strength. Appendix C shows both the first-stage and second-stage results for the robustness checks when using instruments with different base years, together with alternative standard error clustering.

The results presented above are robust. The estimated coefficients for both avoidable mortality and subjective wellbeing remain stable in magnitude and confidence across a wide range of model specifications. The benefits of increased allocations are not driven by model selection or instrument base years, since the links to different measures persists when not controlling for local factors or when using simulated allocations with different base years. This stability provides evidence that allocations have a causal effect on reducing avoidable mortality and improving subjective wellbeing.

[24]The scale for anxiety coefficients are flipped to enable easier magnitude comparisons with other wellbeing measures.

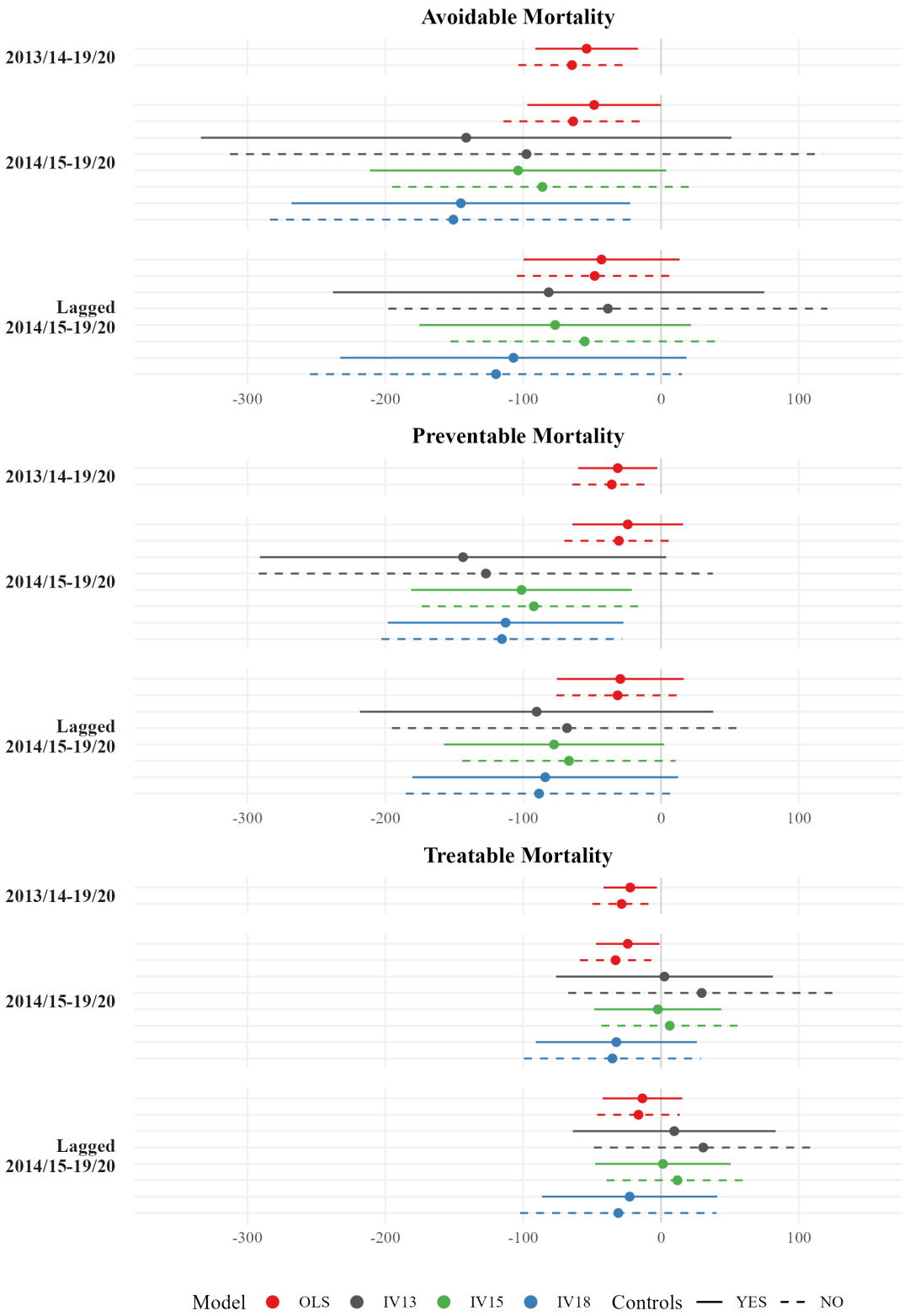


Figure 3: Regression Estimates for Avoidable Mortality Measures (FE Models, 90% CI)

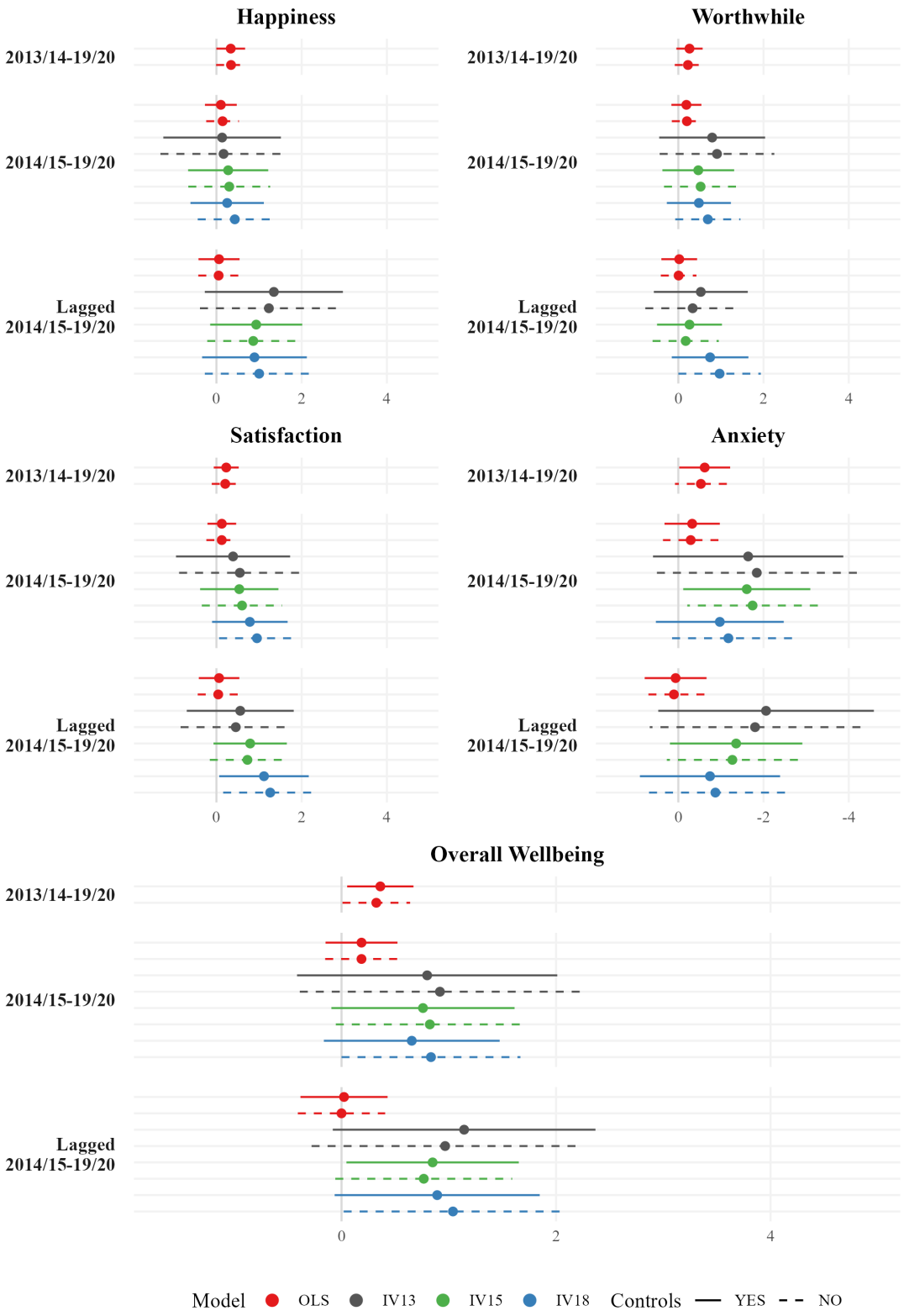


Figure 4: Regression Estimates for Wellbeing Measures (FE Models, 90% CI)

6. Interpretation and Discussion

The results for avoidable mortality and wellbeing collectively validate the proposed conceptual framework. The evidence indicates that budget allocations drive resource availability and are successfully converted into improved healthcare services h_{at}, which in turn generate positive returns on health outcomes o_{at} (measured with avoidable mortality) and subjective wellbeing w_{at}.

The results are not directly comparable with existing literature due to differences in estimation approach. Existing studies typically estimate long-run spending returns using cross-sectional designs that capture structural equilibrium effects. In contrast, this paper estimates short-run allocation returns using a panel design within an intention-to-treat (ITT) framework, focusing on the immediate policy-relevant question: how do outcomes respond to changes in their allocated budgets. Despite these frameworks addressing different questions, I provide descriptive comparisons below to contextualise the magnitude of my estimates.

6.1. Avoidable Mortality Results

Comparison with Existing Literature on Population Mortality

I convert coefficients to elasticities at the sample mean to facilitate comparison with existing literature. Based on the above model selection, I find that a 1% increase in allocation causally reduces preventable mortality by 0.899% and overall avoidable mortality by 0.583% in the short run[25].

In addition, I descriptively convert the elasticity of avoidable mortality into population mortality at -0.13. This figure is a lower-bound estimate, which (unrealistically) assumes the causal impact of funding on total deaths is driven entirely by avoidable deaths (which constitutes approximately 23% of total mortality[26]. This lower-bound estimate (-0.13) is closely aligned with the mean spending elasticity of -0.12 to -0.14 in the global meta regression analysis of 65 studies by Gallet and Doucouliagos (2017).

However, my short-run estimate (-0.13) is significantly smaller in magnitude than those established in UK studies (Claxton et al., 2015; Andrews et al., 2017; Martin et al., 2021a,b, 2023). These studies estimate that the NHS spending elasticity of population mortality ranges from -0.5 to -1.1. This divergence is consistent with two differences in methodology: nature of the treatment variable and time horizons (Kuh, 1959; Angrist and Pischke, 2009; Baltagi and Griffin, 1984).

[25] See footnote 1

[26] Figure based on data from the Office for National Statistics (2025j), Avoidable mortality in England and Wales. The proportion of avoidable deaths relative to total mortality has remained approximately 23% between 2013 and 2020.

First, the nature of the treatment variable used offers insight on the difference in elasticities. Existing literature estimates the actual spending (the structural parameter), using instrumental variables. Their estimates capture the biological return for every pound spent. In contrast, this study estimates the Intent-to-Treat (ITT) effect of the allocation (the policy lever). Allocations are mediated by local decision-making, which can cover existing deficits or administrative frictions before reaching patient care. Consequently, the ITT elasticity is naturally smaller than the structural spending elasticity (in absolute terms). Nonetheless, as discussed previously, the ITT estimate is often more relevant for central government decision-making, as it reflects the return to the budget allocations to healthcare, rather than perfect tracking of actual spending.

Second, existing literature relies mostly on cross-sectional instrumental variable identification, which reflects quasi long-run equilibrium elasticities. Their estimates capture the cumulative effects of historical spending patterns, assuming full factor adjustment as outcomes and funding stabilised. Contrastingly, this paper uses within-area variation over time, isolating short-run response to marginal spending changes. As established in economic theory, short-run supply is less elastic than in the long run due to fixed factors of production (Samuelson, 1983; Varian and Melitz, 2024). In the NHS, highly skilled workforce and capital inputs such as hospital infrastructure and equipment are largely fixed in the short term. Therefore, my short-run estimates are naturally smaller, as it excludes the cumulative long-run adjustment effects.

These findings suggest that policymakers should carefully distinguish between short-run policy effects and long-run equilibrium returns, when forecasting the return on healthcare funding. Relying on cross-sectional spending estimates risks overestimation of the immediate returns of funding on mortality, given that my estimated short-run ITT elasticities are significantly lower than its long-run structural counterparts. As a result, annual spending review, performance targets, or impact assessments should be calibrated to these short-run responses to funding. Rather interpreted as inefficiencies, these results suggest that the full return on investment only materialises after adjustment in the longer term.

Preventable vs Treatable Mortality

Furthermore, preventable mortality is significantly more responsive to funding than treatable mortality. To investigate, I descriptively convert results from existing literature into avoidable mortality results for comparison using a three-stage meta-analysis. I select papers studying NHS cost effectiveness through spending elasticity of mortality in two eras: Primary Care Trust (PCT) era in (2003/04–2012/13) (Martin et al., 2008a,b,c; Claxton et al., 2015; Lomas et al., 2019; Martin et al., 2021b) and Clinical Commissioning Group (CCG) era (2013/14–2019/20) (Martin et al.,

2022, 2023). First, I gather their estimates of total spending elasticities of mortality[27] for different Programme Budget Categories (PBCs), which classifies NHS spending by medical condition or care programme (e.g. cancer, circulatory problems). I pool estimates within each paper into one independent estimate per paper per PBC, using a fixed-effect inverse-variance weighted mean. Second, I conduct a random-effects model[28] to synthesise paper-level estimates into an overall estimate per PBC. This mix strategy (within-paper fixed effects and across-paper random effects) accounts for replicated estimation within-paper across different years or specifications, while addressing between-study heterogeneity (Borenstein et al., 2009). Third, I aggregate PBC elasticities into preventable and treatable mortality elasticity ε_{type}, using a share-weighted summation method as follows:

$$\varepsilon_{type} = \sum_{PBC=1}^{N} (\varepsilon_{PBC} \times w_{PBC,\,type})$$

where ε_{PBC} represents the PBC elasticity for a specific condition (e.g., circulatory problems) and $w_{PBC,type}$ represents the share of total preventable or treatable deaths attributed to the PBC condition[29]. I focus on the five major PBCs that drive most avoidable deaths[30] and for which elasticity estimates exist in both eras. To obtain robust standard errors under unknown covariance between PBC elasticities, I employ a non-parametric bootstrap clustered at the study level with 1000 iterations (Efron and Tibshirani, 1994)[31].

[27] Total spending elasticity of mortality refers to the % change in mortality rates for the PBC resulting from a 1% change in total healthcare spending. Some studies do not directly report total spending elasticities, rather expenditure and outcome elasticities (i.e. the % change in PBC spending from a 1% change in total healthcare spending and the % change in mortality rates from a 1% change in PBC spending). For such studies, I apply chain rule to calculate total elasticities from expenditure and outcome elasticities ($\varepsilon_{Total} = \varepsilon_{Spending} \times \varepsilon_{Outcome}$). Corresponding standard errors are calculated using Goodman's (Goodman, 1960) exact variance formula for the product of independent random variables, ensuring a conservative estimate of uncertainty. Following standard practice in applied settings where covariance matrices are unavailable, I assume expenditure and outcome elasticities are independent (Briggs et al., 2006).

[28] I use Restricted Maximum Likelihood (REML) estimator (Veroniki et al., 2016) to minimise bias and the Knapp-Hartung (Knapp and Hartung, 2003) adjustment to standard errors, accounting for the small number of studies ($N = 6$ for PCT, $N = 2$ for CCG). Additionally, I use Nelder-Mead (Nelder and Mead, 1965) optimizer to ensure robust model convergence under low heterogeneity in PBC elasticity.

[29] The share of total preventable or treatable deaths attributed to the PBC condition is calculated using ONS (Office for National Statistics, 2025c) avoidable mortality by cause in England. The causes are mapped to PBC conditions and shares are calculated for each era. The five selected PBCs are renormalized such that their weights sum to 1 within the closed system of this analysis. I treat these shares as fixed constants, as they are derived from population-level registry data rather than sample estimates.

[30] These five causes of avoidable mortality (related to the five major PBCs) account for around 87% to 90% of all avoidable mortality.

[31] In each bootstrap iteration, papers (not individual estimates) are resampled with replacement, and the entire three-stage meta-analysis is repeated. This procedure empirically derives the standard error of the aggregate elasticity, accounting for both the uncertainty of individual estimates and the covariance structure between diseases, without relying on asymptotic normality assumptions.

Table 11: Outcome Elasticities: Meta-Analysis with Share-Weighted Aggregation

	Outcome Elasticities	
	PCT Era (2003/04–2012/13)	CCG Era (2013/14–2019/20)
	(1)	(2)
Programme Budgeting Category		
Cancer	-0.297***	-0.519***
	(0.043)	(0.061)
Circulatory problems	-1.248***	-1.129***
	(0.196)	(0.134)
Respiratory problems	-1.508***	-2.436***
	(0.214)	(0.147)
Gastrointestinal problems	-1.370***	-3.897***
	(0.261)	(0.098)
Infectious diseases	-0.513	-1.539***
	(0.345)	(0.213)
Aggregated Total Spending Elasticity		
Preventable Mortality	-0.877***	-1.538***
	(0.076)	(0.088)
Treatable Mortality	-0.940***	-1.136***
	(0.094)	(0.101)

Note: Columns 1 and 2 show the results of the meta-regression analysis for the selected Programme Budgeting Categories under different eras. Top panel displays elasticities by the selected PBC; lower panel displays the weighted aggregate elasticity for Preventable and Treatable mortality. Standard errors (in parentheses) for the aggregate elasticities were calculated using non-parametric bootstrapping (n = 1000 iterations). $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

This meta-analysis only serves as descriptive benchmarks, intended to contextualise the primary results against the existing literature. It is by no means rigorous, since PBC conditions do not perfectly relate to the reported classification within preventable or treatable mortality. Table 11[32] presents the results of aggregated elasticities for PBCs and avoidable mortality types.

The meta-analysis results indicate that preventable mortality elasticity (-1.538) is significantly

[32]The standard errors for aggregated preventable or treatable mortality elasticities appear relatively precise, despite limited number of studies. This stability is explained by linear combination of components has lower variance than its individual components, unless components errors are all positively correlated (Wooldridge, 2020). My aggregation relates to the 'portfolio effect' (Koop, 2006) where idiosyncratic noise across distinct disease categories cancels out upon aggregation.

higher than that of treatable mortality (-1.136) in the CCG era (2013/14-2019/20). This difference is consistent, but smaller compared to my panel estimation. This evidence may reflect three potential concurrent mechanisms: fundamental effectiveness, policy focus, and the composition of healthcare services commissioned on the margin.

First, the results suggest a higher marginal productivity for preventative care relative to treatment care in the CCG era, as such difference is observed in both short and long run estimates. Low marginal returns on treatable mortality relate to the "flat of the curve" hypothesis, where curative healthcare funding have reached a point with limited returns (Fisher et al., 2000, 2003; Thornton, 2002, 2011; Hensher et al., 2024; Finkelstein and Gentzkow, 2026). However, high marginal returns on preventable mortality suggest that preventative care likely remains on the steeper section of the healthcare production function (Mays and Smith, 2011; Bradley et al., 2016; Masters et al., 2017). In the context of the English NHS, Martin et al. (2021a) find that the marginal productivity of public health spending (mostly preventative) is around 3.6 times higher than that of CCG expenditure. My analysis extends the evidence, where even within CCG's funding, further marginal funding yield greater survival gains in preventative care (e.g. community health services or mental health services) than in curative care (e.g. elective hospital care or urgent care).

Second, the change in policy focus over time can offer some insights for the observed difference in elasticities. During the PCT era (2003/04–2012/13), such difference in elasticities between preventable and treatable mortality was not significant in the PCT era (-0.877 and -0.940). Comparing with results for the CCG era (2013/14–2019/20), the higher relative responsiveness in treatable mortality in PCT era can reflect the policy focus on acute waiting time targets and hospital capacity expansion, as outlined in the NHS Plan 2000 (Department of Health, 2000). Propper et. al. (Propper et al., 2008) and Gruber et. al. (Gruber et al., 2023) confirm that these policies are effective at reducing waiting time and mortality, hence enhancing treatable elasticities. In contrast, during the CCG era, the Five Year Forward View (NHS England, 2014) and Health and Social Care Act (Health and Social Care Act, 2012) strategically pivots the NHS to increase focus on prevention while enhancing service outsourcing. The comparative lower responsiveness of treatable mortality may reflect the impact of increased outsourcing to for-profit providers since 2013, which Goodair and Reeves (Goodair and Reeves, 2022) link to increased treatable mortality.

Third, the difference in funding elasticities for preventable and treatable mortality is greater than their corresponding difference in spending elasticities from the meta-analysis. This suggests that when marginal allocations increase, the composition of healthcare services commissioned also contributes to the preventable-treatable difference. This composition choice is likely constrained by the national tariff payment system, which establishes distinct payment mechanisms across services (Monitor, 2013; NHS England, 2020). For services that target treatable conditions, such as

elective hospital care and urgent care, payments to providers are primarily activity-based. CCGs pay providers per procedure or admission at prices set under the national tariffs (NHS England, 2022). While local price variation is permitted[33], payments remain largely demand-led, tracking realised activity at predetermined prices and constraining CCGs' ability to redirect resources. In contrast, for preventative services like mental health and community care, payments are set predominantly through fixed annual block contracts negotiated locally (Monitor, 2013). This payment asymmetry creates different responsiveness to marginal funding. Activity-based acute payments automatically absorb increases through rising demand, while block-contracted preventative services provided flexibility for CCGs to direct additional allocations toward upstream interventions. The larger difference in funding elasticity likely reflects this flexibility: CCGs can adjust block contracts when allocations change, while activity-based payments respond to demand fluctuations rather than funding availability. Therefore, marginal funding growth disproportionately enhances prevention-focused services, despite baseline spending remains dominated by acute care.

Limitations of the Allocation Formulae

My results highlight a potential limitation in CCG allocation rules. The formulae are sensitive to changes in preventable mortality but less in treatable mortality. This is because of the unmet needs adjustments in allocation rules are proxied by SMR<75, which is proportionally dominated by preventable causes. Allocations are mechanically prioritised for reducing preventable mortality over treatable mortality in percentage terms. Formulae cannot fully capture the specific resource requirements needed to drive reductions in treatable mortality. This might be problematic since a large proportion of CCG budgets are spent on curative care (i.e. acute or elective care services). Potential updates should ensure that resources are more precisely responsive to both preventable and treatable mortality by considering them separately and proportionally.

6.2. Wellbeing

My wellbeing results contribute to two strands of wellbeing literature. Firstly, they corroborate the well-established association between individual health status and subjective wellbeing (Dolan et al., 2008; Ngamaba et al., 2017; Mathentamo et al., 2024). Secondly, this paper adds to the emerging body of causal evidence linking healthcare funding to subjective wellbeing. My results are consistent with the positive causal effects from increased health spending identified by Kotakorpi and Laamanen (2010) in Finland, and Yang et al. (2022) in China. This paper demonstrates that

[33]There are two types of deviations from national tariffs. Nationally determined variation permits local prices to vary, reflecting regional cost differences, patient complexity, financial risk-sharing and responsibility sharing. Prices can also be set locally for service design innovation or integrated services (Monitor, 2013)

there are welfare gains from increased funding in a formula-based, universal healthcare system (i.e. the English NHS).

I conduct wellbeing cost-benefit analysis with cost per WELLBY, following the methodology proposed Frijters and Krekel (2021). Cost per WELLBY reflects the monetary cost to increase one unit of life satisfaction on a 0-to-10 scale for one person for one year (Frijters et al., 2024). The marginal cost of healthcare spending is estimated at £1,282 per WELLBY.[34] This primary estimate relies on the lagged IV-FE model (Table 10), which yields the most statistically significant coefficient for life satisfaction ($p = 0.131$). To verify these findings, I also estimate the "cost per OWELLBY". It is defined as the marginal cost to increase the overall wellbeing score by one unit on a 0–10 scale (as in equation 2)[35], instead of only in the life satisfaction measure. This is a more stringent standard since the overall index averages four separate indicators, achieving a one-point rise in the overall score effectively requires improving an individual's wellbeing by one point across every measure simultaneously or four points in one measure. My OLS and lagged IV-FE models return estimates at £2,762.4 and £1,195.5 per OWELLBY respectively.

My WELLBY estimate broadly aligns with the estimated marginal productivity of NHS at around £2,500 per WELLBY[36]. My estimate is significantly lower than the societal valuation of a unit of WELLBY at £13,000 per WELLBY, as outlined in UK government's Green Book 2022 (HM Treasury, 2021, 2022). Furthermore, my results verified the predictions in Rutten and Reed (Rutten and Reed, 2009) computable general equilibrium model. The model suggests that a rise in the National Health Service (NHS) budget can yield overall welfare gain through direct increases in population wellbeing. This suggests that reallocating additional resources to the NHS would likely yield a positive net social surplus, which could justify further investment in the NHS.

6.3. Limitations

This paper has several limitations that future work may be able to overcome. First, I limit my analysis to allocations on CCG commissioned, excluding spending for primary care or specialised services. Using total NHS funding enables comprehensive analysis of overall policy effects.

Second, while the instrument relies on the exogenous national formula, strict exogeneity in allocation rule updates cannot be guaranteed, despite such rule changes likely act as exogenous

[34]The WELLBY estimate is calculated by dividing a unit change in allocations (£1000) to the life satisfaction coefficient (Table 10 column 7), and then multiplied by the assumed discount rate at 1.5% (Frijters and Krekel, 2021).

[35]The OWELLBY denotes a one-unit increase in the composite Overall Wellbeing index (0–10). As specified in equation 2, the index is the unweighted mean of the four standardized ONS measures. Therefore, a one-unit change in the index mathematically implies a total of four-unit change across the underlying components (e.g., a one-unit improvement in all four measures).

[36]Derived from the implied cost per QALY at £15,000 into cost per WELLBY (Frijters and Krekel, 2021).

policy shocks at a local level theoretically. This may introduce conservative bias. Under sufficient instrument power, an improvement in the instrument can be achieved by only identifying and selecting strictly exogenous policy changes or using quasi-experimental settings.

Third, only simple descriptive interpretation is conducted within types of avoidable mortality and expenditure. Between-type empirical analysis will help policymakers decide where additional investment in NHS is most beneficial.

Fourth, subjective wellbeing results from IV models hover around marginal significance at the 10% level. Hence, causal interpretation from wellbeing estimates is less secure than for avoidable mortality. The wider intervals reflect the trade-off between efficiency and consistency. Future research can use longer or more granular panel data to potentially increase the confidence in causal inference between healthcare spending and wellbeing (Baltagi, 2021), such as using individual level datasets to difference out variation across individual (Ferrer-i Carbonell and Frijters, 2004).

7. Conclusion

In this paper, I examine the short-run causal effect of increased healthcare funding on avoidable mortality and subjective wellbeing in the English NHS. By exploiting variation driven by national updates to allocation rules, I construct a simulated instrument to overcome endogeneity challenges in panel data analysis. This paper provides new evidence that an increase in healthcare funding improves subjective wellbeing and reduces avoidable mortality, with gains driven primarily by reductions in preventable causes rather than treatable causes.

Methodologically, this paper contributes by adapting the simulated instrument strategy to the literature involving healthcare funding. I demonstrate that the "allocation rule" instruments in cross-sectional studies become weak and endogenous in a panel setting, as they are absorbed by fixed effects or correlate with changing local needs. The simulated instrument approach isolates variation driven solely by national allocation rule changes, which is independent of shifts in local demographics, healthcare needs, or outcomes. This strategy improves instrument power. Estimates are larger in magnitude than those obtained via OLS, highlighting the downward bias present without proper accounting for exogeneity in "needs-based" funding. Future work evaluating returns to formula-based healthcare systems can adapt this methodology for causal identification in panel settings.

Together with existing literature, this paper has three policy implications. First, the higher funding responsiveness of preventable mortality potentially suggests that further marginal returns of funding yields greater benefits in preventative care than curative care. Second, I identify the

limitation of mechanical prioritisation of resources for preventable mortality within formulae in the allocation rules, which could be addressed by future rule updates. Third, wellbeing cost-benefit analysis using WELLBY indicates additional spending on NHS yields net welfare gains for the population, which could justify further investment in the NHS.

References

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics*, 138(1):1–35.

Andrews, M., Elamin, O., Hall, A. R., Kyriakoulis, K., and Sutton, M. (2017). Inference in the presence of redundant moment conditions and the impact of government health expenditure on health outcomes in England. *Econometric Reviews*, 36(1-3):23–41.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, NJ.

Anwar, A., Hyder, S., Mohamed Nor, N., and Younis, M. (2023). Government health expenditures and health outcome nexus: A study on OECD countries. *Frontiers in Public Health*, 11.

Baltagi, B. H. (2021). *Econometric Analysis of Panel Data*. Springer Texts in Business and Economics. Springer International Publishing, Cham.

Baltagi, B. H. and Griffin, J. M. (1984). Short and Long Run Effects in Pooled Models. *International Economic Review*, 25(3):631–645.

Biasi, B. (2023). School Finance Equalization Increases Intergenerational Mobility. *Journal of Labor Economics*, 41(1):1–38.

Bokhari, F. A. S., Gai, Y., and Gottret, P. (2007). Government health expenditures and health outcomes. *Health Economics*, 16(3):257–273.

Bond, T. N. and Lang, K. (2019). The Sad Truth about Happiness Scales. *Journal of Political Economy*, 127(4):1629–1640.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. John Wiley & Sons, Ltd.

Bradley, E. H., Canavan, M., Rogan, E., Talbert-Slagle, K., Ndumele, C., Taylor, L., and Curry, L. A. (2016). Variation In Health Outcomes: The Role Of Spending On Social Services, Public Health, And Health Care, 2000–09. *Health Affairs*, 35(5):760–768.

Brazier, J. and Tsuchiya, A. (2015). Improving Cross-Sector Comparisons: Going Beyond the Health-Related QALY. *Applied Health Economics and Health Policy*, 13(6):557–565.

Briggs, A., Claxton, K., and Sculpher, M. (2006). *Decision Modelling For Health Economic Evaluation*. Oxford University Press.

Brown, T. T., Martinez-Gutierrez, M. S., and Navab, B. (2014). The impact of changes in county public health expenditures on general health in the population. *Health Economics, Policy and Law*, 9(3):251–269.

Card, D., Dobkin, C., and Maestas, N. (2008). The Impact of Nearly Universal Insurance Coverage on Health Care Utilization: Evidence from Medicare. *American Economic Review*, 98(5):2242–2258.

Claxton, K., Martin, S., Soares, M., Rice, N., Spackman, E., Hinde, S., Devlin, N., Smith, P. C., Smith, P., and Sculpher, M. (2015). Methods for the Estimation of the NICE Cost Effectiveness Threshold. *Health Technology Assessment*, 19(14):1–504.

Clement, D. M. (2023). The founding of the NHS: 75 years on – History of government. https://history.blog.gov.uk/2023/07/13/the-founding-of-the-nhs-75-years-on/.

Crémieux, P.-Y., Ouellette, P., and Pilon, C. (1999). Health care spending as determinants of health outcomes. *Health Economics*, 8(7):627–639.

Currie, J. and Gruber, J. (1996). Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women. *Journal of Political Economy*, 104(6):1263–1296.

de Meijer, C., Wouterse, B., Polder, J., and Koopmanschap, M. (2013). The effect of population aging on health expenditure growth: A critical review. *European Journal of Ageing*, 10(4):353–361.

Department for Energy Security and Net Zero (2025). Uk local authority and regional greenhouse gas emissions statistics, 2005 to 2023. Accredited official statistics.

Department of Health (2000). The NHS plan: A plan for investment, a plan for reform. Technical Report Cm 4818, The Stationery Office, London.

Department of Health (2012). The functions of Clinical Commissioning Groups. Technical report, Department of Health.

Department of Health and Social Care (2012). The NHS Constitution for England. Technical report, Department of Health and Social Care.

Dolan, P. (2008). Developing methods that really do value the 'Q' in the QALY. *Health Economics, Policy and Law*, 3(1):69–77.

Dolan, P., Peasgood, T., and White, M. (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology*, 29(1):94–122.

Doyle, J. J. (2011). Returns to Local-Area Health Care Spending: Evidence from Health Shocks to Patients Far from Home. *American Economic Journal: Applied Economics*, 3(3):221–243.

Efron, B. and Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. Chapman and Hall/CRC, New York.

Farahani, M., Subramanian, S. V., and Canning, D. (2010). Effects of state-level public spending on health on the mortality probability in India. *Health Economics*, 19(11):1361–1376.

Ferrer-i Carbonell, A. and Frijters, P. (2004). How Important is Methodology for the Estimates of the Determinants of Happiness? *The Economic Journal*, 114(497):641–659.

Filmer, D. and Pritchett, L. (1999). The impact of public spending on health: Does money matter? *Social Science & Medicine*, 49(10):1309–1323.

Finkelstein, A. and Gentzkow, M. (2026). Re-Examining Geographic Variation in Health and Health Care.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., Baicker, K., and Oregon Health Study Group (2012). The Oregon Health Insurance Experiment: Evidence from the First Year*. *The Quarterly Journal of Economics*, 127(3):1057–1106.

Fisher, E. S., Wennberg, D. E., Stukel, T. A., Gottlieb, D. J., Lucas, F. L., and Pinder, É. L. (2003). The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care. *Annals of Internal Medicine*, 138(4):288–298.

Fisher, E. S., Wennberg, J. E., Stukel, T. A., Skinner, J. S., Sharp, S. M., Freeman, J. L., and Gittelsohn, A. M. (2000). Associations among hospital capacity, utilization, and mortality of US Medicare beneficiaries, controlling for sociodemographic factors. *Health Services Research*, 34(6):1351–1362.

Frijters, P. and Krekel, C. (2021). *A Handbook for Wellbeing Policy-Making: History, Theory, Measurement, Implementation, and Examples*. Oxford University Press.

Frijters, P., Krekel, C., Sanchis, R., and Santini, Z. I. (2024). The WELLBY: A new measure of social value and progress. *Humanities and Social Sciences Communications*, 11(1):736.

Gainsbury, S. (2016). Feeling the crunch: NHS finances to 2020. Technical report.

Gallet, C. A. and Doucouliagos, H. (2017). The impact of healthcare spending on health outcomes: A meta-regression analysis. *Social Science & Medicine*, 179:9–17.

Goodair, B. and Reeves, A. (2022). Outsourcing health-care services to the private sector and treatable mortality rates in England, 2013–20: An observational study of NHS privatisation. *The Lancet Public Health*, 7(7):e638–e646.

Goodman, L. A. (1960). On the Exact Variance of Products. *Journal of the American Statistical Association*, 55(292):708–713.

Goodman-Bacon, A. (2018). Public Insurance and Mortality: Evidence from Medicaid Implementation. *Journal of Political Economy*, 126(1):216–262.

Griliches, Z. and Hausman, J. A. (1986). Errors in variables in panel data. *Journal of Econometrics*, 31(1):93–118.

Gruber, J., Hoe, T. P., and Stoye, G. (2023). Saving Lives by Tying Hands: The Unexpected Effects of Constraining Health Care Providers. *The Review of Economics and Statistics*, 105(1):1–19.

Gruber, J. and Saez, E. (2002). The elasticity of taxable income: Evidence and implications. *Journal of Public Economics*, 84(1):1–32.

Harker, R. (2019). NHS Funding: Clinical Commissioning Groups.

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica*, 46(6):1251–1271.

Health and Social Care Act (2012). Health and social care act 2012. c. 7.

Hensher, M., Blizzard, L., Campbell, J., Canny, B., Zimitat, C., and Palmer, A. (2024). Diminishing marginal returns and sufficiency in health-care resource use: An exploratory analysis of outcomes, expenditure, and emissions. *The Lancet Planetary Health*, 8(10):e744–e753.

Hitiris, T. and Posnett, J. (1992). The determinants and effects of health expenditure in developed countries. *Journal of Health Economics*, 11(2):173–181.

HM Treasury (2021). Wellbeing guidance for appraisal: Supplementary Green Book guidance. Technical report, HM Treasury.

HM Treasury (2022). *The Green Book: Central Government Guidance on Appraisal and Evaluation*. HM Treasury.

Karaman, S., Ürek, D., Bilgin Demir, İ., Uğurluoğlu, Ö., and Işık, O. (2020). The Impact of Healthcare Spending on Health Outcomes: New Evidence from OECD Countries. *Journal of Clinical Practice and Research*, 42(2):218.

Knapp, G. and Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, 22(17):2693–2710.

Koop, G. (2006). *Analysis of Financial Data*. John Wiley & Sons Inc, Chichester ; Hoboken, NJ.

Kornai, J. (2009). The soft budget constraint syndrome in the hospital sector. *International Journal of Health Care Finance and Economics*, 9(2):117–135.

Kotakorpi, K. and Laamanen, J.-P. (2010). Welfare State and Life Satisfaction: Evidence from Public Health Care. *Economica*, 77(307):565–583.

Kuh, E. (1959). The Validity of Cross-Sectionally Estimated Behavior Equations in Time Series Applications. *Econometrica*, 27(2):197–214.

LaVela, S. L. and Gallan, A. (2014). Evaluation and Measurement of Patient Experience.

Lee, H., Vlaev, I., King, D., Mayer, E., Darzi, A., and Dolan, P. (2013). Subjective well-being and the measurement of quality in healthcare. *Social Science & Medicine*, 99:27–34.

Lloyd-Sherlock, P. (2000). Population ageing in developed and developing regions: Implications for health policy. *Social Science & Medicine*, 51(6):887–895.

Lomas, J., Martin, S., and Claxton, K. (2019). Estimating the Marginal Productivity of the English National Health Service From 2003 to 2012. *Value in Health*, 22(9):995–1002.

Luce, B. R., Mauskopf, J., Sloan, F. A., Ostermann, J., and Paramore, L. C. (2006). The Return on Investment in Health Care: From 1980 to 2000. *Value in Health*, 9(3):146–156.

Martin, S., Claxton, K., Lomas, J., and Longo, F. (2022). How Responsive is Mortality to Locally Administered Healthcare Expenditure? Estimates for England for 2014/15. *Applied Health Economics and Health Policy*, 20(4):557–572.

Martin, S., Claxton, K., Lomas, J., and Longo, F. (2023). The impact of different types of NHS expenditure on health: Marginal cost per QALY estimates for England for 2016/17. *Health Policy*, 132:104800.

Martin, S., Lomas, J., Claxton, K., and Longo, F. (2021a). Causal impact of social care, public health and healthcare expenditure on mortality in England: Cross-sectional evidence for 2013/2014.

Martin, S., Lomas, J., Claxton, K., and Longo, F. (2021b). How Effective is Marginal Healthcare Expenditure? New Evidence from England for 2003/04 to 2012/13. *Applied Health Economics and Health Policy*, 19(6):885–903.

Martin, S., Rice, N., and Smith, P. C. (2008a). Does health care spending improve health outcomes? Evidence from English programme budgeting data. *Journal of Health Economics*, 27(4):826–842.

Martin, S., Rice, N., and Smith, P. C. (2008b). Further evidence on the link between health care spending and health outcomes in England. Technical report, The Health Foundation.

Martin, S., Rice, N., and Smith, P. C. (2008c). The link between health care spending and health outcomes for the new English Primary Care Trusts. Technical Report 42, Centre for Health Economics, University of York, York, UK.

Masters, R., Anwar, E., Collins, B., Cookson, R., and Capewell, S. (2017). Return on investment of public health interventions: A systematic review. *J Epidemiol Community Health*, 71(8):827–834.

Mathentamo, Q., Lawana, N., and Hlafa, B. (2024). Interrelationship between subjective wellbeing and health. *BMC Public Health*, 24(1):2213.

Mays, G. P. and Smith, S. A. (2011). Evidence Links Increases In Public Health Spending To Declines In Preventable Deaths. *Health Affairs*, 30(8):1585–1593.

Monitor (2013). 2014/15 National Tariff Payment System. Technical Report Publication code: IRG 34/13; NHS England Publications Gateway Reference 00883, Monitor, London. Presented under the provisions of the Health and Social Care Act 2012.

Moreno-Serra, R. and Smith, P. C. (2015). Broader Health Coverage is Good for the Nation's Health: Evidence from Country Level Panel Data. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 178(1):101–124.

Nakamura, R., Lomas, J. R. S., Claxton, K. P., Bokhari, F., Moreno Serra, R. A., and Suhrcke, M. E. (2016). Assessing the impact of health care expenditures on mortality using cross-country data. https://eprints.whiterose.ac.uk/id/eprint/135816/.

National Audit Office (2018). A review of the role and costs of clinical commissioning groups - NAO report.

National Audit Office (2020). NHS financial management and sustainability 2020 - NAO report. Technical report.

Nelder, J. A. and Mead, R. (1965). A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313.

Ngamaba, K. H., Panagioti, M., and Armitage, C. J. (2017). How strongly related are health status and subjective well-being? Systematic review and meta-analysis. *European Journal of Public Health*, 27(5):879–885.

NHS England (2014). Five Year Forward View. Technical report, NHS England, London.

NHS England (2020). Understanding and using the national tariff. Technical report, NHS England.

NHS England (2022). SUS+ PbR reference manual. Technical report, NHS England Digital. Last edited: July 5, 2022.

NHS England (2025). Previous allocations. Repository of NHS revenue allocations for financial years 2013/14 to 2025/26.

Nolte, E. and McKee, M. (2004). *Does Health Care Save Lives? Avoidable Mortality Revisited.* The Nuffield Trust.

O'Donnell, G. and Oswald, A. J. (2015). National well-being policy and a weighted approach to human feelings. *Ecological Economics*, 120:59–70.

OECD (2023). Avoidable mortality (preventable and treatable): Health at a Glance 2023. https://www.oecd.org/en/publications/health-at-a-glance-2023_7a7afb35-en/full-report/avoidable-mortality-preventable-and-treatable_e7407977.html.

Office for National Statistics (2020). Annual personal well-being estimates: April 2019 to march 2020 – local authority update.

Office for National Statistics (2024). Modelled unemployment rate: 2004 to 2023. Official statistics in development.

Office for National Statistics (2025a). Avoidable mortality by local authorities in england and unitary authorities in wales.

Office for National Statistics (2025b). Avoidable mortality in england and wales.

Office for National Statistics (2025c). Avoidable mortality in England and Wales.

Office for National Statistics (2025d). Avoidable mortality in the UK QMI. Methodology, Office for National Statistics.

Office for National Statistics (2025e). Crime severity score: Year ending march 2025.

Office for National Statistics (2025f). First-time buyer mortgage sales: 2013 to 2024.

Office for National Statistics (2025g). Lower layer super output area population estimates (supporting information). Mid-2022 revised to mid-2024 edition.

Office for National Statistics (2025h). Persistent absences for all pupils: 2006-07 to 2023-24. Department for Education data, available via ONS Explore Local Statistics.

Office for National Statistics (2025i). Personal well-being in the uk QMI. Quality and Methodology Information, last revised 14 Jan 2025.

Office for National Statistics (2025j). Previous releases for avoidable mortality in England and Wales. Index of statistical bulletins; latest release published 28 April 2025.

Office for National Statistics (2025k). Regional gross disposable household income: local authorities.

Propper, C., Sutton, M., Whitnall, C., and Windmeijer, F. (2008). Did 'Targets and Terror' Reduce Waiting Times in England for Hospital Care? *The B.E. Journal of Economic Analysis & Policy*, 8(2):J.

Rahman, M. M., Khanam, R., and Rahman, M. (2018). Health care expenditure and health outcome nexus: New evidence from the SAARC-ASEAN region. *Globalization and Health*, 14(1):113.

Ray, D. and Linden, M. (2020). Health expenditure, longevity, and child mortality: Dynamic panel data approach with global data. *International Journal of Health Economics and Management*, 20(1):99–119.

Riedl, M. and Geishecker, I. (2014). Keep it simple: Estimation strategies for ordered response models with fixed effects. *Journal of Applied Statistics*, 41(11):2358–2374.

Roodman, D. (2009). A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics*, 71(1):135–158.

Rutten, M. and Reed, G. (2009). A comparative analysis of some policy options to reduce rationing in the UK's NHS: Lessons from a general equilibrium model incorporating positive health effects. *Journal of Health Economics*, 28(1):221–233.

Samuelson, P. A. (1983). *Foundations of Economic Analysis*. Harvard University Press.

Sohn, H. (2017). Medicaid's lasting impressions: Population health and insurance at birth. *Social Science & Medicine*, 177:205–212.

Staiger, D. and Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, 65(3):557–586.

Thornton, J. (2002). Estimating a health production function for the US: Some new evidence. *Applied Economics*, 34(1):59–62.

Thornton, J. A. (2011). Does more medical care improve population health? New evidence for an old controversy. *Applied Economics*, 43(24):3325–3336.

Varian, H. R. and Melitz, M. (2024). *Intermediate Microeconomics with Calculus: A Modern Approach*. W. W. Norton & Company.

Veroniki, A. A., Jackson, D., Viechtbauer, W., Bender, R., Bowden, J., Knapp, G., Kuss, O., Higgins, J. P., Langan, D., and Salanti, G. (2016). Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Research Synthesis Methods*, 7(1):55–79.

Wooldridge, J. M. (2020). *Introductory Econometrics: A Modern Approach*. Cengage Learning, Boston, MA, seventh edition edition.

World Health Organization (2025). Global Health Expenditure Database.

Wu, D.-M. (1973). Alternative Tests of Independence between Stochastic Regressors and Disturbances. *Econometrica*, 41(4):733–750.

Yang, Y., Zhao, L., and Cui, F. (2022). How Does Public Health Investment Affect Subjective Well-Being? Empirical Evidence from China. *International Journal of Environmental Research and Public Health*, 19(9):5035.

Appendix

A. CCG Allocation Rules

The mandate to NHS England requires a transparent allocations process based on the objective of equal opportunity of access for equal need. The following information is summarised from the "Technical Guide to Allocation Formulae and Pace of Change" documents published on the NHSE website (NHS England, 2025).

The NHSE Board meets and approves the allocations to CCGs based on the allocation formulae for upcoming fiscal years. An independent, expert, technical committee, the Advisory Committee on Resource Allocation (ACRA), makes recommendations to NHSE on the target formula for NHS allocations. In the study period, there are three major allocation rounds that outline, revise and finalise the principles, formulae and parameters for funding allocations for upcoming years: (1) 2014/15-2015/16, (2) 2016/17-2018/19 and (3) 2019/20. There are also updates, revisions or adjustments within allocation rounds.

As outlined in the technical guides, the four main steps in setting allocations for CCG responsible services (CCG core services) have been constant across allocation rounds, as outlined in section 2.3. This appendix outlines in detail the allocation rules (i.e. the formulae) and their changes over time.

A.1. Target Allocations

Target allocations are calculated with the target formula, which estimates the relative need and relative unavoidable costs between CCG areas for healthcare services.

First, weighted populations for target allocations are calculated for different healthcare services, which are (1) general and acute, (2) mental health, (3) maternity, (4) prescribing, and (5) community services (only introduced in 2019/20). Inputs for the weighted population are CCG population estimates and weights. CCG registered population is estimated from the registered patient list from all CCG's member GP practices[i]. The NHSE or external bodies (e.g. the Nuffield Trust) estimate the weights "need per head" for each CCG-age-sex group through modelling. The weight reflects healthcare needs for the specific age-sex group for a CCG, estimated with explanatory factors such

[i]For the allocation rounds studied, the population level used is one year before the starting year of the allocation round. The overall weighted population is calculated using 2015/16 data when setting 2016/17–2018/19, before being projected forward.

as demographics, ethnicity, past healthcare use, morbidity and costs. Weighted populations for a specific healthcare service (e.g. maternity) are calculated by multiplying the population of each age-sex group by its corresponding weight, then summing these products to the CCG level. Over time, the "need per head" weights are updated with data refreshes and re-running the modelling, introducing endogeneity to allocation which depends on evolving needs.

Second, weighted populations for different healthcare services are then combined for weighted populations for the CCG. The shares for each different service are based on past national outturn spending, which is exogenous.

Third, two adjustments for unavoidable costs are then applied. (1) The Market Forces Factor adjusts for the unavoidable cost differences between areas due to their geographical location (i.e. running a hospital in a city centre faces higher unavoidable input costs than in other areas). (2) Emergency Ambulance Cost Adjustment adjusts for unavoidable variations in the cost of providing emergency ambulance services in different geographical areas, and sparsely populated and metropolitan areas. Since these adjustment factors vary over time, they reflect relative costs and ambulance availability, and therefore are endogenous.

Fourth, weighted populations for unmet needs are calculated and combined. Unmet need is based on the standardised mortality ratio for those under 75 years of age (SMR < 75) at small area (middle layer super output area) level, which considers the inequality in health outcomes within and between CCGs. Based on the SMR < 75 value, each small area is grouped. Each group is given a weight per head, where groups with high SMR < 75 values are given higher weights[ii]. Weighted population is then calculated by multiplying the weight per head with the population for each small area, and then summing to the CCG level. Weighted populations for unmet needs are combined with the above adjusted weighted populations for the CCG. The share of unmet needs remains at 10% within the study period. Because this mortality metric is directly correlated with my outcome variables, I treat this component as endogenous.

Fifth, unavoidable costs of remoteness adjustments are applied. Note that this adjustment is only introduced since 2016/17 allocations. This introduction is an exogenous allocation rule change. CCGs with remote hospitals or relatively low activity levels are identified and given extra adjustment funding. The adjustment funding is then converted back to weighted populations, relative to the size of the total allocations before this adjustment. Total weighted populations including remoteness adjustment is then calculated.

Sixth, the weighted populations are projected forward by applying the estimated percentage growth of CCG registered populations, reflecting endogenous demographics change. This is the

[ii] There are minor adjustments for the calculation of weighted populations for unmet needs over time. There are changes to the number of groups for different SMR < 75 values and the distribution of the weights per head values.

final CCG core services weighted population $wPop_i$.

Exact population and project calculations by allocation rounds varies slightly. Nonetheless, such changes are filtered out by the fixed population level in the simulated instruments.

Finally, the CCGs' monetary (closing) target allocations are calculated by multiplying CCG's share of the total weighted population by the national budget for core services for the financial year.

A.2. Baselines

Baseline for CCG i for the year t are carried forward from previous year's allocations with adjustments.

$$Baseline_{i,t} = SAllocation_{i,t-1} + adjustments_{i,t} \tag{A.1}$$

A.3. Opening Distance from Target

Opening distance from target (DfT) is the proportional difference between baseline and opening target. First, opening target is calculated by multiplying CCG's current share of total weighted population by the sum of baselines (i.e. the total of previous year's allocation in England).

$$OpeningTarget_{i,t} = \frac{wPop_{i,t} \sum_i Baseline_{i,t}}{\sum_i wPop_{i,t}} \tag{A.2}$$

Then, opening distance from target for CCG i is calculated with:

$$DfT_{i,t} = \frac{Baseline_{i,t}}{OpeningTarget_{i,t}} - 1 \tag{A.3}$$

Opening distance from target in per head terms is also calculated in the same way with per head values of target allocations and baselines.

A.4. Pace of Change Policy

Actual allocations are derived from target allocations through the pace of change policy. The policy sets how far CCGs are moved closer to their target allocation each year from their baselines through differential growth. It aims to balance the need for funding stability and the objective of moving furthest under-target CCGs closer to their target. Since there are major changes in the pace of change policy over the study period, I will proceed to describe the pace of change policy for

each allocation round separately[iii]. For notations, a subscript i will indicate CCG-specific variables (endogenous), while nationally determined parameters do not have the subscript (exogenous). Note that all variables below are time-varying (i.e. time subscripts t should be present), but I omit them to avoid redundancy in the following notation.

A.4.1 2014/15-2015/16 Allocation Round

The pace of change policy for this round has three steps. First, a "pace of change" total growth rate $g_{PoC,i}$ for CCG i is calculated using CCG population growth rate $g_{pop,i}$ and opening distance from target per head $DfTph_i$ as follows:

$$g_{PoC,i} = (1 + g_{pop,i})(1 + g_{policy,i}) - 1 \tag{A.4}$$

where:

$$g_{policy,i} = \begin{cases} g_{maxph} & \text{if } DfTph_i < -5\% \\ g_{ph} & \text{if } DfTph_i > g_{threshold} \\ g_{maxph} - 5\% - DfTph_i & \text{otherwise} \end{cases} \tag{A.5}$$

and

$$g_{threshold} = g_{maxph} - 5\% - g_{pc} \tag{A.6}$$

where g_{maxph} refers to the maximum per head allocation growth rate for the most under-target CCGs (with $oDfTph < -5\%$), while g_{ph} is the minimum per head allocation growth for all CCGs. g_{PoC} represents the total allocation growth of CCG considering its population growth and per head growth (depending on $oDfTph_i$).

Second, a floor and celling are introduced for the total actual allocation growth for the CCG. It ensures all CCG receives at least a total allocation growth of $g_{mingrowth}$, while the most over-target CCGs' total allocation growth are capped at $g_{mingrowth}$. The actual growth $g_{actual,i}$ for CCGs are given by:

$$g_{celling,i} = \begin{cases} \min(g_{mingrowth}, g_{PoC,i}) & \text{if } DfTph > 5\% \\ g_{PoC,i} & \text{otherwise} \end{cases} \tag{A.7}$$

$$g_{actual,i} = \max(g_{mingrowth}, g_{celling,i}) \tag{A.8}$$

Where $g_{celling,i}$ caps CCG's growth with opening distance from target greater than 5% at the total

[iii] It is important to note the distinction between total growth rates and per head growth rate. Total growth rates refer to the percentage increase in CCG's total allocations. Per head growth rates refer to the percentage increase in CCG's per head allocations.

allocation growth rate $g_{mingrowth}$. All CCGs are then ensured to have at least a total allocation growth of $g_{mingrowth}$.

Finally, the actual allocation *Allocation* is given by:

$$Allocation_i = (1 + g_{actual,i}) \times Baseline_i \tag{A.9}$$

A.4.2 2016/17-2018/19 Allocation Round

Starting from this allocation round, the pace of change policy not only considers targets and baselines from CCG core services, but also targets and baselines from primary medical care and specialised services. This aims to support co-commissioning policy, where CCGs also take part in commissioning primary care services. This change allows a better reflection of the overall resources that a CCG has, acknowledging that a CCG area can better cope with being below target if some of the other streams are funded above target.

From now on, subscripts of $Core$ are in place for all CCG core services variables (e.g. $Allocation_{Core,i}$ and $DfTph_{Core,i}$, rather than $Allocation_i$ and $DfTph_i$) to ensure it is differentiated with primary care and specialised services variables (e.g. $Allocation_{PC,i}$ and $DfTph_{SPE,i}$).

The pace of change policy for this round has five main steps. For a comprehensive overview, a minimum allocation is first set to meet NHSE's policies for the three service streams (CCG core, primary medical and specialised services). Then, a pace of change is applied to the total resources available to the CCG, based on the sum of the three groups. When total allocation exceeds the sum of minimum allocations, the excess is disaggregated across CCG core services and primary medical services.

1: Minimum Allocation for CCG Core Services

First, per head pace of change growth is applied. A variable, $\theta_{Core,i}$, is calculated for CCGs to map the opening distance from target per head $DfTph_{Core,i}$ onto a unit interval using linear interpolation.

$$\theta_{Core,i} = \begin{cases} 0 & \text{if } DfTph_{Core,i} > DfTph_{Core,max} \\ 1 & \text{if } DfTph_{Core,i} < DfTph_{Core,min} \\ \frac{DfTph_{Core,i} - DfTph_{Core,min}}{DfTph_{Core,max} - DfTph_{Core,min}} & \text{otherwise} \end{cases} \tag{A.10}$$

Where $DfTph_{Core,max}$ refers to the threshold of $DfTph_{Core,i}$ beyond which minimum growth $g_{Core,PoC,min}$ is applied, and $DfTph_{Core,min}$ refers to the threshold of $DfTph_{Core,i}$ below which

maximum growth $g_{Core,PoC,max}$ is applied. When $DfTph_{Core,i}$ is between the thresholds, growth rates are also linearly interpolated. The per head pace of change growth $g_{Core,PoC,i}$ is therefore:

$$g_{Core,PoC,i} = \begin{cases} g_{Core,PoC,min} & \text{if } \theta_i = 0 \\ g_{Core,PoC,max} & \text{if } \theta_i = 1 \\ \frac{\theta_{Core,i}}{(g_{Core,PoC,max} - g_{Core,PoC,min})} + g_{Core,PoC,min} & \text{otherwise} \end{cases} \tag{A.11}$$

The total growth rate from this pace of change growth will therefore be:

$$g_{Core,totalPoC,i} = (1 + g_{pop,i})(1 + g_{Core,PoC,i}) - 1 \tag{A.12}$$

Second, a growth cap is then applied based on total opening Distance from Target $DfT_{Core,i}$. Another variable, $\rho_{Core,i}$, is calculated for CCGs to map the opening distance from target $DfT_{Core,i}$ into a 0-to-1 scale with linear interpolation.

$$\rho_{Core,i} = \begin{cases} 0 & \text{if } DfT_{Core,i} > 10\% \\ 1 & \text{if } DfT_{Core,i} < 5\% \\ \frac{DfT_{Core,i} - 5\%}{5\%} & \text{otherwise} \end{cases} \tag{A.13}$$

To reduce unnecessary notation, I report the minimum threshold (5%) and the maximum threshold (10%) with numbers, since they are fixed for all the years and service types in the allocation round. The total allocation growth after the growth cap $g_{Cap,i}$ is:

$$g_{Core,Cap,i} = \min\left\{g_{Core,totalPoC,i},\ g_{Core,totalPoC,i} - (g_{Core,totalPoC,i} - g_{Core,limit})(1 - \rho_i)\right\} \tag{A.14}$$

Where $g_{Core,limit}$ is the growth cap for CCGs with $DfT_{Core,i} > 10\%$.

Finally, the minimum allocation for CCG core services $Min_{Core,i}$ is given by:

$$Min_{Core,i} = (1 + g_{Core,Cap,i}) \times Baseline_{Core,i} \tag{A.15}$$

2: Minimum Allocation for Primary Medical Care Services (PC)

First, baselines $Baseline_{PC,i}$ and opening distances from target $DfT_{PC,i}$ for primary medical care services are calculated, using their weighted population for $wPop_{PC,i}$ (i.e. same as that for CCG services). Next, the steps of calculating the minimum allocation for primary medical care services are the same as the calculation of the CCG core services. Only the inputs (e.g. $wPop_i$) and the parameters (e.g. $g_{PoC,min}$) of CCG core services from above calculation should be replaced by the

inputs and parameters specific for primary care, which their subscripts are PC rather than $Core$ (e.g. $wPop_{PC,i}$ and $g_{PC,\ PoC,min}$). Hence, I obtain the minimum allocations for primary medical care services $Min_{PC,i}$.

3: Minimum Allocation for Specialised Services (SPE)

Similarly, baselines $Baseline_{SPE,i}$ and opening distances from target $DfT_{SPE,i}$ for specialised services are calculated using their weighted population $wPop_{SPE,i}$.

Minimum allocations for specialised services $Min_{SPE,i}$ are calculated differently than CCG core services and primary care services, where only a uniform per head growth $g_{SPE,\ minph}$ is applied with the following formula:

$$Min_{SPE,i} = \left[(1 + g_{SPE,\ minph})(1 + g_{pop,i}) - 1\right] \times Baseline_{SPE,i} \tag{A.16}$$

4: Total Place-Based Allocation

First, total minimum allocation $MinTotal_i$, total baselines $Baseline_{Total,i}$ and total opening distances from target $DfT_{Total,i}$ are calculated by summing the three services:

$$Min_{Total,i} = Min_{Core,i} + Min_{PC,i} + Min_{SPE,i} \tag{A.17}$$

$$Baseline_{Total,i} = Baseline_{Core,i} + Baseline_{PC,i} + Baseline_{SPE,i} \tag{A.18}$$

$$DfT_{Total,i} = DfT_{Core,i} + DfT_{PC,i} + DfT_{SPE,i} \tag{A.19}$$

Per head values are also calculated accordingly.

Next, the steps for setting total allocation are similar to setting the minimum allocation for CCG in this allocation round. Rather than using values or parameters for CCG core services calculation, input variables $Baseline_{Total,i}$, $DfT_{Total,i}$, $DfTpc_{Total,i}$, and parameters for total place-based allocations are used. Hence, the total allocation growth after the pace of change and growth cap $g_{Total,Cap,i}$ is obtained. Total place-based allocations are calculated by:

$$Allocation_{Total,i} = (1 + g_{Total,Cap,i}) \times Baseline_{Total,i} \tag{A.20}$$

$Allocation_{Total,i}$ is then compared to $MinTotal_i$, and the larger value is selected as the actual total allocation for the CCG $ActualAllocation_{Total,i}$, where:

$$ActualAllocation_{Total,i} = \max(Allocation_{Total,i}, MinTotal_i) \tag{A.21}$$

5: Disaggregation

If the total place-based allocations $Allocation_{Total,i}$ are smaller than the minima $Min_{Total,i}$, all the services use their minimum allocations ($Min_{Core,i}, Min_{PC,i}, Min_{SPE,i}$) as their actual allocations.

If $Allocation_{Total,i}$ are larger than $Min_{Total,i}$, the excess total allocations ($Excess_{Total,i} = Allocation_{Total,i} - Min_{Total,i}$) are distributed back to CCG core services and primary medical care services (not for specialised services). First, minimum distances from target for both services are calculated as follows:

$$MinDfT_{k,i} = Min_{k,i} - Target_{k,i} \quad \text{where } k \in \{Core,\ PC\} \tag{A.22}$$

Where $TargetCore_i$ and $TargetPC_i$ are the target allocation for CCG core services and primary care services calculated at the end of section A.1.

Next, scores $\omega_{Core,i}$ and $\omega_{PC,i}$ are calculated for the two service groups to represent their relative weights for disaggregating the excess total allocations, with the following formula:

$$\omega_{k,i} = \frac{Min_{k,i}}{(1 + MinDfT_{k,i})}/100 \quad \text{where } k \in \{Core,\ PC\} \tag{A.23}$$

Finally, the actual allocations for the two services after the disaggregation of excess total allocations are:

$$-k = \begin{cases} PC & \text{if } k = Core \\ Core & \text{if } k = PC \end{cases} \tag{A.24}$$

$$Excess_{k,i} = \begin{cases} 0 & \text{if } MinDfT_{k,i} > 0 \\ Excess_{Total,i} & \text{if } MinDfT_{k,i} < 0 \text{ and } MinDfT_{-k,i} > 0 \\ \frac{Excess_{Total,i}(\omega_{k,i})}{\omega_{k,i}+\omega_{-k,i}} & \text{if } MinDfT_{k,i} < 0 \text{ and } MinDfT_{-k,i} < 0 \end{cases} \tag{A.25}$$

$$Allocation_{k,i} = MinCore_{k,i} + Excess_{k,i} \tag{A.26}$$

A.4.3 2019/20 Allocation Round

1: Minimum Allocation for CCG Core Services

First, a CCG-specific minimum per capita growth rate $g_{Core,\ minph,i}$ for CCG core services is introduced. It depends on the CCG's opening distance from target per head $DfTph_{Core,i}$. The

growth rate $g_{Core,\ minph,i}$ is given by:

$$g_{Core,\ minph,i} = \begin{cases} g_{Core,\ minph,\ floor} & \text{if } DfTph_{Core,i} > 10\% \\ g_{Core,\ minph,\ cap} & \text{if } DfTph_{Core,i} < 5\% \\ (\tau_{Core,i})g_{Core,\ minph,\ cap} + (1 - \tau_{Core,i})g_{Core,\ minph,\ floor} & \text{otherwise} \end{cases} \tag{A.27}$$

$$\tau_{Core,i} = \frac{10\% - DfTph_{Core,i}}{5\%} \tag{A.28}$$

where $g_{Core,\ minph,\ floor}$ represents the "floor" value for minimum per capita growth for CCGs with $DfTph_{Core,i} > 10\%$, and $g_{Core,\ minph,\ cap}$ represents the "capped" value for minimum per capita growth for CCGs with $DfTph_{Core,i} < 5\%$.

Next, per head pace of change growth is applied. A variable, $\theta_{Core,i}$, is calculated for CCGs to map the opening distance from target per head $DfTph_{Core,i}$ onto a unit interval using linear interpolation.

$$\theta_{Core,i} = \begin{cases} 0 & \text{if } DfTph_{Core,i} > DfTph_{Core,max} \\ 1 & \text{if } DfTph_{Core,i} < DfTph_{Core,min} \\ \frac{DfTph_{Core,i} - DfTph_{Core,min}}{DfTph_{Core,max} - DfTph_{Core,min}} & \text{otherwise} \end{cases} \tag{A.29}$$

where $DfTph_{Core,max}$ refers to the threshold of $DfTph_{Core,i}$ beyond which minimum growth $g_{Core,\ minph,i}$ is applied, and $DfTph_{Core,min}$ refers to the threshold of $DfTph_{Core,i}$ below which a maximise growth value is applied. The per head pace of change growth $g_{Core,PoC,i}$ is therefore:

$$g_{Core,PoC,i} = \max(g_{Core,\ minph,i}, g_{Core,PoC,*,i}) \tag{A.30}$$

$$g_{Core,PoC,*,i} = \begin{cases} 0 & \text{if } \theta_i = 0 \\ \frac{(1+DfTph_{Core,target})CTargetPH_{Core,i}}{Baseline_{Core,i}} - 1 & \text{if } \theta_i = 1 \\ (\theta_{Core,i})(1 + g_{Core,\ quantum}) + (1 - \theta_{Core,i})g_{Core,\ minph,\ cap} & \text{otherwise} \end{cases} \tag{A.31}$$

where $g_{Core,PoC,*,i}$ is intermediate variables used to calculate the final per head pace of change growth value $g_{Core,PoC,i}$. The parameter $minDfT_{CCG,target}$ is the targeted minimum closing distance from target set by NHSE. The variable $CTargetPH_{Core,i}$ is closing target per capita for a CCG, as calculated in section 2.1. The parameter $g_{Core,\ quantum}$ is the overall national funding growth rate in CCG core services. The total growth rate from this pace of change growth will therefore be:

$$g_{Core,totalPoC,i} = (1 + g_{pop,i})(1 + g_{Core,PoC,i}) - 1 \tag{A.32}$$

Third, a minimum program growth at rate $g_{Core,mintotal}$ is ensured.

$$g_{minCore,i} = \max(g_{Core,mintotal}, g_{Core,totalPoC,i}) \tag{A.33}$$

Finally, the minimum allocation for CCG core services $Min_{Core,i}$ is given by:

$$Min_{Core,i} = (1 + g_{minCore,i}) \times Baseline_{Core,i} \tag{A.34}$$

2: Minimum Allocation for Primary Medical Care Services (PC)

Baselines $Baseline_{PC,i}$ and opening distances from target $DfT_{PC,i}$ for primary medical care services are calculated, using their weighted population for $wPop_{PC,i}$ (i.e. same as that for CCG services).

Next, the steps of calculating the minimum allocation for primary medical care services are the same as the calculation of the CCG core services in this allocation round. Only the inputs (e.g. $wPop_i$) and the parameters (e.g. $g_{PoC,min}$) of CCG core services from above calculation should be replaced by the inputs and parameters specific for primary care, which their subscripts are PC rather than $Core$ (e.g. $wPop_{PC,i}$ and $g_{PC,\ PoC,min}$). Hence, I obtain the minimum allocations for primary medical care services $Min_{PC,i}$.

3: Minimum Allocation for Specialised Services (SPE)

Minimum allocations for specialised services $Min_{SPE,i}$ are calculated as in the previous allocation round, where only a uniform per head growth $g_{SPE,\ minph}$ is applied with the following formula (same as past allocation round):

$$Min_{SPE,i} = \left[(1 + g_{SPE,\ minph})(1 + g_{pop,i}) - 1\right] \times Baseline_{SPE,i} \tag{A.35}$$

4: Total Place-Based Allocation

First, total minimum allocation $MinTotal_i$, total baselines $Baseline_{Total,i}$ and total opening distances from target $DfT_{Total,i}$ are calculated by summing the three services:

$$Min_{Total,i} = Min_{Core,i} + Min_{PC,i} + Min_{SPE,i} \tag{A.36}$$

$$Baseline_{Total,i} = Baseline_{Core,i} + Baseline_{PC,i} + Baseline_{SPE,i} \tag{A.37}$$

$$DfT_{Total,i} = DfT_{Core,i} + DfT_{PC,i} + DfT_{SPE,i} \tag{A.38}$$

Per head values are also calculated accordingly.

Next, the steps for setting total allocation are similar to setting the minimum allocation for CCG in this allocation round. Rather than using values or parameters for CCG core services calculation, input variables $Baseline_{Total,i}$, $DfT_{Total,i}$, $DfTpc_{Total,i}$, and parameters for total place-based allocations are used. Hence, the total allocation growth after the pace of change and growth cap $g_{Total,Cap,i}$ is obtained. Total place-based allocations are calculated by:

$$Allocation_{Total,i} = (1 + g_{minTotal,i}) \times Baseline_{Total,\ i} \tag{A.39}$$

$Allocation_{Total,i}$ is then compared to $MinTotal_i$, and the larger value is selected as the actual total allocation for the CCG $ActualAllocation_{Total,i}$, where:

$$ActualAllocation_{Total,i} = \max(Allocation_{Total,i}, MinTotal_i) \tag{A.40}$$

5: Disaggregation

The disaggregation rules for the 2019/20 allocation round are the same as that for the 2016/17-2018/19 allocation round.

A.5. Distinction of CCG's responsibilities

A clear distinction has to be made regarding the exclusion of primary care and specialised services in this study. These services are not considered in our treatment variable: allocations for CCG core services, $Allocation_i$; nor in our simulated instrument: simulated allocations for CCG core services, $SAllocation_i$. Meanwhile, targets and baselines of primary care and specialised services are only used to construct the instrument $SAllocation_i$ for CCG core services through pace of change policy.

Primary care and specialized services are excluded from the analysis due to limited baseline data availability prior to 2016 and funding instability arising from shifts in commissioning responsibilities (NHS England, 2014)[iv]. Consequently, the instrument focuses exclusively on CCG core services to ensure consistent panel stability.

[iv] Primary care co-commissioning models were introduced in 2015 to integrated out-of-hospital services tailored for local needs. CCGs were able to join voluntarily and choose their level of involvement, introducing a change in commissioning responsibilities of CCG on primary care.

B. Simulated Instrument Construction

I construct the instrument "simulated allocation per head", $SAllocation_{at}$. To satisfy exclusion restriction required for causal validity, the instrument has to isolate exogenous variation in allocations from local endogenous responses. As discussed in section 5.5, I assume that changes in the allocation rules and the total budget are exogenous. Subsequently, I explain how the simulated instrument construction isolate variation driven by national allocation rule changes from local endogenous responses, enabling causal identification.

The simulated instrument $SAllocation_{at}$ represents the counterfactual set of allocations a CCG would have received over time, if its demographics and needs had remained fixed at their base year[v] levels. I first explain the general construction methodology across all base years, then explore the minor adjustments required for the construction for different base years.

B.1. General Construction

I repeat the fixed inputs and exogenous changes in my simulated instrument here (as in section 5.4) for easy reference. The fixed inputs for removing endogeneity are:

1. Demographics and population changes
2. Needs weights for each age-CCG matrix
3. Unmet needs adjustments from SMR < 75
4. Supply-side adjustments
5. Any CCG-specific time-varying variables used in the pace of change policy are held constant at the base year.

Exogenous changes include:

1. Changes to the aggregate national budget for CCG allocations
2. Updates to national proportions that combine weighted populations for different areas of services (e.g. the split between maternity and prescription services based on past national expenditure)
3. The pace of change policy (i.e. the convergence of baseline to target allocations over time)
4. The introduction of new national equity adjustments

[v] This can be considered baseline levels of demographics and needs. The word choice of "base year" prevents confusion with "baselines" in allocation formulae, which represents the previous allocation levels.

B.1.1 Target allocations

All components of the CCG core services weighted populations for target allocations are held fixed, which are (1) general and acute, (2) mental health, (3) maternity, (4) prescribing, (5) community services (only introduced in 2019/20, to be discussed in Appendix section 3.2). The fixed components are the CCG population estimates (which evolves every year) and the need weights (which are updated every allocation round). Factors, such as demographics, ethnicity, past healthcare use, morbidity and costs, that explain healthcare needs and hence health outcomes are held constant at the base year to remove endogeneity (fixed inputs 1 & 2).

Next, the weighted populations for different healthcare services are then combined at varying shares. The shares for each different service are based on past national outturn spending (exogenous change 2).

Third and fourth, the adjustments factors for unavoidable costs and unmet needs (SMR < 75) are held constant. This step filters out varying cost factors and mortality rates that confounds allocation, since these variables are likely to directly drive outcomes in wellbeing and avoidable mortality in the area (fixed inputs 3 & 4). I do not consider the minor changes in calculation of weighted populations for unmet needs (SMR < 75) over time. I directly fix the weighted populations to avoid introducing mismatches or uncertainty in simulation, under the limited availability on SMR < 75 data used in the setting allocation by NHSE.

Fifth, unavoidable costs of remoteness adjustments have exogenous and endogenous components. The introduction of the adjustment in 2016/17 is treated as exogenous positive funding shock to the CCGs (exogenous change 4). I argue they are exogenous as the remote and small hospitals are identified using average time to travel and expected higher cost on the cost-curve. These factors are uncorrelated with endogenous factors; such has local evolving health needs or distance from target of these trusts. However, the unavoidable costs of remoteness adjustments after 2016/17 are endogenous. It evolves by uplifting by an amount equivalent to the growth of the CCG core allocation budget over that time period, which the allocation growth is determined by all the factors outlined in Appendix Section 2. To avoid endogeneity, I fix the size of the adjustment funding, which is then converted back to weighted populations relative to the size of the fixed total allocations (fixed input 5).

Sixth, since the weighted populations are fixed, they are not projected forward (fixed input 1). For example, if 2015 populations level are used to calculate 2016/17 allocations (which is projected forward), I fix all endogenous variables in the formulae at 2015 level. Therefore, the simulated CCG core services weighted population $SwPop_i$ has isolated all the exogenous variation.

Finally, the CCGs' monetary (closing) target allocations are calculated by multiplying CCG's

share of the total simulated weighted population $SwPop_i$ (now exogenous) by the national budget for core services for the financial year (exogenous change 1).

B.1.2 Baselines

Baseline for CCG i for the year t are simulated by carrying forward from previous year's simulated allocations with adjustments.

$$SBaseline_{i,t} = SAllocation_{i,t-1} + adjustments_{i,t} \tag{B.41}$$

Using simulated baselines $SBaseline_{i,t}$ for the entire study period ensures no dynamic endogeneity. Not updating the baseline with actual allocations prevents the introduction of endogenous change, as reflected in actual allocations. Hence, the simulated instruments consider the counterfactual set of allocations a CCG would have received if its demographics and needs had remained fixed at their base year levels.

Note that for the first year's actual allocation in this study, 2013/14, has to be used as a starting (simulated) baseline for simulated allocation 2014/15. Simulated allocation per head for 2013/14 is therefore not possible to be constructed, hence resulting in a one-year data loss for my simulated IV strategy.

B.1.3 Opening Distance from Target

Similarly, opening distance from target (DfT) is also simulated, from the proportional difference between simulated baseline and simulated opening target. First, simulated opening target is calculated by multiplying CCG's current share of total weighted population by the sum of baselines (i.e. the total of previous year's allocation in England).

$$SOpeningTarget_{i,t} = \frac{SwPop_{i,t} \sum_i SBaseline_{i,t}}{\sum_i SwPop_{i,t}} \tag{B.42}$$

Then, the simulated opening distance from target for CCG i is calculated with:

$$SDfT_{i,t} = \frac{SBaseline_{i,t}}{SOpeningTarget_{i,t}} - 1 \tag{B.43}$$

Opening distance from target in per head terms is also calculated in the same way with per head values of target allocations and baselines. As in simulated baseline calculations, this prevents dynamic endogeneity.

B.1.4 Pace of Change Policy

Simulated allocations are derived similar to actual allocations, applying the simulated target allocations through the pace of change policy. In the application of pace of change policy for the simulated allocations, all CCG-specific inputs (with subscripts i) used are simulated value to remove endogeneity (fixed input 5). This includes using simulated values of weighted population, baselines, opening targets and closing targets for primary care medical services and specialised services. Weighted populations of the two services are directly fixed as the base year level (fixed inputs 1, 2, 3 & 4). For the simulated baselines, previous year's simulated allocations for the two services are used. For the starting baseline, I use the earliest available data on actual starting baseline for 2016/17 allocations (i.e. NHSE estimated allocations for the two services for 2015/16).

The pace of change policy formulae and parameters are allowed to change over time, since they are set nationally by NHSE (exogenous change 3).

As a result, the simulated instrument isolates all sources of variation in allocations that are driven by the changes in national allocation rules, from local endogenous responses.

B.1.5 Re-scaling

Simulated allocations are first transformed as shares of national budget, then multiplied by the national total budget for the year for re-scaling. This ensures a better fit to the actual allocation levels, enhancing instrument relevance. This re-scaling of simulated allocations does not introduce any bias since only national sums are used in the calculations.

B.2. Base-year-specific Construction

In the following, base-year-specific construction of the instrument is outlined. The strength of the instruments constructed with different base year is discussed and compared. I then explain why base year 2015 is the strongest and selected as the primary instrument.

B.2.1 Base year 2013 (used for allocations 2014/15)

A main limitation of simulated allocations with base year 2013 is the lack of new inputs required for the new pace of change policy. Weighted population for primary medical care and specialised services is not provided for 2013 (for allocations 2014/15 or 2015/16). To construct the instrument, I use newer data (2015, used for allocations 2016/17) and adjust it back to 2013 population level. This causes a potential mismatch in relative needs.

Also, community services weighted population is not available for base year 2013, since it is only introduced in 2019/20. This weighted population component within CCG core services is ignored in the calculation of simulated allocation 2019/20 under base year 2013.

B.2.2 Base year 2015 (used for allocations 2016/17)

Similar as above, community services weighted population remains unavailable for base year 2015, since it is only introduced in 2019/20. This weighted population component within CCG core services is also ignored in the calculation of simulated allocation 2019/20 under base year 2015.

B.2.3 Base year 2018 (used for allocations 2019/20)

In contrast, community services weighted population are available and considered in the calculation of simulated allocation 2019/20 under base year 2018.

B.2.4 Comparing instrument strength

The size of the first-stage coefficient shows the relevance of instrument, with one being perfect 1-to-1 relationship between simulated and actual allocations, while zero being no relevance. Coefficients close to one or zero reflects potential issues of instrument endogeneity and irrelevance respectively. Results (Appendix section 4) show first stage coefficients for base year 2013, 2015 and 2018 are 0.369, 0.531, 0.523 for wellbeing models; 0.380, 0.634 and 0.606 for mortality models.

I select 2015 as the primary base year. It offers the optimal balance between data availability (resolving the missing component issues of the 2013 base) and relevance (minimizing the projection distance compared to the 2018 base). First-stage results confirm this selection, showing that the 2015-based instrument maximises first-stage relevance. Its strength is followed by 2018-based instruments, while 2013-based instruments has the lowest strength.

The nature of how the instrument is simulated explains why base year 2015 is the best relative to other years. On one hand, the instrument has to relate with the starting baseline. An actual starting baseline must be used to apply pace of change policy across the years. Actual allocation in 2013/14 is involved in the simulation as the starting baseline. On the other hand, the instrument has to relate with actual allocations in later years and to have data for new inputs in allocation rule updates. There are new need values and components added to the allocation rules in later years.

For base year 2013, the unavailability of primary care and specialised services allocation data, together with larger difference from actual allocations in later years, explains the base year 2013

having the lowest relevance and highest uncertainty.

Instrument using base year 2015 is the strongest. It balances the needs of relevance of starting baseline, actual allocations in later years and availability of new values used in allocation rule changes.

Instruments using Base year 2018 follows 2015 in terms of strength. Despite an advantage from higher relevance in simulated allocation 2019/20 with community services weighted population, it has a pitfall for being the furthest distance from starting baseline.

Therefore, I therefore use base year 2015 as the main base year for this paper, with 2018 as robustness check. Results from using 2013 base year are deemed unreliable and greyed out in the figures due to larger uncertainty based on the above issues.

C. Robustness Checks

This section presents robustness checks examining the sensitivity of the main results to alternative specifications. For all result tables, I report three sets of standard errors for each coefficient: (i) CCG-intersection clustering, which accounts for the nature of cross-cutting of CCG boundaries, treating LADs with unique CCG combinations as separate clusters; (ii) Single CCG clustering, which assigns each LAD to its primary CCG; and (iii) clustering at the local authority level.

For OLS models, I present results across three time periods: 2013/14-2019/20 (full sample), 2014/15-2019/20 (contemporaneous allocations), and lagged allocations for 2014/15-2019/20. Each specification reports coefficients alongside the three alternative standard errors, allowing direct comparison of inference sensitivity to clustering assumptions.

For IV models, I examine robustness of IV results across base years used to construct the simulated instruments (2013, 2015, and 2018) for both contemporaneous and lagged specifications. I report first-stage and second-stage coefficients with all three clustering methods, alongside corresponding diagnostic tests. The Kleibergen-Paap Wald F-statistics test for weak instruments, while Wu-Hausman statistics test for endogeneity, with both computed separately for each clustering specification. This comprehensive approach ensures that inferences are robust to the choice of clustering level.

Table C1: Avoidable Mortality: Comparing OLS Results with Alternative SE Clustering

	Avoidable Mortality		Preventable Mortality		Treatable Mortality	
	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel 1: 2013/14-2019/20						
Allocation per head	-54.002	-36.431	-31.451	-15.107	-22.376	-21.143
(CCG-intersection)	(22.618)	(21.099)	(17.419)	(15.602)	(11.753)	(11.142)
(Single CCG)	(23.622)	(20.397)	(18.628)	(15.896)	(11.846)	(10.834)
(LAD)	(19.295)	(20.484)	(14.671)	(15.131)	(10.714)	(11.001)
Adjusted R^2	0.980	0.028	0.978	0.011	0.946	0.024
Within R^2	0.060	0.022	0.033	0.011	0.057	0.021
Observations	1,470	1,176	1,470	1,176	1,470	1,176
Panel 2: 2014/15-2019/20						
Allocation per head	-48.543	-33.055	-24.178	-12.993	-24.190	-20.047
(CCG-intersection)	(29.442)	(24.602)	(24.398)	(19.219)	(13.994)	(11.797)
(Single CCG)	(31.622)	(24.334)	(26.607)	(19.978)	(14.575)	(11.510)
(LAD)	(24.552)	(23.446)	(20.418)	(18.329)	(12.076)	(11.785)
Adjusted R^2	0.983	0.024	0.981	0.004	0.957	0.029
Within R^2	0.065	0.025	0.032	0.011	0.066	0.030
Observations	1,176	882	1,176	882	1,176	882
Panel 3: Lagged 2014/15-2019/20						
Lagged allocation per head	-43.210	-33.714	-29.546	-24.482	-13.522	-8.982
(CCG-intersection)	(34.412)	(27.862)	(27.958)	(22.598)	(17.550)	(14.141)
(Single CCG)	(36.973)	(29.731)	(29.181)	(23.452)	(18.172)	(14.438)
(LAD)	(28.827)	(28.293)	(22.917)	(22.587)	(14.390)	(13.967)
Adjusted R^2	0.986	0.034	0.984	0.011	0.966	0.030
Within R^2	0.061	0.037	0.030	0.020	0.058	0.033
Observations	882	588	882	588	882	588

Note: The table presents OLS estimates of the impact of NHS allocations on avoidable mortality measures across three different time panels. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C2: Wellbeing: Comparing OLS Results with Alternative SE Clustering

	Wellbeing		Happiness		Worthwhile		Life Satisfaction		Anxiety	
	FE	FD	FE	FD	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel 1: 2013/14-2019/20										
Allocation per head	0.362	-0.009	0.335	-0.246	0.262	0.045	0.230	-0.137	-0.619	-0.343
(CCG-intersection)	(0.188)	(0.189)	(0.205)	(0.247)	(0.187)	(0.259)	(0.178)	(0.240)	(0.363)	(0.396)
(Single CCG)	(0.192)	(0.191)	(0.217)	(0.229)	(0.166)	(0.271)	(0.175)	(0.237)	(0.389)	(0.434)
(LAD)	(0.180)	(0.224)	(0.197)	(0.277)	(0.173)	(0.256)	(0.181)	(0.265)	(0.330)	(0.437)
Adjusted R^2	0.418	0.062	0.321	0.031	0.399	0.021	0.455	0.046	0.304	0.051
Within R^2	0.015	0.003	0.006	0.002	0.012	0.004	0.016	0.005	0.014	0.002
Observations	2,260	1,933	2,264	1,938	2,264	1,938	2,265	1,940	2,260	1,933
Panel 2: 2014/15-2019/20										
Allocation per head	0.186	-0.009	0.103	-0.251	0.189	0.006	0.126	-0.129	-0.325	-0.385
(CCG-intersection)	(0.204)	(0.193)	(0.227)	(0.245)	(0.214)	(0.266)	(0.205)	(0.246)	(0.394)	(0.410)
(Single CCG)	(0.220)	(0.199)	(0.258)	(0.229)	(0.198)	(0.282)	(0.205)	(0.247)	(0.436)	(0.449)
(LAD)	(0.202)	(0.227)	(0.216)	(0.277)	(0.198)	(0.259)	(0.204)	(0.272)	(0.367)	(0.447)
Adjusted R^2	0.390	0.046	0.300	0.025	0.371	0.007	0.395	0.023	0.309	0.048
Within R^2	0.014	0.005	0.007	0.001	0.014	0.006	0.015	0.007	0.010	0.003
Observations	1,936	1,610	1,940	1,614	1,940	1,614	1,941	1,616	1,936	1,610
Panel 3: Lagged 2014/15-2019/20										
Lagged allocation per head	0.023	-0.034	0.061	0.085	0.021	0.019	0.062	0.212	0.064	0.446
(CCG-intersection)	(0.247)	(0.292)	(0.293)	(0.384)	(0.255)	(0.299)	(0.289)	(0.337)	(0.441)	(0.492)
(Single CCG)	(0.254)	(0.282)	(0.301)	(0.350)	(0.240)	(0.300)	(0.294)	(0.339)	(0.472)	(0.493)
(LAD)	(0.229)	(0.300)	(0.266)	(0.373)	(0.234)	(0.296)	(0.259)	(0.348)	(0.402)	(0.536)
Adjusted R^2	0.392	0.053	0.301	0.035	0.374	0.007	0.395	0.026	0.314	0.059
Within R^2	0.012	0.006	0.005	0.003	0.014	0.005	0.011	0.006	0.006	0.003
Observations	1,612	1,287	1,615	1,289	1,615	1,289	1,616	1,291	1,612	1,287

Note: The table presents OLS estimates of the impact of NHS allocations on wellbeing measures across three different time panels. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C3: Avoidable Mortality: Comparision of IV Base Years with Alternative SE Clustering (2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Avoidable Mortality	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Fitted allocation per head	-141.328	-153.762	-103.681	-111.769	-145.204	-151.814
(CCG-intersection)	(116.959)	(98.949)	(65.306)	(57.650)	(74.658)	(69.500)
(Single CCG)	(119.700)	(98.314)	(67.721)	(58.307)	(78.184)	(71.285)
(LAD)	(101.794)	(97.904)	(57.976)	(57.790)	(64.130)	(68.729)
1st Stage						
Simulated allocation per head	0.380	0.375	0.634	0.596	0.606	0.580
(CCG-intersection)	(0.100)	(0.081)	(0.087)	(0.073)	(0.083)	(0.075)
(Single CCG)	(0.107)	(0.087)	(0.091)	(0.078)	(0.089)	(0.080)
(LAD)	(0.076)	(0.072)	(0.071)	(0.069)	(0.068)	(0.072)
Adjusted R^2	0.983	-0.009	0.983	0.010	0.983	-0.008
Within R^2	0.037	-0.007	0.055	0.012	0.034	-0.006
Observations	1,176	882	1,176	882	1,176	882
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	14.560	21.326	53.721	66.443	52.678	60.421
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	12.610	18.692	48.330	59.034	46.355	51.907
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	24.707	26.832	79.879	74.667	78.879	65.091
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	0.672	1.586	0.895	2.279	1.985	3.338
	(0.412)	(0.208)	(0.344)	(0.131)	(0.159)	(0.068)
Single CCG	0.646	1.606	0.848	2.207	1.827	3.141
	(0.422)	(0.205)	(0.357)	(0.137)	(0.176)	(0.076)
LAD	0.882	1.613	1.102	2.221	2.662	3.379
	(0.348)	(0.204)	(0.294)	(0.136)	(0.103)	(0.066)

Note: The table compares IV estimates of the impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C4: Preventable Mortality: Comparision of IV Base Years with Alternative SE Clustering (2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Preventable Mortality	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Fitted allocation per head	-143.617	-144.820	-101.232	-101.563	-112.740	-106.907
(CCG-intersection)	(89.553)	(74.389)	(48.648)	(42.063)	(51.947)	(45.274)
(Single CCG)	(92.391)	(74.758)	(51.346)	(43.420)	(57.054)	(48.353)
(LAD)	(74.409)	(71.633)	(40.492)	(40.068)	(45.518)	(45.098)
1st Stage						
Simulated allocation per head	0.380	0.375	0.634	0.596	0.606	0.580
(CCG-intersection)	(0.100)	(0.081)	(0.087)	(0.073)	(0.083)	(0.075)
(Single CCG)	(0.107)	(0.087)	(0.091)	(0.078)	(0.089)	(0.080)
(LAD)	(0.076)	(0.072)	(0.071)	(0.069)	(0.068)	(0.072)
Adjusted R^2	0.979	-0.068	0.980	-0.029	0.980	-0.033
Within R^2	-0.055	-0.060	-0.004	-0.021	-0.016	-0.025
Observations	1,176	882	1,176	882	1,176	882
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	14.560	21.326	53.721	66.443	52.678	60.421
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	12.610	18.692	48.330	59.034	46.355	51.907
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	24.707	26.832	79.879	74.667	78.879	65.091
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	1.921	3.365	3.352	5.604	3.729	5.249
	(0.166)	(0.067)	(0.067)	(0.018)	(0.053)	(0.022)
Single CCG	1.822	3.349	3.079	5.278	3.079	4.549
	(0.177)	(0.067)	(0.079)	(0.022)	(0.079)	(0.033)
LAD	2.786	3.624	4.856	6.179	4.739	5.195
	(0.095)	(0.057)	(0.028)	(0.013)	(0.029)	(0.023)

Note: The table compares IV estimates of the impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C5: Treatable Mortality: Comparision of IV Base Years with Alternative SE Clustering (2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Treatable Mortality	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Fitted allocation per head	2.433	-9.229	-2.346	-10.309	-32.451	-45.170
(CCG-intersection)	(47.820)	(40.189)	(28.011)	(24.848)	(35.486)	(34.072)
(Single CCG)	(47.706)	(39.919)	(27.403)	(24.044)	(35.666)	(34.112)
(LAD)	(45.082)	(42.995)	(26.872)	(27.066)	(30.469)	(34.018)
1st Stage						
Simulated allocation per head	0.380	0.375	0.634	0.596	0.606	0.580
(CCG-intersection)	(0.100)	(0.081)	(0.087)	(0.073)	(0.083)	(0.075)
(Single CCG)	(0.107)	(0.087)	(0.091)	(0.078)	(0.089)	(0.080)
(LAD)	(0.076)	(0.072)	(0.071)	(0.069)	(0.068)	(0.072)
Adjusted R^2	0.957	0.028	0.957	0.029	0.957	0.024
Within R^2	0.057	0.029	0.059	0.029	0.065	0.024
Observations	1,176	882	1,176	882	1,176	882
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	14.560	21.326	53.721	66.443	52.678	60.421
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	12.610	18.692	48.330	59.034	46.355	51.907
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	24.707	26.832	79.879	74.667	78.879	65.091
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	0.339	0.079	0.810	0.198	0.064	0.618
	(0.560)	(0.778)	(0.368)	(0.656)	(0.800)	(0.432)
Single CCG	0.344	0.080	0.886	0.213	0.064	0.612
	(0.558)	(0.777)	(0.347)	(0.645)	(0.800)	(0.434)
LAD	0.376	0.068	0.828	0.160	0.087	0.620
	(0.540)	(0.794)	(0.363)	(0.689)	(0.768)	(0.431)

Note: The table compares IV estimates of the impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C6: Wellbeing: Comparision of IV Base Years with Alternative SE Clustering (2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Wellbeing	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Fitted allocation per head	0.798	-0.262	0.759	0.715	0.655	0.667
(CCG-intersection)	(0.737)	(3.586)	(0.519)	(1.776)	(0.498)	(1.056)
(Single CCG)	(0.767)	(3.381)	(0.522)	(1.603)	(0.513)	(0.956)
(LAD)	(0.695)	(3.511)	(0.480)	(1.779)	(0.469)	(1.056)
1st Stage						
Simulated allocation per head	0.369	0.114	0.531	0.220	0.523	0.311
(CCG-intersection)	(0.070)	(0.046)	(0.064)	(0.046)	(0.070)	(0.074)
(Single CCG)	(0.076)	(0.051)	(0.069)	(0.051)	(0.078)	(0.084)
(LAD)	(0.057)	(0.040)	(0.055)	(0.042)	(0.063)	(0.073)
Adjusted R^2	0.386	0.045	0.387	0.042	0.388	0.042
Within R^2	0.007	0.004	0.008	0.000	0.010	0.001
Observations	1,936	1,610	1,936	1,610	1,936	1,610
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	28.192	6.261	68.532	22.924	55.555	17.721
	(0.000)	(0.012)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	23.865	4.984	59.656	18.583	45.427	13.572
	(0.000)	(0.026)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	42.210	8.325	93.368	27.231	69.673	18.312
	(0.000)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	0.746	0.005	1.441	0.168	1.063	0.423
	(0.388)	(0.944)	(0.230)	(0.682)	(0.303)	(0.515)
Single CCG	0.694	0.006	1.465	0.207	1.021	0.522
	(0.405)	(0.940)	(0.226)	(0.649)	(0.312)	(0.470)
LAD	0.848	0.005	1.734	0.168	1.223	0.429
	(0.357)	(0.943)	(0.188)	(0.681)	(0.269)	(0.512)

Note: The table compares IV estimates of the impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses.
$* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C7: Happiness: Comparision of IV Base Years with Alternative SE Clustering (2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Happiness	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Fitted allocation per head	0.133	-3.122	0.275	-0.679	0.251	-0.267
(CCG-intersection)	(0.839)	(4.549)	(0.572)	(2.135)	(0.524)	(1.170)
(Single CCG)	(0.860)	(4.406)	(0.579)	(2.021)	(0.547)	(1.070)
(LAD)	(0.839)	(4.555)	(0.557)	(2.158)	(0.496)	(1.178)
1st Stage						
Simulated allocation per head	0.369	0.114	0.531	0.220	0.523	0.311
(CCG-intersection)	(0.070)	(0.046)	(0.064)	(0.046)	(0.070)	(0.074)
(Single CCG)	(0.076)	(0.051)	(0.069)	(0.051)	(0.078)	(0.084)
(LAD)	(0.057)	(0.040)	(0.055)	(0.042)	(0.063)	(0.073)
Adjusted R^2	0.300	-0.014	0.300	0.025	0.300	0.025
Within R^2	0.007	-0.039	0.006	0.000	0.007	0.001
Observations	1,940	1,614	1,940	1,614	1,940	1,614
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	28.192	6.261	68.532	22.924	55.555	17.721
	(0.000)	(0.012)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	23.865	4.984	59.656	18.583	45.427	13.572
	(0.000)	(0.026)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	42.210	8.325	93.368	27.231	69.673	18.312
	(0.000)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	0.001	0.399	0.107	0.041	0.098	0.000
	(0.971)	(0.527)	(0.743)	(0.840)	(0.754)	(0.989)
Single CCG	0.001	0.426	0.110	0.045	0.094	0.000
	(0.971)	(0.514)	(0.740)	(0.831)	(0.759)	(0.988)
LAD	0.001	0.399	0.112	0.040	0.110	0.000
	(0.971)	(0.528)	(0.738)	(0.842)	(0.741)	(0.989)

Note: The table compares IV estimates of the impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses.
$* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C8: Worthwhile: Comparision of IV Base Years with Alternative SE Clustering (2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Worthwhile	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Fitted allocation per head	0.794	-0.854	0.468	-0.124	0.482	-0.320
(CCG-intersection)	(0.756)	(3.159)	(0.512)	(1.542)	(0.457)	(0.909)
(Single CCG)	(0.720)	(3.062)	(0.452)	(1.425)	(0.432)	(0.869)
(LAD)	(0.644)	(3.192)	(0.452)	(1.602)	(0.415)	(0.957)
1st Stage						
Simulated allocation per head	0.369	0.114	0.531	0.220	0.523	0.311
(CCG-intersection)	(0.070)	(0.046)	(0.064)	(0.046)	(0.070)	(0.074)
(Single CCG)	(0.076)	(0.051)	(0.069)	(0.051)	(0.078)	(0.084)
(LAD)	(0.057)	(0.040)	(0.055)	(0.042)	(0.063)	(0.073)
Adjusted R^2	0.367	0.001	0.370	0.007	0.370	0.006
Within R^2	0.007	0.000	0.012	0.005	0.012	0.005
Observations	1,940	1,614	1,940	1,614	1,940	1,614
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	28.192	6.261	68.532	22.924	55.555	17.721
	(0.000)	(0.012)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	23.865	4.984	59.656	18.583	45.427	13.572
	(0.000)	(0.026)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	42.210	8.325	93.368	27.231	69.673	18.312
	(0.000)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	0.697	0.075	0.360	0.007	0.527	0.141
	(0.404)	(0.785)	(0.548)	(0.932)	(0.468)	(0.708)
Single CCG	0.765	0.080	0.473	0.009	0.583	0.157
	(0.382)	(0.778)	(0.492)	(0.926)	(0.445)	(0.692)
LAD	0.975	0.073	0.472	0.007	0.645	0.125
	(0.323)	(0.787)	(0.492)	(0.934)	(0.422)	(0.724)

Note: The table compares IV estimates of the impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses.
$* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C9: Life Satisfaction: Comparision of IV Base Years with Alternative SE Clustering (2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Life Satisfaction	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Fitted allocation per head	0.390	2.603	0.536	2.044	0.784	1.144
(CCG-intersection)	(0.814)	(4.092)	(0.559)	(2.156)	(0.539)	(1.288)
(Single CCG)	(0.841)	(4.089)	(0.568)	(2.125)	(0.555)	(1.281)
(LAD)	(0.760)	(4.372)	(0.518)	(2.266)	(0.486)	(1.330)
1st Stage						
Simulated allocation per head	0.369	0.114	0.531	0.220	0.523	0.311
(CCG-intersection)	(0.070)	(0.046)	(0.064)	(0.046)	(0.070)	(0.074)
(Single CCG)	(0.076)	(0.051)	(0.069)	(0.051)	(0.078)	(0.084)
(LAD)	(0.057)	(0.040)	(0.055)	(0.042)	(0.063)	(0.073)
Adjusted R^2	0.395	-0.025	0.394	-0.007	0.391	0.013
Within R^2	0.014	-0.043	0.013	-0.025	0.008	-0.004
Observations	1,941	1,616	1,941	1,616	1,941	1,616
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	28.192	6.261	68.532	22.924	55.555	17.721
	(0.000)	(0.012)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	23.865	4.984	59.656	18.583	45.427	13.572
	(0.000)	(0.026)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	42.210	8.325	93.368	27.231	69.673	18.312
	(0.000)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	0.112	0.447	0.622	1.030	1.740	1.015
	(0.738)	(0.504)	(0.430)	(0.310)	(0.187)	(0.314)
Single CCG	0.105	0.448	0.598	1.060	1.626	1.027
	(0.746)	(0.503)	(0.439)	(0.303)	(0.202)	(0.311)
LAD	0.130	0.392	0.742	0.934	2.226	0.957
	(0.719)	(0.531)	(0.389)	(0.334)	(0.136)	(0.328)

Note: The table compares IV estimates of the impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses.
$* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C10: Anxiety: Comparision of IV Base Years with Alternative SE Clustering (2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Anxiety	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Fitted allocation per head	-1.639	1.505	-1.606	-0.883	-0.973	-1.812
(CCG-intersection)	(1.357)	(6.452)	(0.907)	(3.151)	(0.912)	(1.995)
(Single CCG)	(1.418)	(6.213)	(0.944)	(2.908)	(0.933)	(1.896)
(LAD)	(1.220)	(6.170)	(0.789)	(2.912)	(0.839)	(1.907)
1st Stage						
Simulated allocation per head	0.369	0.114	0.531	0.220	0.523	0.311
(CCG-intersection)	(0.070)	(0.046)	(0.064)	(0.046)	(0.070)	(0.074)
(Single CCG)	(0.076)	(0.051)	(0.069)	(0.051)	(0.078)	(0.084)
(LAD)	(0.057)	(0.040)	(0.055)	(0.042)	(0.063)	(0.073)
Adjusted R^2	0.303	0.040	0.303	0.047	0.308	0.043
Within R^2	0.001	-0.006	0.001	0.002	0.008	-0.003
Observations	1,936	1,610	1,936	1,610	1,936	1,610
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	28.192	6.261	68.532	22.924	55.555	17.721
	(0.000)	(0.012)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	23.865	4.984	59.656	18.583	45.427	13.572
	(0.000)	(0.026)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	42.210	8.325	93.368	27.231	69.673	18.312
	(0.000)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	1.023	0.086	2.465	0.025	0.622	0.534
	(0.312)	(0.769)	(0.116)	(0.873)	(0.430)	(0.465)
Single CCG	0.948	0.093	2.342	0.030	0.619	0.600
	(0.330)	(0.760)	(0.126)	(0.862)	(0.431)	(0.438)
LAD	1.276	0.094	3.362	0.030	0.739	0.593
	(0.259)	(0.759)	(0.067)	(0.862)	(0.390)	(0.441)

Note: The table compares IV estimates of the impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses.
$* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C11: Avoidable Mortality: Comparision of IV Base Years with Alternative SE Clustering (Lagged 2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Avoidable Mortality	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Lagged fitted allocation per head	-81.530	-70.162	-76.826	-77.634	-107.079	-120.028
(CCG-intersection)	(95.052)	(78.475)	(59.872)	(50.073)	(76.331)	(64.710)
(Single CCG)	(97.757)	(79.571)	(61.502)	(51.163)	(76.165)	(64.611)
(LAD)	(86.530)	(86.369)	(54.160)	(54.671)	(63.510)	(65.460)
1st Stage						
Lagged simulated allocation per head	0.518	0.508	0.806	0.795	0.684	0.672
(CCG-intersection)	(0.116)	(0.096)	(0.098)	(0.082)	(0.100)	(0.084)
(Single CCG)	(0.125)	(0.102)	(0.102)	(0.085)	(0.106)	(0.088)
(LAD)	(0.080)	(0.080)	(0.071)	(0.073)	(0.071)	(0.074)
Adjusted R^2	0.986	0.031	0.986	0.030	0.986	0.018
Within R^2	0.057	0.034	0.057	0.032	0.049	0.021
Observations	882	588	882	588	882	588
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	19.895	28.270	67.667	93.463	46.515	64.087
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	17.259	24.614	62.011	86.953	41.921	58.294
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	42.373	40.314	128.738	117.536	93.030	82.547
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	0.187	0.247	0.471	1.114	0.879	2.184
	(0.665)	(0.619)	(0.493)	(0.291)	(0.349)	(0.139)
Single CCG	0.179	0.244	0.468	1.113	0.920	2.264
	(0.672)	(0.621)	(0.494)	(0.292)	(0.337)	(0.132)
LAD	0.221	0.199	0.538	0.881	1.274	2.138
	(0.639)	(0.655)	(0.463)	(0.348)	(0.259)	(0.144)

Note: The table compares IV estimates of the lagged impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C12: Preventable Mortality: Comparision of IV Base Years with Alternative SE Clustering (Lagged 2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Preventable Mortality	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Lagged fitted allocation per head	-90.257	-86.812	-77.679	-81.457	-84.023	-90.004
(CCG-intersection)	(77.895)	(64.715)	(48.531)	(40.741)	(58.575)	(48.752)
(Single CCG)	(80.546)	(65.575)	(49.807)	(41.302)	(58.610)	(48.862)
(LAD)	(65.698)	(65.817)	(40.520)	(41.114)	(49.221)	(50.135)
1st Stage						
Lagged simulated allocation per head	0.518	0.508	0.806	0.795	0.684	0.672
(CCG-intersection)	(0.116)	(0.096)	(0.098)	(0.082)	(0.100)	(0.084)
(Single CCG)	(0.125)	(0.102)	(0.102)	(0.085)	(0.106)	(0.088)
(LAD)	(0.080)	(0.080)	(0.071)	(0.073)	(0.071)	(0.074)
Adjusted R^2	0.984	-0.004	0.984	-0.002	0.984	-0.006
Within R^2	0.012	0.006	0.018	0.008	0.015	0.004
Observations	882	588	882	588	882	588
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	19.895	28.270	67.667	93.463	46.515	64.087
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	17.259	24.614	62.011	86.953	41.921	58.294
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	42.373	40.314	128.738	117.536	93.030	82.547
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	0.697	1.056	1.472	2.825	1.120	2.301
	(0.404)	(0.304)	(0.225)	(0.093)	(0.290)	(0.129)
Single CCG	0.654	1.036	1.422	2.809	1.149	2.336
	(0.419)	(0.309)	(0.233)	(0.094)	(0.284)	(0.126)
LAD	0.972	1.017	2.075	2.751	1.564	2.143
	(0.324)	(0.313)	(0.150)	(0.097)	(0.211)	(0.143)

Note: The table compares IV estimates of the lagged impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C13: Treatable Mortality: Comparision of IV Base Years with Alternative SE Clustering (Lagged 2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Treatable Mortality	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Lagged fitted allocation per head	9.548	17.551	1.327	4.370	-22.796	-29.745
(CCG-intersection)	(44.714)	(35.280)	(29.895)	(24.329)	(38.635)	(33.785)
(Single CCG)	(43.610)	(35.955)	(29.782)	(24.815)	(38.724)	(33.890)
(LAD)	(39.859)	(40.117)	(26.174)	(26.563)	(31.764)	(33.837)
1st Stage						
Lagged simulated allocation per head	0.518	0.508	0.806	0.795	0.684	0.672
(CCG-intersection)	(0.116)	(0.096)	(0.098)	(0.082)	(0.100)	(0.084)
(Single CCG)	(0.125)	(0.102)	(0.102)	(0.085)	(0.106)	(0.088)
(LAD)	(0.080)	(0.080)	(0.071)	(0.073)	(0.071)	(0.074)
Adjusted R^2	0.966	0.025	0.966	0.029	0.966	0.027
Within R^2	0.052	0.028	0.056	0.032	0.057	0.030
Observations	882	588	882	588	882	588
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	19.895	28.270	67.667	93.463	46.515	64.087
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	17.259	24.614	62.011	86.953	41.921	58.294
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	42.373	40.314	128.738	117.536	93.030	82.547
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	0.315	0.674	0.376	0.455	0.073	0.458
	(0.575)	(0.412)	(0.539)	(0.500)	(0.788)	(0.499)
Single CCG	0.339	0.649	0.396	0.438	0.074	0.459
	(0.561)	(0.420)	(0.529)	(0.508)	(0.786)	(0.498)
LAD	0.385	0.498	0.461	0.349	0.107	0.454
	(0.535)	(0.480)	(0.497)	(0.555)	(0.743)	(0.500)

Note: The table compares IV estimates of the lagged impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C14: Wellbeing: Comparision of IV Base Years with Alternative SE Clustering (Lagged 2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Wellbeing	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Lagged fitted allocation per head	1.143	2.389	0.849	1.955	0.892	1.437
(CCG-intersection)	(0.744)	(2.420)	(0.488)	(1.174)	(0.581)	(1.417)
(Single CCG)	(0.768)	(2.292)	(0.491)	(1.134)	(0.601)	(1.374)
(LAD)	(0.724)	(2.457)	(0.472)	(1.253)	(0.553)	(1.507)
1st Stage						
Lagged simulated allocation per head	0.527	0.227	0.746	0.415	0.622	0.312
(CCG-intersection)	(0.091)	(0.063)	(0.080)	(0.055)	(0.081)	(0.055)
(Single CCG)	(0.097)	(0.068)	(0.083)	(0.057)	(0.086)	(0.057)
(LAD)	(0.068)	(0.050)	(0.061)	(0.045)	(0.063)	(0.047)
Adjusted R^2	0.382	0.019	0.387	0.030	0.386	0.041
Within R^2	-0.005	-0.031	0.003	-0.019	0.002	-0.008
Observations	1,612	1,287	1,612	1,287	1,612	1,287
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	33.274	12.766	86.448	56.833	58.197	32.185
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	29.367	11.098	79.905	52.532	52.563	29.641
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	60.212	20.525	147.871	85.772	97.618	44.249
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	2.542	1.018	3.841	3.056	2.727	1.126
	(0.111)	(0.313)	(0.050)	(0.080)	(0.099)	(0.289)
Single CCG	2.386	1.135	3.857	3.278	2.542	1.197
	(0.122)	(0.287)	(0.050)	(0.070)	(0.111)	(0.274)
LAD	2.658	0.988	4.001	2.670	2.982	0.992
	(0.103)	(0.320)	(0.045)	(0.102)	(0.084)	(0.319)

Note: The table compares IV estimates of the lagged impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C15: Happiness: Comparision of IV Base Years with Alternative SE Clustering (Lagged 2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Happiness	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Lagged fitted allocation per head	1.347	5.953	0.933	3.606	0.892	2.217
(CCG-intersection)	(0.985)	(3.618)	(0.657)	(1.778)	(0.747)	(2.072)
(Single CCG)	(1.006)	(3.584)	(0.664)	(1.768)	(0.758)	(2.049)
(LAD)	(0.927)	(3.466)	(0.592)	(1.739)	(0.647)	(2.026)
1st Stage						
Lagged simulated allocation per head	0.527	0.227	0.746	0.415	0.622	0.312
(CCG-intersection)	(0.091)	(0.063)	(0.080)	(0.055)	(0.081)	(0.055)
(Single CCG)	(0.097)	(0.068)	(0.083)	(0.057)	(0.086)	(0.057)
(LAD)	(0.068)	(0.050)	(0.061)	(0.045)	(0.063)	(0.047)
Adjusted R^2	0.291	-0.089	0.297	-0.010	0.297	0.019
Within R^2	-0.009	-0.124	-0.001	-0.043	-0.001	-0.014
Observations	1,615	1,289	1,615	1,289	1,615	1,289
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	33.274	12.766	86.448	56.833	58.197	32.185
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	29.367	11.098	79.905	52.532	52.563	29.641
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	60.212	20.525	147.871	85.772	97.618	44.249
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	1.871	2.661	2.204	4.115	1.465	1.096
	(0.171)	(0.103)	(0.138)	(0.043)	(0.226)	(0.295)
Single CCG	1.797	2.706	2.170	4.129	1.429	1.115
	(0.180)	(0.100)	(0.141)	(0.042)	(0.232)	(0.291)
LAD	2.101	2.900	2.722	4.301	1.986	1.147
	(0.147)	(0.089)	(0.099)	(0.038)	(0.159)	(0.284)

Note: The table compares IV estimates of the lagged impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C16: Worthwhile: Comparision of IV Base Years with Alternative SE Clustering (Lagged 2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Worthwhile	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Lagged fitted allocation per head	0.527	-0.406	0.262	0.191	0.745	0.313
(CCG-intersection)	(0.670)	(2.382)	(0.463)	(1.186)	(0.546)	(1.506)
(Single CCG)	(0.646)	(2.413)	(0.437)	(1.173)	(0.550)	(1.537)
(LAD)	(0.664)	(2.416)	(0.458)	(1.268)	(0.514)	(1.580)
1st Stage						
Lagged simulated allocation per head	0.527	0.227	0.746	0.415	0.622	0.312
(CCG-intersection)	(0.091)	(0.063)	(0.080)	(0.055)	(0.081)	(0.055)
(Single CCG)	(0.097)	(0.068)	(0.083)	(0.057)	(0.086)	(0.057)
(LAD)	(0.068)	(0.050)	(0.061)	(0.045)	(0.063)	(0.047)
Adjusted R^2	0.372	0.006	0.374	0.007	0.370	0.007
Within R^2	0.011	0.004	0.014	0.005	0.008	0.005
Observations	1,615	1,289	1,615	1,289	1,615	1,289
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	33.274	12.766	86.448	56.833	58.197	32.185
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	29.367	11.098	79.905	52.532	52.563	29.641
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	60.212	20.525	147.871	85.772	97.618	44.249
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	0.668	0.032	0.390	0.022	2.244	0.040
	(0.414)	(0.857)	(0.532)	(0.881)	(0.134)	(0.842)
Single CCG	0.713	0.032	0.435	0.023	2.142	0.038
	(0.398)	(0.859)	(0.509)	(0.880)	(0.143)	(0.845)
LAD	0.665	0.032	0.375	0.019	2.500	0.036
	(0.415)	(0.859)	(0.541)	(0.890)	(0.114)	(0.850)

Note: The table compares IV estimates of the lagged impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C17: Life Satisfaction: Comparision of IV Base Years with Alternative SE Clustering (Lagged 2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Life Satisfaction	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Lagged fitted allocation per head	0.559	1.381	0.792	1.684	1.116	1.860
(CCG-intersection)	(0.763)	(2.445)	(0.523)	(1.234)	(0.639)	(1.529)
(Single CCG)	(0.804)	(2.437)	(0.523)	(1.211)	(0.644)	(1.528)
(LAD)	(0.703)	(2.679)	(0.477)	(1.368)	(0.580)	(1.753)
1st Stage						
Lagged simulated allocation per head	0.527	0.227	0.746	0.415	0.622	0.312
(CCG-intersection)	(0.091)	(0.063)	(0.080)	(0.055)	(0.081)	(0.055)
(Single CCG)	(0.097)	(0.068)	(0.083)	(0.057)	(0.086)	(0.057)
(LAD)	(0.068)	(0.050)	(0.061)	(0.045)	(0.063)	(0.047)
Adjusted R^2	0.393	0.019	0.391	0.015	0.387	0.012
Within R^2	0.008	-0.001	0.005	-0.005	-0.002	-0.008
Observations	1,616	1,291	1,616	1,291	1,616	1,291
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	33.274	12.766	86.448	56.833	58.197	32.185
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	29.367	11.098	79.905	52.532	52.563	29.641
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	60.212	20.525	147.871	85.772	97.618	44.249
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	0.495	0.233	2.806	1.539	3.424	1.220
	(0.481)	(0.629)	(0.094)	(0.215)	(0.064)	(0.269)
Single CCG	0.441	0.235	2.837	1.604	3.386	1.223
	(0.506)	(0.628)	(0.092)	(0.205)	(0.066)	(0.269)
LAD	0.578	0.194	3.316	1.238	4.124	0.919
	(0.447)	(0.660)	(0.069)	(0.266)	(0.042)	(0.338)

Note: The table compares IV estimates of the lagged impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C18: Anxiety: Comparision of IV Base Years with Alternative SE Clustering (Lagged 2014/15-2019/20)

Dependent Variable:	Base Year: 2013		Base Year: 2015		Base Year: 2018	
Anxiety	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: IV						
2nd Stage						
Lagged fitted allocation per head	-2.059	-2.575	-1.357	-2.330	-0.743	-1.384
(CCG-intersection)	(1.538)	(4.015)	(0.945)	(1.966)	(1.000)	(2.509)
(Single CCG)	(1.552)	(3.690)	(0.965)	(1.947)	(1.052)	(2.465)
(LAD)	(1.313)	(4.216)	(0.857)	(2.211)	(0.966)	(2.544)
1st Stage						
Lagged simulated allocation per head	0.527	0.227	0.746	0.415	0.622	0.312
(CCG-intersection)	(0.091)	(0.063)	(0.080)	(0.055)	(0.081)	(0.055)
(Single CCG)	(0.097)	(0.068)	(0.083)	(0.057)	(0.086)	(0.057)
(LAD)	(0.068)	(0.050)	(0.061)	(0.045)	(0.063)	(0.047)
Adjusted R^2	0.301	0.042	0.309	0.045	0.312	0.053
Within R^2	-0.013	-0.015	-0.003	-0.012	0.003	-0.003
Observations	1,612	1,287	1,612	1,287	1,612	1,287
Diagnostic Tests						
F-statistic (1st stage)						
Refined CCG	33.274	12.766	86.448	56.833	58.197	32.185
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Single CCG	29.367	11.098	79.905	52.532	52.563	29.641
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
LAD	60.212	20.525	147.871	85.772	97.618	44.249
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic						
Refined CCG	2.075	0.575	2.893	2.127	0.808	0.553
	(0.150)	(0.448)	(0.089)	(0.145)	(0.369)	(0.457)
Single CCG	2.061	0.682	2.846	2.172	0.736	0.574
	(0.151)	(0.409)	(0.092)	(0.141)	(0.391)	(0.449)
LAD	2.884	0.522	3.524	1.674	0.844	0.542
	(0.089)	(0.470)	(0.060)	(0.196)	(0.358)	(0.462)

Note: The table compares IV estimates of the lagged impact of NHS allocations on the dependent variable using simulated instruments constructed with different base years. Columns 1 and 2 use the 2013 base year; columns 3 and 4 use the 2015 base year; and columns 5 and 6 use the 2018 base year. Allocation per head is instrumented by the simulated allocation per head. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are reported in parentheses and are clustered at three alternative levels: CCG-intersections, single CCG, and LAD. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under different clustering of standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are given in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

D. Two-Period Models

Table D1: Effect of NHS Allocations (2-Period model) on Avoidable Mortality

	Avoidable Mortality		Preventable Mortality		Treatable Mortality	
	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
Allocation per head	14.032	8.478	38.923	29.589	-25.160	-21.340
	(54.438)	(37.988)	(46.803)	(28.988)	(24.809)	(20.204)
Lagged allocation per head	-53.221	-39.266	-57.316	-43.856	4.428	4.991
	(49.480)	(35.316)	(39.015)	(27.296)	(21.923)	(17.943)
Panel B: IV						
2nd Stage						
Fitted allocation per head	1703.179	1840.521	-148.807	-287.017	1857.728	2151.872
	(7202.608)	(6465.266)	(1960.065)	(1727.535)	(7370.762)	(7203.646)
Lagged fitted allocation per head	-1016.129	-1081.893	4.388	75.150	-1023.210	-1169.775
	(3948.780)	(3507.497)	(1070.642)	(927.521)	(4028.695)	(3900.412)
1st Stage						
Simulated allocation per head	0.160	0.121	0.160	0.121	0.160	0.121
	(0.211)	(0.131)	(0.211)	(0.131)	(0.211)	(0.131)
Lagged simulated allocation	0.458**	0.529***	0.458**	0.529***	0.458**	0.529***
	(0.218)	(0.120)	(0.218)	(0.120)	(0.218)	(0.120)
Local area FE	Yes	-	Yes	-	Yes	-
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.930	-4.268	0.982	-0.251	0.270	-23.091
Within R^2	-3.829	-4.244	-0.074	-0.236	-19.496	-22.981
Observations	882	588	882	588	882	588
Diagnostic Tests						
F-statistic						
Current allocation	0.573	0.851	0.573	0.851	0.573	0.851
	(0.449)	(0.356)	(0.449)	(0.356)	(0.449)	(0.356)
Lagged allocation	4.425	19.451	4.425	19.451	4.425	19.451
	(0.035)	(0.000)	(0.035)	(0.000)	(0.035)	(0.000)
Wu-Hausman statistic	0.104	0.191	1.056	2.013	0.065	0.091
	(0.949)	(0.909)	(0.590)	(0.365)	(0.968)	(0.956)

Note: The table shows OLS and IV estimates of the impact of NHS allocations on mortality outcomes using a two-period model. Columns 1 and 2 report results for Avoidable Mortality; columns 3 and 4 for Preventable Mortality; and columns 5 and 6 for Treatable Mortality. Panel A reports OLS results. Panel B reports IV results where allocation per head and lagged allocation per head are instrumented by the simulated allocation per head and its lag. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are clustered at the CCG-intersection level and given in parentheses. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under clustered standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are reported in parentheses.
$* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table D2: Effect of NHS Allocations (2-Period model) on Wellbeing

	Wellbeing		Happiness		Worthwhile		Life Satisfaction		Anxiety	
	FE	FD	FE	FD	FE	FD	FE	FD	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: OLS										
Allocation per head	0.299	-0.030	0.341	-0.208	0.201	-0.175	0.262	-0.191	-0.387	-0.489
	(0.298)	(0.252)	(0.295)	(0.329)	(0.340)	(0.338)	(0.244)	(0.290)	(0.604)	(0.488)
Lagged allocation per head	-0.149	-0.029	-0.135	0.120	-0.095	0.049	-0.089	0.245	0.286	0.529
	(0.281)	(0.305)	(0.330)	(0.409)	(0.313)	(0.319)	(0.300)	(0.343)	(0.540)	(0.513)
Panel B: IV										
2nd Stage										
Fitted allocation per head	0.773	-4.084	-15.952	-11.650	7.785	0.103	3.459	1.002	-7.262	5.625
	(19.690)	(6.522)	(35.822)	(10.211)	(18.635)	(5.599)	(18.444)	(5.964)	(37.976)	(11.572)
Lagged fitted allocation per head	0.406	3.677	10.076	8.520	-4.200	0.147	-1.191	1.261	2.810	-4.703
	(11.416)	(3.792)	(20.778)	(6.109)	(10.826)	(3.410)	(10.659)	(3.399)	(22.140)	(6.670)
1st Stage										
Simulated allocation per head	0.205^{***}	0.222^{***}	0.205^{***}	0.222^{***}	0.205^{***}	0.222^{***}	0.205^{***}	0.222^{***}	0.205^{***}	0.222^{***}
	(0.049)	(0.044)	(0.049)	(0.044)	(0.049)	(0.044)	(0.049)	(0.044)	(0.049)	(0.044)
Lagged simulated allocation	0.491^{***}	0.385^{***}	0.491^{***}	0.385^{***}	0.491^{***}	0.385^{***}	0.491^{***}	0.385^{***}	0.491^{***}	0.385^{***}
	(0.073)	(0.057)	(0.073)	(0.057)	(0.073)	(0.057)	(0.073)	(0.057)	(0.073)	(0.057)
Local area FE	Yes	-	Yes	-	Yes	-	Yes	-	Yes	-
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.385	-0.120	-0.932	-0.699	0.012	0.006	0.336	0.009	0.206	-0.058
Within R^2	0.002	-0.176	-1.748	-0.754	-0.556	0.005	-0.085	-0.010	-0.150	-0.120
Observations	1,612	1,287	1,615	1,289	1,615	1,289	1,616	1,291	1,612	1,287
Diagnostic Tests										
F-statistic										
Current allocation	17.598	25.346	17.598	25.346	17.598	25.346	17.598	25.346	17.598	25.346
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Lagged allocation	45.785	45.468	45.785	45.468	45.785	45.468	45.785	45.468	45.785	45.468
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wu-Hausman statistic	3.948	3.556	0.496	2.834	0.324	0.202	2.260	3.284	1.874	2.605
	(0.139)	(0.169)	(0.780)	(0.242)	(0.850)	(0.904)	(0.323)	(0.194)	(0.392)	(0.272)

Note: The table shows OLS and IV estimates of the impact of NHS allocations on wellbeing measures using a two-period model. Columns 1 and 2 report results for Wellbeing; columns 3 and 4 for Happiness; columns 5 and 6 for Worthwhile; columns 7 and 8 for Life Satisfaction; and columns 9 and 10 for Anxiety. Panel A reports OLS results. Panel B reports IV results where allocation per head and lagged allocation per head are instrumented by the simulated allocation per head and its lag. Odd-numbered columns (FE) show estimates based on two-way fixed effects models, while even-numbered columns (FD) show estimates based on first difference models. All models include a set of controls, with dummy variables included for missing data. Standard errors are clustered at the CCG-intersection level and given in parentheses. The Kleibergen-Paap Wald F-statistic tests for weak instruments in the first stage under clustered standard errors. The Wu-Hausman statistic tests for endogeneity. P-values for diagnostic tests are reported in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

E. Avoidable Mortality Definition

Source: Office for National Statistics (2025d)

Table E1: Avoidable Mortality Definition

Condition group and cause	ICD-10	Treatable	Preventable
Infectious Diseases			
Intestinal diseases	A00-A09		Yes
Diphtheria, Tetanus, Poliomyelitis	A35, A36, A80		Yes
Whooping cough	A37		Yes
Meningococcal infection	A39		Yes
Sepsis due to streptococcus pneumonia and sepsis due to haemophilus influenzae	A40.3, A41.3		Yes
Haemophilus influenza infections	A49.2		Yes
Sexually transmitted infections (except HIV/AIDS)	A50-A60, A63, A64		Yes
Varicella	B01		Yes
Measles	B05		Yes
Rubella	B06		Yes
Viral Hepatitis	B15-B19		Yes
HIV/AIDS	B20-B24		Yes
Malaria	B50-B54		Yes
Haemophilus and pneumococcal meningitis	G00.0, G00.1		Yes
Tuberculosis	A15-A19, B90, J65	Yes (50%)	Yes (50%)
Scarlet fever	A38	Yes	
Sepsis	A40 (excl. A40.3), A41 (excl. A41.3)	Yes	
Cellulitis	A46, L03	Yes	
Legionnaires disease	A48.1	Yes	
Streptococcal and enterococci infection	A49.1	Yes	

Condition group and cause	ICD-10	Treatable	Preventable
Other meningitis	G00.2, G00.3, G00.8, G00.9	Yes	
Meningitis due to other and unspecified causes	G03	Yes	
Neoplasms			
Lip, oral cavity and pharynx cancer	C00-C14		Yes
Oesophageal cancer	C15		Yes
Stomach cancer	C16		Yes
Liver cancer	C22		Yes
Lung cancer	C33-C34		Yes
Mesothelioma	C45		Yes
Skin (melanoma) cancer	C43		Yes
Bladder cancer	C67		Yes
Cervical cancer	C53	Yes (50%)	Yes (50%)
Colorectal cancer	C18-C21	Yes	
Breast cancer (female only)	C50	Yes	
Uterus cancer	C54, C55	Yes	
Testicular cancer	C62	Yes	
Thyroid cancer	C73	Yes	
Hodgkin's disease	C81	Yes	
Lymphoid leukaemia	C91.0, C91.1	Yes	
Benign neoplasm	D10-D36	Yes	
Endocrine and metabolic diseases			
Nutritional deficiency anaemia	D50-D53		Yes
Diabetes mellitus	E10-E14	Yes (50%)	Yes (50%)
Thyroid disorders	E00-E07	Yes	
Adrenal disorders	E24-E25 (except E24.4), E27	Yes	
Diseases of the nervous system			
Epilepsy	G40, G41	Yes	
Diseases of the circulatory system			
Aortic aneurysm	I71	Yes (50%)	Yes (50%)
Hypertensive diseases	I10-I13, I15	Yes (50%)	Yes (50%)

Condition group and cause	ICD-10	Treatable	Preventable
Ischaemic heart diseases	I20-I25	Yes (50%)	Yes (50%)
Cerebrovascular diseases	I60-I69	Yes (50%)	Yes (50%)
Other atherosclerosis	I70, I73.9	Yes (50%)	Yes (50%)
Rheumatic and other heart diseases	I00-I09	Yes	
Venous thromboembolism	I26, I80, I82.9	Yes	
Diseases of the respiratory system			
Influenza	J09-J11		Yes
Pneumonia due to streptococcus pneumonia or haemophilus influenza	J13-J14		Yes
Chronic lower respiratory diseases	J40-J44		Yes
Lung diseases due to external agents	J60-J64, J66-J70, J82, J92		Yes
Upper respiratory infections	J00-J06, J30-J39	Yes	
Pneumonia, not elsewhere classified or organism unspecified	J12, J15, J16-J18	Yes	
Acute lower respiratory infections	J20-J22	Yes	
Asthma and bronchiectasis	J45-J47	Yes	
Adult respiratory distress syndrome	J80	Yes	
Pulmonary oedema	J81	Yes	
Abscess of lung and mediastinum pyothorax	J85, J86	Yes	
Other pleural disorders	J90, J93, J94	Yes	
Diseases of the digestive system			
Gastric and duodenal ulcer	K25-K28	Yes	
Appendicitis	K35-K38	Yes	
Abdominal hernia	K40-K46	Yes	
Cholelithiasis and cholecystitis	K80-K81	Yes	
Other diseases of gallbladder or biliary tract	K82-K83	Yes	
Acute pancreatitis	K85.0, K85.1, K85.3, K85.8, K85.9	Yes	

Condition group and cause	ICD-10	Treatable	Preventable
Other diseases of pancreas	K86.1, K86.2, K86.3, K86.8, K86.9	Yes	
Diseases of the genitourinary system			
Nephritis and nephrosis	N00-N07	Yes	
Obstructive uropathy	N13, N20-N21, N35	Yes	
Renal failure	N17-N19	Yes	
Renal colic	N23	Yes	
Disorders resulting from renal tubular dysfunction	N25	Yes	
Unspecified contracted kidney, small kidney of unknown cause	N26-N27	Yes	
Inflammatory diseases of genitourinary system	N34.1, N70-N73, N75.0, N75.1, N76.4, N76.6	Yes	
Prostatic hyperplasia	N40	Yes	
Pregnancy, childbirth and the perinatal period			
Tetanus neonatorum	A33		Yes
Obstetrical tetanus	A34		Yes
Pregnancy, childbirth and the puerperium	O00-O99	Yes	
Certain conditions originating in the perinatal period	P00-P96	Yes	
Congenital malformations			
Certain congenital malformations (neural tube defects)	Q00, Q01, Q05		Yes
Congenital malformations of the circulatory system (heart defects)	Q20-Q28	Yes	
Adverse effects of medical and surgical care			
Drugs, medicaments and biological substances causing adverse effects in therapeutic use	Y40-Y59	Yes	

Condition group and cause	ICD-10	Treatable	Preventable
Misadventures to patients during surgical and medical care	Y60-Y69, Y83-Y84	Yes	
Medical devices associated with adverse incidents in diagnostic and therapeutic use	Y70–Y82	Yes	
Injuries			
Transport Accidents	V01-V99		Yes
Accidental Injuries	W00-X39, X46-X59		Yes
Intentional self-harm	X66-X84		Yes
Event of undetermined intent	Y16-Y34		Yes
Assault	X86-Y09, U50.9		Yes
Alcohol-related and drug-related deaths			
Alcohol-specific disorders and poisonings	E24.4, F10, G31.2, G62.1, G72.1, I42.6, K29.2, K70, K85.2, K86.0, Q86.0, R78.0, X45, X65, Y15		Yes
Other alcohol-related disorders	K73, K74.0-K74.2, K74.6		Yes
Drug disorders and poisonings	F11-F16, F18-F19, X40-X44, X85, Y10-Y14		Yes
Intentional self-poisoning by drugs	X60-X64		Yes
Provisional assignment of new diseases			
COVID-19	U07.1-U07.2		Yes

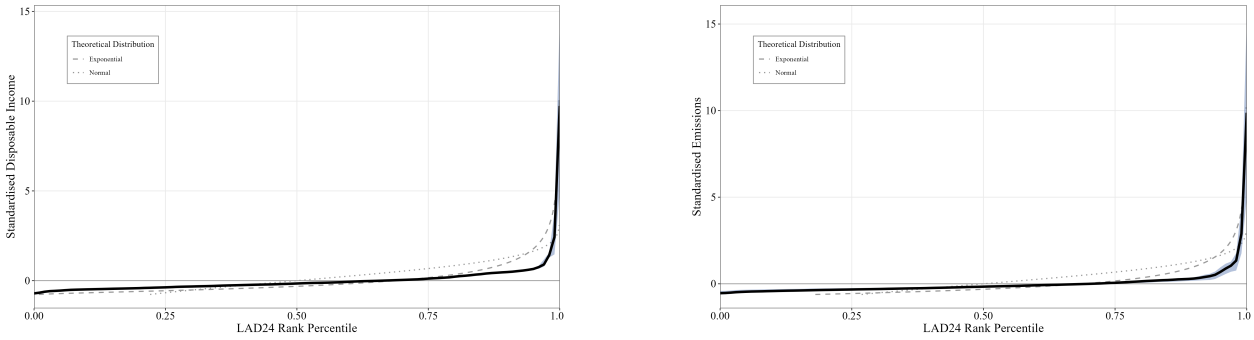
F. Distribution of Variables

This section details the empirical justification for the variable transformations described in Table 3 of the main text. To determine whether control variables should enter the model in levels or logarithmic form, we examined the rank-size distribution of all continuous variables against theoretical benchmarks.

Figures F1 and F2 display the standardised distribution of each variable (solid black line) plotted against its rank percentile. Two theoretical reference distributions are provided for comparison:

1. Normal Distribution (Dotted Line): If the empirical data follows this S-curve, the variable is approximately normally distributed and is included in the model as levels (Figure F2).
2. Exponential Distribution (Dashed Line): If the empirical data tracks this curve or exhibits a sharp convex deviation at the upper percentiles (a "heavy tail"), the variable is right-skewed (Figure F1).

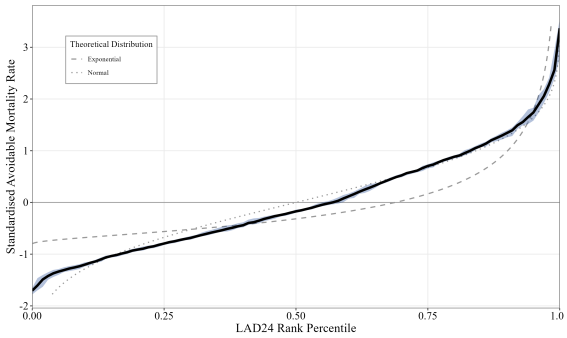
Consistent with the methodology in the main text, we apply log transformations only to variables demonstrating severe right-skewness, specifically those where the upper tail diverges significantly from the normal approximation towards the exponential curve (e.g., Emissions and Crime). This transformation normalises the distribution and prevents outliers from driving the regression estimates. Variables that closely track the normal distribution reference line (such as Wellbeing and Income) are maintained in levels to preserve interpretability and avoid overfitting.



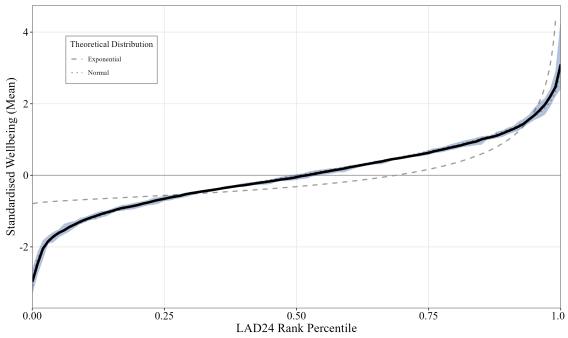
(a) Disposable Income Per Head

(b) Greenhouse Gas Emissions

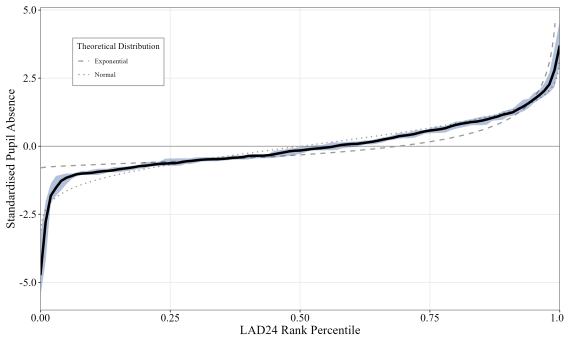
Figure F1: Distribution of Control Variables Transformed to Logs



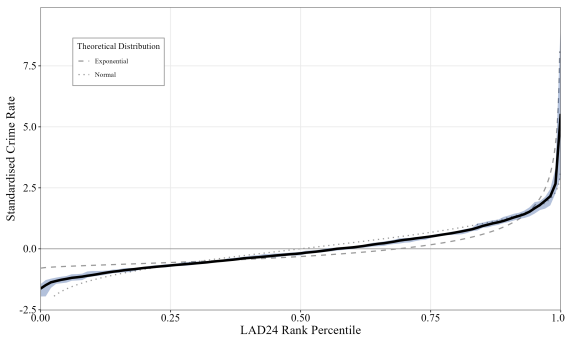
(a) Avoidable Mortality Rates



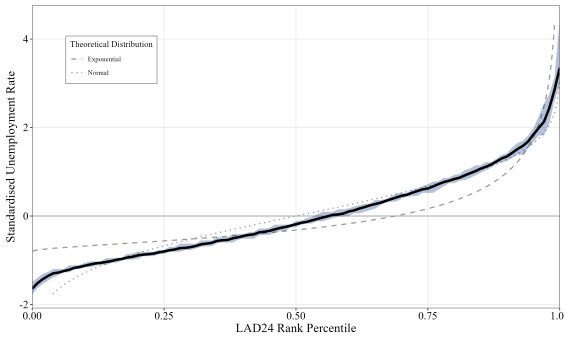
(b) Wellbeing Scores



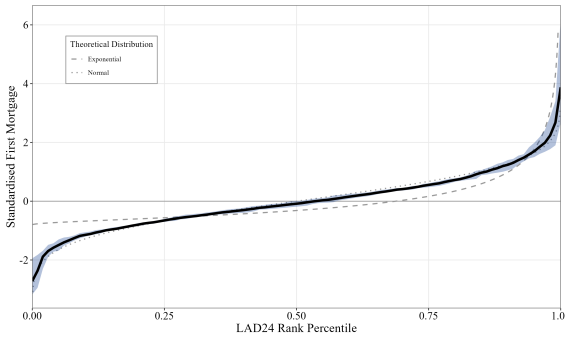
(c) Pupils Absence Rate



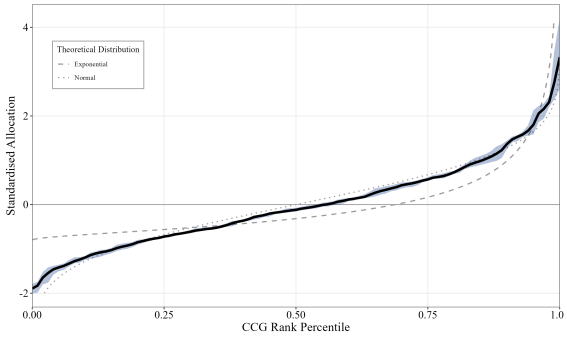
(d) Crime Rates



(e) Unemployment Rate



(f) First Buyer Mortgage Rate



(g) Allocation Per Head

Figure F2: Distribution of Control Variables Kept in Levels