Tracing the Genetic Footprints of the UK National Health Service

Nicolau Martin-Bassols[1], Elisabetta De Cao[1], Pietro Biroli[1], Massimo Anelli[2], Silvia Mendolia[3]

[1]University of Bologna, [2]University of Bocconi, [3]University of Turin

6 May 2025

PRELIMINARY DRAFT – PLEASE DO NOT CIRCULATE OR CITE

Abstract

Estimated treatment effects of early-life shocks or interventions may be biased by selective survival or mortality, which is often difficult to detect and account for. This paper introduces a method for identifying this type of selection using genetic data. In the absence of selection, Polygenic Indices (PGIs), predetermined at birth, should be equally distributed across treatment and control groups. Deviations after a shock or a policy indicate selection. We apply this approach to the introduction of the National Health Service (NHS) in July 1948 in England and Wales, using a regression discontinuity design. We find that the NHS reduced infant mortality by 6.16 per 1,000 births—a decline of 18%—likely due to increased hospital births. This mortality decline altered the genetic composition of surviving cohorts: NHS exposure increased PGIs for adverse conditions (e.g., depression, chronic obstructive pulmonary disease) and decreased PGIs for positive traits (e.g., self-rated good health). Once selective survival is accounted for, estimated long-term health effects of the NHS attenuate, highlighting the importance of correcting for survival bias.

Keywords: Early-life, Health insurance, Survival bias, Genetics, National Health System

JEL classification: I10, I38, C21

[1] nicolau.martinbassols@unibo.it
[1] elisabetta.decao@unibo.it
[1] pietro.biroli@unibo.it
[2] massimo.anelli@unibocconi.it
[3] silvia.mendolia@unito.it

1 Introduction

Extensive evidence shows that early-life conditions can impact an individual's life-long health and human capital outcomes (e.g., Almond & Currie, 2011; Currie et al., 2014).Critical periods extend beyond the in-utero phase, including birth and infancy shaping both physical and cognitive development (Currie & Rossin-Slater, 2015; Dominguez-Bello et al., 2016; Penders et al., 2006).

However, one challenge faced by many of these studies is selective survival or mortality. A severe early-life shock, such as famine, or a public policy, such as hospital care at birth can indeed lead to compositional changes in the samples of individuals observed before and after the event (Nobles & Hamoudi, 2019). This type of selection can bias treatment effects based on the survivors in unclear directions.

In this paper, we propose a novel method for detecting and quantifying selective survival or mortality using genetic data. Specifically, we explore how the distribution of polygenic indexes (PGIs) in affected populations changes after significant events. In particular, we focus on the introduction of the National Health Service (NHS) in the UK in July 1948, which is widely regarded as one of the most significant health interventions globally. It was the first health system to provide universal, free, and tax-funded care for all residents. We exploit the fact that, following the NHS's implementation, pregnant mothers suddenly became more likely to deliver in hospital settings.[1]

We use a regression discontinuity design (RDD) under local randomization. We compare individuals born before and after July 1948, who are otherwise similar in observable and unobservable characteristics. We analyze data from three key sources: the Registrar General's Statistical Review, the UK Biobank (UKB), and the English Longitudinal Study of Ageing (ELSA).

Our results are threefold. *First*, we use district- and year-level infant mortality rates (IMR) constructed from the Registrar General's Statistical Review database [2], and use year of birth as running variable to estimate the impact of the NHS on mortality. We find a decrease of 6.16 deaths per 1,000 births, or a 18.1% reduction compared to the mean. *Second*, building on the substantial effects on IMRs, we explore whether the NHS left a genetic footprint on the affected cohorts, resulting in selective survival. We leverage individual genetic data from UKB and use month of birth as running variable. We consistently find that the NHS implementation increased values in PGIs associated with qualitatively detrimental phenotypes, such as depression and chronic obstructive pulmonary disease (COPD), while reducing those linked to beneficial traits, such as self-rated health and life satisfaction with finances. For both, increases and decreases, effect sizes are as large as 5% of a standard deviation (SD). The results are robust to alternative models that

[1] In this paper, we focus only on England and Wales given that Scotland administered the NHS in a different way, while Norther Ireland established its own version through the Health Services Act.

[2] We thank Stephanie von Hinke for providing us with mortality rates calculated from the Registrar General's Statistical Review database. The rates have been harmonized to account for district boundary changes occurred over time.

examine a subsample of siblings in the UKB, comparing PGIs within families of siblings born within a short window before and after the NHS implementation. They are also consistent when considering PGIs directly linked to early-life conditions, such as gestational duration, birthweight, and head circumference. The analyses lead to similar results when performed using the nationally representative English sample ELSA, which is the only available dataset with genetic data cohorts old enough to be born around the NHS cutoff.[3] *Finally*, we examine the extent to which selective survival biases the estimation of long-term health effects, focusing on the phenotypes that exhibited the strongest evidence of genetic selection, namely, diagnosed depression and self-reported poor health. To this end, we estimate the effect of the NHS on these outcomes among UKB participants aged 45 to 80, controlling for the PGIs previously shown to capture selective survival. We find that the NHS is linked to a one-percentage-point increase in the likelihood of both detrimental health outcomes. However, after adjusting for PGIs, the effect sizes diminish, and standard errors increase, leading to close-to-zero estimates. These findings suggest that survival bias may significantly distort the estimation of long-term health effects.

Our work makes several contributions to the literature. Most importantly, it proposes an empirical approach to test for selective survival or mortality by exploring changes in genetic profiles before and after the implementation of an early-life policy. Two existing studies are worth mentioning. The first is by Furuya et al. (2024) who, using the UKB database, find that in-utero exposure to a higher area-level IMRs is associated with higher values in PGIs associated to better educational attainment. Their results suggest genetic selection in-utero, with more vulnerable individuals failing to surpass the survival threshold in areas with higher IMR, which is reflected in changes in the educational attainment PGI. However, their work is associational, and parents of individuals with different genetic endowments may select into different IMR areas. The second study is by Zhou et al. (2024) who explore whether the Dutch Hunger Winter causally affected the prevalence of PGIs related to BMI.[4] They find close-to-zero effects, which could be driven by their focus on the BMI PGI, which may not be relevant to resilience in-utero, the dimension that would be affected by selective mortality. We contribute by expanding the set of PGIs offering a more comprehensive way to study survival bias. With the decreasing cost of genetic data collection (Muir et al., 2016), this approach will become more feasible in the future, allowing for more precise assessments of selection effects in a broader range of contexts.

Secondly, we contribute to the early-life literature by addressing survival bias in the analysis of treatment effects on survivors, either short- or long-term. Our results show that PGIs as controls leads to meaningful shifts in long-term estimates, suggesting that selective survival can distort treatment effect interpretation. While PGIs capture only a portion of outcome variation, their inclusion helps rebalance treatment and control groups and partially correct for selection. These findings offer the first suggestive evidence that

[3] We aim at replicating the analyses on the UK Household Longitudinal Study if granted access.

[4] Zhou et al. (2024) build their work on the classic study by Stein et al. (2004) and Scholte et al. (2015).

PGIs may be a valuable tool for mitigating survival bias in long-term analyses.

Finally, our work contributes to the literature examining the impact of early childhood health interventions. Research on the US Medicaid shows how this health insurance policy leads to reduction in mortality, as well as, better health and human capital outcomes in the long-run (Currie et al., 2008; Brown et al., 2019; Wherry & Meyer, 2016). Research on universal large-scale health policies, mostly coming from the Nordic countries also highlights important positive impacts (for a recent review see Wüst (2022)). The Well-Child Programs, for example, introduced in the 1930s in Denmark, Norway, and Sweden, targeted infants within their first year of life, leading to significant improvements in infant survival ratesThese early interventions reveal enduring health benefits, including reduced cardiovascular disease incidence and lower age-specific mortality later in life. Although the NHS is one of the most important public health policies of the 20th century (Rivett, 1998), little empirical research has explored its broad effects at the population level. To date, only one study investigates the causal impact of the NHS in the long-run and into the second generation (Jeffrey, 2023). Her results show that individuals exposed to the NHS at birth report better self-reported health in adulthood, are less likely to suffer from long-term limiting illnesses, and are more likely to obtain a degree or a higher qualification. Our long-term findings using ELSA partially align with these results, showing improvements in self-reported health as well. Interestingly, Jeffrey (2023) finds a marginally significant worsening of the second generation's educational outcomes which she explains with the first generation negatively selecting into fertility. Our research contributes new evidence on the NHS immediate effects. It shows that the NHS significantly reduced infant mortality and provided greater protection for the most vulnerable individuals—particularly those with genetic traits associated with poorer health outcomes.

The rest of the paper is organized as follows. The next section describes the NHS and the healthcare system in the UK prior to its implementation. We then provide an introduction to genetics, including a discussion of Genome-Wide Association Studies (GWAS) and PGIs. Sections 3 and 4 respectively detail the data and the methodology used in the analysis. In Sections 5-6 and 7 we respectively present the results on infant mortality, PGIs and long-term health outcomes. Finally, the paper concludes in Section 8 with a discussion of the findings and their implications for both research and policy.

2 Institutional Background and Genetic Measures

2.1 History of the NHS

Before the establishment of the NHS in 1948, healthcare in England and Wales was largely private, requiring patients to pay for medical services, with limited free care available through various fragmented systems.[5] Voluntary hospitals, funded by charitable donations, provided some treatment, while local authority hospitals, remnants of the Poor Law

[5]Scotland and Northern Ireland developed at the same their own NHS. Given some differences in the provision of care and implementation, we will only focus on England and Wales.

system, catered mainly to the indigent. The 1944 Education Act introduced free medical services for schoolchildren, and public health initiatives, such as tuberculosis sanatoria and vaccination programs, aimed to control infectious diseases. Additionally, the National Insurance Act of 1911 created a system of 6,000 Approved Societies that offered basic healthcare to insured workers, covering about 25% to 33% of the population but excluding hospital treatment, medications, and care for dependents, leaving many without adequate medical support (Rivett, 1998).

On 5 July 1948, the UK government introduced the National Health Service (NHS), revolutionizing healthcare by making it universally accessible and free at the point of use, funded through general taxation. Built on three core principles, free provision of healthcare, access based on clinical need rather than financial means, and the equalization of medical services, the NHS aimed to eliminate disparities in care. The response was immediate and overwhelming; by July 6, 84% of the population had already registered with a doctor, rising to 91% by the end of the month. By the close of 1948, an astonishing 96% of English and Welsh people were enlisted with a general practitioner, marking a swift and widespread embrace of the new system (Rivett, 1998).

There were important supply-side changes. The NHS was launched without any significant increase in medical staff - no new doctors or nurses were added. What fundamentally changed was the system through which people accessed and paid for care. In England and Wales, the NHS assumed control of over 2,835 hospitals, including both voluntary and municipal facilities, totaling nearly 388,000 beds, many in mental health institutions (Rivett, 1998). General practitioners (GPs) transitioned to independent contractors, receiving fixed fees per treatment as determined by Executive Councils. By September 1948, 18,165 of the 21,000 GPs had enrolled as NHS contractors. These structural changes facilitated an immediate surge in healthcare utilization, with a 13% increase in consultations and a doubling of prescriptions for proprietary medicines, reflecting the rapid expansion of access to medical services (Rivett, 1998). Major hospital construction remained limited until 1955, with funding restricted mostly to upgrades of specific departments and essential infrastructure.

2.2 NHS Impact on Birth Health: Mechanisms

The introduction of the NHS allowed pregnant mothers to deliver in a hospital setting, and to seak prenatal and postnatal care free of charge. While there is limited evidence regarding the proportion of mothers directly affected by this policy, Jeffrey (2023) finds that female medical consultations increased post NHS.[6] Health improvements for children affected by the policy at the time of birth could be driven by three main channels. First, the shift to hospital-based childbirth has been associated with significant reductions in maternal mortality, as access to skilled medical professionals, sterile environments, and

[6]She used the Survey of Sickness, commissioned by the Registrar General. Conducted between 1943 and 1952, the survey involved monthly interviews with a nationally representative sample of adults (about 4,000 each month) and living in various regions across England and Wales.

emergency interventions greatly improve survival rates (Symonds et al., 2023). This reduced maternal mortality would have benefited children throughout their development. Second, infants born in hospital settings experience lower rates of infections, reduced complications such as neonatal hypoxia, and improved outcomes for preterm births due to improved neonatal care (Backes & Scrimshaw, 2020). Finally, the introduction of universal healthcare mitigates the financial burden on families, reducing out-of-pocket expenses related to childbirth and allowing for broader access to essential maternal and infant care services, which in turn supports long-term economic stability and well-being (Bufe et al., 2021). In our analysis, we are not be able to isolate the effect of each of these channels. Nevertheless, all these mechanisms have likely contributed to improving individuals' health, and the bundled impact of these channels is the policy-relevant one.

2.3 Genetics, GWAS, and PGIs

Genetics is the study of heredity and the variation of inherited characteristics. Deoxyribonucleic acid (DNA), the fundamental molecule found in all living cells, consists of two long chains of nucleotides, each made up of three components: a phosphate group, a sugar molecule, and a nitrogenous base. The four types of nitrogenous bases in DNA are adenine (A), guanine (G), cytosine (C), and thymine (T). These nucleotides form the genetic code, which encodes genes that instruct cells to produce amino acids, the building blocks of proteins. Just as instructions are organized in books, DNA is organized into 22 pairs of chromosomes, one set inherited from the mother and one from the father, along with two sex chromosomes that determine biological sex (Alberts et al., 2015).

While 99% of human DNA is identical across all people, Single Nucleotide Polymorphisms (SNPs) are small genetic variations that occur in the DNA sequence of individuals. SNPs represent the 1% of genetic differences, consisting of single nucleotide changes, either A, T, C, or G. Unlike genes, which are long sequences of DNA that encode the instructions for building proteins and regulating cellular functions, SNPs are single-point variations that can occur within or between genes. While genes provide the blueprint for proteins, SNPs can influence how those proteins are produced or function, affecting traits such as physical appearance, disease susceptibility, and drug response. These variations can occur in non-coding regions or within the coding regions of genes, impacting gene expression or protein function, thereby contributing to phenotypic differences among individuals (Alberts et al., 2015).

Genome-Wide Association Studies (GWAS) are a powerful, hypothesis-free study design that scans the entire genome to identify genetic variants associated with complex traits or diseases. By analyzing large samples and often performing meta-analysis to increase statistical power, GWAS aims to uncover SNPs linked to the phenotypes of interest. The findings from these studies have led to the identification of specific, fixed positions on chromosomes, known as loci, where particular genes or genetic markers are located. These loci are associated with various diseases and traits, offering valuable insights into human

phenotypes. However, while GWAS can highlight associations between genetic variants and traits, the underlying biological mechanisms remain not fully understood, and these studies do not establish causal relationships (Benjamin et al., 2024; Biroli et al., 2025).

PGIs are genetic tools constructed by aggregating the weighted average of effect sizes derived from GWAS. In these indexes, larger effect variants contribute more significantly, reflecting the cumulative genetic contribution to specific phenotypes. PGIs integrate information from multiple loci, offering a comprehensive assessment of an individual's genetic predisposition to traits such as disease risk, educational attainment, or behavior. Each locus (singular of loci) can have different versions of a gene, known as alleles. The major allele is the most common variant in a population, while the minor allele is the less frequent one. An individual's genotype at a given locus, denoted as G_{ij}, can take values of 0, 1, or 2, representing the number of minor alleles they carry (e.g., 0 for no copies, 1 for one copy, and 2 for two copies). The formula for calculating a PGI is represented as:

$$PGI_i = \sum_{j=1}^{n} W_j G_{ij}$$

where W_j represents the effect size of the j-th variant, and G_{ij} denotes the genotype of the i-th individual at the j-th locus (Benjamin et al., 2024; Biroli et al., 2025).

Some of the main limitations of PGIs include their predictive power, as they often explain only a small portion of the variation in complex traits or diseases, with environmental factors and gene-environment interactions playing significant roles that PGIs cannot capture. Additionally, PGIs do not establish causality; they are associated with genetic variants linked to traits, but many SNPs included may be in linkage disequilibrium with other variants that truly influence the phenotype, making direct causality difficult to infer. Population stratification is another limitation, as PGIs developed using datasets from specific populations may lose predictive accuracy when applied to individuals from different populations, reducing their generalizability. Environmental confounding is also a concern, as PGIs may overestimate the genetic contribution to a trait while neglecting complex gene-environment interactions. Finally, ethnic and ancestral bias can affect the performance of PGIs, particularly when they are applied to underrepresented populations, such as non-European groups, leading to biased predictions and limited utility across diverse populations (Biroli et al., 2025).

3 Data

3.1 Registrar General's Statistical Review

The Registrar General's Statistical Review is an annual publication produced by the Registrar General of the UK, which summarized and analyzed vital statistics related to population, health, mortality, births, and other demographic trends. It provides insights into the distribution, size, and characteristics of the population, including information on mar-

riage, death, and disease rates. The review is a key resource for understanding the social and health trends within the UK during its time of publication. The Registrar General is responsible for overseeing the national census and the registration of births, deaths, and marriages, and the Statistical Review was one of the key ways in which the data collected by these activities was disseminated (London & Region, 1952; Carrier, 1960).

We take district-year-level data from the Registrar General's Statistical Review for 1,472 districts from the years 1938 to 1958. We combine data on the number of live births and mortality before age one. This allows us to create a measure for the IMR at a year-district level, which we use to explore whether this measure is affected by the introduction of the NHS.

$$\text{IMR} = \frac{\text{Number of deaths of infants under one year}}{\text{Number of live births}}$$

For our main regressions, considering the years 1946 to 1950, the sample consists of 1,472 district-observation units, with 2,924 observations pre-NHS implementation and 4,410 observations post-NHS implementation. The panel is not fully balanced due to missing observations in some time periods.

3.2 UK Biobank

The UKB is a large-scale biomedical database and research resource containing in-depth genetic, health, and lifestyle information from approximately 500,000 UK participants aged 40 to 69 at recruitment (2006 to 2010). It integrates genetic data, physical and cognitive measures, biochemical assays, and extensive longitudinal health records, making it a valuable resource for studying complex disease etiology and risk factors. With its linkage to national health registries, UKB supports a wide range of research in genetics, epidemiology, and public health, enabling discoveries on disease prevention, treatment, and policy implications.

The UKB recruited participants through 22 assessment centers across the UK, inviting individuals via mailed letters. Approximately 9.2 million invitations were sent, yielding a response rate of around 5%. This low response rate resulted in a sample that is not nationally representative, as it is biased toward individuals with higher education, higher socioeconomic status (SES), and better overall health. Consequently, while UKB provides valuable insights into genetic and health-related research, findings must be interpreted with caution (Swanson, 2012). Recent efforts have aimed to improve national representativeness in the UKB by creating sample weights based on the UK census (van Alten et al., 2024).

The UKB has information on the month of birth of individuals, which allows us to identify if individuals where born before or after the implementation of the NHS. It also collects self-reported health measures (rated on a scale from one to six, with higher values indicating worse health status) and measures for BMI and Body Fat Percentage (BFP), all collected later in life.

The UKB collected blood samples from approximately 500,000 participants, and geno-

typing was performed using the UKB Axiom Array, designed to capture genome-wide genetic variation, including SNPs and short insertions and deletions (indels). DNA was extracted from the stored blood samples for this genotyping process. In addition, the UKB undertook whole exome sequencing for approximately 470,000 participants and whole genome sequencing for all 500,000 participants (Bycroft et al., 2018).

Given the widespread use of the UKB in most GWAS studies to date, a key challenge in constructing PGIs using UKB data is avoiding overfitting. Using the UKB for both GWAS discovery and main analysis risks circularity. To mitigate this, the UKB has been systematically partitioned into three subsets, where GWAS weights are derived from two-thirds of the sample and applied to the remaining third (Becker et al., 2021). This approach has led to the creation of three distinct UKB PGI repositories, which can be used independently but should not be combined.

A challenge in our analysis is deciding which PGIs to examine. Testing all PGIs would require running many regressions, complicating interpretation and increasing the risk of false positives due to multiple hypothesis testing (Benjamini & Hochberg, 1995; Dunn, 1961). On the other hand, selecting only specific PGIs could introduce publication bias if decisions are influenced by significant results (Sterling, 1959). To mitigate this, we focus on PGIs classified into meaningful categories that could capture resilience at early-life, namely, cognition and education, health and health behaviors, and personality and well-being (Becker et al., 2021). We prioritize PGIs based on multiple imputations when available, and as a robustness check, we present results using all PGIs with adjustments for multiple testing. Both approaches confirm the consistency of our findings.

The initial UKB sample consists of 502,336 individuals. We first restrict the sample to individuals born in England and Wales. When the PGIs weights are derived from the same dataset used to perform the GWAS, overfitting can introduce bias. To address this issue, we rely on PGIs from the Polygenic Index Repository, which mitigates overfitting by splitting the UK Biobank sample into three equally sized subsamples. For each subset, summary statistics from the other two are used to construct independent weights. As a result, our estimating sample consists of one-third of the original UK Biobank sample - the first PGI repository, which includes siblings. Although within-family analysis is not central to our work, it is used as a robustness check. These changes reduce our sample to 148,593 individuals. After narrowing the sample to the relevant time window of 18 months pre- and post-NHS, the sample shrinks to 22,748. Finally, when excluding individuals with missing month of birth or outcomes, the final sample consists of 21,344 individuals. Importantly, and as is common practice in the literature (Benjamin et al., 2024), this sample is restricted to individuals of European ancestry, primarily classified as 'whites or Caucasians'. This restriction helps reduce bias from population stratification, systematic differences in genetic ancestry that can confound associations between genetic variants and outcomes, ensuring greater comparability with the GWAS summary statistics used to construct polygenic indices.

3.3 English Longitudinal Study of Ageing

The ELSA is a large, ongoing study that collects data on the health, economic, and social circumstances of over 18,000 individuals aged 50 and older in England. The study, which began in 2002, follows participants over time to examine various factors affecting aging, including physical and mental health, income, employment, and quality of life. ELSA provides valuable insights into the aging process, helping to inform policies related to health, social care, and pensions. Crucially, this database is nationally representative of the English population.

The restricted data version of ELSA includes information on the month of birth of individuals, which we use to identify those born before or after the implementation of the NHS. Furthermore, ELSA includes self-reported health data (rated on a scale from one to five, with higher values indicating worse health status) and BMI measurements taken later in life. This information is used to explore the long-term effects of the NHS.

In 2013 and 2014, genome-wide genotyping was conducted at University College London (UCL) on 7,597 ELSA participants of European ancestry using Illumina HumanOmni2.5 BeadChips, which measure approximately 2.5 million genetic markers with a 2.5% minor allele frequency (MAF). PGIs were constructed using externally valid SNP weights from GWAS. Whenever possible, meta-analyses that excluded ELSA in the discovery analysis were selected to ensure data independence. SNP weights were obtained from consortium webpages, requested from consortium authors, sourced from dbGap, or extracted from published supplemental material. This approach enabled the creation of a broad set of PGIs capturing a wide range of phenotypes (Ajnakina & Steptoe, 2022). We select PGIs across the categories of socio-economic traits, adult mental health and well-being, physical health, anthropometric traits, and intelligence (Ajnakina & Steptoe, 2022). When multiple PGIs exist for the same phenotype from different GWASs, we use the PGI derived from the most recent GWAS.

The initial ELSA sample consists of 21,344 individuals. For our main specification, after restricting the sample to a bandwidth of one and a half years pre- and post-NHS implementation, the number of individuals decreases to 1,110. This further declines to 705 individuals when restricting the sample to those with genotyped information, which constitutes our final sample. Again, this sample is restricted to individuals of European ancestry.

4 Methodology

4.1 Main empirical strategy

Our main identification strategy is a regression discontinuity design (RDD) under local randomization (Cattaneo & Titiunik, 2022). The implementation of the NHS occurred at the daily level, on July 5, 1948. However, given the characteristics of our data, we will consider two different discrete running variables: the year of the interview when

studying the impact of NHS on IMR, and the month of birth for the PGIs and long-term analyses. The RDD under local randomization assumes that within a sufficiently narrow window around the eligibility cutoff, the assignment to NHS exposure is effectively random. Because mothers cannot precisely control the exact day of delivery, some randomness in birth timing around the cutoff date is expected. This local as-if random variation allows for a comparison of individuals born before and after July 1948, who are otherwise similar in observable and unobservable characteristics. The key identifying assumption is that within this narrow bandwidth, the only systematic difference between individuals is their exposure to the NHS. Therefore, differences in outcomes between these groups can be attributed to the causal effect of access to the NHS. Our main specification is given by:

$$Y_i = \beta T_i + \epsilon_i$$

Y_i is the outcome of interest, either IMR or a set of different PGIs. T_i is an indicator equal to one if individual i was born in July 1948 or later (i.e., born under NHS), and zero otherwise. Under the assumption of local randomization within a narrow bandwidth around the cutoff, β captures the causal effect of NHS exposure. ϵ represents the error term. The baseline specification uses a bandwidth of two years for regressions considering IMR as the main outcome and 18 months for those on PGIs. This bandwidth choice includes at least one year before the announcement of the NHS in January 1948. Note that WWII ended in May 1945 and we do not want to select windows that overlap with the war. P-values are estimated through permutation tests, and confidence intervals under interference are constructed (CITE). Results are robust to alternative specifications considering alternative bandwidths, kernels, and multiple hypothesis testing. They are also robust to alternative empirical strategies such as standard sharp RDD, and two-sample t-tests mean comparisons.

Several potential issues threaten our identification. First, individuals might have anticipated the implementation of the NHS and postponed the decision to have children in order to be eligible to deliver in a hospital setting. However, the announcement of the NHS implementation was made in January 1948, and during the six months leading up to its implementation, it faced resistance in Parliament, with its final approval remaining uncertain (Webster, 2002). Therefore, individuals delivering at the cutoff time of the NHS implementation would have already been pregnant at the time of its announcement. Regressions exploring shorter bandwidths in the running variable rule this possibility out. Furthermore, there was only limited availability of contraceptive methods at the time, with significant breakthroughs only occurring in the 1960s when the contraceptive pill became widely accessible (Murphy, 1993; Barban et al., 2025). Accordingly, we assume that such reactions were marginal or non-existent.

A further threat to identification is the occurrence of events parallel to the introduction of the NHS, which could potentially confound its effects. For instance, bread rationing was introduced on 21 July 1946 and ended on 24 July 1948, just weeks after the NHS was

established. However, rationing during the post-war period was complex, with key events including reductions in bacon, fat, and soap rations in 1945; the introduction of bread and potato rationing due to poor harvests in 1946 and 1947; and strikes that disrupted food imports. Meanwhile, petrol rationing underwent repeated adjustments, with restrictions easing in 1948 and ending in 1950. Clothes rationing concluded in 1949, and the gradual dismantling of food rationing culminated in its complete removal in July 1954 (Zweiniger-Bargielowska, 2000; Kynaston, 2008; Zweiniger-Bargielowska, 1994).

To explore how rationing may have affected IMR, we retrieved data from The National Food Survey 1945–1949, focusing on aggregated calorie consumption from 1945 to 1950 (of Food, 1951). Daily calorie intake in 1945, the last year of the war, was 2,375 kcal; it decreased slightly to 2,307 in 1946 and 2,308 in 1947; increased to 2,387 in 1948; and reached 2,425 in 1949. These figures indicate that average calorie intake during those years was well above the minimum survival threshold of 1,500 kcal/day (Fogel, 2004). While changes in the quality of food could have led to health differences, as reported in Gracner et al. (2024), our estimates of IMR and genetic selection should be primarily driven by mortality. In other words, changes in food quality could have not affected individuals genetics, they were driven by selective mortality, and as such any occurrence of parallel events, not affecting mortality, would not confound our results.

4.2 Within-Families Analysis

The UKB includes a sample of siblings, allowing us to perform a within families analysis (Amin et al., 2022). In particular, we explore sibling differences in PGIs before and after the implementation of the NHS, under the assumption that the distribution of PGIs within families remains constant over time. We estimate the following regression model:

$$PGI_{if} = \beta_0 + \beta_1 PostNHS_{if} + \beta_2 X'_{if} + \alpha_f + \epsilon_{if}$$

In this model, PGI is a vector of the different PGI outcomes for sibling i belonging to family f, with the treatment indicator defined as born under the NHS. The vector X includes the following controls: a running numerical variable for month and year of birth (to account for cohort effects), quarter of birth fixed effects (to account for seasonality), gender, date of the interview (in combination with date of birth capturing age). Additional 10 ancestry-specific principal components (PCs) are included. Finally, we additionally control for the age gap between siblings and the gender of each respective sibling to adjust for potential fertility selection. Finally, family-specific fixed effects (α_f) are also included to account for unobserved time-invariant family-level characteristics. This model compares siblings who grew up in similar households and environments, which were likely affected by the NHS in a comparable manner. Standard errors are clustered at the family level.

4.3 Long-term health effects and survival bias

A final model is run to explore the effect of the NHS on the actual observed health outcomes at the time of the initial UKB interview.[7] This model explores how effects on long-term outcomes vary with the systematic inclusion of PGIs as controls. Sensitive results to the inclusion of PGIs as controls will suggest that the bias introduced by selective survival manifest meaningfully in the long-term outcomes estimations. The objective is to examine whether the PGIs that captured survival bias in the first part of the analysis have the potential to mitigate bias, or at least provide insight into its direction or relevance. We will estimate the following model:

$$HealthOutcome_i = \beta_0 + \beta_1 PostNHS_i + \beta_2 X_i^{'} + \beta_3 PGI_i + \epsilon_i$$

In this model, $HealthOutcome_i$ represents the health status of individual i at the time of the initial UKB interview, as measured by self-reported health and depression. The vector X controls for demographic characteristics such as the year-month of birth, quarter of birth, gender, and the date of the interview. PGI is a vector of PGIs and the first 10 ancestry specific PCs. We explore how sensitive the main estimated effect is ($\hat{\beta}_1$) are to the systematic introduction of PGIs as controls, and interpret the changes as mitigating the bias caused by selective survival.

5 The impact of the NHS on infant mortality

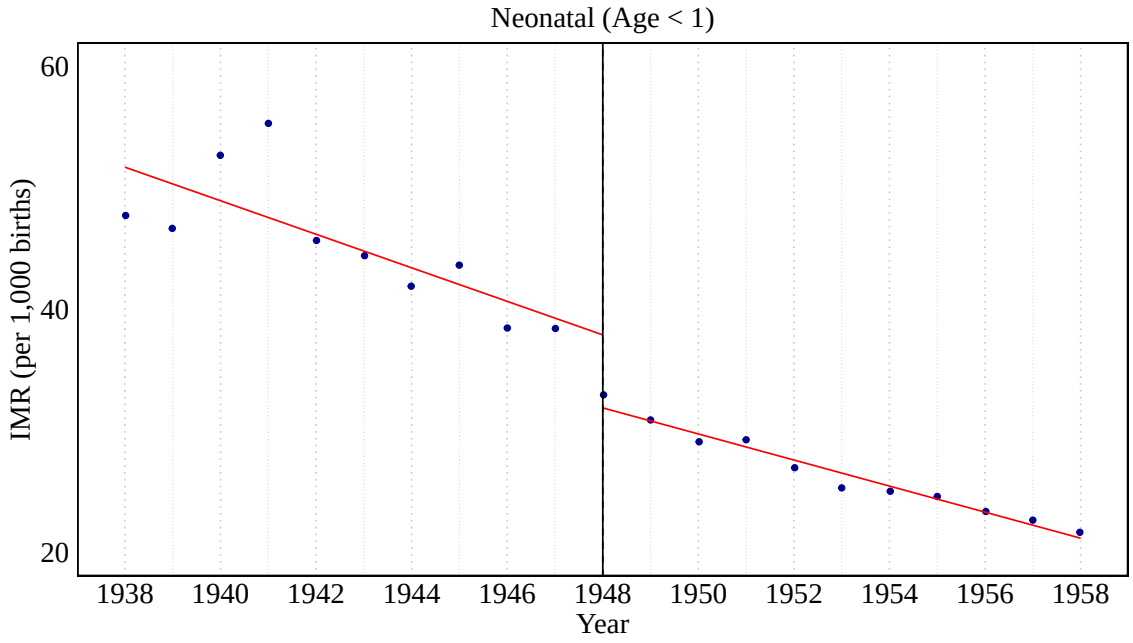
The first contribution of our work is to show of the NHS on infant mortality rates (IMR), which form the foundation for the subsequent analyses in this paper. Identifying a substantial impact on IMR is essential, as it would indicate meaningful demographic shifts that justify investigating potential genetic selection. This analysis relies on data from the Registrar General's Statistical Review, which provides detailed district-year-level mortality and birth statistics for the relevant period. Figure 1 presents compelling descriptive evidence of a significant decrease in IMR following the implementation of the NHS.

Figure 1 not only shows a sharp decline in IMR following the NHS introduction, but also aligns descriptively with major historical events, reinforcing the credibility of the data. For instance, IMR spiked in 1940–41, coinciding with the intense German bombing during WWII,[8] before dropping significantly in 1942, in line with German military focus being redirected toward the Soviet Union.[9] Another decline occurred in 1946 following the war's end, and a final, notable drop appeared in 1948 at the point of NHS implementation. After this, the IMR trend became smooth and steadily declining—consistent with gradual

[7] Some biomarkers and health outcomes are measured multiple times in the data. We consider only the first time they are measured.

[8] The bombing campaigns over British cities in 1940–41 strained healthcare infrastructure and living conditions, contributing to a 15.7% increase in IMR compared to 1939 (Gardiner, 2010).

[9] The so called Operation Barbarossa, which began in June 1941 (Erickson, 2019).



	Bandwidth two years	Bandwidth six years
Coefficient	−6.16	−10.08
p−value	0.00	0.00
Observations pre−NHS	2,942	8,822
Observations post−NHS	4,410	10,293
Average	33.96	34.76

Figure 1: Effects of the NHS on Infant Mortality Rates

Note: Source Registrar General's Statistical Review.

improvements in living standards and public health.[10]

The table at the bottom of Figure 1 presents estimates and p-values from an RDD model under local randomization, examining the causal effect of the NHS on infant mortality. Using a strict bandwidth of two years before and after the NHS implementation, the estimated effect is a reduction of 6.16 deaths per 1,000 births, corresponding to an 18.1% decrease relative to the mean (33.96). Notably, the two-year window excludes the WWII period, thereby avoiding potential contamination of the pre-NHS baseline.

Expanding the time window to six years before and after implementation yields a similar trend, with a larger estimated reduction of 10.08 deaths per 1,000 births, equivalent to a 28.99% decrease relative to the mean (34.76). This estimate is also precise. However, the sharper decline observed with the wider bandwidth appears to be driven by contamination from the WWII years. For this reason, we prefer the specification that restricts the sample to two years on either side of the NHS introduction. The main results are consistent when estimating a sharp RDD model instead of an RDD under local randomization, as shown in Table A.1.

[10]Post-war improvements in nutrition, sanitation, and medical care contributed to the ongoing decline in IMR (Marshall et al., 2019).

Based on the theoretical framework of (Nobles & Hamoudi, 2019), the NHS had a health effect that enabled a significant group of individuals to surpass the threshold for survival at birth. This created important differences between the individuals born before and after that event, differences that could be captured in individuals' genetics. In the next section, we attempt to capture these differences, tracing the genetic footprint of the NHS.

6 Selective Survival

In this section, we explore the effects of the NHS on different PGIs observed in our samples. Genetics are determined at birth and as such PGIs are exogenous to the implementation of the NHS, which should have no effect on them. However, the NHS could have influenced survival rates. If a sufficiently large number of individuals with distinct genetic traits survived after the implementation of the NHS, this could have left a genetic footprint, reflected in changes in the PGIs observed pre- and post-NHS. This way, although the NHS did not directly affected the genetics of individuals, it would have affected the genetic distribution at a population level, and this is what we explore with the models presented in this section.

We begin with the results from the UKB sample. In Figure 2, we present a forest plot illustrating the effects of the NHS on various PGIs. Each dot represents the causal effect from a separate regression, with a different PGI as the outcome, along with its p-values and 95% confidence intervals. The results show a clear trend: the values of PGIs associated with more detrimental conditions, such as depression, COPD, or ADHD, increase after the implementation of the NHS, while the ones for PGIs related to more beneficial conditions, such as self-rated health, life satisfaction with finances, or subjective well-being, decrease. Effect sizes range from -5% to 5% of a standard deviation. Even for PGIs in the middle range, although estimates are imprecise, the main coefficients point in a consistent direction, with detrimental conditions presenting positive effects and beneficial conditions negative ones. Cognitive empathy and being left out of social activities, both of which are estimated with high imprecision, are the only exceptions, two out of the 19 coefficients.

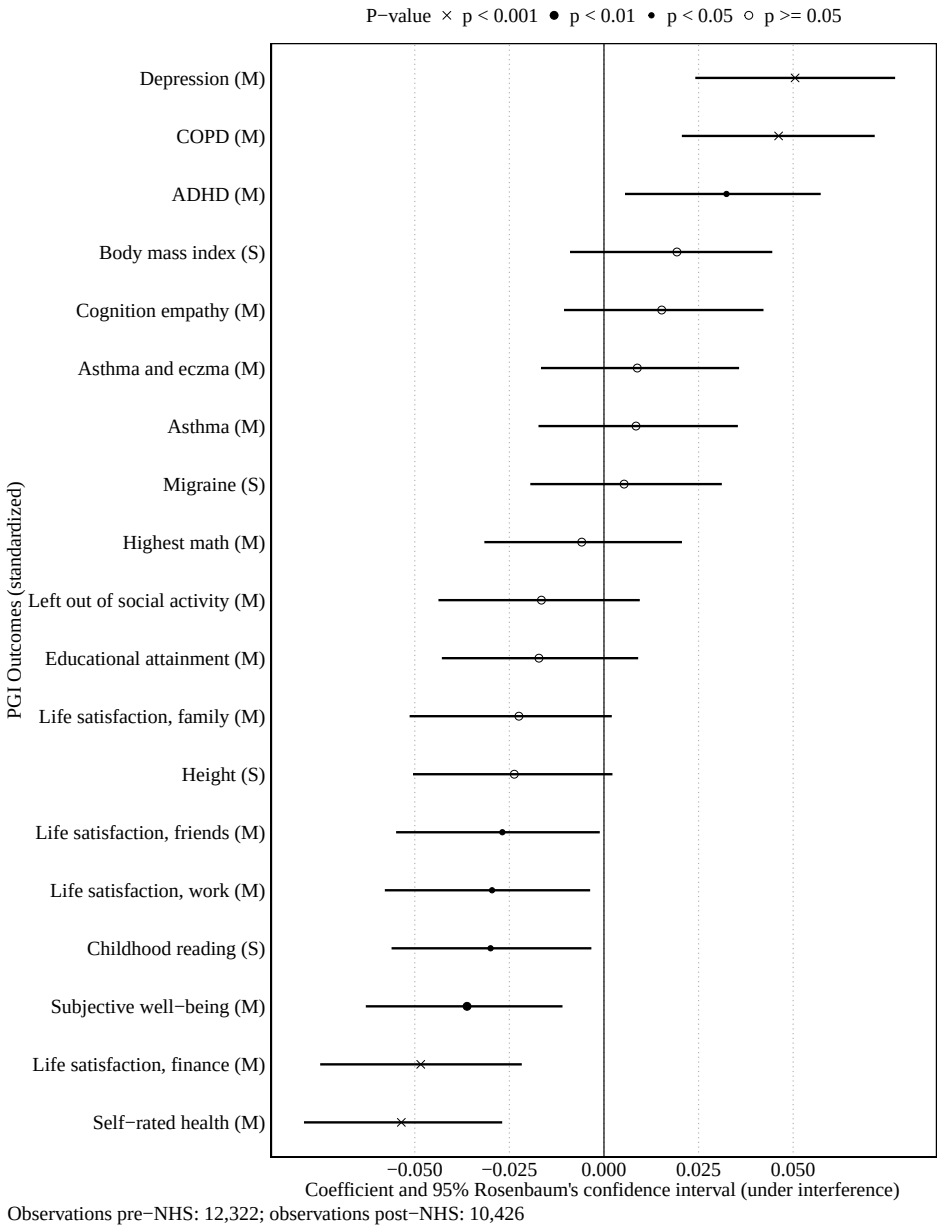


Figure 2: Effects of the NHS on PGIs

Note: Each dot represents the causal effect of a separate regression on a different outcome. P-values under permutation (1,000 permutations) and standard errors under interference. The bandwidth is 18 months pre- and post-NHS implementation. Source: UKB.

These results suggest that a higher proportion of vulnerable individuals survived birth, thanks to the implementation of the NHS. These individuals were genetically different to the pre-NHS samples. We interpret this as evidence of the survival bias caused by the NHS.

6.1 Robustness Checks

We perform multiple robustness checks. First, we replicate the results using different bandwidths, as shown in Figure B.1. This includes bandwidths of three, six, and 12

months, as well as a covariate balance window selection procedure (Cattaneo et al., 2023), which selected a bandwidth of seven months. Overall, the results maintain the same pattern as those estimated using the baseline 18-month bandwidth. The most noticeable trend is that, as the bandwidth narrows, the estimated effects become larger but also more imprecisely estimated due to the loss of statistical power. This is most evident with the three-month bandwidth, where, although the general pattern is preserved, effect sizes reach up to 15% of a SD, and some estimates become statistically insignificant (ADHD, childhood reading, and life satisfaction with friends). For this reason, we favor the more conservative specification that yields the most precisely estimated effects, namely, a bandwidth of 18 months.

Second, in Figure B.2, we present results using different kernels. While, the baseline specification uses uniform kernels, we replicate the analysis using triangular and Epanechnikov ones. Although the results are qualitatively and quantitatively similar across all kernels, the uniform kernel yields slightly smaller effect sizes. We opt for this more conservative specification, although the differences compared to other kernels are negligible.

Third, we address multiple hypothesis testing in Figure B.3 by applying both the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995) and the Bonferroni correction (Dunn, 1961). The Benjamini-Hochberg method controls the false discovery rate (FDR), offering a balance between identifying true effects and limiting false positives. In contrast, the Bonferroni correction controls the family-wise error rate (FWER), providing a more conservative adjustment that reduces the likelihood of any false positives but may increase the chance of overlooking true effects. Applying the Benjamini-Hochberg adjustment, only two outcomes lose statistical significance compared to the baseline specification: the PGIs for height and life satisfaction with family. When using the more conservative Bonferroni correction, several additional outcomes lose significance. However, self-rated health, depression, life satisfaction with finances, and chronic obstructive pulmonary disease (COPD) remain statistically significant at the 99% confidence level. Regardless of statistical significance thresholds, it's important to note the consistent qualitative gradient observed across phenotypes. This pattern suggests a directional selection, reinforcing the validity of our findings beyond mere statistical precision.

Fourth, we replicate the main analysis using a sharp RDD specification. In this approach, we fit separate first-degree polynomial functions on either side of the NHS implementation cutoff. The causal effect is estimated as the difference in the intercepts of these two regression lines at the cutoff point, capturing the discontinuity in the outcome variable attributable to the treatment. The model controls for month-year of birth, interview date, quarter of birth fixed effects, gender, and 10 ancestry-specific principal components (PCs). It is estimated with triangular kernel weights and standard errors using cluster-robust variance estimation with degrees-of-freedom adjustments, combined with nearest neighbor variance estimations. Figure B.4 shows that the sharp RDD estimates closely mirror those obtained from the local randomization RDD. The primary distinction is that the sharp RDD estimates are notably larger, reaching up to 10% of a SD. Given the

structure of our data and that genetic propensities are unlikely to exhibit temporal trends within a three-year bandwidth, we consider the local randomization RDD to be more appropriate.

Fifth, we estimate a multivariate linear regression model where the treatment variable is defined as being born after the implementation of the NHS. The model includes covariates for date of birth, date of the interview, gender, quarter-of-birth fixed effects, and ten ancestry-specific PCs. Additionally, we perform two-sample t-tests to compare the mean differences in PGI values between individuals born before and after the NHS implementation. Figure B.5 illustrates that both approaches yield results consistent with our baseline specifications, exhibiting similar qualitative and quantitative patterns. The similarity in effect sizes suggests that these mean comparison methods are inherently similar to the local randomization RDD approach.

Finally, we extend our analysis to include all available PGIs in the UKB. These encompass PGIs derived from both single- and multiple-input GWAS, as well as phenotypes related to fertility, sexual development, health behaviors, and personality traits. Figure B.6 illustrates that the results obtained using single-input and multiple-input GWAS PGIs are nearly identical, indicating that the choice of GWAS method does not significantly affect the outcomes. Furthermore, the analysis reveals that only a few of the additional outcomes exhibit statistically significant effects. Notably, these significant associations pertain to personality traits, such as loneliness and neuroticism, which exhibit higher values following the NHS implementation. This aligns with our previous findings, considering these phenotypes as detrimental and associated with mental health problems.

6.2 Family Fixed Effects

The UKB does not directly identify siblings, but it infers relatedness between individuals using the 'kinship coefficient'. This coefficient measures the probability that a random allele from one individual is identical by descent to an allele from another individual at the same genetic locus. We utilize this coefficient to identify sibling pairs among participants. We then explore differences in PGIs within siblings, while controlling for family fixed effects.

Figure 3 presents such estimates and it shows that although the coefficients are less precise and most confidence intervals overlap the zero line, the previously observed pattern is still clearly visible. The interpretation of these results is that, within families and sharing the same environments, siblings born after the implementation of the NHS have higher PGI values in phenotypes associated with detrimental conditions, while presenting lower values in conditions associated with beneficial ones. Although the estimates are large in magnitude, reaching up to 30% of the SD, this is driven by a much smaller sibling sample, consisting only of 595 families and 1,190 sibling pairs. As such, the estimates are noisier, and we believe the true effect of the NHS lies closer to our baseline specification ($\pm 5\%$ of a SD).

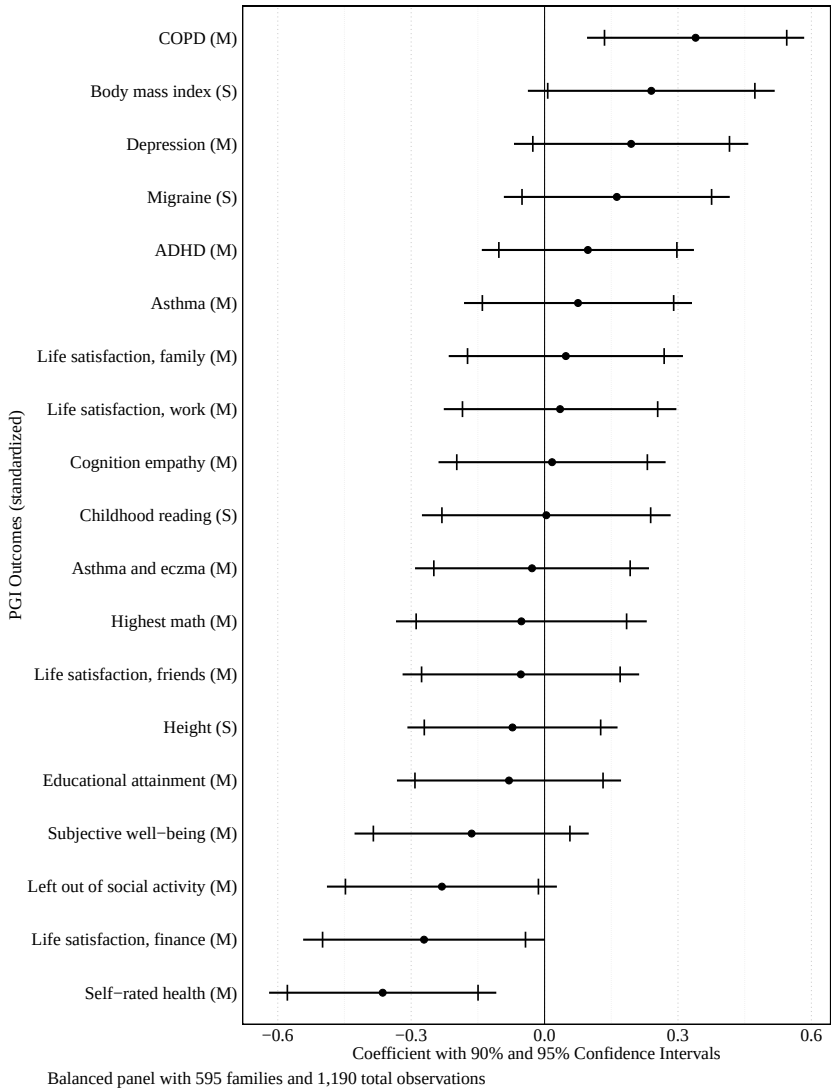


Figure 3: Effects of the NHS on PGIs using the Sibling Sample

Note: Each dot represents the causal effect of a separate regression on a different outcome. Clustered standard errors at the family level. The bandwidth is 18 months pre- and post-NHS implementation. Controls include year-month of birth, interview date (age), gender, quarter of birth fixed effects, and the first 10 ancestry-specific PCs. We additionally control for the age gap between siblings and the gender of each respective sibling. Source: UKB.

6.3 Early Growth Genetics Consortium PGIs

A key assumption of our analysis is that survival at birth, closely tied to early-life vulnerability, can be indirectly captured by examining phenotypes that manifest later in life. This approach relies on the assumption that the SNPs included in the PGIs under consideration are truly relevant to in-utero resilience. However, a more direct strategy is to focus on phenotypes that are explicitly linked to early-life survival.

To this end, we use summary statistics from the Early Growth Genetics (EGG) Consortium to construct PGIs for birthweight (Beck et al., 2021; Warrington et al., 2019), head circumference (Vogelezang et al., 2022), and gestational duration (Solé-Navais et

al., 2023). For each phenotype, we rely on the most recent available GWAS, with the exception of birthweight, where the latest study (Beck et al., 2021) is based on a twin design. Twins differ systematically from the general population, particularly in traits such as birthweight, which is highly correlated within twin pairs. Therefore, we also use birthweight summary statistics from Warrington et al. (2019). Since that GWAS includes the UKB in its discovery sample, we contacted the authors and obtained an alternative version of the summary statistics that excludes the UKB. Importantly, because none of the discovery GWAS used in our analysis include UKB participants, we are able to leverage data from the full UKB sample, rather than restricting ourselves to the first repository. This significantly increases our sample size, nearly tripling the number of observations.

Figure 4 presents the results for the PGIs based on the EGG consortium. While the coefficients for gestational duration are small and imprecisely estimated, we find negative and statistically significant effects of approximately -2.5% of a SD for both birthweight and head circumference. We interpret lower birthweight and smaller head circumference as indicators of smaller babies, capturing early-live vulnerability. As such, these results are consistent both qualitatively and quantitatively with our baseline specifications. It is noteworthy that our main results are not based on these phenotypes, as we aim to use the most well-known and widely accepted PGIs for the baseline models (Becker et al., 2021).

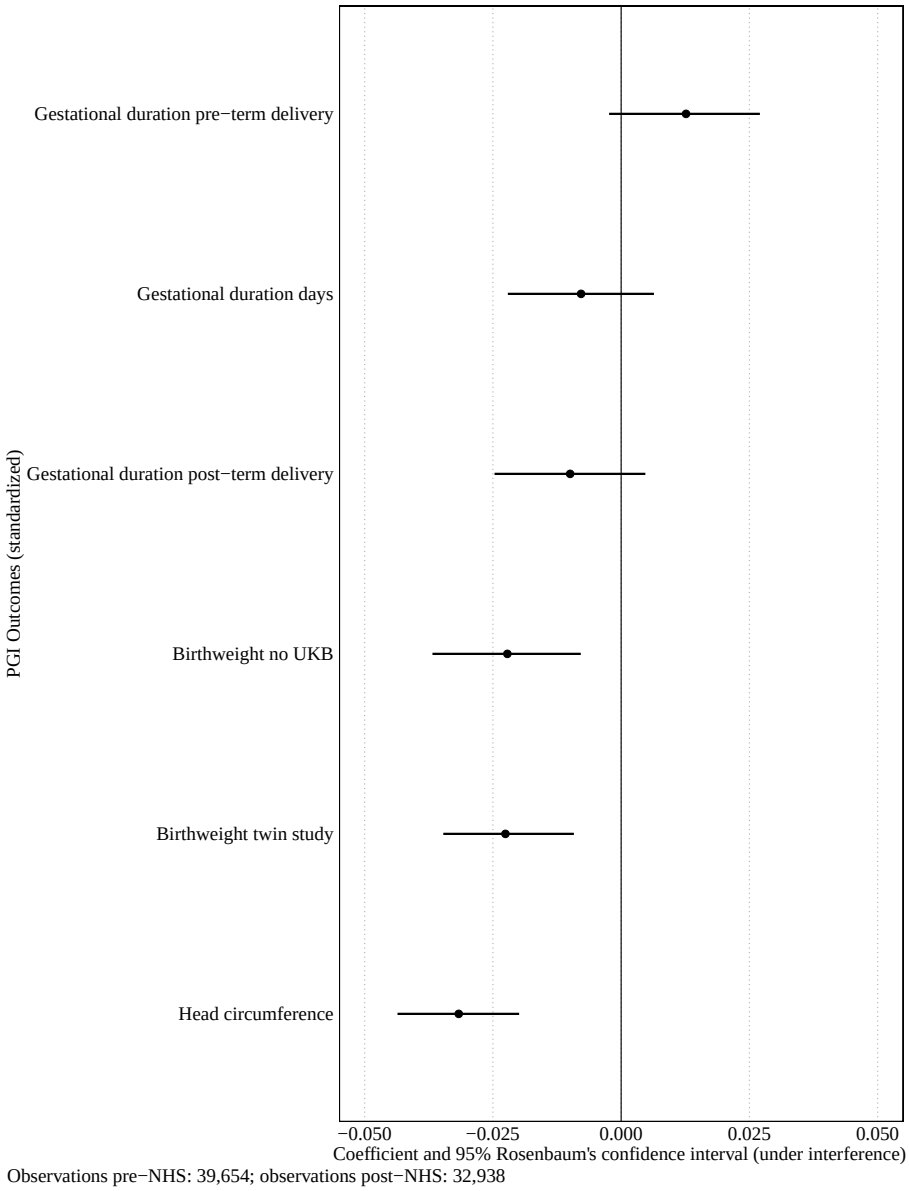


Figure 4: Effects of the NHS on 'early life' PGIs

Note: Each dot represents the causal effect of a separate regression on a different outcome. P-values under permutation (1,000 permutations) and standard errors under interference. The bandwidth is 18 months pre- and post-NHS implementation. Source: UKB.

6.4 ELSA

We replicate our baseline analysis using data from the ELSA database, which, although much smaller in size, is nationally representative of the English population. The set of PGIs in ELSA differs from the one in the UKB and is based on (Ajnakina & Steptoe, 2022). Yet, we select PGIs equivalent to the ones used in the UKB, namely in the categories of socio-economic traits, adult mental health and well-being, physical health, anthropometric traits, and intelligence. Figure 5 shows that although the results with the ELSA sample are much more imprecisely estimated, the previously described qualitative patterns still

remain. PGIs related to detrimental conditions such as migraine and schizophrenia show higher values after the implementation of the NHS, while phenotypes such as IQ, general cognition, and educational attainment (EA3) show lower ones. Estimates are noisy, with coefficient sizes ranging from -25% to 25% of a SD, largely due to the smaller sample size of only 705 observations. Therefore, we believe that the true effect sizes are closer to those in our baseline specification ($\pm$5% of a SD).

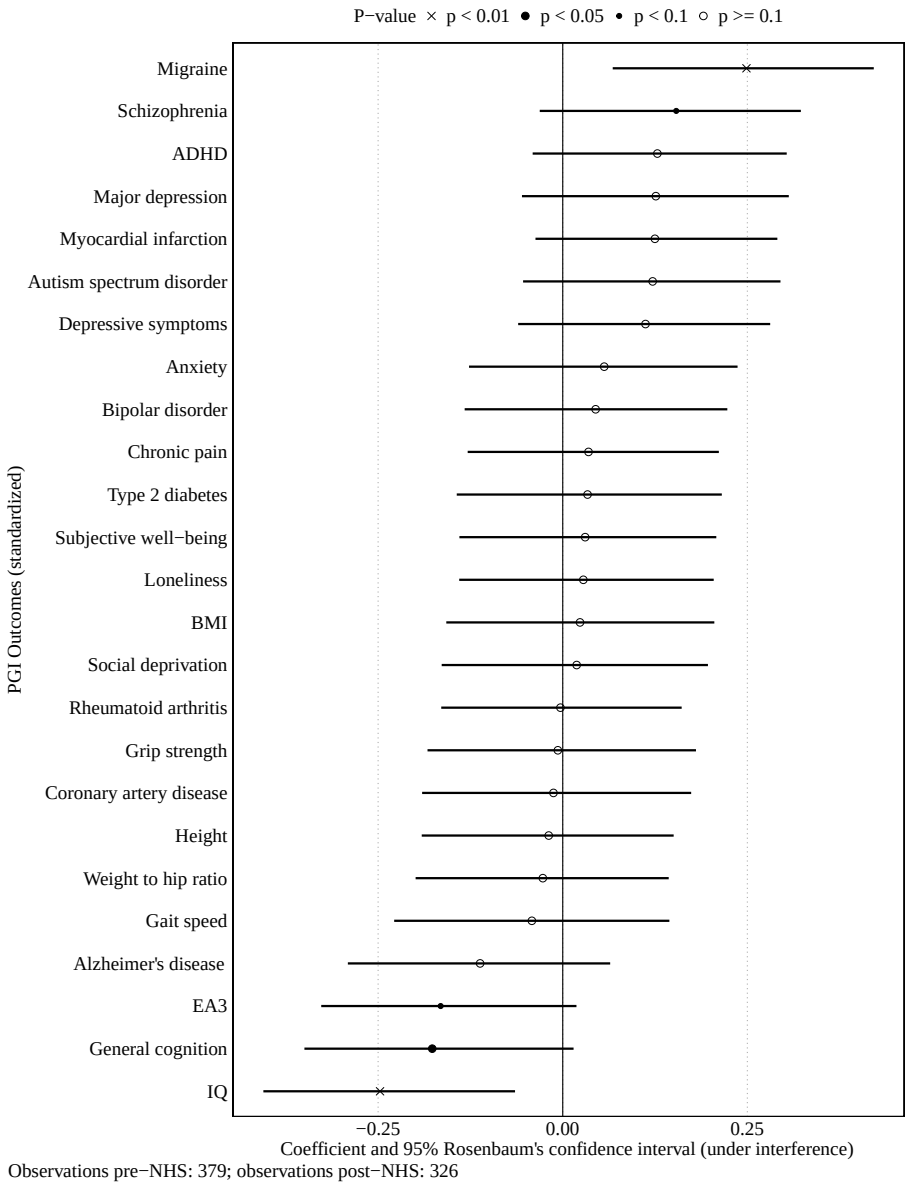


Figure 5: Effects of the NHS on PGIs using ELSA

Note: Each dot represents the causal effect of a separate regression on a different outcome. P-values under permutation (1,000 permutations) and standard errors under interference. The bandwidth is 18 months pre- and post-NHS implementation. Source: UKB.

7 Long-term Effects

In the previous section, we showed that selective survival can be tested using genetic data. But if strong selection is present, can we still trust estimates of short- and long-term effects? Or are they too biased? And can we reduce this bias or at least understand its size and direction? In this section, we explore these questions through an illustrative exercise.

Our approach involves including PGIs as control variables when examining long-term

health effects. The NHS affected survival, which in turn altered the genetic composition of the population. Genetic traits are determined at birth and are exogenous to the NHS. By controlling for PGIs, we are partially rebalancing the samples before and after the NHS implementation and helping to mitigate bias arising from selective survival.

To conduct this analysis, we focused on the phenotypes that showed the strongest effects when testing for selective survival (Figure 1), namely depression and self-rated health. In the UKB participants report whether they have ever sought consultation from a psychiatrist for issues related to nerves, anxiety, tension, or depression (average is 10.9%). Self-rated health is evaluated through a global assessment question, with responses ranging from "Excellent" to "Poor". For our analysis, we created a binary variable to indicate "Poor" health status (average is 4.5%).

We estimate multivariate regression models similar to the ones presented in Figure B.5, where the treatment variable is defined as being born under the NHS, while different sets of controls are sequentially added. Starting with a fully saturated model without controls, both diagnosed depression and self-rated bad health appear to be detrimentally affected by the NHS. Although the effect sizes are small, around a one percentage point increase in the probability of each adverse outcome, they are statistically significant. The results remain robust when basic demographics are added as controls (date of birth, date of interview, gender, and quarter-of-birth fixed effects). These findings are surprising, suggesting that the introduction of the NHS had detrimental long-term health effects.

However, once the outcome-specific PGIs are included, namely, the depression PGI for diagnosed depression and the self-rated health PGI for the self-reported poor health, along with the 10 ancestry-specific PCs, the coefficients shift toward zero and the standard errors increase. Given the small magnitude of these coefficients (around 0.5 percentage points) and the imprecise estimates, we interpret them as close-to-zero effects. Finally, the results remain stable when the full set of 19 PGIs, explored in the previous section, is included.

This seems to be a more plausible finding, suggesting that the NHS did not have discernible effects on long-term health outcomes. It indicates that initial estimates were likely biased towards worse long-term health estimates due to selective survival. Importantly, we refrain from drawing any conclusions about the long-term impact of the NHS, as this is a complex issue on its own and requires a more in-depth and focused analysis. A future, more detailed examination may uncover beneficial long-term health effects. What we highlight in this section is: selective survival can meaningfully influence results, potentially more than previously assumed, leading to biased estimates, flawed conclusions, and ultimately suboptimal policy decisions.

7.1 Robustness Checks Long-term Effects

We perform several robustness checks. First, we examine less severe health outcomes, such as whether individuals have ever consulted a family doctor, rather than a psychiatrist,

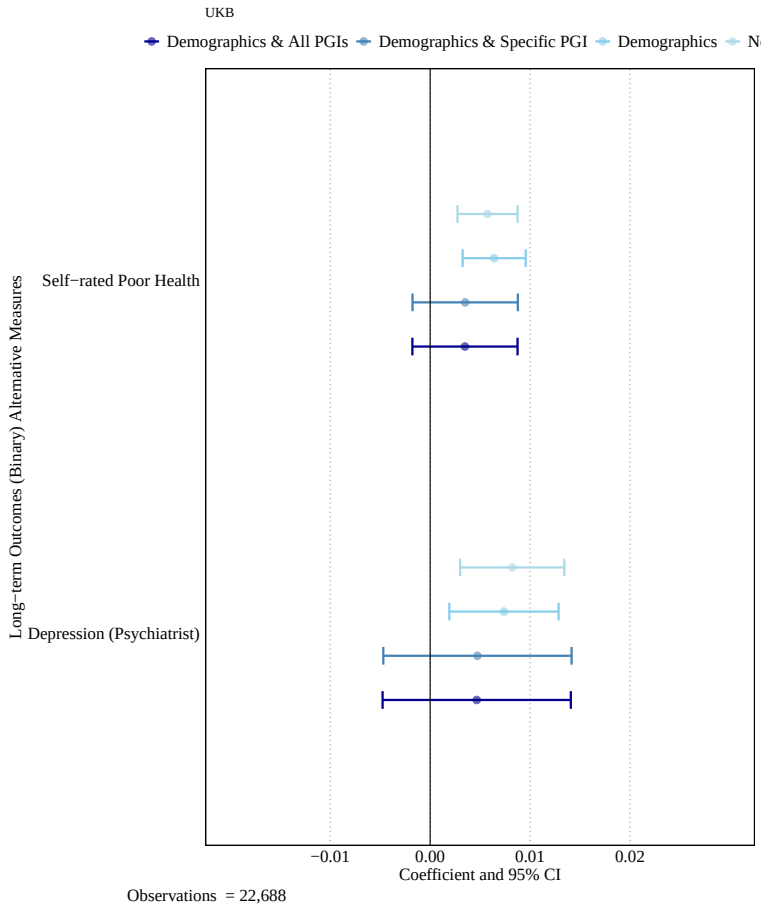


Figure 6: Effects of the NHS on Long-term Health Outcomes

Note: Each dot represents the estimated causal effect from a separate regression on a different outcome. The main effect captures the impact of being born after the implementation of the NHS, sequentially controlling for sex, month and year of birth, interview year, quarter of birth FE, PGIs, and the first 10 genetic principal components (PCs). Estimates are based on an 18-month bandwidth around the NHS implementation date. Source: UKB.

for issues related to nerves, anxiety, tension, or depression (average is 25.5%), as well as whether they self-report "Fair" or "Poor" health (average is 33.9%), rather than just "Poor" health. Figure C.1 shows a very similar trend: the inclusion of PGIs as controls has a significant impact on the estimates, compared to just controlling for basic demographics. It is worth noting that the effect size for self-reported fair and poor health remains very similar, with only the standard errors increasing. However, for depression (family doctor), the change in the main coefficient is substantial enough to reverse its sign, though it remains very close to zero, while the standard errors increase as well.

Second, we take advantage of the repeated observations of health outcomes in the UKB, which include the baseline interview and several follow-ups. These follow-ups introduce potential bias due to attrition and selective participation, so we do not consider them in our main specification. However, they may reflect worsening health outcomes with age, which might captures important variation. To account for this, we use a measure that captures whether individuals have ever reported any of the considered health outcomes,

regardless of the wave, while controlling for age (date of birth and date the interview). Figure C.2 presents the results based on this cumulative health measure, which are very similar to the main specification.

Finally, ELSA includes comparable variables for self-reported health and depression. For self-reported health, individuals are asked to rate their health on a scale from "Excellent" to "Poor"; we classify those reporting "Poor" as having poor health. For depression, the Center for Epidemiologic Studies Depression Scale (CES-D 8) psychometric test is administered, and we classify individuals with scores above three as experiencing depression (Dang et al., 2020). In Figure C.3 we do not observe a clear discontinuity in the main estimates when PGIs are included as controls, relative to including only demographic variables. Nonetheless, the direction of change in the coefficients aligns with UKB baseline results, consistently showing that selective survival biased NHS effects toward worse long-term health outcomes. Estimates are much less precise, due the smaller sample size, which leads to larger apparent effect sizes. It is worth noting that the cumulative modest shifts in the coefficients as we include different sets of controls for self-rated health make the estimate statistically significant once demographics and all PGIs are included. Again, we refrain from interpreting this as evidence of the NHS having beneficial long-term effects. However, it reinforces our point that selective survival can meaningfully bias estimates, altering conclusions and policy implications.

8 Discussion

This section is forthcoming.

References

Ajnakina, O., & Steptoe, A. (2022). *The english longitudinal study of ageing (elsa) polygenic scores. 2020.*

Alberts, B., Bray, D., Hopkin, K., Johnson, A. D., Lewis, J., Raff, M., ... Walter, P. (2015). *Essential cell biology.* Garland Science.

Almond, D., & Currie, J. (2011). Killing me softly: The fetal origins hypothesis. *Journal of economic perspectives*, *25*(3), 153–172.

Amin, V., Fletcher, J. M., Sun, Z., & Lu, Q. (2022). Higher educational attainment is associated with longer telomeres in midlife: Evidence from sibling comparisons in the uk biobank. *SSM-Population Health*, *17*, 101018.

Backes, E. P., & Scrimshaw, S. C. (2020, February). Birth settings in america: Outcomes, quality, access, and choice. In (chap. 6, Maternal and Newborn Outcomes by Birth Setting). Washington, DC: National Academies Press (US). Retrieved from `https://www.ncbi.nlm.nih.gov/books/NBK555483/`

Barban, N., De Cao, E., & Francesconi, M. (2025). Fertility, genes, and the sexual revolution. *Mimeo*.

Beck, J. J., Pool, R., Van De Weijer, M., Chen, X., Krapohl, E., Gordon, S. D., ... others (2021). Genetic meta-analysis of twin birth weight shows high genetic correlation with singleton birth weight. *Human molecular genetics*, *30*(19), 1894–1905.

Becker, J., Burik, C. A., Goldman, G., Wang, N., Jayashankar, H., Bennett, M., ... others (2021). Resource profile and user guide of the polygenic index repository. *Nature human behaviour*, *5*(12), 1744–1758.

Benjamin, D. J., Cesarini, D., Turley, P., & Young, A. S. (2024). Social-science genomics: Progress, challenges, and future directions. *NBER Working Paper*.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289–300.

Biroli, P., Galama, T. J., von Hinke, S., van Kippersluis, H., Rietveld, C. A., & Thom, K. (2025). The economics and econometrics of gene-environment interplay. (*Review of Economic Studies*, forthcoming)

Brown, D. W., Kowalski, A. E., & Lurie, I. Z. (2019, 07). Long-term impacts of childhood medicaid expansions on outcomes in adulthood. *The Review of Economic Studies*, *87*(2), 792-821. Retrieved from `https://doi.org/10.1093/restud/rdz039` doi: 10.1093/restud/rdz039

Bufe, S., Roll, S., Kondratjeva, O., Skees, S., & Grinstein-Weiss, M. (2021). Financial shocks and financial well-being: what builds resiliency in lower-income households? *Social Indicators Research*, 1–29.

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., ... others (2018). The uk biobank resource with deep phenotyping and genomic data. *Nature*, *562*(7726), 203–209.

Carrier, N. (1960). *The registrar-general's statistical review of england and wales for the year.* Wiley Online Library.

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2023). A practical introduction to regression discontinuity designs: Extensions. *arXiv preprint arXiv:2301.08958*.

Cattaneo, M. D., & Titiunik, R. (2022). Regression discontinuity designs. *Annual Review of Economics*, *14*(1), 821–851.

Currie, J., Decker, S., & Lin, W. (2008). Has public health insurance for older children reduced disparities in access to care and health outcomes? *Journal of health Economics*, *27*(6), 1567–1581.

Currie, J., & Rossin-Slater, M. (2015). Early-life origins of life-cycle well-being: Research and policy implications. *Journal of policy Analysis and management*, *34*(1), 208–242.

Currie, J., Zivin, J. G., Mullins, J., & Neidell, M. (2014). What do we know about short- and long-term effects of early-life exposure to pollution? *Annu. Rev. Resour. Econ.*, *6*(1), 217–247.

Dang, L., Dong, L., & Mezuk, B. (2020). Shades of blue and gray: A comparison of the center for epidemiologic studies depression scale and the composite international diagnostic interview for assessment of depression syndrome in later life. *The Gerontologist*, *60*(4), e242–e253.

Dominguez-Bello, M. G., De Jesus-Laboy, K. M., Shen, N., Cox, L. M., Amir, A., Gonzalez, A., ... others (2016). Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer. *Nature medicine*, *22*(3), 250–253.

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American statistical association*, *56*(293), 52–64.

Erickson, J. (2019). *The road to stalingrad: Stalin's war with germany.* Routledge.

Fogel, R. W. (2004). *The escape from hunger and premature death, 1700–2100: Europe, america, and the third world.* Cambridge, UK: Cambridge University Press.

Furuya, S., Zheng, F., Lu, Q., & Fletcher, J. M. (2024). Separating scarring effect and selection of early-life exposures with genetic data. *Demography*, *61*(2), 363–392.

Gardiner, J. (2010). The blitz: the british under attack. *(No Title)*.

Gracner, T., Boone, C., & Gertler, P. J. (2024). Exposure to sugar rationing in the first 1000 days of life protected against chronic disease. *Science*, *386*(6725), 1043–1048.

Jeffrey, G. (2023). *Essays in public provision of healthcare* (Unpublished doctoral dissertation). London School of Economics and Political Science.

Kynaston, D. (2008). *Austerity britain, 1945-1951* (Vol. 1). A&C Black.

London, G., & Region, N. (1952). The registrar general's statistical review of england and wales.[1940–1945 text, vol. i, medical. hmso 6 s. 6 d. 1940–1945 text, vol. ii, civil. hmso 5 s. 0 d. 1946—1947 text, vol. i, medical. hmso 6 s. 6 d.] the war of 1939–45 interrupted the regular publication of the registrar. *JIA*, *78*, 0151–0167.

Marshall, L., Finch, D., Cairncross, L., & Bibby, J. (2019). *Mortality and life expectancy trends in the uk: stalling progress*. Health Foundation London.

Muir, P., Li, S., Lou, S., Wang, D., Spakowicz, D. J., Salichos, L., ... others (2016). The real cost of sequencing: scaling computation to keep pace with data generation. *Genome biology*, *17*, 1–9.

Murphy, M. (1993). The contraceptive pill and women's employment as factors in fertility change in britain 1963–1980: A challenge to the conventional view. *Population studies*, *47*(2), 221–243.

Nobles, J., & Hamoudi, A. (2019). Detecting the effects of early-life exposures: why fecundity matters. *Population research and policy review*, *38*, 783–809.

of Food, M. (1951). *The urban working-class household diet: 1940-1949*. HM Stationery Office.

Penders, J., Thijs, C., Vink, C., Stelma, F. F., Snijders, B., Kummeling, I., ... Stobberingh, E. E. (2006). Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics*, *118*(2), 511–521.

Rivett, G. (1998). *From cradle to grave: fifty years of the nhs*. King's Fund London.

Scholte, R. S., van den Berg, G. J., & Lindeboom, M. (2015, January). Long-run effects of gestation during the Dutch Hunger Winter famine on labor market and hospitalization outcomes. *Journal of health economics*, *39*, 17–30. doi: 10.1016/j.jhealeco.2014.10.002

Solé-Navais, P., Flatley, C., Steinthorsdottir, V., Vaudel, M., Juodakis, J., Chen, J., ... others (2023). Genetic effects on the timing of parturition and links to fetal birth weight. *Nature Genetics*, *55*(4), 559–567.

Stein, A. D., Zybert, P. A., van de Bor, M., & Lumey, L. H. (2004, August). Intrauterine famine exposure and body proportions at birth: The Dutch Hunger Winter. *International Journal of Epidemiology*, *33*(4), 831–6. doi: 10.1093/ije/dyh083

Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. *Journal of the American statistical association*, *54*(285), 30–34.

Swanson, J. M. (2012). The uk biobank and selection bias. *The Lancet*, *380*(9837), 110.

Symonds, N. E., Vidler, M., Wiens, M. O., Omar, S., English, L. L., Ukah, U. V., ... others (2023). Risk factors for postpartum maternal mortality and hospital readmission in low- and middle-income countries: a systematic review. *BMC Pregnancy and Childbirth*, *23*(1), 303.

van Alten, S., Domingue, B. W., Faul, J., Galama, T., & Marees, A. T. (2024). Reweighting uk biobank corrects for pervasive selection bias due to volunteering. *International Journal of Epidemiology*, *53*(3), dyae054.

Vogelezang, S., Bradfield, J. P., Grant, S. F., Felix, J. F., Jaddoe, V. W., Consortium, E. G. G., et al. (2022). Genetics of early-life head circumference and genetic correlations with neurological, psychiatric and cognitive outcomes. *BMC Medical Genomics*, *15*(1), 124.

Warrington, N. M., Beaumont, R. N., Horikoshi, M., Day, F. R., Helgeland, Ø., Laurin, C., ... others (2019). Maternal and fetal genetic effects on birth weight and their relevance to cardio-metabolic risk factors. *Nature genetics*, *51*(5), 804–814.

Webster, C. (2002). *The national health service: A political history.* Oxford University Press, USA.

Wherry, L. R., & Meyer, B. D. (2016). Saving teens: using a policy discontinuity to estimate the effects of medicaid eligibility. *Journal of Human Resources*, *51*(3), 556–588.

Wüst, M. (2022). Universal early-life health policies in the nordic countries. *Journal of Economic Perspectives*, *36*(2), 175–198.

Zhou, J., Indik, C. E., Kuipers, T. B., Li, C., Nivard, M. G., Ryan, C. P., ... others (2024). Genetic analysis of selection bias in a natural experiment: Investigating in-utero famine effects on elevated body mass index in the dutch hunger winter families study. *American Journal of Epidemiology*, kwae376.

Zweiniger-Bargielowska, I. (1994). Rationing, austerity and the conservative party recovery after 1945. *The Historical Journal*, *37*(1), 173–197.

Zweiniger-Bargielowska, I. (2000). *Austerity in britain: rationing, controls, and consumption, 1939-1955*. OUP Oxford.

Appendices

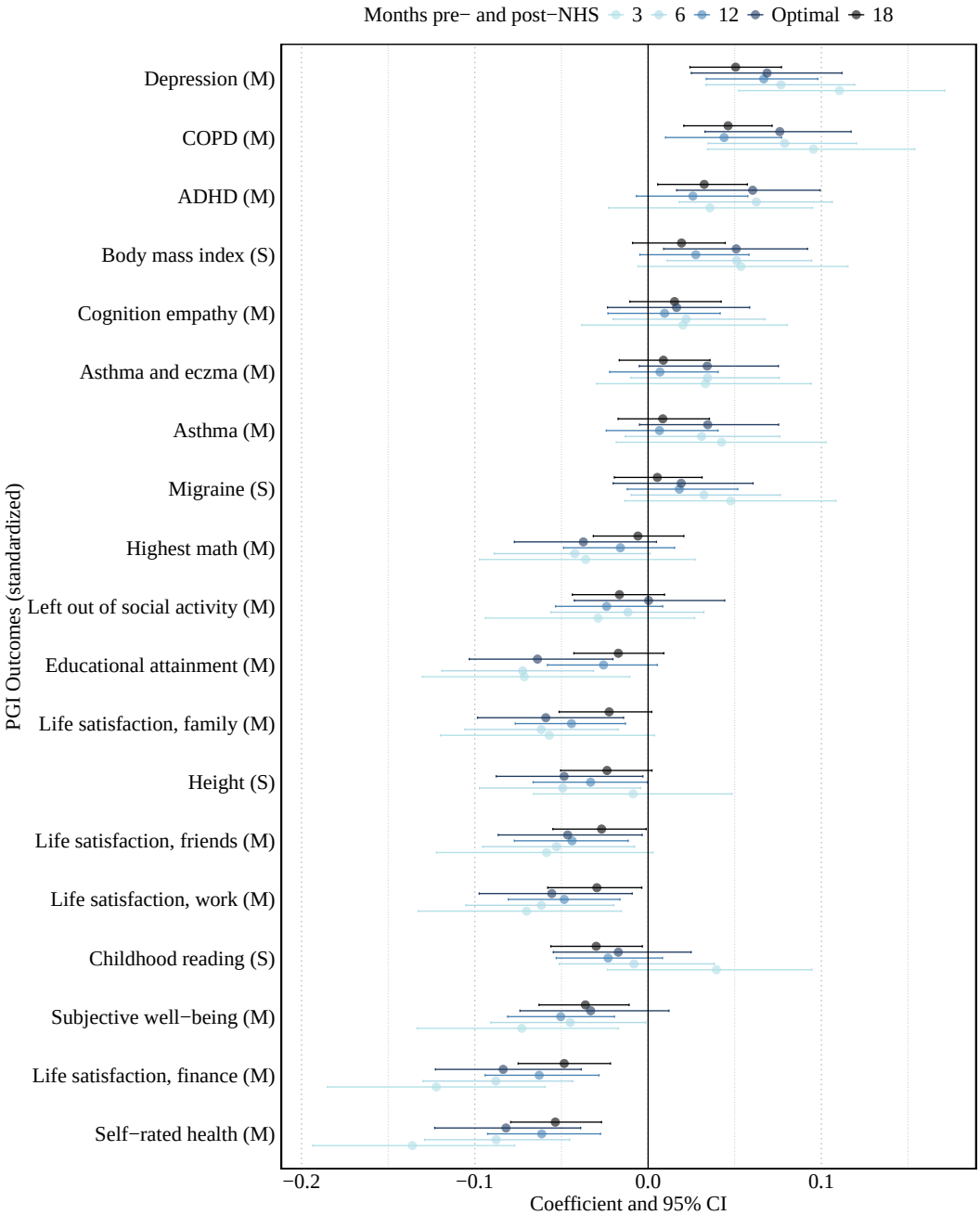
A Extra Results: Effects on Mortality

Table A.1: Effects of the NHS on Infant Mortality using Sharp RDD

	Dependent variable:	
	Infant Mortality Rate	
	BW = 2	BW = 4
	(1)	(2)
NHS Effect	−5.477	−4.145
	(1.094)	(0.661)
Observations	7,352	13,233

Note: Cluster-robust variance estimation with degrees-of-freedom weights, combined with nearest neighbor variance estimation, standard errors in parentheses, *** p<0.001; ** p<0.01; * p<0.05. Linear fit and triangular kernels. *Source*: Registrar General's Statistical Review.

B Extra Results: Survival Bias



Observations: 3 months (m) pre−NHS = 1,897, 3m post−NHS = 2,332; 6m pre−NHS = 3,855, 6m post−NHS = 4,001
12m pre−NHS = 7,743, 12m post−NHS = 7,413; Optimal (7m pre−NHS) = 4,435, Optimal (7m post−NHS) = 4,536
18m pre−NHS = 12,322, 18m post−NHS = 10,426

Figure B.1: Effects of the NHS on PGIs different Bandwidths

Note: Each dot represents the causal effect of a separate regression on a different outcome. P-values under permutation (1,000 permutations) and standard errors under interference. Source: UKB.

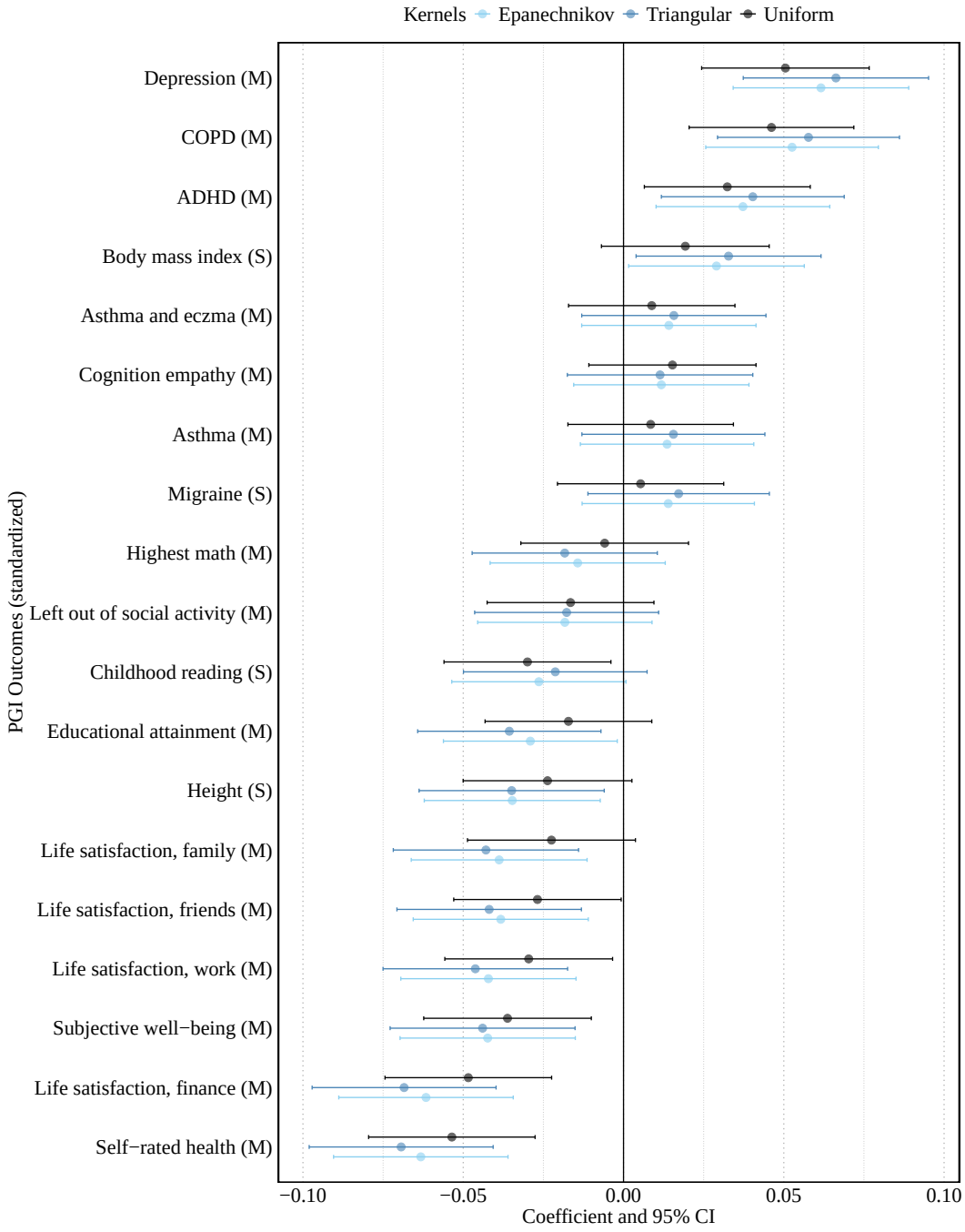


Figure B.2: Effects of the NHS on PGIs different Kernels

Note: Each dot represents the causal effect of a separate regression on a different outcome. P-values under permutation (1,000 permutations) and standard errors under interference. The bandwidth is 18 months pre- and post-NHS implementation. Source: UKB.

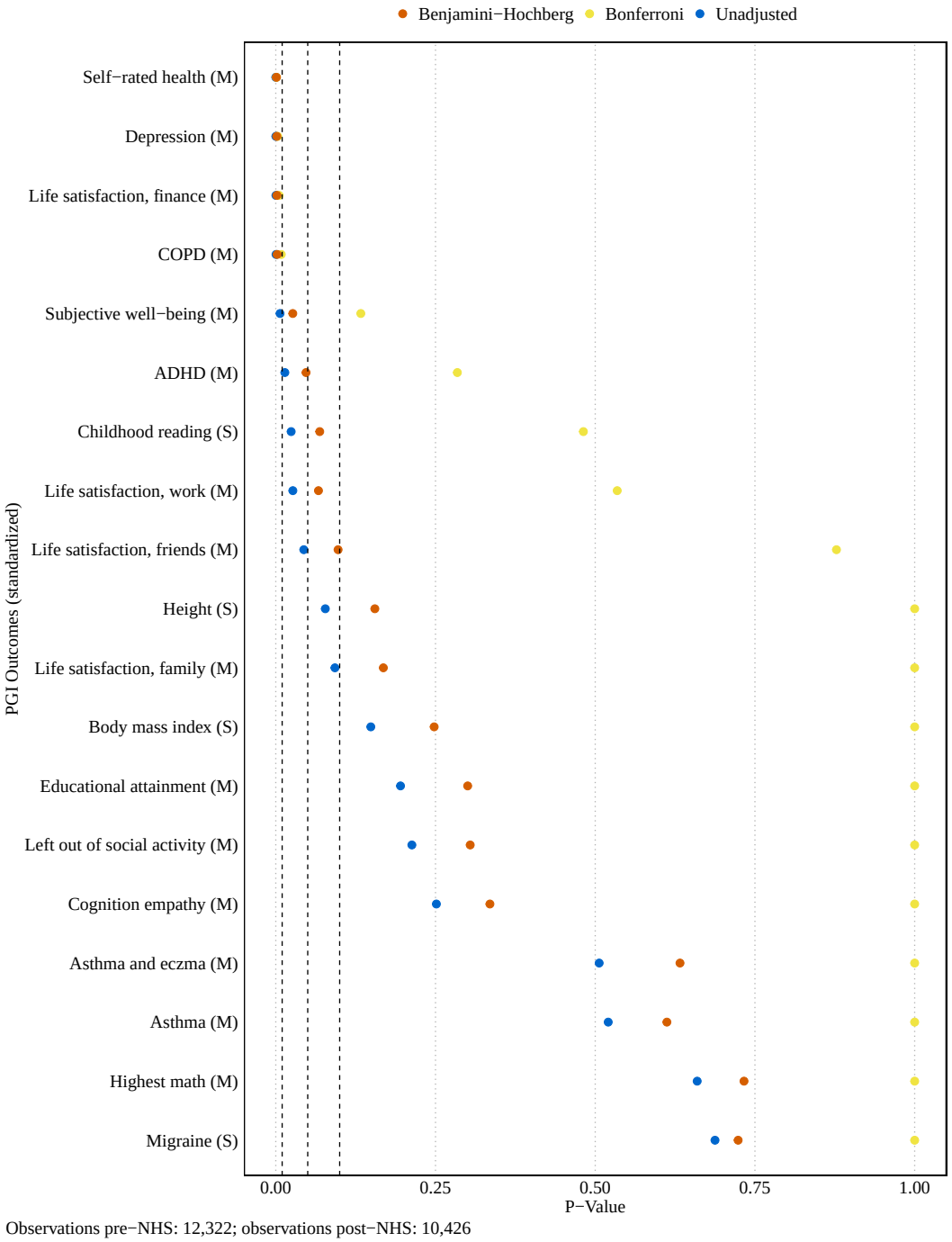


Figure B.3: Multiple Hypothesis Testing Survival Effects of the NHS on PGIs

Note: The bandwidth is 18 months pre- and post-NHS implementation. Source: UKB.

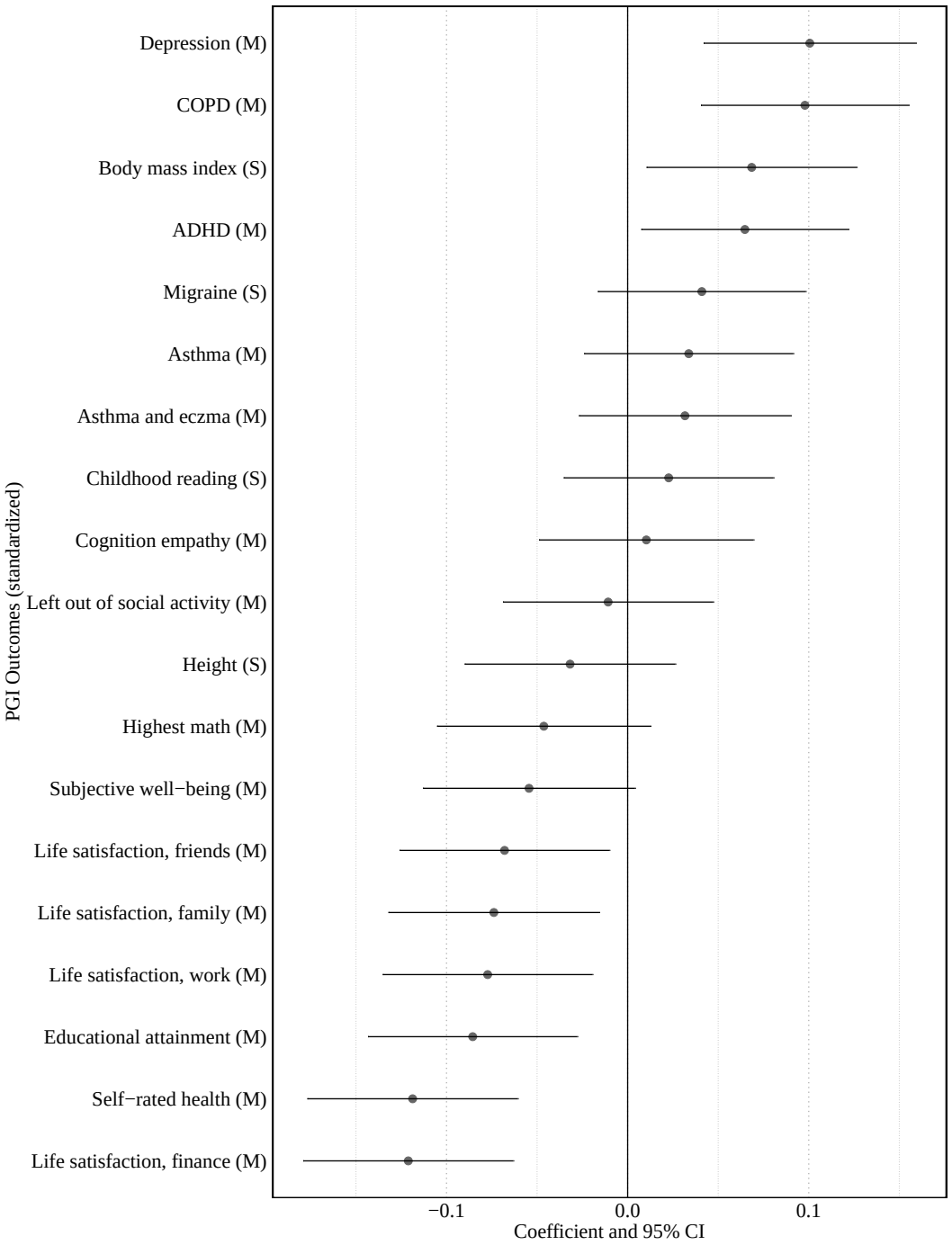


Figure B.4: Effects of the NHS on PGIs using Sharp RDD

Note: Each dot represents the causal effect of a separate regression on a different outcome. 95% confidence intervals. Cluster-robust variance estimation with degrees-of-freedom weights, combined with nearest neighbor variance estimation standard errors. The bandwidth is 18 months per- and post-NHS implementation. Linear fit with triangular kernels. Controls include age, gender, quarter of birth fixed effects, and 10 ancestry specific PCs. Source: UKB.

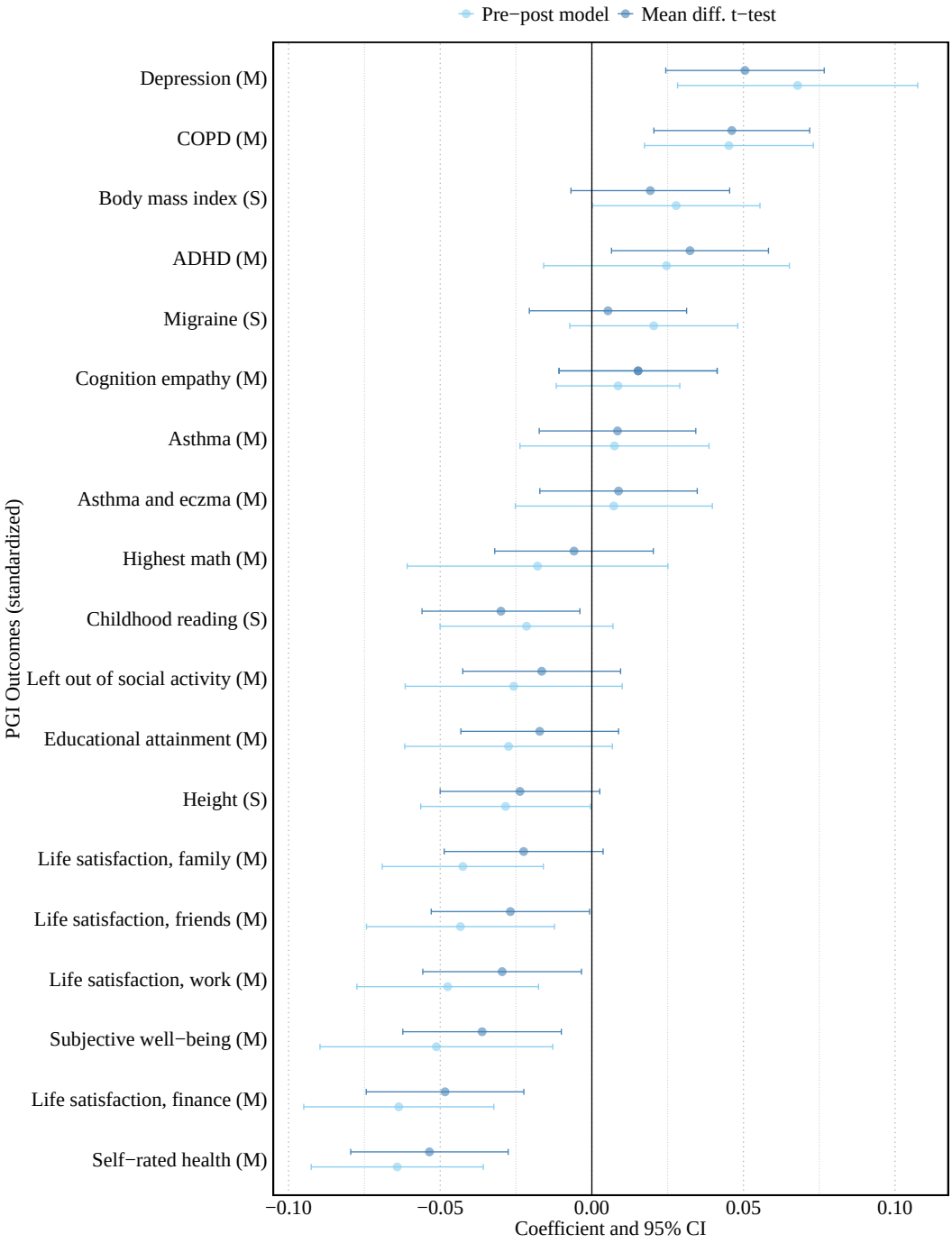


Figure B.5: Effects of the NHS on PGIs using different Bandwidths

Note: Each dot represents the estimated causal effect from a separate regression on a different outcome. The Pre-post regression captures the effect of being born after the implementation of the NHS, controlling for sex, month and year of birth, interview year, and the first 10 genetic principal components (PCs). The Mean difference refers to the difference in average outcomes between pre- and post-NHS births, estimated using a two-sample t-test. Estimates are based on an 18-month bandwidth around the NHS implementation date. Source: UKB

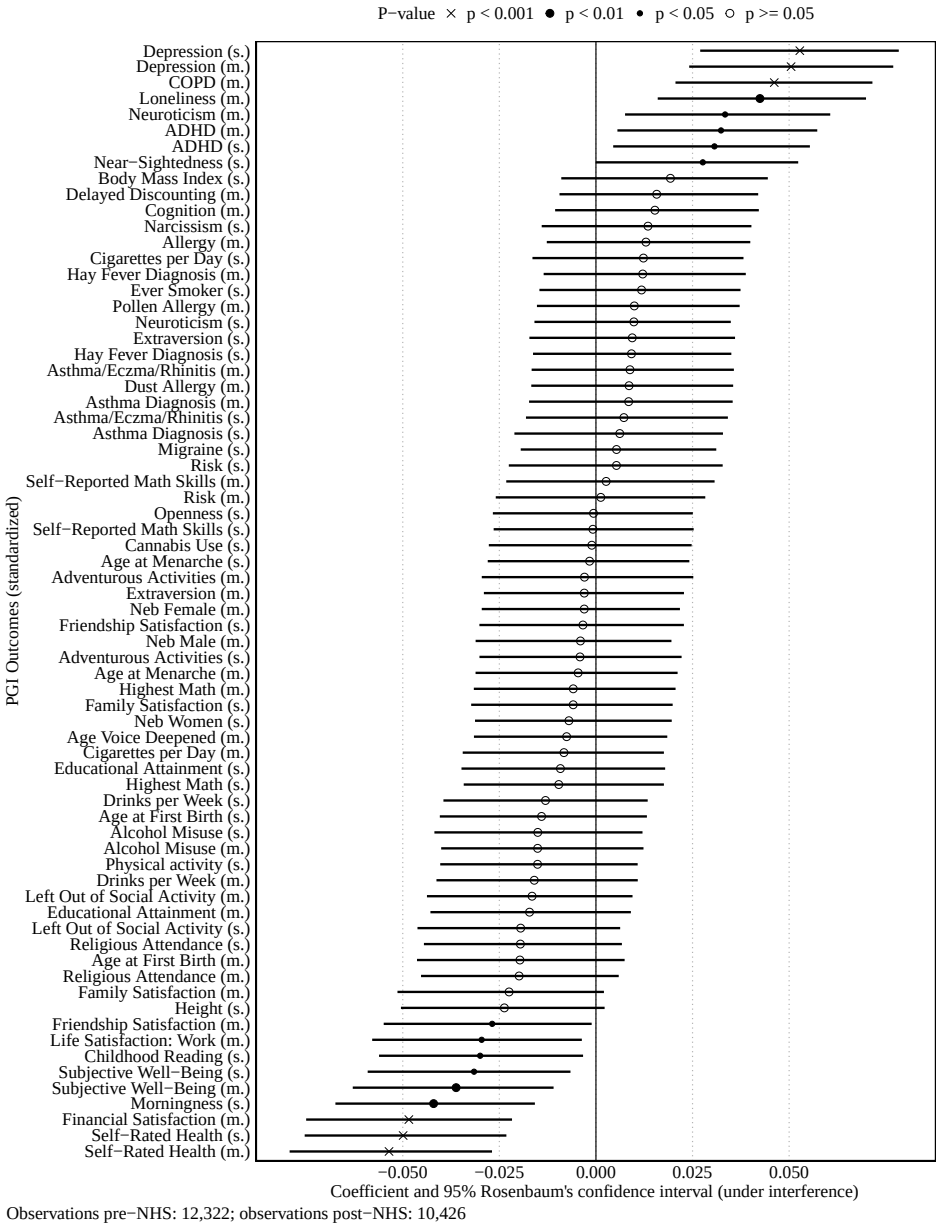


Figure B.6: Effects of the NHS on All PGIs

Note: Each dot represents the causal effect of a separate regression on a different outcome. P-values under permutation (1,000 permutations) and standard errors under interference. The bandwidth is 18 months pre- and post-NHS implementation. Source: UKB.

C Extra Results: Long-term Effects

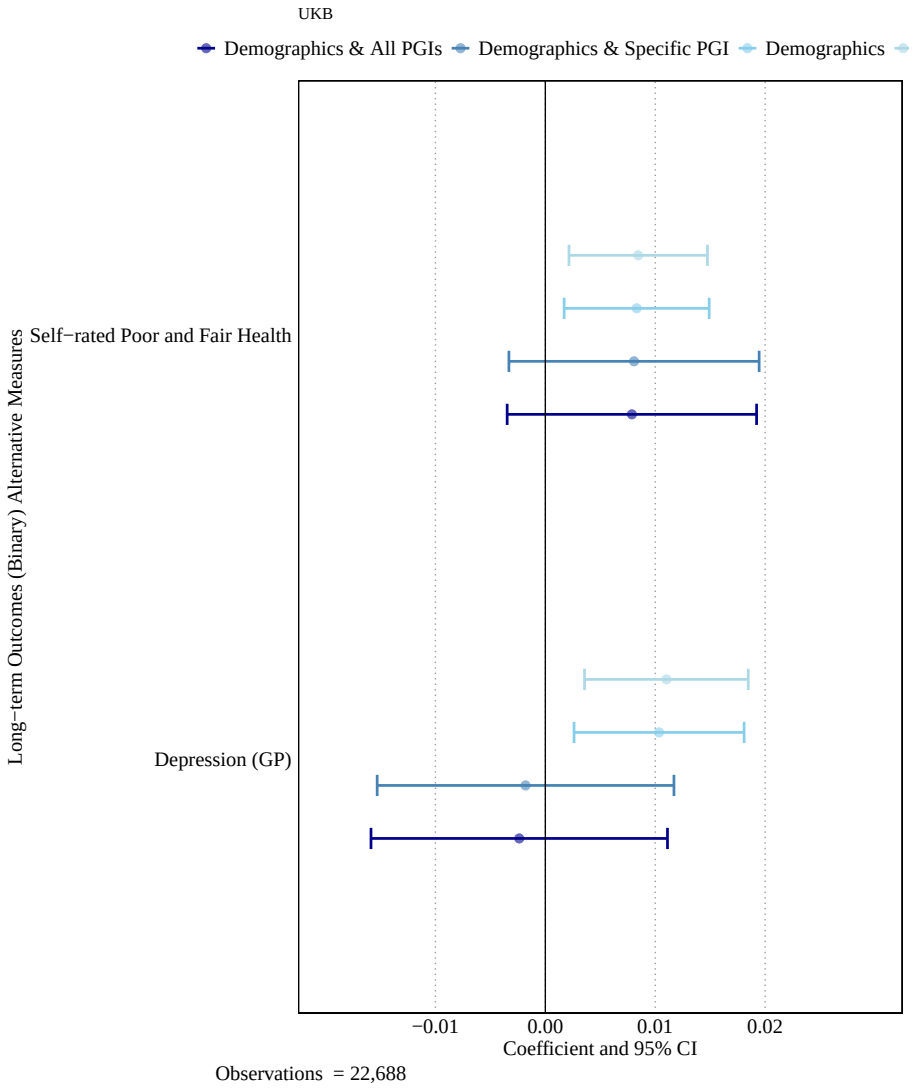


Figure C.1: Effects of the NHS on Long-term Health Outcomes with Alternative Health Measures

Note: Each dot represents the estimated causal effect from a separate regression on a different outcome. The main effect captures the impact of being born after the implementation of the NHS, sequentially controlling for sex, month and year of birth, interview year, quarter of birth FE, PGIs, and the first 10 genetic principal components (PCs). Estimates are based on an 18-month bandwidth around the NHS implementation date. Source: UKB.

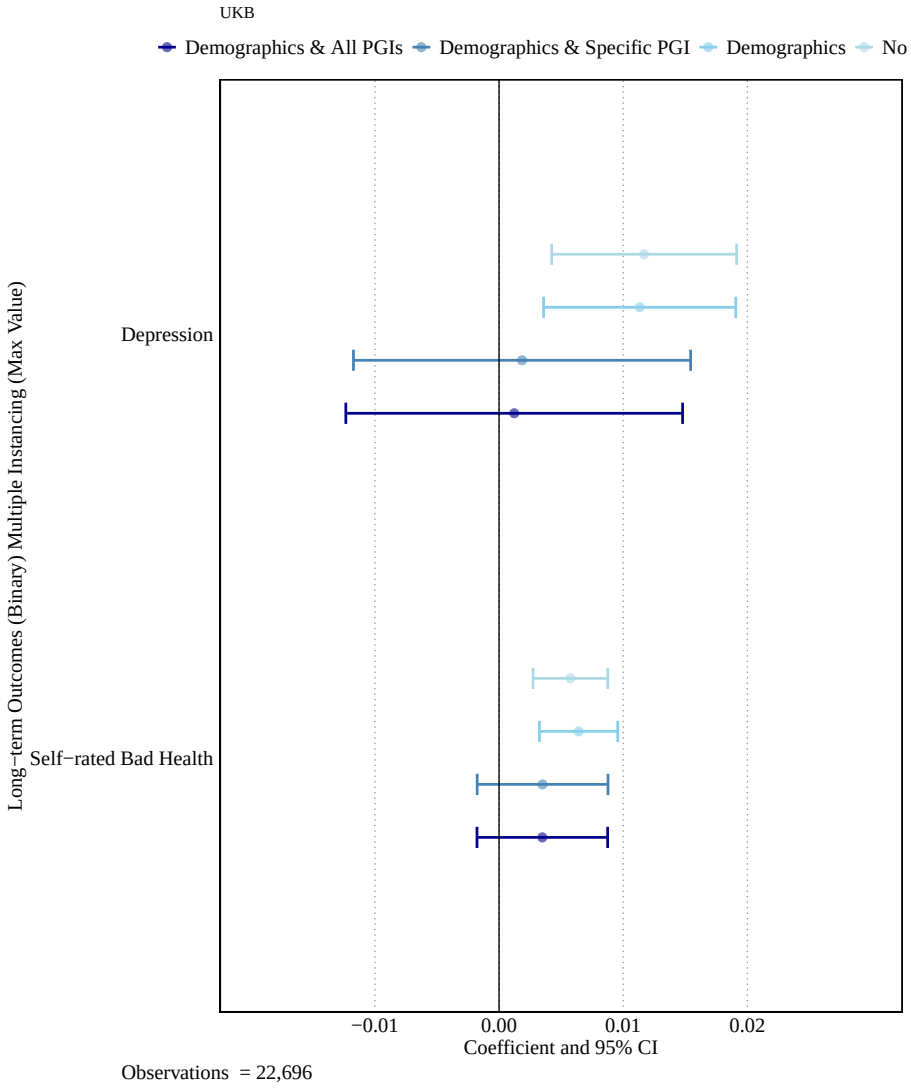


Figure C.2: Effects of the NHS on Long-term Health Outcomes considering all UKB Instances

Note: Each dot represents the estimated causal effect from a separate regression on a different outcome. The main effect captures the impact of being born after the implementation of the NHS, sequentially controlling for sex, month and year of birth, interview year, quarter of birth FE, PGIs, and the first 10 genetic principal components (PCs). Estimates are based on an 18-month bandwidth around the NHS implementation date. Source: UKB.

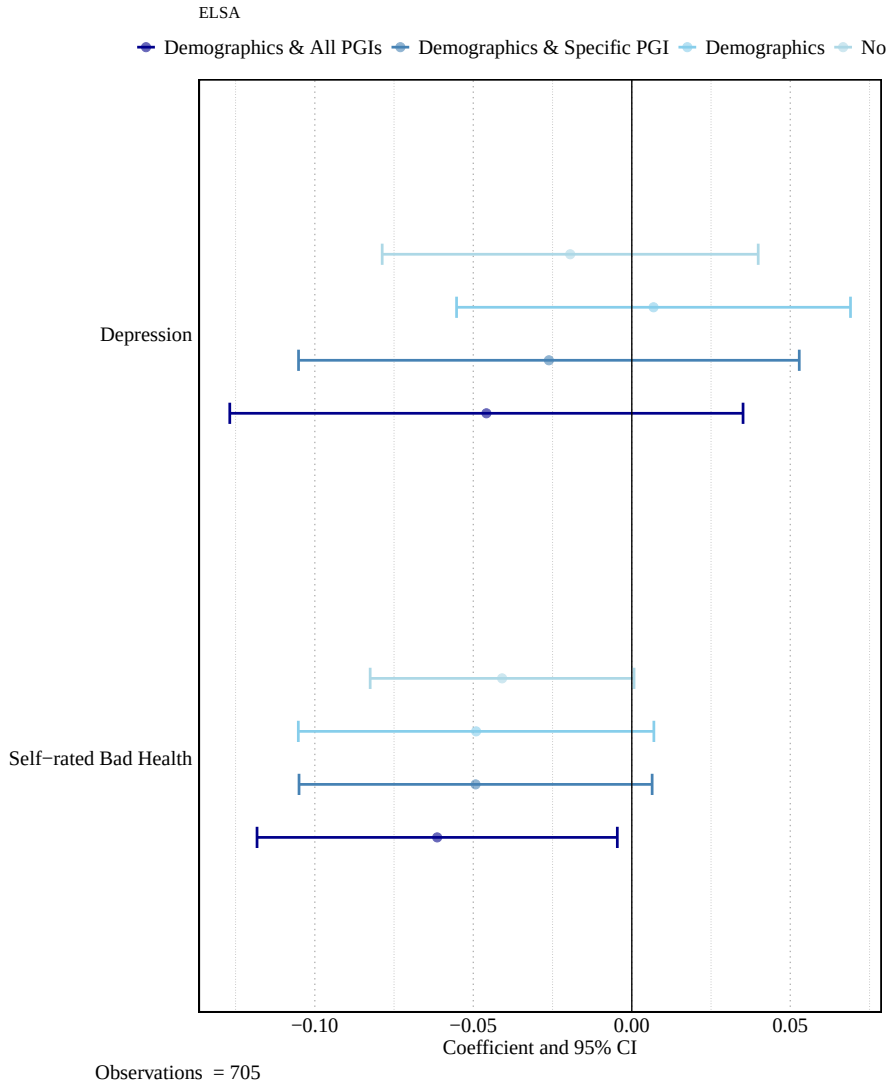


Figure C.3: Effects of the NHS on Long-term Health Outcomes using ELSA

Note: Each dot represents the estimated causal effect from a separate regression on a different outcome. The main effect captures the impact of being born after the implementation of the NHS, sequentially controlling for sex, month and year of birth, interview year, quarter of birth FE, PGIs, and the first 10 genetic principal components (PCs). Estimates are based on an 18-month bandwidth around the NHS implementation date. Source: ELSA.