In partnership with

Econometric modelling undertaken by

NHS Confederation

CREATE Health Foundation

LE London Economics

Women's health economics: investing in the 51 per cent

Appendix: Data sources and expanded methodologies

Contents

A. Full list of data sources

Data set	Definition	Data source	Geographical level	Year
Avoidable mortality	Age-standardised avoidable mortality rate per 100,000 population.	Office for National Statistics	Local authority	2020
CCG expenditure	CCG expenditure data disaggregated by care setting (general and acute, mental health, community, primary, overall).	Freedom of information request	Clinical commissioning groups	2014/15 - 2016/17
CCG needs index	The CCG Needs index is calculated as the ratio of a CCG's weighted population divided by its unweighted population. Weighted populations are the product of the need value and population size for each CCG.	NHS England	Clinical commissioning groups	2014/15 - 2016/17
Cervical screening coverage	Percentage of females aged 25-64 with a recent 'adequate screening' for cervical cancer.	NHS (Cervical Screening Programme)	Local authority	2021-22
Consumer price index	Price indices, percentage changes, and weights for the different measures of consumer price inflation.	Office for National Statistics	National	2014-2023
Contraception use	Number of females using Sexual and Reproductive Health Services for contraception (per 1,000 females aged 16-45).	NHS (Sexual and Reproductive Health Services)	Local authority	2022/23

Disability-free life expectancy (DFLE)	An estimate of the number of years people are expected to spend in different health states.	Office for National Statistics	Local authority	2020
Female employment rate	Female employment rate (for females aged 16 to 64, seasonally adjusted).	Office for National Statistics	Local Authority district	2015/16 - 2018/19
Gross value added (GVA)	Annual gross value added statistics.	Office for National Statistics	Lower Super Output Area (LSOA)	2015-2019
Gynaecological cancer diagnosis waiting times	The percentage of females who received a cancer diagnosis within 28 days of an urgent referral for suspected gynaecological cancer.	NHS (Cancer Waiting Times)	Provider (NHS trust)	2023
Gynaecological cancer treatment waiting times	The percentage of females treated for gynaecological cancers within 31 days of a decision to treat.	NHS (Cancer Waiting Times)	Provider (NHS trust)	2023
Healthy life expectancy (HLE)	An estimate of the proportion of a lifetime spent in 'very good' or 'good' health, based on how individuals perceive their general health.	Office for National Statistics	Local authority	2020
Human papillomavirus (HPV) vaccination coverage	Percentage of Year 9 females who have received two doses of the HPV vaccination.	GOV.UK	Local authority	Academic year (2021/22)
Hormone replacement therapy (HRT) prescriptions	HRT prescriptions per 100,000 females over 40.	NHS (Hormone Replacement Therapy)	ICB (Integrated Care Board)	2023
Index of multiple deprivation	Statistics on relative deprivation in small areas of England.	Office for National Statistics	Lower Super Output Area (LSOA)	2015

Mental health, learning disability and autism services	Number of people in contact with NHS funded secondary mental health, learning disabilities and autism services per 100,000 women.	NHS (Mental Health Bulletin)	Sub-ICB (integrated care board)	2021-22
Midwife staffing	The number of full-time equivalent (FTE) midwives at trust level per 100,00 females aged 15-44.	NHS Workforce statistics	Provider (NHS trust)	2023
Mid-year population estimates	National and subnational mid-year population estimates for the UK and its constituent countries by administrative area, age and sex (including components of population change, median age and population density).	Office for National Statistics	Local authority districts (LAD)	2014, 2015, 2016, 2017
Neonatal mortality	The number of stillbirths and neonatal deaths (before 28 days old) per 1,000 live births and stillbirths.	NHS (Outcomes Framework indicators)	Local authority	2022
NHS staff data	Specialty group for medical staff and area for non-medical staff.[1]	NHS England	NHS trust	Monthly (September 2014 – September 2016)
NHS earnings estimates	Grade for medical staff and staff group for non-medical staff.	NHS England	National	Monthly (September 2014 – September 2016)
NHS providers sites (care trusts, trusts)	Information on NHS trusts, foundation trusts, and healthcare providers licensed by NHSE.	NHS England	Provider	2019
Obstetrics and gynaecology staffing	The number of full-time equivalent (FTE) obstetrics and gynaecology consultants, associate specialists, and specialty doctors per 100,000 females.	NHS Workforce statistics	Provider (NHS trust)	2023

[1] Staff in Obstetrics and gynaecology is defined as the FTE in 'Obstetrics & gynaecology' in the Medical Staff table by CCG of NHS England. Staff in maternity services is defined as the FTE in 'Maternity Services' in the Non-Medical Staff table by CCG of NHS England. All three staff categories – obstetrics, gynaecology, and maternity – are included in this proxy.

Referral to treatment waiting times for gynaecology	The average waiting time (in weeks) for a completed pathway in the gynaecology service.	NHS (Referral to Treatment Waiting Times)	Sub-ICB (Integrated Care Board)	2023
Repeat abortion rate	Repeat abortions as a percentage of total abortions.	Office for Health Improvement and Disparities	Local authority	2021-22
Rural-urban classification	Rural/urban view of datasets at output area (OA), super output area (SOA) and ward level.	Office for National Statistics	Lower super output area (LSOA)	2011
Sickness absence in the UK labour market	The percentage of working hours that are lost because of sickness absence amongst females.	Office for National Statistics	Regional	2022
NHS sickness absences	NHS sickness absence rates.	NHS England	NHS trust	Monthly (October 2015 - October 2019)
Smoking status at time of delivery	Percentage of females who are known smokers at time of delivery.	NHS	Sub-ICB (integrated care board)	2023
Women's health conditions' impact on formal and informal labour market participation and productivity	Population level data in England reflecting women's reproductive health and experiences through the life-course.	Reproductive Health Survey	UK	2023
Women's health conditions' impact on labour market participation	Tracks the lives of approximately 17,000 individuals who were born in Great Britain in a single week in 1970. Alongside information on education, economic activity, income, and family circumstances, the BCS surveys a broad range of health conditions, including several that are specific to women.	British Cohort Study	UK	1970 - present

B. Expanded methodology: a longitudinal assessment of select conditions on labour market participation across the UK

I. Econometric approach

The key challenge in identifying the impact of select health conditions on women involved the likely differences between those that experienced the condition (the 'treatment group') and those that did not (the 'control' or 'untreated group') in ways other than simply whether they experienced the condition. These differences had the potential to impact the outcomes considered, and so a simple comparison between the economic activity, mental health and physical health of the control and treatment groups before and after the start of the condition could have resulted in incorrectly estimating the effects of the conditions.[2] To tackle this problem, **propensity score matching** (PSM)[3] was implemented to identify a set of control respondents that closely resembled the treated respondents and minimised these differences. Then, a comparison was made between this set of control respondents and the treated respondents using the **difference-in-differences method**. These models were estimated using STATA statistical software.

II. Propensity score matching

PSM allowed us to match those who experience the relevant condition with those who do not, based on historical characteristics.[4] The approach involves matching the treatment respondents to similar respondents in the control group based on a 'propensity score'.[5] The propensity score was calculated using a logit model: a standard economic model used when the outcome of interest is binary (whether the respondent experiences the given condition in the next wave). This model is represented by the following equation for respondent *i*:

$$\Pr(T_i = 1) = F(\beta C_i)$$

[2] For example, one might expect that those who experienced poor health before the onset of menopause would be more likely to be impacted by the menopause than those in good health.
[3] For further information on the method more generally, an introduction to PSM can be found in Caliendo, M., & Kopeinig, S.
[4] For example, within the analysis of the menopause, we matched respondents that do experience menopause symptoms when they are 46 to those that do not when they are 46, based on control variables recorded in the previous wave of the assessment. Control variables used for the PSM approach include education level, partner's economic status, education level, number of children, income band, occupation, and health status.
[5] The propensity score is the probability of a respondent being in the treatment group, estimated based on their characteristics in the previous wave.

- T_i is a dummy variable that took the value 1 if the respondent was in the treatment group and 0 otherwise.
- F is the cumulative standard logistic distribution, which is defined such that $F(x) = \frac{1}{1+e^{-x}}$.
- C_i is a vector of characteristics of the respondent.
- β represents the associated coefficients for these characteristics.[6]

Once the propensity scores were calculated, each treated respondent was matched to the respondent with the closest propensity score (ie the 'nearest neighbour'). Respondents were only considered for either the treatment or the control group if they did not experience the condition in the first wave of analysis. This ensured that the difference-in-differences models only included those respondents who did not experience the condition to begin with.

First, the model was run without allowing for replacement, meaning that each control respondent could only be matched to one respondent in the treatment group. Then, the algorithm was run again to allow for replacement. To ensure that each treatment respondent was matched to only one control respondent after the completion of the 'with replacement' model,[7] the one control respondent matched with the given treatment respondent was chosen at random, with the weighting for the regressions based on this random matching. To ensure that the random matching did not alter the results, this process was completed ten times.

As a result, eleven different PSM weights were calculated separately for the menopause and dysmenorrhea models. The results presented in the main report relate to the difference-in-differences regressions run using the PSM weights without replacement.

With regard to the menopause, the two waves of the BCS survey included in the analysis were wave 9 and wave 10, as these waves occurred when women would theoretically be approaching or within average menopausal age.[8] The PSM sample consists of those who responded to the survey in both wave 9 and wave 10 and who did not show any sign of menopause in wave 9. Treatment and control observations were then matched based on their characteristics in wave 9. Questions pertaining to dysmenorrhea are included in numerous waves. The difference-in-differences analysis pooled data from waves 6, 9 and 10

[6] Details on the characteristics included in the Propensity Score Matching can be found later in this Appendix.
[7] A 'with replacement' model means that each respondent in the control group can be matched to more than one respondent in the treatment group. This was employed to ensure that the weights generated are integers, so that frequency weights could be used in the difference-in-differences regressions.
8 This analysis operates under the notion that those experiencing perimenopause and menopause symptoms are typically aged 46 – 55 and therefore does not capture the approximate 5 per cent of the population who enter the menopause prematurely
due to certain cancer treatments, surgery, or certain health conditions.

of the British Cohort Study (BCS), as these are the only waves in which detailed questions related to secondary dysmenorrhea are asked. The PSM was therefore conducted three times for dysmenorrhea, each time using data from the previous wave (ie waves 5, 8 and 9) to match treatment and control observations.

III. Difference-in-differences model

The difference-in-differences approach compared outcomes for those that do and do not experience the condition, allowing us to account for any pre-existing differences in the outcome variables[9] between the treatment and control groups that may not have been eliminated in the propensity score matching process.[10] The core specification for the difference-in-differences model is described by the following equation:

$$\Pr(y_{it} = 1) = F(\gamma T_{it} + \lambda I_t + \mu E_{it} + \beta H_i + \delta W_t)$$

- y_{it} denotes the outcome variable. For economic activity, this took the value of 1 if the respondent was in employment and 0 otherwise. For mental health, this took the value of 1 if the respondent reported that they regularly feel miserable or depressed, and 0 otherwise. For physical health, this took the value of 1 if the respondent reported that their health was poor or fair, and 0 otherwise.
- T_{it} are a set of dummy variables indicating whether the respondent was in the treatment group, and if so, the type of condition they had. For the menopause, these are separate dummy variables for whether the respondent had 1-2 symptoms or 3-4 symptoms. For dysmenorrhea, these dummy variables indicate whether the respondent had primary or secondary dysmenorrhea. As all respondents did not have the condition in the previous wave, these dummy variables equal 0 for all respondents in the first wave of analysis.
- $\boldsymbol{I_t}$ indicates whether the observation was pre- or post-treatment.
- $\boldsymbol{E_{it}}$ are a set of dummy variables controlling for the education level of the respondent. Further details on these variables can be found in Table 1.
- $\boldsymbol{H_i}$ is a dummy variable indicating the respondent's health status in the first wave of analysis. For analysis relating to waves 5-6 and 9-10, information on whether the

[9] For example, there may have been differences in labour market outcomes between those that experience menopause symptoms when they are 46 and those that do not, before their menopause symptoms begin.
[10] Ideally, the treated and control samples would be similar to the extent that their outcomes before the condition begins for the treatment group would be identical. However, given that it is not possible to observe all relevant characteristics and incorporate these into the PSM process and because the analysis is based on a (limited) sample, it is possible that differences in outcomes remain between both groups even after the matching.

individual experienced a long-term health condition in wave 5 or wave 9 is used. Due to data availability, the analysis relating to waves 8-9 uses a question which asks whether the respondent's health limited their day-to-day activities in wave 8. Therefore, for the pooled dysmenorrhea regression, this variable included a mix of both questions depending on the wave of analysis. This variable was time-invariant and did not consider the value post-treatment, as this is highly correlated with whether the respondent experienced dysmenorrhea or the menopause post-treatment. For a similar reason, this variable was not included in the regressions where physical health was the outcome variable.

- W_t is a set of dummy variables indicating whether the observation was in the given wave. This is only applicable for the dysmenorrhea model, as the regressions were pooled using the weights from three separate PSM runs (i.e. for waves 5-6, 8-9 and 9-10).
- For the menopause model, $\boldsymbol{I_t}$ indicates whether an observation was pre- or post-treatment, therefore showing the wave for which the observation referred to.
- γ, λ, μ, β and δ are individual scalars/vectors of coefficients. λ is the estimate for the effect of experiencing the condition.

The difference-in-differences approach was repeated twelve times for each condition and outcome combination. The results in the main body refer to the regressions that used weights from the PSM ran without replacement. As a sensitivity analysis, the models were rerun using the PSM weights from the ten with replacement models. A further sensitivity analysis ran the difference-in-differences unweighted, but including the control variables that were included in the PSM as further controls within the difference-in-differences.

IV. Data

It is important to note that the analysis used to estimate the impact of select health conditions on absenteeism and presenteeism does not use an econometric approach due to a lack of longitudinal data. While a causal relationship between health conditions and outcomes was possible to estimate with the BCS data (as it is longitudinal), this was not the case for analysis using the Reproductive Health Service (RHS). The BCS does not collect data on workplace productivity; however, the RHS collected data on women who would have liked to take days off work due to specific health reasons but did not. To estimate the impact of presenteeism based on the conditions assessed, we made two assumptions. First, if those who wanted to take days off work had done so, they would have taken an amount equal to the average of those who did take days off due to their health condition. Second, on

the days that they worked but would have preferred to have taken off, they are 20 per cent less productive as a result of their symptoms, as based on 2023 employee wellbeing survey. We also assume that these data are representative of the population experiencing the condition.

Table 1: Variables used in the PSM

Variable	Explanation
Experiences condition in next wave	The outcome variable for the PSM. This variable was set to missing if the respondent experienced the condition in the first wave, to ensure that all those with a positive weight did not have the condition originally. This variable was not split by severity of menopause or by type of dysmenorrhea.
Economic activity	Economic activity, split into the following broad categories: full-time employment, part-time employment, and not in the labour market. This variable was used as an exact match, meaning that treatment observations were only matched to control observations which shared the same value as this variable.
Education level	The highest level of qualification obtained by the date of the survey. The question from the survey was aggregated for the PSM analysis into the following groups: No qualifications or pre-GCSE qualifications, GCSE equivalent, AS/A Level equivalent, degree or equivalent, higher degree or equivalent.
Partner's economic status	Split into the following categories: does not live with a partner, lives with a partner who does not work, lives with a partner who does work.
Number of children by age-band	Three dummy variables indicating whether the respondent lived with a child within the age band. The age bands considered are 0-4 years old, 5-11 years old and 12-15 years old.
Income band	Weekly take home pay, split into the following bands: £0 (not in employment), £1-£199, £200-£399, £400-£599, £600 or more.
Occupation	Occupation of the respondent based on SOC codes. These were aggregated for the PSM into the following categories: 'Not applicable/not in employment', 'Managers and professional occupations', 'Other desk-based occupations' and 'Manual occupations'.

Region	Region of residence at survey date at the government office region level.
Health	Subjective health status. For PSM models that relied on data from waves 5 and 9 (ie where the main analysis was conducted using waves 6 and 10), a variable asking whether the respondent experienced a long-term health condition was used. For the dysmenorrhea PSM model relying on data from wave 8, where the main analysis was conducted using wave 9, a variable asking the respondent whether their health limited their day-to-day activities was used. This difference is due to data availability.
Caring for children	A dummy variable that equals 1 if the respondent reported that they do most of the caring for children in their household, and 0 otherwise. Due to data availability, this was only included in the models relying on data from wave 9, where the main analysis was conducted for wave 10.

V. Econometric results

The tables below show the results from the econometric analysis as coefficient estimates; these are presented as marginal effects in the report. The results outlined in the report relate to the columns entitled: PSM without replacement.

For the tables related to dysmenorrhea, the rows of interest relate to ‘primary dysmenorrhea’ and ‘secondary dysmenorrhea’. For the tables related to the menopause, the key rows of interest are ‘1-2 menopause symptoms’ and ‘3-4 menopause symptoms’. In each table, the unweighted regression includes all variables that were included in the PSM (see Table 1), but the results relating to these are not included in the tables for brevity. The ‘sweep’ coefficients refer to the effect compared to the response being from sweep 5. The ‘highest qualification’ coefficients refer to the effect compared to the respondent having no formal qualifications.

Table 2: Estimated impact of dysmenorrhea on employment

Variable	PSM without replacement	PSM with replacement	Unweighted
Primary dysmenorrhea	-0.110	-0.114 - -0.036	-0.207***

	(0.083)		(0.073)
Secondary dysmenorrhea	0.024 (0.166)	-0.011 - 0.046	-0.050 (0.166)
Sweep 6	0.137 (0.145)	-0.082 - 0.152	0.078 (0.064)
Sweep 8	-0.276** (0.130)	-0.328 - -0.311	0.072 (0.076)
Sweep 9	0.141 (0.136)	0.065 - 0.133	0.190** (0.083)
Sweep 10	0.648*** (0.158)	0.579 - 0.662	0.259*** (0.096)
Highest qualification – GCSE equivalent	0.358*** (0.096)	0.276 – 0.361	0.446*** (0.076)
Highest qualification – AS or A Level equivalent	0.747*** (0.169)	0.759 - 0.838	0.724*** (0.118)
Highest qualification – degree or equivalent	0.738*** (0.106)	0.710 – 0.769	0.892*** (0.084)
Highest qualification – higher degree or equivalent	1.309*** (0.250)	1.591 – 1.755	1.447*** (0.194)
Has a pre-existing health condition	-0.970*** (0.088)	-0.926 - -0.875	-0.973*** (0.069)
Observations	8,562	6,753 – 6.759	18,184

Source: London Economics analysis of BCS data

Note: *** statistically significant at the 1% level, ** statistically significant at the 5% level, * statistically significant at the 10% level

Table 3: Estimated impact of dysmenorrhea on mental health

Variable	PSM without replacement	PSM with replacement	Unweighted
Primary dysmenorrhea	0.382*** (0.129)	0.274 - 0.443	0.306*** (0.071)
Secondary dysmenorrhea	0.140 (0.241)	-0.207 - -0.046	0.422*** (0.145)
Sweep 6	-0.378** (0.157)	-0.368 - -0.215	-0.276*** (0.065)
Sweep 9	-0.331** (0.141)	-0.413 - -0.264	0.038 (0.085)
Sweep 10	-0.448*** (0.155)	-0.412 - -0.278	-0.045 (0.095)
Highest qualification – GCSE equivalent	-0.227* (0.134)	-0.223 - -0.12	-0.167** (0.081)
Highest qualification – AS or A Level equivalent	-0.556*** (0.212)	-0.469 - -0.376	-0.317** (0.130)
Highest qualification – degree or equivalent	-0.576*** (0.148)	-0.585 - -0.491	-0.367*** (0.095)
Highest qualification – higher degree or equivalent	-0.262 (0.261)	-0.399 - -0.3	-0.246 (0.185)
Has a pre-existing health condition	0.787*** (0.110)	0.808 - 0.884	0.710*** (0.065)
Observations	3,179	2,678 – 2,683	12,418

Source: London Economics analysis of BCS data

Note: *** statistically significant at the 1% level, ** statistically significant at the 5% level, * statistically significant at the 10% level

Table 4: Estimated impact of dysmenorrhea on physical health

Variable	PSM without replacement	PSM with replacement	Unweighted
Primary dysmenorrhea	0.218** (0.089)	0.184 - 0.221	0.234*** (0.080)
Secondary dysmenorrhea	0.503*** (0.160)	0.419 - 0.455	0.607*** (0.152)
Sweep 6	0.514*** (0.163)	0.306 - 0.512	0.454*** (0.086)
Sweep 8	0.103 (0.177)	-0.159 - -0.002	0.459*** (0.102)
Sweep 9	0.637*** (0.194)	0.079 - 0.212	0.683*** (0.105)
Sweep 10	-0.479*** (0.108)	0.452 - 0.636	1.028*** (0.110)
Highest qualification – GCSE equivalent	-0.800*** (0.186)	-0.625 - -0.561	-0.364*** (0.086)
Highest qualification – AS or A Level equivalent	-0.948*** (0.119)	-1.003 - -0.866	-0.502*** (0.150)
Highest qualification – degree or equivalent	-0.903*** (0.268)	-0.944 - -0.866	-0.555*** (0.103)
Highest qualification – higher degree or equivalent	0.218** (0.089)	0.184 - 0.221	-0.475** (0.227)
Observations	8,592	6,780 – 6,783	16,501

Source: London Economics analysis of BCS data

Note: *** statistically significant at the 1% level, ** statistically significant at the 5% level, * statistically significant at the 10% level

Table 5: Estimated impact of the menopause on employment

Variable	PSM without replacement	PSM with replacement	Unweighted
1-2 menopause symptoms	0.013 (0.156)	0.045 - 0.122	-0.043 (0.121)
3-4 menopause symptoms	-0.352** (0.164)	-0.336 - -0.259	-0.303** (0.124)
Post-treatment dummy	0.364*** (0.090)	0.387 - 0.464	0.075 (0.095)
Highest qualification – GCSE equivalent	0.200 (0.137)	0.099 - 0.128	0.376*** (0.100)
Highest qualification – AS or A Level equivalent	0.695*** (0.250)	0.632 - 0.64	0.869*** (0.192)
Highest qualification – degree or equivalent	0.575*** (0.142)	0.653 - 0.697	0.877*** (0.107)
Highest qualification – higher degree or equivalent	0.813*** (0.290)	1.042 - 1.053	1.427*** (0.247)
Has a pre-existing health condition	-0.957*** (0.113)	-1.109 - -1.094	-0.923*** (0.082)
Observations	4,170	3,653 – 3,655	7,749

Source: London Economics analysis of BCS data

Note: *** statistically significant at the 1% level, ** statistically significant at the 5% level, * statistically significant at the 10% level

Table 6: Estimated impact of the menopause on mental health

Variable	PSM without replacement	PSM with replacement	Unweighted
1-2 menopause symptoms	0.561*** (0.142)	0.597 - 0.655	0.623*** (0.120)
3-4 menopause symptoms	0.975*** (0.151)	0.831 - 0.888	0.977*** (0.125)
Post-treatment dummy	-0.399*** (0.100)	-0.459 - -0.404	-0.550*** (0.104)
Highest qualification – GCSE equivalent	-0.241* (0.128)	-0.35 - -0.315	-0.089 (0.096)
Highest qualification – AS or A Level equivalent	-0.328 (0.222)	-0.247 - -0.156	-0.292 (0.182)
Highest qualification – degree or equivalent	-0.520*** (0.131)	-0.703 - -0.664	-0.270** (0.108)
Highest qualification – higher degree or equivalent	-0.605** (0.250)	-0.429 - -0.367	-0.124 (0.193)
Has a pre-existing health condition	0.838*** (0.108)	0.978 - 1.021	0.818*** (0.077)
Observations	3,827	3,373 – 3,386	6,153

Source: London Economics analysis of BCS data

Note: *** statistically significant at the 1% level, ** statistically significant at the 5% level, * statistically significant at the 10% level

Table 7: Estimated impact of the menopause on physical health

Variable	PSM without replacement	PSM with replacement	Unweighted
1-2 menopause symptoms	0.226 (0.153)	0.139 - 0.173	0.471*** (0.125)
3-4 menopause symptoms	1.080*** (0.151)	1.022 - 1.057	1.081*** (0.127)
Post-treatment dummy	0.092 (0.102)	0.004 - 0.032	-0.248** (0.106)
Highest qualification – GCSE equivalent	-0.431*** (0.138)	-0.5 - -0.484	-0.257*** (0.099)
Highest qualification – AS or A Level equivalent	-0.963*** (0.275)	-0.763 - -0.699	-0.422** (0.209)
Highest qualification – degree or equivalent	-1.039*** (0.146)	-1.202 - -1.142	-0.507*** (0.113)
Highest qualification – higher degree or equivalent	-1.286*** (0.346)	-0.886 - -0.863	-0.339 (0.232)
Observations	4,186	3,665 – 3,667	6,741

Source: London Economics analysis of BCS data

Note: *** statistically significant at the 1% level, ** statistically significant at the 5% level, * statistically significant at the 10% level

VI. UK population estimates

In the report, the percentage point estimates surrounding the impact of dysmenorrhea and the menopause on employment, mental health and physical health are converted into estimates in terms of the number of females across the UK. This section outlines the assumptions made and sources used to estimate the number of people in the UK with each

condition, and the conversion between the number of jobs lost and the direct economic impact associated with the menopause.

Dysmenorrhea

The analysis assumes that dysmenorrhea[11] can impact any woman between the ages of 16 and 45. Using the latest population estimates from the Office for National Statistics,[12] across the UK, there were approximately 13.1 million females in this age bracket in 2022. Data from the RHS was used to estimate the proportion of women that are likely to experience dysmenorrhea. While this survey is not representative, it is preferred to the BCS as it captures a wider range of ages, so is likely to better account for how these conditions change as people age. The table below shows the proportion of women who experience dysmenorrhea and the conversion into a UK-wide population estimate.

Table 8: Percentage of women in the RHS and number of females in the UK estimated to experience dysmenorrhea

Category	Percentage	UK population
Dysmenorrhea	55.4 per cent	7,250,000
Primary dysmenorrhea	39.7 per cent	5,190,000
Secondary dysmenorrhea	15.7 per cent	2,060,000

The population estimates shown above were then multiplied by the percentage point changes associated with experiencing the conditions to estimate the impact of the conditions in terms of the UK population.

Menopause

Based on wider literature, we assumed that those experiencing the menopause are typically between the ages 45-55.[13] As stated in the report, this age range does not account for the approximate five per cent of those assigned female at birth who enter the menopause prematurely (before 45 years of age) due to certain cancer treatments, surgery, or certain

[11] For the purposes of this analysis, dysmenorrhea is defined is experiencing heavy periods or severe pain during periods. Primary dysmenorrhea refers to those experiencing heavy periods or severe pain during periods but not reporting a diagnosed reproductive health condition, whilst secondary dysmenorrhea refers to experiencing heavy periods or severe pain during periods and also having a diagnosed reproductive health condition.
[12] ONS estimates of the population for the UK, England, Wales, Scotland, and Northern Ireland.
[13] For example, see here.

health conditions. Again, using population estimates from the ONS, there were approximately 4.5 million women in this age bracket in 2022.

Then, data from the BCS was used to determine the proportion of these individuals that experience the menopause, and within this group, those that experience less severe (one-to-two symptoms) or more severe (three-to-four symptoms) menopause. Capturing menopause severity at age 46 might be a cautious estimate as many females at this age may not have yet started the menopause. However, due to a lack of better data, and to remain conservative with our estimates, the BCS was used. These findings are shown in Table 9 alongside the conversion into an estimate of the UK population as a whole.

Table 9: Percentage of females in the BCS and number of females in the UK estimated to experience the menopause

Category	Percentage	UK population
0 menopause symptoms	31.0 per cent	1,390,000
1-2 menopause symptoms	40.6 per cent	1,830,000
3-4 menopause symptoms	28.5 per cent	1,280,000

The main analysis found that approximately 60,000 females were not in employment as a result of experiencing more severe menopause. To convert this to a direct economic impact estimate, median pay for females of menopausal age was calculated using ONS earnings data.[14] Earnings data by age is only published in wide age bands, so an average of the median take home pay within the 40-49 and 50-59 age brackets was used. This is a weighted average, adjusted by the number of women in employment across the two groups, resulting in a median annual gross pay for females of menopausal age of £23,900.

Waves of the British Cohort Study

The table below shows each wave of the British Cohort Study alongside the year(s) that fieldwork was undertaken for each wave and the ages covered. Each wave is a specific survey undertaken of all those studied within the British Cohort Study (excluding those that could not be contacted or refused to answer the survey). Survey waves 6, 9 and 10 were conducted at ages 29-30, 42-43 and 46-48.

[14] Office for National Statistics. 2023. Earnings and hours worked, place of residence by local authority: ASHE Table 6.

Table 10: Waves of the British Cohort Study (Source: Centre for Longitudinal Studies)

Wave	Years	Ages
5	1996	26
6	1999-2000	29-30
8	2008-2009	38-39
9	2012-2013	42-43
10	2016-2018	46-48

C. Expanded methodology: the case for investment in obstetrics and gynaecology services in England

I. Statistical approach

To shed light on the association between the NHS investment in obstetrics and gynaecology services and the three selected performance indicators – gross value added (GVA), female sickness absences, and NHS workforce absences – a statistical approach comparing the performance of high and low investors was adopted. The designed approach consists of five steps, outlined in detail below.

Step 1: Measurement of the change in the yearly NHS investment in obstetrics and gynaecology services at the CCG level for the two periods 2014/15 to 2015/16 and 2015/16 to 2016/17.

In the first step, the yearly change in NHS expenditure in obstetrics and gynaecology relative to need was computed for all CCGs, both in percentage and absolute terms, for the periods 2014/15 to 2015/16 and 2015/16 to 2016/17. Outliers were then identified and excluded. Specifically:

- The three-standard deviation rule was applied to detect outliers in the cross-sectional distribution of NHS expenditure in obstetrics and gynaecology relative to need for each financial year.
- CCGs with a year-on-year change exceeding 100 per cent in NHS expenditure in obstetrics and gynaecology relative to need were also excluded.

Step 2: Measurement of the change in the indicator of economic performance at CCG level in the following three tax years.

In the second step, the three-year change in each economic performance indicator (in both percentage and absolute terms) was calculated for all CCGs for the periods 2014/15 to 2017/18 and 2015/16 to 2018/19. For the GVA indicator, measured by calendar year, the changes were computed for the periods 2015 to 2018 and 2016 to 2019. Outliers were then identified and excluded for each indicator.[15] Specifically:

- The three-standard deviation rule was applied to detect outliers in the cross-sectional distribution of each indicator for each financial year.
- CCGs with a year-by-year change exceeding 100 per cent in the value of the indicator were also excluded.

For the NHS sickness absences indicator, the sample was further restricted to CCGs with an NHS parent site in their geography,[16] enabling direct measurement of the NHS sickness absence rate.

Step 3: Identification of CCGs with high and low change in investment in obstetrics and gynaecology services.

The third step involved categorising the CCGs into four groups based on their investment in obstetrics and gynaecology (relative to need):

- CCGs with high investment in obstetrics and gynaecology between 2014/15 and 2015/16.
- CCGs with high investment in obstetrics and gynaecology between 2015/16 and 2016/17.
- CCGs with low investment in obstetrics and gynaecology between 2014/15 and 2015/16.
- CCGs with low investment in obstetrics and gynaecology between 2015/16 and 2016/17.

A CCG was classified as a high investor for a specific period if it belonged to the top 33 per cent of the distribution of changes in NHS expenditure in obstetrics and gynaecology

[15] To minimize the impact of outlier detection on the sample, identified outliers have been excluded only for the analysis of the specific indicators they affect. This means that if a CCG was considered an outlier for the GVA indicator but not for the female employment rate, it was excluded from the sample for the former analysis but retained for the latter.
[16] A more detailed overview of NHS staff apportioning is provided in the subsequent section.

services relative to need during that period. Similarly, it was classified as a low investor if it fell into the bottom 33 per cent of the distribution.

Step 4: Use of propensity score matching (PSM) to refine the definition of the group of low investors to match the characteristics of high investors.

The fourth step involved refining the group of low investors to match the characteristics of high investors for the same period. To achieve this, propensity score matching (PSM) with replacement and without a caliper[17] was used, avoiding a reduction in the sample size. As described above, PSM is a statistical technique used to estimate the effect of a treatment, policy, or intervention by accounting for covariates.

In this context, PSM was applied to identify a subset of low investors whose initial characteristics closely matched those of the high-investors group. Each CCG in the high-investors group was matched with the most similar CCG in the low-investors group based on their propensity scores. The propensity score was calculated using the following features, all measured at the beginning of the period: total NHS expenditure relative to need, total population, the proportion of the population living in a rural environment, and the average deprivation level of the area (using the index of multiple deprivation). This matching process ensured that the comparison between high and low investors was made between CCGs with similar baseline characteristics, thereby providing a more accurate estimate of the impact of investment levels in obstetrics and gynaecology services.

Step 5: Statistical analysis of differences in the economic performances of CCGs in the groups of high and low investors.

After consolidating the high investors and low investors from both periods, the final step involved employing a Wilcoxon signed-rank test under the one-side alternative hypothesis to examine differences in the distribution of changes in economic performance indicators (as identified in step 2) between CCGs classified as high and low investors. The Wilcoxon signed-rank test is a non-parametric[18] statistical hypothesis test utilised to compare two paired samples. When considering a one-sided alternative hypothesis (as in this case), the Wilcoxon signed-rank test evaluates whether the median of the differences between paired observations is greater than or less than zero.

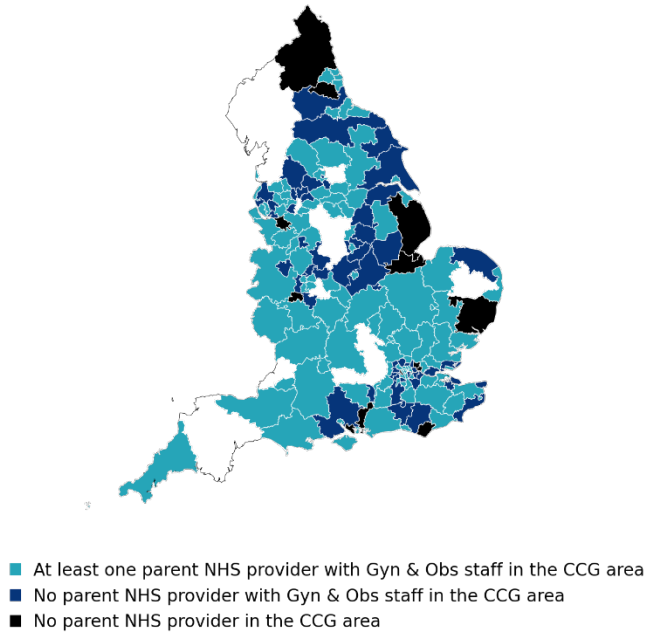
[17] Specifies the maximum distance at which two observations are a potential match.
[18] Not involving any assumptions as to the form or parameters of a frequency distribution.

II. Apportioning of NHS staff: from NHS providers to CCGs

One of the key steps in defining the proxy involved measuring the cost of medical and non-medical staff in the various specialties for each CCG. However, while CCGs directly employed a small proportion of the overall NHS staff, the majority were employed by NHS care trusts and NHS trusts (ie NHS providers). NHS providers directly delivered healthcare activities and were commissioned by the CCGs. Each provider typically operated through several sites, that is, locations and facilities where health services were delivered.

As no official historical information on commissioning relationships across CCGs and NHS providers exists in the public domain, an approach based on geographical proximity was adopted to map across these two levels. The underlying assumption behind this approach is that patients are likely to use services near their residential area rather than travelling outside of their CCG's geographical region to see a doctor. In the NHS staff data, the information is organised at the level of the headquarters (or parent-site) of each provider, with no breakdown of the number of staff by each site.

Figure 1: Availability of parent-sites of NHS providers by clinical commissioning group



Source: London Economics' analysis of NHS, ONS and GOV.UK data.

As depicted in Figure 1, CCGs can be classified into three groups, based on the presence of NHS parent sites within their geography and their staff:

- CCGs with at least one obstetrics and gynaecology specialist employed at an NHS provider parent-site within their geography (light blue).
- CCGs with NHS providers' parent-sites within their geography but without obstetrics and gynaecology specialists (dark blue).
- CCGs without NHS providers' parent-sites within their geography (black).

- The CCGs with no shading (white) had no relevant NHS parent/provider data.

Two different approaches were used to measure the ratio between obstetrics, gynaecology, and maternity services staff cost to the overall staff cost, depending on the geographical relationship between CCGs and NHS providers.

For the first group (light blue areas), the ratio between the cost of obstetrics, gynaecology, and maternity services staff to the overall staff cost was constructed using staff information from all NHS providers (parents) located in the corresponding geography.

For the second and the third groups (dark blue and black), the ratio between the cost of obstetrics, gynaecology, and maternity services staff to the overall staff cost was constructed using information on NHS providers sites (not parent), according to the following approach:

- First, for each CCG, we identified all parent sites of the NHS provider sites located in the CCG area.[19]
- Then, for each CCG, the ratio was constructed by taking the sum of the cost of obstetrics, gynaecology and maternity service staff in the parent organisations with at least one site in the CCG area and dividing it by the overall staff cost of these.

This is equivalent to a weighted average, with weights equal to the relative size (measured in staff cost) of each parent organisation. Using select ICB expenditure data, we were able to compare actual expenditure with our proxy by looking at the value of the proxy for those CCGs that were merged to define the ICB. The comparison results validated the ability of the proxy to capture the volume of expenditure in obstetrics and gynaecology services.

III. Calculations

The portion of expenditure across obstetrics and gynaecology services was determined based on the ratio of NHS staff costs across these service areas to the overall NHS staff costs:[20]

$$TI_{ccg} = Exp_{ccg}\, x\, \frac{G\&O\_staffcost_{ccg}}{All\_staffcost_{ccg}}$$

[19] Each NHS Trust or NHS Care Trust might have multiple sites where services are provided along with the main site (parent-site).

[20] Exp_{ccg} indicates the total expenditure of the CCG (excluding expenditure in the Primary Care); $G\&O\_staffcost_{ccg}$ is the total cost of medical staff in Obstetrics & Gynaecology and non-medical staff in maternity services working at NHS providers in the geography of the CCG; and $All\_staffcost_{ccg}$ is the total cost of staff working at NHS providers in the geography of the CCG

The proxy created also represents expenditure in obstetrics and gynaecology relative to need[21] across the female population. To account for differences in the population and demographic composition of the various CCGs, we then computed the NHS investment in obstetrics and gynaecology per need as:

$$I_{ccg} = \frac{TI_{ccg}}{Need_{ccg}} \, x \frac{Total\ population_{ccg}}{Female\ population_{ccg}}$$

After attributing NHS staff to each CCG based on the location of the NHS provider employing the staff member the total NHS investment in obstetrics and gynaecology services for each CCG was computed as:

$$TI_{ccg} = Exp_{ccg} \, x \, \frac{G\&O\_staffcost_{ccg}}{All\_staffcost_{ccg}}$$

where Exp_{ccg} indicates the total expenditure of the CCG (excluding expenditure in the Primary Care); $G\&O\_staffcost_{ccg}$ is the total cost of medical staff in obstetrics and gynaecology services alongside non-medical staff in maternity services working at NHS providers in the geography of the CCG; and $All\_staffcost_{ccg}$ represents the total cost of staff working at NHS providers in the geography of the CCG.

To account for differences in the population and demographic composition of the various CCGs, we then computed the NHS investment in obstetrics and gynaecology services per need as:

$$I_{ccg} = \frac{TI_{ccg}}{Need_{ccg}} \, x \frac{Total\ population_{ccg}}{Female\ population_{ccg}}$$

where $Need_{ccg}$ is the total age-adjusted population in the CCG.

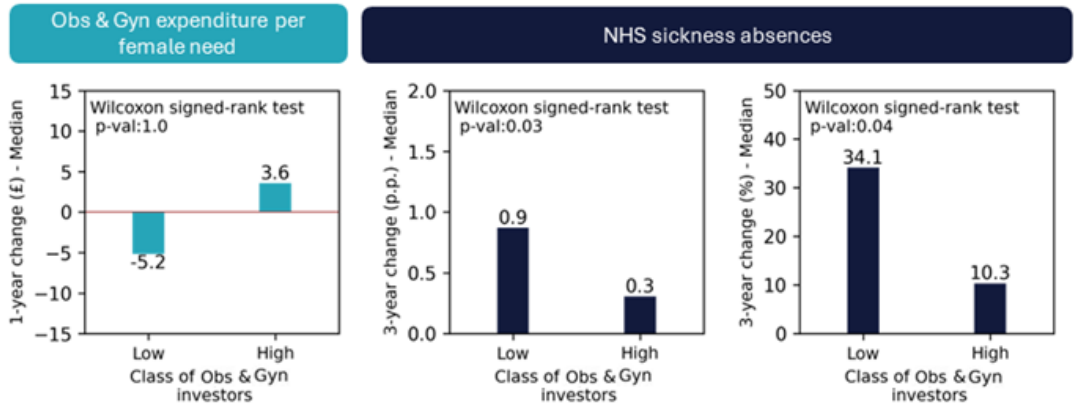
[21] NHS funding allocation is based on the needs assessed by NHSE for each geographical unit. These needs are measured by considering factors such as population size and demographics, prevalence of diseases, socioeconomic conditions, access to health services, and health outcomes. To account for these elements, rather than measuring the expenditure in obstetrics and gynaecology per woman, we measure the expenditure per woman need by normalising the NHSE need of each area based on the ratio of women to the total population in each geographical unit.

IV. Results

Comparing median values,[22] the additional GVA per capita over the following three-year period amounted to £184.30 (£451.0-£266.70), or equivalently £61.4 per year.[23] Given a difference in the median change in obstetrics and gynaecology expenditure per need between low and high investors of £11 [measured as £6.8 (-£4.2)] we found that low investors could have benefited from an additional £5.58 GVA per capita[24] for each additional pound relative to need spent in obstetrics and gynaecology services.

Analysing the data received from ICBs we identified that the median ICB had an obstetrics and gynaecology service expenditure of £103.30 per female need in 2022/23. Given the overall female population in England in 2022/23 amounted to 29.1 million, this suggests an overall expenditure in obstetrics and gynaecology services of approximately £3.0 billion. To calculate the GVA per capita for the overall population we performed the following: additional GVA per capita (£5.58) times the overall English population (57.1 million). Finally, to calculate the ROI, we performed the following: Multiplier (£5.58) x (Total population / Total female population) = 5.58 / .51.

Figure 2: Impact of NHS investment in obstetrics and gynaecology on NHS sickness absences

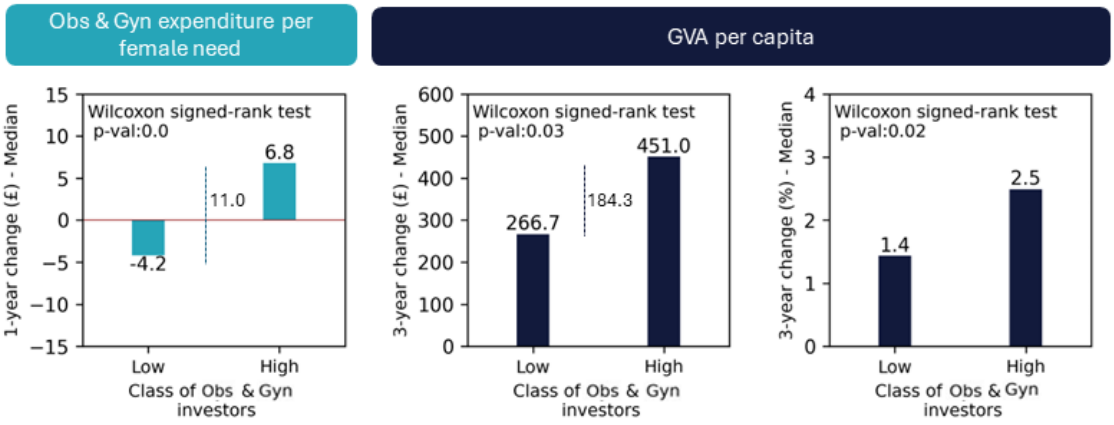


Note: The chart on the left displays the median one-year absolute change in obstetrics and gynaecology investment per female need of low and high investors. The middle chart displays the

[22] We chose to look at median values instead of average values to have a more conservative estimate (stripping out the effect of residual extreme values, if any).
[23] This means that in high-investing CCGs, the median increase in GVA per capita in the three years following the investment period was £184.30 greater than in low-investing, averaging £61.4 more per year.
[24] The additional GVA per capita is measured for both males and females to account for potential spillover effects of investment in Gynaecology and Obstetrics on men.

median three-year absolute change in the NHS sickness absences of low and high investors. The third chart displays the median three-year percentage change in the NHS sickness absences of low and high investors. The Wilcoxon signed-rank test tests the null hypothesis that two related paired samples come from the same distribution. A one-side alternative was used to evaluate whether the median of the differences between matched observations (each low investor is matched to a high investor) is less than zero.

Figure 3: Impact of NHS investment in obstetrics and gynaecology on gross value added per head



Note: The chart on the left displays the median one-year absolute change in obstetrics and gynaecology service investment per female need of low and high investors. The middle chart displays the median three-year absolute change in the GVA per capita of low and high investors. The third chart displays the median three-year percentage change in the GVA per capita of low and high investors. The Wilcoxon signed-rank test tests the null hypothesis that two related paired samples come from the same distribution. A one-side alternative was used to evaluate whether the median of the differences between matched observations (each low investor is matched to a high investor) is less than zero.

D. Expanded methodology: an assessment of women's health inequalities throughout England

I. Creating an index for each domain

To combine the selected indicators into a separate index for each of the two domains (access to healthcare and health outcomes), we performed the following:

1. Each indicator was standardised such that they could be combined on the same scale. This included remapping any data that is not reported at the local authority level and redefining some indicators so that a higher value is always 'better'.

2. Weights were assigned to the indicators based on a) the overall assessment of the data quality, b) overlap with other indicators, and c) relevance and discussion with the steering group.

3. All indicators in each domain were combined to produce a single index for each of the domains.

In general, indicators were treated as equally important, except in instances that the data was assessed to be of lower quality and a reduced weighting was agreed with the steering group. Tables 11 and 12 provide details surrounding the weighting allocated to each indicator in the index for each domain.

II. Access domain

Table 11: Indicator weights for the access to healthcare index

Indicator	Definition	Weight
Gynaecological cancer Diagnosis waiting times	The percentage of females who received a cancer diagnosis within 28 days of an urgent referral for suspected gynaecological cancer.	0.0514
Gynaecological cancer treatment waiting times	The percentage of females treated for gynaecological cancers within 31 days of receiving a decision to treat.	0.0514
Referral to treatment waiting times for gynaecology	The average waiting time (in weeks) for a completed pathway in the gynaecology service.	0.1028
Midwife staffing	The number of full-time equivalent (FTE) midwives at trust level.	0.1028
Obstetrics and gynaecology staffing	The number of FTE obstetrics and gynaecology consultants, associate specialists, and specialty doctors.	0.1028
HRT prescriptions	HRT prescriptions per 100,000 females over 40.	0.1028
Repeat Abortion Rate	Repeat abortions as a percentage of total abortions.	0.0750
Mental health, learning disability and autism services	The number of people in contact with NHS funded secondary mental health, learning disabilities and autism services per 100,000 women.	0.1028
Cervical screening coverage	Percentage of females aged 25-64 with a recent adequate test for cervical cancer.	0.1028

Human papillomavirus (HPV) vaccination coverage	Percentage of Year 9 females who have received two doses of the Human Papillomavirus (HPV) vaccination	0.1028
Contraception use	Number of females using Sexual and Reproductive Health Services for contraception (per 1,000 females aged 16-45)	0.1028

III. Outcomes domain

Table 12: Indicator weights for the health outcomes index

Indicator	Definition	Weight
Avoidable mortality	Age-standardised avoidable mortality rate per 100,000 population.	0.2333
Neonatal mortality	The number of stillbirths and neonatal deaths (before 28 days old) per 1,000 live births and stillbirths.	0.2333
Smoking status at time of delivery	Percentage of females who are known smokers at time of delivery.	0.1500
Healthy Life Expectancy (HLE)	An estimate of the proportion of lifetime spent in 'very good' or 'good' health, based on how individuals perceive their general health.	0.1167
Disability-free life expectancy (DFLE)	An estimate of the proportion of lifetime free from a limiting persistent illness that limits day-to-day activities.	0.1167
Sickness absence in the UK labour market	The percentage of working hours that are lost because of sickness absence.	0.1500