Investigating the fall in teenage pregnancy rates in the UK from 1999 onwards using systematic review and natural experimental methods

Mr. Andrew James Baxter, BSc (Hons), MSc

Submitted in fulfilment of the requirements for the Degree of

Doctor of Philosophy

MRC/CSO Social and Public Health Sciences Unit

Institute of Health and Wellbeing

College of Medical, Veterinary & Life Sciences

University of Glasgow

August 2021

Abstract

Teenage pregnancy is associated with numerous health risks, both to young women and to infants, and may contribute to entrenched social inequalities. Reducing high rates of teenage pregnancy may be an effective target for action to reduce inequalities. Between 1990 and 2016, UK countries saw relatively high rates of teenage pregnancy compared with other European high-income countries. A prominent intervention in response to these observations was England's Teenage Pregnancy Strategy (henceforth referred to as the Strategy), which was launched in 1999 and cost £280m across 10 years of implementation. Across the Strategy period there was a 51% reduction in under-18 pregnancy rates between 1998 and 2014. The Strategy was hailed as a success and as a model for other countries to follow. However, researchers have questioned this conclusion of effectiveness. My analyses first examine the range of plausible hypothesised causes of the UK's falling rates. I construct an initial causal tree of interventions and changes in social determinants presented in prominent sexual health literature and hypothesised to be causative of falling pregnancy rates. I use this tree to inform a systematic literature search of prior reviews of teenage pregnancy prevention. I synthesise the evidence presented of both the effectiveness of each cause and the exposure of the UK population using modified Bradford Hill criteria. I use this evidence to construct an updated causal tree. I present this causal tree as a summary of potentially testable causal pathways to determine the contribution of each to the observed falling rates in the UK. I then conduct an evaluation of the effect of the Strategy on teenage pregnancy rates. I use two natural experimental methods to address potential biases and triangulate between datasets, outcomes and methods. Using controlled interrupted time series models, I compare England under-18 pregnancy rates with Scotland and Wales from 1992-2016. Changes in trends in England at 1999 closely match those in Scotland (0.08 fewer pregnancies per 1000 women per year in England; -0.74 to 0.59) and Wales (0.14 more pregnancies per 1000 women per year in England; -0.48 to 0.76). In my second analysis, I use synthetic control models to construct predicted under-18 birth rate and under-20 pregnancy rate trends in England from data for 15 comparable countries. Across all models and sensitivity analyses, controls saw similarly falling teenage birth and pregnancy rates in the absence of the Strategy. I conclude that the Strategy had little or no effect on England's teenage pregnancy rates. Future policy approaches to preventing teenage pregnancy should not rely on reproduction of the Strategy to achieve similar results. Further investigation should seek to understand the true causes of falling pregnancy rates. This thesis demonstrates the value of robust public health policy evaluation and the importance of achieving accurate estimates of the effects of a policy on its target outcomes.

Table of Contents

List of tables

List of figures

List of accompanying material

All code, protocols and interactive apps used in my analyses are published openly online with URLs listed in Appendix A and linked at **phd.andybaxter.me**.

The analyses reported in Chapters 5 and 6 are published as a journal article:

> **Baxter, A. J.**, Dundas, R., Popham, F. and Craig, P., (2021) 'How effective was England's Teenage Pregnancy Strategy? A comparative analysis of High-Income Countries', *Social Science & Medicine*, 270. doi: 10.1016/j.socscimed.2021.113685.

Acknowledgements

A great number of people have supported and helped me through the years of completing this thesis. Naomi joined with my excitement at being accepted to undertake this task and similarly with my excitement in nearing the finish line. Without her support and her joining with me in the vision of completing a PhD, I certainly could not have done it. Sophia and Martha also have been a joy throughout it all and have helped me in the closing year to keep focussed on finishing my thesis within the bustle of a lockdown household. This thesis is primarily dedicated to the three Baxter girls for their love and support.

I'd like to thank Mum and Dad for cheering me on through all stages of my education and for their help in looking after kids when it was needed to have rest, keep the house in order, or get some thesis writing done. Also, a thank you to wider family and in-laws for support and help in these years.

I would like to extend a huge expression of gratitude to my supervisors Peter, Ruth and Frank who helped shape this project from the very start and believed in me and my ability to do it well. They have helped me to get the most out of the project and the resources of the University to maximise my learning and development throughout. Alongside them, the researchers, staff and fellow students at the Social and Public Health Sciences Unit have been a marvellous team to work with and have additionally helped and encouraged me through everything. I'd like to thank Cindy, Vittal, Hillary and Emily, who through my master's year helped me gain the confidence that I could take on a PhD. I'd also like to thank Marissa, Mirjam, Mark and Danny for the many opportunities to work with them on the range of exciting and interesting projects in the Unit and beyond and the chance to keep learning and developing as a researcher through them. A further thankyou to Richard, Mike and Uncle Andy for reading over my thesis and providing valuable feedback.

Throughout the stresses, strains and joys of these last four years in Glasgow, our church family have consistently supported and cared for us. For all those who rallied round us through the difficult periods, helping us move house, look after kids, and mourn the loss of friends and family, I want to express my thanks. In particular, I have benefitted from the ongoing reminders of my calling to love the Lord with all my heart, soul mind and strength. I present this thesis as an expression of my gifts used in gratitude to fulfil this calling and I present my contribution to public health research as an attempt to show love for my neighbours.

Author's declaration

I declare that I am the sole author of this thesis, except where the contribution of others has been acknowledged. The work contained in this thesis has not been submitted in any form for another degree or professional qualification at the University of Glasgow or any other institution.

Andrew James Baxter

August 2021

Chapter 1 Introduction

1.1 Overview of chapter

In my introductory chapter, I lay out how the topic of teenage pregnancy is conceived as a public health priority leading to interventions with pregnancy rate reduction as a goal. I then draw out how concerns with the impact of high pregnancy rates on health and social inequalities prompted the introduction of the Teenage Pregnancy Strategy, henceforth referred to as the Strategy. I highlight some of the questions asked of the Strategy's approach and its evaluation, and how these prompt wider questions of what the hypothesised causes of changing rates of teenage pregnancy are. I then discuss outcome evaluation and the natural experiment methods which I will use to evaluate the Strategy. I close by summarising the proposed value and impact of the output of this thesis, in both setting out the plausible hypothesised causes to be tested and providing a robust estimate of the Strategy's contribution to the observed fall in teenage pregnancy rates.

1.2 Teenage pregnancy as a public health priority

In recent decades, across several high-income countries, teenage pregnancy has been conceptualised as a public health problem requiring preventative action. Childbearing at adolescence is associated with numerous health risks, both to mothers and infants. Teenage pregnancies are more likely to be unintentional than are adult pregnancies (Wellings *et al.*, 2013). Unplanned pregnancies are more likely to end in abortion, may present greater health problems and disadvantages to young women, and result in greater cost to society (Glasier and Wellings, 2012). Additionally, teenagers undergoing pregnancy and maternity face greater risks of to their own and their children's health than do adults. Such problems include maternal anaemia, pre-eclampsia, infant mortality, pre-term labour, and longer and more difficult labour (Irvine *et al.*, 1997; Social Exclusion Unit, 1999). Teenage mothers experience worse rates of poor mental health, suicide, and substance use problems than non-parenting teenagers (Hodgkinson *et al.*, 2014).

Early pregnancy is more common among women from poorer families, single-parent households, areas of greater deprivation, and those born to teenage parents (Irvine *et al.*, 1997; Social Exclusion Unit, 1999; Information Services Division Scotland, 2018a). Teenagers with a previous pregnancy are up to five times more likely to experience rapid repeat pregnancies (Falk *et al.*, 2006). Teenage parents are more likely to face barriers to

further education, employment or training, and may require greater social support for parent and child health, positive family relationships, and income and housing support (Bradley, Cupples and Irvine, 2002; Letourneau, Stewart and Barnfather, 2004). Advocates of teenage pregnancy prevention as a public health priority propose that reduction in rates could improve the health outcomes of young women and infants and reduce health and social inequalities (Social Exclusion Unit, 1999; World Health Organization, 2020).

1.3 Measuring and comparing pregnancy rates

In pregnancy prevention responses, prevention of the conception itself is commonly the primary goal (Social Exclusion Unit, 1999; Glasier and Wellings, 2012; Mitchell, Wellings and Zuurmond, 2012). To measure the effectiveness of an intervention or approach, an accurate measure of the rate of conceptions to teenagers would be a preferred outcome. However, as a conception may result in maternity (live birth of one or more child), still birth, spontaneous miscarriage or abortion, conception rates themselves are difficult to measure accurately across a population. Across vital statistics reports, global estimates and journal published measurements of teenage pregnancy, I encountered several approaches to recording prevalence. Summations of pregnancy-related outcomes across a population were often stratified by age to produce outcome rates across the broad category of 'teenager' (e.g. categorising by aged 'Under 16', 'Under 18' and 'Under 20'; Information Services Division Scotland, 2018a; Office for National Statistics, 2019). Reports of pregnancies in cross-country comparisons most commonly used counts of births to teenage women to calculate a 'birth rate'. This metric does not take account of pregnancies ending in abortion and stillbirth, which may vary between countries and therefore may give biased estimates of true rates of conception. However, as abortion and stillbirth data may not be equally reliable or complete across countries, disregarding them and using solely measures of births as an outcome may allow more reliable international comparisons (Sedgh *et al.*, 2015; Max Planck Institute for Demographic Research, 2019). Where reliable data recording abortions and still births are available, these can be used alongside birth statistics to estimate total pregnancies and produce a 'pregnancy rate'. Within the UK, for instance, recorded dates of birth of women and weeks of gestation at birth, still birth or abortion are used to calculate pregnancy rates by age at conception (Information Services Division Scotland, 2018a; Office for National Statistics, 2019). The UK approach can account for differing patterns of pregnancy outcomes across time and between countries. This measure allows a closer look at the rate of conceptions by accounting for the most common observable results of a teenage pregnancy (Sedgh *et al.*, 2015).

When estimating pregnancy rates globally across recent decades, low-income countries tend to show higher rates, whilst high-income European countries have generally shown lower rates (Sedgh *et al.*, 2015; Paton, Bullivant and Soto, 2020; World Health Organization, 2020). Amongst high-income countries, several English-speaking countries, including the UK, saw sustained high rates of teenage pregnancy across the 1990s and early 2000s (Paton, Bullivant and Soto, 2020). At their most recent peak in 1998, England saw 47 pregnancies per 1,000 women aged 15-17, Scotland saw 45 per 1,000 and Wales 55 per 1,000 (Information Services Division Scotland, 2018a; Office for National Statistics, 2019).

When comparing under-20 birth rates across countries, the teenage pregnancy report presented to parliament by the Social Exclusion Unit (1999) showed that rates of births to women aged under 20 in the UK from 1973 to 1996 were consistently higher than those in all other European high-income countries (HICs). Several of these European HICs had seen rapid declines in numbers of births to teenagers through the 1970s and 1980s. This effect was not seen to the same extent in the UK (Social Exclusion Unit, 1999, fig. 3). The report additionally highlighted that other English-speaking HICs – Australia, Canada, New Zealand and the U.S.A. – experienced similarly high rates (Social Exclusion Unit, 1999, fig. 2). In my own data examining both births and abortions to under-20s, I also observed that the UK had shown distinctly high rates across the period 1990 to 2000 compared to similar countries, with no clear explanation (Figure 1-1). The comparably higher rates in UK countries have been cited as a reason for action to prevent teenage pregnancies (Social Exclusion Unit, 1999; Hadley, Ingham and Chandra-Mouli, 2016; Ma, 2016).

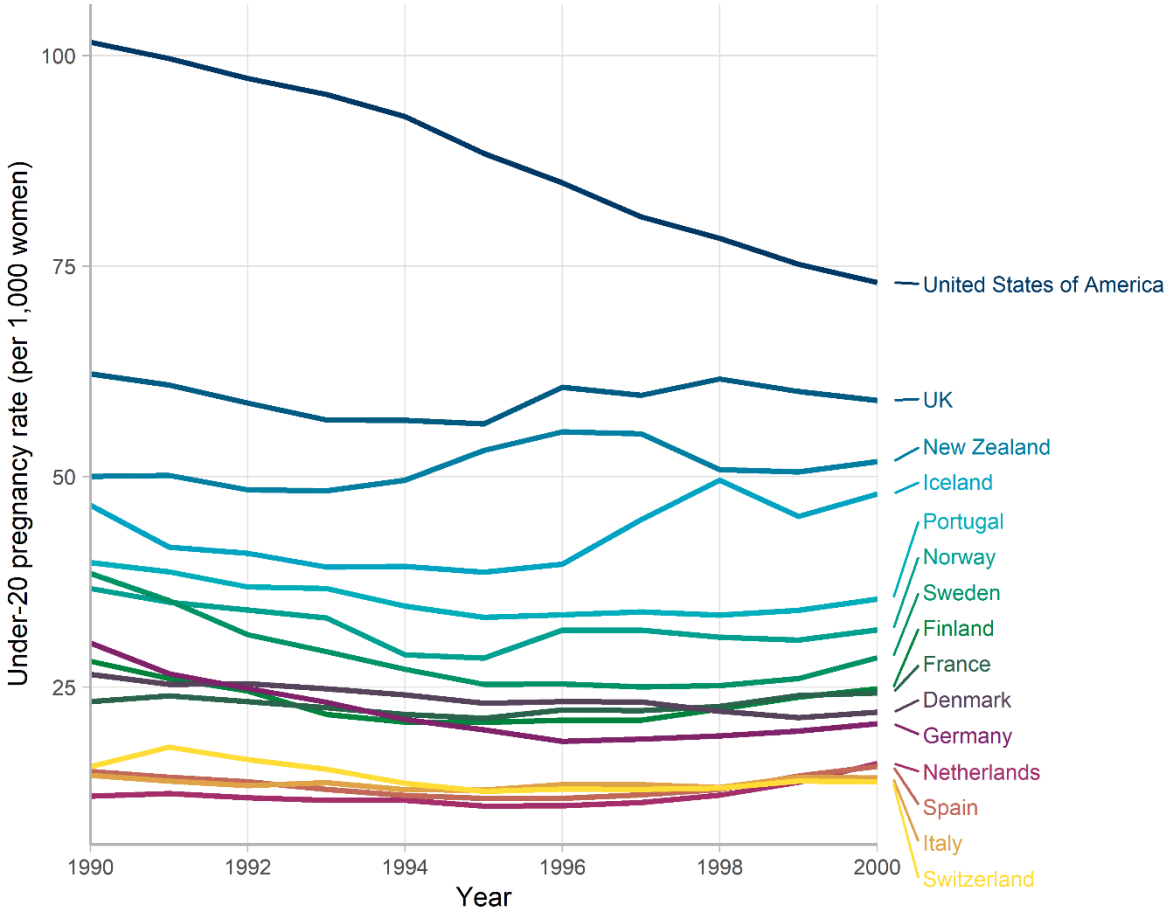

Figure 1-1 – Under-20 pregnancy rates in high-income countries from 1990 to 2000. Pregnancy rates are calculated from summing total births and abortions, divided by number of women aged 15-19 in each population. These rates were calculated from data estimating births by age of woman and population by age of woman for each country from the Human Fertility Database and the Human Mortality database (Max Planck Institute for Demographic Research, 2019; University of California and Max Planck Institute for Demographic Research, 2019) and numbers of abortions to women under-20 from the WHO Health for All Explorer (WHO Regional Office for Europe, 2019), and supplemented with data from other sources for countries not included in these datasets (Kost, Maddow-Zimet and Arpaia, 2017; Statistics New Zealand, 2019b, 2019a).

1.4 England's Teenage Pregnancy Strategy

One prominent policy approach to reducing high rates of teenage pregnancies was England's 1999 'Teenage Pregnancy Strategy' (Social Exclusion Unit, 1999). The Strategy was developed following the 1998 peak in pregnancy rates and combined several approaches to reducing rates into one coordinated policy response. The research underpinning its development was carried out by the Social Exclusion Unit, which was established by the Labour Government in 1997 shortly after their election victory. The Unit was founded with the express aims of tackling a range of identified societal inequalities (Blair, 1997b, 1997a). In light of this goal, the Unit published a report to parliament, launching the Strategy (Social Exclusion Unit, 1999). The report aimed to summarise the problems presented by teenage pregnancy in England, to present reasons for action, to synthesise evidence of previous successful interventions and to set in place a ten-year

strategy to lower pregnancy rates. This action was presented with the goal of "combat[ting] the risk of social exclusion for vulnerable teenage parents and their children" (Social Exclusion Unit, 1999, p. 2).

The Strategy aimed to reduce England's under-18 pregnancy rates by 50% in ten years, whilst also aiming to address the effects of parenthood on those who became teenage mothers (Social Exclusion Unit, 1999; Hadley, Ingham and Chandra-Mouli, 2016). To achieve this, it took a multifaceted approach to reducing rates of teenage pregnancy and addressing associated health and social problems. To reduce pregnancies, the Strategy report set out a coordinated action plan linking national and local implementation. Within schools and local clinics, it recommended improvements in sex education and pregnancy prevention resources, including interventions to improve adolescents' access to contraceptives. It also recommended national media-campaigns aimed at teenagers and parents. To address challenges faced by teenagers who give birth, it recommended greater support for young parents to remain in education and access housing, and other measures to improve health and healthcare. To coordinate all efforts and to ensure equal effectiveness in all areas, the Strategy prompted structured and 'joined up' action at national and local level – manifested as local co-ordinators and strategies, local implementation funds, and a national task force and advisory group to support and monitor progress. The Strategy was claimed by its development and evaluation teams to be the first of its kind, coordinating local and national action to reduce pregnancies nationwide (Social Exclusion Unit, 1999; Teenage Pregnancy Strategy Evaluation, 2005).

Implementation of the Strategy prompted increased funding directed at Strategy initiatives. Beginning with local implementation grants awarded in 1999, £60m of funding was allocated for the first three years of the Strategy (Social Exclusion Unit, 1999; Teenage Pregnancy Strategy Evaluation, 2005). These additional resources were targeted to areas with high pregnancy rates (Social Exclusion Unit, 1999; Hadley, Chandra-Mouli and Ingham, 2016). A mid-term review in 2005 led to significant changes in implementation aiming to intensify Strategy efforts. These changes included publication of new guidance for local authorities, a redesigned media campaign, new health and education programmes, and further measures to increase access to contraception (Hadley, Ingham and Chandra-Mouli, 2016). Expenditure on the Strategy from central government, local authorities, health authorities, other government programmes and charities, came to £167.6m by the mid-term review in 2005 (Teenage Pregnancy Strategy Evaluation, 2005) and reached an estimated £280m by the end of Strategy activity in 2010 (Billingsley, 2011).

Observations of under-18 pregnancy rate changes across the Strategy period allowed examination of the Strategy's progress in reducing these (Figure 1-2). A mid-term evaluation observed small decreases in rates in England relative to pre-1998 baseline rates, but little difference in comparison to changes seen in Scotland and Wales which were not exposed to the Strategy (Wilkinson *et al.*, 2006). A later analysis, conducted after the end of the Strategy implementation period, noted a fall in rates of teenage pregnancy across this period and beyond, from 47.1 pregnancies per 1,000 women aged under 18 in 1998, to 22.9 per 1,000 women by 2014 – a reduction of 51% (Wellings, Palmer, Geary, *et al.*, 2016). This was compared with a mean reduction of 22% in under-18 births across 28 European comparison countries (Wellings, Palmer, Geary, *et al.*, 2016). Both the mid-term and post implementation period studies observed greater decreases in rates in areas with greater Strategy-related spending (Wilkinson *et al.*, 2006; Wellings, Palmer, Geary, *et al.*, 2016).

The Strategy was deemed a success following the observed declining rates (Wilkinson *et al.*, 2006; Hadley, Chandra-Mouli and Ingham, 2016; Hadley, Ingham and Chandra-Mouli, 2016; Ma, 2016; Skinner and Marino, 2016). The authors of the final analysis concluded that the Strategy, "alongside other social and educational changes, has probably contributed to a substantial and accelerating decline in [under-18] conceptions" (Wellings, Palmer, Geary, *et al.*, 2016, Abstract).

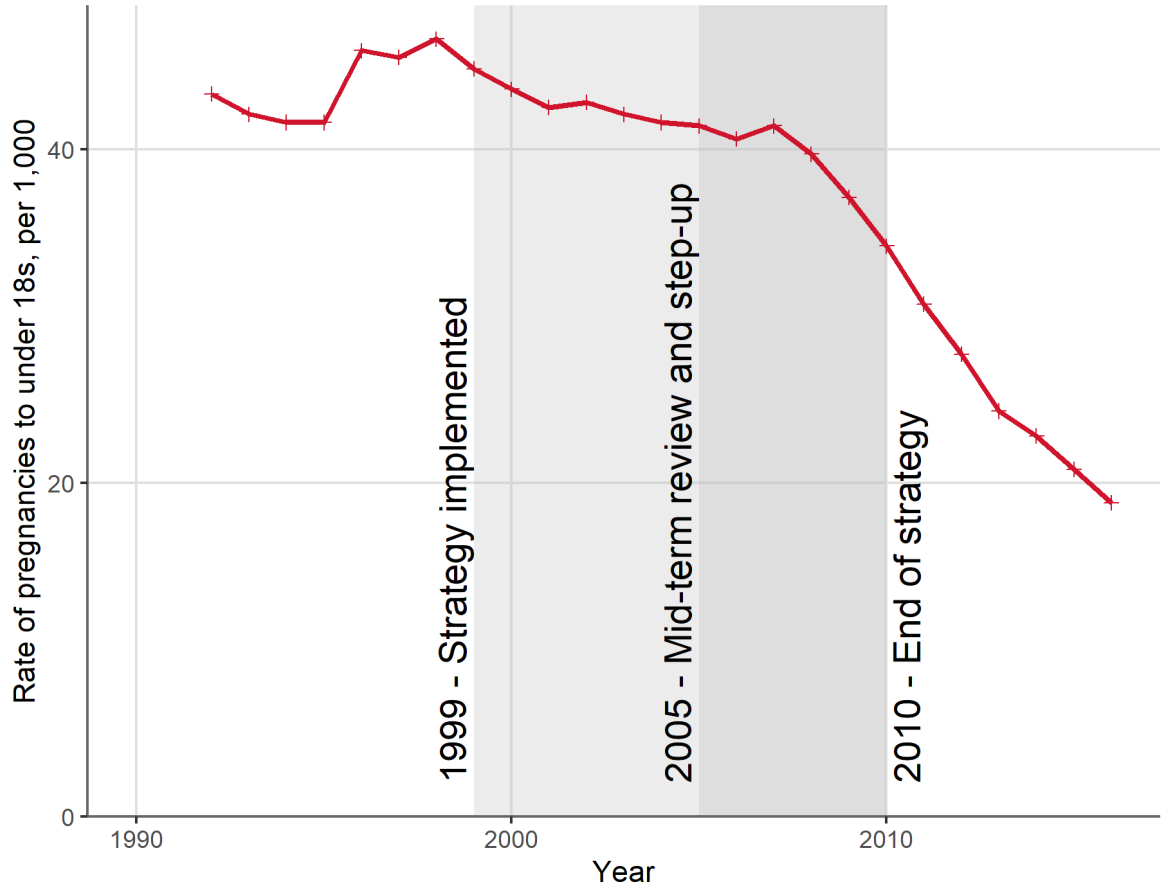

Figure 1-2 – England under-18 pregnancy rates from 1992 to 2016, surrounding the Strategy period highlighted from 1999 to 2010. The 2005 mid-term review prompted suggested improvements in the Strategy and a hypothesised period of more intense exposure, highlighted from 2005 to the Strategy's end in 2010. The continuation of the trend after 2010 was interpreted by evaluators of the ongoing effectiveness of "mainstreamed" Strategy elements (Wellings, Palmer, Geary, *et al.*, 2016). Data from ONS Conception Statistics report (Office for National Statistics, 2019).

The Strategy has been promoted as a unique, national approach, whose substantial cost was justified by the observed fall in pregnancies (Hadley, Ingham and Chandra-Mouli, 2016; Skinner and Marino, 2016). It has been held up as a replicable model for implementation in countries with similarly high rates and for ongoing government action in the UK (Hadley, Ingham and Chandra-Mouli, 2016; UNESCO Education Sector, 2017; Public Health England and Local Government Association, 2018).

1.5 Querying the Strategy's apparent success

A number of studies have queried the approach taken by the Strategy and claims of its success. Several authors have questioned whether framing teenage pregnancy prevention as a public health priority is justified (Lawlor, Shaw and Johns, 2001; Lawlor and Shaw, 2002; Arai, 2009). Others have questioned whether becoming pregnant as a teenager is a major contributor to the range of poor social and health outcomes discussed above, or whether confounding due to other socioeconomic factors may be responsible for the bulk

of the inequalities associated with teenage pregnancy (Hoffman, 1998). Other action to tackle the societal structures perpetuating these inequalities may have greater effect (Lawlor, Shaw and Johns, 2001).

In developing the Strategy, the report authors discussed various conflicting and overlapping approaches advocated for and used in pregnancy prevention (Social Exclusion Unit, 1999). Commonly implemented approaches use varying routes to engage with teenagers, such as school lessons, clinics and healthcare workers, media, and peers (Collumbien and Macdowall, 2012; Forrest and Wellings, 2012). Often interventions may target pregnancy alongside other outcomes, such as sexual morality, sexually transmitted infections, substance use, mental health and other areas of sexual health (Mitchell, Wellings and Zuurmond, 2012). The goals and messages of such programmes may vary according to the aims and values of the bodies implementing them. Varying approaches include promoting abstinence from sex, prompting more consistent or more effective contraceptive use, or restricting interventions to providing education and information. Whilst ostensibly sharing the common target outcome of pregnancy prevention, such a variety of approaches produces conflicting sets of practices and competing claims of effectiveness. Whilst the Strategy report sought to use prior evidence to identify effective responses, more recent reviews have queried the effectiveness of these approaches. Programmes to improve sex education and contraception provision, prominent components of Strategy delivery, were not found by later reviews to clearly reduce pregnancy rates (Marseille *et al.*, 2018; Andrzejewski, Liddon and Leonard, 2019).

Queries have also been raised as to whether the focus of the Teenage Pregnancy Strategy was appropriate (Arai, 2009). In the report setting out the Strategy, the Social Exclusion Unit (1999) identifies three mediating factors it sets out as modifiable targets: ‘low expectations’, ‘ignorance’ and ‘mixed messages’ (Social Exclusion Unit, 1999, p. 7). Ultimately, however, the Strategy appears to have considered ‘ignorance’ to be the most influential contributor to high rates of pregnancy, presenting it as the easiest, most feasible, most economical and most acceptable to remedy (Carabine, 2007; Arai, 2009). This approach reflects a view that teen pregnancy is “an educational or medical problem to be solved by increased access to contraception, abortion, and sex education” (Geronimus, 1997, p. 408). This assumption was manifested in the Strategy’s emphasis on information and education alongside provision of contraception. It is possible that more effective approaches were missed as other concerns outweighed the stated goal of producing change through evidence-based strategies.

Later studies have queried whether England's decrease in teenage pregnancy rates is due to the Strategy. Craig et al. (2016) note that pregnancy rates followed a similar pattern in England and the other UK countries (Figure 1-3). Paton et al. (2020) show that across several countries, teenage pregnancy prevention policies with components similar to the Strategy do not explain the observed trends in pregnancy rates. Cuts in spending in areas of England, effectively halting Strategy-related activity, did not lead to an increase in pregnancy rates in an expected dose-response relationship (Paton and Wright, 2017). A recent study that shows continued decreases in rates beyond the ending of the Strategy further calls into question whether the earlier fall was attributable to the Strategy (Heap, Berrington and Ingham, 2020).

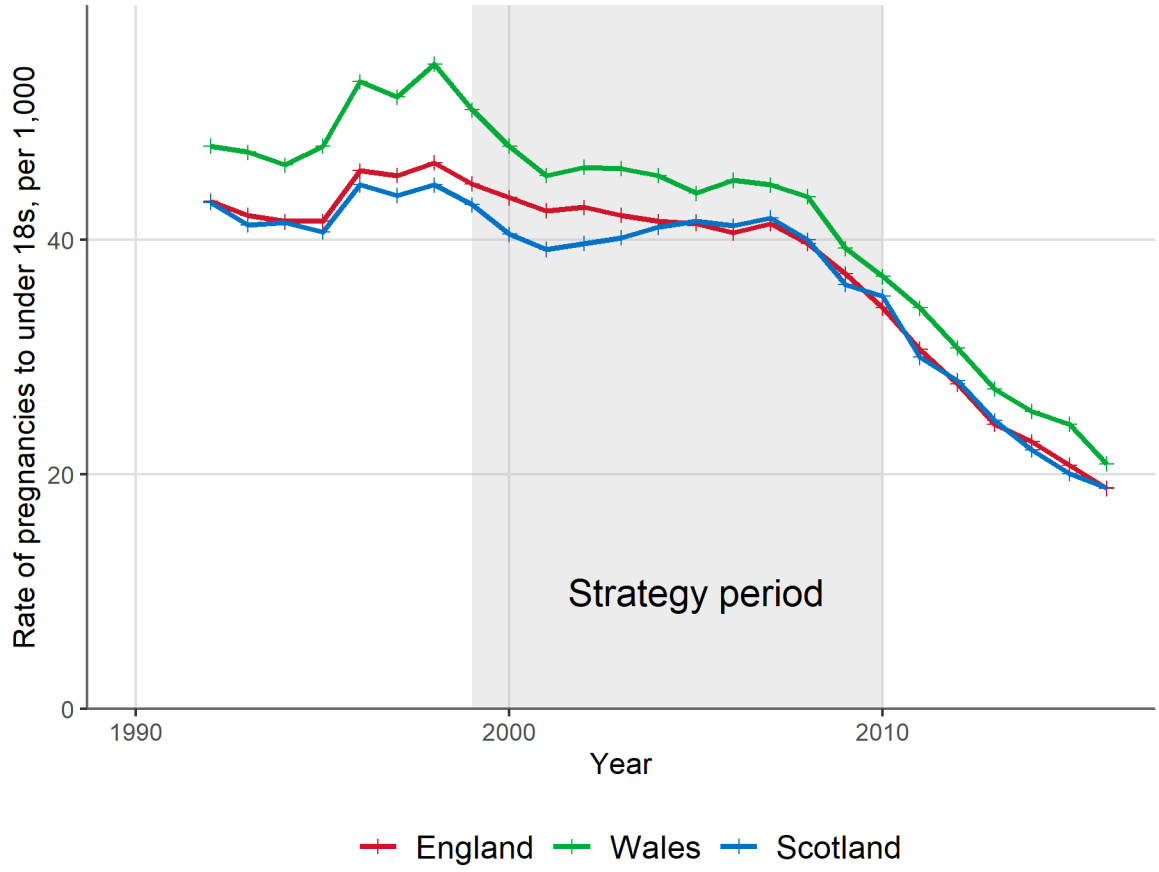

Figure 1-3 – Under-18 pregnancies across England, Scotland and Wales from 1992 to 2016, with the Strategy period highlighted in grey from 1999 to 2010. Data from ONS Conception Statistics report (Office for National Statistics, 2019) and ISD Scotland Teenage Pregnancy report (Information Services Division Scotland, 2018a).

Large declines in teenage pregnancy rates amongst high-income countries suggests the influence of common causes across many countries, including the UK (Sedgh *et al.*, 2015). Observers have proposed that several interventions intended to reduce risk of pregnancy amongst teenagers, such as improved sex education in schools and contraception distribution programmes, rather than a single effective policy approach, may have each

contributed to this overall decrease (Bache and Wickham, 2014; Kearney and Levine, 2015a; Wellings, Palmer, Geary, *et al.*, 2016; Sipsma *et al.*, 2017). Other potential candidates for explaining the effects, not arising from explicitly targeted actions, have been suggested. Changes in employment, increased participation in education, wider changes in welfare and social services provision, or increased use of personal electronic devices have been suggested as contributing to lower pregnancy rates (Girma and Paton, 2015; Kearney and Levine, 2015a; Sipsma *et al.*, 2017; Twenge, 2017). The relative contributions of each of these hypothesised causes remain unknown.

1.6 The methodological challenges of policy evaluation

Outcome evaluation of a policy or intervention is a key element of evidence based policy making (Craig *et al.*, 2019). However, in many cases of public health interventions, robust outcome evaluation has not been carried out (House of Commons Health Committee, 2009; Theis and White, 2021). Neglecting policy evaluation may give rise to repetition of expensive, ineffective, or even counterproductive interventions. Such interventions may incur an 'opportunity cost' by taking the place of more effective approaches. The House of Commons Heath Committee report (2009) note that “time and time again” (ibid. para 360) evaluation has focussed on aspects of process evaluation or implementation but has omitted outcome evaluation. Process and implementation evaluations do not incorporate tests for whether a health policy worked at addressing the issues it was launched to target, and thus “are not a substitute for an outcome evaluation” (Craig *et al.*, 2019, p. 12). Ongoing efforts to address health inequalities are hampered by a lack of evidence assessing what works, preventing evidence-based development of effective approaches. Thus, in directing large sums of public money to improve population health, “Evaluation is an ethical imperative” (House of Commons Health Committee, 2009, para. 361).

In other cases, outcome evaluation may be intended, but may be inadequately planned for or carried out. Theis and White (2021) note that often a policy’s design may not lend itself to evaluation of its effectiveness. A common occurrence appears to be a failure to identify measurable outcomes or setting targets (Theis and White, 2021). Additionally, though it is recognised that randomised trials are not always suitable (Craig *et al.*, 2019), opportunities to use these robust methods to evaluate complex interventions may be overlooked in the pathway from development to implementation. The pre-specification of targets and conducting of robust effectiveness evaluations are key elements in producing effective health policies and learning from past successes or failures (Baggott, 2012).

The Strategy was a large-scale intervention, costing an estimated £280m. Presenting the policy as an evidence-based programme, and reflecting the importance of testing its effectiveness, developers set clear goals and timeframes for planned evaluation in the Strategy report action plan (Social Exclusion Unit, 1999, sec. 11). In reducing pregnancy rates, the report set a target of halving rates of pregnancy in under-18s by 2010 (Social Exclusion Unit, 1999, p. 8). This target allowed comparisons across the period to measure the change associated with exposure to the Strategy. The fall in rates of 51% by 2014 was presented as evidence that the Strategy met its ambitious target, albeit after a few additional years, and therefore it was considered to be a successful intervention in achieving its pregnancy reduction goals (Toynbee, 2013; Hadley, 2014; Ma, 2016; Skinner and Marino, 2016).

Although evaluation was planned and conducted on the Strategy, the findings of falling pregnancy rates themselves do not provide an adequate indication of effectiveness. Methods previously used in evaluation may have given biased estimates of effect sizes. In both the midpoint evaluation and research conducted after the Strategy's end, pre-post estimates of differences in annual pregnancy rates were used as a primary indicator of effect sizes (Wilkinson *et al.*, 2006; Wellings, Palmer, Geary, *et al.*, 2016). The comparisons drawn are between an unexposed population pre-1999 and an exposed population several years later. This method relies on the assumption that the Strategy exposure is the only cause of change in the outcome (Craig *et al.*, 2017). Although the authors do note that similar falling rates in other countries indicate broader secular trends, they conclude that the Strategy has probably contributed to some of the effect seen in England (Wellings, Palmer, Geary, *et al.*, 2016). Further researchers and commentators attribute all or most of the effect to the Strategy (Toynbee, 2013; Hadley, Chandra-Mouli and Ingham, 2016). To test and quantify the true effects of the Strategy however, methods that can rule out other possible explanations are needed. The inability of previous evaluations to account for the effects of other possible exposures gives rise to the need to re-evaluate the Strategy's effectiveness and fulfil the ethical imperative of robust evaluation using better research approaches.

1.7 Methods for testing causation

In picking apart the relative effects of implemented interventions and other possible exposures upon a population, several problems are presented as to the validity of causal inference. Attributing observed changes in populations to policies occurring at the same

time as these changes, or to other feasible-sounding causes, is in danger of assuming effects where no effects exist. Reverse causation, common causes or other confounding variables such as time may give rise to spurious causal inference (Hernán, Hernández-Díaz and Robins, 2004). A true measure of the causal effects of an intervention or exposure is the difference between two potential outcomes, which cannot both be observed and measured for a single unit. These outcomes are the observed effects on a unit under one condition, and the unobservable effects on that same unit under the opposite condition (Hernán and Robins, 2020). In comparing two units, one exposed and one unexposed, and estimating the effects from the difference in outcomes, one must ensure that they are exchangeable – that exposure would cause the same change in both units (Craig *et al.*, 2017; Hernán and Robins, 2020). The mechanisms of assignment to exposed and unexposed conditions, for example by living in different countries with different policy approaches, may also contribute to differences in outcome independent of the exposure itself. By understanding and accounting for such assignment mechanisms it may be possible to correct for these and expect exchangeability between exposed and unexposed groups, allowing testing of the hypothesised effects of exposure (Craig *et al.*, 2017; Hernán and Robins, 2020).

Randomised controlled trials (RCTs) are one approach to ensuring exchangeability, by breaking the link between assignment and other characteristics which may be associated with outcomes (Hernán and Robins, 2020). The findings of a well-conducted RCT are often read as the true measurement of an intervention's effects, where different measures given by observational data are presumed to be biased (Guyatt *et al.*, 2000; Dagan *et al.*, 2021; Hernán, 2021). Evidence appraisal tools, such as the GRADE guidelines, recommend that randomised trials should be assigned a higher initial rating of quality of evidence (Guyatt, Oxman, Vist, *et al.*, 2008; Balshem *et al.*, 2011). Cochrane review guidelines recommend randomised trials, cautioning that non-randomised study designs are likely to give greater potential for biases and that they "vary in their ability to estimate a causal effect" (Reeves *et al.*, 2020). In most iterations of 'evidence hierarchy' pyramids, randomised controlled trials are a discrete category placed above observational study designs, with only systematic reviews of evidence from trials assigned a higher status (Murad *et al.*, 2016).

Randomised trials may be held as a preferred route for evaluating health interventions, but such evaluations may not always be possible (Craig *et al.*, 2019). In the case of pregnancy prevention, issues of sufficient sample size to detect a rare event, contamination between

participants and non-participants, and the ethical problems of withholding treatment from control groups may limit the feasibility of randomised trials in assessing certain interventions (Black, 1996). Additionally, in cases of complex behaviour change interventions, issues of blinding of participants, assessors and evaluators; generalising from trial implementation to roll-out effectiveness; heterogeneous treatment-as-usual controls; and differential loss to follow-up can also introduce bias and limit the power of the trial to detect true effect sizes. In cases of natural changes of exposure, such as cultural changes, experimental manipulation is impossible and randomised trials are a wholly unsuitable set of methods to test causal relationships which may exist.

When RCTs are not possible, not ethical or not desirable, several other methods can be used to investigate hypothesised causes and provide estimates of effect sizes (Glasziou *et al.*, 2007; Craig *et al.*, 2012, 2017, 2019; Hernán and Robins, 2020). In some cases, natural experimental methods can provide an alternate route to achieving confidence of causal inference (Craig *et al.*, 2012, 2017, 2019). In a natural experiment, as opposed to a classical experiment, exposure and non-exposure to a hypothesised cause are not under the control of the researcher (Craig *et al.*, 2012). In most cases exposure is unlikely to be randomised and so introduces bias and confounding to uncorrected comparisons between units (Craig *et al.*, 2012). However, if mechanisms assigning units to exposed and unexposed conditions are understood and can be modelled, this can help to correct for the resulting biases. This allows conditional exchangeability and an estimate of the exposure's contribution to observed differences between units (Craig *et al.*, 2017).

Natural experimental methods use a variety of approaches to account for potential confounders. These methods propose treating an event which divides a population into exposed and unexposed groups as an experiment with conditional assignment. Accounting for assignment allows comparison with an exchangeable control and thereby sets these methods apart from classical 'observation' studies. Both may employ similar statistical tests of association; however, natural experimental methods aim to test causal hypotheses rather than assume causal relationships. Natural experiment methods have additional advantages over randomised trials. They are able to measure the real-world effect of an intervention, removing it from the 'laboratory conditions' of an RCT, can examine effects on whole populations, and are relatively inexpensive and convenient if data are already available (Craig *et al.*, 2012, 2019).

Natural experimental methods may be an option for investigating some of the varied hypothesised causes of reducing teenage pregnancy rates suggested by observers, where other methods are unable to do so. In circumstances where RCTs are impossible and traditional observational studies provide little confidence of causal inference on their own, natural experiments may provide us with a third way of testing some of these hypotheses. In the case of the Strategy, it is clearly no longer possible to randomly assign the 1999 population of England to exposed and unexposed conditions. The England-only scope of the policy and the well-defined starting point may allow us to treat this event as a natural experiment and gain a clearer understanding of the Strategy's effects.

1.8 This thesis

In this thesis I aim to investigate the conflicting claims of the contribution of the Strategy to the observed fall in pregnancy rates from 1999 to 2016 in England as a contribution to the area of policy evaluation and the Strategy. The substantial cost of the Strategy, and its promotion as a model for other countries, mean that reliable estimates of its impact are important for future policy making. I investigated the comparative contribution of the Strategy itself to the change in England alongside other unexamined explanations for the reduction. Alongside the Strategy, I hypothesise that several other interventions and cultural changes may have contributed to falling rates.

In my literature review chapter (Chapter 2), I begin by discussing and critiquing the Strategy's development within the wider questions within public health of how to address teenage pregnancy. I identify several gaps in the process of developing it as an effective intervention to achieve maximum impact on pregnancy rates. From this literature I draw out the range of suggested contributors to falling teenage pregnancy rates and construct an initial causal tree of prominent hypothesised causes.

In my methods chapter (Chapter 3), I describe how I develop and test this causal model. I report my results in my empirical chapters (Chapters 4, 5 and 6), answering my research questions in two stages:

- First, in noting the similar falls in rates in Scotland, England and Wales across the period, I examine the evidence for plausible hypothesised causes of the trends presented in the research literature. By systematically searching for and synthesising existing reviews, I propose an updated causal tree of all hypothesised

causes, summarising evidence currently available and prompting further investigation. This includes interventions with some evidence of effectiveness, changes in social determinants with hypothesised effects on pregnancy, and policies which have been implemented, whether or not their impact has been evaluated. It excludes interventions or changes for which there is clear evidence suggesting their ineffectiveness in reducing teenage pregnancy rates. My evidence-informed causal tree provides a backdrop of potential contributors to observed change, which can be considered alongside the Strategy. It further suggests avenues for ongoing causal investigation.

- Second, to examine the Strategy's contribution to reducing rates of teenage pregnancy, I apply natural experiment methods to compare England with exchangeable comparators. To account for potential limitations in methods and to triangulate my results, I test the effectiveness of the Strategy in two ways. In my first analysis, I choose Scotland and Wales as comparators given their similarity to England in other characteristics which may affect teenage pregnancy rates. I use interrupted time series methods to compare each country with England across the implementation period and up to most recent observations. In my second analysis I compare birth and pregnancy rates in England with those of a wider pool of potential comparator countries using synthetic control methods. Considering both approaches together, I produce a robust estimate of the Strategy's effects, triangulated across two methods and datasets.

The questions addressed by this thesis are important because of the substantial cost of the Strategy and the imperative to adequately evaluate the outcomes of health interventions. Understanding whether the Strategy made a substantial contribution to the falling rates, or whether these results were achieved through action elsewhere or were driven by social changes independent of the Strategy, can inform ongoing action in the UK and globally. My analyses seek to contribute a clearer understanding of the complex area of the multiple possible causes of falling teenage pregnancy rates and demonstrate a robust policy evaluation of a key intervention design.

Chapter 2 Literature Review

2.1 Overview of chapter

In this literature review I set out key research findings on teenage pregnancy prevention and questions prompted for further investigation. I identify several uncertainties in the evidence underpinning the development and evaluation of teenage pregnancy prevention interventions which may limit their effectiveness in reducing pregnancy rates. To explore the proposed mechanisms of action of preventative efforts, I further examine the theoretical approach taken by the Teenage Pregnancy Strategy and construct a framework for mapping preventative efforts to hypothesised causes of teenage pregnancy.

Setting the stage for my research, I discuss the range of advocated approaches to preventing teenage pregnancy, how these interventions map on to the framework, and the conflicting claims of the efficacy and ethics of their approaches. This is followed by a discussion of hypothesised environmental changes which may have contributed to falling teenage pregnancy rates alongside interventions. Together, I use these interventions and environment changes to construct an initial causal tree, presenting the most prominent hypothesised causes of lowering teen pregnancy rates. I close by setting out the contribution my thesis will make to answering the questions raised from the literature.

2.2 Approach

To critically engage with previous discussion and evidence connected to teenage pregnancy and the Teenage Pregnancy Strategy, I first examine the broader literature surrounding sexual health intervention development. I consult prominent sexual health literature and evidence cited in the discussion of the Strategy. To assess and structure the queries raised regarding intervention development, I additionally consult the MRC guidance on complex interventions and the ‘six steps in quality intervention development’ (6SQuID) framework to determine the use of recommended practices (Wight *et al.*, 2015; Craig *et al.*, 2019). I use causal diagrams to present the issues highlighted within the development and evaluation process. I further examine the Strategy literature and published critiques to determine whether there are similar concerns in its development.

Responding to the range of alternative causes suggested in the discussion of falling rates, I examine the competing hypothesised pathways and reasons for proposed effectiveness or ineffectiveness. Reflecting on the predominance of theories of behaviour change

interventions throughout the literature, I use the 'Capability, Opportunity and Motivation – Behaviour' (COM-B) model (Michie, van Stralen and West, 2011) as a method for structuring the elements which the Strategy and similar interventions may seek to manipulate in order to produce behaviour changes preventing pregnancy.

To summarise the discussion surrounding competing approaches and hypotheses, I extract details from the cited literature and aim to structure these into a diagram of alternative explanations. This takes the form of a 'causal tree', with the trunk of 'reducing pregnancy rates' following on from the range of 'branches' representing alternative causal pathways. The resulting tree presents a summary of prominently proposed causal pathways and is not sensitive to every hypothesised cause. However, I present it as a framework for developing a sensitive, systematic search strategy to detect causes and assess supporting evidence.

Finally, I lay out how my thesis will expand the causal tree through a systematic literature search and synthesis to identify and test plausible causes of observed changes in rates. I then briefly outline how natural experimental methods can address the questions raised by this tree by testing the effects of the Strategy.

2.3 Identifying uncertainties in the development of teenage pregnancy interventions

As noted in Chapter 1, the UK has observed a relatively steep decline in teenage pregnancy rates across England, Scotland and Wales in recent years (Figure 1-3). At recent observations, from peaks in 1998 Scotland had seen a 58% decrease to 19 pregnancies per 1,000 women aged 15-17 in 2016, England a 60% decrease to 19 pregnancies per 1,000 women, and Wales a 62% decrease to 21 pregnancies per 1,000 women (Information Services Division Scotland, 2018a; Office for National Statistics, 2019). The observed falling rate in England has been suggested as evidence of the effectiveness of the Teenage Pregnancy Strategy by several researchers (Toynbee, 2013; Ma, 2016; Skinner and Marino, 2016; Wellings, Palmer, Geary, *et al.*, 2016). However, this explanation has been questioned on the grounds that it cannot account for the very similar patterns in Scotland and Wales shown in Figure 1-3 (Craig, Dundas, *et al.*, 2016). Researchers evaluating the Strategy suggested that the similarity of these patterns may be due to a 'contamination' effect of the Strategy's media campaign, producing similar effects in other UK countries, or simultaneous, effective independent strategies in all three (Wellings, Palmer, Wilkinson, *et al.*, 2016).

Following apparent successes, teenage pregnancy reduction remains a goal in UK health policy (Scottish Government, 2016; Public Health England and Local Government Association, 2018; RSHP, 2019). Alongside a wider scope of improving sexual health and reducing STI risk, lowering pregnancy rates amongst adolescents is presented as a target for improving health and social outcomes, highlighting that such pregnancies are more likely to be unplanned than are adult pregnancies (Public Health England and Local Government Association, 2018).

Recent calls for action to reduce teenage pregnancy rates are in line with widespread agreement in recent decades of the necessity of such interventions. Despite the similar language of these concerns, it is often unclear how subsequently proposed actions understand the problem and seek to address it. Several commentators have suggested that the attention given to teenage pregnancy prevention as a public health priority may incorrectly identify problems to be addressed (Lawlor, Shaw and Johns, 2001; Lawlor and Shaw, 2002; Arai, 2009). Whilst prevention of negative health and social outcomes may be an accepted goal, there is disagreement over the causes of these harms, which causes have greatest scope for change, and how this change might be delivered (Steps 1-3 of '6 Steps for Quality Intervention Development: 6SQuID'; Wight *et al.*, 2015). Additionally, evaluation of public health policies and interventions is a complex task and methods which fail to take account of background trends, overlapping interventions or other events are in danger of producing biased estimates of effect sizes (Craig *et al.*, 2019).

The 'problem' of teenage pregnancy has been conceptualised in diverse ways across recent decades, as tracked by Lisa Arai (2009). The term 'teenager' is understood to refer to adolescents aged 13-19 and centres the age of the woman as the key concern. In Western contexts, calls for public health action in the 1980s and 1990s focused on the outcome of parenthood and the relative unlikelihood of the mother being married at a young age, using the phrase "out-of-wedlock births" to describe these observations (Corcoran, 1999, p. 1). This presents teenage pregnancies as representing moral and social decline, which some have argued is the view expressed in the Conservative Government's Health of the Nation report (Dryfoos, 1990; Adler, 1997; Arai, 2009). Whilst the parenthood is the observed outcome, the root problem is presented as the sexual activity of young, unmarried people.

From focus on the topic of sexual morality, attention changes towards the end of the 20th century to the health and social consequences of teenage motherhood (Arai, 2009). The Social Exclusion Unit (1999) report presents high risks of pregnancy and births to

teenagers as both a result of social disadvantage and a cause of further problems to those who become mothers. Other literature points to the health consequences more tied to the age of the mother than her social circumstances. Highlighted health outcomes include developmental delays to infants born to teenage mothers, risks during birth, and ongoing maternal physical and mental health problems (Irvine *et al.*, 1997; Hodgkinson *et al.*, 2014). Terms such as 'epidemic' used to describe high rates of teenage pregnancies and births reflect the anxiety and urgency in public health proposals to address this problem (Arai, 2009). Early or unwanted parenthood is seen as the root of the problem, with pregnancy a necessary, and preventable, step in the causal pathway.

In proposals of actions to address teenage pregnancy, combinations of these conceptualisations of the problem of such pregnancies are often presented to win support for proposed measures (Arai, 2009, chap. 4). Arai highlights the issues generated by an unclear definition of a problem: where "there is often little distinction between sexual activity, pregnancy and parenthood; they are often attributed to the same factors" (Arai, 2009, p. 65). Lack of clarity of which outcomes should be addressed and therefore which causes should be tackled may lead to ineffective interventions (Wight *et al.*, 2015). Resulting propositions for effective approaches are likely to disagree on whether to prioritise social and cultural change to maintain 'family values', or to promote 'sexual health' through education and contraception provision (Arai, 2009, chap. 4).

Further to this, researchers have questioned whether it is accurate to portray experiences of teenage motherhood as determinants of poor health outcomes (Mcdermott, Graham and Hamilton, 2004). Problems observed amongst teenage mothers may be ameliorable by modifying societal structures rather than preventing pregnancy or parenthood (Arai, 2009, chap. 8). The social and health problems themselves may be produced by confounding through the mothers' socio-economic statuses (Lawlor, Shaw and Johns, 2001; Lawlor and Shaw, 2002).

Potentially further limiting the effectiveness of preventative efforts, the causal pathway linking teenage pregnancy to health and social outcomes may be less clear than is often assumed. Negative outcomes for teenage mothers compared to non-parenting teens are often presented to justify action to prevent pregnancies (e.g. Social Exclusion Unit, 1999, sec. Summary). Presenting these outcomes as preventable through reducing pregnancy rates suggests that all negative outcomes associated with teenage pregnancy are directly caused by it, as shown in Figure 2-1a. Alternatively, Lawlor, Shaw and Johns (2001)

propose that the lower socio-economic status of young women at high risk of pregnancy may account for the majority of their poor ongoing outcomes, regardless of pregnancy, (Figure 2-1b). This confounding may be compounded by a higher likelihood of rapid, repeat pregnancies amongst this group, exposing young people to higher risks of poverty, lower likelihood of completing education and poor maternal and infant health (Lawlor, Shaw and Johns, 2001; Barnet *et al.*, 2010; Aslam *et al.*, 2017). In reviewing several studies seeking to control for background circumstances, by comparing teenage mothers to non-mothering sisters, twins or miscarrying teenagers, Hoffman suggests that teenage childbearing is unlikely to be "a devastating event" (1998, p. 239), with negative outcomes more likely caused by the prior circumstances (Lawlor and Shaw, 2002). Such evidence is likely not strong enough to rule out additional harms as a result of the pregnancy itself (Irvine *et al.*, 1997; Hoffman, 1998; Arai, 2009). This indicates that the relationship in Figure 2-1a plays a role, but that it does not rule out the confounding relationship in Figure 2-1b by accounting for all pathways from antecedents to negative outcomes. The direct effect of pregnancy and the confounding role of the young mother's circumstances in Figure 2-1c may both contribute to health and social harms.

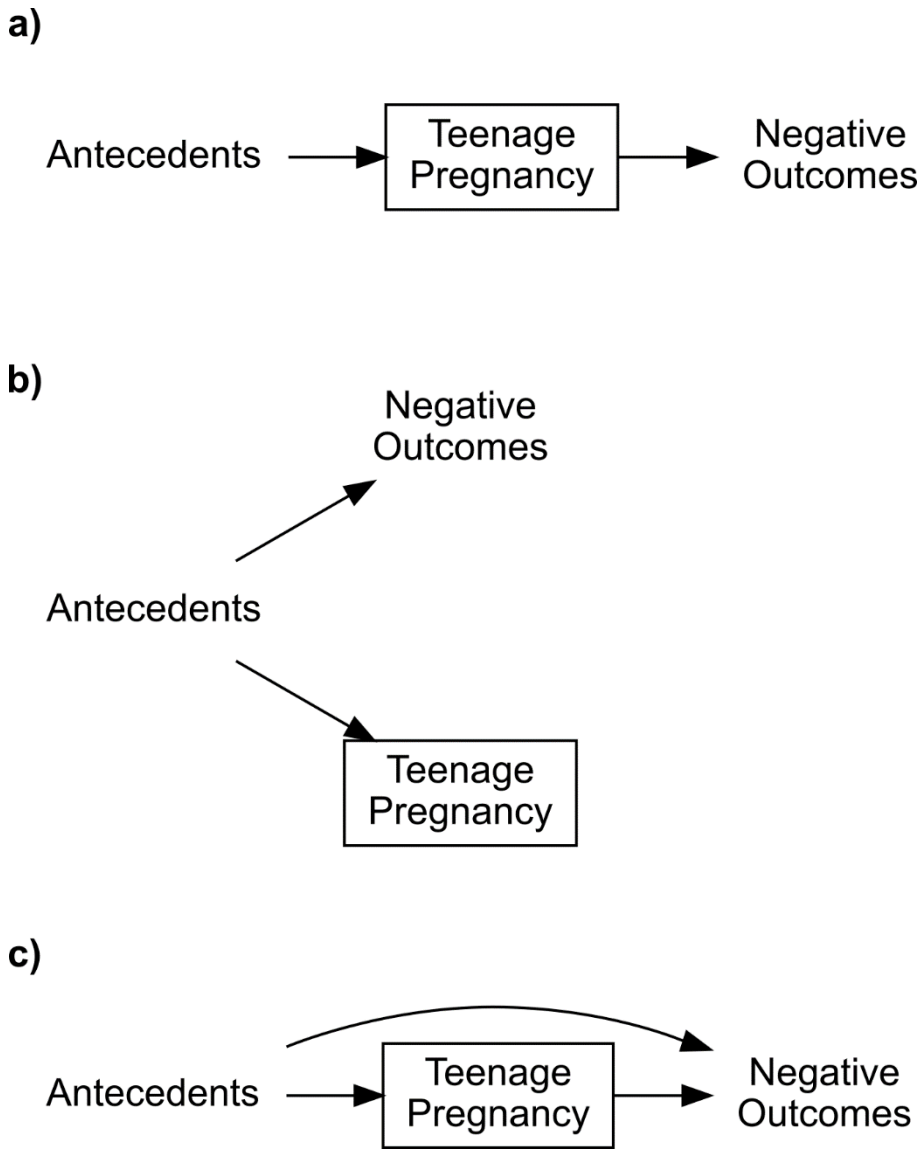

Figure 2-1 – Three simplified hypothesised causal relationships linking Teenage Pregnancy to Negative (health and social) Outcomes. In Figure a), all resulting negative outcomes are presumed to result from the pregnancy, positing that pregnancy prevention would prevent these. Figure b) suggests the circumstances predictive of teenage pregnancy risk (Antecedents) as additionally accounting for all negative health outcomes as proposed by Lawlor, Shaw and Johns (2001). All association between pregnancy and outcomes is due to confounding. Figure c) suggests that pregnancy as a teenager may contribute to negative outcomes but does not explain all of these as confounding is still present.

A further possibility is that negative outcomes for pregnant teenagers are predominantly produced by modifiable structural and cultural factors, unequally impacting teenage parents (Figure 2-2a). This centres the problem not with the teenager's pregnancy and parenthood, but with wider social determinants and society's response to the pregnancy. Unemployment and reduced years of education amongst young mothers for example may be avoidable through support structures, and relatively late engagement with health services amongst pregnant teenagers may also be ameliorated (Social Exclusion Unit, 1999). This proposes that positive health and social outcomes may be achievable even after

the pregnancy event has occurred. Several approaches detailed in the Strategy are aimed at addressing these mediators amongst teenage mothers (Social Exclusion Unit, 1999).

A fuller model of causal factors proposes that the negative health and social outcomes associated with pregnancy are both caused by antecedents such as socio-economic status and are causally linked to pregnancy itself via modifiable mediators (Figure 2-2b). This presents the antecedents, risk of pregnancy and mediators as potential targets for preventative action. However, this raises questions of which causal factors are malleable, which present greatest scope for change, and which changes would be most effective (Wight *et al.*, 2015).

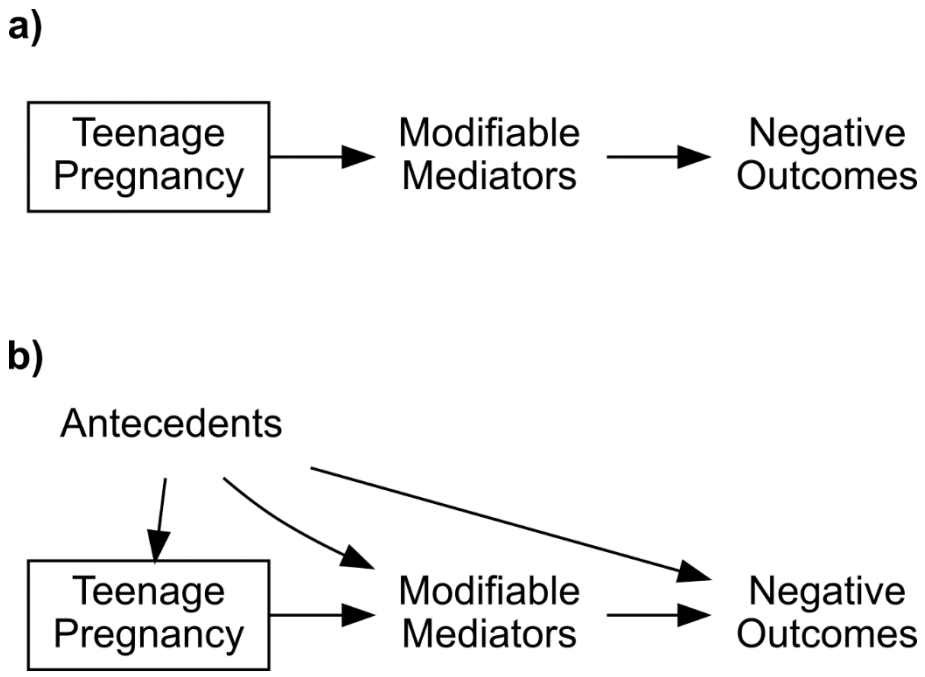

Figure 2-2 - Simplified causal diagrams demonstrating the role which modifiable mediators may play in causation and prevention of negative outcomes in cases of teenage pregnancy. Figure a) presents an alternative explanation of the pathway between pregnancy and negative outcomes and suggests modifiable mediators may be a target for intervention to address harms. Figure b) combines a) with Figure 2-1c) to give an overall hypothesised view of the causal problem and illustrate all potential targets for intervention. In both diagrams, non-modifiable mediators and direct causal links from teenage pregnancy to negative outcomes are omitted for clarity.

A further suggested oversight of some prevention efforts is that not all teenage pregnancy is ‘unintended’ or ‘unwanted’ (Nettle, Coall and Dickins, 2009). A study of under-20 pregnancies in the U.S.A. in 2013 found that, whilst the majority of pregnancies were ‘unintended’, proportions of births which were ‘intended’ pregnancies (a planned or ‘wanted’ pregnancy, or where the woman was unsure whether she wanted to become pregnant) were 32% in under-18s and 42% in women aged 18-19 (Kost, Maddow-Zimet and Arpaia, 2017). Whilst births are a subset of pregnancies, this still presents to

pregnancy prevention efforts a subgroup of women whose choices differ from expectations. Young people may feel positively about the prospect of becoming parents (Bell *et al.*, 2004), which may influence the decisions of teenagers. Wrong assumptions about intentions may contribute to incorrect action to prevent pregnancies. Differing approaches can be expected to have differing effects on each group of teenagers, depending on whether they are motivated to put into action the resources and information given.

2.4 Examining causal pathways affecting pregnancy rates

2.4.1 Critiquing the Strategy's approach

Several of these concerns in the process of developing an intervention were noted to have occurred in development of the Strategy (Arai, 2009). Arguably, by focussing on the high pregnancy rate to quantify the problem, the Strategy may have assumed all negative outcomes were ameliorable primarily by pregnancy prevention. Assumptions of the causes of high rates and how best bring about change may also have hindered effective action. These concerns, discussed below, highlight uncertainties as to the Strategy's likely ability to achieve desired changes.

In the report setting out the Teenage Pregnancy Strategy, the Social exclusion Unit (1999) categorise three suggested factors that contribute to the UK's sustained high rates of teenage pregnancy: 'low expectations', 'ignorance' and 'mixed messages' (Social Exclusion Unit, 1999, p. 7). 'Low expectations' is linked to young people's lack of educational and career prospects, particularly concentrated in low socio-economic status areas, giving less reason to take action to avoid pregnancy. 'Ignorance' combines lack of knowledge about the costs of parenting alongside understanding of the risk of pregnancy in sexual relationships and knowledge about contraception (Social Exclusion Unit, 1999, p. 7). Contraception knowledge is expanded to describe aspects of understanding how to access contraceptives, understanding rights to medical advice, technical capabilities in using contraceptives correctly and awareness of available methods (Social Exclusion Unit, 1999, sec. 7). The final category, 'mixed messages' presents the teenagers' cultural environment as confused and disorientating, as the "adult world" gives two simultaneous but contradictory perspectives: media output promoting sexual expression contrasted with disapproval from parents and institutions. These opposing messages may hinder discussion of sexual health with adults and discourage seeking information from trusted sources

(Social Exclusion Unit, 1999). The relevant influential messages are assumed to have a sexual theme and other cultural messages which may affect sexual behaviour are not considered. “The net result is not less sex, but less protected sex.” (Social Exclusion Unit, 1999, p. 7).

Arai (2009) summarises these three factors as ‘structural’, ‘technical/educational’, and ‘social/cultural’, centring the problem on the influences on teenagers rather than the faults of the teenagers themselves (Arai, 2009, chap. 4). ‘Structural’ influences are mapped to ‘low expectations’, summarising the social circumstances of vulnerable teenagers which present few viable aspirations for adult life and therefore little cost in disrupting further education and employment prospects. She traces out how ‘structural’ influences predominantly affect fertility decisions – whether to become parents or not (Arai, 2009, chap. 4) – thus theoretically mainly affecting pregnancies where a young person is at most ‘ambivalent’ about the risk of becoming pregnant (rather than active in planning to avoid pregnancy; Barrett, Smith and Wellings, 2004). Systemic ‘technical/educational’ deficiencies, mapping to ‘ignorance’ factors, contribute more towards sex and contraception decisions, rather than conscious decisions of whether to become a parent or not prior to intercourse (Arai, 2009, chap. 4; though with some influence over decisions ‘to continue with a pregnancy’). ‘Social/cultural’ influences, perceived as ‘mixed messages’, are likewise assumed to have more influence over sexual behaviour leading to pregnancy, rather than fertility decisions.

Arai (2009) criticises the Strategy report as not clearly identifying which of these interacting and overlapping explanations best explain the high rates of teenage pregnancy, risking focussed action on approaches with minimal impact (Arai, 2009, chap. 4). She suggests this oversight is partially driven by a desire to “avoid controversy” by not aligning with the ideological positions proposing each as a primary cause (Arai, 2009, p. 65). Additionally, there may have been a level of naivety in considering all outcomes (sexual behaviour, pregnancy and parenthood) as one problem and equally influenced by all factors. Ultimately, however, the Teenage Pregnancy Strategy appears to have generally considered the technical/educational factors as the most influential contribution to high rates of pregnancy, simultaneously presenting them as the easiest, most feasible, most economical and most acceptable to remedy (Carabine, 2007; Arai, 2009, chap. 4). This assumption was manifested in the Strategy’s activities as information and education outputs alongside provision of contraception. The focus on technical/educational factors

reflects a view that teen pregnancy is “an educational or medical problem to be solved by increased access to contraception, abortion, and sex education” (Geronimus, 1997, p. 408).

Arai (2009) suggests that the Strategy’s technical/educational focus resulted in the (then observed) slow progress of the Strategy in lowering rates, pointing to evidence that sex education programmes rarely produce long-term effects on behaviours leading to pregnancy (Arai, 2009, chap. 6). This observation of little impact may be partially due to a ‘saturation’ effect, where teenagers are already knowledgeable in accessing and using contraception and have free access to key resources to prevent their pregnancy (Arai, 2009, chap. 6). The Strategy may have failed to take into account the complicated nature of the decision-making of teenagers surrounding pregnancy risk, presuming that deficient knowledge and tools to address ‘ignorance’ was the only factor to be remedied (Barrett and Wellings, 2000; Arai, 2009).

Ingham (2012) responds to Arai’s critique of the Strategy, highlighting that the limited evidence given for the relative ineffectiveness of short-term educational interventions is not sufficient to conclude that there is no value to “longer term and more comprehensive sex and relationships programmes” (Ingham, 2012, p. 256). Additional reasons for education, including the right of the child to information regarding their health and the potential to drive cultural change for greater societal benefits are also presented in support of such technical/educational-focussed interventions (Ingham, 2012). However, it remains unclear to what extent the ‘ignorance’ factor highlighted in the 1999 report contributed to population-level pregnancy rates and whether this had the greatest scope for change.

The misidentification of a factor as a focal point in prevention efforts would hamper potential intervention effects (Wight *et al.*, 2015). Structural influences, giving disadvantaged teenagers “no reason not to get pregnant” (Social Exclusion Unit, 1999, p. 7), and social/cultural influences on teenagers’ behaviour may have had greater influence on producing the higher pregnancy rates observed in the UK. They therefore may have presented targets with more scope for greater effectiveness. It is possible that the Strategy missed opportunities to reduce pregnancy rates by not fully unpacking causal pathways to identify areas for greater potential impact. Conversely, structural and cultural factors are generally less malleable and corresponding interventions are often less acceptable (Frieden, 2010). The focus on technical/educational factors may be justifiable as having identified causes which had greatest scope for change (Wight *et al.*, 2015).

2.4.2 Using the COM-B model to compare explanations of relative effectiveness of competing approaches

As highlighted above in discussion concerning the Strategy, a complex set of factors may be influencing the behaviour of teenagers. Intervention-implemented actions can target several determinants of behaviour in aiming to produce change but may not always be effectively identifying and targeting all or most areas of action with greatest potential for impact. Additionally, wider changes in social determinants may contribute to changes in pregnancy risk amongst a population over time. To re-examine the potential causal pathways, both designed and natural, and to assess the targeting mechanisms of various intervention types, I aimed to use a theoretical framework for understanding precursors of behaviour change. The chosen framework is the 'Capability, Opportunity and Motivation – Behaviour' model (COM-B, Figure 2-3), representing the centre of the Behaviour Change Wheel (Michie, van Stralen and West, 2011; Michie, Atkins and West, 2014), as an approach for understanding the context in which a behaviour or behaviour change occurs. This further links into the intervention functions of the Behaviour Change Wheel (described in Michie, Atkins and West, 2014, Table 2.1) and suggests which forms of intervention are able to prompt the desired change.

I selected the COM-B model as it represents a synthesis of and improvement over previous behaviour change frameworks, it aims to be comprehensive and incorporate diverse approaches and interventions, and its use is related to intervention development, reflecting the Strategy's development (Michie, van Stralen and West, 2011). The model presents the three factors of Capability, Opportunity and Motivation as a summary of interacting elements which are hypothesised as essential for a change in behaviour to occur. The Capabilities element incorporates both the 'psychological' and 'physical' capabilities of an individual, representing both the skills and the tools needed to engage in a behaviour. Opportunity represents the 'physical' and 'social' factors outside of an individual which may provide or restrict the opportunity to engage in a behaviour or prompt the behaviour. The Motivation element incorporates both 'reflective' motivation – how individuals plan behaviour in response to knowledge and values – and 'automatic' motivation – driven by emotion, impulse and habits (Michie, van Stralen and West, 2011).

I use this model to examine the difference in potential outcomes between the theoretically exchangeable subset of adolescent women who would have become pregnant at the stage when rates were high, but not when rates were low. I assume that, during periods of high

pregnancy rates, the behaviour necessary to prevent pregnancy is hindered through blocking of parts of the COM-B pathway. I thus use this model to examine how interventions and changes in social determinants might reduce rates of teenage pregnancy.

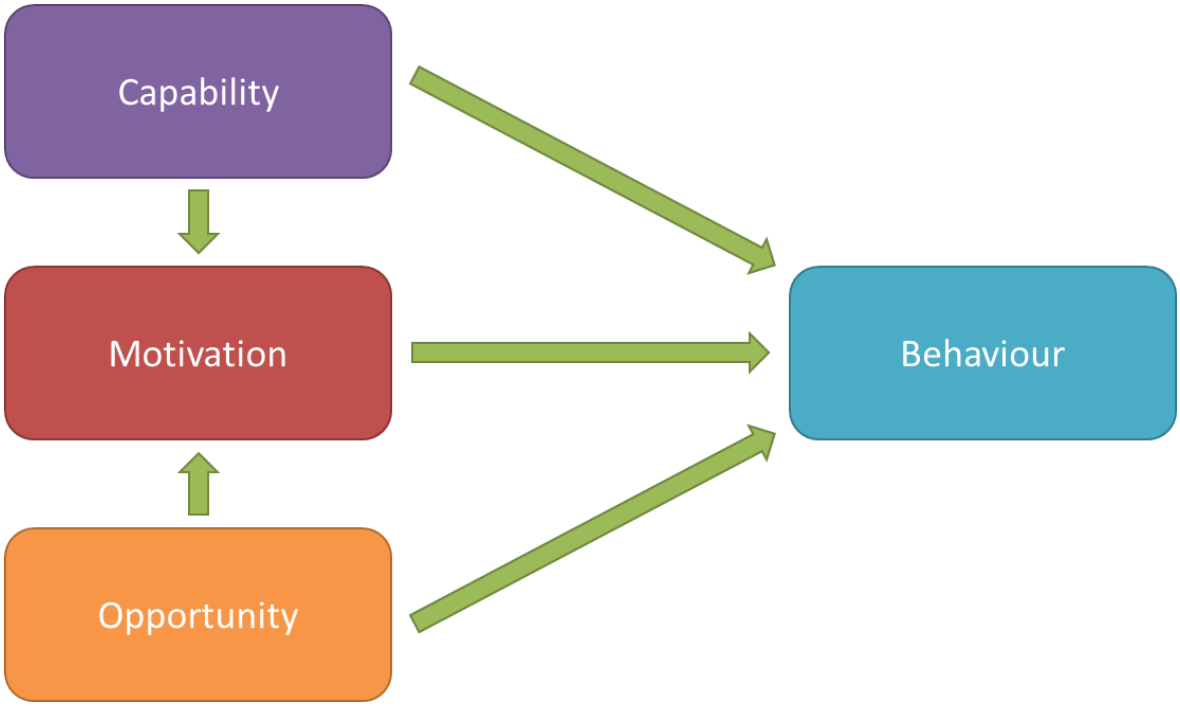

Figure 2-3 – the COM-B (Capability, Opportunity, Motivation - Behaviour) model of behaviour change. (Michie, van Stralen and West, 2011; Michie, Atkins and West, 2014)

In this analysis, all difference in pregnancy risk will be taken as mediated by the mechanisms of frequency of sexual intercourse and frequency and efficacy of contraceptive use. These elements form the Behaviour component and assume that all changes in observed pregnancy rates result from changes in behaviours affecting pregnancy risk. Two limitations of using the COM-B model for this purpose are that it does not account for non-cognitive behaviour change (for example the hypothesis that atmospheric lead exposure causes risky behaviour in teenagers, with reductions in leaded petrol use leading to less sexual risk behaviour over time; Reyes, 2015; Paton, 2020), nor does it allow for changes in infertility across a population over time. If either or both of these elements are the main causes of the observed changes then the approach of examining behaviour change through the COM-B model will not be able to fully explain the phenomenon. In this thesis I will assume that the majority of observed effects are attributable to reflective or habitual behaviour change. This pathway reflects the assumptions of policies and interventions tackling teenage pregnancy as discussed above.

To unpack the potential causes of change, I map elements of teenage behaviour discussed above to the three precursors of behaviour change in the pregnancy risk COM-B model. A

clear element highlighted is the understanding strategies to reduce the risks of sexual behaviours and obtained skills for using these to reduce the likelihood of pregnancy. I represent this element using the Capabilities element, predominantly through the psychological capabilities component (with a limited involvement of the 'physical' capabilities element, for example through condom application skills; Michie, Atkins and West, 2014, Table 1.3). The psychological capability of a person to avoid pregnancy might be given through an intervention providing information and understanding, or through training in emotional, cognitive or behavioural skills (Michie, van Stralen and West, 2011; Table 2). An adolescent's learned capability to negotiate sexual intercourse and contraceptive use with a partner, or to select and use more suitable contraceptive methods (independent of the accessibility of such contraceptives), would be elements of improved Capability able to produce behaviour change lowering pregnancy risk.

I use the Opportunity component in the pregnancy risk COM-B model to represent both the availability of the resources used in pregnancy-prevention methods (i.e., contraceptives; 'physical opportunity') as well as any reduction or restriction of opportunities to engage in sexual activity ('social opportunity'). As both elements are external to the individual, changes in an adolescent's environment are necessary to influence the Opportunity component. The physical opportunity to avoid pregnancy (within a consensual sexual relationship) would primarily be provided through interventions giving an adolescent access to effective contraceptives (Michie, van Stralen and West, 2011; Table 2). Reduction in social opportunity to engage in risky sexual behaviours could be a result of changes to the physical or social environments of teenagers, for example a change in behaviour of young men being less likely to pressure women into engaging in risky sexual behaviour, or through rules reducing the opportunity to engage in sexual risk behaviour, for example by extending compulsory education (Michie, van Stralen and West, 2011 Table 2).

In addressing Motivation in behaviour change affecting pregnancy risk, I connect this to a primary driver represented by 'reflective motivation', whereby an adolescent perceives the cost of pregnancy, desires not to become pregnant, and is therefore given greater reason to take action to reduce risk. Additionally, behaviours which are preventive of pregnancy may be driven by 'automatic motivation' – habits or learned anxieties over risky sexual behaviours which may contribute to the consistency of precautions taken. Both types of motivation may be modifiable by targeted interventions (Michie, van Stralen and West, 2011). Alongside interventions, changing social influences may contribute to reflective

motivation to avoid pregnancy, by presenting pregnancy as a more undesirable event, or automatic motivation to use contraceptives, by normalising contraceptive use and eliciting anxieties over unprotected sex.

In using this framework, I aim to draw upon its key strength in classifying mechanisms for change. I use it to theorise the explanations for instances of success and failure of various approaches to changing teenage pregnancy.

Using the COM-B model, Arai's proposed factors of 'structural', 'technical/educational' and 'social/cultural' contributions to high rates of teenage pregnancy can be reframed to highlight problems and pathways with scope for change (Arai, 2009, chap. 4). In an environment where not becoming pregnant is a possible, accessible behavioural option, each of these factors increasing the risk of pregnancy can be mapped as interfering with the pathways from Capability, Opportunity or Motivation to relevant Behaviour. This model is presented in Figure 2-4. Using this approach, I aim to examine the targets of prevention efforts, to clarify how they are assumed to produce change (Wight *et al.*, 2015), and to allow critical thinking of possible successes and failures in implementation of each intervention. I also aim to account for environmental changes, examined or unexamined, which may need to be considered as potential future avenues of action, or confounders to address in intervention evaluations.

Structural factors as conceived by Arai (2009) encompass primarily the 'low expectations' causal pathways from the Social Exclusion Unit report (1999). 'Structural' as used here describes chiefly the effects of an individual's prospects of unemployment and no access to further education giving "no reason not to get pregnant" (p. 7), rather than the macro-level influences of socio-economic, environmental and cultural influences underlying a wider set of determinants (as proposed by Dahlgren and Whitehead, 1991). These individual-level effects of structural factors would primarily block the Motivation pathway as it would reduce motivation to act to avoid pregnancy. Teenagers who are ambivalent about becoming pregnant would primarily fall into this category, whereby lack of motivation to avoid pregnancy is a causal factor overwhelming all other prevention methods targeting them. Differential access to education and employment may also impact on Opportunity, as teenagers engaging in either may have fewer physical opportunities to engage in sexual risk behaviours.

The second category of social/cultural factors may similarly block the Motivation pathway, but more by increasing the cost of action (e.g., declining sex in contrast to perceived norms, possessing and using contraceptives at the risk of stigma). Additionally, it may impact on social opportunity, as changing sexual norms contribute to changes in the availability of sexual relationships to the young person, independent of their motivation to pursue them. This would be depicted as blocking the Opportunity pathway.

The final category of technical/educational factors is linked to both the Capability pathway and Opportunity pathway, hypothesising that pregnancies thus affected are primarily caused by lack of skills and knowledge in preventing pregnancy, or lack of access to contraceptives. Effective action in these areas may primarily compose educational and informational interventions, raising awareness and giving skills and knowledge to take effective action, coupled with improving access to contraception, or encouraging use of contraceptive options which are more effective (or less prone to human error).

These categories further allow hypothesising of non-intervention changes which act on the same pathways. These will be classed as changes to the teenagers' 'environment' to highlight the causal nature of these influences whilst distinguishing them from designed interventions with pregnancy reduction goals in mind. Such changes might work through cultural influences, determining sexual activity and contraceptive use, or changing perceptions of pregnancy and motivating action. Further, changing sources of information might address technical/educational needs through pathways other than sexual education programmes.

Visualising the problem and proposed solutions using the COM-B model highlights that effective action requires all three pathways to function. In this manner it takes on the character of a three-legged stool – removal of one of the COM legs would result in the stool 'falling over', and the desired behaviour change not occurring. Evidence-based and well-funded action targeting one branch of the model and producing encouraging intermediate results (such as greater knowledge, or greater contraception distribution), may yet fail to produce the intended outcomes through the Behaviour component if other aspects are missing. A teenager with contraceptive access and skills but desiring to become pregnant will become no less at risk through further contraceptive provision. Another teenager who wants to avoid pregnancy and knows what action to take, but does not have access to contraceptives, will similarly remain at a higher risk. A third teenager who is provided with contraceptive pills but is unaware of the relative ineffectiveness of sporadic

use will also still face risks. These three scenarios demonstrate the hypothesised problems of blocking the Motivation, Opportunity and Capability pathways respectively.

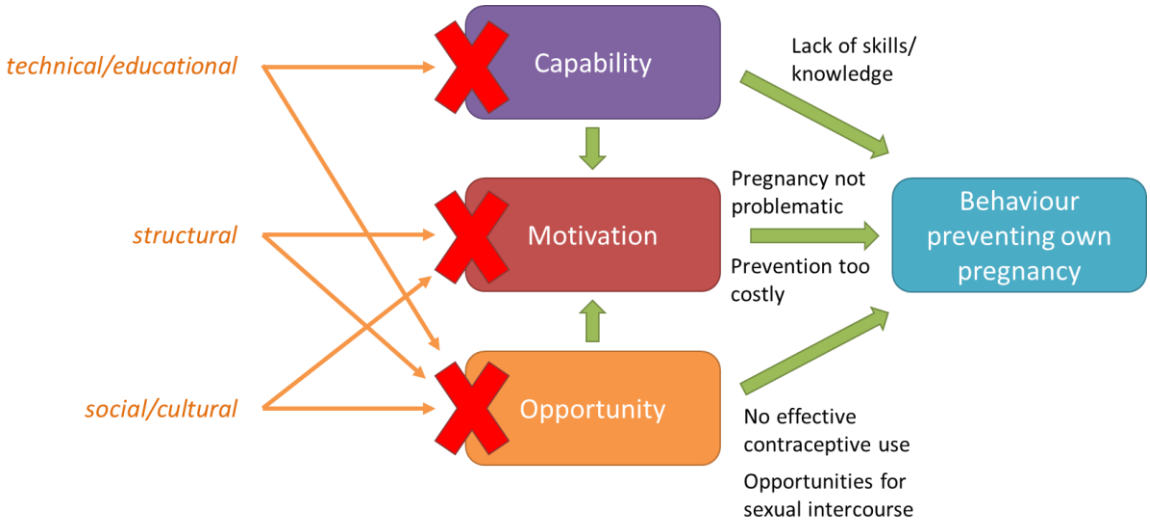

Figure 2-4 – COM-B model as a framework for understanding behaviours leading to pregnancy. Developed from Michie, van Stralen and West (2011). Arai's (2009) categorisations of the explanations of high rates are shown on the left and joined to the pathways they theoretically inhibit (orange arrows) 'Structural' factors denotes the individual-level effects of low education and employment aspirations produced by circumstances. 'Technical/educational' factors encompass inhibited access to tools and knowledge to prevent unwanted pregnancy. 'Social/cultural' factors represent messages motivating and demotivating action, disseminated through media and peers. Pathways to effective behaviour (green arrows left to right) are each blocked by example theorised problems which interventions seek to modify.

2.5 Proposed explanations of reducing pregnancy rates

Amongst the range of approaches proposed for reducing teenage pregnancy rates, there is little consensus on effective methods of achieving this. To compare these methods, I will use the COM-B framework developed above to explore how they differentially aim to address different theorised deficiencies in the pathways underlying a teenager's sexual behaviour (Figure 2-4). Differing intervention approaches may target various pathways in the COM-B model based on ideologies or assumptions about the area of greatest deficiency or most appropriate intervention (Arai, 2009, chap. 4; Collumbien and Macdowall, 2012; Forrest and Wellings, 2012). Here I compare intervention types and other influences on sexual health behaviour of teenagers outlined by Wellings, Mitchell and Collumbien (2012, sec. 4), the World Health Organisation Regional Office for Europe (2004) and the Social Exclusion Unit at Strategy launch (Social Exclusion Unit, 1999). I aim to draw out commonly highlighted causal pathways of proposed and implemented interventions and environment changes which are hypothesised to contribute to lowering teenage pregnancy rates. To these previously highlighted interventions, I add other suggested factors raised in research which are hypothesised as causative, but not designed as interventions. Highlighting shared behavioural mediators, I aim to construct an initial

causal model that encompasses these prominent hypothesised pathways. I conceptualise this as a ‘casual tree’, depicting hypothesised causes as branches connected to the trunk of reduced pregnancy rates, to further inform inquiry into evidence supporting each hypothesis.

2.5.1 Interventions aimed at lowering teenage pregnancy rates

Sexual Health and Relationships Education

A prominent proposal in pregnancy prevention is mandatory or semi-mandatory (i.e., with parental opt-out) school-based Sex and Relationships Education (SRE). Several researchers have proposed increased SRE as a foundational approach in reaching all youth and affecting change in risk behaviours (Kirby *et al.*, 1994; Kane and Wellings, 1999; Cheesbrough, Ingham and Massey, 2002; WHO Regional Office for Europe, 2004). Others argue that these approaches rely on behaviour change theories which situate sexual activity as a rational choice based on beliefs about the risks and benefits of engaging in sex (Oettinger, 1999). In this framework, adolescent pregnancy is understood as a predominantly unwanted outcome, resulting from lack of knowledge concerning the risk of becoming pregnant, the cost of pregnancy and childbearing, or strategies to reduce risk (Geronimus, 1997; Arai, 2009, chap. 4). Methods of action in such interventions predominantly seek to improve knowledge and skills about the risks involved in sexual activity, teaching strategies to reduce these risks. This approach targets the teenagers’ Capability for avoiding pregnancy (Figure 2-4). Additionally, they may also include aspects of Motivation as they aim to portray pregnancy as an undesirable event, delaying sexual activity or prompting greater precaution.

The relationship between SRE and pregnancy is complex, as shown in Figure 2-5. Several authors suggest that SRE interventions have unintended negative consequences. They may normalise earlier sexual engagement, increase awareness amongst sexually inactive teenagers and increase sexual activity at a rate greater than their successes in preventing resulting pregnancies (Oettinger, 1999; Paton, Bullivant and Soto, 2020). However, other studies comparing interventions with controls conclude there is no increase in sexual activity (Cheesbrough, Ingham and Massey, 2002; Kirby, 2002a). Additionally, they may be ineffective in achieving positive change (Arai, 2009). Whilst some evidence points to positive change from specific interventions (Kirby *et al.*, 1994; Kirby, 2002a), several studies have concluded that there is little certainty that increased SRE results in lowered rates of teenage pregnancy (Henderson *et al.*, 2007; Elliott *et al.*, 2013; Mason-Jones *et al.*,

2016). When observed across several countries, SRE appears ineffective in reducing national pregnancy rates and may increase them in some cases (Paton, Bullivant and Soto, 2020).

A recent study examining SRE provision in some schools in Scotland found that even though the intervention when evaluated was no more effective than treatment-as-usual, later induced reductions in SRE were associated with reduced knowledge in contraception use (Nixon, Wight and Henderson, no date; Henderson *et al.*, 2007). It is possible that a baseline provision of SRE is a key aspect in raising awareness of sexual health risks amongst teenagers, but in modern contexts such awareness has reached a saturation point. Greater provision of and expenditure on SRE may not produce consistent scaled improvements in teenage pregnancy prevention (Elliott *et al.*, 2013).

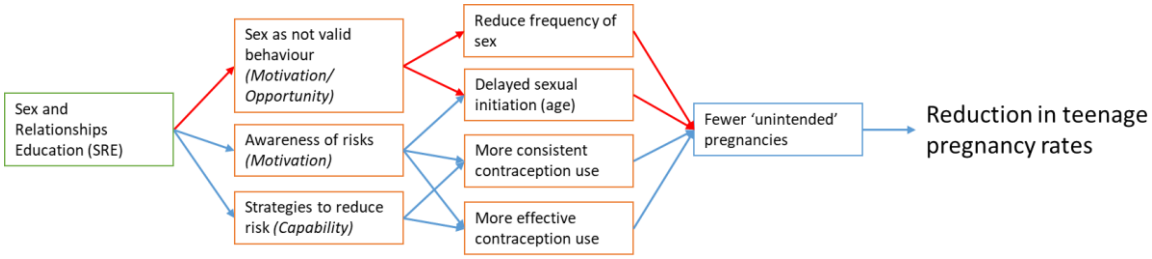

Figure 2-5 – Hypothesised causal pathway linking Sex and Relationships Education to reducing teenage pregnancy rates. Blue arrows denote hypothesised positive causes. Red arrows denote hypothesised negative causes or causes with conflicting claims of direction of effect. A full legend is provided in Figure 2-14.

Abstinence education

Abstinence-based education within schools has drawn much interest and critique. Discussion of this approach is centred predominantly within the U.S.A., with most studies conducted there (Underhill, Operario and Montgomery, 2007a; Underhill, Montgomery and Operario, 2008). Abstinence-only education can be defined as having "as its exclusive purpose, teaching the social, psychological, and health gains to be realized by abstaining from sexual activity" (Santelli *et al.*, 2006; Table 1). An alternative form, abstinence-plus programmes, present abstinence as best practice whilst additionally promoting strategies to reduce risk for those who do engage in sexual activity (Underhill, Montgomery and Operario, 2008). Often, abstinence-only and abstinence-plus interventions are built from religious foundations. Such programmes advocate for abstinence until marriage and occasionally include expressly faith-based messages and participant interaction to position this as a moral action (Kirby, 2002a). These aims, if successful, would result in reduced frequency of sex and delayed sexual initiation amongst adolescents (Figure 2-6). This model can be seen as targeting both the Motivation element, portraying sexual intercourse

as undesirable and pregnancy as a more alarming outcome, and the Opportunity element by reducing perceptions of sex as a morally available option and reducing approaches from partners who have received similar messages.

The abstinence approach has drawn many strong criticisms of the ethics of prescribing a moral stance, the poor practices often associated with its implementation, and the lack of evidence of its effectiveness (Santelli *et al.*, 2006; Ott and Santelli, 2007). There appears little clear evidence of its effectiveness in reducing teenage pregnancy, but similarly no consistent indications of adverse effects directly attributable to abstinence-based teaching (Kirby, 2002a; Underhill, Operario and Montgomery, 2007a; Underhill, Montgomery and Operario, 2008). However, there may be an opportunity cost as effective programmes providing other strategies to reduce risk (Capability) which contravene the abstinence-only messages are withheld from adolescents. This exclusion may limit provision of knowledge and resources allowing consistent and effective contraception use (Santelli *et al.*, 2006; Ott and Santelli, 2007). The aims of abstinence programmes alongside hypothesised negative side effects are shown in Figure 2-6.

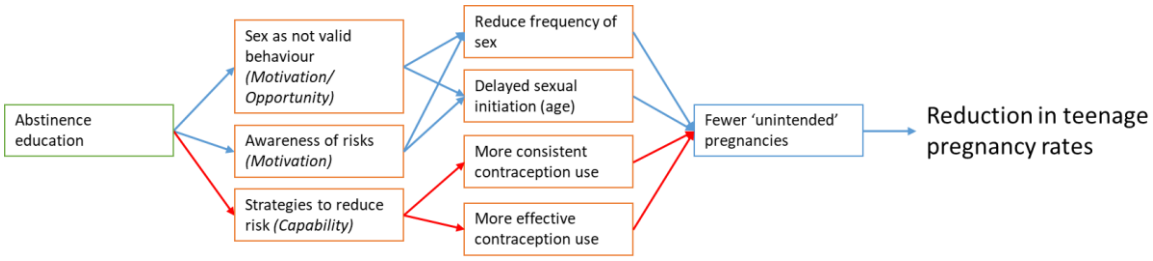

Figure 2-6 – Hypothesised causal pathway linking Abstinence Education (including Abstinence-only and Abstinence-plus approaches) to reducing teenage pregnancy rates.

Clinic-based interventions

Clinic-based interventions present a further opportunity to engage with teenagers, provide effective contraception and promote behaviour change (Figure 2-7; Social Exclusion Unit, 1999; French and Mitchell, 2012). Linking into education through school-based clinics can provide discreet access to contraception and advice to teenagers as needs arise. Outside of education, dedicated sexual health clinics and provision of services and counselling at local health clinics may also promote such behaviour change (French and Mitchell, 2012). These approaches may improve Opportunities to avoid pregnancy through provision of more effective contraceptives, Capabilities through taught skills, and Motivation through raised awareness of risks. Evidence on the effectiveness of each at a population or community-level is mixed (Goesling *et al.*, 2014). Targeted one-to-one counselling and

'individualised' clinic services may be effective for those reached but may be too narrow focussed to realistically affect rates in a local population (Goesling *et al.*, 2014).

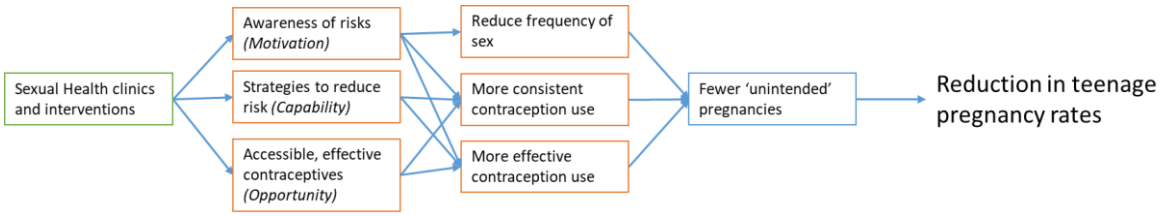

Figure 2-7 – Hypothesised causal pathways linking clinic-based interventions to reducing teenage pregnancy rates

Improved contraception access

Improving access to or changing methods of contraception are approaches which aim to reduce pregnancies without expressly reducing frequency of or delaying onset of sexual activity (Glasier and Wellings, 2012). Condom distribution through clinics, schools and community outreach is a commonly implemented intervention (Social Exclusion Unit, 1999). The proposing of this approach arguably assumes that the most influential modifiable factor in behaviour change is the lack of access to the contraceptives which the teenager has already the skills and motivation to use (the lower arrow blocking 'Opportunity' in Figure 2-4). A key criticism of this type of intervention is that high levels of contraceptive uptake may present an impression of a successful intervention but be mostly driven by substitution of already available and used contraception, thereby not increasing contraceptive use overall (Kirby, 2001). They may also contribute to normalisation of early teenage sexual activity whilst aiming to increase contraceptive use or give access to more effective contraceptive methods (Figure 2-8).

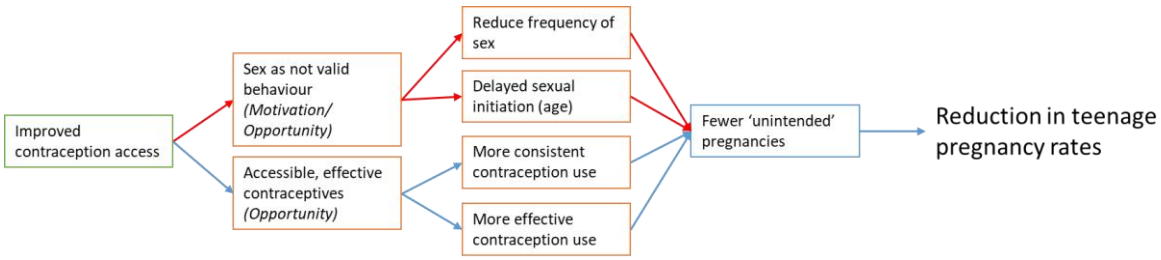

Figure 2-8 – Hypothesised causal pathway linking Contraception access interventions to reducing teenage pregnancy rates.

Better contraceptive technologies

Changing contraceptive technologies is an additional potential area for impact on pregnancy rates. Increased use of non-condom contraceptives such as hormonal contraceptive pills, post-coital contraceptives, and Long-Acting Reversible Contraceptives (LARCs) may prompt more consistent contraceptive use and more effective contraceptive

use (Figure 2-9). Additionally, unlike condoms, they do not rely on male skill and motivation to be effective, and their use may be more driven by young women for whom unplanned pregnancy is a bigger concern (Connolly *et al.*, 2014; Wellings, Palmer, Geary, *et al.*, 2016). Both pathways would primarily affect the Opportunity component, both through allowing greater practical access to contraceptives and making contraceptive use more reliable. Potential critiques of such methods are that they may also prompt 'replacement' of already-effective condom use, leaving both partners more at risk of STIs, or give users confidence beyond their real-world effectiveness (Girma and Paton, 2015).

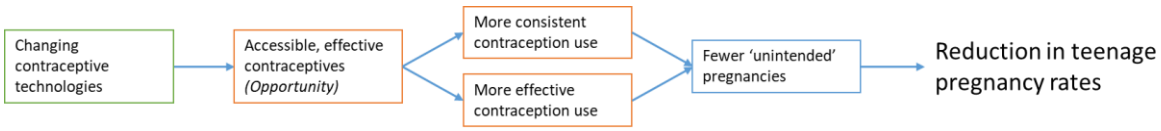

Figure 2-9 – Hypothesised causal pathway linking Improving contraceptive technologies to reducing teenage pregnancy rates.

2.5.2 Environmental factors

Beyond the effects of interventions expressly targeting teenage pregnancies, there are a number of other policy changes and changes to teenagers' environment across the period of observed decline which are hypothesised to have contributed to these falls.

Increased time in full-time education

More time spent in full-time education during adolescence is associated with lower risk of teenage pregnancy. Increases in years spent in education across populations likely contributes as a cause of lower pregnancy rates (Wilson, 2014; Girma and Paton, 2015). These increases may be structural changes produced by policy, and possibly enacted with teenage pregnancy prevention as a goal. Here I consider them as environment changes to allow for non-designed changes with similar effects. This pathway may target Opportunity or Motivation pathways. Policies mandating more years of schooling or increasing access to further education amongst young people at greater risk of pregnancy may have an 'incarceration effect' on physical and social opportunity. This effect hypothesises that as teenagers are forced to spend more time in environments where sexual risk behaviours are not possible (in this case in educational premises under adult supervision), opportunities to engage in sexually risky behaviour are reduced (Wilson, 2014). This causal pathway predominantly targets the Opportunity component in changing behaviour, delaying onset of sexual activity and reducing the frequency of sexual contact (Figure 2-10). Alternatively, improved educational attainment in early years, which may be influenced by or influence educational and career aspirations, is shown to strongly affect pregnancy risk

(Girma and Paton, 2015). This pathway could be seen as predominantly targeting the Motivation element of behaviour change as it increases the perceived cost of pregnancy and parenthood, prompting preventative actions (Figure 2-10). Change in education aspiration may be driven by policy designed to achieve this, or by non-designed changes in culture from other national or global events.

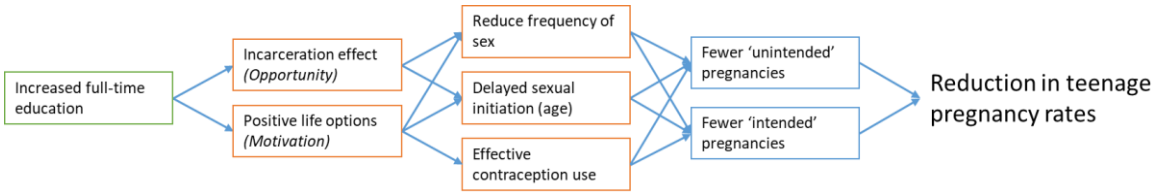

Figure 2-10 – Hypothesised causal pathway linking cultural changes in increased full-time education amongst teenagers to reducing teenage pregnancy rates.

'Later parenting' culture

A broader cultural change in perceptions of parenting may have influenced decisions and behaviour of teenagers in this period. Relative to 1990, pregnancy rates across England and Wales increased for women aged 30 and over and decreased for younger women (Figure 2-11). Teenagers ('Under 20', 'Under 18' and 'Under 16' age groups) saw the greatest declines (Office for National Statistics, 2019). These changes could represent broader changing social norms in timings of sexual initiation and pregnancy. As a driver of this trend, changes in how society perceives teenage pregnancy may have been driven by media presentations of teenage parents (Kearney and Levine, 2015b), motivating greater precautions by increasing perceived cost. Teenage pregnancies which are not 'unplanned' may also have been affected, as other opportunities for career, education and social mobility become more attractive options for young women than the positive aspects of parenthood. A further possibility is that advances in the empowerment of young women may affect pregnancy timings through greater capability to refuse sexual intimacy, to negotiate contraceptive use, and to determine timings of pregnancy (Mitchell *et al.*, 2021). These cultural changes may additionally present in individuals not as conscious or reasoned behaviour change but as habitual delay of initiation of intercourse or use of contraceptives, influenced by such cultural trends. Increasing life expectancy may also lead to conscious decisions to delay pregnancies for longer amongst more women across the observed period (Nettle, 2011).

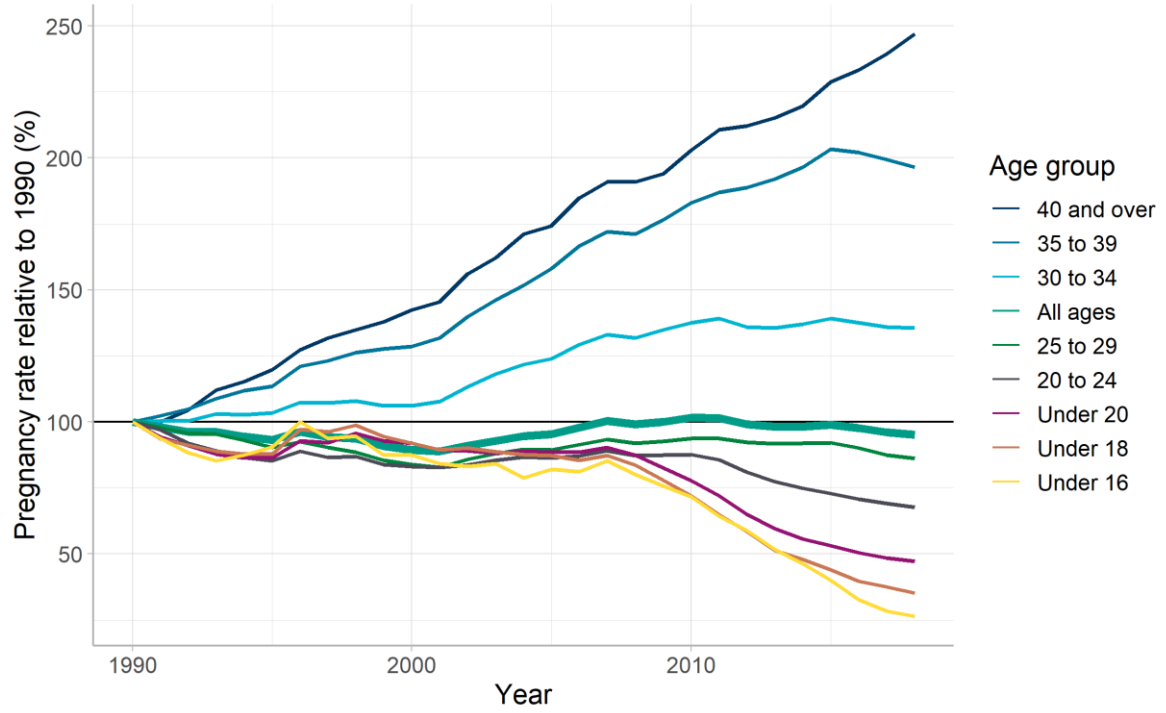

Figure 2-11 – Relative changes in age-specific conception rates in England and Wales, 1990 to 2017 (Office for National Statistics, 2019, fig. 2)

Two suggested pathways mapping how changing cultural perceptions of parenting ages may reduce teenage pregnancy rates are shown in Figure 2-12. Presentation of positive life options for female teenagers may prompt behaviour change. Alongside this positive message, negative perceptions of teenage parenthood may also restrict social opportunity by increasing perceptions of it as not a valid option. Together these pathways would result in decreases of both planned and unplanned pregnancies.

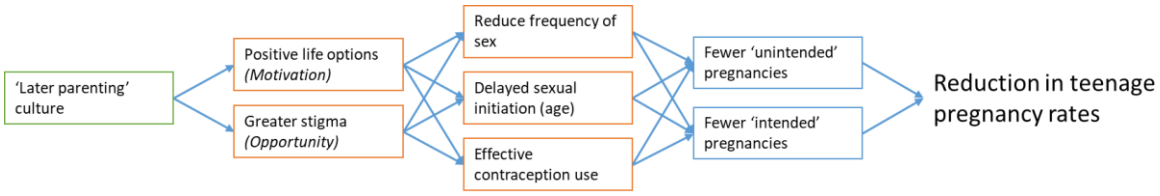

Figure 2-12 – Hypothesised causal pathway linking shifting cultural attitudes towards parenting and age to reducing teenage pregnancy rates.

Digital media use

A further suggested influence on changing culture and behaviour through the early 2000s is increasing digital media access and use. Expansion of broadband access through 1999 to 2007, for example, was found to account for a small proportion of the reduction in teenage births in the USA (Guldi and Herbst, 2017). Less formal hypotheses link the introduction of smartphones in 2007 to the observed sharp falls in pregnancy rates across subsequent years (Twenge, 2017; Triggle, 2018). Both initial causes may affect teenagers' behaviour

through one or more pathways in the COM-B model. Greater access to global media output may have influenced teenagers' life and career goals and aspirations, as attractive roles are more prominently featured in their homes and through media devices (The Harris Poll, 2019). Perceived positive and achievable life options may motivate action to avoid risks which threaten these. Alternatively, an 'incarceration effect' may prompt teenagers to engage with socialising and alternative hobbies at home and thereby shield teenagers from opportunities to engage in risky sexual behaviour (Guldi and Herbst, 2017).

A further mechanism by which smart phone ownership and internet access may impact on pregnancy risk is through sexual health information access (Patterson, Hilton, *et al.*, 2019). Adolescent internet access is continually expanding and use of smartphone and tablet devices is increasingly common (Ofcom, 2020). Searching for health information, and sexual health information in particular, may improve adolescents' psychological capabilities through increased knowledge (Borzekowski and Rickert, 2001; Michie, Atkins and West, 2014, Table 1.3; Patterson, Hilton, *et al.*, 2019; Patterson, McDaid, *et al.*, 2019). These three hypothesised pathways are combined in Figure 2-13.

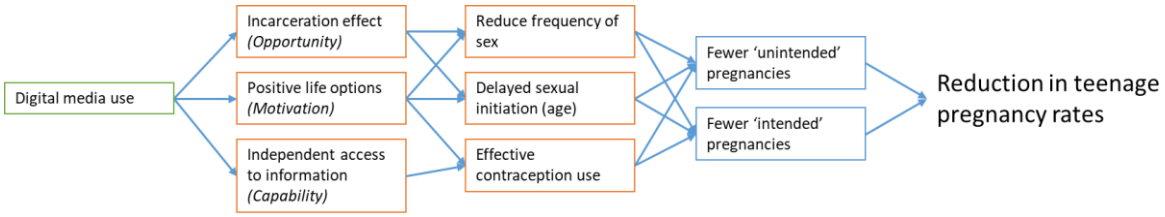

Figure 2-13 – Hypothesised causal pathway linking increasing adolescent digital media use to reducing teenage pregnancy rates.

2.5.3 Drawing a causal tree from identified hypothesised pathways

To compare these suggested causes of reducing teenage pregnancy rates I constructed an initial 'causal tree' – a combination of hypothesised pathways mapping common mechanisms and the range of exposure types. Figure 2-14 below shows the five proposed intervention types as five 'branches' of the tree in the upper part of the diagram and the three environment changes in the lower part. All suggested mediators are shown in the middle section of the diagram. Highlighted potential mechanisms of action forming the pathways are drawn across the diagram with red arrows signifying hypothesised negative impacts (increasing risk of pregnancy) and blue arrows positive impacts (reducing risk). Outcomes are split into 'intended' and 'unintended' pregnancies to highlight issues identified above in how to frame and target teenage pregnancy (p.22). All outcomes are

combined in the single measurement of ‘Reduction in Teenage Pregnancy Rates’ to note the observed overall rate changes which encompass both types of outcome.

Notably, though highlighted by previous commentators that not all adolescent pregnancy is unplanned or unwanted (Nettle, Coall and Dickins, 2009; Nettle, 2011), I found no explicitly declared aims of interventions to address and reduce intended pregnancies. The upper section of the causal tree omits the causal step of “Fewer ‘intended’ pregnancies”. As argued by Arai (2009), the Teenage Pregnancy Strategy considers young women with planned pregnancies under the umbrella cause of “low expectations” (Social Exclusion Unit, 1999). However, the focus of the action on technical/educational aspects does not appear to address this factor.

The causal tree presents potentially testable pathways to assess each branch’s contribution to the UK’s fall in teenage pregnancy rates. Other published literature, not identified here, may present data sufficient to eliminate branches with little likelihood of effect or add branches not previously considered.

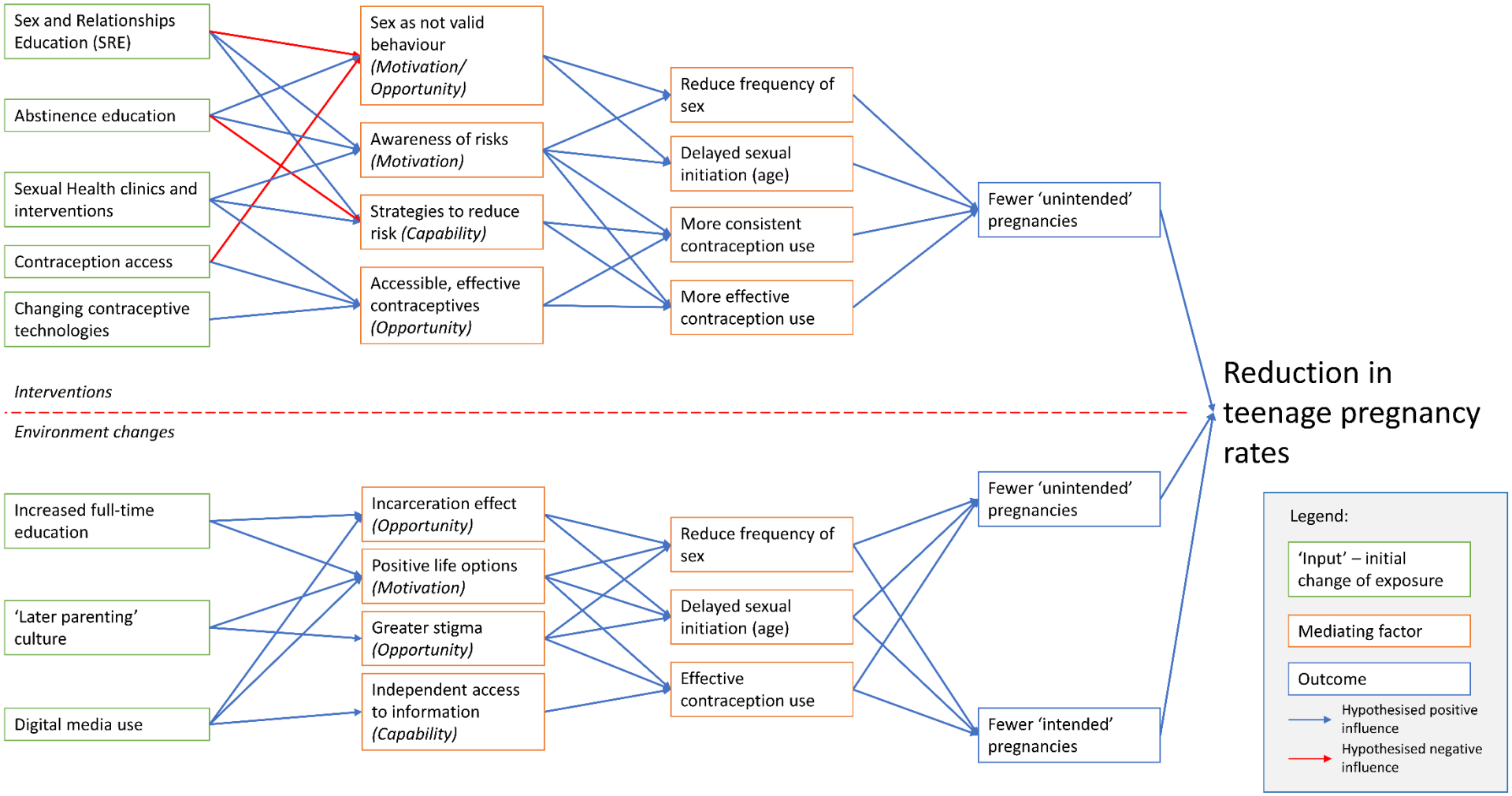

Figure 2-14 – Causal tree depicting the range of hypothesised causes for the observed reduction in teenage pregnancy rates. Constructed from pathways identified in the literature in sections 2.5.1 and 2.5.2.

2.6 Addressing methodological gaps

As discussed above, the Strategy, alongside other factors, may have contributed to falling rates in England from 1999 onwards, but substantial uncertainties remain. The causal tree in Figure 2-14 highlights a range of potential causes of reducing teenage pregnancy rates. The Strategy's prominent components of improved sex education in schools, improved clinic access and distribution of contraceptives would map most clearly onto the three corresponding branches in the upper part of the tree. Investigation of the evidence of the effects of each of these could provide a clearer picture of the Strategy's contribution. Simultaneously, it could indicate whether other interventions were more effective, or if observed successes were down to other, non-designed changes. However, as the initial tree is drawn only from prominent literature and the development and critique of the Strategy, it is potentially blind to evidence supporting stronger causal pathways and may produce bias in informing further research. Little literature was found addressing macro-level changes. The changing economic climate including the banking crash in 2008 may have impacted pregnancy rates, positively or negatively, during the Strategy period. Other political changes enacted by the New Labour government and common to the whole UK may have contributed also. A further limitation to this tree is the missing elements of parental engagement and a national media campaign – considered core Strategy elements (Social Exclusion Unit, 1999). I found insufficient evidence to assess these in the prominent literature I used to construct the tree. More substantial treatment of these elements may be published elsewhere.

My first approach in this thesis is to test and expand the initial causal tree from a sensitive synthesis of existing literature. I aim to achieve this by systematically reviewing the literature to detect and summarise the evidence testing the causal effects of each branch on teenage pregnancy rates. These results are reported in Chapter 4, highlighting key branches for further investigation. In this analysis, I combine logic modelling and causal analysis with systematic review methodology by searching the literature for data able to test the hypothesised pathways in the tree. I produce an updated causal tree in Chapter 4, summarising published knowledge and highlighting key hypotheses for further testing.

In testing the causal impact on observed changes in UK teenage pregnancy rates, several challenges are presented. As discussed in Chapter 1, changes in outcomes must be measured in comparison with an 'exchangeable' control, either identical to the exposed group ('marginal' or 'unconditional exchangeability') or able to be corrected to account for

differences (‘conditional exchageability’; Craig *et al.*, 2017; Hernán and Robins, 2020). The ‘gold standard’ approach of conducting randomised controlled trials may not always be acceptable or possible, particularly in the cases of non-designed environment changes and other policies highlighted above not specifically targeting pregnancy as an outcome (Guyatt, Oxman, Kunz, *et al.*, 2008; Craig *et al.*, 2012, 2019; Abadie and Cattaneo, 2018). As introduced in Chapter 1, natural experimental methods provide one possibility for overcoming these challenges.

The Strategy, as an intervention with a well-defined start point and to which only teenagers in England were exposed, may then allow a natural experimental evaluation. The Strategy, as a coherent and holistic intervention, incorporated several elements which can be mapped to branches in Figure 2-14. Previously highlighted items which map to already identified tree branches are: improved sex and relationships education, contraception access, and improved sexual health clinic access and use (Hadley, Ingham and Chandra-Mouli, 2016). Testing the Strategy would allow assessment of the effects of these combined approaches on teenage pregnancy rates and give an estimate of the extent to which the Strategy contributed to falling rates. I present my findings on two methods for testing the Strategy’s contribution to falling rates in Chapter 5 and 6.

In this thesis I aim to contribute to our understanding of the possible causes of the observed falling rates of teenage pregnancy across the UK from 1999 to most recent observations. I aim to achieve this both by synthesising the range of hypothesised causes with evidence presented for their effects on pregnancy rates and by testing a prominent hypothesised cause in an evaluation of the Strategy. Together, these analyses aim to explore the value of such research and methods in giving a clearer understanding of the observed data and allowing us to test the causal impact of the Strategy.

Chapter 3 Methods

3.1 Overview of chapter

This chapter lays out the methods for the two parts of my analyses. In the opening section, I set out how the methods were chosen as robust approaches to achieving the aims of the thesis and the novel contribution they bring to the area of teenage pregnancy prevention and public health policy evaluation.

In the second section, I describe the stages in the conducting of the systematic review to uncover plausible causes of the observed decrease in UK teenage pregnancies in published literature. I use the approach of reviewing reviews to detect the most prominent hypothesised causes amongst an anticipated vast body of literature. I assess the quality of reviews and the strength of evidence for a causal effect of each exposure on a population. Finally, I synthesise the evidence in an updated causal tree of plausible contributors to the falling UK rates.

In the following sections, I describe the natural experimental methods I use to evaluate the effects of England's Teenage Pregnancy Strategy on teenage pregnancy rates. I test the effectiveness of the Strategy using two methods and datasets, aiming to use these two approaches to triangulate the results and give a robust estimate of effect sizes. In my first analysis, I choose Scotland and Wales as comparators given their similarity to England in other characteristics which may affect teenage pregnancy rates. I use interrupted time series methods to compare each country with England across the implementation period and up to most recent observations. To account for potential contamination among neighbouring UK countries, in my second analysis I compare birth and pregnancy rates in England with those of a wider pool of potential control countries using synthetic control methods.

3.2 Selecting methods and defining research questions

As I demonstrated in Chapter 2, concerns over 'high' rates of pregnancy amongst teenagers have prompted a range of actions in high-income countries. Observations of falling pregnancy rates have been cited as indications of the effectiveness of these approaches (Hadley, Ingham and Chandra-Mouli, 2016; Ma, 2016; Wellings, Palmer, Geary, et al., 2016). Changes in the social, cultural or economic circumstances (summarised as changes in 'environmental factors'; section 2.5.2 above) of the teenagers across the last three

decades can also be hypothesised to contribute to decreasing rates. In summarising the prominent literature in section 2.5 I produced Figure 2-14, hypothesising the overlapping mechanisms of action and the trade-offs of suggested positive and negative effects. This represents the initial causal tree, allowing a brief but incomplete overview of the complexity of potentially interacting hypothesised causes. This tree can inform further inquiry, but cannot reliably summarise the evidence for all plausible hypothesised causes.

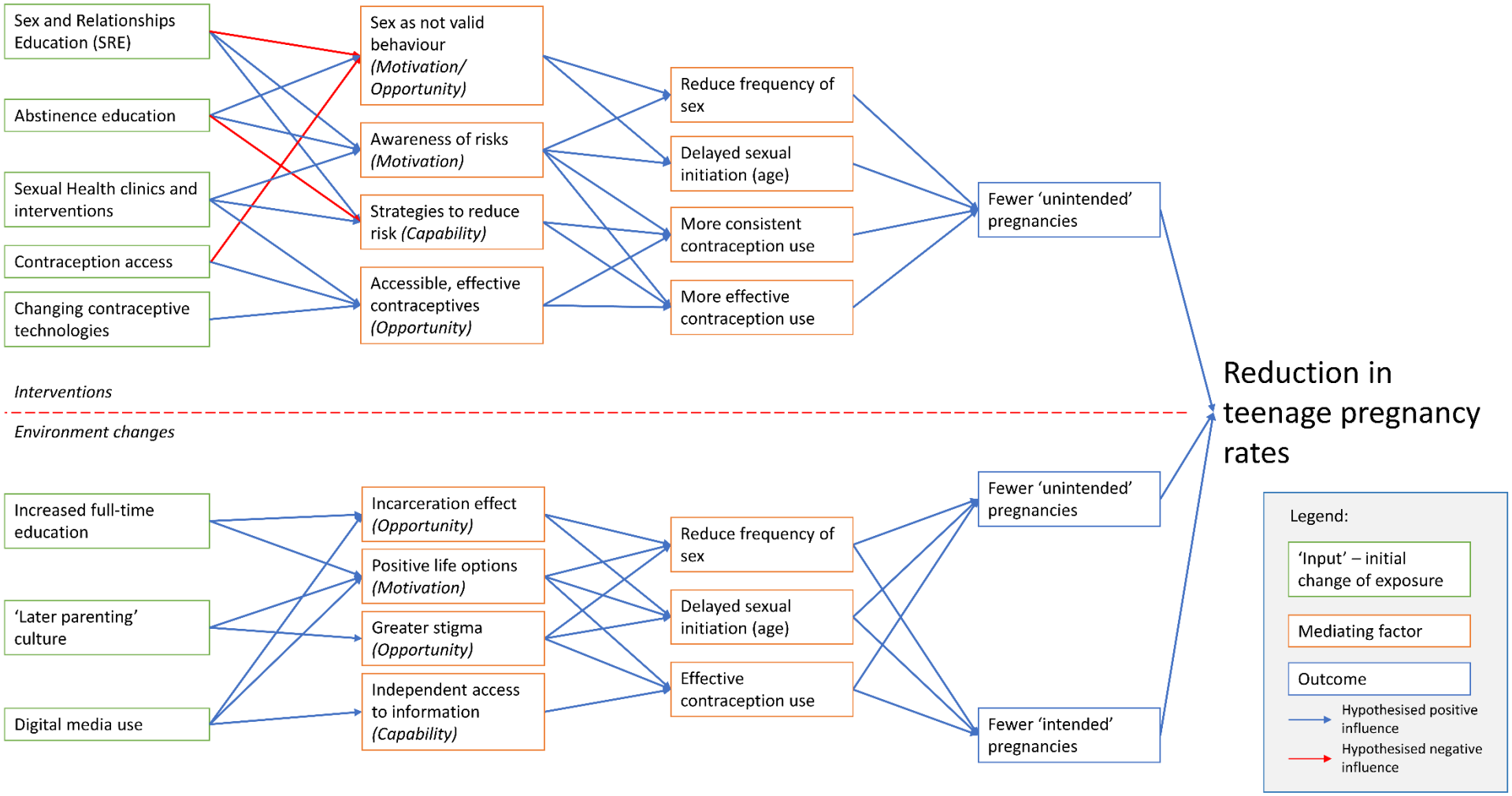

Figure 3-1 – The initial causal tree showing a range of potential causes of reducing teenage pregnancy rates and overlapping pathways and mediators. This was constructed from a scoping review of the literature Chapter 2 to identify initial pathways for investigation (reproduced from Figure 2-14).

As stated in Chapter 1, in this thesis I aim to explore the UK's falling teenage pregnancy rates, first by investigating plausible hypothesised causes of the falling rates and second by simultaneously evaluating the contribution of the Teenage Pregnancy Strategy to England's rates. I aim to both review the literature for the best evidence presented for all hypothesised causes and to use robust methods to test the Strategy's effectiveness.

In investigating the wide range of causes likely to have contributed, I use systematic review methods to assess published evidence for the presence or absence of an effect of each cause. I begin with a search strategy informed by the initial causal tree of Figure 2-14, sensitive to detect other interventions and exposures which I had not previously encountered. I aim to include evidence which tested the causal influence of an exposure on a real or representative sample population across the time-period corresponding with observed decreases in the UK. Anticipating that plausible causes may have been tested outside of the UK, I include a selection of countries similar to the UK to be sensitive to evidence assessing the scalability of exposures to whole populations. To appraise evidence of causation, I use modified Bradford Hill criteria (Hill, 1965) to query the reviews for evidence of the effectiveness of a hypothesised cause, and evidence of the exposure of a population to this cause within the time period of interest. I construct an updated causal tree which allows for evidence-based exclusion of ineffective branches and inclusion of branches with supporting evidence or branches requiring further testing. I present this evidence-informed causal tree as a summary of prominent hypothesised causes of falling teenage pregnancy rates to inform further investigation.

In testing the Strategy's contributions to the declining rates, I approach this aim as an outcome evaluation of the intervention and sought to use appropriate policy evaluation methods. I select natural experiment methods as a set of approaches to estimating the effects of such an intervention. This allows for evaluation of an already implemented intervention, where re-creating the exposed and unexposed conditions in a trial setting was not possible.

Natural experimental methods provide several ways to emulate a comparison study between exchangeable groups. They can thus assess exposures as hypothesised causes, but have specific susceptibilities to bias and may depend on data availability for their use (Craig *et al.*, 2012, 2017). To address these concerns, I aimed to use multiple methods and datasets to compare estimates with different potential biases. Comparison of consistent or conflicting results across study types can give a cautious assessment of certainty of effects

from approaches which are susceptible to bias (Craig *et al.*, 2019). This allows me to assess the certainty of results by triangulating findings across several methods and datasets (Munafò and Davey Smith, 2018). As noted earlier (section 2.2), several concurrent events may have contributed to observed effects. I sought to account for such events by using similar countries to estimate a counterfactual for England. My two selected methods are interrupted time series and synthetic control methods (Abadie, Diamond and Hainmueller, 2010; Lopez Bernal, Cummins and Gasparrini, 2017).

Interrupted time series methods are widely used for evaluating public health interventions (Craig *et al.*, 2017; Lopez Bernal, Cummins and Gasparrini, 2017). These models use statistical methods to estimate the effects of an intervention or exposure by comparing outcome trends in time series data before and after an implementation cut-off point. This correction aims to account for underlying trends producing a difference in outcome in the event of a null intervention effect and can incorporate controls to attempt to adjust for other population changes over time (Craig *et al.*, 2017; Lopez Bernal, Cummins and Gasparrini, 2018).

In selecting controls for interrupted time series analyses, I limit the set of comparison countries to other nations within the UK (excluding Northern Ireland due to incomplete abortion data). All UK populations would be expected to undergo many similar exposures and may give close estimates of the counterfactual outcomes of Strategy-exposed England. My observations in Chapter 1 (section 1.5) of the trends similar to England followed by Scotland and Wales in the pre-Strategy period suggest that these make good predictors of England.

In response to previous observations of similarities in rates across the UK, researchers discussed concerns about using Scotland and Wales as comparators to identify the effects of the English Strategy (Craig, Dundas, *et al.*, 2016; Craig, Leyland, *et al.*, 2016; Wellings, Palmer, Wilkinson, *et al.*, 2016). A potential limitation may be that they may have been exposed to elements of the Strategy, such as the media campaign (Craig, Dundas, *et al.*, 2016; Craig, Leyland, *et al.*, 2016). Additionally, both Scotland and Wales saw some changes in SRE provision and other published reports across the Strategy period (Teenage Pregnancy Strategy Evaluation, 2005; Wilkinson *et al.*, 2006; Wellings, Palmer, Geary, *et al.*, 2016). These are summarised in Table 3-1. A national strategy with a teenage pregnancy reduction focus was unique to England, whilst both Wales and Scotland both stated reduction in teenage pregnancy rates as a goal in published reports. Scotland saw

some changes to sex education provision with the launch of the SHARE programme in 1993 which underwent a randomised controlled trial, showing little effect (Wight *et al.*, 2002; Henderson *et al.*, 2007). A later programme, 'Healthy Respect', was piloted in the Lothian region from 2000. Although Wilkinson *et al.*, (2006) report that "it was thought to have a major contribution to make to the whole of Scotland", later evaluations measuring sexual behaviour outcomes against unexposed comparison areas concluded that Healthy Respect had no impact on sexual risk behaviours or pregnancy rates (Elliott *et al.*, 2010, 2013). Wales saw the publication of a report in 2000 and an action plan in 2001, recommending similar approaches to the Strategy (Table 3-1). The £1 million recurrent funding allocated to this was split between national and local implementation, and with the wider remit of all sexual health outcomes across the wider population (The National Assembly for Wales, 2000; Wilkinson *et al.*, 2006). This in turn represented a lower level of total investment in sexual health in Wales than in England (Williams, 2006). Further recommendations to improve school education and a condom distribution programme followed this. Policy changes which addressed areas other than teenage pregnancy, not considered here, may have impacted pregnancy rates differentially across the UK countries during the Strategy period. Whilst not reproducing the same holistic and joined-up approach of the Strategy, the interventions seen in Scotland and Wales may reduce the estimation of Strategy effects if they produced trend changes in control countries.

A further limitation of the interrupted time series methods is that they rely on testing linear trend changes in a pre-specified 'impact model' (Lopez Bernal, Cummins and Gasparrini, 2017) – the Strategy with ongoing changes in its implementation may not have followed a linear pattern (Hadley, Ingham and Chandra-Mouli, 2016).

Policy actions	England	Scotland	Wales
National teenage pregnancy strategy	The teenage pregnancy strategy (TPS) – 1999	Towards a Healthier Scotland* – 1999; Respect and Responsibility* – 2005	A strategic framework for promoting sexual health in Wales* – 2000; Post-consultation action plan* – 2001
Funding allocated to teenage pregnancy prevention	TPS – 1999: est. £280m over 10 years	-	Post-consultation action plan – 2001: £1m recurrent funding from 2001
Local implementation strategies	TPS – 1999: "local implementation fund for integrated and innovative programmes"	-	Local health boards developed action plans – 2001
Changes to sex and relationships education	TPS – 1999: "new guidance for schools on sex and relationships education"	SHARE – 1993; Healthy Respect (Lothian-based pilot) – 2000/2005	Post-consultation action plan – 2001: "Ensure that all young people in Wales receive effective education about sex and relationships"; Sex and Relationships Education in Schools – 2002: new guidance
Improved clinic access	TPS – 1999: "new health service standards for effective and responsible contraceptive advice and treatment for young people"	-	Post-consultation action plan – 2001: "Ensure that all sexually active people in Wales have access to good quality sexual-health advice and services"
Contraceptive provision	TPS – 1999: "Better prevention… including access to contraception"	-	Condom Grants Initiative – 2002
Media campaign	TPS – 1999: "a national publicity campaign"	-	-

Table 3-1 – Comparison of national sexual health and teenage pregnancy policies enacted across UK countries from 1993-2005. * Reductions in teenage pregnancy not exclusive focus. Sources: Social Exclusion Unit (1999); Wilkinson *et al*., (2006).

To account for limitations due to data, in my second analysis I aim to compare birth and pregnancy rates in England with those of a wider pool of control countries. Selecting appropriate controls outside of the UK presents additional challenges. As highlighted in the development of the Strategy, and in previous evaluations, many similar high-income countries had substantially lower birth rates in the period immediately before the Strategy launch in 1999 (Social Exclusion Unit, 1999; Teenage Pregnancy Strategy Evaluation, 2005). Validity of comparison studies for causal inference rely on the exchangeability of exposed and unexposed units (Hernán and Robins, 2020). Manually selecting comparator countries by assumed likeness to England presents the danger of biased results and would not produce a reliable estimate of effect size. Additionally, countries outside of the UK may be subject to a different set of exposures and not be good predictors of England's changing rates of pregnancy and birth.

To address these concerns and conduct a comparative analysis with a larger set of countries, I selected the synthetic control method for a second empirical test of the Strategy's effects (Bouttell *et al.*, 2018). The synthetic control algorithm uses a data-driven approach to construct a control from a set of comparator units by taking a weighted mean of pre-intervention observations to match the outcome trend of the exposed unit most closely. These same weights are applied to post-intervention observations across unexposed countries to calculate an assumed trend in the absence of the intervention (Abadie, Diamond and Hainmueller, 2010, 2011; Craig *et al.*, 2017). Additionally, as this method does not apply a linear model, it can allow prediction and assessment of non-linear trends in the pre- and post-implementation periods. Several sensitivity analyses allow for testing of the stability and confidence of synthetic control models (Abadie, Diamond and Hainmueller, 2010).

Using the methods outlined above, the six research questions (RQs) that this thesis will aim to answer are:

RQ1: What changes in interventions, culture, policy and environment are hypothesised to cause reductions in teenage pregnancy in the UK and similar countries?

RQ2: For which causes is there evidence of their effects in reducing pregnancy, and their exposure across populations?

RQ3: Which causes have been robustly examined but have failed to produce evidence of their effects across a population, indicating the absence of an effect?

RQ4: What are the core elements to include in a causal tree which may have contributed to reductions in teenage pregnancy rates in the UK?

RQ5: Did implementation of the Teenage Pregnancy Strategy cause a change in level or trend in yearly teenage pregnancy rates in England, compared with other UK countries?

RQ6: Did implementation of the Teenage Pregnancy Strategy reduce yearly teenage pregnancy rates in England, compared with a wider set of control countries?

3.2.1 Open access code and data

Throughout my analyses, I aimed to follow the principles of open science in order to ensure my research was accessible, transparent, reproducible and accessible (UKRI, 2020). For statistical analyses I used R as a free, open source tool (R Core Team, 2020) and used publicly available data. I published open access protocols for each stage of my research before conducting them. I continuously published the code used in my analyses and where possible aimed to present findings as interactive online apps to allow reproduction of my analyses without prior R knowledge.

Code for conducting analyses and producing graphs is shared at the GitHub repositories **github.com/andrewbaxter439/teen_preg_review_docs** and **github.com/andrewbaxter439/teen-preg-project**. All apps, interactive plots and interactive data tables are collected online at **phd.andybaxter.me**. A full table of URLs and DOIs is included in Appendix A.

3.3 Systematic review of reviews of hypothesised causes of reducing teenage pregnancy rates

3.3.1 Developing the approach of the review

The protocol for the systematic review is published at PROSPERO (ID: **CRD42018085759**).

I restricted searches to published reviews of relevant studies to address the expectedly large volume of literature of all types on the topic. I anticipated that well designed reviews

would produce the strongest evidence testing hypotheses and more manageable presentations of findings across a diverse range of study topics. I also expected that the most important hypothesised causes would be addressed in such reviews.

I considered data presented in reviews as evidence of the effects of an exposure if they reported data suitable for testing causal relationships. As the primary aim of this review was to test for the presence or absence of an effect of a hypothesised cause, rather than to estimate effect sizes, I extracted data to assess the strength of evidence and not effect estimates. I assessed the strength of the evidence presented in each review using markers of review quality, consistency and strength of evidence reported, and indicators of population-wide exposure. To assess review quality, I developed an adapted critical appraisal tool to assign scores weighted by domains of bias critical to this review's questions (section 3.3.5.1 below).

I grouped hypothesised causes by similarity of the elements used in their implementation or the context of their delivery (school, community, etc.). The overall evidence of the effects attributable to each hypothesis was assessed to evaluate the plausibility of its effects on recent teenage pregnancy trends. I used this data to update the causal tree, removing branches where the evidence indicated no effect, adding branches representing newly uncovered hypotheses with some evidence of effect, and highlighting branches presenting strong evidence of effect.

3.3.2 Aim and research questions

In this review, I aimed to assess the evidence for previously identified hypothesised causes of reducing teenage pregnancy rates and detect and assess other causes presented in the literature. I discovered several hypothesised causes in a non-systematic review of the literature in Chapter 2, which I presented as an 'initial' tree to inform a sensitive and systematic search. The review aimed to identify evidence that would answer research questions 1 to 4 (section 3.2 above):

RQ1: What changes in interventions, culture, policy and environment are hypothesised to cause reductions in teenage pregnancy in the UK and similar countries?

RQ2: For which causes is there evidence of their effects in reducing pregnancy, and their exposure across populations?

RQ3: Which causes have been robustly examined but have failed to produce evidence of their effects across a population, indicating the absence of an effect?

RQ4: What are the core elements to include in a causal tree which may have contributed to reductions in teenage pregnancy rates in the UK?

In this analysis, the 'core elements' of the causal tree will be defined as causal pathways with evidence of effectiveness, alongside hypothesised causal pathways with insufficient evidence to test effectiveness (retained for further investigation), excluding branches with evidence indicating their lack of effect on pregnancy rates.

3.3.3 Search strategy

I searched four databases (Medline, Embase, Scopus and Cochrane Database of Systematic Reviews) for reviews of adolescent/teenage pregnancy and prevention or reduction in rates, published in 1990 or later. Where reviews of reviews were found in search results, I manually added all cited reviews to screening. Full search strategies, detailing key words, regular expressions and filtering terms used, are presented in Appendix B.

3.3.4 Screening

3.3.4.1 Inclusion criteria

Population

The focus of this review is on female adolescents, aged 13-19 years living in high-income countries (HICs). I aimed to allow for generalisability across similar countries and account for the potential of interventions being tested in one country but implemented in another. I selected countries for inclusion from two sources. I initially selected the 31 countries included in the Euro-Peristat network (Euro-Peristat, 2018) representing European countries. To these added four English speaking HICs: Australia, Canada, New Zealand and the United States of America (see Box 3-1).

Austria	Germany	Norway
Australia	Greece	Poland
Belgium	Hungary	Portugal
Bulgaria	Iceland	Romania
Canada	Ireland	Slovakia
Croatia	Italy	Slovenia
Cyprus	Latvia	Spain
Czech Rep.	Lithuania	Sweden
Denmark	Luxembourg	Switzerland
Estonia	Malta	UK
Finland	Netherlands	United States of America
France	New Zealand	

Box 3-1 – A list of the 35 countries to be included in this review. This list was compiled from Euro-Peristat countries plus other English speaking high-income countries.

Intervention/Exposure

I included reviews evaluating interventions which aimed to reduce teenage pregnancies or to address associated sexual health risk behaviours. Reviews which tested the effects of social, cultural or economic changes on pregnancies and risks were also included. These include reviews evaluating interventions which are designed to be applied to a population or a vulnerable subset of a population and are transferable across countries.

Control

Reviews reporting a comparison between exposed and unexposed populations were included. Reviews examining the effects of other broader cultural changes on pregnancy rates, measured between populations or across time, were included.

Outcome

The primary outcome was rates of pregnancies or births to women aged under 20. Other behaviours directly affecting pregnancy risk – sexual activity and contraceptive use – were also included.

Study type

Only reviews of other published literature were included.

3.3.4.2 Exclusion criteria

I excluded reviews which only reported outcomes such as abortions, maternal and infant health, or other outcomes after conception. Reviews which evaluated associations of hypothesised determinants of pregnancy risk across a population, with no analysis of changes of exposure over time, were not included. Additionally, I did not include reviews which only measured changed attitudes or knowledge as an intervention outcome, rather than behaviour change. I treated these changes as mechanisms of behaviour change rather than outcomes, as they may be poor predictors of actual risk of pregnancy (Wight *et al.*, 2015).

3.3.5 Data extraction and synthesis

3.3.5.1 Quality assessment and risk of bias

To grade the quality of each review, I developed a series of four domains to grade as 'high' or 'low' risk of bias. These were adapted from ROBIS tool for assessing risk of bias in systematic reviews and supplemented with questions from the AMSTAR 2 tool for systematic reviews of studies evaluating healthcare interventions (Whiting *et al.*, 2016; Shea *et al.*, 2017). As this review included non-systematic literature reviews, I chose to adapt and expand on the assessment questions used in each tool. I added a further criterion to the domains presented in the ROBIS and AMSTAR 2 tools to distinguish between systematic and non-systematic literature search and synthesis methods. Further, as my aim in this review was to assess the overall evidence for the presence or absence of an effect of an exposure, I assigned several domains less weight than as presented in the existing tools. Within the AMSTAR 2 domains, item 14 assessing explanations of observed heterogeneity was not relevant for this review as I intended to note and discuss heterogeneity across reviews rather than use it to assess review quality. Questions assessing smaller contributions to bias and error in data extraction and effect estimates would be unlikely to fully mask a present effect or conclude the effectiveness of an ineffective exposure (items 5, 6, 10-12, 15 and 16; Shea *et al.*, 2017).

In assessing the domains to be measured in light of areas highlighted in existing tools, I anticipated five Concerns (C1-C5) where evidence presented in my included reviews may give biased assessments of causal links:

C1. Reviewers have purposefully selected from the pool of evidence addressing the hypothesis, producing a biased sample of studies.

C2. Non-systematic or poorly executed literature searching is likely to have missed out key publications.

C3. Reviewers have synthesised evidence with a non-systematic method of weighting studies, potentially producing biased results.

C4. Included studies which have aimed to control for bias have not done so sufficiently and risk of bias remains high.

C5. Included studies have not addressed confounding or established directions of cause; associations between exposure and outcome may have been produced by factors other than the causal effect of exposure.

The four domains listed in Table 3-2 were used to query all included reviews, answering 'yes' or 'no' to each domain query.

Domains of risk of bias	**Low quality evidence if:**	**Mapping to existing tools**		**Assessing anticipated concerns (above)**
		ROBIS domains	*AMSTAR 2 criteria*	
1 – Was it systematic, defining objectives and setting appropriate study eligibility criteria?	Concerns regarding selectiveness of included studies or unclear objectives	1	1, 2, 4	C1
2 – Were search and screening methods appropriate, clear and not likely to have missed key studies?	Concerns regarding methods used to identify and/or select studies	2	3, 7, 8	C2
3 – Were findings synthesised systematically, presenting all available evidence?	Concerns regarding the synthesis and findings	4		C3
4 – Was risk of bias and confounding adequately assessed, and were results presented to take account of this?	Concerns regarding strength of evidence for causation	3, 4	9, 13	C4 and C5

Table 3-2 – Domains of risk of bias and questions used to assess key concerns in included literature. Where domains of existing tools were incorporated into domains for this review, these are noted in the 3rd and 4th column. Several domains listed in these tools were not used as I judged them not relevant for this review.

I assigned each review a score of 0 to 4 for how many queries were answered positively for addressing risk of bias. High overall quality reviews, with zero or one areas of concern (scoring 3 or 4) were used in synthesis as primary sources of evidence. Low quality (scoring less than 3) reviews were referred to in cases of few or no high-quality reviews and in overall comparisons of conflicting findings to assess potential sources of heterogeneity.

3.3.5.2 Extracting hypotheses and coding reviews for evidence of efficacy and exposure

I constructed a table with a row for each hypothesised cause identified in Figure 3-1 and separate columns for each distinct cause addressed in at least one review. New rows were added as additional suggested causal pathways were identified. The finalised table was published online (**phd.andybaxter.me/review_data**). In this table I aimed to detect and to list all hypothesised causes, blind to strength of evidence provided, to answer RQ1: 'What changes in interventions, culture, policy and environment are hypothesised to cause reductions in teenage pregnancy in the UK and similar countries?'.

I aimed to assess the impact of each hypothesised cause as a combination of measurements of its effects and the extent of the exposure of whole target populations. To construct questions to assess the evidence presented I used the RE-AIM framework (Glasgow, Vogt and Boles, 1999). This tool queries five domains of intervention implementation – reach, efficacy, adoption, implementation and maintenance – to expand the evaluation of interventions beyond only examining effectiveness. I regrouped the criteria for an intervention's hypothesised impact as 'efficacy' ('Efficacy' or 'Effectiveness' from the RE-AIM framework) and 'exposure' ('Reach', 'Adoption', 'Implementation' and 'Maintenance'). I coded each review's evidence for the causes they examined as presenting either no evidence of efficacy or exposure, or strong evidence (evidence of both efficacy and exposure) or weak evidence (evidence for one but not the other) of a positive or negative effect on a population's pregnancy rates.

Evidence for efficacy was appraised using a set of guidelines for testing causation developed by Howick, Glasziou and Aronson, (2009), as modified from the Bradford Hill guidelines (Hill, 1965). These modified guidelines are reproduced in Table 3-3, arranged to summarise three 'types' of evidence, with revised guidelines matching Hill's original terms.

Type of evidence	Revised, structured guidelines	Hill's original guidelines
Direct	Size of effect not attributable to plausible confounding	Experiment
	Appropriate temporal and/or spatial proximity (cause precedes effect and effect occurs after a plausible interval; cause occurs at the same site as the intervention)	Strength
	Dose-responsiveness and reversibility	Temporality
Mechanistic	Evidence for a mechanism of action (biological, chemical, mechanical)	Biological gradient Biological plausibility
	Coherence	Coherence
Parallel	Replicability	Consistency
	Similarity	Analogy

Table 3-3 – Howick, Glasziou and Aronson's (2009; Table 1) proposed revisions to Bradford Hill's original guidelines to assess efficacy of hypothesised causes. 'Coherence' is moved from the 'Parallel' subheading to 'Mechanistic', in accordance with the paper's summary of the guidelines and other tables and figures.

From these, I developed a series of questions to evaluate the evidence given in a review for each hypothesised cause. Reviewed causes were marked with a '+' for effectiveness, denoting consistent evidence of causing a reduction in pregnancy rates in exposed samples, if at least one question regarding evidence strength was answered positively:

1. Are data presented which shows an effect while controlling for plausible confounding? ('Direct' evidence – see Table 3-3)

2. Are data presented which shows an effect whilst NOT controlling, but:

    a. A logic model or narrative synthesis of mechanisms is presented? ('Mechanistic'); AND

    b. The model is coherent with what is currently known? ('Mechanistic – Coherence')

3. Are results seen consistently in different national contexts? ('Replicability')

Similarly, if reviews presented evidence of negative effects of an intervention or environment change, increasing pregnancy rates, the evidence was categorised by '-'s.

To account for studies which may present strong evidence of the effectiveness of an intervention, but where the intervention may not have been scaled up to a whole population, I assessed evidence for the extent of exposure presented in reviews. A hypothesis was assigned a '+' for evidence of exposure if it was plausibly presented as contributing to a whole population reduction in rates, or a '-' if contributing to an increase in pregnancies. Reporting of exposure may include reviews testing the effectiveness of an intervention in a trial or pilot context, whilst noting its (unassessed) population-wide implementation elsewhere. Exposures examined within the UK were weighted similarly to exposures seen in other countries in assigning a score – this exposure score indicated evidence for the potential for scalability and population-wide implementation. The following question was used to determine exposure:

4. Is this hypothesised cause:

    a. Observed across the population of a whole nation (as included in this analysis – see Box 3-1); *OR*

    b. Observed amongst a large proportion of the population, or a high risk group (e.g., targeting prevention of repeat pregnancies or socio-economic subgroup); *OR*

    c. Cited by reviewers as having been implemented across a population of an included country after 1990, but with observations not reported in the review? (Direct evidence)

In cases where a hypothesis was considered by a review, but no clear evidence was found of either effects or population-wide exposure, review findings were marked '0' for that hypothesised cause. The resulting levels of categorisation assigned for each review are presented in Table 3-4.

Grade	Interpretation		Reviews classed if:		
++	Strong evidence for reduction in pregnancy rates	:	Studies show effective in decreasing rates	AND	Reported national or international implementation or exposure
+	Weak evidence for reduction in pregnancy rates	:	Studies show effective in decreasing rates BUT Not applied to population	OR	Associated with decreasing population rates BUT Not causally assessed
0	No clear evidence either way	:	Conflicting study results	OR	No association found
–	Weak evidence for increased pregnancy rates	:	Studies show effective in increasing rates BUT Not applied to population	OR	Associated with increasing population rates BUT Not causally assessed
– –	Strong evidence for increased pregnancy rates	:	Studies show effective in increasing rates	AND	Reported national or international implementation or exposure

Table 3-4 – Grades of strength of evidence presented in a review for a hypothesised cause of lowered teenage pregnancy rates

3.3.5.3 Synthesising within causal branches and updating causal tree

From the full set of identified causal pathways, I synthesised these into branches to be assessed, grouping by similarities in the environments in which they were delivered or similarities in programme elements. Where data were presented in high-quality reviews which were deemed sufficient to accurately assess the effects of a single branch, I synthesised data exclusively from these reviews. Where only low-quality reviews, or insufficient data from high-quality reviews were identified, I synthesised across all available reviews and noted this as a limitation due to study quality.

Within each branch I narratively synthesised available evidence across reviews assessing the effect and exposure of the hypothesised cause and constructed a summary table (Table 4-2) of evidence by branch. Where I found evidence for a branch in the reviews for each of the three 'efficacy' questions identified above, I recorded this as '-', '0' or '+' representing negative effects, no effects or positive effects respectively. Evidence for a causal effect across a population was similarly categorised under 'exposure'. Final grades were assigned to each branch using the same coding as Table 3-4 above. Where stronger evidence

conflicted with weaker evidence – for example where plausible mechanisms of effect were presented but randomised trials found no effect – I assigned the final grade to reflect the findings of the stronger evidence. Where strong or weak evidence was presented for the effects of hypothesised causes, I reported these in answer to RQ2: ‘For which causes is there evidence of their effects in reducing pregnancy, and their exposure across populations?’.

Where high quality reviews of an exposure consistently failed to find evidence of a direct effect on a sample population, or a population-wide implementation was not found to produce an effect, the branch was labelled ‘0’ overall and interpreted as evidence of the absence of an effect. These branches were used to answer RQ3: ‘Which causes have been robustly examined but have failed to produce evidence of their effects across a population, indicating the absence of an effect?’.

I used the summary table to update the causal tree. Where weak or strong evidence was presented for an existing branch, the branch was retained, and causal lines (negative or positive) were updated. Where weak or strong evidence was presented for a new branch, this was added to the tree. In cases where a hypothesised cause was extensively reviewed by high quality reviews, but no clear evidence of effect was found, this was considered evidence of the absence of an effect. Existing branches and new hypothesised branches with no effect were removed or omitted from the tree. In cases where insufficient or inconclusive evidence was identified across reviews, such branches were retained or added and highlighted as requiring further investigation. I reported this updated tree to answer RQ4: ‘What are the core elements to include in a causal tree which may have contributed to reductions in teenage pregnancy rates in the UK?’.

3.4 Evaluating the Teenage Pregnancy Strategy using interrupted time series methods

3.4.1 Developing the approach of the interrupted time series analysis

The combined protocol for the interrupted time series and synthetic control analyses is published at **osf.io/tdbr8**.

Time series analyses are useful approaches for estimating future trends from past data. In the area of policy evaluation, trends established in the pre-implementation period can be

used to project a counterfactual trend in outcomes to represent the assumed observations in the absence of the intervention (Craig *et al.*, 2017; Lopez Bernal, Cummins and Gasparrini, 2017). Numerous approaches with varying complexity have been suggested to test the assumptions of the time series model in robust ways (Kontopantelis *et al.*, 2015). In this analysis I applied a segmented regression model, which fits linear trends pre- and post-intervention and measures a change in level and trend at the point of implementation to estimate the effects of exposure (Kontopantelis *et al.*, 2015; Craig *et al.*, 2017; Lopez Bernal, Cummins and Gasparrini, 2017). A potential limitation of such methods is the possibility of time-varying confounders, such as simultaneous events or other interventions which may impact on trends. To account for potential unanticipated population changes across the time period which may produce spurious effects, I used control country observations to estimate the null effect level and trend changes expected in the absence of the Strategy (Lopez Bernal, Cummins and Gasparrini, 2018).

A further strength of the interrupted time series model is its ability to correct for autocorrelation. Sequential data points measured for the same observed unit are unlikely to be independent of each other – for example, an unusually high rate of pregnancies in one year may be expected to predict an unusually high rate in the next. This correlation of sequential error terms, termed autocorrelation, violates a requirement of standard linear regression models (Nelson, 1998; Lopez Bernal, Cummins and Gasparrini, 2017). Autocorrelation may occur over a series of one or more 'lags', meaning the previous one or more time points may have an effect on later outcomes. The two types of autocorrelations – autoregression (AR) and moving average (MA) – can be detected using a series of tests and coefficients can be added to the linear model to correct for each using the 'nlme' package tools in R (Pinheiro *et al.*, 2019).

To simulate the launch of the Teenage Pregnancy Strategy, in each analysis I set the intervention start as 1999. Across all comparisons, the primary outcome was under-18 pregnancy rates, defined as counts of conceptions per 1,000 women aged under 18 years, ending in a birth, abortion or stillbirth and recorded in the UK.

3.4.2 Aim and research question

In the Interrupted Time Series analysis, I aimed to evaluate the effectiveness of the Teenage Pregnancy Strategy by comparing England's change in rates to Scotland and

Wales as unexposed controls. This was done to answer research question 5 (section 3.2 above):

RQ5: Did implementation of the Teenage Pregnancy Strategy cause a change in level or trend in yearly teenage pregnancy rates in England, compared with other UK countries?

3.4.3 Data collection

I hypothesised that recorded rates of pregnancy across the period 1999 onwards would be affected by the Strategy. I sought to estimate rates of pregnancies to women aged under-18 rather than births, reflecting the Strategy's target outcome and consistent with previous evaluations (Social Exclusion Unit, 1999; Wilkinson *et al.*, 2006; Wellings, Palmer, Geary, *et al.*, 2016). This metric would include all conceptions ending in birth, still birth or abortion, reflecting Strategy goals and action plan (Social Exclusion Unit, 1999). I aimed to collect data for England, Scotland and Wales for the exposed period and from at least 1990 onwards to capture a sufficient time period leading up to the Strategy and until most recent reporting of data from relevant sources (to at least 2016 at time of data collection). These data would allow estimates of pre-implementation and post-implementation trends. All data sources used are listed in Appendix C.

I extracted estimated rates of teenage pregnancy directly from the Office for National Statistics (ONS) report for England and Wales (Office for National Statistics, 2019) and the Information Services Division Scotland (ISD) report for Scotland (Information Services Division Scotland, 2018a) for all reported age groups (under-16, under-18 and under-20). Both sources used the calculations in Equation 3-1, summing recorded births, still births and abortions in each age group, correcting for date of conception and for multiple births, and dividing by the estimated female population of a representative age range (Office for National Statistics, 2017; Information Services Division Scotland, 2018a). Conceptions to UK teenagers ending in miscarriages, births outside the UK, and unrecorded abortions are not included in these counts (Office for National Statistics, 2017). I did not include Northern Ireland due to the unreliability of estimates of abortions (The Family Planning Association, 2015).

$$a)\ Under\ 16\ rate = \frac{births_{<16} + abortions_{<16} + still\ births_{<16}}{population\ 13-15}$$

$$b)\ Under\ 18\ rate = \frac{births_{<18} + abortions_{<18} + still\ births_{<18}}{population\ 15-17}$$

$$c)\ Under\ 20\ rate = \frac{births_{<20} + abortions_{<20} + still\ births_{<20}}{population\ 15-19}$$

Equation 3-1 – Calculation of pregnancy rates for three adolescent age groups, used by ISD Scotland and ONS in calculation of reported rates. Each count was a count of events, rather than number of babies to better represent conceptions. Counts of births, abortions and still births were corrected to age at conception. Population counts are units of 1,000 women, estimated from census data. Different age ranges are used as denominators across the three measures, excluding 13 and 14-year-olds in b) and c). These groupings would represent a comparable population of most at-risk young women who are under the threshold age.

Rates were reported by age group as pregnancies per 1,000 women per year. All rates are reported in this analysis as 'under-16', 'under-18' or 'under-20' to represent these calculated values. Data were available for under 18s for England and Wales separately from 1992 to 2016, England and Wales combined from 1987-2016, and Scotland from 1994-2016, but were sparse for other age categories (Figure 3-2). To allow for comparisons of pregnancy rates across all age categories from an earlier start date, I sought data from other sources to allow for a fuller range of comparisons.

Scottish rates were only reported by ISD for 1994 onwards; to supplement these I calculated Scottish under-18 and under-20 pregnancy rates from 1987 to 1993 to match the earliest data available for England and Wales. I extracted counts of Scottish births by age of mother (National Records of Scotland, 2018), abortions to under-20s (Information Services Division Scotland, 2018b), and estimates of female population by age (from the Human Mortality Database; University of California and Max Planck Institute for Demographic Research, 2019). Total births to women aged 12 to 18 and 12 to 20 were calculated for the years 1987 to 2015. These were divided by estimated populations in units of 1,000 women grouped in the denominators in Equation 3-1b) and c) to give uncorrected estimates of births resulting in conceptions for women aged under 18 and under 20. This total included multiple births so it overestimated numbers of pregnancies ending in registered births and did not include still births or abortions.

I compared these rates with ISD Scotland reported birth rates for under-18s and under-20s for the years 1994 to 2015 to determine corrections for each group to account for multiple births, still births and births at a higher age and later year than conception. I calculated a proportion of births to 18-year-olds and 20-year-olds to include in counts of conceptions to under-18s and under-20s respectively, and a proportion of births in each calendar year to

count as conceptions in the previous year. I calculated a correction factor to account for multiple births and still births by performing linear regression, using my calculated rates as a predictor of ISD reported rates. This produced a good fit predictor for both age groups (R^2_{U18} = 0.996 and R^2_{U20} = 0.997). Records of abortions to under-20s were similarly corrected for year of conception added to the under-20 age group to give total estimates of under-20 conceptions. Finally, I used under-20 abortion rates in a linear model as a predictor of under-18 abortion rates which produced a good fit model (R^2 = 0.991).

I compared final estimates of rates using these calculations to observed years across 1994-2015 to assess the accuracy of the calculations. These methods produced good fit predictors of Scottish pregnancies for under-18 and under-20 age groups across the overlapping years, and so I assumed they provided accurate proxy measures of pregnancies when recorded rates were not available. I used these correction factors to impute under-18 and under-20 pregnancy rates for the unobserved years 1987 to 1993. I also estimated Under-16 pregnancy rates using these methods but I found they were poorly matched to reported pregnancies and so I did not use these calculated rates in my analyses.

As shown in Figure 3-2, data for England only are only available for 2001-2016 for under-16 and 2009 to 2016 for under-20. I was unable to find data to supplement these. Therefore, I used aggregated England and Wales rates as a proxy for rates in exposed England in these age group comparisons. I compared England only data with England and Wales combined data for years where both rates were reported to assess the suitability of the combined data as a proxy for England. England contributed around 95% to both counts of population and pregnancy outcomes and rates were very similar across all years. This suggests that the aggregated values for England and Wales rates are a good indicator of exposed England's rates in the absence of England-only data. In analyses with these data, I used only Scotland as a comparator.

The full set of available data allowed primary comparisons of England's under-18 pregnancy rate against Scotland and against Wales for the years 1992-2016 and secondary comparisons of combined England and Wales rates with Scotland for 1994-2016 (under-16 rates), 1987-2016 (under-18 rates) and 1990-2016 (under-20 rates). I aimed to test the sufficiency of this data through checking for variability, linearity and whether correction for other confounders produced good fit pre-intervention trends (Lopez Bernal, Cummins and Gasparrini, 2017).

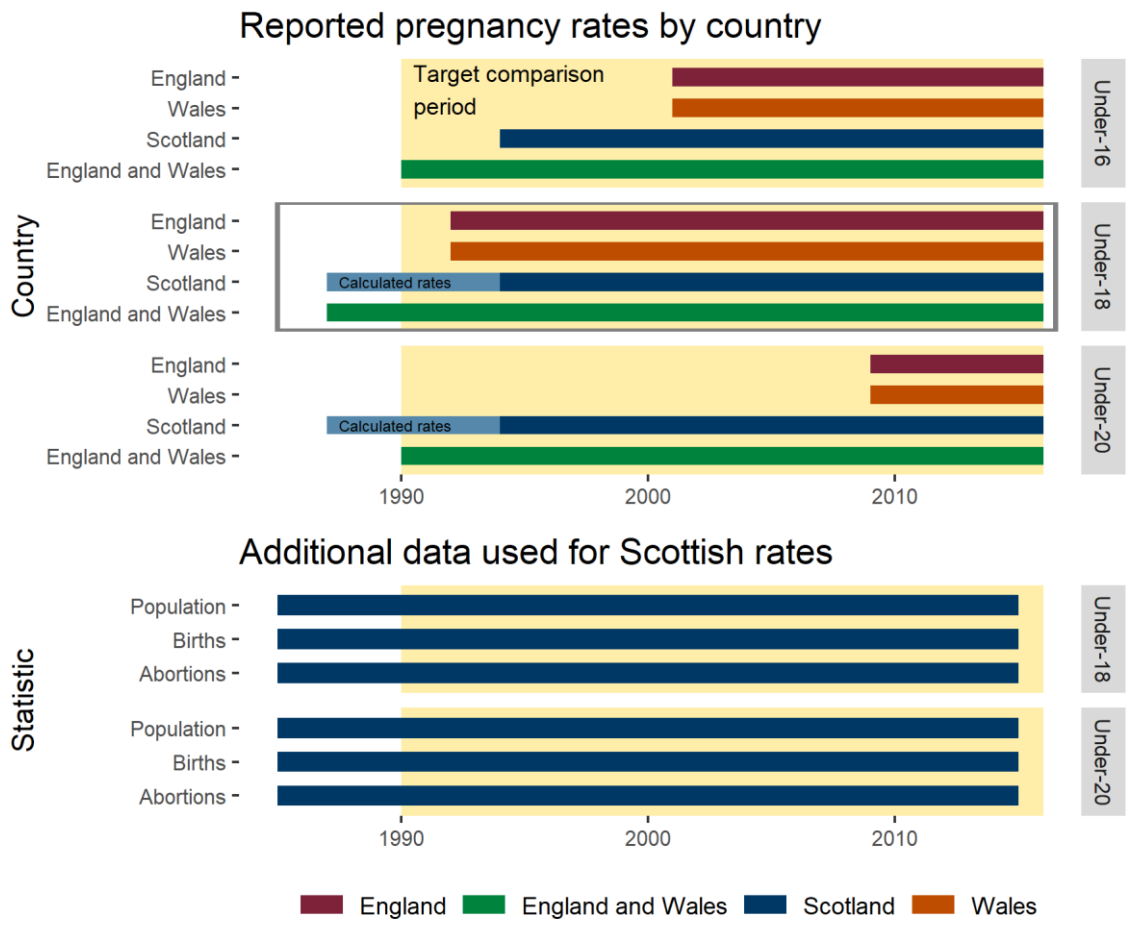

Figure 3-2 – Availability of data by years for England, Wales and Scotland, highlighting year range 1990-2016 (yellow boxes) and age group 'Under-18' (grey rectangle) as primary targets for data extraction. Data were available for Scotland for calculations of under-16 rates but were not used due to poor power to predict pregnancy rates in known years. Under-18 pregnancy rates were the target outcome, and England rates alone were the primary measure of these.

3.4.4 Statistical analysis

3.4.4.1 Preparatory models

All analyses used R (v. 3.6.3; R Core Team, 2020) and RStudio (RStudio Team, 2015). I built a Shiny app to carry out the analyses, allowing for interactive adjusting of parameters, reproducibility of results and standardised model outputs (Chang *et al.*, 2019). The app is published online at **andybaxter.shinyapps.io/teen_preg_uk_its**. All R packages used are listed in Appendix C.

In my preparatory models, I fitted a linear model using ordinary least squares to England observations from 1992 to 1998, before the start of the Strategy. This was used to estimate the baseline trend as a hypothesis of the trajectory England would have followed in the absence of the Strategy. When I visually inspected the pre-intervention trend in England data, there appeared a relatively stable linear trend from 1992 to 1998, with an increase in

level at 1996 (see Figure 5-1 in Chapter 5 reporting analysis results). This has previously been attributed to the effects of a warning issued concerning the safety of oral contraceptive pills in 1995 and the subsequent fall in contraceptive use (the 'pill scare'; Furedi, 1999; Teenage Pregnancy Strategy Evaluation, 2005; Wellings, Palmer, Geary, *et al.*, 2016). I concluded that, pre-intervention, these observations could be assumed to be following a linear trend with a mid-point level increase. I thus aimed to test in the full model whether a corrector for the 'pill scare' would improve model fit by accounting for the jump in rates at 1996.

Using a dummy variable for Strategy exposure ('Intervention' = 1 for all years from 1999 onwards) and a further count for years from implementation (starting at 'Trend' = 1 in 1999), I then fitted an intervention trend line to data from 1999 to 2016 to estimate the changes in trend and level from the start of the Strategy. This allowed years beyond the 2010 end of the intervention to contribute to estimates of its effects, consistent with previous evaluations (Wellings, Palmer, Geary, *et al.*, 2016). I visually inspected the pregnancy rates across this period to determine if any changes immediately after the 2010 end indicated a temporary effect of the Strategy, requiring exclusion of later data. This trend remained consistent and so these time points were used in all analyses as assumed ongoing effects of the Strategy.

I tested England-only data for autocorrelation using Durbin-Watson tests, and autocorrelation and partial autocorrelation function plots. If a significant autocorrelation was seen at a lag time point in the Durbin-Watson test, I then examined the two plots to determine if this appeared consistently and whether it represented an autoregression or a moving average correlation. I repeated these analyses across models with added predictors and compared results. There were repeated, strong indicators of autoregression at lag 1 and inconsistent evidence of both autoregression and moving average components at lag 3. This result indicated that a higher-than-expected rate in year y would predict a higher-than-expected rate in year y+1 but would have little additional predictive power over subsequent years. To achieve consistency across all models, and to apply parsimony in adding components, I set all models to correct for autoregression at lag 1. I hypothesised that this would be consistent with the nature of the data, as these were grouped by calendar year, spanning two school years. One-off events or changes in behaviour or culture could feasibly produce a more homogenous effect across a school year. An increase over the expected pregnancy rates in one calendar year may have a common cause with similar increases in the next year. I produced a final fitted model using maximum likelihood

estimation and autoregression correction set to lag 1 and moving average correction set to 0. This model was termed AR1, MA0 in reporting.

In my initial model, I recorded fitted terms representing the baseline rate at 1991, the pre-Strategy yearly change in pregnancy rates, the level change in 1999 at the introduction of the Strategy, and the post-implementation difference in trend, along with standard errors and confidence intervals. These were calculated using Equation 3-2. Level and trend change coefficients (β_2 and β_3) were taken to represent the naïve uncontrolled estimate of extent of change at 1999, associated with Strategy implementation.

$$Rate = \beta_0 + \beta_1 * Time + \beta_2 * Intervention + \beta_3 * Trend + \epsilon$$

Equation 3-2 – Linear model estimating yearly pregnancy rates using England data with no comparators. 'Time' records the number of years from the start of observations (with 1992 coded as Time = 1), 'Intervention' is a binary dummy variable for exposure/non-exposure to the Strategy, and 'Trend' is a variable counting the years since Strategy implementation (with 1999 coded as Trend = 1). β_0 represents the intercept at Time 0 (in 1991, before first observation), ϵ is an error term.

To improve the fit of the pre-intervention trend, I added a 'pill scare' dummy variable for all dates from 1996 onwards (generating an additional fitted term; equation not shown here). This aimed to account for the hypothesised jump in rates caused by the decreased use of oral contraceptives. This parameter improved model fit (likelihood ratio = 10.7; p = 0.001), and so I retained this for all further models. I extracted level and trend changes as a closer estimate in an England-only model.

3.4.4.2 Comparison models

My controlled models used Scotland and Wales as control populations to estimate the expected changes at 1999 in the absence of the Strategy. In primary analyses I used data from 1992 onwards to match the earliest data available for all countries. Upon inspection, all three countries exhibited the same 'pill scare' level increase in 1996, which was consistent with the UK-wide exposure to this hypothesised cause (Furedi, 1999). This corrector was applied to all countries in subsequent analyses. No other abnormalities were seen across the pre-intervention trends from 1992 and so this date was kept as the beginning of analyses across primary comparative models. One further parameter was added to the model ('Group'; Equation 3-3) to note England as the exposed unit. This was allowed to interact with baseline trend and variables for intervention exposure and trend change. Changes in level and trend seen in Scotland and Wales were subtracted from those seen in England to give estimates of the Strategy's effects, corrected for background changes common to all three countries. I extracted difference in level and trend changes at

1999 (β_6 and β_7 below) as estimates of the Strategy effectiveness whilst controlling using estimates from unexposed units.

$$Rate = \beta_0 + \beta_1 * Time + \beta_2 * Group + \beta_3 * Group * Time + \beta_4 * Intervention + \beta_5 * Trend + \beta_6 * Intervention * Group + \beta_7 * Trend * Group + \epsilon$$

Equation 3-3 – Linear model estimating trend and level change differences from unexposed control. As in Equation 3-2, 'Time' records the number of years from the start of observations (with 1992 coded as Time = 1), 'Intervention' is a binary dummy variable for exposure/non-exposure to the Strategy, and 'Trend' is a variable counting the years since Strategy implementation (with 1999 coded as Trend = 1). To compare with a control, the 'Group' variable codes 1 for the exposed unit (England) and zero for controls. Used alone (β_2), this accounts for a difference in baseline level between England and control. Interactions with Time, Intervention and Trend give coefficients estimating the difference in baseline trend (β_3), difference level change at the start of the intervention (β_6) and the difference in trend across the intervention from control observations (β_7). ϵ is an error term.

Inspection of pre-intervention trends between England and controls indicated that all three countries closely followed the same pattern before the Strategy. I removed the interaction term for difference in baseline trend and compared model fit. Removal of the parameter did not significantly impact model fit (likelihood ratio = 0.20; p = 0.66) and the term allowing for baseline trend changes did not therefore have significant explanatory power. In its absence, pre-intervention observations from both control and exposed groups were used in calculation of baseline trends, reducing the standard errors of trend estimates. Therefore, in the primary model I used the assumption of pre-intervention parallel trends, allowing more stable predictions from the limited pre-intervention data. Coefficients β_5 and β_6 (Equation 3-4) represented changes associated with intervention exposure.

$$Rate = \beta_0 + \beta_1 * Time + \beta_2 * Group + \beta_3 * Intervention + \beta_4 * Trend + \beta_5 * Intervention * Group + \beta_6 * Trend * Group + \epsilon$$

Equation 3-4 – Linear model of controlled time series with assumption of parallel trends.

3.4.4.3 Sensitivity analyses

In further testing my models, I examined the effects of adding further parameters to improve model fit to determine if these produced conflicting effects estimates. After examining rates across all three countries, I saw a similar trend change occurring from 2008 onwards, dividing the post-intervention period into two segments. As this was a pattern repeated across UK nations and not coinciding with a Strategy-related event, I hypothesised that it represented a UK-wide exposure for which my model was unable to account. In sensitivity analyses, I treated 2008 as a common shock across all countries and allowed a common trend-change to better fit the observations (giving β_7 in Equation 3-5 below).

$$Rate = \beta_0 + \beta_1 * Time + \beta_2 * Group + \beta_3 * Intervention + \beta_4 * Trend_1 + \beta_5 * Group * Intervention + \beta_6 * Group * Trend_1 + \beta_7 * Trend_2 + \epsilon$$

Equation 3-5 – Linear equation of time series with common shock (trend change) at 2008. The variable '$Trend_1$' denotes the trend change at intervention start date, allowed to interact with group definition and produce a difference in trend associated with Strategy-exposure. '$Trend_2$' denotes a second, common change in trend at 2008, hypothesised as an unknown UK-wide exposure from observation of the data. It is similarly coded as years from start of shock period (with 2008 as $Trend_2 = 1$).

The inclusion of a common trend-change at 2008 produced an improved fit (likelihood ratio 51.3; $p < 0.0001$). Visual inspection of linear regression fits both with and without this corrector did appear to show that the data points most closely followed a linear trend with all included terms. However I did not replace the primary model with the expanded model. I judged that the parsimony of fewer terms, the lack of a satisfactory explanation for this trend change, and the assumption that the controlled model could account for such a common shock justified presenting the simpler model as a primary analysis. I included this model in sensitivity analyses and examined model fit, consistence of parameter estimates and power to detect change.

To test whether allowing for a phase-in period improved model fit, I excluded data for the years immediately following the start of the intervention. This made no difference to fit or prediction, so data recording all years were retained in final analyses.

To assess whether other age groups or longer time periods were more able to detect an effect, I conducted robustness checks on secondary data sets. Data for England alone was only available for 1992 onwards, giving seven pre-intervention time points (Figure 3-2). To examine longer pre-intervention time periods, I used combined England and Wales data, available from 1987, to compare with Scotland for a longer period. Additionally, as England and Wales data were not available separately for pre-intervention periods for the age groups under-16 and under-20, I used combined data to conduct secondary analyses of these. I repeated the models in Equation 3-4 and Equation 3-5 across the three age groups with aggregated England and Wales data. In visual inspection of the under-18 and under-20 pregnancy rates in these analyses, Scotland exhibited a large divergence from England and Wales observations prior to 1991, and linear trends were less clearly defined before this date. I set observations start dates to 1991 in these analyses, providing only one additional observation per country. Scotland under-16 data was only available from 1994, and so this comparison observed only from 1994 onwards. I compared level and trend

change estimates across these models with the primary model to check for consistency of results.

3.4.4.4 Testing power of models

To test the power of the models to detect change, I used estimates of the slope change in under-18 pregnancies at 1999 in Scotland and Wales for the three forms of controlled model: the single trend change model with no correctors, the model with a 1996 pill scare corrector (the primary comparison model) and the model with both the 1996 corrector and the 2008 trend change corrector (the sensitivity analysis model). I used these values and standard errors to estimate the probability of detecting a greater negative trend change in England at the 95% significance level, assuming equal standard errors between control and England. I tested for small, medium and large differences in trend change of 0.5, 1 and 2 fewer pregnancies per 1,000 women per year. A reduction of 2.33 pregnancies per 1,000 women per year across the ten years of the Strategy's implementation would reach the target of a 50% reduction from the 1998 pregnancy rate (46.6 pregnancies per 1,000 women per year). A reduction of 1 pregnancy per 1,000 women per year would represent a rate change similar to the 2% annual reduction estimated from the mid-term evaluation (Wilkinson *et al.*, 2006).

To calculate power, I adapted the equation presented in McLeod and Vingilis (2005) to use a one-tailed approach (aiming to detect a hypothesised decrease in pregnancy rates, rather than a change in a non-specified direction) using observed standard errors of Scotland and Wales' trend changes. Using the standard normal distribution function, I calculated the one-tailed probabilities of observing a non-significant effect estimate (an estimated trend change for England smaller than 1.64 standard deviations lower than control) in the cases of the true effect being the small, medium or large tested values. These represented the false-negative rates (the probability of not detecting a true trend change). I subtracted the probabilities from 1 to calculate the probability of observing a significant effect if these reductions were the true effects. I used these figures to represent the power of the models to detect change at each level of effect size (McLeod and Vingilis, 2005, 2008).

The power of the models to detect a medium size change in trend was 56% using Scotland data and 62% using Wales data. This power increased by a small amount to 60% and 63% respectively in the primary models using the pill scare corrector. Further addition of a common trend change term at 2008 in sensitivity models increased the power of the models to 95% and 94% to detect a trend change at 1999. All models had high power to

detect a large change (2 fewer pregnancies per 1,000 women per year), ranging from 98% to ~100% (Table 3-5).

Model	Power to detect a difference in trend change of magnitude: Small: 0.5 fewer pregnancies per 1,000 women per year	Medium: -1 fewer pregnancies per 1,000 women per year	Large: -2 fewer pregnancies per 1,000 women per year
Scotland – no correctors	23%	56%	**98%**
Wales – no correctors	25%	62%	**99%**
Scotland – pill scare corrector	24%	60%	**98%**
Wales – pill scare corrector	26%	63%	**99%**
Scotland – pill scare and 2008 corrector	50%	**95%**	**100%**
Wales – pill scare and 2008 corrector	48%	**94%**	**100%**

Table 3-5 – Power calculations for small, medium and large effect sizes across comparative model types. Power values greater than 80% are in bold and underlined.

3.4.4.5 Reporting results

From each model, I extracted coefficients and 95% confidence intervals for difference in level and trend change seen in England over controls at 1999 and used these as markers of change due to exposure to the Strategy. I extracted measurements of R^2, Mean Squared Prediction Error and log-likelihood as indicators of model fit. I produced a scatter plot showing the observations across all time points used and linear trends fitted for observed periods and predicted trends.

3.5 Evaluating the Teenage Pregnancy Strategy using synthetic control methods

3.5.1 Developing the approach of the Synthetic Control method

The combined protocol for the interrupted time series and synthetic control analyses is published at **osf.io/tdbr8**.

In a synthetic control analysis, an exposed unit can be compared to a set of comparator units whilst selecting from this pool of comparators to construct a best fit series of

predictions. Rather than taking a mean of all comparators to predict the outcome at each time point, an algorithm can apply different weights to each unit consistently across the observation period to construct a synthetic control as a series of weighted means (Abadie, Diamond and Hainmueller, 2011; Bouttell *et al.*, 2018). In the pre-intervention comparison period, weights are assigned to comparators to produce a series of predictions which minimises the error and therefore best fits the observations of the exposed unit in the unexposed period. If the weightings can be shown to produce a good predictor, then these same weights can be applied to controls to construct a post-implementation time series which can be posited as predicting the outcomes in an unexposed setting. This approach ensures a 'data driven' selection of controls to account for potential bias in use of controlled analysis and corrects for common confounders in closely-matched units (Abadie, Diamond and Hainmueller, 2010). An additional advantage which is crucial for this analysis, is that the construction of a control from multiple units means that comparison can still be carried out when no single control unit closely matches the exposed outcomes. This may be the case when comparing England to a wider set of countries outside of the UK.

The key assumptions of synthetic control model methodology are: treated and potential control units are similar; there is no contamination of exposure effects in potential control units; there are no idiosyncratic shocks in potential control units (Bouttell *et al.*, 2018). My selection of controls from European and high-income European countries sought to satisfy the first assumption. As discussed above, contamination should be minimal, but I additionally conducted analyses using only countries external to the UK to allow for this. Addressing the third assumption, several countries were excluded due to known historical events, not applicable to the UK.

To enable a synthetic control analysis, I sought data from a range of comparison countries with pre-Strategy pregnancy or birth rates both higher and lower than England. I began by collecting data to estimate each country's rates and aimed to construct a synthetic England from these data alone. I supplemented this analysis using predictors which I hypothesised to be associated with pregnancy or birth rates as available for all included countries. I sought data to test if adding these to the model produced a closer fit synthetic control in the pre-Strategy period, indicating a more robust predictor of post-Strategy outcomes.

I conducted placebo tests as described by Abadie, Diamond and Hainmueller (2010, 2011) to assess the comparative magnitude of difference from control for non-exposed countries

compared to exposed England. I conducted time-based placebo tests by setting placebo intervention times and robustness tests by removing top-weighted countries to test the model (Abadie, Diamond and Hainmueller, 2015).

3.5.2 Aim and research questions

In the synthetic control analysis, I aimed to evaluate the effectiveness of the Teenage Pregnancy Strategy by comparing England's change in rates to a control constructed from countries most closely matching England's rates and trends. This was used to answer research question 6 (section 3.2 above):

RQ6: Did implementation of the Teenage Pregnancy Strategy reduce yearly teenage pregnancy rates in England, compared with a wider set of control countries?

3.5.3 Data collection

For the synthetic control analyses, I selected countries for comparison based on cultural, political, geographical and economic similarity to England (Abadie, Diamond and Hainmueller, 2015). I used the same set of countries as for the systematic review as an initial point for selecting a donor pool (section 3.3.4.1 above). I sought data on teenage births and pregnancies for all Euro-peristat nations (Euro-Peristat, 2018) and other high-income Anglophone countries. I excluded eight European countries that were either in the former Yugoslavia or the USSR, or were previously USSR-backed, as they had turbulent histories around this time (Hungary, Estonia, Lithuania, Slovenia, Czechia, Poland, Croatia, and Bulgaria). These changes may produce idiosyncratic shocks in exposures which are not generalisable to England (Abadie, Diamond and Hainmueller, 2015; Bouttell *et al.*, 2018). I aimed to collect data recording births and pregnancies for a sufficient number of time points before and after the intervention. Data were available for included countries for the years 1990 to 2013, which I judged sufficient. All data sources used are listed in Appendix C.

As a primary outcome, I aimed to match the measures of teenage pregnancy used in the interrupted time series analysis as closely as possible. However, there were no consistent records of under-16, under-18 or under-20 pregnancy rates available for the full set of countries. I sought data on populations, births and abortions to estimate these rates for the relevant time period. I extracted data estimating births by age of mother from the Human Fertility Database (Max Planck Institute for Demographic Research, 2019), populations

from the Human Mortality Database (University of California and Max Planck Institute for Demographic Research, 2019), and numbers of abortions to women under-20 from the WHO Health for All Explorer (WHO Regional Office for Europe, 2019). For countries not included in the Human Fertility Database (New Zealand & USA), I sought data on births, abortions and pregnancies from national statistics websites. I extracted counts of births and abortions New Zealand from Statistics New Zealand reports on (Statistics New Zealand, 2019b, 2019a). I extracted pre-calculated pregnancy and birth rates for the USA from a report by the Guttmacher Institute (Kost, Maddow-Zimet and Arpaia, 2017). The authors calculated these using population, birth and abortion data from the National Centre for Health Statistics and the Center for Disease Control.

After conducting searches for data, I excluded five countries for which complete data were not available (Austria, Australia, Canada, The Republic of Ireland and Northern Ireland). This was due to different age groupings, insufficient time points or no reliable records of abortions. The final selection of fifteen control countries with England and Wales and dates of data availability are shown in Table 3-6.

For included countries, data were available to calculate under-18 birth rates and under-20 pregnancy rates for 1990-2013. I recalculated England and Wales' and Scotland's rates from these datasets to make them comparable. England and Wales were used as a single unit as only combined data were available. As seen in section 3.4.3 for the interrupted time series, this was a suitable proxy for England only.

Country	Dates data available by outcome				
	Population	**Under-18 births**	**Under-18 abortions**	**Under-20 births**	**Under-20 abortions**
Countries with counts extracted from datasets					
Denmark	1990-2013	1990-2013		1990-2013	1990-2013
England and Wales	1990-2013	1990-2013	1990-2013*	1990-2013	1990-2013
Finland	1990-2013	1990-2013		1990-2013	1990-2013
France	1990-2013	1990-2013		1990-2013	1990-2013
Germany	1990-2013	1990-2013		1990-2013	1990-2013
Iceland	1990-2013	1990-2013		1990-2013	1990-2013
Italy	1990-2013	1990-2013		1990-2013	1990-2013
Netherlands	1990-2013	1990-2013		1990-2013	1990-2013
New Zealand	1990-2013	1990-2013		1990-2013	1990-2013
Norway	1990-2013	1990-2013		1990-2013	1990-2013
Portugal	1990-2013	1990-2013		1990-2013	1990-2013
Scotland	1990-2013	1990-2013	1990-2013*	1990-2013	1990-2013
Spain	1990-2013	1990-2013		1990-2013	1990-2013
Sweden	1990-2013	1990-2013		1990-2013	1990-2013
Switzerland	1990-2013	1990-2013		1990-2013	1990-2013
Countries with rates reported elsewhere					
USA		1990-2013†	1990-2013†	1990-2013†	1990-2013†
Countries with insufficient data (Excluded from analysis)					
Austria, Australia, Canada, Ireland, Northern Ireland					

Table 3-6 – Data availability for all countries across years searched. *Under-18 abortion rates for Scotland and England and Wales were available separately from ISD Scotland and ONS data but were not used. †USA data were reported as calculated rates, and these were extracted directly.

I calculated under-18 birth rates by summing all births to women aged under-18 and dividing by total populations aged 15-17 (Equation 3-6a), matching the age group denominators used by ONS and ISD Scotland (Office for National Statistics, 2017; Information Services Division Scotland, 2018a). As counts of abortions were not available for the under-18 age category, I was only able to estimate birth rates for under-18s, rather than total pregnancies (Table 3-6). The outcome ‘under-18 birth rate’ was then used as a proxy for under-18 pregnancies. I calculated under-20 pregnancy rates by summing all under-20 births, adding total abortions to women under 20 and dividing by total populations aged 15-19 (Equation 3-6b). As numbers of pregnancies to under-16s was expectedly smaller and such pregnancies are often observed to end with abortion more

often than birth, I concluded that calculations of under-16 rates would likely produce biased and uncertain estimates, and so excluded this outcome from my synthetic control analysis.

$$a)\ Under\ 18\ birth\ rate = \frac{births_{<18}}{population\ 15-17}$$

$$b)\ Under\ 20\ pregnancy\ rate = \frac{births_{<20} + abortions_{<20}}{population\ 15-19}$$

Equation 3-6 – Calculations of under-18 birth rates and under-18 pregnancy rates. These calculated rates did not account for multiple births, still births, miscarriages, illegal abortions or conceptions within the age category leading to birth or abortions after turning 18 or 20 respectively. The age ranges used to sum population as denominators are taken from the equivalent rates in Equation 3-1 used by ONS and ISD Scotland. These ranges use the groups of women under 18 and under 20 who are in theory at most risk of pregnancy and birth.

Both measures used in my synthetic control analyses are proxies of true pregnancy rates. Counts of births and abortions do not allow for correction for multiple births, still births or illegal abortions. Further, each count or estimate represented a birth or abortion to a woman of a given age in a calendar year, irrespective of whether the conception occurred that year or after the woman turned that age. Some conceptions to 17- and 19-year-olds will be missing from the under-18 birth rates and under-20 pregnancy rates respectively. Additionally, the rates reported for the USA were derived using different numerators (total pregnancies and abortions to women aged from 15 years old rather than 12 years old; Kost, Maddow-Zimet and Arpaia, 2017) and are not directly comparable. However, I judged that within the synthetic control analysis itself these concerns would be minimal, as biases resulting from these would be unidirectional and would feasibly only account for a negligible number of pregnancies across each country. The methods of weighting in the synthetic control analysis minimise the potential effects of these data limitations, as control countries are only allowed to contribute to the weighted control via a scaled proportion of their total rates.

I hypothesised that several population characteristics which vary across time may be predictors of pregnancy rates and sought data on these to use to match countries by similarity to England. Estimates of yearly gross domestic product (GDP), mobile phone ownership and proportion of population resident in urban settings for years 1990 to 2013 were extracted from World Bank open data as predictor variables for the synthetic control models (The World Bank Group, 2019). Public spending on education as a proportion of GDP for the years 1990 to 2013 was extracted from OECD data (Organisation for Economic Co-operation and Development, 2018). These data were not reported separately

for countries within the UK. Consequently, I used UK-wide data as an estimate of England and Wales and excluded Scotland from analyses using these data.

All countries' under-18 birth rates, under-20 pregnancy rates and predictors were compiled into a data frame to be used in each analysis. These data frames are available at **phd.andybaxter.me/Synth-analysis**.

3.5.4 Statistical analysis

3.5.4.1 Optimising model fit and fitting models

All analyses used R (v. 3.6.3; R Core Team, 2020) and RStudio (RStudio Team, 2015). I primarily used the Synth package as developed by Abadie et al. (2011) in constructing synthetic control models. After I conducted all analyses, I built the data and functions running synthetic control models and producing graphs into a Shiny app to allow for reproduction of each model (available at **phd.andybaxter.me/synth-app**). All other packages used in data preparation and presentation are listed in Appendix C.

I used synthetic control methods to construct a comparison unit from a weighted average of other countries' rates, fitted to pre-intervention England and Wales observations. I used under-18 birth rates as a primary outcome and under-20 pregnancy rates as a secondary outcome to get a clearer estimate of effect on pregnancies rather than births.

In constructing models, I focussed on the pre-intervention period to assess and improve the closeness of the synthetic control predictor and observed England and Wales. I set England and Wales as the exposed unit and 1990-1998 as the time period for optimisation of fit. Initial models used each country's mean rate across the whole pre-intervention period (1990-1998) as a single predictor to construct the synthetic England. I recorded the pre-intervention mean squared prediction error (MSPE) as a measure of the average closeness of predicted values to the yearly observed rates. Visual inspection of the observed and predicted trend lines showed that this produced a poor fit (and a high MSPE of 1.97) as the England and Wales rates did not follow a linear pattern.

To improve the pre-intervention control fit, I divided the nine pre-intervention years into continuous groupings. I used a data-driven approach to finding an optimal grouping by iterating over all permutations up to four clusters of years. I directed the Synth functions to calculate means for each period as a predictor, to allow for the non-linear pattern of the yearly rate changes.

From all tested combinations of predictor years, I aimed to determine an optimal grouping as a combination of as few groups as possible and a minimised mean squared prediction error. The final grouping for the under-18 births model was four clusters of years: 1990-1993, 1994, 1995 and 1996-1998, with the synthetic predictor simultaneously aiming to match England in each period. By attempting to match the pattern seen in England across these four groups of years, the Synth function produced an improved fit over the method of using the whole period mean as a single predictor (MSPE = 0.092). I repeated this process with under-20 pregnancy rates. The final grouping for this model was three clusters of years: 1990, 1991-1995 and 1996-1998 (MSPE = 4.27, reduced from 23.0 with one mean taken across all years).

After selecting the best pre-intervention fit for each using only birth and pregnancy rates, I tested the effects of adding other predictors on the overall model fit. I used data recording GDP per capita, public spending on education, percentage of population in urban dwellings and rate of mobile phone ownership to construct controls which were close to England in other factors which may influence pregnancy rates. In each data source for further predictors, only aggregate UK figures were available. I used these as estimates of England and Wales' metrics and excluded Scotland from these multi-predictor models, as using identical data would have over-estimated Scotland's suitability as a comparator. I constructed synthetic control fits to pre-intervention periods and examined England's observed values across predictors alongside synthetic calculated values to check how closely these were able to be replicated and to assess model fit. I compared the pre-intervention MSPE to that of the simple model (with no predictors besides pregnancy or birth rate). Finally, I constructed synthetic control trends across the comparison period with all predictors included.

3.5.4.2 Testing for Strategy effects

To test for differences in pregnancy rates from control which were consistent with Strategy exposure, I used the methods described in Abadie, Diamond and Hainmueller (2010, 2011). Removing England and Wales data, I repeated the synthetic control analyses for each of the other countries to test their deviation from predicted controls in the absence of an exposure and recorded observed and predicted values. Across the observation period I calculated yearly prediction gaps as the difference between observed and predicted values. I calculated mean squared prediction errors (MSPE) for the pre-intervention and post-implementation periods as the mean of the squares of the prediction gaps for each country. These trends for unexposed countries were taken as null-effect placebos, that is, if my

assumptions were correct, then these should show the distribution of deviations from predicted controls consistent with noise in the absence of the Strategy, whilst accounting for common trends. To test whether the effects of the Strategy could be detected as producing greater deviations than the range of produced by noise in placebo countries, I plotted yearly England and Wales prediction gaps alongside corresponding differences calculated for the other countries. I restricted this comparison to placebo series of countries which had a similarly close-fitting pre-intervention period by using countries with less than 5-times the pre-intervention MSPE of England and Wales. This removed placebos for which large post-implementation gaps may have been due to poorly fit models.

Using all comparison countries' placebo predictions and predictors of England and Wales, I calculated MSPE ratios for each country by dividing the MSPE of the post implementation period by the MSPE of the pre-intervention period. I used this to standardise estimates of post-implementation deviation by goodness of pre-intervention fit. This post/pre-MSPE ratio would be greater than 1 if the predictions of post-implementation period showed a greater deviation from observed values than the pre period, with larger ratios representing larger deviations and/or better fit models. I examined the distribution of MSPE ratios and the relative ranking of the England and Wales ratio alongside placebos. This allowed me to check whether England and Wales saw a comparatively large deviation from predicted post-intervention rate, indicative of an exposure effect.

3.5.4.3 Sensitivity analyses

I performed several sensitivity analyses to test the reliability of my models. In initial models, Scotland was consistently highly weighted in the synthetic control. As I highlighted in the introduction of this analysis (section 3.2 above), concerns had previously been raised with using Scotland as a control due to the potential spill over of the Strategy media campaign (Craig, Dundas, *et al.*, 2016; Wellings, Palmer, Wilkinson, *et al.*, 2016). As the aim of this analysis was to provide a comparison with a wider set of countries, I ran additional models using only birth and pregnancy rates for each data set with only Scotland data absent to test whether the hypothesised contamination was masking or reducing the effects of the Strategy. To additionally test for over-reliance on other countries' data, I re-ran models with countries removed from the donor list. I iteratively removed the top-weighted country in each analysis, plotting yearly differences between England and Wales and the new synthetic control, and extracting pre-intervention MSPE for each to test whether results remained consistent as donor countries were removed. Finally, I constructed plots of observed and synthetic rates for models fitted to dummy intervention

dates across 1995-1998 to examine whether the model was robust to shocks in pre-intervention years.

3.5.4.4 Reporting results

For each model I produced line graphs of the observed trends in England and Wales alongside the synthetic control predictions. I additionally plotted the yearly gaps between observed and control for England and Wales compared with placebos, and the distribution of post/pre-intervention MSPE ratios of England and Wales compared with placebos. Animated time-based placebo graphs were published online (available at **phd.andybaxter.me/Synth-analysis**), and frames were combined in facets of static graphs for reporting in this thesis.

For each model I reported the MSPE of the pre-intervention period as a metric of goodness of model fit, the absolute difference in numbers of pregnancies between exposed and synthetic England and Wales, and a p-value calculated from the proportion of countries showing post/pre-intervention MSPE ratios as large as or larger than England and Wales. This represented the probability that a ratio at least as large would be produced in the event that the Strategy had no effect on pregnancy or birth rates.

3.6 Conclusion

Across these three analyses I will be able to draw robust conclusions about the existing evidence for the effects of interventions and changing social circumstances on pregnancy rates, and the effectiveness of the Teenage Pregnancy Strategy as a prominent example of such an intervention. The review will indicate the range of potential common causes which need to be considered when comparing across countries and highlight plausible hypothesised explanations of observed decreasing pregnancy rates. The two natural experimental methods can give a more reliable assessment of the Strategy's contribution than a pre-post analysis of England only by accounting for changes common to England and comparator countries. The use of two methods across two datasets can additionally give an indication of the reliability of the results obtained by allowing for comparison of estimates of effect direction and size to test for consistency or conflict. Overall, these approaches offer a novel combination of research methods to answer important questions about effective approaches to teenage pregnancy prevention and to demonstrate the use of dependable policy evaluation methods.

Chapter 4 Systematic Review Results

4.1 Overview of chapter

In this chapter I report the results of the systematic review identifying plausible hypothesised causes of the observed decrease in teenage pregnancy in the UK.

From 591 articles retrieved through literature searching and 40 additional articles identified through citations, I identify 89 relevant reviews, 62 of which are rated as high-quality – presenting zero or one concerns of risk of bias. All reviews present either weak evidence of the effect of a hypothesised cause on a population's pregnancy rates, or no evidence of an effect.

These reviews are synthesised to produce 20 causal branches, grouping together 34 distinct exposures. Comparing new evidence to the initial tree constructed in Chapter 2 (section 2.5.3), I assess these branches for inclusion in an updated causal tree. After assessment of evidence presented in reviews, I retain 13 branches for further analysis, with some showing stronger evidence of potential to have contributed to the UK's falling teenage pregnancy rates.

4.2 Results

4.2.1 Search and screening

Database searches returned 591 articles after deduplication. Forty further articles were added from internet searches and citations or references to prior reviews (Figure 4-1). From this body of search results, I identified 97 review articles, five of which were reviews of reviews (Swann *et al.*, 2003; Trivedi *et al.*, 2007; Bowring, Lim and Hider, 2016; Shackleton *et al.*, 2016; Bowring *et al.*, 2018). To avoid duplication of findings, all citations from reviews of reviews were screened, and the findings of the reviews of reviews themselves were not used in analyses (one review, Fullerton *et al.*, 1997, conducted searches including systematic reviews but extracted and synthesised study data alongside other idenified primary studies and so was included). Three papers (Underhill, Montgomery and Operario, 2007; Underhill, Operario and Montgomery, 2007b; Aslam *et al.*, 2017) reported the results of fuller reviews published elsewhere (Underhill, Operario and Montgomery, 2007a; Underhill, Montgomery and Operario, 2008; Whitaker *et al.*, 2016) – these pairings of published papers and fuller reviews were each treated as one

review. This produced a total of 89 reviews for inclusion in analysis. All included reviews are listed in Appendix D.

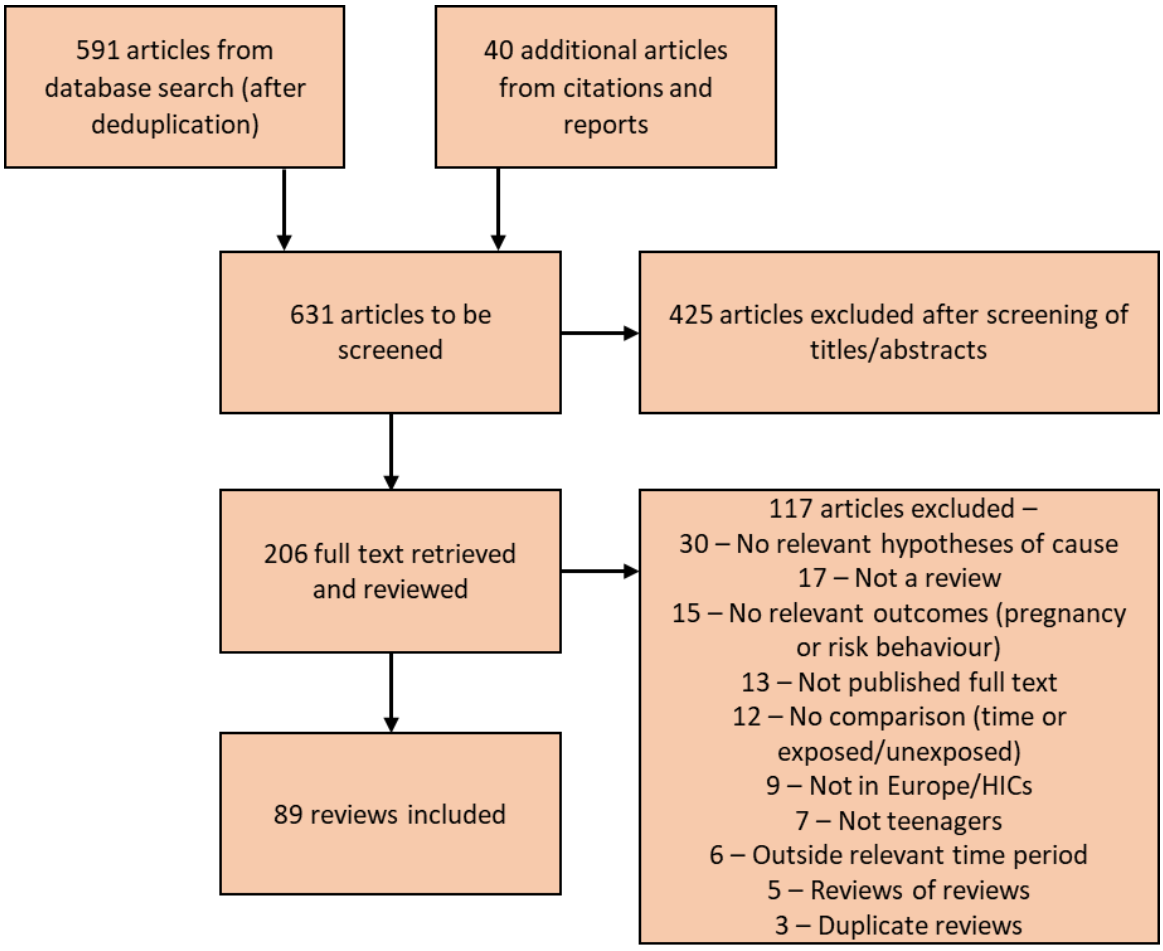

Figure 4-1 – Prisma diagram of search results and screening. Eighty-nine reviews met inclusion criteria.

4.2.2 Quality assessment

Sixty-one of the 89 reviews were rated as ‘high’ quality – reporting zero or one concerns across the four domains of risk of bias. Twenty-eight were rated ‘low’ quality (Figure 4-2). High quality reviews were often explicitly described as systematic reviews, meta-analyses, or another review design using systematic search and synthesis methods. Most low-quality reviews did not use systematic methods for searching, screening and synthesising evidence. Amongst high-quality reviews the most common concern limiting confidence in concluding causal relationships was inadequate accounting for confounding and bias across diverse study types.

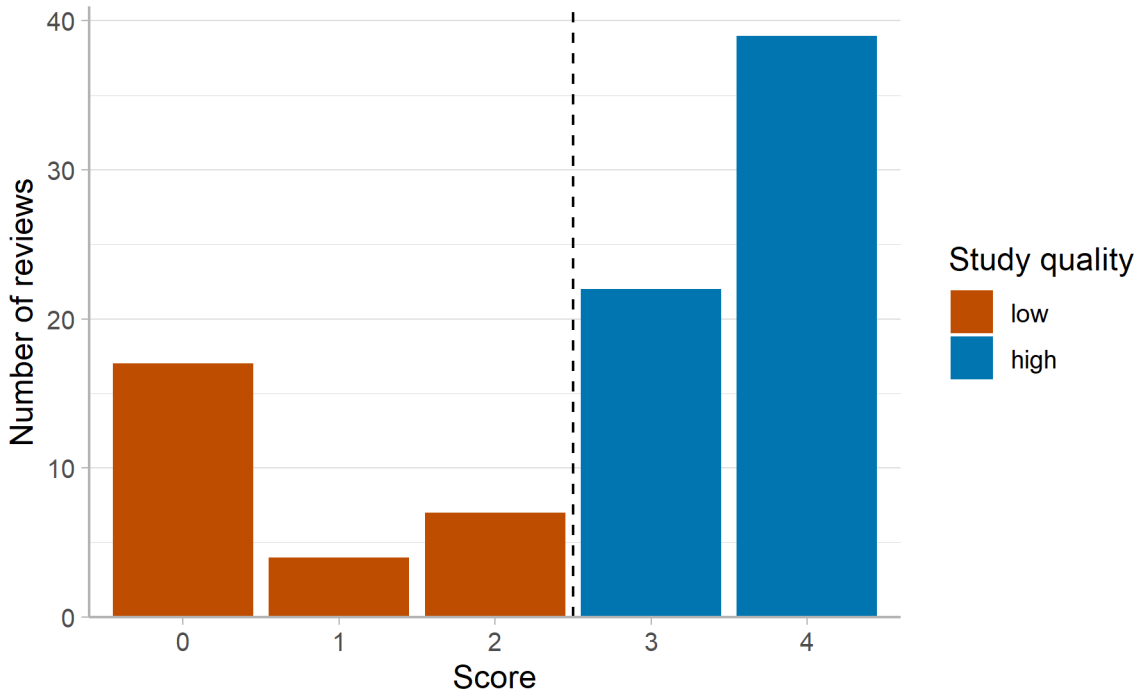

Figure 4-2 – Distribution of study quality across all included reviews. Each review was marked across the four domains of risk of bias and awarded a point if the concern was addressed.

Of reviews rated as ‘low’ quality, the majority were presented as literature reviews which selectively included and synthesised evidence. Sources of evidence were commonly unclear, and the lack of an explicit and systematic search strategy had the potential to miss relevant studies. In three cases, reviews described as systematic were given a low score due to concerns of bias in at least two domains. All quality assessment outcomes for included studies are presented in Appendix E.

4.2.3 Overview of findings

All review conclusions were rated either ‘+’ (weak evidence of positive effect) or ‘0’ (no evidence of effect). No reviews found a consistent negative effect of an intervention. In most reviews presenting weak positive evidence, evidence of some effect was presented in case-control or RCT studies of sample populations (effectiveness), but no assessment of the population reach of an implemented intervention (exposure). In three reviews, exposure was measured across a population, but weak evidence was presented for causation. Ratings of evidence presented in all reviews by branch are presented online at **phd.andybaxter.me/review_data.**

Reviews reported on a total of 30 interventions or environment changes associated with differing pregnancy risks. When compared with the initial causal tree (section 2.5.3), seventeen new elements were identified, and four previously described branches were

divided into sub-categories to reflect distinctions used in included reviews. Four environment changes which were previously highlighted as hypothesised causes were not addressed in the literature. The resulting 32 intervention and environment changes were categorised into 20 distinct branches to be assessed for inclusion in or exclusion from the causal tree (Table 4-1).

Hypothesised Cause – Intervention	High quality reviews reporting:				Low quality reviews reporting:			
	Positive effect	No effect	Negative effect	Total reviews	Positive effect	No effect	Negative effect	Total reviews
School-based interventions								
School-based pregnancy education	2	10	0	12	8	4	0	12
School-based STI-focussed education	9	3	0	12	4	1	0	5
Abstinence-based education	5	15	0	20	4	11	0	15
Clinic-based interventions								
Counselling/medical staff one-to-one	8	3	0	11	4	1	0	5
School-based SH clinic	0	8	0	8	6	2	0	8
Teenager SH clinic access and use	3	7	0	10	6	2	0	8
Contraception interventions								
Advance supply of EC	0	3	0	3	1	0	0	1
Condom promotion/ distribution	2	5	0	7	2	3	0	5
Contraception access (other)	2	2	0	4	0	1	0	1

Hypothesised Cause – Intervention	High quality reviews reporting:				Low quality reviews reporting:			
	Positive effect	No effect	Negative effect	Total reviews	Positive effect	No effect	Negative effect	Total reviews
Contraception initiation follow-up	0	1	0	1	0	0	0	0
Changing contraceptive technologies	3	2	0	5	3	1	0	4
Community-based interventions								
Community-based pregnancy education	1	4	0	5	3	1	0	4
Community-based STI education	2	2	0	4	3	0	0	3
Adolescent development interventions								
Early-years intervention†	3	0	0	3	0	0	0	0
Personal development (inc. volunteer work)†	7	4	0	11	7	1	0	8
Vocational/academic training†	1	1	0	2	3	1	0	4
Digital media-based interventions								
Digital media-based intervention (targeted)	0	2	0	2	0	0	0	0

Hypothesised Cause – Intervention	High quality reviews reporting:				Low quality reviews reporting:			
	Positive effect	**No effect**	**Negative effect**	**Total reviews**	**Positive effect**	**No effect**	**Negative effect**	**Total reviews**
Digital media-based SH intervention	4	1	0	5	0	0	0	0
Education policies (promoting)†	2	0	0	2	3	0	0	3
Family/Community engagement	9	4	0	13	4	2	0	6
Public information/ media campaign	0	0	0	0	2	0	0	2
Peer-contact sexual health intervention	4	9	0	13*	3	2	0	5
Targeting rapid-repeat pregnancies	3	5	0	8	5	2	0	7
Social support interventions								
Targeting vulnerable youth	2	0	0	2	1	2	0	3
Social support†	0	0	0	0	0	1	0	1
Virtual infant simulator	0	2	0	2	0	1	0	1

Hypothesised Cause – Environment change	High quality reviews reporting:				Low quality reviews reporting:			
	Positive effect	No effect	Negative effect	Positive effect	Negative effect	Positive effect	No effect	Negative effect
Later parenting culture	*No reviews*							
Change in education								
Educational/career aspiration/planning	0	0	0	0	2	1	0	3
More years in school across population	0	0	0	0	1	0	0	1
Digital media-driven culture change	*No reviews*							
Alcohol use – changing patterns	0	0	0	0	1	0	0	1
Employment outside school hours†	0	0	0	0	1	0	0	1

*denotes newly added hypothesised branch; * Fourteen high quality reviews looked at evidence for 'peer-contact sexual health interventions'; however, one review* (Blank *et al.*, 2010b)**,** *assessing only one study, concluded that the results were unreliable, and so this review's findings were not included for this outcome. † These interventions targeted antecedents of teenage pregnancy, rather than setting pregnancy prevention as a primary aim. STI – Sexually-Transmitted Infection; SH – Sexual Health; EC – Emergency Contraceptives*

Table 4-1 – Twenty identified branches of interventions or exposures hypothesised to cause reductions in teenage pregnancy rates, subdivided into 32 distinct modes of delivery and summarising number of reviews presenting evidence of positive, negative or no effects on pregnancy. Hypotheses not previously identified are highlighted by blue borders preceding relevant rows. Branches were identified prior to literature searches (n=8 branches) or through the review literature (n=12).

4.2.4 Existing branches with supporting evidence

4.2.4.1 School-based sex and relationships education (SRE)

Thirty-two reviews, 19 of which were of high-quality, looked at school-based sex and relationships education programmes. During extraction, these were divided into pregnancy-focussed programmes (where avoiding pregnancy was a stated goal, alongside associated sexual risk behaviours; reported in 12 high-quality reviews and 12 low-quality reviews; Table 4-1) and STI-focussed programmes (where STI prevention was stated and presented to participants as an alternate or additional risk besides pregnancy; 12 high-quality reviews and 5 low-quality reviews).

Two high-quality reviews presented meta-analyses of pregnancy-focussed interventions, with all studies conducted in the USA and Canada. These reviews gave the most comprehensive overview of effectiveness. DiCenso et al. (2002) reported no difference in age of sexual initiation, contraceptive use or pregnancy in studies published from 1987 and 2001. The more recently published review, Marseille et al. (2018), similarly found no clear evidence of effectiveness attributable to school-based SRE in studies published from 1987-2016. This review did identify several potentially effective adaptations which could be attached to school-based programmes: female-only sessions and linked school and community approaches. Several reviews identified isolated effective examples of school-based pregnancy education. Lopez et al (2009; 2016), focussing exclusively on RCTs of theory-based interventions, found several approaches which were effective in improving contraceptive use. Manlove et al. (2015) and DiCenso et al. (1999) highlighted several interventions which produced positive results in contraceptive use (in each case as a minority of included studies), though with few clear indicators of elements contributing to their success and few observations of effects on pregnancy itself.

In reviews of STI-focussed interventions, there was more consistent evidence of effectiveness and a greater number of effective programmes identified, all conducted in the USA and the UK (Kirby, 2001, 2002a; Blank *et al.*, 2010a; Cardoza *et al.*, 2012; Manlove, Fish and Moore, 2015; Lopez, Bernholc, *et al.*, 2016). This finding is consistent with previous commentators suggesting STI interventions are more effective at addressing a range of sexual health outcomes (Kirby, 2003). This outcome may be an indicator of more sensitive or robust study design, better designed and funded programmes, or the greater effects of the presentation of a health risk of universal aversion to both males and females (Kirby, 2002a, 2003).

Overall, there is some evidence that school-based approaches can reduce pregnancy risks, however this category represents a diverse range of programme contents, and interventions were tested in relatively few countries. Whilst some programmes may have delivered positive change, no reviews presented evidence that school-based interventions had contributed to population-level changes in teenage pregnancy rates.

4.2.4.2 Clinic-based interventions

Clinic-based interventions were classified into three types: improved access to sexual health clinics, appointments with medical staff, and school-based clinics. These intervention types exhibited varying levels of specificity of targeting teenagers and intensity of engagement. Thirty-four reviews, twenty-two of which were high quality, examined at least one form of clinic-based intervention (Table 4-1).

Interventions and policies to provide more general access to sexual health clinics for teenagers in community and primary care settings were assessed in 11 high-quality reviews with mixed findings. These interventions were assessed for effects on pregnancy rates and sexual health behaviours. A meta-analysis across clinic-based interventions showed no evidence of positive effect on contraception use or reducing risk of pregnancy (DiCenso *et al.*, 2002; five studies, all conducted in the USA). Several further reviews assessing only one or two studies also found little additional benefit of new clinics or improved access or engagement in preventing pregnancy across several European and North American countries (Lazarus *et al.*, 2010; Wakhisi *et al.*, 2011; Blank *et al.*, 2012). However, two reviews did find positive effects in increasing contraceptive use and reducing pregnancies of clinic-based interventions in the USA and Canada (Franklin *et al.*, 1997; Manlove, Fish and Moore, 2015). Narrative synthesis of qualitative studies of the views of adolescents by Baxter et al. (2011) indicated that improvements to the accessibility and confidentiality of clinics could lead to greater use, supporting their hypothesised contribution to lowered risk of pregnancy.

Several reviews distinguished general clinic contact from more targeted and intense engagement with participants (Kirby, 2001). Appointments or interventions involving medical staff engaging one-to-one with teenagers at risk were assessed in eleven high-quality reviews. Most reviewers found evaluations of successful programmes conducted in several countries, with some overlap across reviews. These reviews concluded that there was some evidence of their effectiveness in reducing teenage pregnancy risk (Kirby, 2001, 2002a; Lin *et al.*, 2008; Blank *et al.*, 2012; Cooper *et al.*, 2014; O'Connor *et al.*, 2014;

Manlove, Fish and Moore, 2015; Zapata *et al.*, 2015). Such interventions were highly targeted, and evidence was predominantly presented at the level of individual impact.

As a contrasting, population-wide intervention, eight high-quality reviews reported on the effectiveness of school-based sexual health clinics. None found consistent evidence of a positive effect on sexual behaviours or pregnancy (Kirby *et al.*, 1994; Fullerton *et al.*, 1997; Kirby, 2001; Matthias, 2002; Blank *et al.*, 2010a; Lazarus *et al.*, 2010; Wakhisi *et al.*, 2011; Mason-Jones *et al.*, 2012). Whilst some studies observed increasing contraception provision through schools, this was hypothesised by authors of reviews as constituting a "substitution effect", with teenagers obtaining similar levels of contraception from the clinic as they previously had from other sources (Kirby *et al.*, 1994; Kirby, 2001).

Overall, there was mixed evidence to support the hypothesis that increasing clinic-based interventions and contacts with adolescents contributed to declining pregnancy rates. Interventions targeting teenagers at greater risks of pregnancy in one-on-one medical appointments may produce behaviour change leading to lowered risk of pregnancy. However, no evidence was presented for the scale of exposure across a population. Population-wide interventions, such as school clinics, had comparatively little effect.

4.2.4.3 Changing contraceptive technologies (substituting for non-condom methods)

Nine reviews, five of which were high quality, assessed the impact of the promotion and use of non-condom contraceptive methods (Table 4-1). These methods include hormonal contraceptives (pills, patches or injections), various forms of long-acting reversible contraceptives, and emergency contraceptives (Glasier and Wellings, 2012). I excluded advance supply of emergency contraceptives, which I classed in section 4.2.5.2 below as contraception-promotion rather than a contraceptive method exchange intervention. Reviews considered evidence from a range of countries in Europe and North America. These reviews reported three forms of changes in contraceptive use: effectiveness of specific non-condom contraceptive methods when used by adolescents, promotions of these methods as more effective alternatives to condom use and uptake of emergency contraceptive use.

Two high-quality reviews compared alternative contraceptive methods. Tang et al. (2012) were not able to draw conclusions regarding efficacy of these methods due to small sample sizes (studies conducted in USA, Sweden and Finland). Usinger et al. (2016) found that

intrauterine devices had higher continuation rates than oral contraceptives and hormonal injections and patches, though similar to hormonal implants (studies conducted in USA and Sweden).

Interventions to promote use of contraceptives other than condoms were examined in two high quality reviews. Blank et al. (2012) found weak evidence that they contributed to greater contraceptive adherence and lower pregnancy rates (across the USA, UK, Canada and Sweden). Dean et al. (2014) reported no clear effects of school-based provision of improved contraception (USA only).

In one high-quality review, observations of emergency contraceptive (EC) uptake across three populations (France; British Columbia, Canada; and the UK) indicated that policy changes allowing over-the-counter access to these resulted in greater uptake from pharmacies and increased use (Denno, Chandra-Mouli and Osman, 2012). In the UK, the increase in pharmacy provision was matched by decreases in prescription of ECs, and so may not represent greater use. It is also uncertain from this data whether this contributed to greater or more effective overall contraceptive use or reduction in pregnancy rates.

Across all examined contraceptive technologies, though direct comparison of pregnancy rates with users of condoms alone was not possible, these results do indicate that promotion of more effective contraceptive methods may result in reduced risk of pregnancy. Similarly, evidence of population exposure was found. Together, these reviews satisfy both categories for strong overall evidence of the contribution of this hypothesised cause to observed falling population rates (Table 4-2), although causal effects on pregnancy were not tested at a population level.

4.2.4.4 Change in education aspiration and involvement

Three reviews, all of which were low-quality, reported environment changes involving education and their hypothesised effects on teenage pregnancy (Table 4-1). One review noted that increases in years of schooling amongst adolescent girls across populations over time were associated with falling teenage pregnancy rates (Kirby, 2002b), suggesting a population-wide effect but not directly testing causation. Three reviews looked at educational and career aspirations' effects on pregnancy, with two finding some evidence of an effect on pregnancy risk (Kerr and Matlak, 1998; Kirby, 2002b). The third review assessed qualitative studies and highlighted that teenagers may contemplate education and career risks when making decisions to avoid risk of pregnancy, but did not directly test

association (Spear and Lock, 2003). One additional review, Harden et al. (2006), presented narrative syntheses of teenagers' views expressed in included studies which link low future aspirations and expectations to greater risk of pregnancy. However, the review did not present evidence of a change in education goals over time.

Across all reviews, these findings are compatible with the hypothesised contribution of increasing education aspiration and involvement to lowering pregnancy rates. However, the presented evidence did not test a causal relationship.

4.2.5 Existing branches with evidence of little effect or not assessed in literature

4.2.5.1 Abstinence-based education

Thirty-five reviews, twenty of which were high-quality, examined evidence for interventions with an abstinence promotion aspect (Table 4-1). Most studies were conducted in the USA. This category included both abstinence-only (promoting abstinence from sex until marriage as the only positive option) and abstinence-plus interventions (promoting safer sex for sexually active teenagers as secondary to abstinence as best practice; Santelli *et al.*, 2006; Ott and Santelli, 2007; Underhill, Operario and Montgomery, 2007b). This review focused on whether the common abstinence-promotion component was effective, as not all reviews presented separate syntheses for each intervention type.

Meta-analyses of abstinence-based programmes were in several cases consistent with null effects of abstinence programmes when compared with treatment as usual (DiCenso *et al.*, 2002; Marseille *et al.*, 2018). Chin et al. (2012), however, reported significant reductions in sexual activity only in non-randomised trials and significantly increased risk of pregnancy across nine studies. Some evidence suggesting positive effects on sexual risk behaviours and pregnancy in individual programmes was found in several reviews (Blank *et al.*, 2010a; Cardoza *et al.*, 2012; Manlove, Fish and Moore, 2015; Lopez, Bernholc, *et al.*, 2016). However, the majority of reviews reported results consistent with no effect of abstinence-based education (Franklin *et al.*, 1997; DiCenso, Guyatt and Willan, 1999; Kirby, 2001; Blank *et al.*, 2010a; Johnson *et al.*, 2011). Several reviews found stronger evidence for the effectiveness of abstinence-plus programmes than abstinence-only (Bennett and Assefi, 2005; Underhill, Operario and Montgomery, 2007a; Underhill, Montgomery and Operario, 2008). This differentiation would indicate that the additional

component including contraception education is responsible for the majority of improved outcomes and that the abstinence component itself has little impact on pregnancy risk behaviours.

4.2.5.2 Contraception access interventions

Four types of intervention were categorised as promoting and providing access to contraceptives. These were classified into three types: condom promotion and distribution, advance provision of Emergency Contraceptives (ECs), improved access to other contraceptives, and medical staff following-up contraception initiation. Eighteen reviews, eleven of which were high-quality, examined at least one contraception access intervention (Table 4-1).

Seven high-quality reviews looked at condom promotion and distribution interventions. A meta-analysis of four youth-focussed studies in the USA found positive but uncertain evidence of their effectiveness on condom use (Pooled OR: 1.35; 95%CI: 0.94-1.96; I^2 = 83%; Charania *et al.*, 2011). Other studies across the USA and Canada showed mixed results, with some cases of decreased use of condoms or replacement of other contraceptives, and several cases of no effect on contraception use or pregnancy (Kirby, 2001; Andrzejewski, Liddon and Leonard, 2019). Successful condom-promotion interventions were found in North America when combining with counselling sessions (Blank *et al.*, 2012) or HIV prevention sessions (Cardoza *et al.*, 2012). These reviews indicate that condom distribution itself has little effect on risk of pregnancy and that any effects may be attributable to other elements of such interventions.

Three high-quality reviews examined the effects of advance provision of ECs on adolescent use of EC, other sexual risk behaviours and pregnancy rates (Meyer, Gold and Haggerty, 2011; Blank *et al.*, 2012; Rodriguez *et al.*, 2013; all studies conducted in North America, UK or Sweden). These reviews were published in successive years and there was a substantial overlap of included studies, with four of eight randomised trials reviewed across all three. I classed the advance provision of ECs as distinct from the promotion and use of more effective methods (section 4.2.4.3 above), as advance provision interventions were proposed as supplementing rather than replacing other contraceptive methods (Meyer, Gold and Haggerty, 2011). Across all three reviews there was consistent evidence that advance supply of ECs increased their likelihood of use and reduced the time between intercourse and use of ECs (thus theoretically increasing their efficacy). Despite this, no studies reporting pregnancy rates found a clear difference between EC-supplied groups and

controls, suggesting little effect on risk of pregnancy overall. Most studies reported little evidence of negative effects on other contraceptive use or risk behaviours, with improvement in some cases. One study, however, suggested that adolescents supplied with EC may engage in unprotected sex more often (Rodriguez *et al.*, 2013). The study experienced substantial loss to follow up at 12 months and was rated poor quality evidence.

Interventions to promote use of or improve access to other contraceptives (not including condoms; including EC access upon request but not advance provision) were reviewed in four high-quality reviews. Policies increasing access to emergency contraceptive pills in pharmacies, not requiring a prescription, led to increased uptake of ECs in several countries (Denno, Chandra-Mouli and Osman, 2012). However, in the UK context this was offset by a decrease in ECs prescribed to adolescents and so doesn't provide clear evidence of increased use. In several cases, access programmes led to greater use when combined with other approaches, such as counselling in clinic settings (Kirby, 2001; Blank *et al.*, 2012), or targeting vulnerable youth (Denno, Chandra-Mouli and Osman, 2012). There was little overall evidence that improved access to other contraceptive methods led to reduced risk of pregnancy.

One review looked at follow-up interventions to enhance contraceptive-use continuity but found no clear evidence of any effects across four studies (Steenland *et al.*, 2013).

Across the four types of interventions aiming to improve contraception access and use, there was little evidence of their effectiveness in decreasing pregnancies across a population. Where interventions were effective, they were commonly combined with other intensive and targeted counselling from medical staff. This suggests that improved contraceptive provision itself has little explanatory power for population-level observations of decreasing teenage pregnancy rates. This has been suggested by some as due to a substitution effect, whereby adolescents who were motivated to use contraceptives were already capable of accessing these and new interventions tended to alter the source of contraceptives rather than increase their use.

4.2.5.3 Digital media use (environment change)

No studies examined associations between environmentally driven changes in digital media use and teenage pregnancy rates.

4.2.5.4 Later parenting culture

No studies examined associations between cultural changes in perceptions of teenage parents or norms of parenting at later ages and teenage pregnancy rates.

4.2.6 Additional branches presenting evidence of effectiveness

4.2.6.1 Community-based interventions

Twelve reviews, six of which were high-quality, examined interventions conducted in a community setting (Table 4-1), with targeted or invited youth participants attending events or meetings. Similar to school-based interventions, these were categorised into two groups of pregnancy-focussed and STI-focussed (with or without combined pregnancy-focus) interventions. All included studies were conducted in North America.

Interventions which focussed exclusively on pregnancy prevention were examined in five reviews, with limited evidence of effectiveness. Four reviews identified successful programmes (DiCenso, Guyatt and Willan, 1999; Kirby, 2001; Wakhisi *et al.*, 2011; Cardoza *et al.*, 2012) but the majority of studies (and earlier versions of similar programmes) produced inconsistent effects or no observable change. Two reviewers noted that interventions with intense and prolonged exposure appeared to be more successful and suggested that these factors may be necessary for continuing effectiveness (DiCenso, Guyatt and Willan, 1999; Kirby, 2001). In one review, several community approaches were combined with school-based approaches (categorised in this review as both school-based and community-based) but produced inconsistent results, with some interventions reporting reductions in contraception use (Wakhisi *et al.*, 2011).

Four reviews included studies evaluating community-based interventions focussing on STI prevention or combining STI and pregnancy prevention and found stronger overall evidence of effectiveness. Lopez et al. (2009) focussed exclusively on RCTs evaluating theory-based approaches and reported interventions resulting in fewer pregnancies and increased contraceptive use. Cardoza et al. (2012) reported increased condom ownership and increased sexual abstinence from one study. Chin et al (2012), aggregating both intervention types, reported overall effectiveness in several areas of reducing sexual risk behaviours but no significant effects on pregnancy.

Overall, these reviews suggest it is possible to deliver effective interventions in community settings, though with mixed evidence of the benefits of the community approach itself.

Similar to school-based interventions, more consistent evidence was found for the effectiveness of STI-focussed than pregnancy-focussed interventions. No evidence was presented in reviews assessing population-wide exposure to community-based interventions. The suggested requirement of high intensity and long duration for community programmes to achieve their goals however may make such interventions less feasible for population-wide implementation.

4.2.6.2 Adolescent development interventions

Twenty reviews, twelve of which were high quality, reported on interventions intended to address wider goals of improving adolescents' social circumstances (Table 4-1). In several cases pregnancy prevention was noted as an aim, but predominantly such interventions had other primary goals.

In eleven high-quality reviews, evidence was assessed for interventions to improve adolescents' personal development involving such elements as skills building, volunteering experience, future planning, self-esteem and assertiveness training. The review with the widest scope, conducted by Gavin et al. (2010), focussed exclusively on youth development programmes addressing more distal social determinants, with less than 50% of participant contact time focussed on addressing proximal causes of sexual risk behaviour. They identified a large number of programmes which were effective in reducing sexual risk or pregnancy (conducted in the USA and Australia; Gavin *et al.*, 2010). Other reviews included smaller subsets of these studies, similarly concluding that there was evidence of effectiveness in some programmes (DiCenso, Guyatt and Willan, 1999; Kirby, 2001, 2002a; DiCenso *et al.*, 2002).

Two high-quality reviews looked at studies evaluating vocational and academic training programmes. Parts of these were often combined with more general development interventions (addressed above). There was some positive evidence for positive impacts of interventions including vocation and academic training and support across the USA, Canada and the UK (Fullerton *et al.*, 1997; Harden *et al.*, 2006), but no testing of the specific contributions of these elements. Further low-quality reviews highlighted these and similar programmes and studies, drawing positive conclusions about their potential to be effective (Kirby and Coyle, 1997; Nitz, 1999; Fletcher *et al.*, 2007).

Three high-quality reviews looked at the long-term effects of early years interventions, including studies in the USA, Canada and the UK. Three programmes were identified and

meta-analysed by Harden et al. (2006), with two of these included in Kirby et al. (2001) and Zoritch et al. (1998). In meta-analyses across all studies there was some evidence of effectiveness in reducing the risk of subsequent teenage pregnancies (Harden *et al.*, 2006). Two pre-school programmes (the Abecedarian Project and the Perry Preschool Project) were reported as reducing rates of teenage pregnancy amongst relatively small sample sizes (n = 111 and n = 123 respectively; neither showing significant results). One programme targeted school-aged children with a larger sample (the Seattle Social Development Program; n = 808) found positive, but not significant, reductions in pregnancy rates. All studies reported pregnancy rates amongst teenagers who had undergone interventions outside of the review period (in the 1960s, 1970s and 1980s). Such interventions may have contributed to lowering rates within the review period (after 1990) due to delayed effects or repeated programmes, but this was not assessed in included reviews.

Whilst no evidence was presented for a population-wide exposure or impact of such adolescent development interventions on teenage pregnancy, there is some evidence for the positive impacts of these interventions upon the antecedent social determinants of pregnancy. Wider policies targeting these but not specifically addressing pregnancy, therefore not assessed here, could plausibly have contributed to observed decreases in pregnancy rates.

4.2.6.3 Digital media-based interventions

Seven reviews, all of which were high-quality, examined sexual health interventions conducted using digital media devices to communicate with or disseminate information to adolescents (Table 4-1). Examined studies were conducted in the UK, USA, Australia and the Netherlands.

Five reviews looked at a range of interventions using digital media to deliver sexual health promotion messages to groups of teenagers using games and interactive activities, widely-distributed messages or enabling access to information (Guse *et al.*, 2012; Hieftje *et al.*, 2013; McLellan and Dale, 2013; Jones *et al.*, 2014; Widman *et al.*, 2018). Several interventions were identified across these reviews which were effective in promoting behaviour change. The majority of studies reported more proximal measures of knowledge or attitude changes, which were not relevant for this review. One review (Jones *et al.*, 2014) highlighted several contradictory effects on behaviours and was classed as not showing consistent evidence of benefits. Widman et al. (2018) meta-analysed condom use

and abstinence, finding positive effects on both outcomes. However, countries not relevant to this review were included in both analyses, limiting the generalisability of the effect size to target countries.

Two reviews looked at interventions targeting vulnerable youth in high-engagement settings or through enrolment in a sexual health messaging service (DeSmet *et al.*, 2015; L'Engle *et al.*, 2016). Neither review found consistent evidence of a positive effect in pregnancy prevention. No data was reported indicating wider population rollout or exposure to these interventions.

4.2.6.4 Policies promoting education

Policies promoting educational engagement, achievement or prolonged attendance were discussed in two high quality reviews (Harden *et al.*, 2006; Dean *et al.*, 2014) and three low quality reviews (Table 4-1; Kirby, 2002b; Beltz *et al.*, 2015; Fleming *et al.*, 2015). Amongst high-quality reviews, Harden et al. (2006) found evidence that early-years educational interventions were successful in preventing pregnancy risk, and hypothesised from qualitative interviews with teenagers that this was mainly due to their motivation to avoid pregnancy (in the USA, UK and Canada). Dean et al. (2014) looked at two studies focussing on providing access to education for teen mothers after birth, which were both effective in reducing the risk of rapid repeat pregnancies. To supplement these limited findings, data from three low-quality reviews was further considered. Kirby (2002b), hypothesising several possible mechanisms of action, highlighted evidence that time in education is linked to reduced pregnancy risk and discussed a study evaluating a successful intervention aiming to increase school attachment (resulting in lower pregnancy risk). Beltz et al. (2015) found five studies evaluating education policy changes and concluded that these present reasonable evidence for their effects on lowering pregnancies across state populations. Fleming et al. (2015) additionally discussed supported return to education as a protective factor against rapid repeat pregnancies.

Taken together, the above evidence from both high- and low-quality reviews suggests a positive effect of policies promoting education in reducing teenage pregnancy risks. They draw out several possible mechanisms of action which strengthen this causal inference. Data recording effects of population-wide exposure further add to the strength of evidence.

4.2.6.5 Family and community engagement

Twenty reviews, thirteen of which were high quality, reported a range of family and community engagement interventions (Table 4-1). These interventions were considered distinct from community-based education interventions as they were conducted as open, community-engagement focussed programmes rather than attendance at a series of sessions. The aim of achieving pregnancy reduction through improving long-term support from parents and through community relationships was often explicitly stated (Kirby, 2001).

Three reviews which specifically aimed to assess parent-focussed interventions found these to be effective in 50% to 100% of included studies (all conducted in the USA; Wight and Fullerton, 2013; Sutton *et al.*, 2014; Gavin *et al.*, 2015), whilst one reported inconsistent results (Downing *et al.*, 2011). Further reviews with wider scope highlighted evidence of effective family and parent programmes (Wakhisi *et al.*, 2011; Manlove, Fish and Moore, 2015). Parent-focussed appeared to be more effective than family-engagement (Downing *et al.*, 2011; Sutton *et al.*, 2014), and culturally-targeted approaches may be more effective than generic programmes (Cardoza *et al.*, 2012; Sutton *et al.*, 2014). Improving parent-adolescent communication was highlighted as a mechanism with evidence presented for the effects of the interventions via this proposed mediator (Downing *et al.*, 2011; Gavin *et al.*, 2015). In one review, it was suggested that comparisons of parent-delivered interventions with healthcare worker-delivered interventions indicated similar levels of effectiveness and were more effective than no intervention (Downing *et al.*, 2011). Though community-engagement programmes were reported with positive results in several cases (Cardoza *et al.*, 2012), the approach of building community relationships was less commonly used than family-focussed interventions and had mixed results (Kirby, 2001).

Overall, reviews reported relatively strong evidence for the potential effectiveness of parent- and family- focussed interventions, through the hypothesised mechanism of improving parent-adolescent communication to avoid sexual risks. All studies showing effects on pregnancy and pregnancy-risk behaviour were conducted in the USA, and no evidence was presented for a population-wide effect of an implemented intervention.

4.2.6.6 Targeting rapid repeat pregnancies

Fifteen reviews, eight of which were of high quality, looked at interventions targeting rapid repeat pregnancies by engaging with adolescent women with a previous pregnancy (Table 4-1). Aslam et al. (2017) conducted meta-analyses across nine studies, split by intervention

type (‘home visit’, ‘community’ and ‘telephone in person’) and saw no clear effects on repeat pregnancy in any of the three intervention types or in any individual study. Dean et al. (2014) saw a clear significant pooled reduction in repeated pregnancies across sixteen studies (two studies overlapping with Aslam *et al.*, 2017), though this analysis included papers published outside of the time scale of this review. They additionally highlighted three examples of effective programmes in narrative synthesis, all evaluated after 1990 (Dean *et al.*, 2014). Some evidence of effective programmes in the USA was presented in two additional reviews (Blank *et al.*, 2010a; SmithBattle *et al.*, 2017), but results were mixed or unclear in other reviews and in studies conducted in other countries (Blank *et al.*, 2012; Rodriguez *et al.*, 2013; Lopez, Grey, *et al.*, 2016).

Overall, there was insufficient evidence to include or exclude interventions targeting repeat adolescent pregnancies as a plausible explanation of decreasing pregnancy rates. Although targeted at a high-risk subgroup rather than all teenagers, repeat pregnancies were estimated to account for one fifth of adolescent pregnancies in the UK (Aslam *et al.*, 2017); effective interventions targeting this vulnerable group could potentially produce detectable effects across a population. However, no evidence was presented in these reviews for population-wide implementation of such interventions.

4.2.6.7 Employment whilst attending school

One low-quality review examined the effects of youth employment during school holidays and outside school hours on pregnancy (Table 4-1; Nitz, 1999). Within this review, one study presented the results of a programme conducted in the USA to provide jobs to vulnerable youth with the primary aim of addressing employment and education difficulties as determinants of pregnancy risk. They found a decreased risk of pregnancy amongst participants. However, study quality was not assessed. No reviews looked at employment rates as a non-intervention environment change.

4.2.7 Additional suggested branches with limited evidence or evidence of little effectiveness

4.2.7.1 Peer-contact intervention

Eighteen reviews, fourteen of which were of high quality, examined studies reporting peer-contact interventions’ effects on pregnancy (Table 4-1). These interventions aimed to disseminate information, promote health behaviour through cultural change, or enable access to contraception through a recruited and trained adolescent ‘peer supporter’. Meta-

analyses in Marseille et al. (2018) and Kim and Free (2008) found no clear evidence of the effectiveness of peer-led interventions observed across several countries (the USA, UK, Canada and Italy). Across several reviews, some examples of effective peer led programmes were highlighted. These were assessed either as parts of multicomponent interventions compared to usual care or as peer-led compared to adult-led interventions (Kirby *et al.*, 1994; Wakhisi *et al.*, 2011; Brittain *et al.*, 2015; Lopez, Bernholc, *et al.*, 2016; Lopez, Grey, *et al.*, 2016). Overall, however, there appears little consistent evidence of the effectiveness of the peer-led component itself across most studies and contexts (DiCenso, Guyatt and Willan, 1999; DiCenso *et al.*, 2002; Lazarus *et al.*, 2010; Tolli, 2012).

4.2.7.2 Virtual infant simulators

Interventions using a computerised infant simulator, conducted in the USA, were examined in two high-quality reviews (Table 4-1; Blank *et al.*, 2010a; Cardoza *et al.*, 2012). This approach uses simulators (robotic dolls) to mimic the tasks of parenthood and develop adolescents' understanding of the costs associated with pregnancy. The reviews reported on three studies using the "Baby Think It Over" programme. Blank et al. (2010a) found no conclusive evidence of effectiveness across programme aims; measured outcomes were not explicitly stated in the review. Cardoza et al. (2012) reviewed a study recording increased reported use of contraceptives (from 24 to 31 of 109 participants), but this was not measured against control or tested for significance. A low-quality review also reported from this study an increase in participants wanting to be a teenage parent from 12% before intervention to 15% after (Hoyt and Broom, 2002), though did not report on measures of behaviour change. There is little evidence across all reviews of substantial effects of this intervention on pregnancy and pregnancy-risk behaviours and none reported wider implementation across a population or amongst high-risk groups.

4.2.7.3 Public information or media campaign

Two reviews, both low-quality, reported evidence for public health information or public media campaigns to reduce pregnancy risk (Table 4-1). Grunseit et al. (1997) highlighted two studies evaluating public information campaigns (conducted after 1990), one showing a large increase in condom use, but neither compared with control. Kirchengast (2012) highlights a population wide intervention which included a national information campaign (the Teenage Pregnancy Strategy) but concludes there is little evidence of the intervention's overall effectiveness. Overall, little evidence, of low quality, was found for the impact of public information and media campaigns.

4.2.7.4 Social support interventions

One low-quality review examined the effects of policies providing benefits support for vulnerable families on state-level teenage pregnancy rates (Table 4-1; Beltz *et al.*, 2015). This review did not find strong or consistent evidence of the effectiveness of these policies, concluding that results were mixed. Limited evidence from two high-quality reviews suggests the targeting of prevention efforts at vulnerable subsections of the adolescent population may have enhanced effectiveness. Kirby (2002a) reports that studies targeting more at risk youth with HIV-prevention and sexual health interventions had potentially more impact. Johnson et al. (2011) also report greater effectiveness amongst institutionalised youth. Neither review presented evidence that effective approaches in targeted groups was able to contribute to observed population-level changes in rates.

4.2.7.5 Adolescent culture – alcohol use

One low-quality review examined US-based surveys recording alcohol use and sexual behaviour (Table 4-1) and reported evidence that alcohol use amongst teenagers was associated with both engagement in sexual activity and lack of contraceptive use (Kerr and Matlak, 1998). The mechanism hypothesised for this was impaired risk perception informing sexual behaviour. The review did not present evidence of changing patterns in reducing alcohol use over time or test causal assumptions, so was considered weak evidence of this factor's effects on population pregnancy rates.

4.2.8 Reviewing the causal tree

4.2.8.1 Summarising the evidence

Evidence for the effects of each branch on a population's teenage pregnancy rate are presented in Table 4-2. Overall scores were determined from consistent evidence of an effect (across domains of Direct, Mechanistic/Coherence and Replicability), giving one '+', and evidence of their scalability to a whole population similar to the UK (Exposure), giving an additional '+'. Direct evidence was given a greater weight, as robust tests of direct effects would provide a reliable estimate of effectiveness even in the presence of conflicting evidence assessing Mechanistic/Coherence and Replicability domains. Where evidence was found indicating no effect, a '0' was recorded for overall evidence. Two interventions were assigned '++' denoting strong overall evidence of the potential of such approaches to have contributed to changes in a population's pregnancy rates. Changing contraceptive technologies and policies promoting education showed both strong evidence of effectiveness and some observations of population-wide exposure.

Branch	Direct *Data are presented which show an effect while controlling for plausible confounding*	Mechanistic/ Coherence *A logic model or narrative synthesis of mechanisms is presented; AND The model is coherent with what is currently known*	Replicability *Results are seen consistently in different national contexts*	Exposure *Observed across the population of a whole nation; OR Observed amongst a large proportion of the population, or a high-risk group; OR Noted by authors to have been applied to a population other than the observed groups*	Overall*
School-based interventions	+	+	+		+
Abstinence-based education	**0**				**0**
Clinic-based interventions	+	+			+
Contraception access interventions	**0**	+	+	+	**0**
Changing contraceptive technologies	+	+	+	+	++
Community-based interventions	+	+			+
Adolescent Development interventions	+	+	+		+

Branch	Direct *Data are presented which show an effect while controlling for plausible confounding*	Mechanistic/ Coherence *A logic model or narrative synthesis of mechanisms is presented; AND The model is coherent with what is currently known*	Replicability *Results are seen consistently in different national contexts*	Exposure *Observed across the population of a whole nation; OR Observed amongst a large proportion of the population, or a high-risk group; OR Noted by authors to have been applied to a population other than the observed groups*	Overall*
Digital media-based interventions	+	+	+		+
Education policies (promoting)	+	+	+	+	++
Family/Community engagement	+	+			+
Public information/ media campaign	**0**			+	**0**
Peer-contact sexual health intervention	**0**	+			**0**
Targeting rapid-repeat pregnancies	+	+			+
Social support interventions					**0**
Virtual infant simulator	**0**	+			**0**

Branch	Direct *Data are presented which show an effect while controlling for plausible confounding*	Mechanistic/ Coherence *A logic model or narrative synthesis of mechanisms is presented; AND The model is coherent with what is currently known*	Replicability *Results are seen consistently in different national contexts*	Exposure *Observed across the population of a whole nation; OR Observed amongst a large proportion of the population, or a high-risk group; OR Noted by authors to have been applied to a population other than the observed groups*	Overall*
Later parenting culture	*No evidence presented*				**NA**
Change in education	+				+
Digital media use (environment change)	*No evidence presented*				**NA**
Alcohol use – changing patterns	+				**0**
Employment outside school hours	+				+

*denotes newly added hypothesised branch; *Branches are assigned one '+' for presentation of clear, positive evidence in at least one of 'Direct', 'Mechanistic/Coherence' or 'Replicability', denoting evidence suggesting the effectiveness of the approach, and one '+' if presenting clear evidence in the 'Exposure' category, suggesting that a whole-country effect may be present in literature. If strong evidence suggesting a null effect was presented in any category across reviews, this was marked as a '0' in that column and the overall effectiveness rated '0'.*

Table 4-2 – Summary of overall evidence presented, categorised by causal tree branch. Hypotheses not previously identified are highlighted by blue borders preceding relevant rows.

4.2.8.2 Branches removed

Two intervention branches presented little overall evidence for effectiveness across a large number of reviews and studies. A large number of reviews assessed abstinence-based education programmes and together presented relatively little evidence of their effectiveness in reducing pregnancy risk. Where abstinence-based programmes were individually found to be effective, they tended to incorporate elements shared with other causal branches ('abstinence-plus', including contraceptive education). Consequently, I deemed the abstinence component to provide no explanatory power for observed effects and removed this branch.

Across all interventions targeting promoting contraceptive access and use, the majority of reviews did not find clear evidence of effects on pregnancy risk. As this was a relatively widely reviewed intervention type, lack of consistent evidence led me to remove this branch also as I judged it unlikely to explain drops in population rates.

Such extensive evaluation of well-established programmes could be expected to show indications of positive outcomes were they able to achieve their goals. I removed these two branches from the causal tree as I considered them unlikely to have contributed to observed falling rates. Other branches with no evidence of effectiveness, but not considered by reviews, were retained due to uncertainty.

4.2.8.3 Branches retained for further consideration

I retained three branches representing intervention pathways, as I found evidence of their potential effectiveness across reviews. Policies promoting school-based sex education were effective in some cases, indicating that increasing efficacy of such programmes may contribute to lowering pregnancy rates. Clinic-based interventions were more effective when delivered in more personalised, individualised settings. Whilst no evidence was presented for the increased exposure of a population to such changes in healthcare delivery, it is feasible that greater accessibility and engagement through clinics may have contributed to falling rates.

Changing contraception technologies presented comparatively strong evidence for the potential effects upon a population. Reviews reported effects in trials in different national contexts (demonstrating replicability) and these effects were also observed when examined across whole populations in the case of emergency contraception provision.

Two hypothesised cultural causes of lowering pregnancy rates were not assessed in any reviews – later parenting culture and changes in digital media use. These were retained in the causal tree, as lack of evidence may represent an insufficient search strategy, or the relative unlikelihood of reviews to consider environment changes alongside interventions. Further investigation may produce clearer evidence of their effects on teenage pregnancy.

4.2.8.4 New branches added

Several suggested causes presented some evidence of their effectiveness in reducing teenage pregnancy rates. Community-based interventions focussing on STI prevention, adolescent development interventions, family and community engagement, and interventions targeting rapid repeat pregnancies were each considered in several reviews and saw moderate evidence for their effects on pregnancy rates.

I found some evidence for the effectiveness of digital media-based sexual health interventions when delivered to a group or a class of adolescents. Interventions using digital media to deliver content to vulnerable adolescents showed little evidence of effects. Relatively few studies examined this pathway. Such interventions may have only begun to be studied recently due to the developing technologies giving rise to new approaches. I added the branch for digital media interventions to represent the possibility of the impact of greater uptake of such an intervention in recent years.

Policies and interventions promoting education were also added to the model. These causes were additionally considered in relatively few reviews but presented fairly consistent evidence of positive impacts and were also assessed across populations.

A final addition to the model was increased employment of teenagers outside of school hours. My initial tree proposed this branch as a change in culture, however it was only addressed in one review as an intervention. I added this pathway to ‘Environment changes’ to allow for wider investigation of ongoing shifts in adolescent employment, whilst still allowing the tree to be sensitive to interventions aiming to influence this.

Further modification of the tree included combining the behavioural mediators in the ‘Interventions’ section into one box (Figure 4-4). These pathways were not clearly distinguished across reviews. The pathway of decreasing ‘intended’ pregnancies was added to the upper section to reflect the explicit intervention element of motivating an adolescent to avoid a pregnancy they might not have otherwise considered as unwanted.

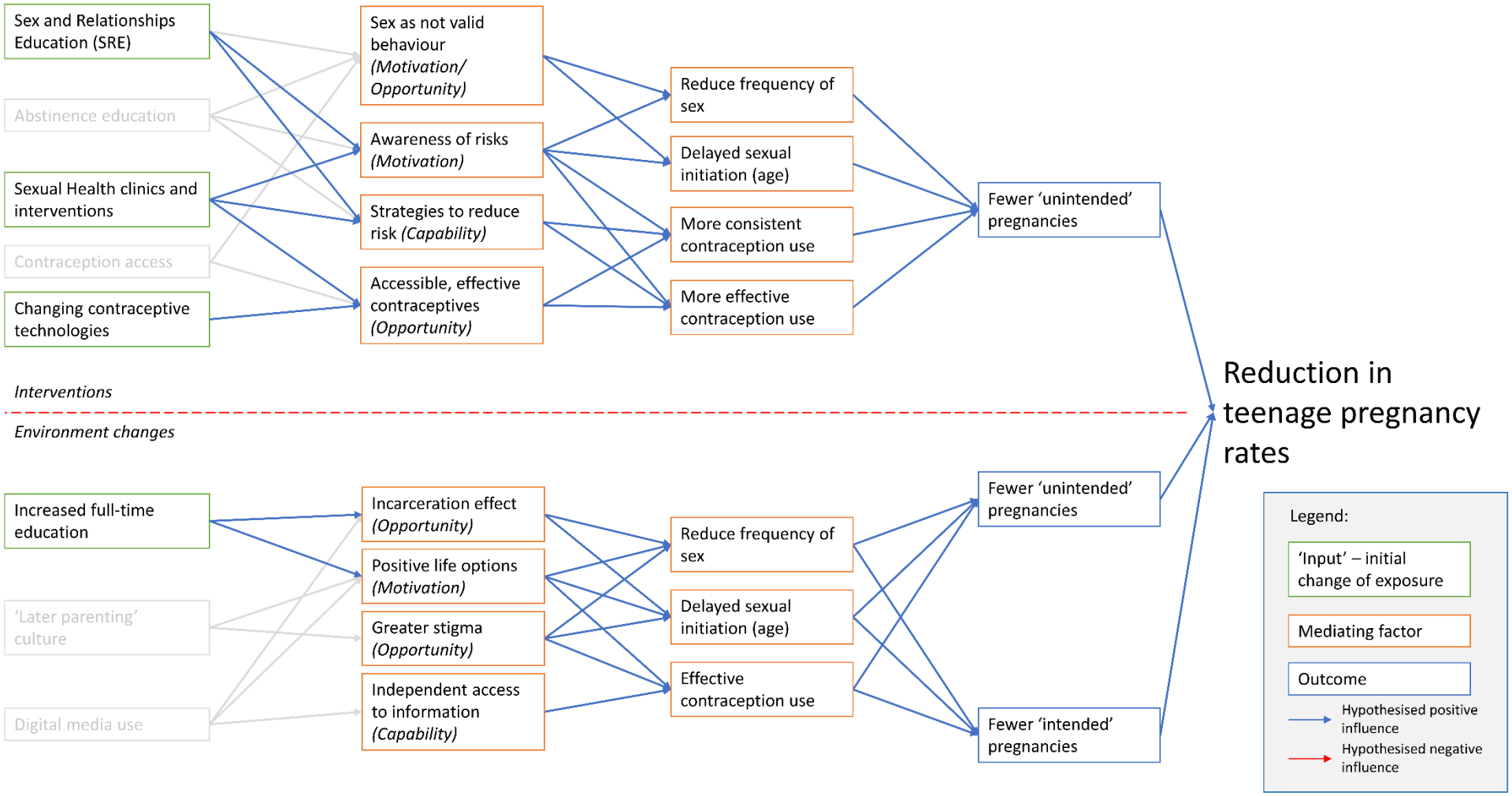

Figure 4-3 – Interim causal tree greying out branches and pathways with little evidence of effectiveness. No evidence for negative effects were found across any intervention or environment change. Two intervention types – abstinence education and contraception access – were assessed across several reviews with evidence presented consistent with null effects. Two environment changes were not addressed in reviews and so were considered untested.

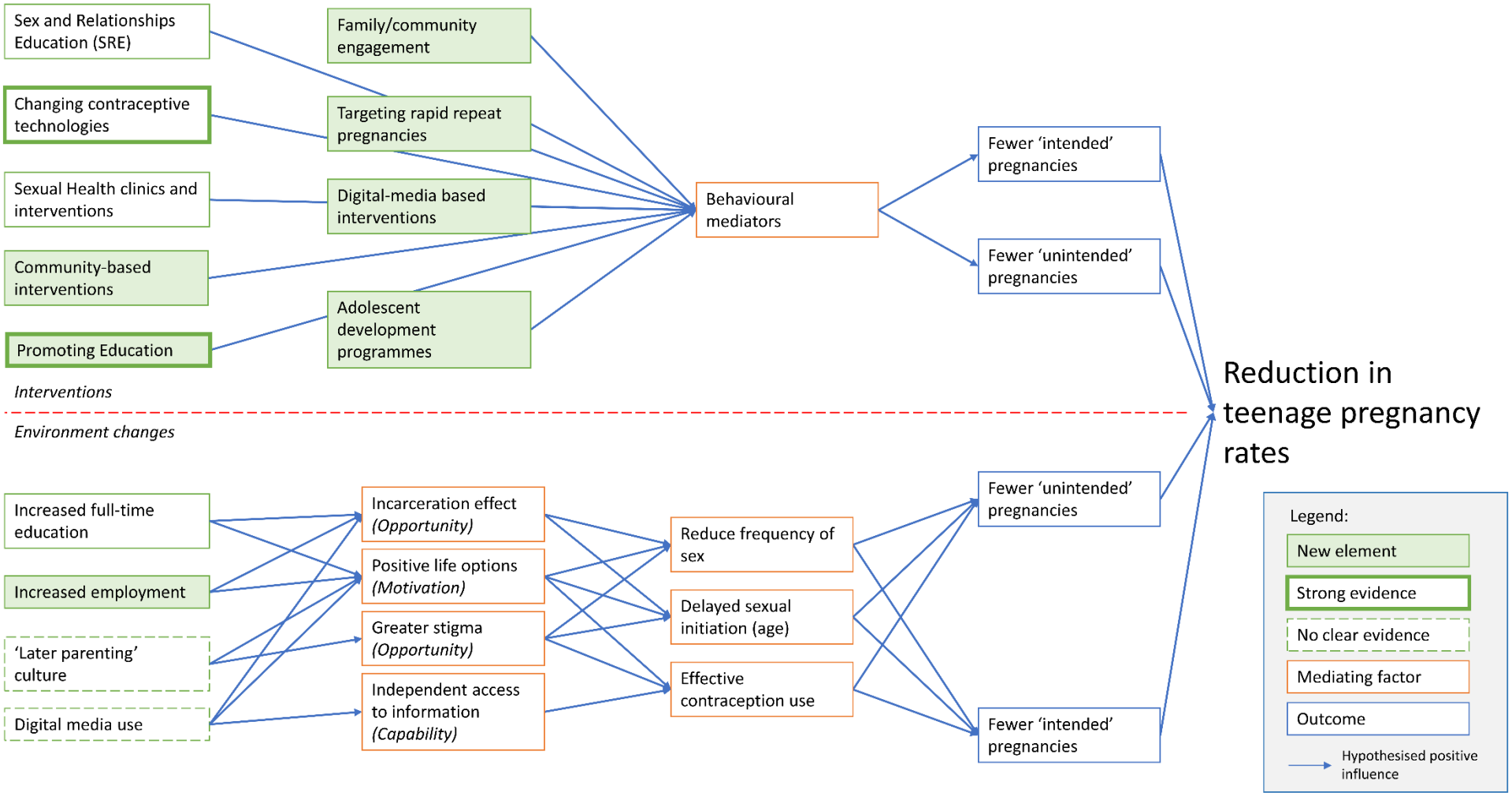

Figure 4-4 – Updated causal tree synthesised from the literature, highlighting newly added branches. Branches identified from the literature with supporting evidence are marked by shaded boxes. Branches with strong evidence of an effect on population are highlighted with thick borders. Branches not adequately assessed are marked with dashed borders. Behavioural mediators in the upper section of the diagram have been simplified as these were not clearly and consistently mapped across the literature addressing the interventions.

4.3 Discussion

My review of reviews found evidence assessing several hypothesised causes of reductions in teenage pregnancy and highlighted a number of additional pathways. I was able to assess the strength of evidence for the effects of the majority of identified causal pathways on the pregnancy risk of those exposed to each intervention or environment change. For several causal pathways there was evidence of a population-wide exposure, indicating their scalability to whole populations and potential to have similarly affected the UK. Although two hypothesised causes showed evidence of both efficacy and population-wide exposure – changing contraception technology and policies promoting education involvement – no reviews causally tested the effects of an exposure on the UK's teenage pregnancy rates from 1999. Consequently, the data presented here is unable to answer the question of what contributed to observed falling rates in the UK. Clear directions for further causal investigation are identifiable from the reviewed literature.

Most reviews explicitly aimed to search for and synthesise evidence for the effectiveness of interventions, usually restricting inclusion to papers reporting similar methods or locations of delivery. Environment changes were rarely considered by reviews. Consequently, my review of reviews may not have been sensitive to capture other published research outputs assessing these environment changes, as they may not have been included in the scope of identified reviews.

4.3.1 Assessing the updated causal tree

The updated causal tree synthesised from the reviewed literature incorporates a broader range of interventions designed to lower rates of teenage pregnancy risk by adding 6 intervention branches, and includes one further hypothesised environmental change. The final suggested causes in the tree (Figure 4-4) represent both pathways with at least some evidence of effectiveness, alongside relatively unexamined pathways with few studies testing their contribution to observed changes.

Several interventions were reported in a large number of reviews with inconsistent findings. The heterogeneity of results in several cases may have arisen from diverse programme designs addressed under the same headings. For example, amongst school-based interventions, there were overlapping aims across programme types – e.g., postponing sexual involvement, reducing number of partners, reducing pregnancy rates, reducing STIs. However, there were potentially differences in programme elements

implementation, producing heterogeneous impact on pregnancy risk. This was partially addressed using split categories of pregnancy-focussed and STI-focussed education programmes. STI programmes are likely to focus on condom use to prevent transmission, whereas pregnancy programmes may recommend other contraceptives with less user-error risk. STI programmes motivate action through raising awareness of infection, whereas pregnancy programmes inform about pregnancy and childbearing. Adolescent pregnancy is not always 'unintended' (Kost, Maddow-Zimet and Arpaia, 2017); motivation to avoid pregnancy may be less universal than motivation to avoid infection. Additionally, motivation to protect against STIs may be more equally shared across both participants. Male partners may be more prompted to take action to avoid infection than pregnancy. Differences in programme focusses may contribute to differences in effectiveness (Kirby, 2003). In recent implementations of school-based SRE, both sets of goals and messages are likely to be incorporated and communicated (e.g. RSHP, 2019).

Studies evaluating school-based interventions may have produced diminishing results over time. Reviews conducted later included more recent primary research showing little effect, masking the effectiveness of interventions tested in the 1990s (Marseille *et al.*, 2018). As all studies reported comparisons with (relatively heterogeneous) sexual health provision at that time and within that country and state, gradual improvements incorporated from previously developed interventions may reach a saturation point. More substantial overhauls of educational approaches compared to usual care in early studies, representing a paradigm shift in public health approaches to sexual health (Arai, 2009), may give little indication of the potential effects of further incremental improvements in SRE in later periods. Other reviewers have identified several areas which may improve effectiveness (Kirby, 2001; Swann *et al.*, 2003; Trivedi *et al.*, 2007), however, this review did not examine which components of SRE were tested for effectiveness. The inclusion of SRE in the final causal tree allows for this uncertainty but prompts more robust causal analysis of the effects of education-focussed programmes to reduce pregnancy.

In several cases, no or little evidence was found addressing environmental exposures, potentially either reflecting unconsidered hypotheses or questions not prioritised in reviews.

In two cases, branches may be mis-specified as intervention or environment change. The decision to classify increased employment as an environment change allows for its investigation as a natural trend, although evidence was only presented in one case for its

manipulation in an intervention. Additionally, trends moving towards use of other contraceptives may be driven by secular trends in changing norms of contraceptive method choice rather than by interventions designed to promote such changes. Though strong evidence was found for interventions promoting such changes, unexamined cultural trends may still have contributed more to observed population changes. Caution should be taken, due to the potential of misspecification, in interpreting this evidence as assessing the effects of interventions to promote changing contraceptive technologies. Likewise, further investigation of this pathway's contribution to falling rates should take care to account for non-designed changes in contraceptive use across the time period.

4.3.2 Strengths and limitations

A clear strength of this review was its wide search of up-to-date literature, focusing on reviews which may have been influential on ongoing practice. Data from a large number of reviews, examining a large number of studies and a broad range of causal pathways, were synthesised to give a concise graphical summary in the updated causal tree. Hypothesised pathways were identified which were not previously detected from non-systematic literature review approaches. This synthesis of hypothesised causes in the causal tree presents the competing models as testable branches.

One limitation is that the search strategy used, which focussed on reviews, was potentially more sensitive to intervention-related hypothesised causes than environment and culture changes. The aim of most reviews was to evaluate intervention approaches, so this limited their sensitivity to detect evidence testing environmental changes. Additionally, the methods used to evaluate interventions are ideally protocol-driven, using robust causal methods and presented clearly. Exposure to interventions can sometimes feasibly be manipulated by evaluators to compare outcomes (Craig *et al.*, 2019). Causal investigation of environment changes is unlikely to be possible through manipulated exposure, and so similarly robust assessment of non-designed trends may be less commonly conducted and less commonly reviewed. This review is likely to have both included environment changes with no effects on teenage pregnancy (but not clearly tested in included literature), and to have omitted potential candidates for significant impacts on pregnancy rates (Paton, 2020). It is possible that the updated causal tree will not be able to explain the full effects on teenage pregnancy seen in the UK. A further review of primary studies could add to the evidence in this review by focussing on environmental exposures and including a wider range of literature testing associations alongside causation.

Previously used tools for assessing risk of bias in reviews were not suitable for use in this analysis. My approach was to develop a new tool to give a simple grading of study quality, however this tool has not undergone tests of reliability and validity. Domains may have been open to interpretation and have produced bias in rating study quality. The predominant split in my rating and use of 'high' and 'low' quality studies was between a set of well-conducted and well-reported systematic reviews and non-systematic literature reviews. Potential unreliability of the tool or subjective judgements in using it may have unfairly weighted influential reviews in deriving conclusions.

A further limitation is the exclusion of published primary studies, known about from other sources, which were not encountered in review literature. In one instance, a more recently-published RCT of virtual infant simulators found an increased risk of pregnancy amongst intervention participants over control, suggesting a negative effect (Brinkman *et al.*, 2016). This result, however, was consistent with my review findings of no significant reduction of adolescent pregnancy rates from virtual infant simulator interventions. For branches with little or conflicting evidence of effectiveness, further synthesis of primary studies could be more sensitive to more up to date and relevant evidence assessing effectiveness.

4.4 Conclusion

The aim of this review was to assess the evidence for hypothesised causes of reducing teenage pregnancy rates in the UK and detect and assess other causes presented in the literature. I present a summary of hypothesised pathways which could inform further testing to determine their contribution to the observed reductions in teenage pregnancy rates in the UK from 1999 to 2016.

An initial proposed intervention to test is the Teenage Pregnancy Strategy, applied to the whole population of England from 1999 to 2010. As highlighted in Chapter 2, the Strategy has been proposed as an explanation of the observed decrease in rates and comprised several of the branches identified above. Strategy elements included school-based sex and relationships education, engaging with parents, and sexual health clinic access and use – branches included in the updated tree as potentially effective approaches (Hadley, Ingham and Chandra-Mouli, 2016). The Strategy also included branches with little evidence of effectiveness: contraception access and a national media campaign. This intervention as a well-defined programme with a clearly identifiable exposure period would be one potential opportunity to test as a hypothesised cause.

This review presents a broadly scoped review of actions taken to reduce teenage pregnancy and hypothesised environmental changes affecting pregnancy rates. It uses a novel approach of developing a search strategy from an initial tree of hypothesised causes and using the literature to update the tree to synthesise available evidence. I successfully eliminated branches which had been repeatedly tested but for which there was evidence of their ineffectiveness, and therefore were unlikely to have contributed to falling rates observed in the UK and provided stronger evidence for branches with the potential to have had an effect. My updated tree highlights causal pathways for further investigation.

Chapter 5 Comparing England with Scotland and Wales using Interrupted Time Series Models

5.1 Overview of chapter

In this chapter, I report the results of my interrupted time series analyses of effects of the Teenage Pregnancy Strategy and discuss the implications of the findings. I use a segmented regression approach, assigning England and comparator countries a pre-intervention and post-implementation period either side of the 1999 Strategy launch and test for changes in level and trend attributable to the Strategy. Comparison models are used to account for other policy, cultural and environmental changes. I use several sensitivity tests to determine the consistency and reliability of my models.

5.2 Results

5.2.1 Exploratory visualisation of pregnancy rates in England, Wales and Scotland

Figure 5-1 shows the yearly under-16, under-18 and under-20 pregnancy rates for the years 1987-2016 (as available) for England, Scotland and Wales. The years before the 1999 Strategy launch were considered the pre-intervention period (unexposed to the Strategy), and 1999 and following years were considered the post-implementation period (exposed). To allow comparison to detect changes in teenage pregnancy trends occurring at the introduction of the Strategy, I required data recording the same measure of pregnancy rates for England and a comparator across both the pre-intervention and post-implementation periods. I obtained sufficient data to allow analysis of the primary outcome measure, under-18 pregnancy rates, as separate comparisons with both Scotland and Wales (Figure 5-1).

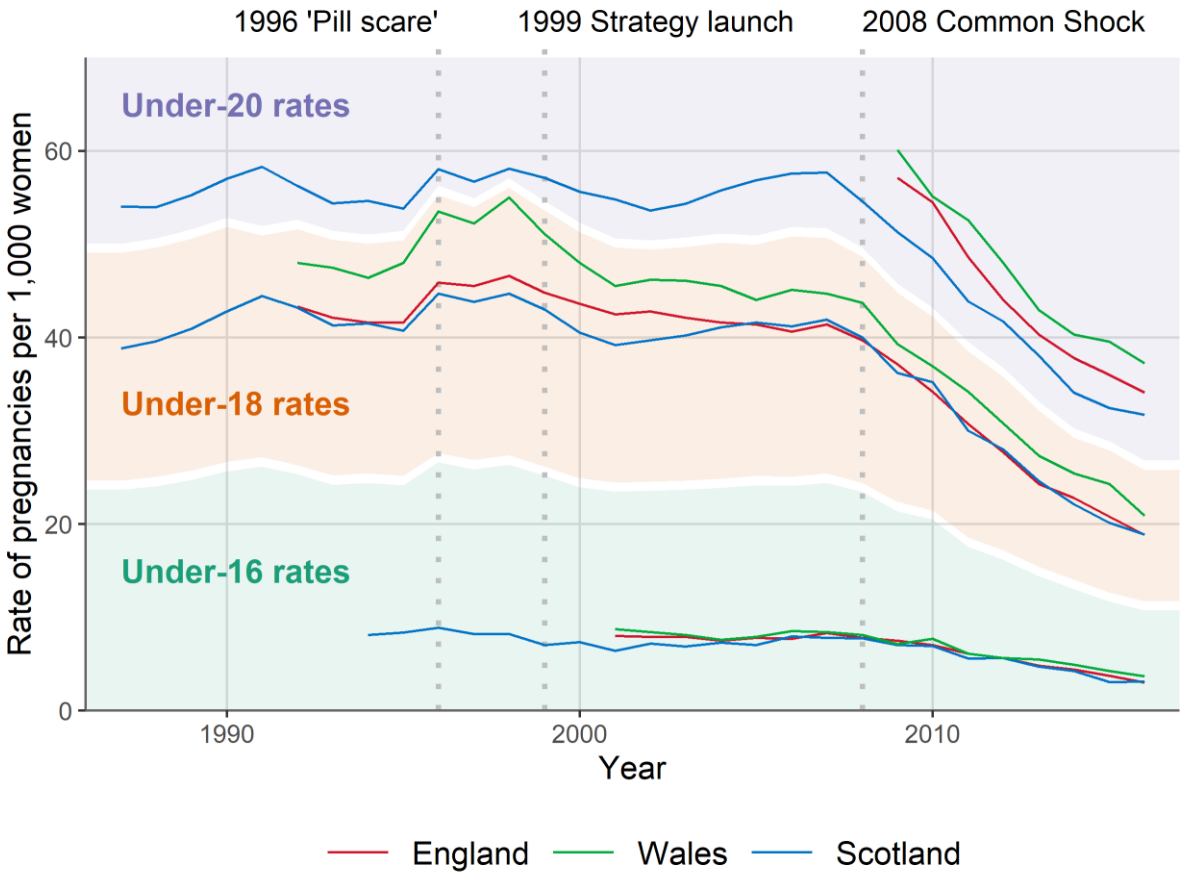

Figure 5-1 – Yearly pregnancy rates across age groups 'Under 16', 'Under 18' and 'Under 20' for England, Wales and Scotland, 1987 to 2016. The range of pregnancy rates did not overlap between age groups; the three extracted age groups are highlighted by shaded sections stacked across the graph. Three time points discussed above – the 1996 'pill scare', the Strategy launch in 1999 and the 2008 common shock – are marked by dashed vertical lines.

As data recording under-16 and under-20 pregnancy rates for both England and Wales were only available for the post-implementation period, comparisons using England and Wales separately were not possible for these age categories. Using combined England and Wales data to compare with Scotland rates, I was able to examine longer pre-intervention periods across all three age groups (Figure 5-2).

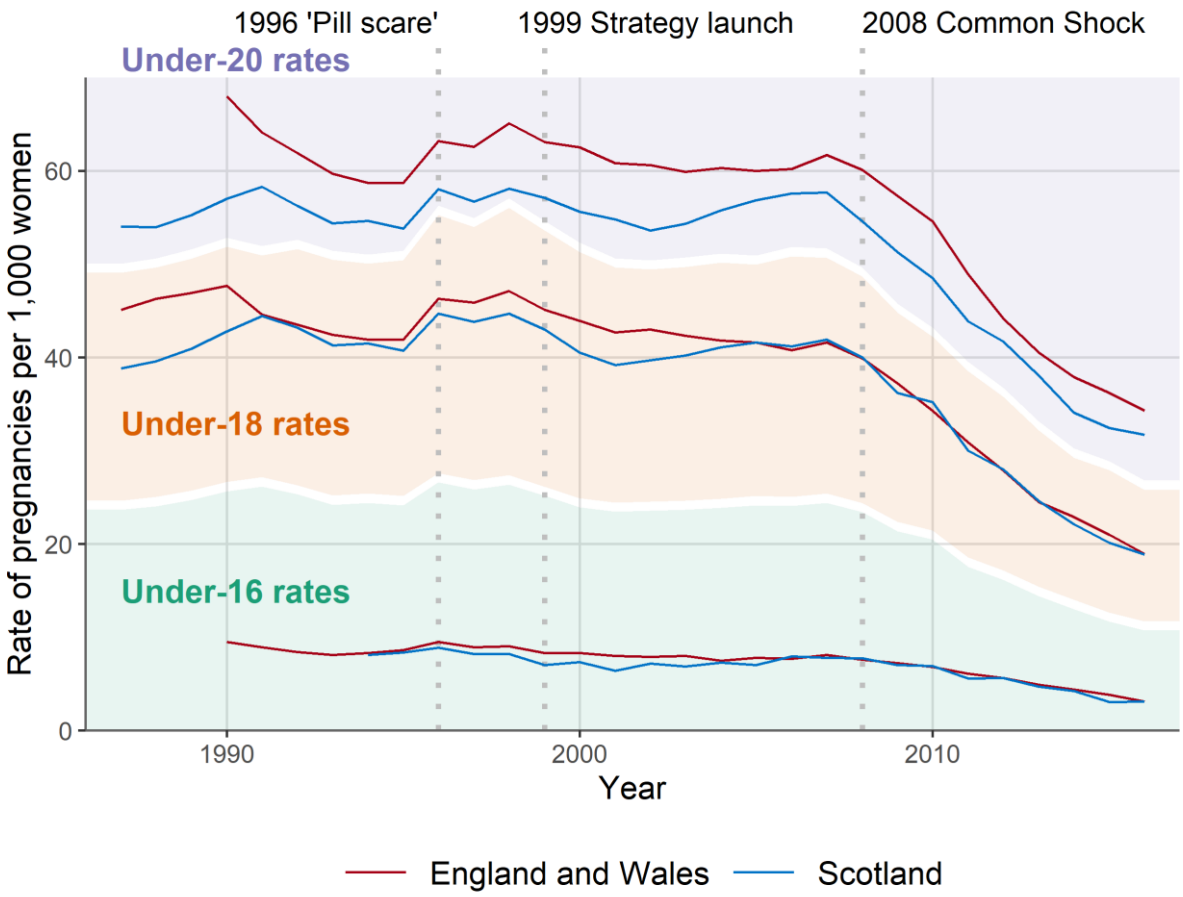

Figure 5-2 – Yearly pregnancy rates across age groups 'Under 16', 'Under 18' and 'Under 20' for Scotland and combined England and Wales, 1987 to 2016.

Examining yearly pregnancy rates before the Strategy launch in 1999, all three countries showed similar rates and similar trends from 1991 onwards. Data available for Scotland for dates before 1991 showed larger gaps between Scotland and combined England and Wales than across the remaining pre-intervention period from 1991 to 1998 (Figure 5-2). This suggests that Scotland's rates across the years before 1991 followed a distinct trend not matching England's. Assigning a level and a trend change at 1991 to Scotland data to correct for a potential bias would explicitly exclude 1987-1990 Scotland data points from calculation of a fitted line to 1991-1998 data. Consequently, I did not use these data and restricted analyses to 1991 onwards. Overall, I judged Scotland and Wales to be good controls for England in this analysis.

Across the 1991 to 1998 pre-intervention period, I noted potential common events affecting pregnancy rates across all three countries to be accounted for. All countries across all age groups showed a sharp increase in rates at 1996, consistent with the 'pill scare' event (Furedi, 1999). As there did not appear to be a similarly shaped drop in rates in later years indicating the end of a defined exposure period, I treated all years from 1996 onwards as exposed to the 'pill scare' corrector. During the post-implementation time

period from 1999 to 2016, a clear change in trend is visible in all countries from 2008 onwards (Figure 5-1 and Figure 5-2). The later trend in England has previously been attributed to mid-term improvements in Strategy implementation (Hadley, 2014; Hadley, Ingham and Chandra-Mouli, 2016; Wellings, Palmer, Geary, *et al.*, 2016). However, as it appeared in all three countries and produced very similar changes at the same time point, I considered it a common shock to be adjusted for in sensitivity analyses.

Across the post-implementation period from 1999 to 2016, all three countries saw similarly decreasing pregnancy rates. In England, pregnancies among women aged under 18 fell by 60% between 1998 and 2016, from 46.6 to 18.8 pregnancies per 1,000 women. Across the same period, Scotland saw a reduction in pregnancies of 58% (from 44.7 to 18.9 pregnancies per 1,000 women) and Wales of 62% (from 55.0 to 20.9 pregnancies per 1,000 women). Similar trends can be seen among women aged under 16 and under 20 across all countries (Figure 5-2). In the interrupted time series models, I aimed to quantify the differences whilst adjusting for background trends to test if any effects are observable.

5.2.2 Preparatory models – England only pre-post intervention

All interrupted time series models were run using a shiny app, published at **andybaxter.shinyapps.io/teen_preg_uk_its**. This was used to generate graphs and reports of model fit and fitted terms. Under each model graph I have included links to view these analyses interactively using the shiny app to ensure repeatability and transparency.

In preparatory models to test fit for England-only data, I fitted a simple segmented regression model to pre- and post-Strategy trends in under-18 pregnancy rates. Model 1 was generated from England-only under-18 pregnancy rates across the years 1992 to 2016. I included three independent variables – a count of years from the start of observation to estimate a pre-intervention trend (beginning at 1992: Time = 1), exposure to Strategy to estimate a level change at Strategy launch (coded as pre-intervention years = 0 and post-implementation years = 1), and years from Strategy start to estimate a trend change at Strategy launch (beginning at 1999: Trend = 1).

Model 1 generated a relatively good fit, able to explain 93% of the observed variance (R^2 = 0.93) and relatively low average prediction error of 3.3 pregnancies per 1,000 women per year (Mean Squared Prediction Error, MSPE = 10.6). The pre-intervention period had a lower R^2 of 0.56, suggesting a relatively poor pre-intervention fit (MSPE 6.06 for the pre-

intervention period). The best fit regression model is represented in Figure 5-3, with points marking the yearly pregnancy rates, solid lines showing the fitted linear model to the pre-intervention and post-implementation phases, and a dashed line representing the continuation of the pre-intervention regression line under the assumption of no Strategy.

Figure 5-3 – England under-18 pregnancies, 1992-2016 with Strategy launch in 1999. The dashed line shows the continuation of the pre-intervention trend, representing the hypothesised continuation of England's pregnancy rate trends in the absence of the Strategy. The time series model incorporates autocorrelation correction: AR1, MA0. MSPE = 10.6; R^2 = 0.93; Log-likelihood = -43.5. Analysis of this model can be viewed using the shiny app at **this link** (with all parameters set to repeat model 1).

Across tests for autocorrelation, there was a consistent indicator of an association between yearly rates and the previous year's rate after adjusting for trend, suggesting autoregression needed to be adjusted for in these models (Stock and Watson, 2020). I applied first-order autoregression correction (AR1) to this and subsequent models. As I re-applied tests for autocorrelation to this corrected model, there did not appear any clear indicators of further moving average (MA) or autoregression autocorrelation.

When I visually examined the yearly data points and regression lines, there were two clear periods (pre- and post-Strategy), some evidence of linear trends and no suspicious outliers. There did however appear two elements of systematic differentiation from the regression

line: a step up between 1995 and 1996 observations, and trend change at 2008 suggesting that the post-Strategy period experienced different linear trends across two distinct periods (Figure 5-3).

The four coefficients for model 1 are shown in Table 5-1. The first two fitted terms suggest a pre-intervention period with an increasing trend (0.70 more pregnancies per 1,000 women per year; 95%CI: -0.34 to 1.74) from an estimated baseline rate of 43.1 pregnancies per 1,000 women per year (95%CI: 36.2 to 49.8). This trend is interrupted at the launch of the Strategy in 1999, with a clear change in slope of -2.22 pregnancies per 1,000 women per year (95%CI: -3.49 to -0.95). The level change associated with Strategy exposure appeared to be very small, with a wide 95% confidence interval incorporating zero (-0.03 pregnancies per 1,000 women; 95%CI: -3.08 to 3.01). The outputs of this simple model are consistent with previous observations of decreasing pregnancy rates from a high pre-Strategy baseline (Wellings, Palmer, Geary, *et al.*, 2016).

Coefficient	Estimate	Standard Error	Lower 95% CI	Upper 95% CI
England (est.) rate at 1991	43.06	3.46	36.27	49.85
England pre-intervention trend	0.70	0.53	-0.34	1.74
England change in level at 1999	-0.03	1.55	-3.08	3.01
England change in trend at 1999	-2.22	0.65	-3.49	-0.95

Table 5-1 – Fitted terms and confidence intervals of model 1 fitted to England-only under-18 pregnancy rates with no additional correction coefficients. These were generated from England-only data with three independent variables – years from observation start, exposure to Strategy and years from Strategy start – and the dependent variable: under-18 pregnancies per 1,000 women.

To improve the fit of the model, in model 2 I tested the effects of adding a 'pill scare' corrector to the years 1996 onwards (of value = 1 in all exposed years) to simulate the effects of medical advice concerning health risks of contraceptive pills on teenage pregnancy risk (Furedi, 1999). This additional component significantly increased the likelihood of the model (likelihood ratio between models = 10.7; p = 0.001) and improved the fit of the pre-intervention period (with the MSPE reduced from 6.06 to 3.56 and R^2 increased from 0.56 to 0.92). The fitted pre-Strategy line in Figure 5-4 more closely fits the observed data points compared to the uncorrected model in Figure 5-3, suggesting that a linear trend with a level increase at 1996 is a better predictor of post-implementation trends.

Figure 5-4 – England under-18 pregnancies, 1992-2016 with Strategy launch in 1999 and 'pill scare' corrector allowing a jump in rates from 1996 onwards. This corrector represents a jump of 4.5 pregnancies per 1,000 women. I applied this to all subsequent years as there was no clear return to a low rate before 1999. Autocorrelation correction: AR1, MA0; MSPE = 10.8; R^2 = 0.94; Log-likelihood = -38.1. Model 2 can be viewed at **this link**.

The fitted terms for model 2 show a clear increase of 4.5 more pregnancies per 1,000 women at 1996, consistent with the hypothesised effects of the pill scare (Table 5-2; 95%CI: 1.84 to 7.17). The fitted trend to the pre-intervention period was estimated to be marginally lower than in the uncorrected model (from a 0.70 to -0.11 change in pregnancy rate per year), but with both estimates within each other's confidence intervals. The post-implementation trend is not itself affected by this added corrector, but the magnitude of the trend change is smaller as its difference from the pre-intervention trend is decreased. This model also showed a significant change in trend from 1999 and the start of the Strategy, of 1.41 fewer pregnancies per year across each year from Strategy implementation (95%CI: -2.58 to -0.24).

Coefficient	Estimate	Standard Error	Lower 95% CI	Upper 95% CI
England (est.) rate at 1991	44.11	3.34	37.57	50.65
England pre-intervention trend	-0.11	0.51	-1.10	0.88
England change in level at 1999	-0.14	1.26	-2.62	2.34
England change in trend at 1999	-1.41	0.60	-2.58	-0.24
'Pill Scare' corrector	4.51	1.36	1.84	7.17

Table 5-2 – Fitted terms and confidence intervals of model 2 fitted to England-only under-18 pregnancy rates 1992-2016, with a 'Pill scare' corrector. This model used the coefficients of model 1 and added a fourth variable – a dummy corrector for exposure to the pill scare coded as 0 in years before 1996 1 in all other years.

In inspecting Figure 5-4, I could not confidently identify a discrete end to the pill scare corrected period, which might appear as a similarly sized immediate fall in rates. If the effects were indeed produced by the hypothesised impact of publicised concerns over the health impact of the oral contraceptive pill, these alternatively could be expected to decrease steadily over time rather than end at a distinct point. However, incorporating a trend change as well as a level change to account for pill scare effects from 1996 would then require calculation of a Strategy effect on trends from only three pre-intervention observations, which would likely be insufficient to detect any effects. If a decrease in rates was produced by decreasing concerns with oral contraceptive pills or gradual replacement of methods, this model would be unable to disentangle these effects from the impact of the Strategy. As the pill scare was also observable in both Scotland and Wales, I therefore assumed that controlled models would be able to account for this. Controlled models would provide adjustment for such background trends and allow identification of the effects of the Strategy.

5.2.3 Comparison models – Scotland and Wales as controls

In the first comparative model 3, I plotted Scotland observations alongside England, fitting a regression line to Scotland using the same predictors as England in the preparatory model. To test the effects on England, I assigned additional variables to England data to represent the difference in baseline rate and the difference in level and trend change from control. I fitted regression lines to visualise these differences from Scotland pre- and post-Strategy (Figure 5-5). As discussed in the Methods chapter, I assumed pre-intervention parallel trends across UK countries as this produced a relatively stable and parsimonious model.

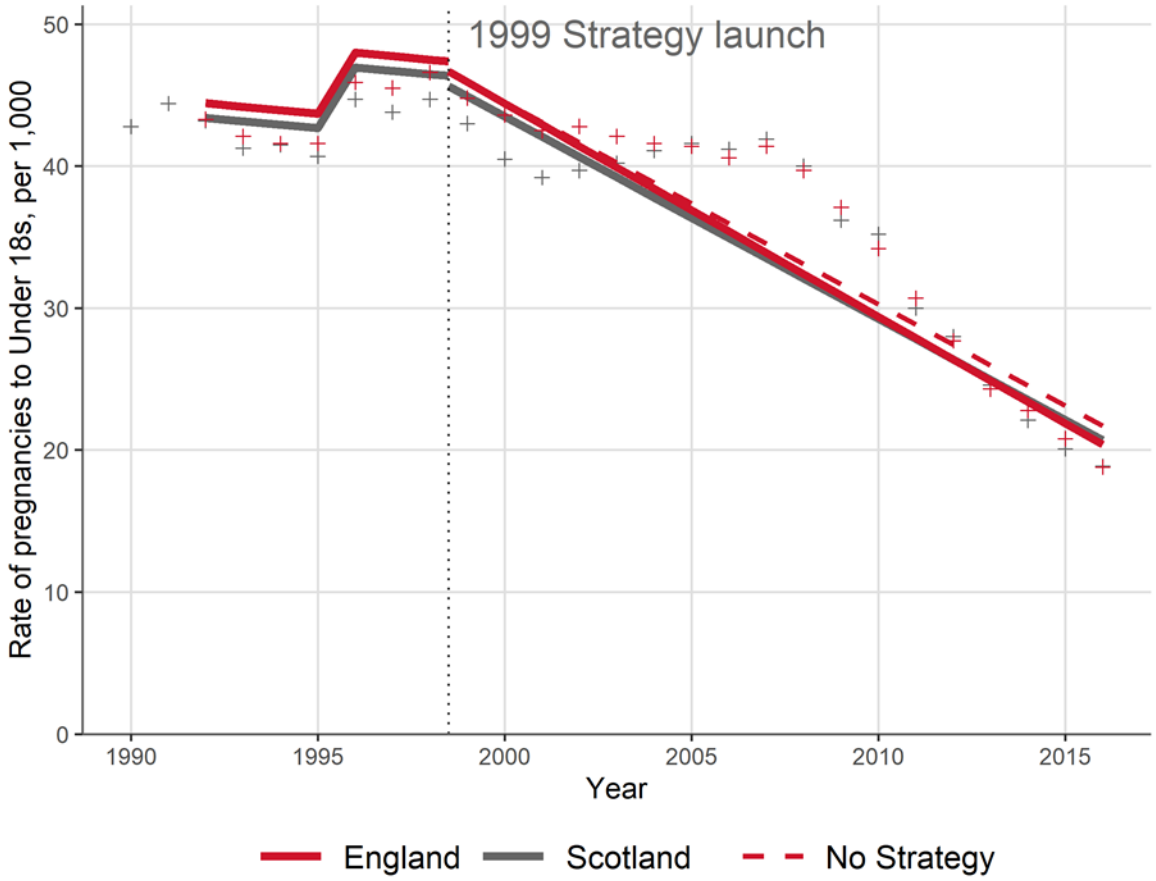

Figure 5-5 – England under-18 pregnancies compared with Scotland, 1992-2016 with Strategy launch in 1999 and 'pill scare' corrector allowing a jump in rates from 1996 onwards. Scotland observations and trends are shown in grey. England observed trends are represented by solid lines. The dashed line shows the hypothesised trend in the absence of the Strategy. This is constructed by applying Scotland level and trend changes to England's 1992-1998 pre-intervention trends to project England's hypothesised post-implementation trends from 1999 to 2016. Autocorrelation correction: AR1, MA0; MSPE = 12.4; R^2 = 0.92; Log-likelihood = -85.0. Model 3 can be viewed at **this link**.

In this model, I used Scotland's estimated changes at 1999 applied to England's pre-intervention trend to estimate the trend change that would have occurred in England in the absence of the Strategy. This adjusted trend is shown in Figure 5-5 as a dashed line and can be seen to closely follow the observed trend in the post-implementation period. The coefficients for a difference in level change (0.06 more pregnancies per 1,000 women per year; 95%CI: -4.03 to 4.16) and a difference in trend change (-0.08 decrease in pregnancy rate per year; 95%CI: -0.74 to 0.59) seen in England in 1999 when compared to Scotland are both estimated as small, non-significant values, consistent with a null effect of the Strategy (Table 5-3).

Coefficient	Estimate	Standard Error	Lower 95% CI	Upper 95% CI
Scotland (est.) rate at 1991	43.65	3.29	37.21	50.09
Scotland pre-intervention trend	-0.24	0.41	-1.05	0.57
England difference in rate at 1991	1.02	4.12	-7.06	9.09
Scotland change in level at 1999	-0.14	1.48	-3.04	2.76
Scotland change in trend at 1999	-1.19	0.52	-2.19	-0.18
England difference in level from control at 1999	0.06	2.09	-4.03	4.16
England difference in trend from control at 1999	-0.08	0.34	-0.74	0.59
'Pill Scare' corrector	4.51	1.13	2.30	6.72

Table 5-3 – Fitted terms and confidence intervals of model 3 fitted to England and Scotland under-18 pregnancy rates 1992-2016. Three more independent variables were used to estimate the difference in pregnancy rates across the pre-intervention period between Scotland and England, the difference in level change at 1999 and the difference in trend change from 1999 onwards. The two fitted terms representing England's difference in level and trend change at 1999 were taken as measurements of the effects of exposure.

A visual inspection of the fitted lines for both Scotland and England suggests that the pill scare corrector appears to remain a valid parameter to include as similar jumps were seen in both countries. The coefficient for the step up in rates at 1996 remained constant between preparatory (England only; model 2) and comparative (England vs Scotland; model 3) models (at 4.51 more pregnancies per 1,000 women per year), supporting the conclusion that a similar effect was seen across both countries. Scotland's rates in the absence of the Strategy saw a significant change in trend from 1999 of an accumulating 1.19 fewer pregnancies per 1,000 women per year (95%CI: -2.19 to -0.18).

In my second comparative model, model 4, I used Wales as a control and performed the same analyses on rate and trend changes seen across England and Wales (Figure 5-6).

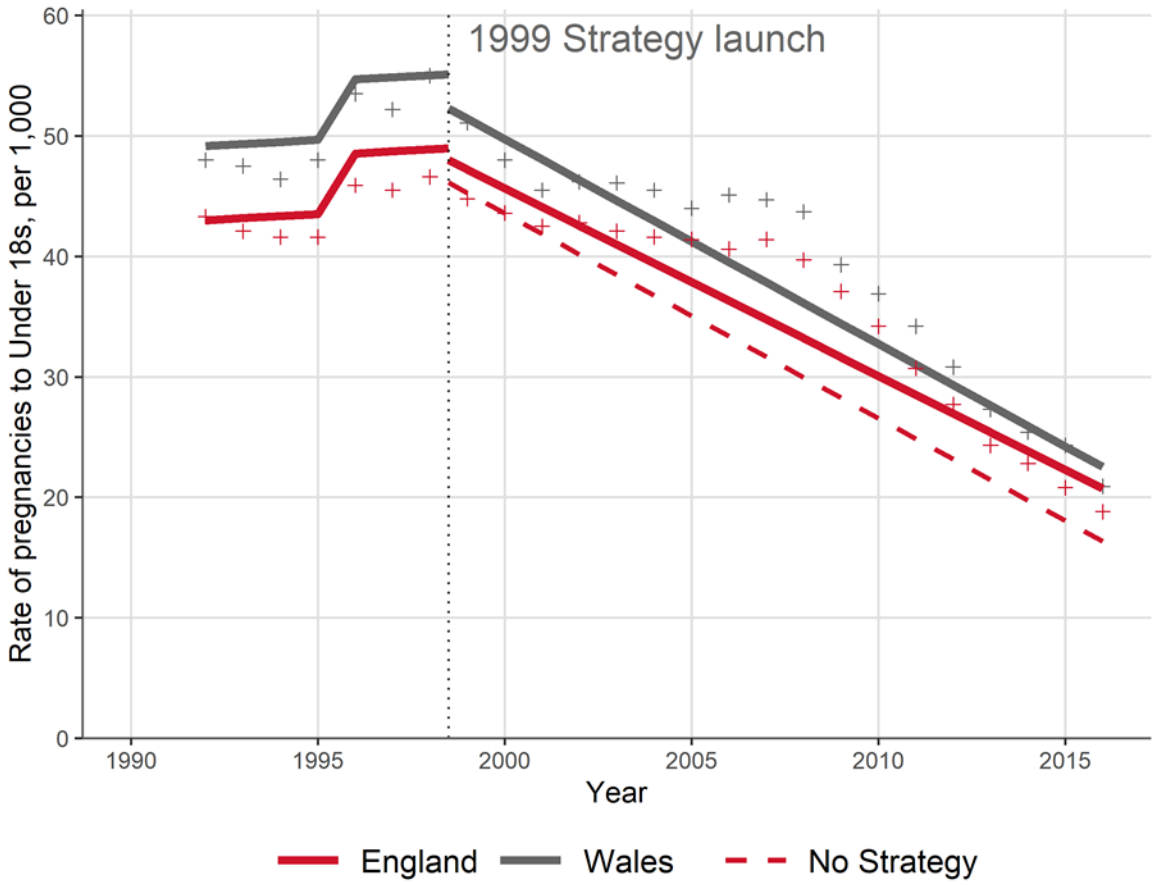

Figure 5-6 – England under-18 pregnancies compared with Wales, 1992-2016 with Strategy launch in 1999 and 'pill scare' corrector allowing a jump in rates from 1996 onwards. Wales observations and trends are shown in grey. England observed trends are represented by solid lines and the dashed line shows the hypothesised trend in the absence of the Strategy. Autocorrelation correction: AR1, MA0; MSPE = 9.42; R^2 = 0.95; Log-likelihood = -84.7. Model 4 can be viewed at **this link**.

England saw lower rates across the pre-intervention period than Wales (averaging -6.16 fewer pregnancies per 1,000 women across 1992-1998; 95%CI: -13.21 to 0.89) and the common pre-intervention trend estimated a small yearly increase of 0.17 more pregnancies per 1,000 women per year, with a confidence interval spanning zero and consistent with previous estimates (95%CI: -0.63 to 0.97; Table 5-4).

Coefficient	Estimate	Standard Error	Lower 95% CI	Upper 95% CI
Wales (est.) rate at 1991	48.99	2.95	43.21	54.77
Wales pre-intervention trend	0.17	0.41	-0.63	0.97
England difference in rate at 1991	-6.16	3.60	-13.21	0.89
Wales change in level at 1999	-1.90	1.48	-4.81	1.00
Wales change in trend at 1999	-1.87	0.50	-2.86	-0.89
England difference in level from control at 1999	1.81	2.09	-2.30	5.91
England difference in trend from control at 1999	0.14	0.32	-0.48	0.76
'Pill Scare' corrector	4.86	1.14	2.63	7.10

Table 5-4 – Fitted terms and confidence intervals of model 4 fitted to England and Wales under-18 pregnancy rates 1992-2016. This uses the same coefficients as model 3, with England's change in level and trend at 1999 representing estimates of the Strategy's effects.

The pill scare corrector in the England and Wales model was again similar to previous estimates, suggesting a common effect of a similar magnitude affecting Wales also. Wales, similar to Scotland, saw a decreasing trend in the absence of the Strategy, with a trend change of 1.87 fewer pregnancies per 1,000 women per year accumulating over the Strategy period (95%CI: -2.86 to -0.89). England's change in level at 1999 was smaller than Wales', with a difference of 1.81 more pregnancies per 1,000 women than predicted (95%CI: -2.30 to 5.91). The trend change showed a very small difference of 0.14 pregnancies per 1,000 women per year (95%CI: -0.48 to 0.76). Wales may have experienced greater reductions in pregnancy rate from 1999 than Strategy-exposed England but observations are consistent with no difference from 1999 between countries and a null effect of the Strategy.

5.2.4 Sensitivity and robustness analyses

5.2.4.1 2008 Common shock corrector

In my sensitivity analyses, to examine the common shock to the trends in each of the three countries shown in the figures above, I added a variable to allow for a common trend change from 2008, simulating an unknown event which appears to affect all UK rates. This was coded as 1 in 2008 and incrementing in following years to estimate a change in linear trend across this period. Figure 5-7 shows an initial model, model 5, fitted using only England data to test the strength of this predictor.

Figure 5-7 – England under-18 pregnancies, 1992-2016 with Strategy launch in 1999, pill scare corrector at 1996 and additional 'common shock' corrector from 2008-2016. The common shock variable was coded as 1 in 2008 and incremented by year across the remaining period. Autocorrelation correction: AR1, MA0; MSPE = 0.453; R^2 = 0.997; Log-likelihood = -21.9. Model 5 can be viewed at **this link**.

In comparison with preparatory model 2 fitted to England-only data, the combination of correctors for pre-intervention and post-implementation trends in model 5 produced a better fit. Fitting linear trends to grouped years 1992-1995, 1996-1998, 1999-2007 and 2008-2013 accounted for over 99% of the variance from the mean across all time periods (R^2 = 0.997, increased from R^2 = 0.935 in model 2), and a lower MSPE of 0.453 (decreased from 10.8 in model 2). The common shock coefficient estimated a large change in trend of -2.17 pregnancies per 1,000 women per year from 2008 onwards (95%CI: -2.56 to -1.79). England's observed rates from 1999 saw a small drop in level of -1.44 fewer pregnancies per 1,000 women (95%CI: -2.94 to 0.06), but a minimal trend change with a narrow confidence interval spanning zero (-0.08 fewer pregnancies per year than predicted trends; 95%CI: -0.64 to 0.48). When using these predictors on England-only data, there was no clear indicator of a Strategy effect at the 1999 launch date, with a seemingly consistent reducing trend across the period 1996 to 2007 and a small but uncertain level change at 1999.

Coefficient	Estimate	Standard Error	Lower 95% CI	Upper 95% CI
England (est.) rate at 1991	43.01	0.88	41.30	44.73
England pre-intervention trend	-0.28	0.24	-0.75	0.20
England change in level at 1999	-1.44	0.76	-2.94	0.06
England change in trend at 1999	-0.08	0.29	-0.64	0.48
Common change in trend at 2008 shock	-2.17	0.20	-2.56	-1.79
'Pill Scare' corrector	4.62	0.90	2.85	6.39

Table 5-5 – Fitted terms and confidence intervals of model 5 fitted to England under-18 pregnancy rates 1992-2016, with a further corrector for a trend change at 2008. This model used the same variables as model 2 with an added trend variable, counting from 2008 onwards, to estimate a change in trend occurring in 2008.

This slow change over the initial period of Strategy implementation is consistent with earlier evaluations, attributing this to a smaller than anticipated Strategy effect (Wilkinson *et al.*, 2006). Model 5 suggests that this downward trend is a continuation of the trend from 1992-1998 (with pill-scare corrector to account for a jump in rates).

Re-introduction of Scotland and Wales as controls in Figure 5-8 and Figure 5-9 produced similarly well-fitted models and measures of trend changes associated with a common shock at 2008 which were consistent with the England-only model (-2.17 change in trend in England model, compared with -2.40 and -2.11 in Scotland and Wales models respectively). This closeness of fit across all three models suggests that a similarly sized effect was seen in all three contexts, and that the fitting of a single predictor to both England and controls in each case was able to fully account for this in the models.

Figure 5-8 – England under-18 pregnancies compared with Scotland, 1992-2016 with Strategy launch in 1999, pill scare corrector at 1996 and 2008 common shock corrector. Autocorrelation correction: AR1, MA0; MSPE = 0.793; R^2 = 0.99; Log-likelihood = -59.3. Model 6 can be viewed at **this link**.

Coefficient	Estimate	Standard Error	Lower 95% CI	Upper 95% CI
Scotland (est.) rate at 1991	42.65	0.87	40.94	44.36
Scotland pre-intervention trend	-0.46	0.22	-0.90	-0.02
England difference in rate at 1991	1.04	0.76	-0.44	2.52
Scotland change in level at 1999	-2.38	0.90	-4.13	-0.62
Scotland change in trend at 1999	0.36	0.26	-0.16	0.88
England difference in level from control at 1999	0.90	1.10	-1.27	3.06
England difference in trend from control at 1999	-0.13	0.09	-0.30	0.04
Common change in trend at 2008 shock	-2.40	0.17	-2.74	-2.06
'Pill Scare' corrector	4.61	0.86	2.93	6.29

Table 5-6 – Fitted terms and confidence intervals of model 6 fitted to England and Scotland under-18 pregnancy rates 1992-2016, with a further corrector for a trend change at 2008.

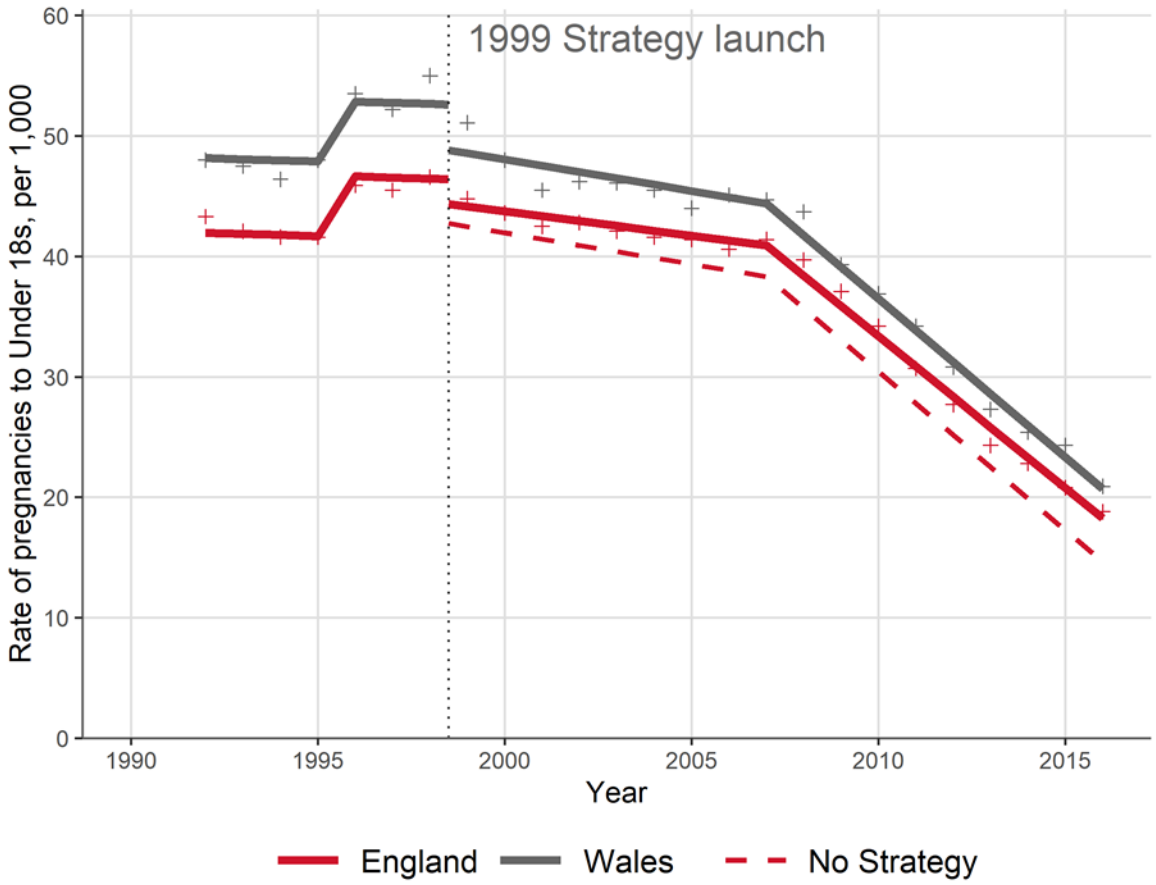

Figure 5-9 – England under-18 pregnancies compared with Wales, 1992-2016 with Strategy launch in 1999, pill scare corrector at 1996 and 2008 common shock corrector. Autocorrelation correction: AR1, MA0; MSPE = 0.863; R^2 = 0.995; Log-likelihood = -63.5. Model 7 can be viewed at **this link**.

Coefficient	Estimate	Standard Error	Lower 95% CI	Upper 95% CI
Wales (est.) rate at 1991	48.24	0.89	46.49	49.99
Wales pre-intervention trend	-0.09	0.25	-0.57	0.39
England difference in rate at 1991	-6.20	0.74	-7.65	-4.75
Wales change in level at 1999	-3.58	0.94	-5.42	-1.73
Wales change in trend at 1999	-0.43	0.28	-0.98	0.12
England difference in level from control at 1999	1.67	1.13	-0.54	3.87
England difference in trend from control at 1999	0.12	0.09	-0.05	0.29
Common change in trend at 2008 shock	-2.11	0.17	-2.45	-1.78
'Pill Scare' corrector	5.05	0.96	3.17	6.92

Table 5-7 – Fitted terms and confidence intervals of model 7 fitted to England and Wales under-18 pregnancy rates 1992-2016, with a further corrector for a trend change at 2008.

Across these models, differences in level and trend changes at 1999 at Strategy launch in England are small, consistent with previous models without a 2008 common shock, and all with confidence intervals crossing zero (Table 5-6 and Table 5-7). These observations are similarly consistent with a null effect of the Strategy.

5.2.4.2 Comparisons across all age groups using England and Wales combined data

In further analyses, I compared Scotland with combined England and Wales data. This allowed me to repeat comparisons of under-18 pregnancy rates with Scotland as control over a longer time period, and to conduct similar comparisons on under-16 and under-20 age groups. I based these models on my primary comparison model (models 3 and 4), with pill scare corrector included but no common shock at 2008.

In model 8 I used the under-18 age group and compared Scotland with combined England and Wales data over a longer period to compare with model 3 using Scotland and England only. England and Wales combined under-18 rates were available from 1987 to 2016, but I restricted my analysis period to 1991 onwards as the earlier years appeared to show different trends for Scotland (Figure 5-2). This allowed inclusion of one more year to calculate pre-intervention trends for each unit. This change resulted in little difference from previous models, with a slightly steeper decreasing trend and a small reduction in the standard error of the pre-intervention trend (from 0.41 to 0.36; Table 5-8; Figure 5-10). Changes attributable to the Strategy remained close to zero with confidence intervals spanning the null effect point for both level and trend changes at 1999 (Table 5-8). This was consistent with results for model 3 presented in Table 5-3 above.

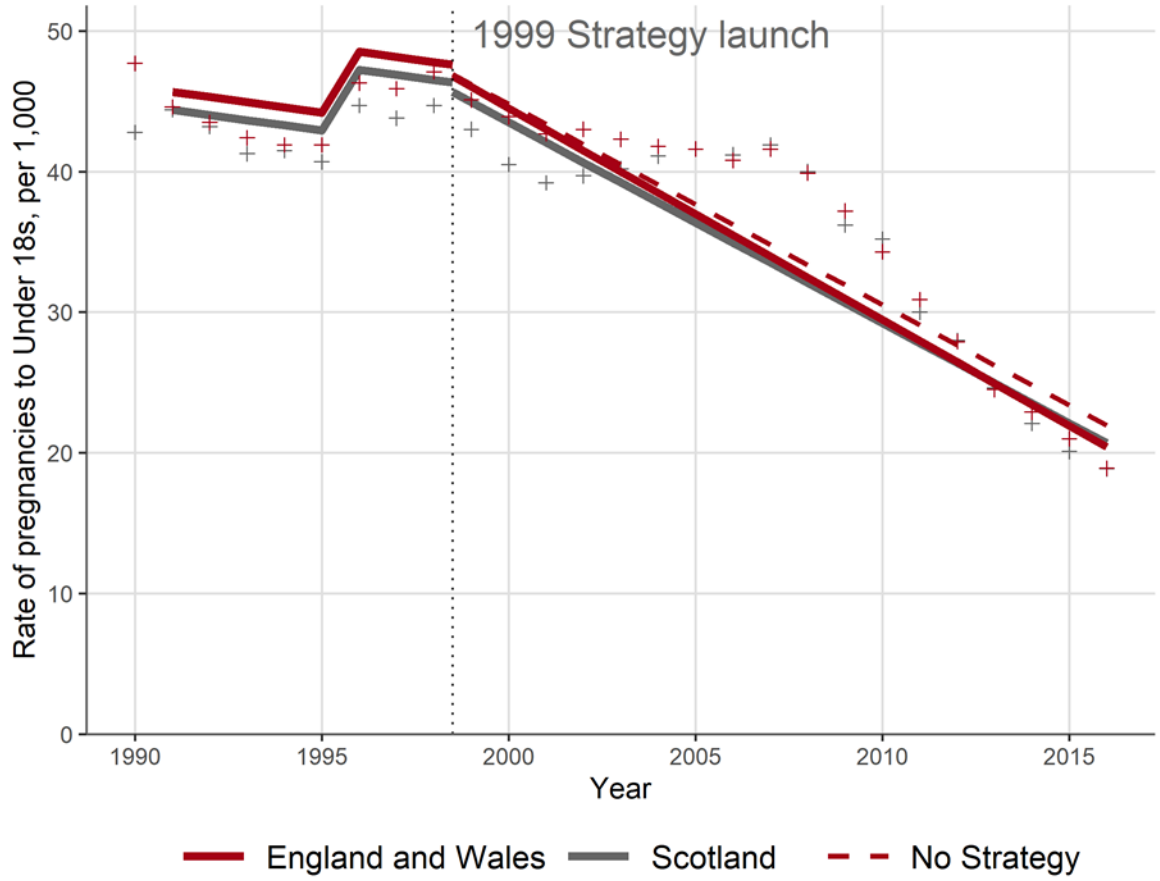

Figure 5-10 – England and Wales combined under-18 rates compared with Scotland for the period 1991-2016. Autocorrelation correction: AR1, MA0; MSPE = 12.4; R^2 = 0.92; Log-likelihood = -87.6. Model 8 can be viewed at **this link**.

Coefficient	Estimate	Standard Error	Lower 95% CI	Upper 95% CI
Scotland (est.) rate at 1990	44.76	3.24	38.40	51.10
Scotland pre-intervention trend	-0.37	0.36	-1.08	0.35
England and Wales difference in rate at 1990	1.28	4.05	-6.65	9.21
Scotland change in level at 1999	-0.15	1.45	-3.00	2.70
Scotland change in trend at 1999	-1.06	0.47	-1.98	-0.14
England and Wales difference in level from control at 1999	-0.12	2.05	-4.14	3.91
England and Wales difference in trend from control at 1999	-0.08	0.34	-0.74	0.58
'Pill Scare' corrector	4.69	1.09	2.54	6.83

Table 5-8 – Fitted terms and confidence intervals of model 8 fitted to Scotland and combined England and Wales under-18 pregnancy rates 1991-2016.

I used data recording under-16 and under-20 pregnancy rates for Scotland and combined England and Wales in robustness checks of different populations. I compared these with model 3 as the primary comparison between England only and Scotland data. The under-16 and under-20 age categories were not noted as Strategy targets (Social Exclusion Unit, 1999). Both, however, could be expected to be affected by interventions to lower teenage pregnancy, and changes seen in these groups may be indicative of Strategy effects.

Across both models examining other age groups, I saw similar results, all consistent with a null effect of the Strategy. In the under-16 model 9, model fit was poorer than model 3 (R^2 = 0.88, lower than R^2 = 0.92 above) and the coefficient of the pill scare corrector was smaller than estimated in the initial model, and not significant (0.79 more pregnancies per 1,000 women per year; 95%CI: -0.06 to 1.65; Table 5-9; cf. 4.51 more pregnancies in the under-18 model; Table 5-3). The observed rates across the period were approximately 20% of the under-18 rates (Figure 5-2), but as the standard error was comparatively large, this reduced the sensitivity of this comparison to detect a change induced by the Strategy. The estimated changes around 1999 appear small and consistent with a null effect of the Strategy (Figure 5-11 and Table 5-9).

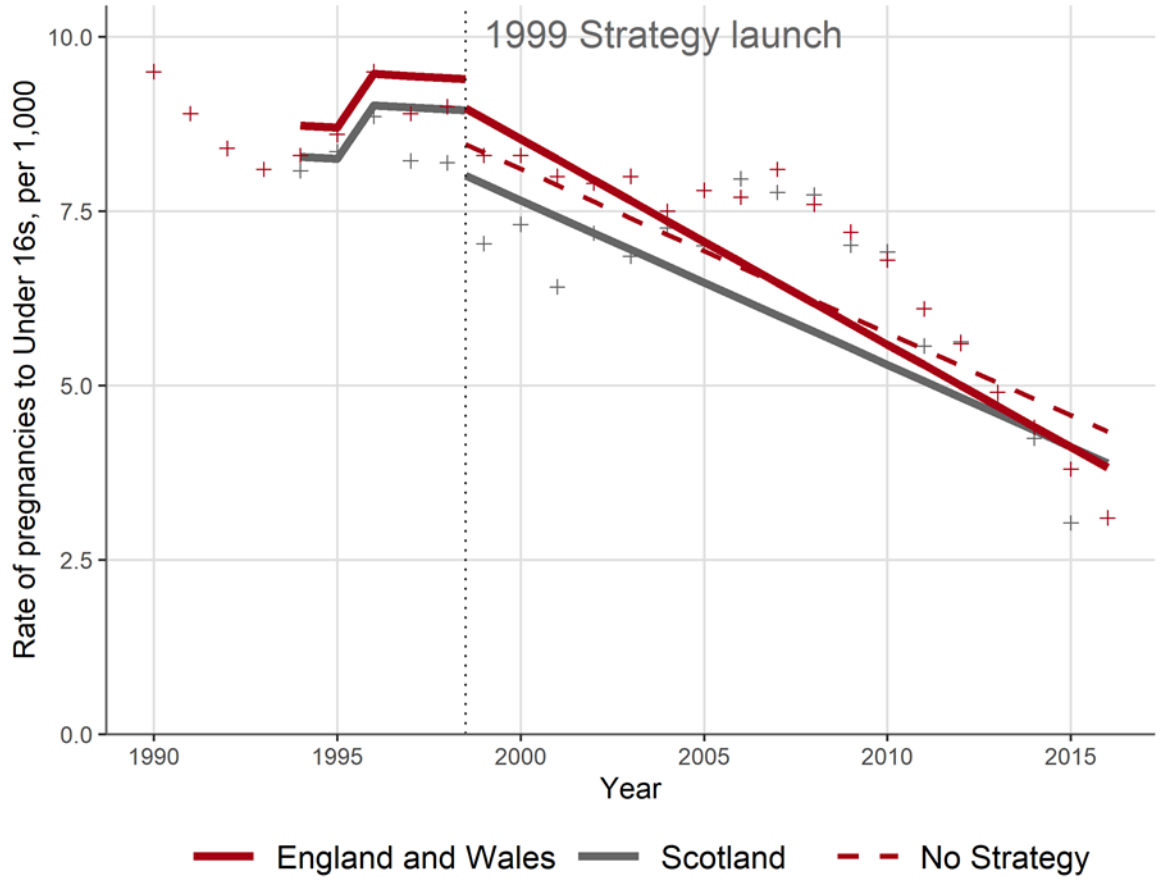

Figure 5-11 – England and Wales combined under-16 pregnancies compared with Scotland, 1994 to 2016. Autocorrelation correction: AR1, MA0; MSPE = 0.725; R^2 = 0.88; Log-likelihood = -29.5. Model 9 can be viewed at **this link**.

Coefficient	Estimate	Standard Error	Lower 95% CI	Upper 95% CI
Scotland (est.) rate at 1993	8.31	0.90	6.55	10.06
Scotland pre-intervention trend	-0.03	0.19	-0.41	0.35
England and Wales difference in rate at 1993	0.45	1.04	-1.59	2.50
Scotland change in level at 1999	-0.83	0.53	-1.87	0.20
Scotland change in trend at 1999	-0.21	0.22	-0.63	0.22
England and Wales difference in level from control at 1999	0.55	0.74	-0.91	2.01
England and Wales difference in trend from control at 1999	-0.06	0.10	-0.25	0.13
'Pill Scare' corrector	0.79	0.44	-0.06	1.65

Table 5-9 – Fitted terms and confidence intervals of model 9 fitted to Scotland and combined England and Wales under-16 pregnancy rates 1994-2016.

In comparisons of under-20 pregnancy rates, model 10 fit was also poorer than primary models (R^2 = 0.88; Figure 5-12). England and Wales saw a decreasing trend in pregnancy rates slightly greater than predicted from control, but still consistent with a null effect (-0.23 further reductions in under-20 pregnancy rates per year; 95%CI: -1.09 to 0.62; Table 5-10).

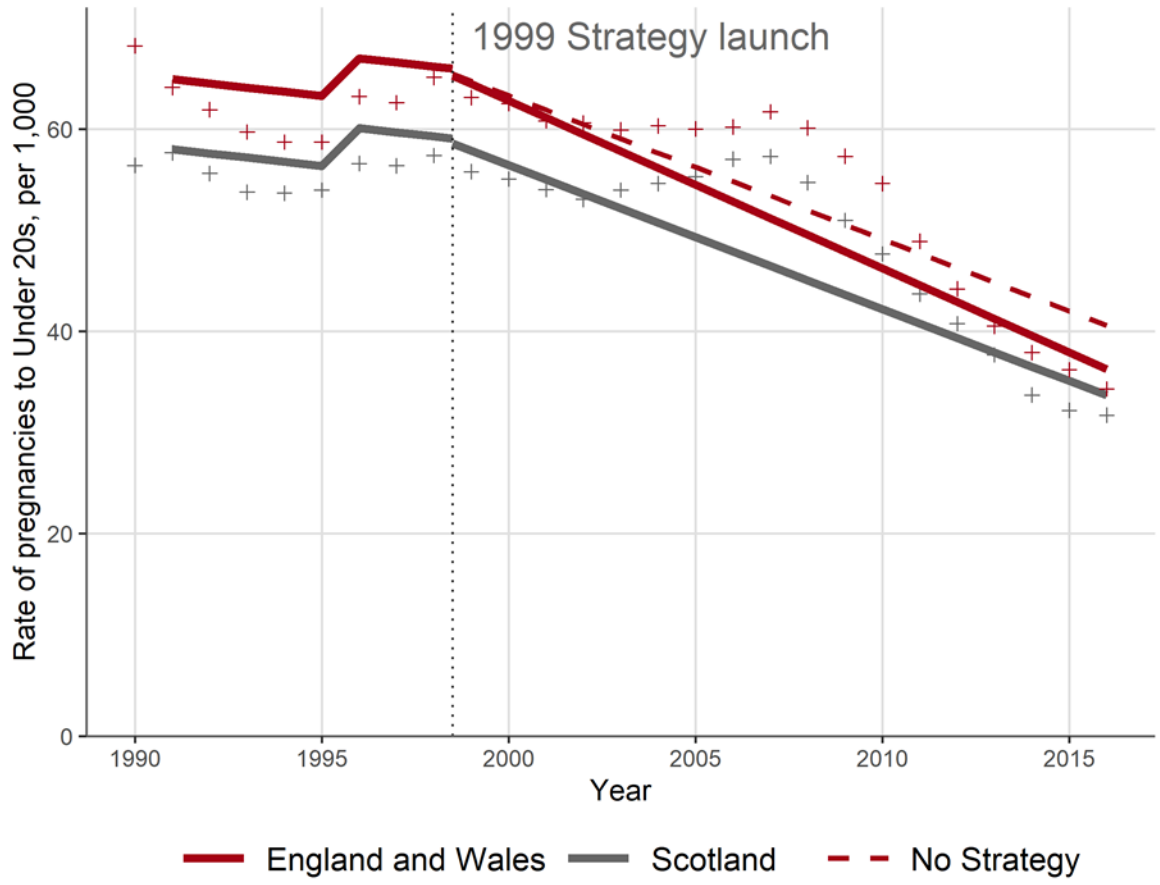

Figure 5-12 – England and Wales combined under-20 pregnancies compared with Scotland, 1994 to 2016. Autocorrelation correction: AR1, MA0; MSPE = 22.2; R^2 = 0.88; Log-likelihood = -101.0. Model 10 can be viewed at **this link**.

Coefficient	Estimate	Standard Error	Lower 95% CI	Upper 95% CI
Scotland (est.) rate at 1990	58.40	4.28	50.02	66.79
Scotland pre-intervention trend	-0.41	0.47	-1.33	0.51
England and Wales difference in rate at 1990	6.91	5.38	-3.63	17.46
Scotland change in level at 1999	0.00	1.86	-3.64	3.64
Scotland change in trend at 1999	-1.01	0.61	-2.20	0.18
England and Wales difference in level from control at 1999	-0.08	2.63	-5.23	5.07
England and Wales difference in trend from control at 1999	-0.23	0.44	-1.09	0.62
'Pill Scare' corrector	4.13	1.40	1.40	6.87

Table 5-10 – Fitted terms and confidence intervals of model 10 fitted to Scotland and combined England and Wales under-20 pregnancy rates 1991-2016.

5.3 Discussion

5.3.1 Summary of results

Across all comparative models, my results were consistent with a null effect of the Strategy. In preparatory models using only England data, I found evidence of significant changes of trends across the Strategy period, with a steeper downward trend from 2008 onwards. There were no clear indications of a level change in 1999 at Strategy launch. These observations were consistent with previous hypotheses attributing these results to the effects of launching, incrementally implementing, and later revising the Strategy (Wellings, Palmer, Geary, *et al.*, 2016). In comparisons with Scotland and Wales however, I found that all changes appeared to occur across all three countries, suggesting that England may have seen these decreasing rates in the absence of the Strategy. These results were consistent across all sensitivity and robustness checking using both comparators, all three age groups and all models with additional predictors (Figure 5-13 and Figure 5-14).

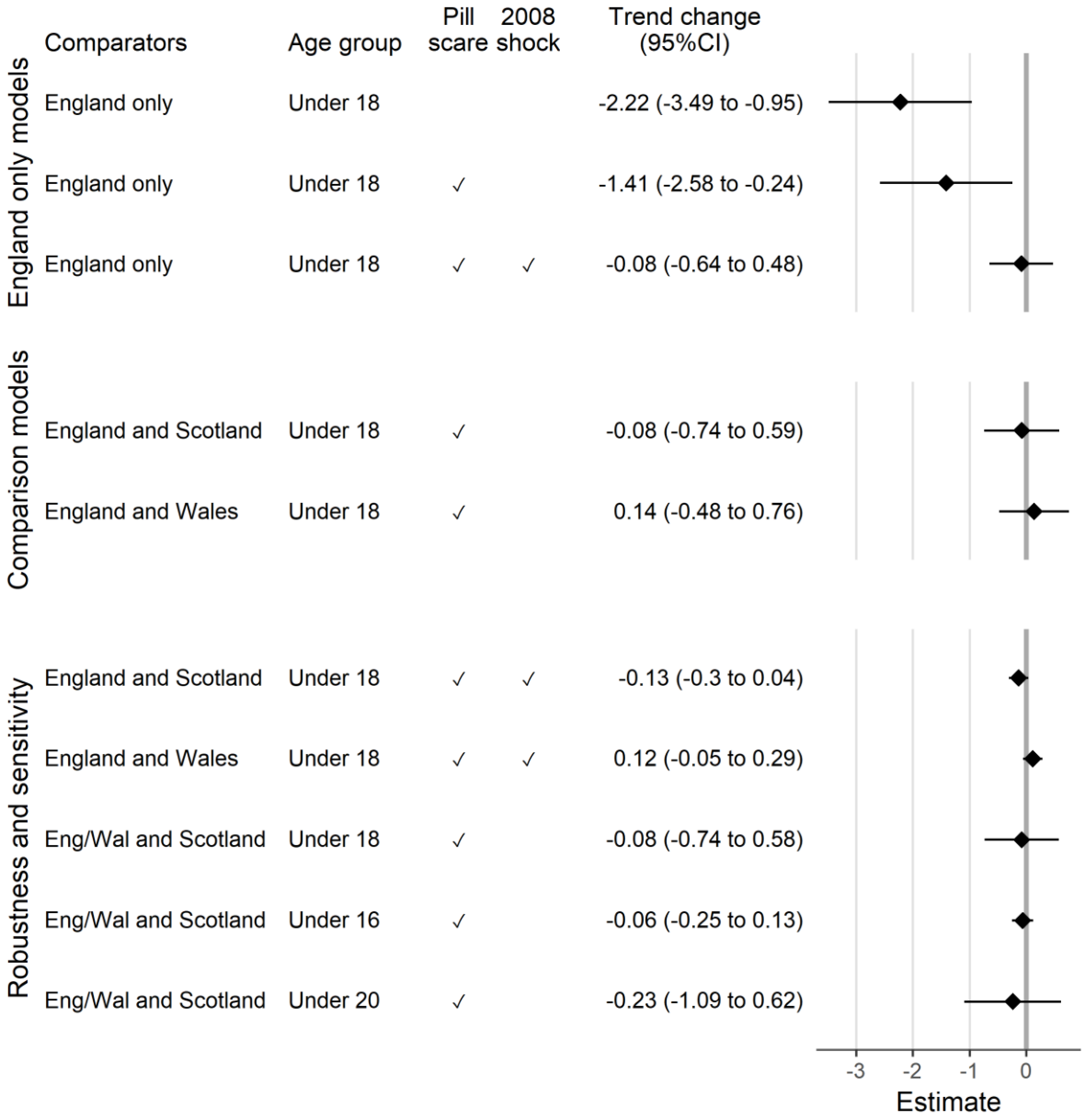

Figure 5-13 – Forest plot of trend change coefficients estimating Strategy effect size at 1999, arranged by model and model type. Figures and graphs in the right-hand columns represent the fitted terms for a change in trend at 1999 in England-only models. In comparison models, figures represent differences in trend change at 1999. England only models converge on zero as further correctors are added. All comparison models, including robustness and sensitivity analyses are close to zero and consistently estimate a small effect size, consistent with a hypothesised null effect of the Strategy.

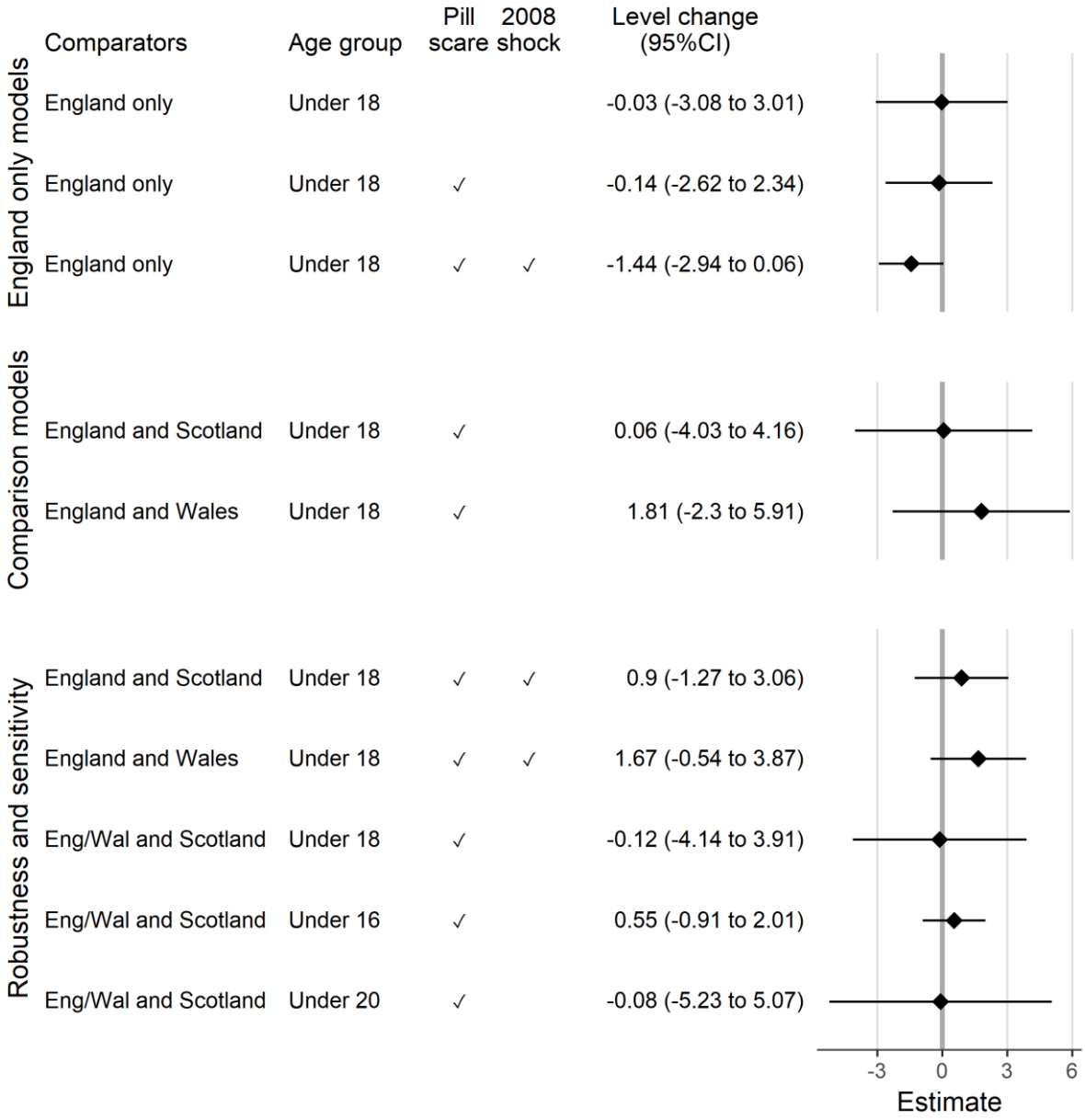

Figure 5-14 – Forest plot of level change coefficients estimating Strategy effect size at 1999, arranged by model and model type Figures and graphs in the right-hand columns represent the fitted terms for level change or difference in level change from control at 1999. No clear indications of a level change or difference in level change at 1999 due to the Strategy are evident across all models.

The similarity of change at 1999 experienced by all three countries suggests that these changes were the result of a common cause, independent from Strategy exposure. Likewise, the common shock predictor appears to represent a change seen across all three UK countries producing a similar effect on pregnancy rate trends, independent of Strategy exposure.

5.3.2 Strengths and limitations

I used publicly available, reliable data from several sources to produce comparable measures of effect across Scotland, England and Wales. I outlined my methods in an openly accessible protocol to ensure that my analyses followed a pre-specified and

transparent framework for testing my hypotheses (**osf.io/tdbr8**). I used two online repositories for all data and code used – GitHub and OSF – to ensure my analyses were clearly presented, publicly accessible and repeatable, following the principles of open research (UKRI, 2020). I created and published an interactive web app for all aspects of running models and creating outputs to ensure each finding could be replicated and verified without prior knowledge of R (see Appendix A for all links to online material).

My models produced well fitting predictions of pre-intervention pregnancy rates across England, Wales and Scotland using interrupted time series methods. These analyses suggest that the segmented linear regression model approach is a good representation of existing trends in observed data in the lead up to the Strategy and is sufficiently sensitive to detect exposure-induced level and trend changes in teenage pregnancy rates.

Data were limited in terms of periods of observation for each age group. Under-18 pregnancy rates for England alone represented my primary outcome. However, in published data, these data were only available for seven pre-intervention time points (Office for National Statistics, 2019). A minimum of eight time points for interrupted time series analyses are usually recommended; my primary models may have lacked power to detect small changes (Zhang, Wagner and Ross-Degnan, 2011; Penfold and Zhang, 2013; Lopez Bernal, Cummins and Gasparrini, 2017). Sensitivity analyses using England and Wales data with more pre-intervention time points were used to account for this limitation and results were consistent with primary models.

The outcome measures for each analysis had several limitations. Rates calculated for the UK using the ONS methods of adding births, still births and recorded abortions are not able to account for miscarriages and illegal abortions (Office for National Statistics, 2017). As all three countries had similar laws, healthcare and access to abortion clinics, the assumption that the vast majority of abortions were recorded and legal is likely equally valid in each case (as applied to England and Wales in recording rates; Office for National Statistics, 2017). Previous estimates of miscarriages from birth and abortion data suggest they equal approximately one in eight confirmed pregnancies (Office for National Statistics, 2017; NHS, 2018). In accounting for miscarriages from U.S.A. data, Sedgh et al. estimate these as 20% of births plus 10% of abortions (2015). Whilst likely not generalisable to the U.K., this calculation suggests that large changes in proportions of pregnancies ending in abortion may affect the proportion of pregnancies resulting in miscarriage. Proportions of pregnancies ending in abortion for under-18s remained

relatively stable across the period (with a slight increase in England and Wales, not seen in Scotland; Wellings, Palmer, Geary, *et al.*, 2016; Information Services Division Scotland, 2019). This suggests that all three countries have seen a relatively consistent and similar rate of pregnancies ending in miscarriage if these assumptions were valid. As illegal abortions were likely to be rare events, and rates of miscarriage were likely to be stable, I judged that rates of pregnancies not ending in recorded abortion or live birth would therefore be unlikely to have been differentially distributed across countries and years. Such missing data therefore would produce negligible bias.

Comparative analysis can allow us to pick out the effects of exposure to the Strategy from other changes occurring over a wider range of populations. The use of Scotland and Wales as comparators can account for potential unmeasured confounders that affect all three countries. These comparisons allow us to suggest that remaining differences between exposed England and unexposed control are a measure of the additional effects of the Strategy. The lack of remaining differences suggests that the targeted exposure of the Strategy on England produced little or no detectable difference than would have happened otherwise. The observed changes were likely a result of other wider-scoping interventions or changes in social determinants.

Previous estimates using England-only data could potentially have missed other underlying causes. Interventions outside of the Strategy which were applied to the whole of the UK, or other cultural and economic changes may have affected England's rates during the Strategy period. These may have produced changes which could be mistaken for Strategy effects.

In previous evaluations of the Strategy, researchers posited that Scotland and Wales may show changes in pregnancy rates similar to those observed in England in the Strategy period due to simultaneous interventions or spill over effects of Strategy-related activity across the whole of the UK (Wilkinson *et al.*, 2006; Wellings, Palmer, Geary, *et al.*, 2016; Wellings, Palmer, Wilkinson, *et al.*, 2016). This could be the case, for example, if the Strategy's media campaign was predominantly responsible for the effects in England, with Wales and Scotland being exposed to these messages, whilst other elements of the Strategy had little or no effect. Contamination is a possibility, but any spill over effects should be weaker than the effect of direct exposure to the Strategy, consistent with an expected dose-response effect of more intense action and focus on England. My analyses may not have been sensitive to detect a small additional effect in England associated with full exposure to the Strategy. However, Wellings et al. (2016) report on the differential effects associated

with Strategy spending between local authorities in England. Every additional £100 of local implementation funding spent per girl on Strategy activity was associated with a 6.2% reduction in pregnancy rates across the Strategy period (Wellings, Palmer, Geary, *et al.*, 2016; adjusting for region and deprivation). If this association represented a causal relationship, such differences should also be evident in cross-border comparisons, with higher spending in England resulting in greater decreases than in Scotland and Wales. Yet the trends in England, Scotland and Wales are all very similar and appear to be independent of Strategy activity. These observations contradict the suggestion of a strong spill over effect.

A further suggestion that simultaneously launched interventions in Scotland and Wales may have produced similar changes masking Strategy effectiveness relies on all three experiencing the same size of effect (see Table 3-1 for a comparison of policies across the UK). An evaluation of Scotland's co-occurring SHARE programme to prevent teenage pregnancies from 1996 to 1999 showed 26.4 more pregnancies per 1,000 women in SHARE schools than control (unadjusted; 95%CI: -33.3 to 86.2; Henderson *et al.*, 2007). These results do not suggest that the programme could cause a large decrease in pregnancy rates in Scotland across 1999-2016. It is unlikely that SHARE as a simultaneously launched program, comparable to the Strategy, could have produced a magnitude of effect in Scotland that could mask a substantial effect of the English Strategy on English rates. Further interventions in Scotland during the Strategy period also saw little change associated with exposure (Elliott *et al.*, 2010). Wales launched a simultaneous sexual health strategy, albeit without teenage pregnancy as an exclusive focus and with no set target for rate reduction (Wilkinson *et al.*, 2006, Webappendix). As a number of interventions did occur, and effect estimates were not available for all exposures, it cannot definitively be shown that Scotland and Wales show rates reflecting a null-effect England. Other methods, comparing England with a wider pool of comparable countries, would be able to test whether Scotland and Wales as comparators have produced a biased result through other unobserved changes, independent of the Strategy and not applied to England.

Early evaluations of the Strategy reported a change in trend at 1999 which was not sufficient to achieve the 10-year goal of reducing pregnancies by 50% and suggested this slow start was partly due to incremental implementation across initial years (Wilkinson *et al.*, 2006). Later evaluations finding a delayed achievement of Strategy goals highlight the 2008 change in trend, suggesting that this was a result of improvements in Strategy output following the earlier review (Hadley, Ingham and Chandra-Mouli, 2016; Wellings, Palmer,

Geary, *et al.*, 2016). Model 5, in the absence of controls, similarly showed small changes in 1999 and large changes in 2008, however it was unable to test whether this change is attributable to the Strategy. In comparison models allowing a common trend change in 2008, a similarly sized change was seen in Scotland and Wales. This similarity of effect size suggests this unexplained trend change was also a common effect across all three countries and unrelated to the Strategy.

5.4 Conclusion

Using controlled interrupted time series methods to detect change attributable to Strategy exposure, I was unable to detect any effects of the intervention on teenage pregnancy rates. The large decline of more than 50% across the observed period (1999-2016) was experienced across Scotland, England and Wales. The decline cannot therefore be attributed to the Teenage Pregnancy Strategy. Spill over effects would have to be very strong to account for very similar trends in Scotland and Wales, and their hypothesised equal effect on Scotland and Wales is inconsistent with the differences seen across local authorities in England.

Chapter 6 Comparing England and Wales with synthetic control

6.1 Overview of chapter

In this chapter, I report the results of my synthetic control analyses. I use a selection of European and English-speaking high-income countries to compare with England and Wales to determine the effects of Strategy exposure. A weighted mean of other countries' rates is fitted to England and Wales' pre-intervention observations. The same set of weights is then used to combine post-implementation observations and predict England and Wales' rates in the absence of the Strategy. I examine the exposed and control trends for differences attributable to the Strategy. I also construct synthetic controls for each unexposed country as placebos. I compare the post intervention difference between England and its synthetic control with the distribution of differences across the set of placebo comparisons. I use this comparison to test whether exposure to the Strategy produced a greater deviation from control than that expected due to random variation in rates. I use several sensitivity analyses to test the models and examine the consistency of the results.

6.2 Results

6.2.1 Exploratory visualisation of teenage birth and pregnancy rates across pre-intervention years 1990-1998 and across all years of data

I extracted data and calculated estimates of under-18 birth rates and under-20 pregnancy rates for the 16 high-income countries listed in Box 6-1. Data was available across the years 1990 to 2013, giving multiple time points either side of the 1999 Strategy launch.

Denmark	Norway
Finland	Portugal
France	Scotland
Germany	Spain
Iceland	Sweden
Italy	Switzerland
Netherlands	England and Wales
New Zealand	The United States of America

Box 6-1 – The sixteen countries with data included in synthetic control analyses recording under-18 birth rates and under-20 pregnancy rates. I used the 'England and Wales' combined rate as an estimate of rates in England which was exposed to the Strategy.

In assessing the suitability of the data which I collected for the synthetic control analysis, I first visualised the data for all countries across the pre-intervention years 1990 to 1998 to assess similarities in the unexposed period, blind to post-intervention rates. Figure 6-1 shows the initial plot of under-18 birth rates for all included countries by year from 1990 to 1998. England and Wales were seen to have comparatively high birth rates across the whole pre-intervention period, with only New Zealand, Scotland and the United States of America producing higher rates. As England and Wales do not have the highest or lowest rates in any pre-intervention year, these data provide observations of countries with both higher and lower rates; this forms a 'convex hull' around the exposed unit (Abadie, Diamond and Hainmueller, 2010). These data allowed me to construct a synthetic control from positive weights summing to 1. The model did not need to assign any country a negative weight, hypothesising an inverse effect on England of a common event, or assign a weight greater than 1, thus extrapolating the effects from a low rate country to estimate the impact on England's higher rate. This thereby allows more confidence in the validity of the predicted control (Abadie, Diamond and Hainmueller, 2011, 2015).

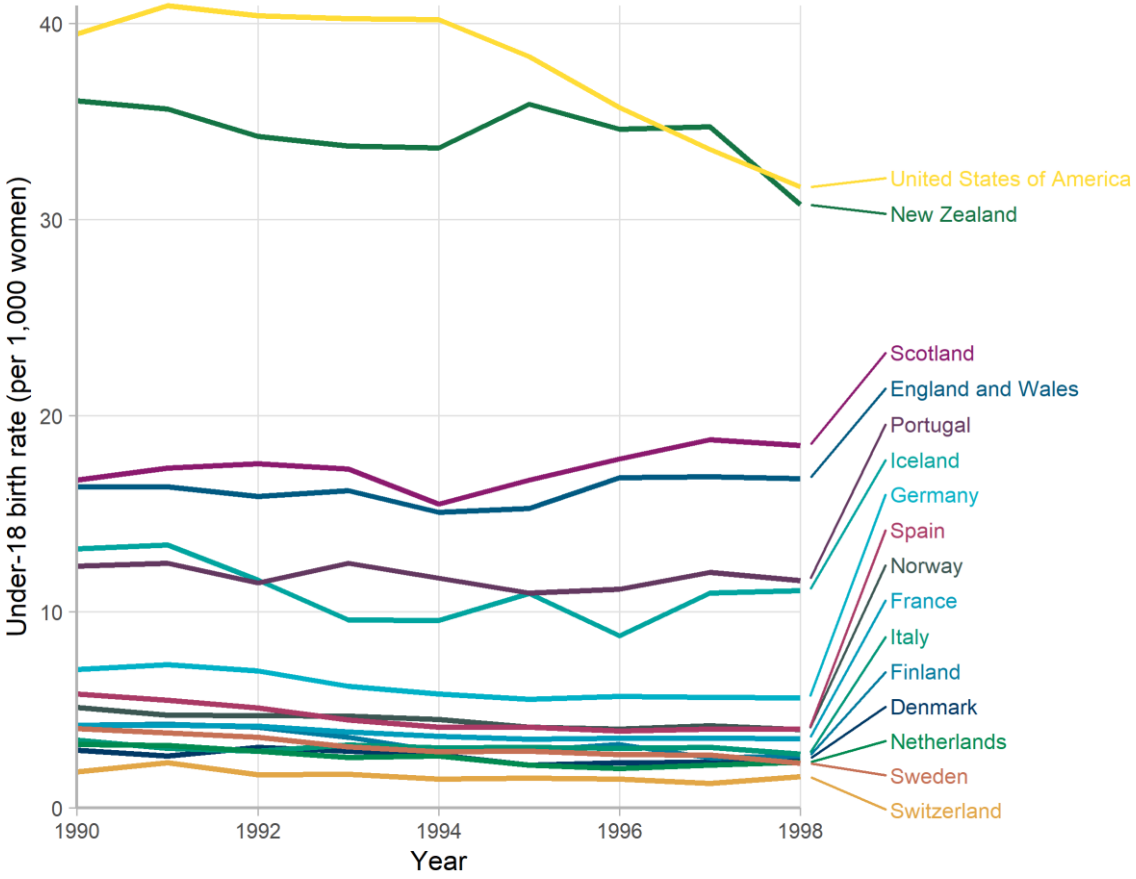

Figure 6-1 – Under-18 birth rates across 16 high-income countries, 1990 to 1998

In Figure 6-2 presenting the under 20 pregnancy rates for all countries across pre-intervention years, England and Wales similarly showed relatively high rates, with only the United States of America showing higher pregnancy rates. England and Wales were still within the convex hull of control countries across the entire pre-intervention period.

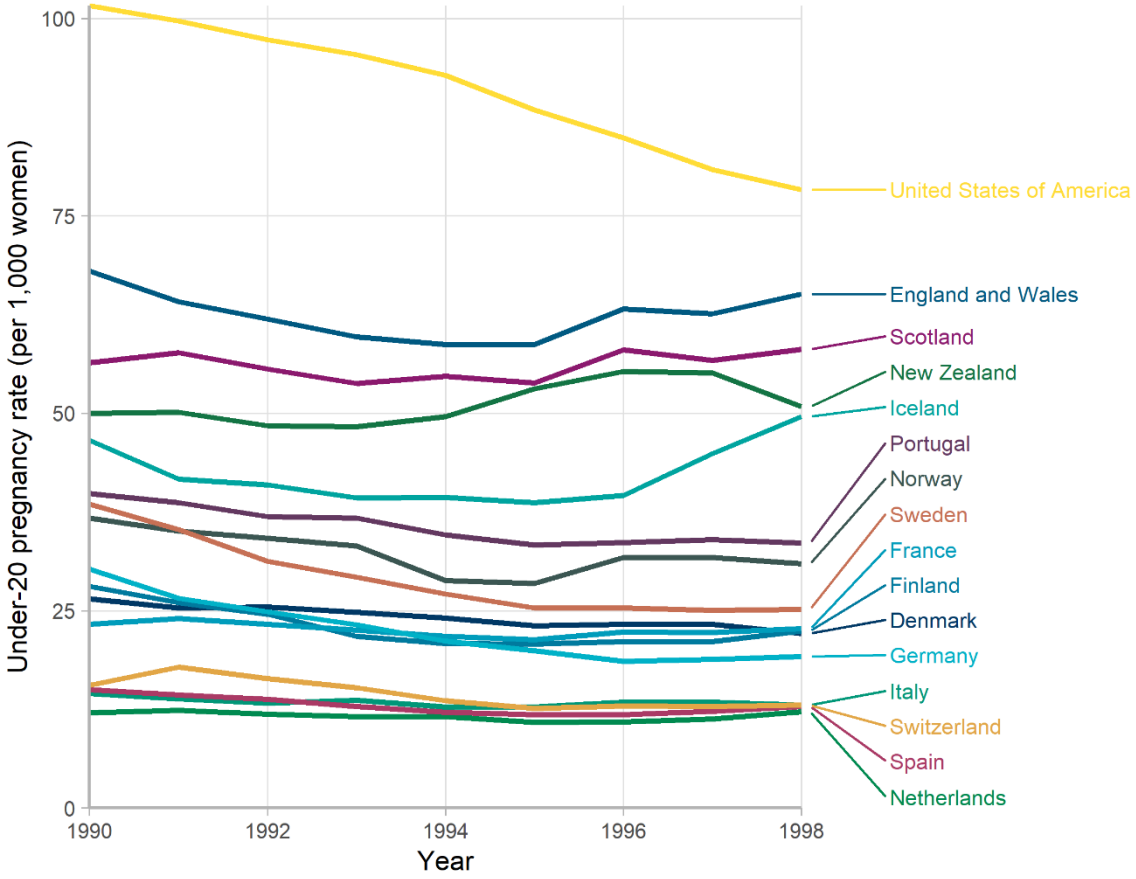

Figure 6-2 – Under-20 pregnancy rates across 16 high-income countries, 1990 to 1998

Having examined the pre-intervention period, I then produced Figure 6-3 and Figure 6-4 graphing all outcome data for all countries used in analyses. In comparing under-18 birth rates, New Zealand shows a distinct rise in rates from 2002 to 2008. In comparing under-20 pregnancy rates, England and Wales showed higher rates than all comparators across 2007 to 2013.

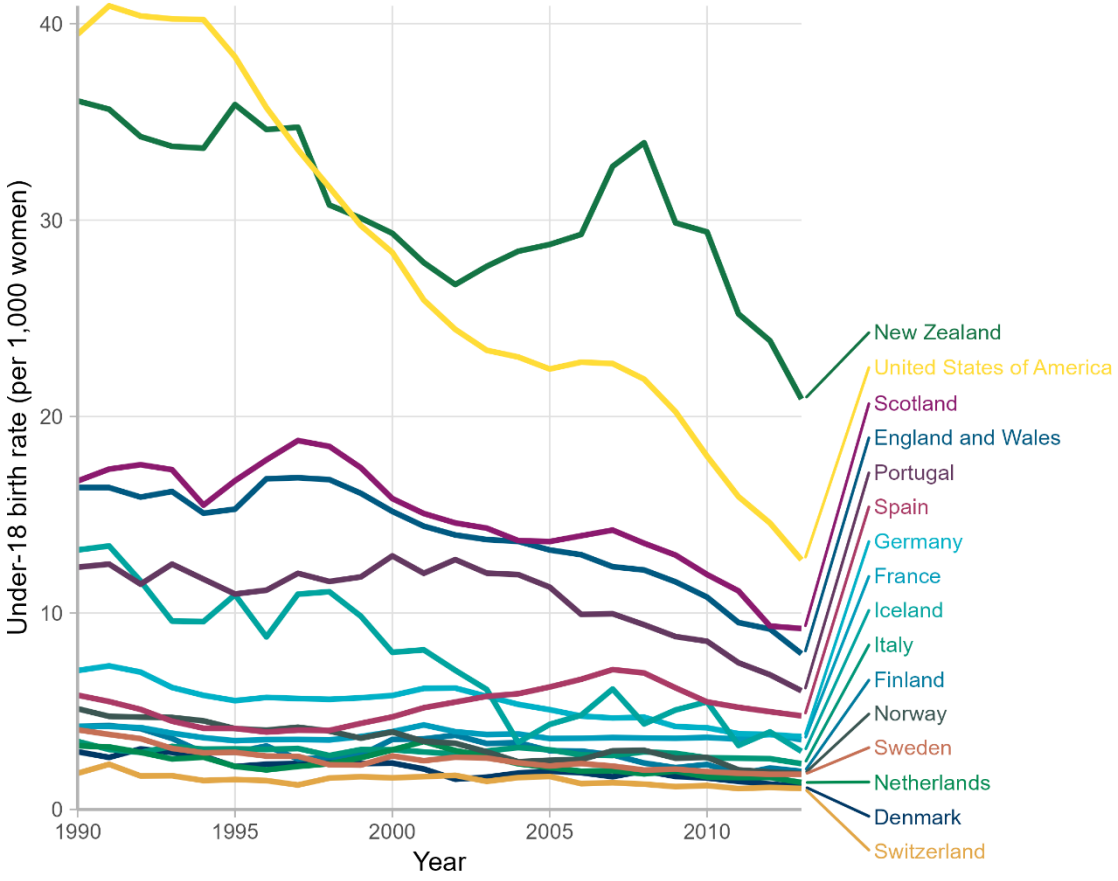

Figure 6-3 – Under-18 birth rates across England and Wales and all donor countries, 1990-2013

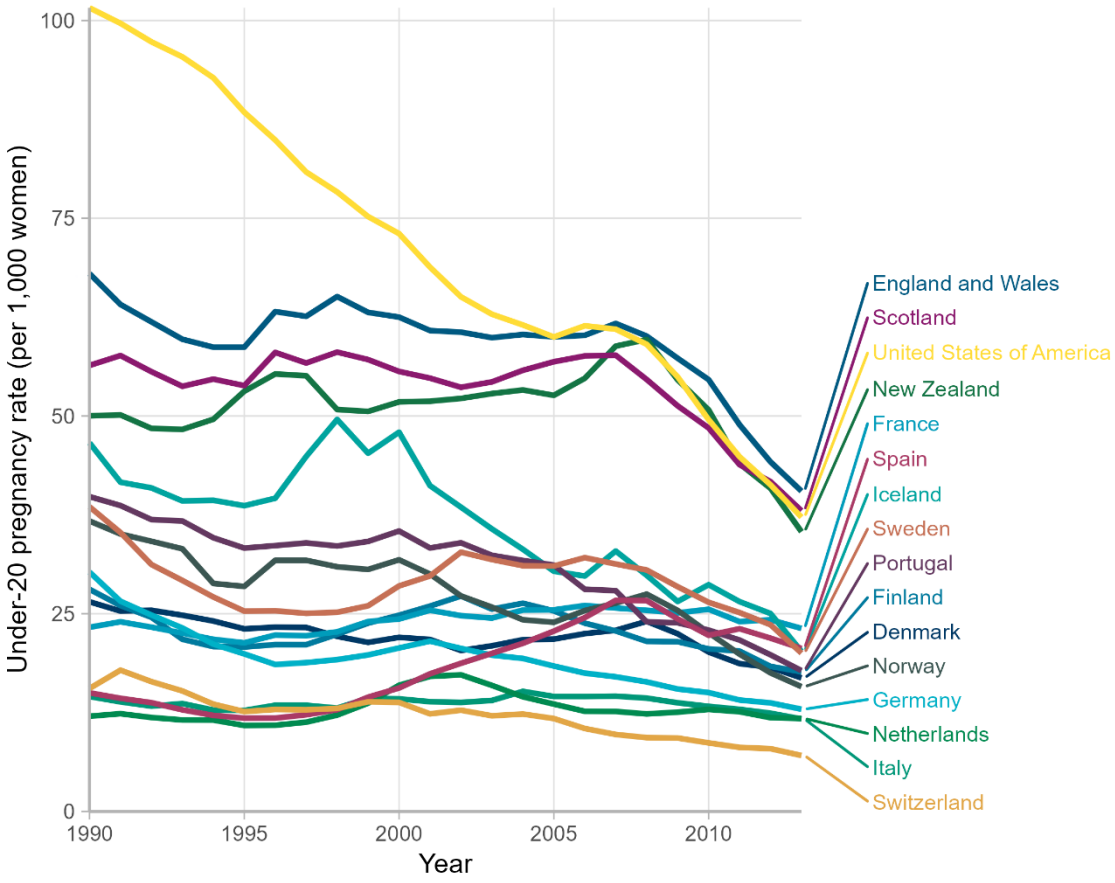

Figure 6-4 – Under-20 birth rates across England and Wales and all donor countries, 1990-2013

6.2.2 Interactive presentations of analyses

To enable interactive repetition of my synthetic control models, I created a shiny app published online at **phd.andybaxter.me/synth-app** with all data used in these analyses. All model graphs and tables were produced using R and combined into an html document published at **phd.andybaxter.me/Synth-analysis** (with links to interactive graphs).

6.2.3 Model 1 – Under-18 birth rates across all countries

In my primary synthetic control model, I fitted a control to the pre-intervention period, matching controls using pre-intervention birth rates as the only predictor. Iterating across clusters of pre-intervention years, I was able to construct a good-fit synthetic control using the means of four time periods as predictors (1990-1993, 1994, 1995, 1996-1998), minimising the error in prediction of yearly birth rates. The prediction error of this model in the pre-intervention period was small in proportion to actual observed rates, averaging 0.30 births per 1,000 women per year around a mean of 16.2 births across the 9-year period (MSPE = 0.092). The control was constructed predominantly from Scotland (weight of

67.2%), Portugal (29.5%), the U.S.A. (1.6%) and New Zealand (1.2%). Using these weights to construct post-implementation trends, I observed that across this period the predicted control closely followed observed England and Wales rates (Figure 6-5).

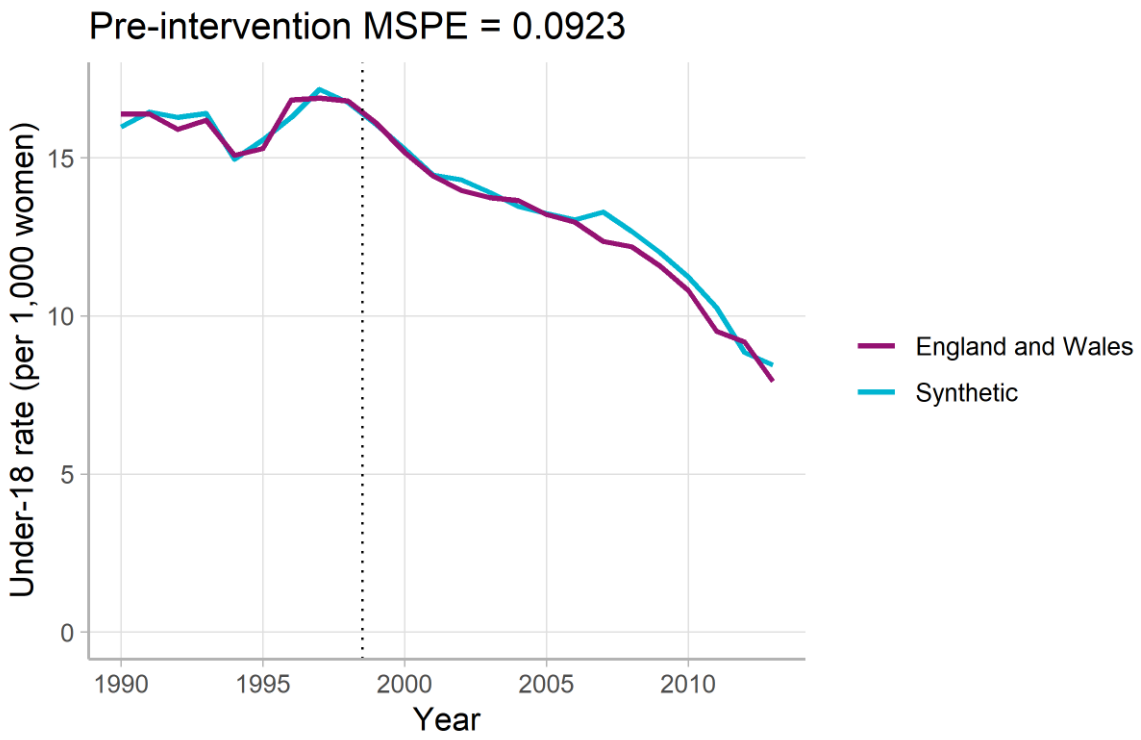

Figure 6-5 – England and Wales under-18 birth rates compared with synthetic control, 1990-2013; model 1 featuring all countries and only pre-intervention rates as predictors. The dotted line marks the introduction of the Strategy in 1999.

The *synth* function assigned similar weights of 33.3%, 29.5% and 36.2% to each of the latter three clusters of years respectively, assigning negligible weight to the period 1990-1993 (Table 6-1). Alongside weighted Scotland, Portugal, the U.S.A. and New Zealand, all other countries were given zero or negligible weighting.

Donor weights		Predictor weights	
Country	**Weight**	**Predictor**	**Weight**
Scotland	0.672	Birth rates 1990-1993	0.01
Portugal	0.295	Birth rates 1994	0.333
United States of America	0.016	Birth rates 1995	0.295
New Zealand	0.012	Birth rates 1996-1998	0.362
Iceland	0.001		
Sweden	0.001		
Switzerland	0.000		
Germany	0.000		
Denmark	0.000		
Spain	0.000		
Finland	0.000		
France	0.000		
Italy	0.000		
Netherlands	0.000		
Norway	0.000		

Table 6-1 – Weightings applied to control countries' observations to calculate predicted rates, and weightings of predictors used in model 1 comparing under-18 birth rates across 1990-2013. Donor weights were calculated to minimise the MSPE from observed England and Wales in the pre-intervention period 1990-1998 then applied to post-implementation rates 1999-2013 to calculate a weighted mean birth rate as a predictor of unexposed England and Wales. Means across clusters of years were tested for their predictive power of the outcome and weighted to minimise MSPE across the pre-intervention period. Each column is independent and sums to 1.

Removing England and Wales from the group of included countries, I repeated synthetic control analyses for each donor country, setting it as the exposed unit. I calculated gaps between observed and predicted yearly rates to test the deviation from control expected under placebo conditions. I plotted the gaps for the set of countries with similarly well-fit predictors (less than 5 times the pre-intervention MSPE of England and Wales) alongside exposed England and Wales' gaps to compare the range of deviation from predictions in the post-intervention period. If England and Wales saw a large effect in the Strategy period then the trend would appear below the spread of placebo countries' gaps. When I compared England and Wales post-intervention gaps with placebo results from other countries, England and Wales trends showed comparatively little difference from control (Figure 6-6).

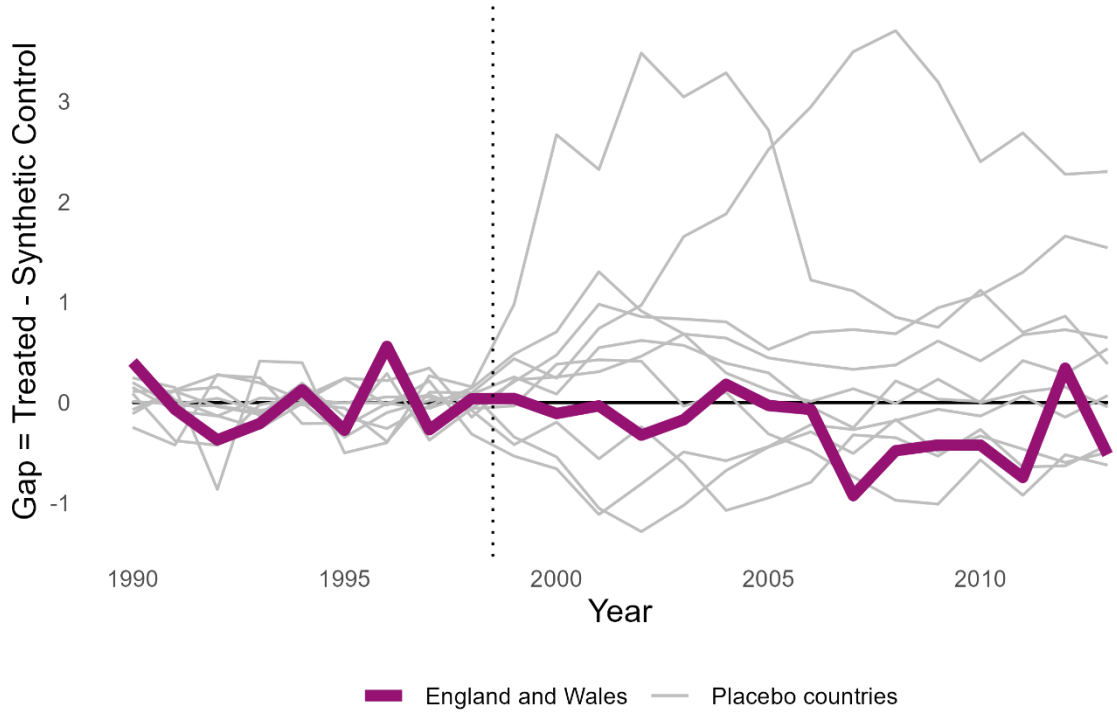

Figure 6-6 – Gaps plot for model 1, comparing difference from control at each time point across well fit donor countries. The bold coloured line plots the difference between England and Wales and its synthetic control at each time point. The light grey lines plot the differences for the placebo countries. The Strategy launch is marked in 1999 with a dotted line. Countries with a pre-intervention MSPE greater than 5 times that of England and Wales were excluded: Switzerland, Scotland, Iceland, United States of America and New Zealand. Published interactively at **chart-studio.plotly.com/~andrewbaxter439/43**

I calculated post/pre-intervention MSPE ratios for each country to compare its magnitude of deviation from zero in post-implementation years. Using pre-intervention MSPE to quantify model fit and post-implementation MSPE to quantify the deviation from predictions produced by (placebo) exposure, this ratio produced a comparable set of effect estimates, standardised by goodness of pre-intervention fit. A larger ratio would indicate a relatively large effect size from a relatively well-fit predictor. I used these ratios to rank countries by effect size to test if England and Wales with the defined Strategy exposure produced a ratio larger than the range of null effect deviations experienced in unexposed countries. England and Wales saw a post/pre-intervention MSPE ratio of 1.86. Thirteen of fifteen comparison countries (87%) showed greater post/pre-MSPE ratios than England and Wales, indicating that the probability of a country experiencing a difference from control larger than England and Wales in the absence of the Strategy is at least $p = 0.87$ (Figure 6-7).

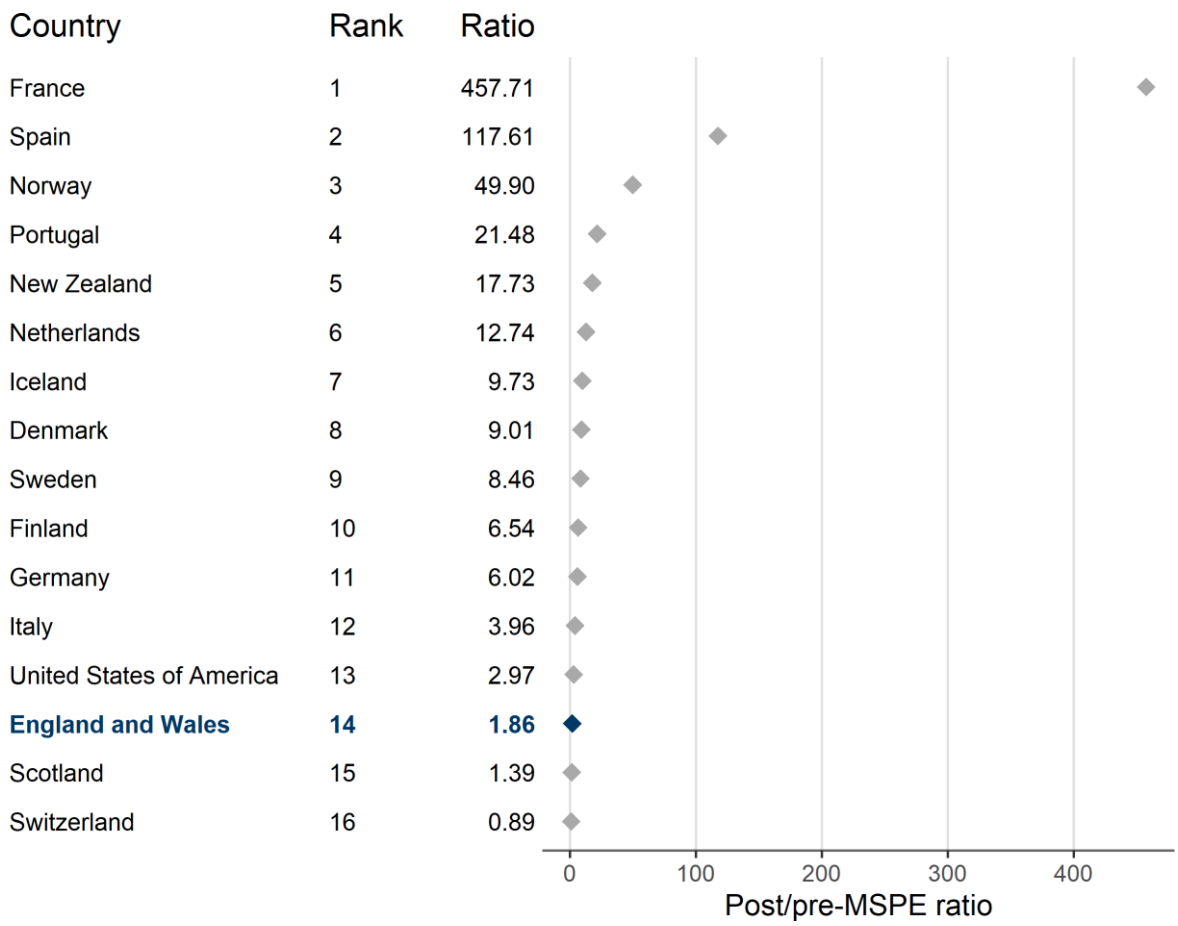

Figure 6-7 – England and Wales' post/pre-intervention MSPE ratio compared with the ratios of control countries for under-18 births; model 1. Thirteen countries which were not exposed to the Strategy during the period 1999-2013 showed a greater relative deviation from pre-intervention fit than England and Wales (p=0.87).

In time-based placebo tests, intervention dates set at 1995 and 1996 produced synthetic controls following much lower trends than observed England and Wales (Figure 6-8). All later years tested from 1997 to 2005 produced relatively stable synthetic control predictions, in line with true date predictions in Figure 6-5. This suggests that, aside from an event occurring in the exposed unit in 1996 leading to an increase in pregnancies, other causes in changes of rates were not unique to England and Wales. This is consistent with the pill scare hypothesis corrected for in ITS analyses in the previous chapter (Furedi, 1999). The inclusion of 1997 and 1998 as pre-intervention years is able to account for the discrepancy, potentially by assigning similarly exposed Scotland a large weight (Table 6-1).

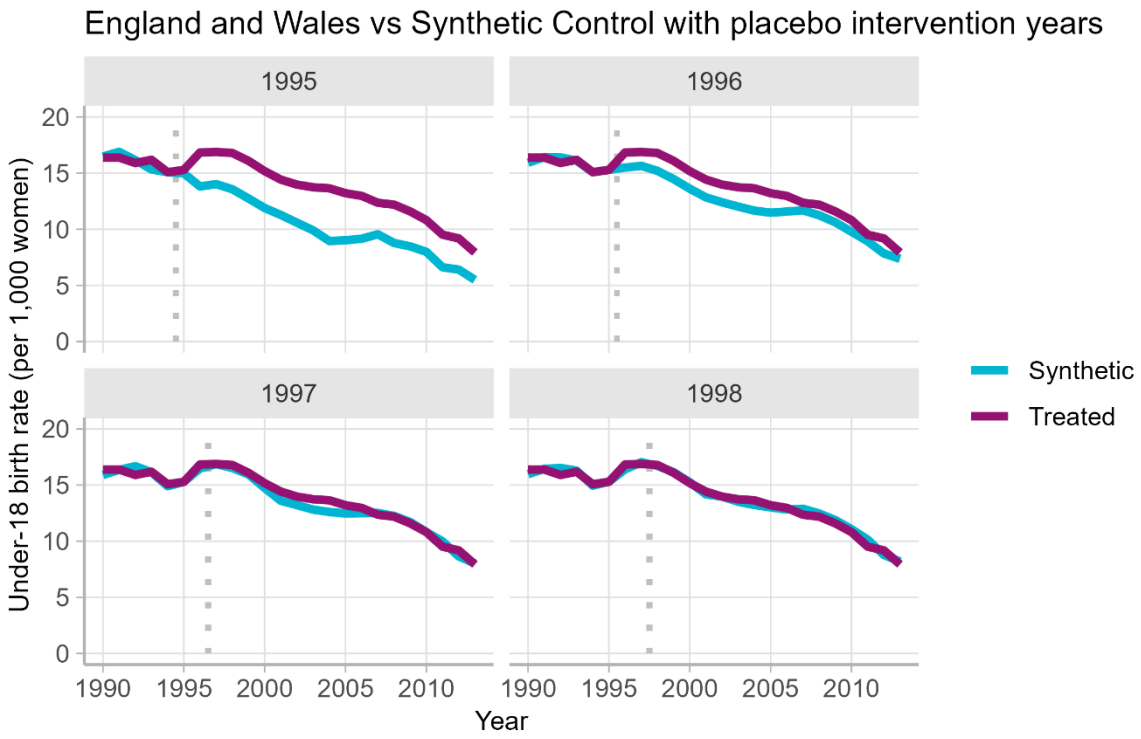

Figure 6-8 – Time-based placebo graphs for placebo years set to 1995-1998 in model 1. Dotted lines mark the simulated Strategy launch in each tested year. Further placebo years (2000 – 2005) produced stable trends matching the true Strategy year (1999). All years can be viewed interactively at **chart-studio.plotly.com/~andrewbaxter439/15**.

6.2.4 Model 2 – Under-20 pregnancy rates across all countries

In my second model, I used under-20 pregnancy rates as the outcome of interest. This metric has the advantage of including pregnancies ending in abortion – data which weren't available for under-18s. This outcome also includes pregnancies at older ages which were not considered Strategy priorities (outside the Strategy goal of 'halving the rate of conceptions among under 18s by 2010'; Social Exclusion Unit, 1999, p. 8). In the initial model fit I used only pre-intervention pregnancy rates as predictors and used a best fit grouping of years into three clusters: 1990, 1991-1995 and 1996-1998. The fitted synthetic control saw a poorer fit across pre-intervention years than the previous model, with an average prediction error of 2.1 pregnancies per 1,000 women across the nine pre-intervention years (around a mean rate of 62 pregnancies per 1,000 women; MSPE = 4.27; Figure 6-9). The control was constructed from Scotland (63.9%), the United States of America (22.7%) and Iceland (13.4%; Table 6-2).

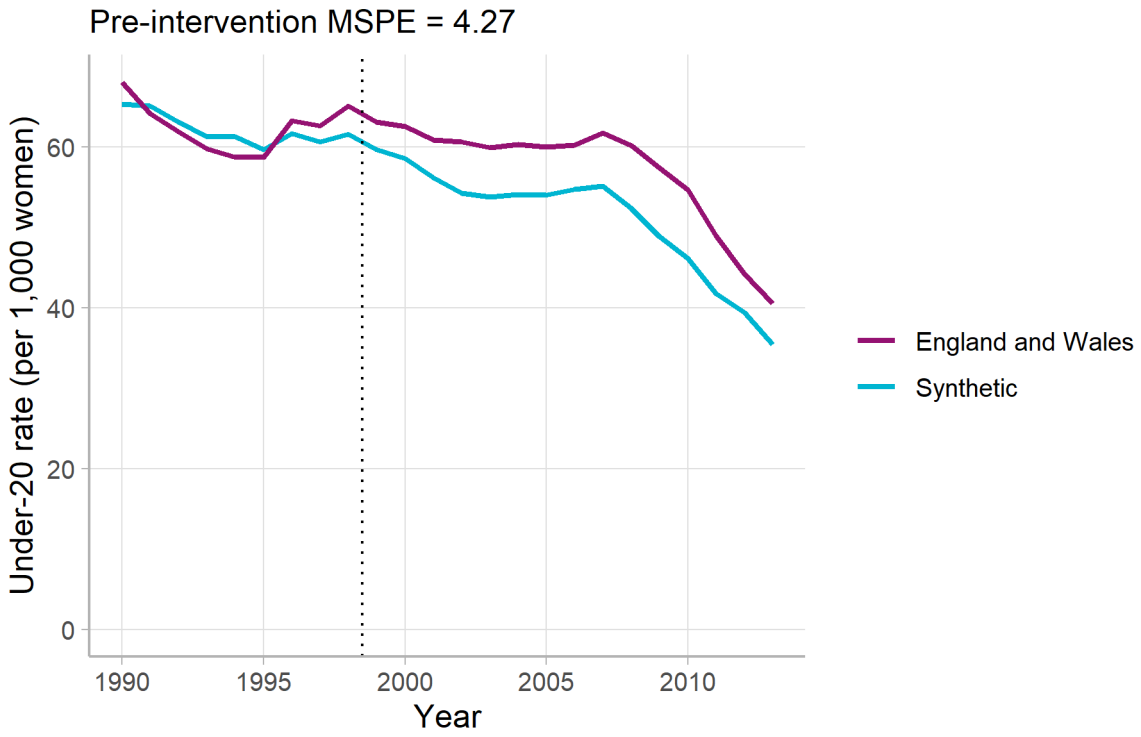

Figure 6-9 – England and Wales under-20 pregnancy rates compared with synthetic control, 1990-2013; model 2 featuring all countries and only pre-intervention rates as predictors. The dotted line marks the introduction of the Strategy in 1999.

Donor weights	
Country	**Weight**
Scotland	0.639
United States of America	0.227
Iceland	0.134
Switzerland	0.000
Germany	0.000
Denmark	0.000
Spain	0.000
Finland	0.000
France	0.000
Italy	0.000
Netherlands	0.000
Norway	0.000
Portugal	0.000
New Zealand	0.000
Sweden	0.000

Predictor weights	
Predictor	**Weight**
Pregnancy rates 1990	0.152
Pregnancy rates 1991-1995	0.578
Pregnancy rates 1996-1998	0.271

Table 6-2 – Weightings applied to control countries' observations to calculate predicted rates, and weightings of predictors used in model 2 comparing under-20 pregnancy rates across 1990-2013. Means across clusters of years were tested for their predictive power of the outcome and weighted to minimise MSPE across the pre-intervention period. Each column is independent and sums to 1.

During the Strategy period, England and Wales' rates were consistently higher than rates predicted from control. Figure 6-10 shows England and Wales' calculated differences from control alongside similarly calculated gaps from placebo analyses across unexposed countries. England and Wales' rates were higher than predicted across the intervention period. Most placebo countries saw lower rates relative to control than England and Wales. This comparatively large deviation from the null difference was balanced by a similarly large pre-intervention MSPE, as demonstrated by England and Wales' relatively low post/pre-MSPE ratio of 9.1 (Figure 6-11). Ten of the fifteen control countries saw larger ratios, meaning the probability of achieving a ratio larger than England and Wales in the absence of the Strategy is at least p = 0.67.

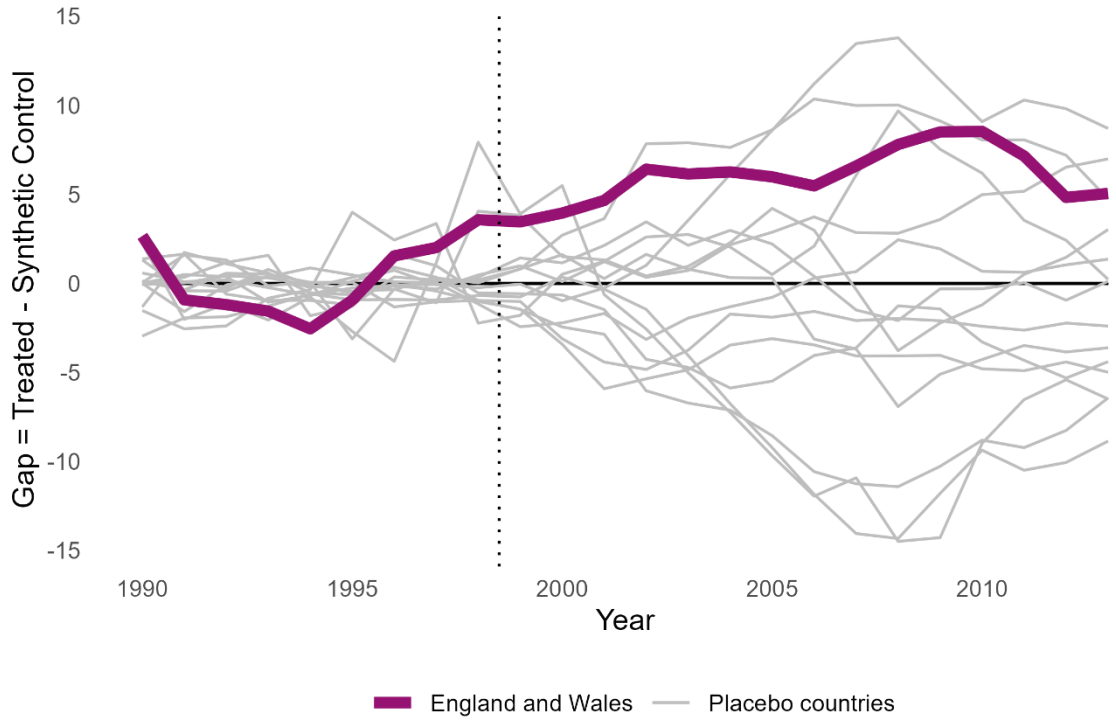

Figure 6-10 – Gaps plot for model 2, comparing difference from control at each time point across all donor countries. The Strategy launch is maked in 1999 with a dotted line. Countries with a pre-intervention MSPE greater than 5 times that of England and Wales were excluded: United States of America. Published interactively at **chart-studio.plotly.com/~andrewbaxter439/49**.

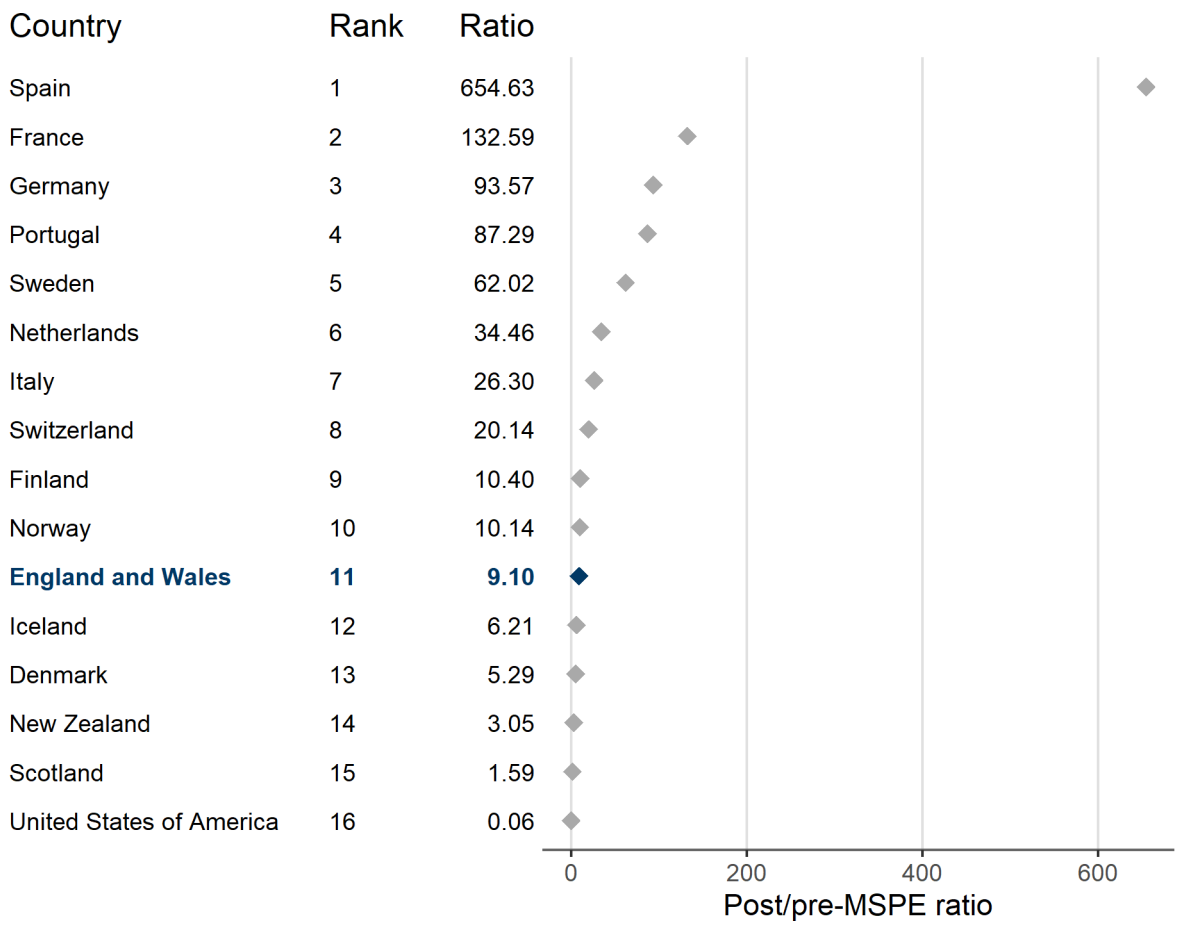

Figure 6-11 – England and Wales' post/pre-intervention MSPE ratio compared with the ratios of control countries for under-20 pregnancies; model 2. Ten of fifteen control countries had a higher post/pre-MSPE ratio than England and Wales (p = 0.67).

Time-based placebo tests showed similar patterns to model 1, with hypothesised intervention years earlier than the 1996 pill scare producing large gaps and higher pregnancy rates in England and Wales than in predicted controls (Figure 6-12). Post-implementation years gave relatively stable predictions from 2000 to 2005.

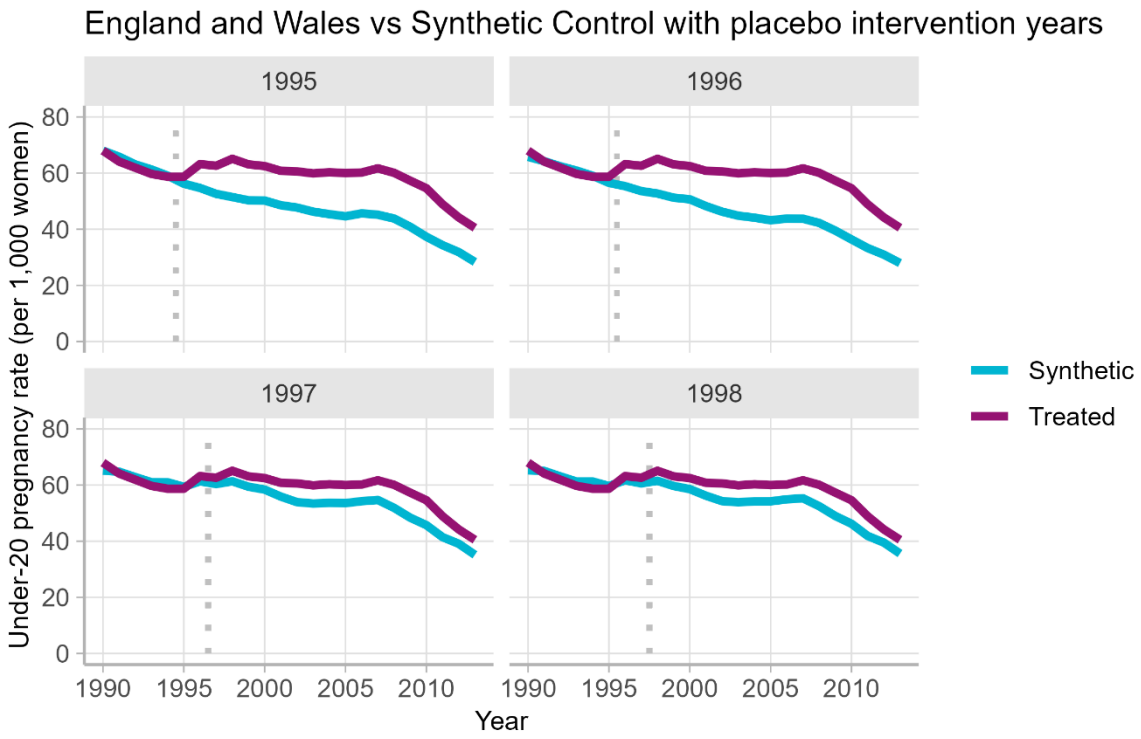

Figure 6-12 – Time-based placebo graphs for placebo years set to 1995-1998 in model 2. Dotted lines mark the simulated Strategy launch in each tested year. Similar to model 1, further placebo years (2000 – 2005) produced stable trends matching the true Strategy year (1999). All years can be viewed interactively at **chart-studio.plotly.com/~andrewbaxter439/104**.

6.2.5 Sensitivity analyses excluding Scotland

As highlighted previously, there are concerns regarding use of Scotland as a comparator. Scotland's falling rates during the Strategy period may be partially due to contamination from UK-wide effects of Strategy activity, or simultaneously implemented interventions (Wilkinson *et al.*, 2006; Craig, Dundas, *et al.*, 2016; Wellings, Palmer, Wilkinson, *et al.*, 2016). To test for my model's reliance on Scotland for construction of controls, I removed Scotland from the donor pool and repeated my analyses. This generated two further models for under-18 birth rates and under-20 pregnancy rates.

In model 3, the reduced set of included countries produced a poorer fit synthetic control in the pre-intervention period than the full model 1. The predictions were constructed from Portugal (80.7%) and New Zealand (19.3%; Table 6-3). The average prediction error rose to 0.78 births per 1,000 women per year around a mean of 16.2 births across the 9-year period (MSPE = 0.62; Figure 6-13). Across the majority of the Strategy period, England and Wales saw lower rates of birth than predicted from synthetic control.

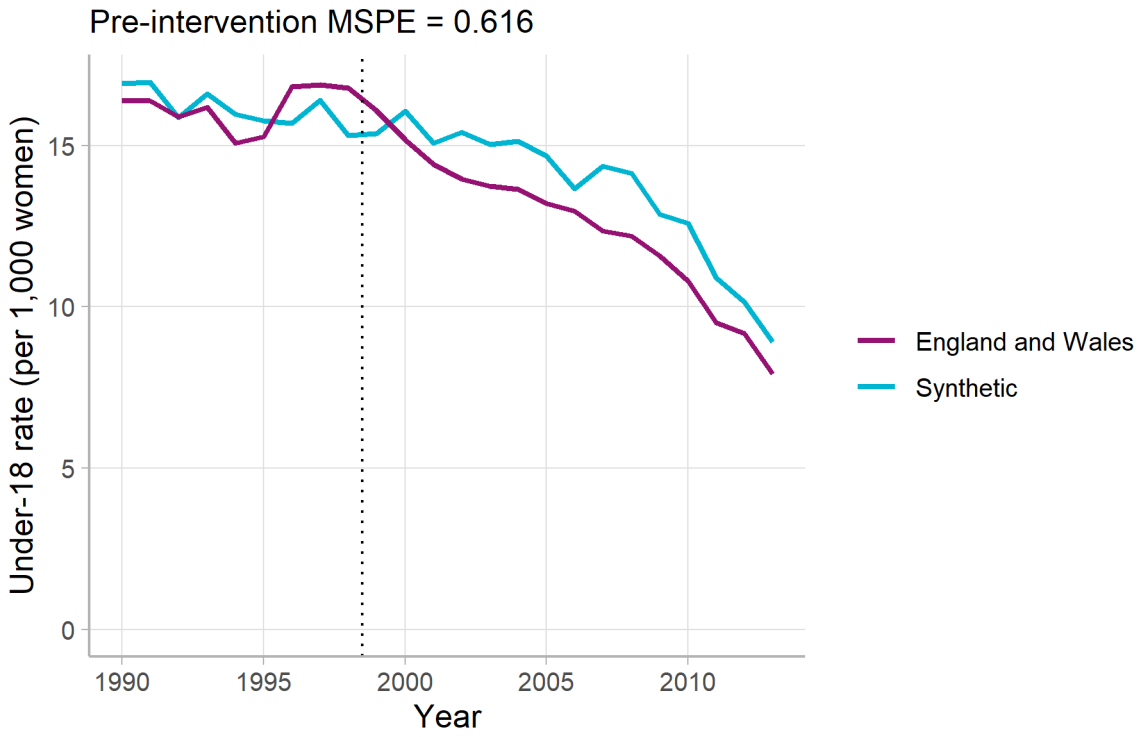

Figure 6-13 – England and Wales under-18 birth rates compared with synthetic control, 1990-2013; model 3 with Scotland removed and only pre-intervention rates as predictors. The dotted line marks the introduction of the Strategy in 1999.

Donor weights	
Country	**Weight**
Portugal	0.807
New Zealand	0.193
Switzerland	0.000
Germany	0.000
Denmark	0.000
Spain	0.000
Finland	0.000
France	0.000
Iceland	0.000
Italy	0.000
Netherlands	0.000
Norway	0.000
Sweden	0.000
United States of America	0.000

Predictor weights	
Predictor	**Weight**
Birth rates 1990-1993	0.206
Birth rates 1994	0.156
Birth rates 1995	0.318
Birth rates 1996-1998	0.320

Table 6-3 – Weightings applied to control countries' observations to calculate predicted rates, and weightings of predictors used in model 3 comparing under-18 birth rates across 1990-2013. Clusters of years used to calculate predictor weights were left unchanged from model 2. Each column is independent and sums to 1.

Gaps between observed and predicted rates are plotted for England and Wales alongside placebo countries in Figure 6-14. England and Wales' rates were consistently lower than predicted. Most other countries saw higher rates relative to their predictors across the period than England and Wales. However, in comparing post/pre-MSPE ratios to adjust for goodness of model fit, it appears that these difference of the observations from predicted are consistent with noise in the absence of an effect from the Strategy (Figure 6-15).

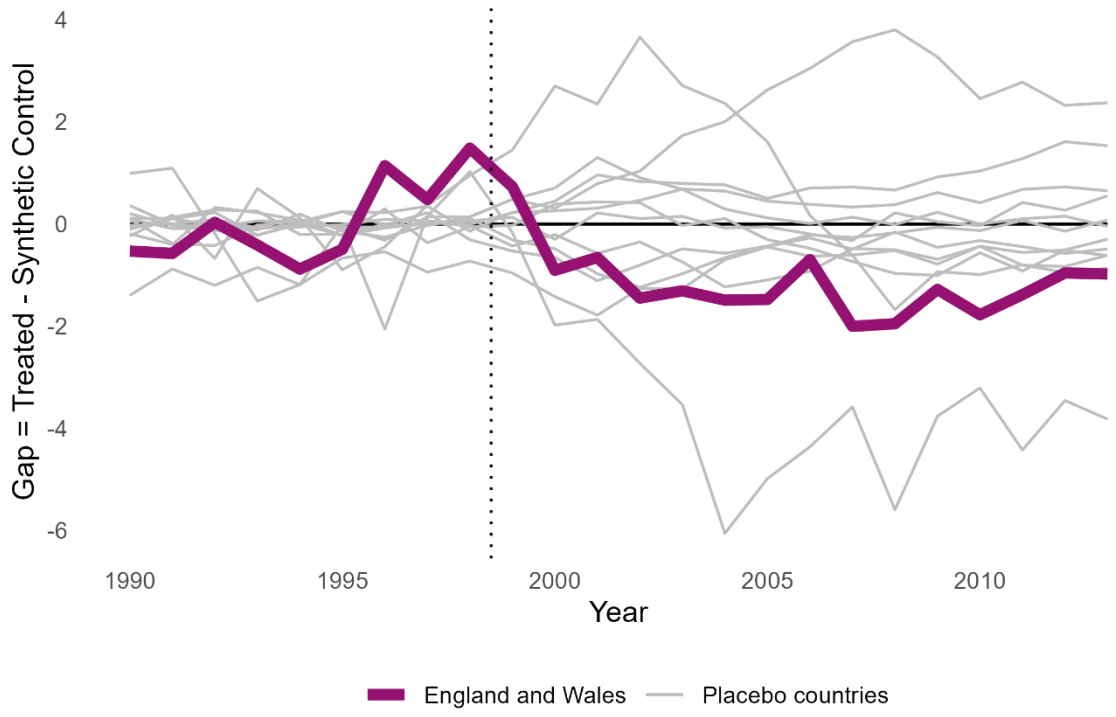

Figure 6-14 – Gaps plot of model 3, showing differences in under-18 birth rates of England and Wales from control, alongside placebo country tests The Strategy launch is maked in 1999 with a dotted line. Countries with a pre-intervention MSPE greater than 5 times that of England and Wales were excluded: New Zealand and United States of America. Published interactively at **chart-studio.plotly.com/~andrewbaxter439/61**

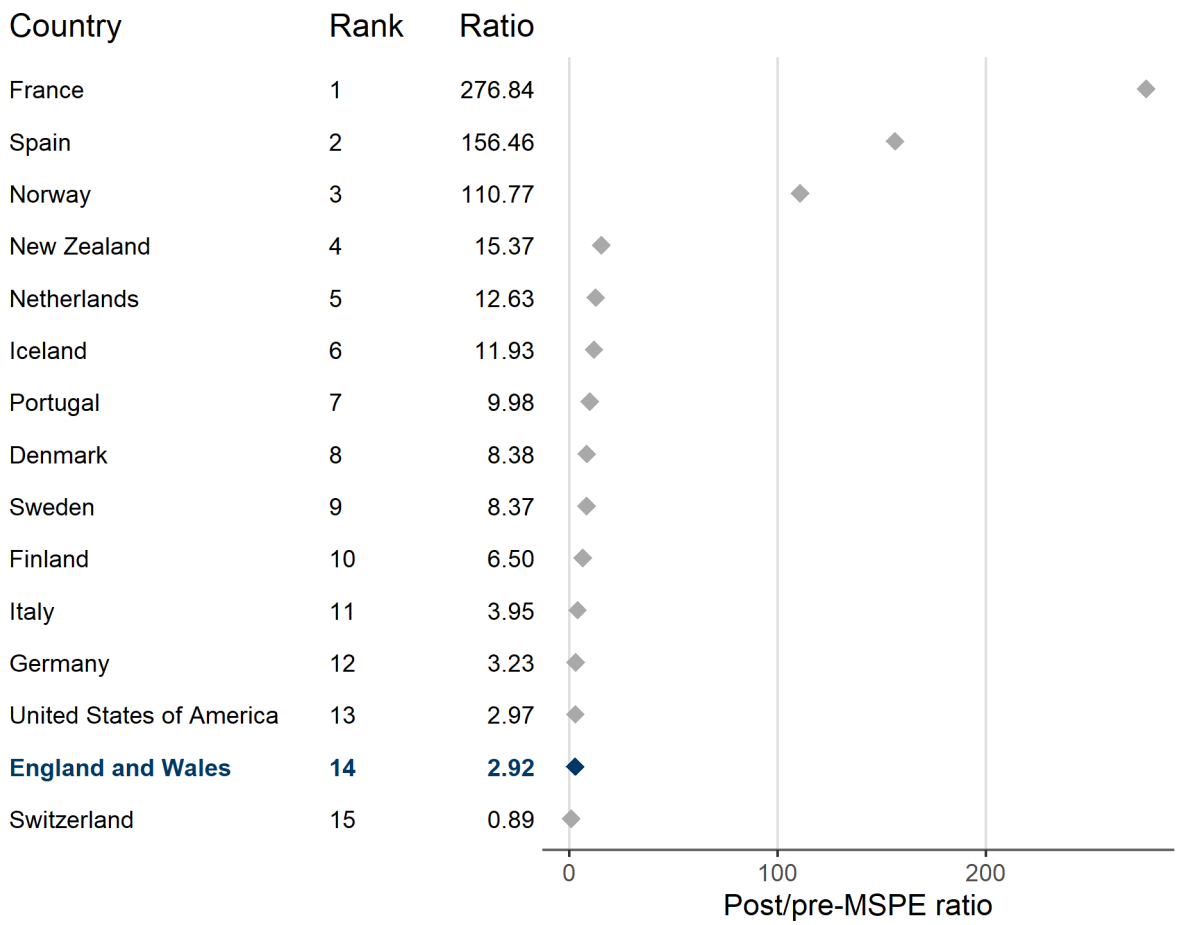

Figure 6-15 – England and Wales' post/pre-intervention MSPE ratio compared with the ratios of control countries for under-18 births; model 3. Scotland was excluded from the list of control countries. Thirteen of fifteen control countries had higher MSPE ratios than England and Wales (p=0.87).

Removing Scotland from the list of control countries for under-20 pregnancies, I refitted a synthetic control to produce model 4. The MSPE of the pre-intervention period increased, indicating a marginally poorer fit (MSPE = 6.03, compared to 4.27 from model 2 above). The control predictions were constructed from Iceland (35.9%), the United States of America (35.3%) and New Zealand (28.8%). Across the post-implementation period, the synthetic control predicted consistently lower rates than seen in England and Wales (Figure 6-16).

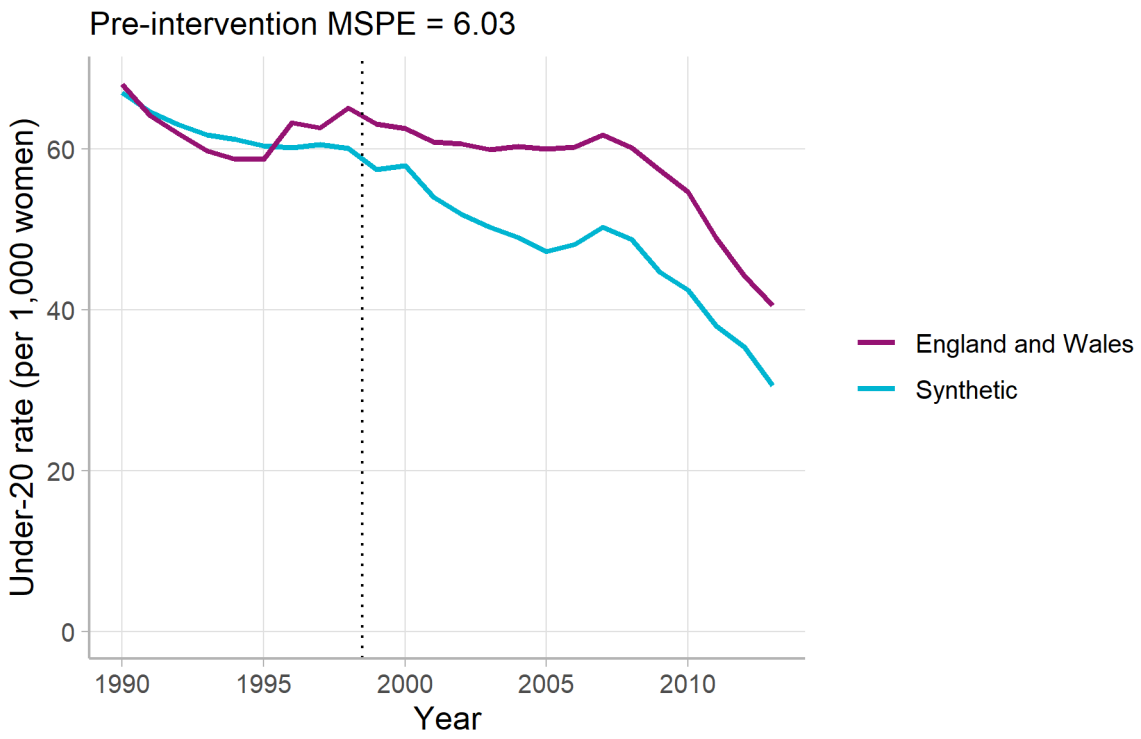

Figure 6-16 – England and Wales under-20 pregnancy rates compared with synthetic control, 1990-2013; model 4 with Scotland removed and only pre-intervention rates as predictors. The dotted line marks the introduction of the Strategy in 1999.

Donor weights	
Country	**Weight**
Iceland	0.359
United States of America	0.353
New Zealand	0.288
Switzerland	0.000
Germany	0.000
Denmark	0.000
Spain	0.000
Finland	0.000
France	0.000
Italy	0.000
Netherlands	0.000
Norway	0.000
Portugal	0.000
Sweden	0.000

Predictor weights	
Predictor	**Weight**
Pregnancy rates 1990	0.309
Pregnancy rates 1991-1995	0.498
Pregnancy rates 1996-1998	0.193

Table 6-4 – Weightings applied to control countries' observations to calculate predicted rates, and weightings of predictors used in model 4 comparing under-20 pregnancy rates across 1990-2013. Clusters of years used to calculate predictor weights were left unchanged from model 2. Each column is independent and sums to 1.

Gaps between observed and predicted England and Wales are plotted in Figure 6-17 alongside gaps calculated for placebo countries. Most countries saw lower rates relative to their predictions than England and Wales across the post-implementation period, similar to model 2. The post/pre-MSPE ratio of England and Wales was larger than in previous models (17.4), but still remained relatively small in comparison with countries with no exposure to the Strategy. Eight of fifteen countries (53%) saw greater ratios, indicating that the probability of seeing a higher MSPE ratio in the absence of the Strategy was at least p=0.53 (Figure 6-18).

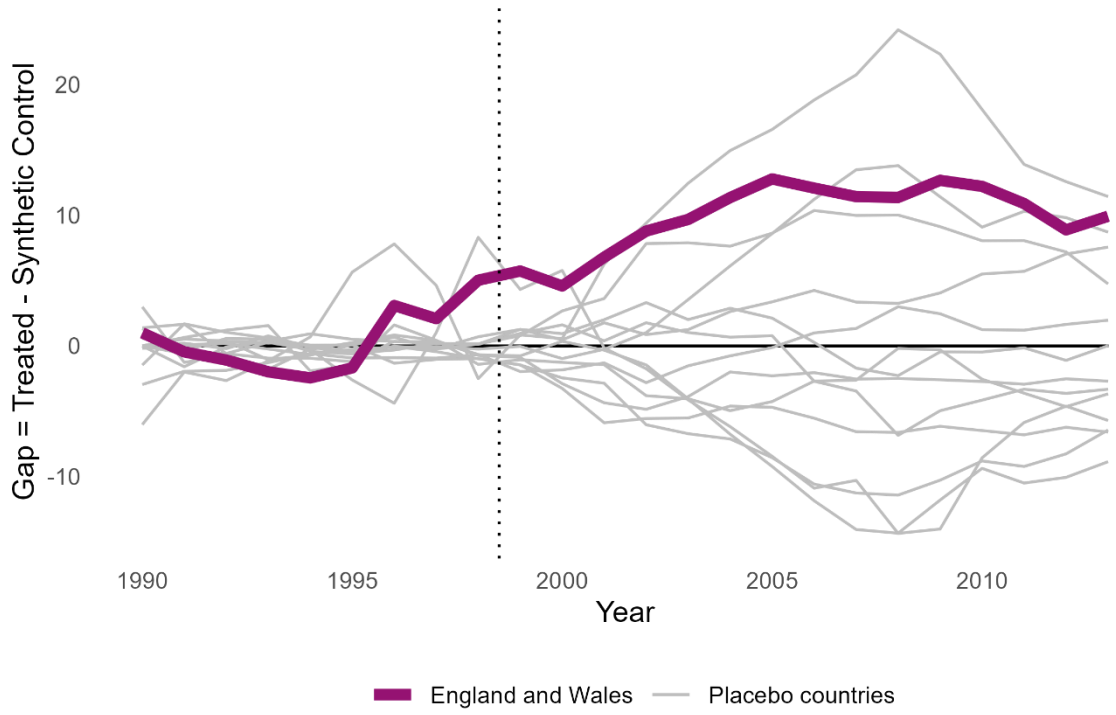

Figure 6-17 – Gaps plot of model 4, showing differences in under-20 pregnancy rates of England and Wales from control, alongside placebo country tests. The Strategy launch is marked in 1999 with a dotted line. Countries with a pre-intervention MSPE greater than 5 times that of England and Wales were excluded: United States of America. Published interactively at **chart-studio.plotly.com/~andrewbaxter439/102**

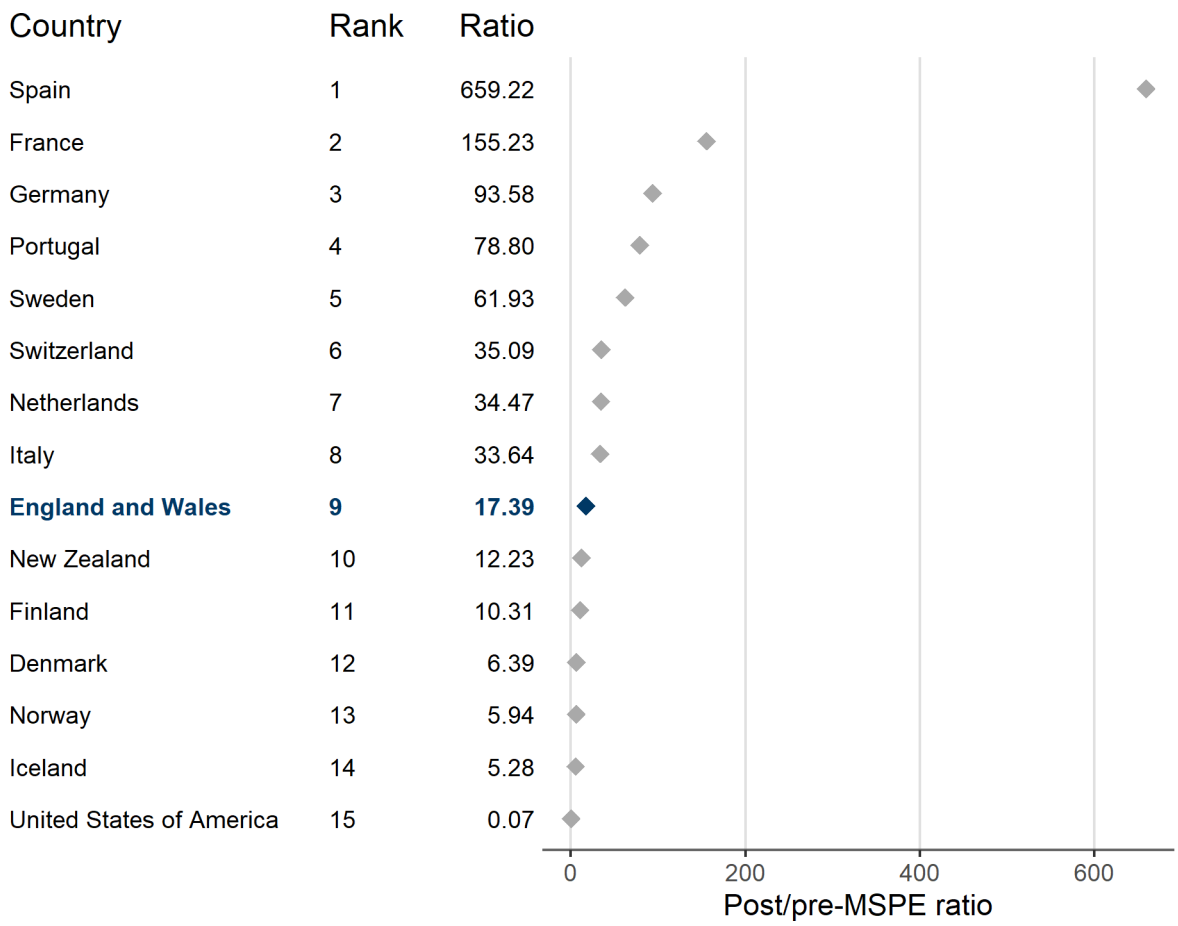

Figure 6-18 – England and Wales' post/pre-intervention MSPE ratio compared with the ratios of control countries for under-20 pregnancies; model 4. Scotland was excluded from the list of control countries. Eight of fifteen control countries had higher MSPE ratios than England and Wales (p=0.53).

6.2.6 Sensitivity analyses adjusting for all predictors

To test whether the addition of further predictors improved model fit and produced consistent results, I repeated analyses of both outcomes with all hypothesised predictors of birth and pregnancy rates. Observations of the predictors GDP per capita, public spending on education as a proportion of GDP, percentage of population in urban dwellings, and percentage of population owning mobile phones were used across the entire observation period, 1990-2013. These were grouped by clusters of years to similarly estimate the best grouping to predict outcomes. In both models, pre-intervention under-18 birth rates and under-20 pregnancy rates were included as predictors. I excluded Scotland from each analysis, as separate data were not consistently available for different countries within the UK. New Zealand was also excluded from these analyses due to missing data. Where UK-wide data were extracted, these were used as proxies of England and Wales alone.

Figure 6-19 shows the outputs of model 5 using all predictors alongside under-18 births. Plotted trends of England and Wales and the synthetic control are shown alongside the plotted gaps for England and Wales compared to all control countries.

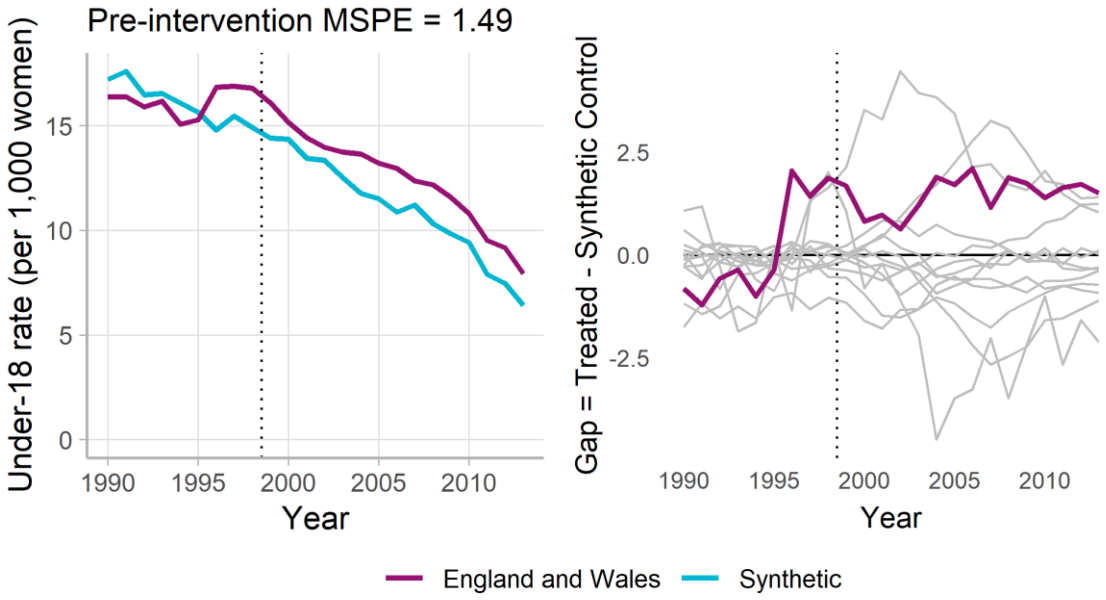

Figure 6-19 – Plotted trends in under-18 birth rates across 1990-2013 and gaps plot for model 5, constructed from 13 donor countries and all predictors. The Strategy launch is marked in 1999 with a dotted line. Countries with a pre-intervention MSPE greater than 5 times that of England and Wales were excluded: United States of America. Gaps plot published interactively at **chart-studio.plotly.com/~andrewbaxter439/106**.

The model fit was poorer than model 3 (MSPE = 1.49, from model 3 MSPE = 0.62). This resulted in an average pre-intervention prediction error of 1.2 births per 1,000 women per year (around a mean of 16.2 births). In the post-implementation period, the synthetic control predicted consistently lower rates in each year than observed in England and Wales. England and Wales' post/pre-MSPE ratio of 1.56 was relatively low, with 11 of 13 (85%) control countries experiencing higher ratios.

Table 6-5 shows the weights given to donor countries and to predictors in construction of the synthetic control. The predicted birth rates were calculated from a weighted mean of Portugal (58.2%), Iceland (24.5%), and the United States of America (17.2%). The predictors GDP per capita and the proportion of the population living in urban environments were not used as predictors, and level of spending on education was given a low weight overall (0.4%). The strongest predictor was percentage of population owning mobile phones in 1996-1997, which was assigned a weight of 52.2% (Table 6-5).

Donor weights		Predictor weights	
Country	**Weight**	**Predictor**	**Weight**
Portugal	0.582	(under 18) Birth rates 1990-1993	0.165
Iceland	0.245	(under 18) Birth rates 1994	0.049
U.S.A.	0.172	(under 18) Birth rates 1995	0.12
Switzerland	0.000	(under 18) Birth rates 1996-1998	0.016
Germany	0.000	(under 20) Pregnancy rates 1990	0.01
Denmark	0.000	(under 20) Pregnancy rates 1991-1995	0.008
Spain	0.000	(under 20) Pregnancy rates 1996-1998	0.031
Finland	0.000	GDP per capita 1990-2003	0
France	0.000	GDP per capita 2004	0
Italy	0.000	GDP per capita 2005-2010	0
Netherlands	0.000	GDP per capita 2011-2013	0
Norway	0.000	Spending on education 1995-1998	0
Sweden	0.000	Spending on education 1999-2007	0.002
		Spending on education 2008-2013	0.002
		Mobile phone ownership 1990-1992	0.05
		Mobile phone ownership 1993-1995	0
		Mobile phone ownership 1996-1997	0.522
		Mobile phone ownership 1998-2013	0.025
		Proportion of urban pop 1990	0
		Proportion of urban pop 1991	0
		Proportion of urban pop 1992	0
		Proportion of urban pop 1993-2013	0

Table 6-5 – Weightings applied to control countries' observations to calculate predicted rates, and weightings of predictors used in model 5 comparing under-18 birth rates across 1990-2013. Birth rate and pregnancy rate pre-intervention predictors from previous models were combined. Further predictors were added as clustered years fitted to maximise pre-intervention MSPE: GDP per capita, public spending on education as a proportion of GDP, percentage of population in urban dwellings, percentage of population owning mobile phones. Each column is independent and sums to 1.

Model 6 was constructed using the same predictors for the outcome of under-20 pregnancies. The outputs are shown in Figure 6-20, plotting the predicted and observed trends and the gaps produced for England and Wales compared with placebo countries. This similarly saw a poorer fit than model 4 using only pregnancy rates as a predictor (MSPE = 11.6, increased from 6.03; average prediction error of 3.41 pregnancies per 1,000 women per year). England and Wales' rates were consistently higher than synthetic control predictions across the post-implementation period and produced large positive gaps across 2001-2013 compared to control countries. The post/pre-MSPE ratio for England and Wales

was relatively high at 18.7, with seven of the 13 control countries (54%) producing higher ratios. This indicates that the relatively large gaps are still consistent with a null effect of the Strategy, where a country with no Strategy exposure still had at least a p=0.54 chance of producing a greater relative deviation from predicted values.

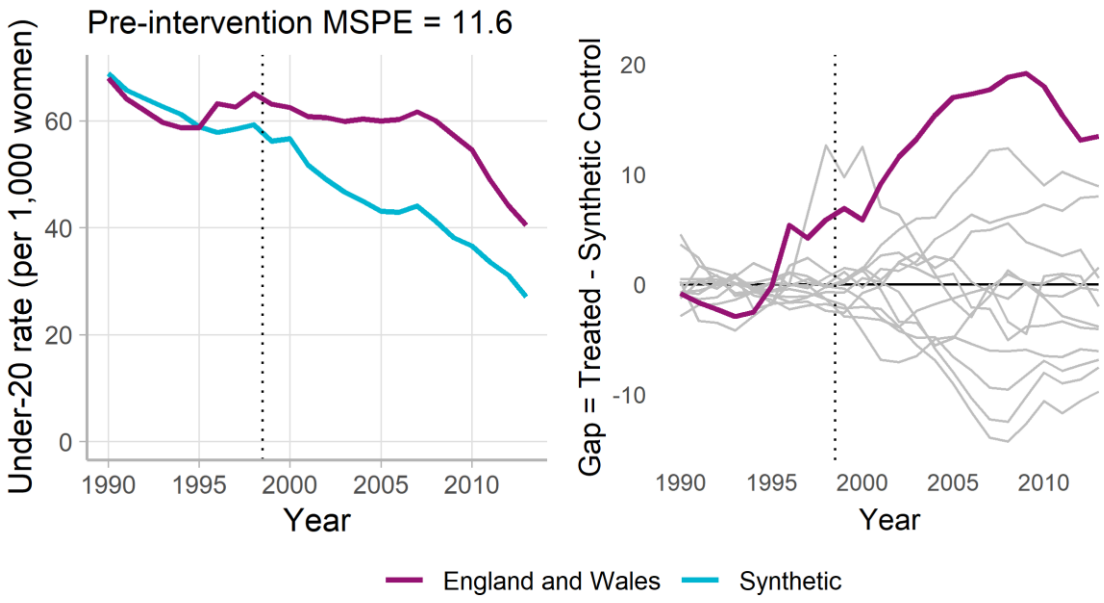

Figure 6-20 – Plotted trends in under-20 pregnancy rates across 1990-2013 and gaps plot for model 6, constructed from 13 donor countries and all predictors. The Strategy launch is maked in 1999 with a dotted line. Countries with a pre-intervention MSPE greater than 5 times that of England and Wales were excluded: United States of America. Gaps plot published interactively at **chart-studio.plotly.com/~andrewbaxter439/100**.

The fitted weights for donor countries and predictors are shown in Table 6-6. Mobile phone ownership and public spending on education were given relatively little weight (1.2% and 0.2% of the total weight respectively). Pre-intervention pregnancy rates were predominantly used as predictors, consistent with previous models using under-20 pregnancy rate data.

Donor weights		Predictor weights	
Country	**Weight**	**Predictor**	**Weight**
U.S.A.	0.424	(under 20) Pregnancy rates 1990	0.202
Iceland	0.420	(under 20) Pregnancy rates 1991-1995	0.179
Portugal	0.156	(under 20) Pregnancy rates 1996-1998	0.345
Switzerland	0.000	(under 18) Birth rates 1990-1993	0.011
Germany	0.000	(under 18) Birth rates 1994	0.016
Denmark	0.000	(under 18) Birth rates 1995	0.017
Spain	0.000	(under 18) Birth rates 1996-1998	0
Finland	0.000	Mobile phone ownership 1990-1992	0.001
France	0.000	Mobile phone ownership 1993-1995	0.002
Italy	0.000	Mobile phone ownership 1996-1997	0.009
Netherlands	0.000	Mobile phone ownership 1998-2013	0.002
Norway	0.000	Spending on education 1995-1998	0.002
Sweden	0.000	Spending on education 1999-2007	0
		Spending on education 2008-2013	0
		GDP per capita 1990-2003	0.028
		GDP per capita 2004	0.024
		GDP per capita 2005-2010	0.006
		GDP per capita 2011-2013	0.011
		Proportion of urban pop 1990	0.038
		Proportion of urban pop 1991	0.036
		Proportion of urban pop 1992	0.036
		Proportion of urban pop 1993-2013	0.034

Table 6-6 – Weightings applied to control countries' observations to calculate predicted rates, and weightings of predictors used in model 6 comparing under-20 pregnancy rates across 1990-2013. Birth rate and pregnancy rate pre-intervention predictors from previous models were combined. Further predictors were added as clustered years fitted to maximise pre-intervention MSPE: GDP per capita, public spending on education as a proportion of GDP, percentage of population in urban dwellings, percentage of population owning mobile phones. Each column is independent and sums to 1.

6.2.7 Robustness testing of models

In robustness testing models, I iteratively removed one country at a time and re-ran synthetic control models for England and Wales for the remaining 14 donor countries. I recorded the gaps between observed and predicted values across the yeas 1990-2013 and the calculated pre-intervention MSPE to compare model fit. The resulting graphs for under-18 birth rates and under-20 pregnancy rates are shown in Figure 6-21 and Figure 6-22.

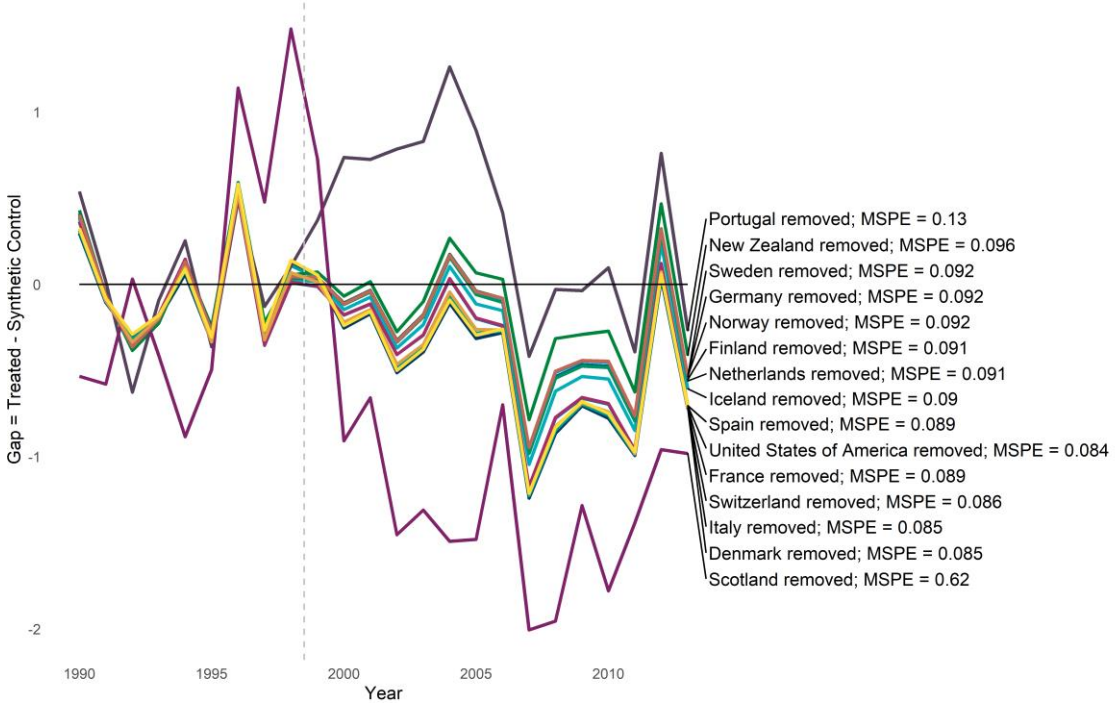

Figure 6-21 – Gaps plot for England and Wales under-18 birth rates compared to a series of controls with one donor country removed. The Strategy launch is maked in 1999 with a dotted line. Countries removed from the donor pool and the resulting pre-intervention MSPE are noted on the right hand side. All models used pre-intervention under-18 birth rates as predictors and were fitted with the remaining fourteen donor countries.

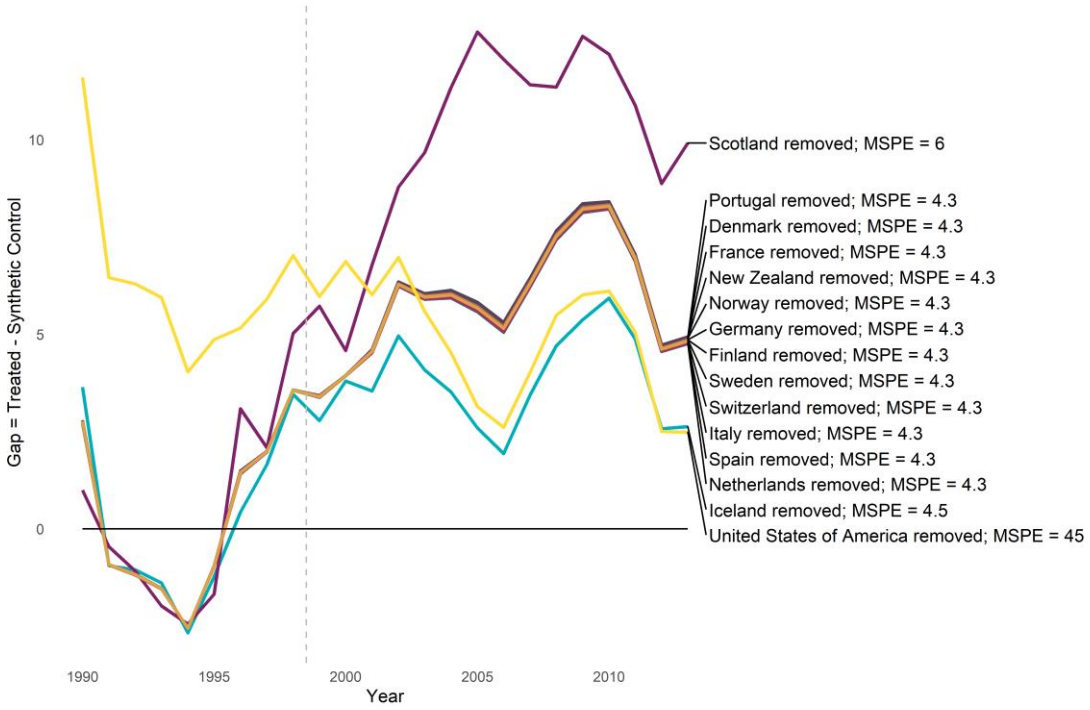

Figure 6-22 – Gaps plot for England and Wales under-20 pregnancy rates compared to a series of controls with one donor country removed. The Strategy launch is maked in 1999 with a dotted line. Countries removed from the donor pool and the resulting pre-intevention MSPE are noted on the right hand side. All models used pre-intervention under-20 pregnancy rates as predictors and were fitted with the remaining fourteen donor countries.

In both tests, removal of Scotland produced a large jump in pre-intervention MSPE (shown by the purple line; Figure 6-21 and Figure 6-22), replicating the models 3 and 4 above. In

most other cases, the pre-intervention MSPE and the post-implementation pathways remained stable, as removed countries had previously been assigned little weight when included in models 1 and 2. Removal of the United States of America from the calculation of under-20 pregnancy rates resulted in a very large increase of MSPE (to MSPE = 45, from MSPE = 4.3 across most other models, consistent with model 2). As can be seen from Figure 6-2, removal of the U.S.A. data leaves England and Wales as showing the highest pregnancy rates across the pre-intervention period, and therefore no longer within a convex hull. The remaining valid models appear to show a range of trends consistent with the previously tested models 1-6. Scotland was relied upon to produce well-fitted pre-intervention periods for both under-18 birth rates and under-20 pregnancy rates. Removal of Scotland did not subsequently show clear indications of lowered rates in England and Wales when compared with synthetic control.

6.3 Discussion

6.3.1 Summary of results

Using publicly available data from a range of high-income European and English-speaking countries, I was able to construct synthetic controls to make robust predictions of England and Wales' birth and pregnancy rates in the absence of the Strategy. Across all models (summarised in Table 6-7), I saw consistent predictions of downward trends across the Strategy period. In most cases, the predicted birth or pregnancy rates in the absence of the Strategy were close to or lower than those observed in England and Wales. In one model, excluding Scotland data, England and Wales may have seen pregnancy rates 1-2 pregnancies per 1,000 women lower than expected across each year of the Strategy period. However, this trend was predicted from a poorer-fit model and differences from control were within the range of effects seen in control countries in the absence of the Strategy.

England and Wales' decreasing teenage birth and pregnancy rates appear to be consistent with trends in similar countries and may be at least partially attributable to common events, external to the Strategy. In particular, the English-speaking countries of the UK, New Zealand and the United States of America showed the highest rates in the pre-intervention period and each model included one or more of Scotland, New Zealand or the U.S.A. in calculating predicted trends. Changes in other exposures predominantly affecting English-speaking countries may account for some of the change seen in the Strategy period.

Model	Outcome	Predictors	Number of donors	Pre-1999 MSPE	Post/pre-MSPE ratio and rank
Primary analyses					
Model 1 – all countries	Under-18 births	Pre-Strategy births	15	0.092	1.86 14 of 16
Model 2 – all countries	Under-20 pregnancies	Pre-Strategy pregnancies	15	4.27	9.1 11 of 16
Sensitivity analyses					
Model 3 – Scotland removed	Under-18 births	Pre-Strategy births	14	0.62	2.92 14 of 15
Model 4 – Scotland removed	Under-20 pregnancies	Pre-Strategy pregnancies	14	6.03	17.4 9 of 15
Model 5 – Additional predictors	Under-18 births	All	13	1.49	1.56 12 of 14
Model 6 – Additional predictors	Under-20 pregnancies	All	13	11.6	18.7 8 of 14

Table 6-7 – Summary figures of primary comparative and sensitivity models. England and Wales' post/pre-MSPE ratio was used to rank the relative deviance of the post-implementation period controlled by goodness of pre-intervention fit, comparing across all countries. A higher rank represents a greater deviation in the post-implementation period than produced by noise in unexposed comparator countries, consistent with a tested exposure's large positive or negative effects.

6.3.2 Strengths and limitations

As previously highlighted, reservations concerning the use of Scotland and Wales as comparators to test Strategy effects on England arise from both the potential for contamination of Strategy effects or co-occurring interventions during the Strategy period (Wilkinson *et al.*, 2006; Wellings, Palmer, Geary, *et al.*, 2016). Both initial models including Scotland as a comparator assigned Scotland a majority of the weighting (67.2% in model 1, 63.9% in model 2). Removing Scotland in both cases gave a slightly poorer pre-intervention fit, but the remaining donor countries still predicted a downward trend, similar to observed England and Wales rates. The suggestion of similarly sized contamination effects across the remaining weighted donors of New Zealand, the United States of America, Iceland and Portugal appears extremely unlikely. Hypothesised simultaneous interventions in all weighted comparators could potentially have masked Strategy effectiveness. The similarity of effect sizes – possibly predicting greater effects in controls – and similarity of shapes of trends more strongly indicate a common cause. This may represent a widespread policy response to teenage pregnancy in countries with high

rates producing diverse actions (of which the Strategy was one), although this would in turn contradict the uniqueness of the Strategy and query the necessity of the extra allocated spending. The remaining explanation, that the Strategy had little or no effect on under-18 birth and under-20 pregnancy rates, seems best suited to these observations.

The poorness of model fit after removal of Scotland may be due to a combination of factors. Though England and Wales was still within a convex hull in both cases, there were relatively few countries with higher rates or rates as close to England and Wales as Scotland. This limitation in the data restricted the synthetic control algorithm's ability to match a pre-intervention synthetic trend as accurately as prior models. Further data from other English-speaking countries (e.g., Ireland, Canada, Australia) may be able to address this, but were not available for this analysis. Additionally, in the immediate pre-intervention period, across the years 1996-1998 there appears a raised rate in England and Wales, which Scotland is able to account for in the under-18 births model (and to partially account for in the under-20 pregnancy model). This change in trend is not repeated in non-UK countries and is consistent with the hypothesised effects of the 1996 'pill scare' temporarily increasing pregnancies and births (Furedi, 1999). The poorer fit after removal of Scotland again indicates the potential limitations of the data and the limitations of the method in accounting for this idiosyncratic UK event.

After removing several countries from the donor pool to account for known historical events, in visualisation of rates I further discovered a large spike in New Zealand under-18 birth rates which wasn't seen in other countries (Figure 6-3 – not replicated in under-20 pregnancy rates). This violates the assumption that no external shocks occurred in control units (Bouttell *et al.*, 2018). The primary model, assigning New Zealand a very small weight, and the analysis of under-18 birth rates with all predictors, excluding New Zealand, both showed consistent trends and may be more reliable predictions of unexposed England and Wales.

The Strategy's target was to reduce under-18 pregnancies (Social Exclusion Unit, 1999). Previous evaluations have used this metric as a primary marker of Strategy effects (see also Chapter 5; Wilkinson *et al.*, 2006; Wellings, Palmer, Geary, *et al.*, 2016). When scoping available data from target countries, I was able to generate estimates of births to women aged in both under-18 and under-20 age groups but was only able to obtain abortion estimates for under-20s. My two outcome measures, under-18 birth rates and under-20 pregnancy rates do not precisely match the target outcome. Each measure may contain

errors and biases. Under-18 births would not be able to account for changing proportions of conceptions ending in abortion. Under-20 pregnancies may include conceptions to older women which the Strategy's purpose does not count as problematic (cf. the report's highlighting of unplanned pregnancies, pregnancies to unmarried mothers and rates in deprived areas as prompts for action; Social Exclusion Unit, 1999).

The methods for calculating rates additionally have important limitations in their ability to measure total pregnancies. I was unable to correct counts of births for age of mother at conception (rather than birth), conceptions ending in still births, or pregnancies ending in multiple births (see Office for National Statistics, 2017; Information Services Division Scotland, 2019). Methods for collecting counts of abortions varied across countries and may not all be reliable records of all abortions (WHO Regional Office for Europe, 2019). Miscarriages were not able to be counted but were assumed to be stably related to birth and abortion rates and unlikely to introduce systemic bias (Sedgh *et al.*, 2015).

Using both birth and pregnancy rates together, however, I expected that an effective Strategy could produce detectable changes to either outcome in either age group. I was able to use two measures and, in testing for consistency, to assess their reliability. The relatively similar downward trends seen in all synthetic controls across both outcomes indicates that these measures are sensitive to detect common effects on both pregnancy and birth rates for the whole combined age group. The consistency of the closeness of observed results to the predicted null-effect trend indicates that the Strategy saw similarly small or null effects on its target of under-18 pregnancies.

If, however, the Strategy had produced a small effect, the synthetic control method may not have been able to detect the changes with sufficient confidence. Across all models, differences from predicted rates varied widely, indicating that the synthetic control predictions may not be precise estimates of what would have been seen in England and Wales with no Strategy exposure. However, the closeness of model fit in model 1 followed by relatively small gaps in the comparison period suggests that these observations are the result of a null-effect Strategy. In four of the five subsequent models, the observed rates remained consistently higher across the Strategy period than predicted. Whilst they did not provide confident evidence of the negative effect of the Strategy, the tendency of models to predict unexposed rates as lower than observed suggest that strong, positive Strategy effects are inconsistent with these observations.

6.4 Conclusion

Using the synthetic control method, I was able to construct matching controls to estimate England and Wales' teenage birth and pregnancy rates from 1999 to 2013 in the absence of the Strategy. I tested my models for robustness and sensitivity. Although model fit deteriorated with removal of Scotland data, all models were similarly able to replicate a downward trend across the Strategy period in the absence of the Strategy. Better-fit models in the pre-intervention period showed smaller gaps across the post-implementation period than poorer-fit models. Other models predicted faster declines after 1999 than seen in Strategy-exposed England. All models were consistent with the null effect of the Strategy. These models were able to address concerns of contamination and confounding previously raised over questions of comparison between England and unexposed countries.

Chapter 7 Discussion

7.1 Overview of chapter

In this concluding chapter I summarise the results of my empirical chapters and present my overall findings on falling teenage pregnancy rates and the contribution of the Teenage Pregnancy Strategy. I compare these findings with previous research highlighted in Chapters 1 and 2 and discuss how my analyses contribute to our understanding of the topic of teenage pregnancy. I further discuss the strength of my methods and the reliability of my findings. Finally, I present the implications of my findings for further research and policy development.

7.2 Summarising and synthesising results from empirical chapters

7.2.1 Summary of results

In this thesis I aimed to identify the plausible hypothesised causes of falling teenage pregnancy rates in the UK and to examine the Strategy's contribution to these. In pursuing this aim I combined two approaches. I first systematically reviewed the literature for evidence testing the contribution of hypothesised causes to the observations of declining teenage pregnancy rates in the UK. I began with an initial causal tree highlighting the prominent hypothesised causes discussed in recent teenage pregnancy prevention literature. I assessed the available evidence for these causes and detected and assessed further hypotheses using a causal framework to produce an updated tree. I thus summarised plausible hypothesised causes of falling teenage pregnancy rates. I then applied robust policy evaluation methods to the Strategy to better estimate its effects.

My empirical chapters (Chapters 4, 5 and 6) addressed the six research questions:

RQ1: What changes in interventions, culture, policy and environment are hypothesised to cause reductions in teenage pregnancy in the UK and similar countries?

RQ2: For which causes is there evidence of their effects in reducing pregnancy, and their exposure across populations?

RQ3: Which causes have been robustly examined but have failed to produce evidence of their effects across a population, indicating the absence of an effect?

RQ4: What are the core elements to include in a causal tree which may have contributed to reductions in teenage pregnancy rates in the UK?

RQ5: Did implementation of the Teenage Pregnancy Strategy cause a change in level or trend in yearly teenage pregnancy rates in England, compared with other UK countries?

RQ6: Did implementation of the Teenage Pregnancy Strategy reduce yearly teenage pregnancy rates in England, compared with a wider set of control countries?

In my systematic review (Chapter 4), I reviewed and assessed the evidence measuring the effects of previously identified causal pathways and discovered several additional hypothesised causes (addressing RQ1). I used my initial causal tree, constructed from prominent hypotheses presented in the literature, to create a broadly scoped search strategy. This strategy was designed to retrieve evidence for these branches and detect other hypothesised causes. This evidence allowed me to assess the strength of evidence for each causal pathway as a contributor to the observed rates in the UK. I was able to use the results of this analysis to both remove branches which had been comprehensively tested but had produced little observable effect, and to add branches which had supporting evidence (RQs 2 and 3). In the updated causal tree, I summarised a wide range of possible contributors to the falling rates which either had supporting evidence or were not previously assessed (RQ4). This updated tree incorporated both interventions to address sexual risk behaviours and reduce teenage pregnancy rates, and changes in social, cultural and environmental determinants that may have an impact on teenage pregnancy rates. This causal tree provides a framework for further testing of the causal effects of both interventions and wider changes in culture and social environment.

My review suggested a range of effective interventions for reducing teenage pregnancy. The Strategy incorporated several elements of the tree, with developers seeking to take a multifaceted approach and combine interventions with evidence of effectiveness (Social Exclusion Unit, 1999). These included: improved school-based sexual health education, engaging with parents, and improving access to and use of sexual health clinics (Figure 7-1). My review found no evidence of a population-wide exposure, limiting these findings to ‘weak evidence’ of a causal effect. However, the national scope of the Strategy is in

itself confirmation of a whole population exposure, strengthening the evidence when combined with the systematic review's findings. The Strategy employed several other approaches which my review suggested were ineffective: increasing contraception access and public information and media campaigns. Previous evaluations of the Strategy had highlighted these combined elements as contributing to its success (Wilkinson *et al.*, 2006; Hadley, Ingham and Chandra-Mouli, 2016, fig. 1).

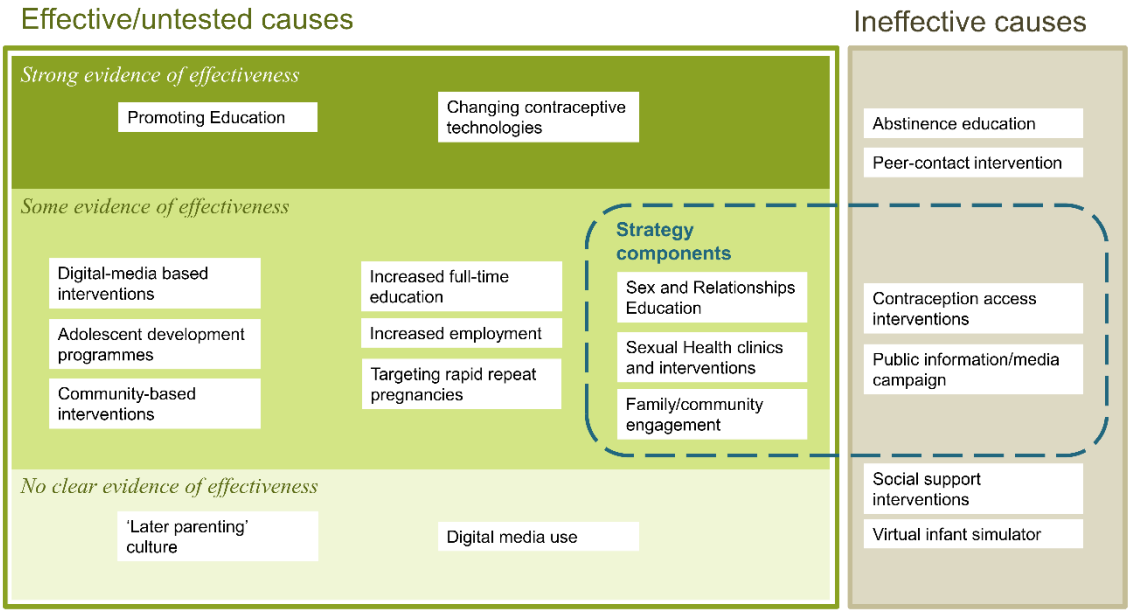

Figure 7-1 – The range of hypothesised causes, including both interventions and changes in social determinants, considered in the systematic review, arranged by strength of evidence and with Strategy components highlighted. Components hypothesised to have an effect but not tested in the reviewed literature (bottom box of left-hand column) are distinguished from hypotheses which were tested but for which no effect could be found (right-hand column).

My evaluation of the Strategy (Chapters 5 and 6) demonstrated an approach to testing a hypothesised cause, whilst conducting more robust outcome evaluation of the intervention. I took a natural experimental approach to assessing a well-defined intervention which was applied to a whole population.

In my interrupted time series models (Chapter 5), I compared England to Scotland and Wales to determine the Strategy's effects, whilst accounting for background trends and events common to UK countries. I hypothesised that the full exposure of an effective Strategy in England would have caused a greater reduction in under-18 pregnancy rates from 1999 onwards. I found no detectable difference in trends in pregnancy rates when comparing exposed England with unexposed Scotland and Wales (RQ5). This inability to detect a difference was consistent across all sensitivity analyses. All confidence intervals were narrow, allowing me to confidently rule out a large effect in England (Table 7-1).

Control	Outcome	Effect direction and certainty		
Interrupted time series models		*Level/trend*		
Scotland	Under-18 pregnancies	0/0		
Wales	Under-18 pregnancies	∧ / ∨		
Synthetic control models		*Difference*		
European and English-speaking countries	Under-18 births	0		
European and English-speaking countries	Under-20 pregnancies	∧		
European and English-speaking countries (excl. Scotland)	Under-18 births	∨		
European and English-speaking countries (excl. Scotland)	Under-20 pregnancies	⋀		
Key: ⋀ or ⋁ – large increase or decrease in pregnancy rates, ∧ or ∨ – small increase or decrease; 0 – very small difference	***Significance:***	*Sig –*	*Non-sig*	*Sig +*

Table 7-1 – Summary of model fit, effect direction and certainty across all comparative models Only primary models from chapters 5 and 6 are included. Sensitivity analyses across all comparisons showed similar results, with no evidence of a reduction in pregnancy rates in exposed England compared to controls.

In a separate test of the Strategy's effects (Chapter 6), I used synthetic control methods to compare England and Wales to a wider set of comparison countries. Researchers had previously highlighted that certain Strategy elements, such as the media campaign, may have had an effect in Scotland and Wales, creating contamination in these countries as controls (Wellings, Palmer, Wilkinson, *et al.*, 2016). In my synthetic control comparison, the wider range of included countries meant that I was able to correct for potential confounding. I used the two measures of under-18 birth rates and under-20 pregnancy rates. I hypothesised that although these measures were not primary Strategy focusses, they both would likely have seen similarly sized reductions in the presence of an effective Strategy. In both models, I found that controls constructed from unexposed countries showed trends very similar to those of England and Wales, with or without Scotland included in the control (RQ6). The changes in countries similar to England may have been produced by a common effect leading to similarly falling rates across most high-income countries, independent of the Strategy. This was consistent across all robustness and sensitivity analyses.

Whilst I identified a range of intervention approaches not used in the Strategy which could have contributed, in this analysis I did not assess the exposure of the UK population to these. Interventions addressing more distal social determinants, and shaped as adolescent development interventions, may be able to achieve pregnancy risk reduction through a primordial prevention approach (Kirby, 2002a). Whilst contraception distribution to prevent rapid repeat pregnancies may be an element of standard post-partum or post-abortion care (Healthcare Improvement Scotland, 2021), my review did not identify evidence of how provision of this might have changed across the observed period (1990-2016). Additionally, whilst digital media-based interventions may be a promising novel approach with the potential for increased use as mobile technologies developed in the last 20 years, my review did not uncover evidence of their roll out in the UK. Further investigation could identify widely implemented interventions and test their contributions to falling teenage pregnancy rates.

7.2.2 Setting in the literature

My findings concerning the Strategy are in contradiction to previous estimates of a large effect of the Strategy on pregnancy rates (Toynbee, 2013; Hadley, Ingham and Chandra-Mouli, 2016; Ma, 2016; Wellings, Palmer, Geary, *et al.*, 2016). A comparison of England's change in under-18 birth rates to the change in European average birth rates across the period measured a 51% fall in England from 2004 to 2014 compared with a 22% reduction in Europe (Wellings, Palmer, Geary, *et al.*, 2016). The authors suggested that this control represented a relatively small change due to secular trends (Wellings, Palmer, Geary, *et al.*, 2016). These reductions were from a baseline of 13.6 births per 1,000 women aged 15-17 in England and 7.7 births per 1,000 women in other European countries. The use of the European average as a comparator relies on the assumption that common events outside the Strategy would produce a similarly sized relative reduction of 22% from both the high English and low European baseline rates. My comparative models selected comparable countries and controlled for differences in baseline rates and trends. Both my comparisons with UK countries and my synthetic controls constructed from a larger pool of countries produced controls with pre-Strategy rates closer to England than the European average used above. The use of controls that were as similar as possible to England avoids the extrapolating of changes to low European pregnancy rates as representing a similarly sized relative background effect on England's high pregnancy rates (Abadie, Diamond and Hainmueller, 2011; Lopez Bernal, Cummins and Gasparrini, 2018). All my comparative models suggest that a similar fall in rates would have occurred in the absence of the

Strategy. This finding indicates that the observed effect previously attributed to the success of the Strategy was produced by other more widespread changes causing similarly falling rates in demonstrably comparable countries (Table 7-1). Thus, my analyses are more likely to represent the true effect size than previous evaluations. These findings suggest that the Strategy was a relatively ineffective intervention for reducing teenage pregnancy rates.

In my review I detected a range of effective interventions for reducing teenage pregnancy as potential explanations of the falling rates. Several of these were incorporated into the Strategy – improved school based sexual health education, improved sexual health clinic access, and engaging with parents and family (Figure 7-1). Strategy development had sought to use prior evidence to inform effective approaches (Social Exclusion Unit, 1999). Previous evaluations of the Strategy had highlighted the evidence-based approach and selection of effective interventions as potential explanators of its success (Hadley, Ingham and Chandra-Mouli, 2016). The Strategy as a combination of interventions with prior evidence of effectiveness does not appear to have achieved the effects that the evidence in my review suggested were possible. The similarity of pattern across the controls in my analyses suggests that, rather than England-implemented interventions incorporated in the Strategy, the causes are likely to have been shared across several countries.

7.2.3 Reviewing the COM-B model

Using the pregnancy risk COM-B model, developed in section 2.4.2, I was able to map how various sexual health interventions and environment changes were hypothesised to affect pregnancy rates and highlight the potential for the differential successes of competing causes. The testing of the Strategy as one prominent approach, combined with the highlighting in the literature of changing contraceptive technologies and increased education as showing strong evidence, may contribute more detail to the model.

In light of the seeming ineffectiveness of the Strategy's approach, it is unclear whether the problems in the Strategy development previously highlighted in the literature (section 2.4.1) contributed to this unexpected outcome. The focus of the Strategy on addressing 'technical/educational' factors (Arai, 2009), e.g., through school-based sex education which in the literature had shown potential effectiveness, could be presented as directing most resources to improving the psychological capability of young people to avoid pregnancy (section 2.4.2; Figure 2-4). My systematic review suggested some evidence that educational approaches, addressing the psychological capability and reflective motivation

of teenagers to engage in behaviour change (Michie, van Stralen and West, 2011), may in some contexts be effective at lowering pregnancy rates. However, if Capability was already a 'saturated' component of the COM-B model – if teenagers were already equipped with sufficient skills and knowledge for preventing pregnancy – then such interventions may be expected to produce little change through this mechanism (Figure 7-2). In the reviewed literature, the comparative studies indicating the effectiveness of improved school-based sex education were predominantly studies conducted between 1990 and 2000 in the USA and Canada (Marseille *et al.*, 2018). It is possible that the control groups in these cases were not comparable with the pre-Strategy conditions in England; pre-Strategy sex education may already have been more effective than the treatment as usual conditions in the studies. Observations of similar approaches to improve sex education in Scotland found little effects of replacement of existing programmes (Henderson *et al.*, 2007; Elliott *et al.*, 2013), but indications that reductions in education reduced their beneficial effects (Nixon, Wight and Henderson, no date).

Other Strategy components may have targeted different pathways, but not implemented effective approaches. These pathways may have seen greater influence from other changes over the period. My literature review indicated that contraception distribution interventions, addressing a 'technical' factor and providing the physical opportunity to avoid pregnancy (Arai, 2009), were unlikely to be effective. Conversely, the potential effects of improving contraceptive technology, also a 'technical' factor but not addressed by the Strategy, may have been more effective at improving the Opportunity pathway to avoid pregnancy (Figure 2-4). In the literature I found strong evidence for the potential effects of changing contraceptive technologies on reducing teenage pregnancy rates (considered separately from interventions to improve access to contraception). This echoed the later Strategy evaluation which highlighted two UK-wide changes in the availability of different contraceptive methods: over the counter provision of emergency contraceptives and recommendations of long acting reversible contraceptives (Wellings, Palmer, Geary, *et al.*, 2016, fig. 1). These were independent of the Strategy and implemented in tandem. The UK-wide implementation of these changes, alongside global trends in improving contraceptive technologies, may account for some of the similarities of predicted rates between England and controls. However, previous investigation of the effects of changing contraceptive technologies on population pregnancy rates suggested that the overall effect size is likely to be small (Girma and Paton, 2015).

Other interventions or non-designed environment changes may have additionally addressed 'structural' and 'social/cultural' factors outside of the Strategy. Removal of these inhibiting factors may in turn produce the Motivation driving the behaviour change to avoid pregnancy. In particular, increasing education and employment amongst young women, driven either by cultural changes or interventions, was found to have positive effects on lowering pregnancy rates. By empowering young women with more perceived control over their futures, greater access to education, and improved capabilities to take on desired careers, these changes may have produced greater reflective motivation to avoid the risk of pregnancy (Figure 7-2). Additionally, increased education of both sexes may reduce opportunities for pregnancy risk behaviour as both partners have less physical opportunity for sexual intercourse. It may be expected that global trends in increasing time in education and access to further education could have produced the effects observed across comparison countries. This finding is consistent with previous research (Girma and Paton, 2015; Wellings, Palmer, Geary, *et al.*, 2016).

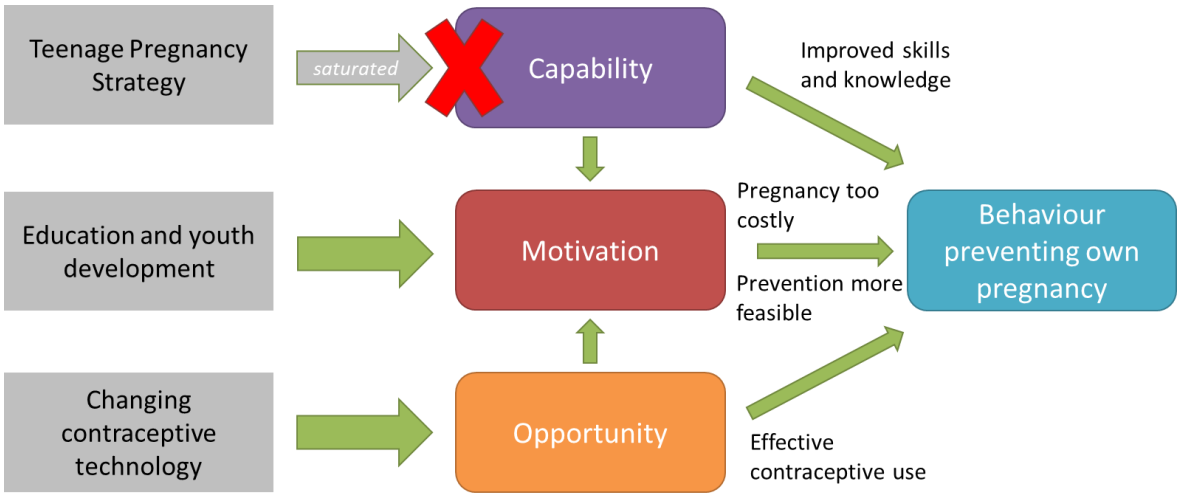

Figure 7-2 – Hypothetical effects of a sample of identified hypothesised causes on the COM-B model of behaviours leading to pregnancy. One tested and ineffective causal branch, the Strategy, is depicted as not affecting the targeted Capability element due a hypothesised saturation effect. Two elements with strong evidence – education and development policies and changing contraceptive technologies – are depicted as having a positive effect on other components.

My analyses give a theoretical model for further understanding reducing pregnancy rates. My systematic review highlighted a range of approaches and potential causes with varying levels of evidence in support. My comparative analyses are compatible with the hypotheses that implementation of one or more of these interventions or changing exposures across several countries globally may have contributed to falling rates across the UK. My analyses additionally cast doubt on the conclusion that the Strategy itself was a significant contributor to the observed effects (Wellings, Palmer, Geary, *et al.*, 2016). I hypothesise from these findings that teenagers at risk of becoming pregnant in the UK in 1998 had a

greater need of motivation and opportunities to avoid pregnancy, which changes over the following period provided. However, my methods were not able to test these assumptions. Future efforts to further address teenage pregnancy or other similarly behaviour-mediated health concerns could make similar use of the COM-B model to identify target components with more scope for change.

7.2.4 Strengths and limitations

I combined three approaches – a systematic review with a causal framework, an interrupted time series analysis and a synthetic control analysis – to provide the clearest and most robust evidence testing the effectiveness of interventions to reduce teenage pregnancy. I synthesised the findings of these three methods to 'triangulate' results and strengthen the inferences of Strategy effects on pregnancy rates (Munafò and Davey Smith, 2018). The practice of using multiple methods in a study can strengthen causal inference (Craig *et al.*, 2017).

I first developed a systematic review approach which was able to identify and assess the evidence for a range of hypothesised causes of an observed effect. The approach I used in my review was designed to detect plausible hypothesised causes; this differs from previous reviews, which primarily aimed to identify effective and reproducible approaches for further action. Whilst synthesising the evidence for the effectiveness of these interventions, I was able to add to this a sensitivity to detect evidence of hypothesised non-intervention causes. My method began with an initial causal tree constructed from the literature to inform systematic literature searching. I then used this evidence to update the initial tree and produce an updated causal tree. This output tree summarises the evidence of hypothesised causes of reducing teenage pregnancy rates presented in the literature. My tree highlighted three practices incorporated into the Strategy with evidence of potential effectiveness (school-based sex education, contraceptive distribution, clinic access and a national media campaign). It additionally highlighted potentially more effective causes outside of the Strategy which may have produced simultaneous changes. This tree can be used to inform further investigations of falling rates by laying out the range of possible confounders and scoping a number of branches potentially contributing to the overall observed effect.

By taking a 'review of reviews' approach, my review may have missed published evidence for other hypothesised causes. The reviews themselves predominantly focussed on

evaluating interventions in order to inform future action, and so changes in environmental and cultural social determinants may have been of less interest to reviewers. Within the wider published literature addressing hypothesised links between various non-designed exposures and pregnancy risk, the evidence is likely to be of mixed quality for assessing causation.

In testing the contribution of the Strategy, I selected methods primarily aimed at assessing the effects of the Strategy against a comparator. Previous evaluations have highlighted the challenges of identifying valid comparators (Wilkinson *et al.*, 2006; Wellings, Palmer, Geary, *et al.*, 2016). Each method used different approaches to overcome these challenges. In applying several methods, datasets and outcome measurements, I was able to use these to check for consistency of effect direction and size. A potential limitation of my triangulation approach is the use of Scotland data in both analyses, as violations of the set of assumptions that Scotland represents a no-effect England would equally undermine both sets of findings. I attempted to correct for this possibility in my synthetic control sensitivity analyses. As all comparison models produced consistent findings, this added strength to the reliability of their estimates. Conversely, inconsistent results would have indicated the potential for error and bias to obscure true effect sizes in one or more approach. An additional element which I could have included within the analysis protocol would have been a pre-specified estimation of the directions of bias across each method. This transparency could have pre-empted likely inconsistencies in findings and given a clearer platform to assess the impacts of biases on effect direction and size.

My outcome evaluation used robust natural experimental methods for testing the hypothesised effects of the Strategy's combination of approaches on England's teenage pregnancy rates. In all analyses I demonstrated the suitability of the controls and the comparative reliability of my effect estimates as a route to giving more accurate measures of change across the Strategy period than uncontrolled observations. The methods I used are able to produce robust causal inferences about the effect of an exposure when a randomised controlled trial is not a feasible option. In the area of post-implementation policy evaluation, as demonstrated in my interrupted time series and synthetic control chapters, these methods were able to give robust comparative estimates of the real-world effectiveness of an implemented intervention. This allowed an outcome evaluation of a complex intervention where a randomised controlled trial was not previously conducted (Craig *et al.*, 2019). However, a key limitation of these methods is the relative lack of control in selecting comparison conditions. In my analyses I was unable to assign control

countries a true ‘treatment as usual’ which would represent England’s sexual health provision in the absence of the Strategy. The effect estimates produced here are estimates of what change the Strategy produced over and against the standard actions taken across comparator countries. My results do not identify the effect of specific components of the Strategy. The results of the two sets of analyses do however cast doubt on whether the Strategy, as a well-funded policy drive to improve existing interventions and supplement them with a national campaign, produced additional effects over the other changes to environment and interventions that were experienced in common with comparator countries.

In my first approach I used Scotland and Wales as similar comparators to correct for UK-wide background trends. These populations were likely to have undergone many of the same non-Strategy changes in policy and social environment. These data were collected consistently across three countries directly representing the target outcome and presented a clear predicted counterfactual. They did, however, rely on assumptions of negligible contamination effect of the Strategy and no confounding from simultaneous interventions, and correctors for other events (principally the 1996 ‘pill scare’). I selected interrupted time series methods as an approach to control for pre-Strategy trends in pregnancy rates which may produce differences across the ten-year Strategy period even in the absence of the Strategy. When including controls, I further used these methods to adjust for differences in pre-Strategy pregnancy rates between countries to ensure these were comparable. The crucial element which I was able to address using these models was the possible effects of UK-wide events, independent of the Strategy, which affected pregnancy rates.

Scotland and Wales have been noted to have developed initiatives independent from the Strategy across the years 1993 to 2005, which may have produced similar results (Wilkinson *et al.*, 2006, Appendix; Wellings, Palmer, Geary, *et al.*, 2016). My conclusions from the interrupted time series analyses rely on the assumption that the range of interventions and policies across Scotland and Wales would not have produced changes in rates which would not be seen in a no-Strategy England. Scotland and Wales both implemented interventions which overlapped with Strategy components, though on a notably smaller scale than the English Strategy (Table 3-1). These included improved sex and relationships education and contraceptive access interventions, highlighted in Figure 7-1 as Strategy elements. Scotland additionally saw later interventions, limited to selected areas, to improve sexual health clinic access and launch media campaigns (Elliott *et al.*,

2013). Both pre- and post-implementation outcome evaluations of Scotland's interventions appear to show little effectiveness in producing change compared with prior approaches (Henderson *et al.*, 2007; Elliott *et al.*, 2010, 2013). These uncertainties over the exposures of 'treatment as usual' limit my UK comparisons' ability to test the individual Strategy elements noted in Figure 7-1. Rather, the results of the interrupted time series analyses may be best read as testing the specific nationally coordinated and nationally funded Strategy approach over and against incremental and ongoing change in sexual health education and services. Within the results of this analysis, it is possible that the range of UK interventions – prompted by UK-wide concern over similarly high teenage pregnancy rates – produced varying effects in each population. However, the range of less funded and seemingly ineffective interventions in Scotland and Wales implemented around the time of the Strategy would be unlikely to have produced large effects at exactly the same time as the Strategy. The spill-over effects of the media campaign, though not noted as an effective approach in my review, has been hypothesised as responsible for some of the similarities of effect size and timing across England, Scotland and Wales (Wellings, Palmer, Wilkinson, *et al.*, 2016). However, the ability of the media campaign to explain all of or the majority of the effect is contradicted by previous observations of differing effect sizes in England corresponding to local implementation grant spending (Wellings, Palmer, Geary, *et al.*, 2016).

In my preparatory models using only England data with no comparators, I observed a substantial change in trend coinciding with the introduction of the Strategy, but these models were unable to test whether this effect was produced by the Strategy or was a result of other events. The comparator models conclusively showed that these changes were seen simultaneously across all three countries and were therefore not unique to the Strategy.

As a further approach I sought to compare England with a wider range of countries, by using a synthetic control. My synthetic control models used a wider pool of countries to correct for contamination of Strategy effects, co-occurring interventions and wider changes in social determinants as confounders. This method presented an additional challenge of finding comparable countries from a wide range of possible controls. Further, it required demonstrating the validity of a weighted composite of these countries for producing a counterfactual estimate of the teenage pregnancy rates that would be observed in England in the absence of the Strategy. The synthetic control models allowed a data-driven selection of countries from which to construct a close-fit comparator. However, the available data greatly reduced the number of countries from which the synthetic control

algorithm could draw in its construction of control. Most models assigned weight to only 2-3 comparator countries. This was undoubtedly driven by the large differences in rates between England and the majority of the European countries I had included. The pattern predicted by controls may have been heavily dependent on the assumption that the few weighted countries – particularly the U.S.A. – closely followed the true counterfactual pattern of an unexposed England. My robustness analyses, sequentially removing control countries, attempted to test reliance on each. The consistency of findings of no clear effects across all robustness models could be an indicator that the observed falling rates were indeed due to a common effect on all weighted comparators.

The primary comparison models assigned Scotland a majority of the weighting, which could be expected given the relative similarity in observable pre-intervention rates. These primary models are susceptible to the same limitations as the interrupted time series comparisons with Scotland. This weakens the triangulation approach of my two analyses as they do not use truly non-overlapping comparators. One approach to addressing this overlap could have been to exclude Scotland from primary models.

In sensitivity analyses I was able to fit a predictive model without Scotland data, in order to check for UK contamination and confounding from Scotland-specific interventions. These were less able to match the pattern seen in England and Wales in the pre-intervention period and so are less certain indicators of the pregnancy rate changes in the absence of the Strategy. In models with Scotland included, I observed little difference from unexposed countries' rates produced in England and Wales by exposure to the Strategy when compared with control. In most other models, England and Wales saw higher rates than controls over the Strategy period, with one model showing a small difference producing lower rates in England and Wales. Each of these were within the range of the distribution of non-Strategy effects produced by placebo models, and so these small differences may have been the result of random chance and poor model fit and not a result of Strategy exposure. Thus, whilst later models may have been less sensitive to detect change due to the effects of the Strategy, there were no clear inconsistencies in model findings over and against random noise. The primary models including Scotland can be interpreted as good fit predictors, as similar as possible to exposed England, with minimal impact of the possible contamination effect discussed above.

As discussed in Chapters 5 and 6, the outcomes used for my comparative analyses were primarily driven by the availability of data for the selected countries. These data and

outcomes had certain limitations. Estimates of under-18 pregnancies as reported by ISD Scotland and ONS most closely matched the target outcome of the Strategy (Social Exclusion Unit, 1999; Information Services Division Scotland, 2018a; Office for National Statistics, 2019). These data, however, did not include pregnancies ending in unrecorded abortions and miscarriages. Counts of births used to calculate under-18 births did not include pregnancies to 17-year-olds ending in birth at age 18, or abortions or miscarriages. Under-20 pregnancy rates included an estimate of abortions to teenagers but included 18- and 19-year-olds who may have been less directly targeted by the Strategy. All three measures do represent valuable public health outcomes. The use of all three additionally allowed comparison to check for consistency of results.

In both my approaches, across all outcomes and across all my comparative models, I saw results which were consistently compatible with a hypothesised minimal specific effect of the Strategy on reducing teenage pregnancy rates. As the estimates of 1999 trend change produced in interrupted time series approaches have relatively small confidence intervals which span zero, my observations rule out all but a very small effect of the Strategy.

7.3 Implications

Teenage pregnancy rates have been falling since 1999 in England, Scotland and Wales. My review gives a summary of reliable evidence used to test a range of hypothesised contributors to these observed effects. However, this evidence is unable to account for the large changes seen in previous decades. Other existing research, not addressed in my review, may give clearer assessments of several of these hypotheses. My causal tree provides a basis for further investigation of hypothesised causes and evidence of the effects of each of these upon the UK population across the years recording falling rates. In the case of several hypothesised causes, in particular the non-designed changes in social determinants discussed in the review, a randomised controlled trial may not be possible for testing causation. In several of these situations, natural experiment methods may be able to assess the causal impact of a non-intervention exposure.

As discussed in section 7.2 above, the hypothesised causes of improved contraceptive technology or increased employment and education, possibly driven by child and adolescent development interventions, warrant further investigation. Such investigation could initially review existing literature not included in my review for each cause. Further investigation of the range of approaches taken across countries, NHS regions and local

authorities could help to identify where and when such interventions may have been implemented and use population data to assess evidence of the exposure of populations to these branches. From this it could be determined if it were possible to treat events leading to differential exposures to hypothesised causes as a natural experiment to test it causally.

Discovering the true explanation of falling rates is crucial in future policy development. As further policies aim to build on perceived successes, or ensure that rates stay low, causal analyses of the hypothesised branches of the tree may provide clarity of direction. Understanding the causes may help to divert attention and spending away from ineffective interventions or highlight new approaches which seek to replicate the benefits of non-designed changes in social environments.

My thesis adds to the mounting evidence that the Strategy was not primarily responsible for falling teenage pregnancy rates (Paton and Wright, 2017; Heap, Berrington and Ingham, 2020; Paton, Bullivant and Soto, 2020). I used robust analyses to conduct a first comparison of England with exchangeable controls to quantitatively estimate the Strategy's effects. I triangulated across two methods and constructed controls from a range of countries to account for potential biases and confounders. My findings of no effect, combined with prior observations suggesting the same, would indicate that the combined elements of the Strategy were not effective, despite the 1999 report seeking evidence for previously successful approaches (Social Exclusion Unit, 1999).

Public health policy developers and researchers looking to further lower England's teenage pregnancy rates, or address high rates in other populations, have previously pointed to the Strategy's elements as a replicable approach to produce similar reductions (Hadley, Chandra-Mouli and Ingham, 2016; UNESCO Education Sector, 2017; Public Health England and Local Government Association, 2018). In my systematic review of plausible causes, I found little evidence of the effectiveness of contraception distribution and media campaigns in reducing teenage pregnancy. Other elements used in the Strategy – improved sexual health education in schools, improved clinic access and parental engagement – showed evidence of effectiveness in the literature, in line with the Strategy's recommendations from prior evidence (Social Exclusion Unit, 1999). However, my evaluation of the Strategy suggests that this holistic approach, combining potentially effective elements, did not produce the changes seen in England. Should future research and policy development seek to learn from the observed falling rates, further investigation

should endeavour to find the true causes. It is important to be ‘right for the right reason’ (Smith and Ebrahim, 2001).

Outcome evaluation of health interventions should remain a crucial part of continuing development and action in improving public health (House of Commons Health Committee, 2009; Theis and White, 2021). In conducting such evaluations, care must be taken to use methods suitable for detecting causal effects and excluding alternative explanations (Craig *et al.*, 2019). My interrupted time series and synthetic control analyses have demonstrated two natural experimental methods which can achieve this in the absence of a randomised controlled trial (Craig *et al.*, 2017). The design of interventions, the collection of both administrative and survey data and the reporting of outcome measures should similarly be informed by the range of suitable methods to clearly evaluate outcomes (Craig *et al.*, 2019).

Often interventions which appear feasible, acceptable and potent to produce change may fail to achieve their goals. Theis and White (2021) for example note that in the implementation of a series of nationwide anti-obesity policies occurring across a similar period to the Strategy (1991 to 2020) there are numerous examples of poor intervention development, implementation and evaluation. In spite of several consecutive strategies prompting 689 anti-obesity policies, these actions have “not yet successfully and consistently reduced obesity prevalence or health inequities” (Theis and White, 2021, p. 38). They highlight the repeated failure of these policies to plan effective action as potential explanators of this lack of success. The Strategy was a costly intervention, considering both the potential for money wasted on an ineffective intervention, estimated at £280m (Billingsley, 2011), and an opportunity cost, as potentially more effective approaches were replaced or delayed in the process of implementing the Strategy. The Strategy, similar to the obesity policies discussed in Theis and White (2021), may have misidentified factors with the greatest scope for change, or adopted ineffective approaches for achieving this change (Wight *et al.*, 2015). Addressing other modifiable mediators between socioeconomic circumstances leading to pregnancy risk, or between teenage maternity and poor health outcomes, may have achieved more as public health interventions with clear health goals.

A final key element which may prompt further investigation is the rapid uptake of smartphones and social media amongst the teenage population in the post-2007 period (Toseeb and Inkster, 2015; Royal Society for Public Health, 2017; Ofcom, 2018). Beyond

the hypotheses raised in the literature review of Chapter 2, I did not uncover any further research investigating this causal link. My interrupted time series and synthetic control analyses both however saw notable changes in trends from 2008 onwards. This was consistent with the launch of the first smartphone; incremental uptake of devices amongst younger teenagers could feasibly have contributed to a sharply declining trend in the following years. My analyses did not state this as a prior hypothesis and this trend change thus cannot be taken as a testing of the causal impact of smartphone use and social media access. Further research could aim to test the hypothesised effects of changing digital media technologies and culture on pregnancy rates. The three pathways I suggested in my literature review – the 'incarceration effect' of greater social media use, autonomous access to sexual health information, and the influence on life and career aspirations (Guldi and Herbst, 2017; Patterson, Hilton, *et al.*, 2019; The Harris Poll, 2019) – could additionally be investigated to identify further public health concerns or opportunities for health improvement.

7.4 Conclusion

My review highlighted several plausible hypothesised causes of falling teenage pregnancy rates in the UK. In some cases – e.g., changing contraceptive technologies and increased education and employment – I found evidence for their potential contributions to the observed falling rates, but no causal assessment of their effects on the UK population during the period. Several other hypotheses were presented without clear evidence to accept or dismiss the possibilities of their contributions. My constructed causal tree presents an evidence-based and further modifiable set of hypotheses to test and to further understand the causes of falling teenage pregnancy rates.

In my analyses using two complementary natural experimental methods, I applied robust testing to one prominent hypothesised cause: England's Teenage Pregnancy Strategy. In spite of review evidence indicating that the combined approaches may have been effective interventions, I was unable to detect an effect of the Strategy. Other potential causes highlighted in the review may have contributed to the overall fall in pregnancy rates. My analyses are the first quantitative outcome evaluations of the £280m Strategy which use strong causal analysis techniques and demonstrably comparable controls. They demonstrate the opportunity for and importance of better policy evaluation. Future policies and interventions seeking to further reduce teenage pregnancies, in the UK or other high-income countries, should first seek to understand the true causes of falling rates.

Appendices

Appendix A Links to open science components published online

Host	Review code etc.	ITS shiny app	Synth app and results	Code used across parts
GitHub	**github.com/andrewbaxter439/teen_preg_review_docs**	**github.com/andrewbaxter439/ITS_shinyapp**	**github.com/andrewbaxter439/synth-app**	**github.com/andrewbaxter439/teen-preg-project**
Protocol	**www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42018085759**			**osf.io/tdbr8**
OSF	**osf.io/dr42f** (10.17605/OSF.IO/DR42F)	**osf.io/j2efx** (10.17605/OSF.IO/J2EFX)	**osf.io/cdqxn** (10.17605/OSF.IO/CDQXN)	**osf.io/8u9jp** (10.17605/OSF.IO/8U9JP)
Zenodo		**doi.org/10.5281/zenodo.5141547**		**doi.org/10.5281/zenodo.5141538**
Plotly			Multiple graphs at **chart-studio.plotly.com/~andrewbaxter439#**	
andybaxter.me	**phd.andybaxter.me/review_data**		**phd.andybaxter.me/Synth-analysis**	
Shiny app		**phd.andybaxter.me/ITS_shinyapp**	**phd.andybaxter.me/synth-app**	

Appendix B Systematic review search strategy

Embase

	Search term	Results
1	adolescent pregnancy/	8984
2	Pregnancy, Unwanted/ or Pregnancy/ or Pregnancy, Unplanned/ or Pregnancy Rate/	705252
3	pregnan*.ti,ab,kw.	639370
4	2 or 3	946303
5	Adolescent/	1499054
6	adolescen*.ti,ab,kw.	326348
7	teenage$.ti,ab,kw.	26617
8	teens.ti,ab,kw.	7861
9	youth.ti,ab,kw.	68386
10	young adult$.ti,ab,kw.	109718
11	underage.ti,ab,kw.	1277
12	5 or 6 or 7 or 8 or 9 or 10 or 11	1683719
13	exp contraception/	163689
14	exp contraceptive agent/	154162
15	family planning/	37597
16	sexual education/	11505
17	contracepti*.ti,ab,kw.	81252
18	condom$.ti,ab,kw.	21520
19	birth control.ti,ab,kw.	5057
20	(sex$ adj2 education).ti,ab,kw.	14350
21	sexual behavior/	100266
22	(sex$ adj2 behavio?r).ti,ab,kw.	28534
23	((reduc* or prevent*) adj4 pregnan*).ti,ab,kw.	21589
24	13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23	448274
25	12 and 24	52701
26	4 and 25	14889
27	1 or 26	20164
28	exp animal/	24927529
29	exp nonhuman/	5471043
30	exp human/	19879391
31	(28 or 29) not 30	6633468
32	27 not 31	20028
33	limit 32 to (english language and yr="1990 -Current")	15060
34	review.kw,ti.	489221
35	exp Meta-Analysis/	145487
36	((meta adj analy$) or meta-analys$ or metaanalys$).ti,ab,kw.	172366
37	34 or 35 or 36	640187
38	33 and 37	423
39	Europe/	138385
40	North America/	33047
41	"Australia and New Zealand"/ or Australia/	159437

42	developing country/	90421
43	"Africa south of the Sahara"/ or Africa/	66062
44	india/	130716
45	39 or 40 or 41	316203
46	42 or 43 or 44	267875
47	46 not 45	252525
48	38 not 47	381

Medline

	Search term	**Results**
1	Pregnancy in Adolescence/	7457
2	Pregnancy/	819348
3	pregnan*.ti,ab,kw.	430537
4	2 or 3	882317
5	Adolescent/	1868798
6	adolescen*.ti,ab,kw.	223529
7	teenage$.ti,ab,kw.	17364
8	teens.ti,ab,kw.	5062
9	youth.ti,ab,kw.	49716
10	young adult$.ti,ab,kw.	70569
11	underage.ti,ab,kw.	858
12	5 or 6 or 7 or 8 or 9 or 10 or 11	1944846
13	Contraception/	18580
14	Contraceptive Agents/	4413
15	Sex Education/	8436
16	Family Planning Services/	23894
17	contracepti*.ti,ab,kw.	64061
18	condom$.ti,ab,kw.	17823
19	birth control.ti,ab,kw.	4595
20	sex education.ti,ab,kw.	5323
21	((reduc* or prevent*) adj4 pregnan*).ti,ab,kw.	14439
22	Sexual Behavior/	51039
23	(sex$ adj2 behavio?r).ti,ab,kw.	23354
24	13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23	160482
25	12 and 24	46273
26	4 and 25	15010
27	1 or 26	18998
28	exp Animals/	21589270
29	exp Humans/	17121896
30	28 not 29	4467374
31	27 not 30	18963
32	limit 31 to (english language and yr="1990 -Current")	12902
33	exp Review Literature as Topic/	9915
34	review.kw,ti.	329599
35	Meta-Analysis/	89399
36	Meta-Analysis as Topic/	16259

37	((meta adj analy$) or meta-analys$ or metaanalys$).ti,ab,kw.	105546
38	33 or 34 or 35 or 36 or 37	431572
39	32 and 38	324
40	EUROPE/	94777
41	NORTH AMERICA/	18673
42	AUSTRALIA/	91048
43	New Zealand/	35800
44	Developing Countries/	70191
45	"AFRICA SOUTH OF THE SAHARA"/ or AFRICA/	34060
46	india/	91914
47	40 or 41 or 42 or 43	227278
48	44 or 45 or 46	182344
49	48 not 47	173482
50	39 not 49	286

SCOPUS

TITLE-ABS-KEY (pregnan*) AND TITLE-ABS-KEY (adolescen* OR teenage* OR teen OR youth OR "young adult" OR underage) AND TITLE-ABS-KEY (contracepti* OR condom OR {birth control} OR "sex education" OR (reduc* OR prevent* W/4 pregnan*)) AND PUBYEAR > 1989 AND SRCTYPE (j) AND TITLE-ABS-KEY ("systematic review" OR "meta analy*" OR metaanaly*)

283 results

Cochrane Database of Systematic Reviews

	Search term	**Results**
#1	MeSH descriptor: [Pregnancy in Adolescence] this term only	202
#2	MeSH descriptor: [Pregnancy] this term only	64
#3	pregnan*:ti,ab,kw	38064
#4	MeSH descriptor: [Adolescent] this term only	95495
#5	(adolescen* or teenage or teen or youth or "young adult" or underage):ti,ab,kw	158206
#6	MeSH descriptor: [Contraception] explode all trees	441
#7	MeSH descriptor: [Sex Education] this term only	256
#8	MeSH descriptor: [Sexual Behavior] this term only	1800
#9	(contracepti* or condom or "birth control" or "sex education" or ((reduc* or preven*) near/4 pregnan*)):ti,ab,kw	9709
#10	#1 or ((#2 or #3) and (#4 or #5) and (#6 or #7 or #8 or #9)) - Publication Year from 1990 to 2018, in Cochrane Reviews (Reviews only) and Other Reviews (Word variations have been searched)	50

Appendix C Data sources and R packages used in analyses

Data sources

Data used	Source	URL
Numbers of abortions to under-20s in Europe	European Health Information Gateway - Health for all explorer	gateway.euro.who.int/en/hfa-explorer
Populations of countries by age	Human mortality database	www.mortality.org
Birth totals of countries by age	Human fertility database	www.humanfertility.org
Numbers of births by age, Scotland	NRS Scotland	www.nrscotland.gov.uk/statistics-and-data/statistics/stats-at-a-glance/registrar-generals-annual-review/2017
Number of conceptions by age, Scotland 1994-2016	ISD Scotland	www.isdscotland.org/Health-Topics/Sexual-Health/Publications/2018-07-03/mat_tp_table1.xls
Estimates of yearly gross domestic product (GDP), mobile phone ownership, proportion of females in population, proportion of population resident in urban settings	World Bank Open Data	data.worldbank.org
Conceptions in England and Wales to under-20s	Office for National Statistics	www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/conceptionandfertilityrates/bulletins/conceptionstatistics/2017
Pregnancy and birth rates to teenagers in the USA	The Guttmacher Institute	www.guttmacher.org/report/us-adolescent-pregnancy-trends-2013
New Zealand births by age of mother	Statistics New Zealand	stats.govt.nz
New Zealand abortions by age of mother	Statistics New Zealand	stats.govt.nz
Public spending on education	OECD Data	data.oecd.org/eduresource/public-spending-on-education.htm

Packages used in R

- shiny (Chang *et al.*, 2019)
- nlme (Pinheiro *et al.*, 2019)
- tidyverse (Wickham, 2017)
- svglite (Wickham *et al.*, 2019)
- broom (Robinson and Hayes, 2020)
- car (Fox and Weisberg, 2019)
- XLConnect (Mirai Solutions GmbH, 2018)
- officer (Gohel, 2020a)
- rvg (Gohel, 2020b)
- patchwork (Pedersen, 2019)
- export (Wenseleers and Vanderaa, 2018)
- gganimate (Pedersen and Robinson, 2019)
- readxl (Wickham and Bryan, 2019)
- Synth (Abadie, Diamond and Hainmueller, 2011)
- SPHSUgraphs (Baxter, 2019)
- ggrepel (Slowikowski, 2020)
- ggpubr (Kassambara, 2020)
- gtools (Warnes, Bolker and Lumley, 2020)
- foreach (Microsoft and Weston, 2020)
- doParallel (Corporation and Weston, 2019)
- plotly (Sievert, 2018)
- magrittr (Bache and Wickham, 2014)
- patchwork (Pedersen, 2019)
- gt (Iannone, Cheng and Schloerke, 2020)

Appendix D Included studies with quality rating and reviewed pathways

Reference	Allen-Meares, P., Montgomery, K. L. and Kim, J. S. (2013) 'School-based social work interventions: A cross-national systematic review', Social work, 58(3), pp. 253–262. doi: 10.1093/sw/swt022.
Published Year	2013
Review quality score	2 - low
Pathways reviewed	Abstinence-based education, School-based SH clinic

Reference	American Academy of Pediatrics (2001) 'Condom use by adolescents', Pediatrics. 2001/06/05, 107(6), pp. 1463–1469. doi: 10.1542/peds.107.6.1463.
Published Year	2001
Review quality score	0 - low
Pathways reviewed	Abstinence-based education, School-based STI-focussed education, Counselling or medical staff one-to-one, Adolescent SH clinic access and use, Condom promotion/distribution, School-based pregnancy education, Community-based STI education

Reference	Andrzejewski, J., Liddon, N. and Leonard, S. (2019) 'Condom availability programs in schools: A review of the literature', American Journal of Health Promotion, 33(3), pp. 457–467. doi: 10.1177/0890117118791118.
Published Year	2018
Review quality score	3 - high
Pathways reviewed	Condom promotion/distribution

Reference	Aslam, R. W. et al. (2017) 'Intervention Now to Eliminate Repeat Unintended Pregnancy in Teenagers (INTERUPT): A systematic review of intervention effectiveness and cost-effectiveness, and qualitative and realist synthesis of implementation factors and user engagement', BMC Medicine, 15(1), p. 155. doi: 10.1186/s12916-017-0904-7.
Published Year	2017
Review quality score	4 - high
Pathways reviewed	Targeting rapid-repeat pregnancies

Reference	Baldwin, M. K. and Edelman, A. B. (2013) 'The effect of long-acting reversible contraception on rapid repeat pregnancy in adolescents: A review', Journal of Adolescent Health, 52(4 SUPPL.), pp. S47–S53. doi: 10.1016/j.jadohealth.2012.10.278.
Published Year	2013
Review quality score	0 - low
Pathways reviewed	Targeting rapid-repeat pregnancies, Changing contraceptive technologies

Reference	Baxter, S. et al. (2011) 'Views of contraceptive service delivery to young people in the UK: A systematic review and thematic synthesis', Journal of Family Planning and Reproductive Health Care, 37(2), pp. 71–84. doi: 10.1136/jfprhc.2010.0014.
Published Year	2011
Review quality score	3 - high
Pathways reviewed	Adolescent SH clinic access and use

Reference	Beltz, M. A. et al. (2015) 'State policy and teen childbearing: A review of research studies', Journal of Adolescent Health, 56(2), pp. 130–138. doi: 10.1016/j.jadohealth.2014.11.001.
Published Year	2015
Review quality score	2 - low
Pathways reviewed	Abstinence-based education, Adolescent SH clinic access and use, Policies promoting education, Social support

Reference	Bennett, S. E. and Assefi, N. P. (2005) 'School-based teenage pregnancy prevention programs: A systematic review of randomized controlled trials', Journal of Adolescent Health, 36(1), pp. 72–81. doi: 10.1016/j.jadohealth.2003.11.097.
Published Year	2005
Review quality score	4 - high
Pathways reviewed	Abstinence-based education

Reference	Blank, L. et al. (2010) 'Systematic review and narrative synthesis of the effectiveness of contraceptive service interventions for young people, delivered in educational settings', Journal of Pediatric and Adolescent Gynecology, 23(6), pp. 341–351. doi: 10.1016/j.jpag.2010.03.007.
Published Year	2010
Review quality score	4 - high
Pathways reviewed	Targeting rapid-repeat pregnancies, Abstinence-based education, School-based STI-focussed education, School-based SH clinic, Personal development (inc. volunteer work), Virtual infant simulator, Peer-contact sexual health intervention

Reference	Blank, L. et al. (2012) 'Systematic review and narrative synthesis of the effectiveness of contraceptive service interventions for young people, delivered in health care settings', Health Education Research, 27(6), pp. 1102–1119. doi: 10.1093/her/cys056.
Title	Systematic review and narrative synthesis of the effectiveness of contraceptive service interventions for young people, delivered in health care settings
Published Year	2012
Review quality score	4 - high
Pathways reviewed	Targeting rapid-repeat pregnancies, Counselling or medical staff one-to-one, Adolescent SH clinic access and use, Condom promotion/distribution, Contraception access (other), Changing contraceptive technologies, Advance supply of emergency contraception

Reference	Brittain, A. W. et al. (2015) 'Youth-friendly family planning services for young people: A systematic review', American Journal of Preventive Medicine, 49(2), pp. S73–S84. doi: 10.1016/j.amepre.2015.03.019.
Published Year	2015
Review quality score	3 - high
Pathways reviewed	Peer-contact sexual health intervention

Reference	Card, J. J. (1999) 'Teen pregnancy prevention: Do any programs work?', Annual Review of Public Health, 20(1), pp. 257–285. doi: 10.1146/annurev.publhealth.20.1.257.
Published Year	1999
Review quality score	0 - low
Pathways reviewed	Targeting rapid-repeat pregnancies, Abstinence-based education, School-based STI-focussed education, School-based SH clinic, Personal development (inc. volunteer work), Peer-contact sexual health intervention, Counselling or medical staff one-to-one, Adolescent SH clinic access and use, Changing contraceptive technologies, School-based pregnancy education, Community-based pregnancy education, Community-based STI education, Family/Community engagement, Vocational/academic training, Targeting vulnerable youth

Reference	Cardoza, V. J. et al. (2012) 'Sexual health behavior interventions for U.S. Latino adolescents: A systematic review of the literature', Journal of Pediatric and Adolescent Gynecology, 25(2), pp. 136–149. doi: 10.1016/j.jpag.2011.09.011.
Published Year	2012
Review quality score	3 - high
Pathways reviewed	Abstinence-based education, School-based STI-focussed education, Virtual infant simulator, Condom promotion/distribution, Community-based pregnancy education, Community-based STI education, Family/Community engagement

Reference	Charania, M. R. et al. (2011) 'Efficacy of structural-level condom distribution interventions: A meta-analysis of U.S. and international studies, 1998–2007', AIDS and Behavior, 15(7), pp. 1283–1297. doi: 10.1007/s10461-010-9812-y.
Published Year	2011
Review quality score	4 - high
Pathways reviewed	Condom promotion/distribution

Reference	Cheyne, K. L. (1999) 'Adolescent pregnancy prevention', Current Opinions in Pediatrics, 11(6), pp. 594–597. doi: 10.1097/00008480-199912000-00022.
Published Year	1999
Review quality score	0 - low
Pathways reviewed	Abstinence-based education, Condom promotion/distribution, School-based pregnancy education

Reference	Chin, H. B. et al. (2012) 'The effectiveness of group-based comprehensive risk-reduction and abstinence education interventions to prevent or reduce the risk of adolescent pregnancy, human immunodeficiency virus, and sexually transmitted infections: Two systematic reviews for the g', American Journal of Preventive Medicine, 42(3), pp. 272–294. doi: 10.1016/j.amepre.2011.11.006.
Published Year	2012
Review quality score	4 - high
Pathways reviewed	Abstinence-based education, School-based STI-focussed education, Peer-contact sexual health intervention, School-based pregnancy education, Community-based pregnancy education, Community-based STI education

Reference	Cooper, B. et al. (2014) 'Brief sexuality communication - A behavioural intervention to advance sexually transmitted infection/HIV prevention: A systematic review', BJOG: An International Journal of Obstetrics and Gynaecology, 121(s5), pp. 92–103. doi: 10.1111/1471-0528.12877.
Published Year	2014
Review quality score	3 - high
Pathways reviewed	Counselling or medical staff one-to-one

Reference	Coyne, C. A. and D'Onofrio, B. M. (2012) 'Some (but not much) progress toward understanding teenage childbearing. A review of research from the past decade', Advances in Child Development and Behavior, 42, pp. 113–152. doi: 10.1016/b978-0-12-394388-0.00004-6.
Published Year	2012
Review quality score	0 - low
Pathways reviewed	Personal development (inc. volunteer work)

Reference	Dean, S. V. et al. (2014) 'Preconception care: Promoting reproductive planning', Reproductive Health, 11(Suppl 3). doi: 10.1186/1742-4755-11-S3-S2.
Published Year	2014
Review quality score	3 - high
Pathways reviewed	Targeting rapid-repeat pregnancies, Abstinence-based education, Peer-contact sexual health intervention, Condom promotion/distribution, Changing contraceptive technologies, Family/Community engagement, Policies promoting education

Reference	Denno, D. M., Chandra-Mouli, V. and Osman, M. (2012) 'Reaching youth with out-of-facility HIV and reproductive health services: A systematic review', Journal of Adolescent Health, 51(2), pp. 106–121. doi: 10.1016/j.jadohealth.2012.01.004.
Published Year	2012
Review quality score	3 - high
Pathways reviewed	Contraception access (other), Changing contraceptive technologies

Reference	DeSmet, A. et al. (2015) 'A systematic review and meta-analysis of interventions for sexual health promotion involving serious digital games', Games for health journal, 4(2), pp. 78–90. doi: 10.1089/g4h.2014.0110.
Published Year	2015
Review quality score	3 - high
Pathways reviewed	Digital-media based intervention (targeted)

Reference	DiCenso, A. et al. (2002) 'Interventions to reduce unintended pregnancies among adolescents: systematic review of randomised controlled trials.', BMJ, 324(7351), p. 1426. doi: 10.1136/BMJ.324.7351.1426.
Published Year	2002
Review quality score	4 - high
Pathways reviewed	Abstinence-based education, Personal development (inc. volunteer work), Peer-contact sexual health intervention, Adolescent SH clinic access and use, School-based pregnancy education, Family/Community engagement

Reference	DiCenso, A., Guyatt, G. and Willan, A. (1999) 'A systematic review of the effectiveness of adolescent pregnancy primary prevention programs'. Ontario: Public Health Branch Ontario Ministry of Health.
Published Year	1999
Review quality score	4 - high
Pathways reviewed	Abstinence-based education, Personal development (inc. volunteer work), Peer-contact sexual health intervention, Counselling or medical staff one-to-one, Adolescent SH clinic access and use, School-based pregnancy education, Community-based pregnancy education

Reference	Downing, J. et al. (2011) 'A systematic review of parent and family-based intervention effectiveness on sexual outcomes in young people', Health Education Research, 26(5), pp. 808–833. doi: 10.1093/her/cyr019.
Published Year	2011
Review quality score	3 - high
Pathways reviewed	Family/Community engagement

Reference	Eshre Capri Workshop Group (2015) 'Emergency contraception. Widely available and effective but disappointing as a public health intervention: a review', Human Reproduction, 30(4), pp. 751–760. doi: 10.1093/humrep/dev019.
Published Year	2015
Review quality score	0 - low
Pathways reviewed	Advance supply of emergency contraception

Reference	Fleming, N. et al. (2015) 'Adolescent Pregnancy Guidelines', Journal of Obstetrics and Gynaecology Canada, 37(8), pp. 740–756. doi: 10.1016/S1701-2163(15)30180-8.
Published Year	2015
Review quality score	2 - low
Pathways reviewed	Targeting rapid-repeat pregnancies, Changing contraceptive technologies

Reference	Fletcher, A. et al. (2007) 'Interventions addressing the social determinants of teenage pregnancy', Health Education, 108(1), pp. 29–39. doi: 10.1108/09654280810842111.
Published Year	2008
Review quality score	0 - low
Pathways reviewed	Personal development (inc. volunteer work), Vocational/academic training, Policies promoting education

Reference	Franklin, C. et al. (1997) 'Effectiveness of prevention programs for adolescent pregnancy: A meta-analysis', Journal of Marriage and the Family, 59(32488), pp. 551–567. doi: 10.2307/353945.
Published Year	1997
Review quality score	3 - high
Pathways reviewed	Abstinence-based education, Adolescent SH clinic access and use, School-based pregnancy education, Family/Community engagement

Reference	Franklin, C. and Corcoran, J. (2000) 'Preventing adolescent pregnancy: a review of programs and practices', Social work, 45(1), pp. 40–52. doi: 10.1093/sw/45.1.40.
Published Year	2000
Review quality score	0 - low
Pathways reviewed	Abstinence-based education, School-based SH clinic, Peer-contact sexual health intervention, Adolescent SH clinic access and use, Condom promotion/distribution, School-based pregnancy education, Community-based pregnancy education

Reference	Fullerton, D. et al. (1997) 'Preventing unintended teenage pregnancies and reducing their adverse effects', Quality in Health Care, 6, pp. 102–108. doi: 10.1136/qshc.6.2.102.
Published Year	1997
Review quality score	4 - high
Pathways reviewed	Abstinence-based education, School-based SH clinic, Personal development (inc. volunteer work), School-based pregnancy education, Family/Community engagement, Vocational/academic training

Reference	Gavin, L. E. et al. (2010) 'A review of positive youth development programs that promote adolescent sexual and reproductive health', Journal of Adolescent Health, 46(3), pp. S75–S91. doi: 10.1016/j.jadohealth.2009.11.215.
Published Year	2010
Review quality score	4 - high
Pathways reviewed	Personal development (inc. volunteer work)

Reference	Gavin, L. E. et al. (2015) 'Programs to strengthen parent-adolescent communication about reproductive health: A systematic review', American Journal of Preventive Medicine, 49(2), pp. S65–S72. doi: 10.1016/j.amepre.2015.03.022.
Published Year	2015
Review quality score	3 - high
Pathways reviewed	Family/Community engagement

Reference	Goesling, B. et al. (2014) 'Programs to reduce teen pregnancy, sexually transmitted infections, and associated sexual risk behaviors: A systematic review', Journal of Adolescent Health, 54(5), pp. 499–507. doi: 10.1016/j.jadohealth.2013.12.004.
Published Year	2014
Review quality score	4 - high
Pathways reviewed	Abstinence-based education, Personal development (inc. volunteer work), Adolescent SH clinic access and use, School-based pregnancy education

Reference	Grunseit, A. et al. (1997) 'Sexuality education and young people's sexual behavior: a review of studies', Journal of Adolescent Research, 12(4), pp. 421–453. doi: 10.1177/0743554897124002.
Published Year	1997
Review quality score	1 - low
Pathways reviewed	Abstinence-based education, School-based STI-focussed education, School-based SH clinic, Peer-contact sexual health intervention, Adolescent SH clinic access and use, School-based pregnancy education, Family/Community engagement, Public information/media campaign

Reference	Guse, K. et al. (2012) 'Interventions using new digital media to improve adolescent sexual health: A systematic review', Journal of Adolescent Health, 51(6), pp. 535–543. doi: 10.1016/j.jadohealth.2012.03.014.
Published Year	2012
Review quality score	3 - high
Pathways reviewed	Digital media-based SH intervention

Reference	Harden, A. et al. (2006) Young people, pregnancy and social exclusion: a systematic synthesis of research evidence to identify effective, appropriate and promising approaches for prevention and support, EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. London: The Evidence for Policy and Practice Information and Co-ordinating Centre. Available at: https://eppi.ioe.ac.uk/cms/Portals/0/PDF reviews and summaries/pregnancy_social_exclusion.pdf.
Published Year	2006
Review quality score	4 - high
Pathways reviewed	Personal development (inc. volunteer work), Vocational/academic training, Policies promoting education, Early-years intervention

Reference	Hawes, Z. C., Wellings, K. and Stephenson, J. (2010) 'First heterosexual intercourse in the United Kingdom: A review of the literature', Journal of Sex Research, 47(2), pp. 137–152. doi: 10.1080/00224490903509399.
Published Year	2010
Review quality score	0 - low
Pathways reviewed	School-based pregnancy education

Reference	Hieftje, K. et al. (2013) 'Electronic media-based health interventions promoting behavior change in youth: A systematic review', JAMA Pediatrics, 167(6), pp. 574–580. doi: 10.1001/jamapediatrics.2013.1095.
Published Year	2013
Review quality score	4 - high
Pathways reviewed	Digital media-based SH intervention

Reference	Hoyt, H. H. and Broom, B. L. (2002) 'School-based teen pregnancy prevention programs: a review of the literature.', The Journal of School Nursing, 18(1), pp. 11–17. doi: 10.1177/10598405020180010401.
Published Year	2002
Review quality score	0 - low
Pathways reviewed	Abstinence-based education, Personal development (inc. volunteer work), Virtual infant simulator, School-based pregnancy education

Reference	Jackson, C. et al. (2012) 'Interventions to prevent substance use and risky sexual behaviour in young people: a systematic review', Addiction, 107(4), pp. 733–747. doi: 10.1111/j.1360-0443.2011.03751.x.
Published Year	2012
Review quality score	4 - high
Pathways reviewed	Personal development (inc. volunteer work), Family/Community engagement

Reference	Johnson, B. T. et al. (2011) 'Interventions to reduce sexual risk for human immunodeficiency virus in adolescents: A meta-analysis of trials, 1985-2008', Archives of Pediatrics & Adolescent Medicine, 165(1), pp. 77–84. doi: 10.1001/archpediatrics.2010.251.
Published Year	2011
Review quality score	4 - high
Pathways reviewed	Abstinence-based education, Targeting vulnerable youth

Reference	Jones, K. et al. (2014) 'The impact of health education transmitted via social media or text messaging on adolescent and young adult risky sexual behavior: A systematic review of the literature', Sexually Transmitted Diseases, 41(7). doi: 10.1097/OLQ.0000000000000146.
Published Year	2014
Review quality score	3 - high
Pathways reviewed	Digital media-based SH intervention

Reference	Kao, T.-S. et al. (2012) 'A comparison of family interventions to address adolescent risky behaviors: A literature review', Western Journal of Nursing Research, 35(5), pp. 611–637. doi: 10.1177/0193945912465021.
Published Year	2012
Review quality score	1 - low
Pathways reviewed	Family/Community engagement

Reference	Kerr, D. L. and Matlak, K. A. (1998) 'Alcohol use and sexual risk-taking among adolescents: A review of recent literature', Journal of HIV/AIDS Prevention and Education for Adolescents and Children, 2(2), pp. 67–88. doi: 10.1300/J129v02n02_05.
Published Year	1998
Review quality score	1 - low
Pathways reviewed	Educational/career aspiration/planning, Adolescent alcohol use

Reference	Kim, C. R. and Free, C. (2008) 'Recent evaluations of the peer-led approach in adolescent sexual health education: a systematic review', International Family Planning Perspectives. 2008/07/23, 34(2), pp. 89–96. doi: 10.1363/ifpp.34.0089.08.
Published Year	2008
Review quality score	4 - high
Pathways reviewed	Peer-contact sexual health intervention

Reference	Kirby, D. et al. (1994) 'School-based programs to reduce sexual risk behaviors: A review of effectiveness', Public Health Reports, 109(3), pp. 339–360. Available at: https://www.jstor.org/stable/4597597.
Published Year	1994
Review quality score	3 - high
Pathways reviewed	Abstinence-based education, School-based STI-focussed education, School-based SH clinic, Peer-contact sexual health intervention, Condom promotion/distribution, Contraception access (other), School-based pregnancy education

Reference	Kirby, D. (2001) 'Emerging Answers: Research findings on programs to reduce teen pregnancy (Summary)', American Journal of Health Education, 32(6), pp. 348–355. doi: 10.1080/19325037.2001.10603497.
Published Year	2001
Review quality score	3 - high
Pathways reviewed	Abstinence-based education, School-based STI-focussed education, School-based SH clinic, Personal development (inc. volunteer work), Counselling or medical staff one-to-one, Adolescent SH clinic access and use, Condom promotion/distribution, Contraception access (other), Community-based pregnancy education, Community-based STI education, Family/Community engagement, Early-years intervention

Reference	Kirby, D. (2002a) 'Effective approaches to reducing adolescent unprotected sex, pregnancy, and childbearing', Journal of Sex Research, 39(1), pp. 51–57. doi: 10.1080/00224490209552120.
Published Year	2002
Review quality score	3 - high
Pathways reviewed	Abstinence-based education, School-based STI-focussed education, Personal development (inc. volunteer work), Counselling or medical staff one-to-one, Targeting vulnerable youth

Reference	Kirby, D. (2002b) 'The impact of schools and school programs upon adolescent sexual behavior', Journal of Sex Research. 2002/12/12, 39(1), pp. 27–33. doi: 10.1080/00224490209552116.
Published Year	2002
Review quality score	0 - low
Pathways reviewed	Abstinence-based education, School-based STI-focussed education, School-based SH clinic, Personal development (inc. volunteer work), Condom promotion/distribution, Policies promoting education, Educational/career aspiration/planning, More years in school across population

Reference	Kirby, D. and Coyle, K. (1997) 'Youth Development Programs', Children and Youth Services Review, 19(5), pp. 437–454. doi: 10.1016/S0190-7409(97)00026-1.
Published Year	1997
Review quality score	1 - low
Pathways reviewed	Personal development (inc. volunteer work), Vocational/academic training

Reference	Kirchengast, S. (2012) 'Teenage-pregnancies from a human life history viewpoint - an updated review with special respect to prevention strategies', Current Women's Health Reviews, 8(3), pp. 248–255. doi: 10.2174/157340412803760685.
Published Year	2012
Review quality score	0 - low
Pathways reviewed	Abstinence-based education, Contraception access (other), Changing contraceptive technologies, School-based pregnancy education, Public information/media campaign

Reference	L'Engle, K. L. et al. (2016) 'Mobile phone interventions for adolescent sexual and reproductive health: A systematic review', Pediatrics, 138(3). doi: 10.1542/peds.2016-0884.
Published Year	2016
Review quality score	3 - high
Pathways reviewed	Digital-media based intervention (targeted)

Reference	Lazarus, J. V et al. (2010) 'Systematic review of interventions to prevent the spread of sexually transmitted infections, including HIV, among young people in Europe', Croatian Medical Journal, 51(1), pp. 74–84. doi: 10.3325/cmj.2010.51.74.
Published Year	2010
Review quality score	3 - high
Pathways reviewed	School-based STI-focussed education, School-based SH clinic, Peer-contact sexual health intervention, Adolescent SH clinic access and use

Reference	Lin, J. S. et al. (2008) 'Behavioral counseling to prevent sexually transmitted infections: A systematic review for the U.S. Preventive Services Task Force', Annals of Internal Medicine, 149(7), pp. 497–508. doi: 10.7326/0003-4819-149-7-200810070-00011.
Published Year	2008
Review quality score	4 - high
Pathways reviewed	Counselling or medical staff one-to-one

Reference	Lopez, L. M. et al. (2009) 'Theory-based strategies for improving contraceptive use: a systematic review', Contraception, 79(6), pp. 411–417. doi: 10.1016/j.contraception.2008.12.006.
Published Year	2009
Review quality score	4 - high
Pathways reviewed	School-based STI-focussed education, Community-based STI education

Reference	Lopez, L. M., Bernholc, A., et al. (2016) 'School-based interventions for improving contraceptive use in adolescents', Cochrane Database of Systematic Reviews, 6(CD012249). doi: 10.1002/14651858.CD012249.
Published Year	2016
Review quality score	4 - high
Pathways reviewed	Abstinence-based education, School-based STI-focussed education, Peer-contact sexual health intervention, School-based pregnancy education

Reference	Lopez, L. M., Grey, T. W., et al. (2016) 'Theory-based interventions for contraception', Cochrane Database of Systematic Reviews, 11(CD007249). doi: 10.1002/14651858.CD007249.pub5.
Published Year	2016
Review quality score	4 - high
Pathways reviewed	Targeting rapid-repeat pregnancies, Abstinence-based education, Peer-contact sexual health intervention, School-based pregnancy education

Reference	Manlove, J., Fish, H. and Moore, K. A. (2015) 'Programs to improve adolescent sexual and reproductive health in the US: a review of the evidence', Adolescent Health, Medicine and Therapeutics, p. 47. doi: 10.2147/AHMT.S48054.
Published Year	2015
Review quality score	3 - high
Pathways reviewed	Abstinence-based education, School-based STI-focussed education, Counselling or medical staff one-to-one, Adolescent SH clinic access and use, School-based pregnancy education, Family/Community engagement

Reference	Maravilla, J. C. et al. (2016) 'The Role of Community Health Workers in Preventing Adolescent Repeat Pregnancies and Births', Journal of Adolescent Health, 59(4), pp. 378–390. doi: 10.1016/j.jadohealth.2016.05.011.
Published Year	2016
Review quality score	4 - high
Pathways reviewed	Targeting rapid-repeat pregnancies, Counselling or medical staff one-to-one

Reference	Marseille, E. et al. (2018) 'Effectiveness of school-based teen pregnancy prevention programs in the USA: A systematic review and meta-analysis', Prevention Science, 19(4), pp. 468–489. doi: 10.1007/s11121-017-0861-6.
Published Year	2018
Review quality score	4 - high
Pathways reviewed	Abstinence-based education, School-based STI-focussed education, Personal development (inc. volunteer work), Peer-contact sexual health intervention, School-based pregnancy education

Reference	Mason-Jones, A. J. et al. (2012) 'A systematic review of the role of school-based healthcare in adolescent sexual, reproductive, and mental health', Systematic Reviews, 1(49). doi: 10.1186/2046-4053-1-49.
Published Year	2012
Review quality score	3 - high
Pathways reviewed	School-based STI-focussed education, School-based SH clinic

Reference	Matthias, K. (2002) Youth-specific primary health care – access, utilisation and health outcomes: a critical appraisal of the literature. Christchurch, New Zealand: New Zealand Health Technology Assessment.
Published Year	2002
Review quality score	4 - high
Pathways reviewed	School-based SH clinic

Reference	McLellan, J. and Dale, H. (2013) 'Can technology be effective in interventions targeting sexual health and substance use in young people; a systematic review', Health and Technology, 3(3), pp. 195–203. doi: 10.1007/s12553-013-0059-2.
Published Year	2013
Review quality score	3 - high
Pathways reviewed	Digital media-based SH intervention

Reference	Meyer, J. L., Gold, M. A. and Haggerty, C. L. (2011) 'Advance provision of emergency contraception among adolescent and young adult women: A systematic review of literature', Journal of Pediatric and Adolescent Gynecology, 24(1), pp. 2–9. doi: 10.1016/j.jpag.2010.06.002.
Published Year	2011
Review quality score	4 - high
Pathways reviewed	Advance supply of emergency contraception

Reference	Nitz, K. (1999) 'Adolescent pregnancy prevention: A review of interventions and programs', Clinical Psychology Review, 19(4), pp. 457–471. doi: 10.1016/S0272-7358(98)00087-7.
Published Year	1999
Review quality score	0 - low
Pathways reviewed	Targeting rapid-repeat pregnancies, Abstinence-based education, School-based SH clinic, Personal development (inc. volunteer work), Peer-contact sexual health intervention, Counselling or medical staff one-to-one, Adolescent SH clinic access and use, School-based pregnancy education, Community-based pregnancy education, Family/Community engagement, Vocational/academic training, Targeting vulnerable youth, Employment outside school hours

Reference	O'Connor, E. A. et al. (2014) 'Behavioral sexual risk-reduction counseling in primary care to prevent sexually transmitted infections: A systematic review for the U.S. Preventive Services task force', Annals of Internal Medicine, 161(12), pp. 874–883. doi: 10.7326/M14-0475.
Published Year	2014
Review quality score	4 - high
Pathways reviewed	Counselling or medical staff one-to-one

Reference	Ott, M. A. and Santelli, J. S. (2007) 'Abstinence and abstinence-only education', Current Opinion in Obstetrics and Gynecology, 19(5), pp. 446–452. doi: 10.1097/GCO.0b013e3282efdc0b.
Published Year	2007
Review quality score	0 - low
Pathways reviewed	Abstinence-based education

Reference	Pierre, N. and Cox, J. (1997) 'Teenage pregnancy prevention programs', Current Opinion in Pediatrics, 9, pp. 310–316. doi: 10.1097/00008480-199708000-00003.
Published Year	1997
Review quality score	0 - low
Pathways reviewed	Targeting rapid-repeat pregnancies, Abstinence-based education, School-based SH clinic, Peer-contact sexual health intervention, Counselling or medical staff one-to-one, Adolescent SH clinic access and use, Condom promotion/distribution, School-based pregnancy education, Family/Community engagement

Reference	Robin, L. et al. (2004) 'Behavioral interventions to reduce incidence of HIV, STD, and pregnancy among adolescents: A decade in review', Journal of Adolescent Health, 34(1), pp. 3–26. doi: 10.1016/S1054-139X(03)00244-1.
Published Year	2004
Review quality score	2 - low
Pathways reviewed	Abstinence-based education, School-based STI-focussed education, Personal development (inc. volunteer work), Adolescent SH clinic access and use, School-based pregnancy education, Community-based STI education, Family/Community engagement, Targeting vulnerable youth

Reference	Rodriguez, M. I. et al. (2013) 'Advance supply of emergency contraception: A systematic review', Contraception, 87(5), pp. 590–601. doi: 10.1016/j.contraception.2012.09.011.
Published Year	2013
Review quality score	4 - high
Pathways reviewed	Targeting rapid-repeat pregnancies, Advance supply of emergency contraception

Reference	Sedlecky, K. and Stanković, Z. (2016) 'Contraception for adolescents after abortion', European Journal of Contraception and Reproductive Health Care, 21(1), pp. 4–14. doi: 10.3109/13625187.2015.1082175.
Published Year	2016
Review quality score	0 - low
Pathways reviewed	Targeting rapid-repeat pregnancies, Counselling or medical staff one-to-one

Reference	SmithBattle, L. et al. (2017) 'An umbrella review of meta-analyses of interventions to improve maternal outcomes for teen mothers', Journal of Adolescence, 59, pp. 97–111. doi: 10.1016/j.adolescence.2017.05.022.
Published Year	2017
Review quality score	4 - high
Pathways reviewed	Targeting rapid-repeat pregnancies

Reference	Spear, H. J. and Lock, S. (2003) 'Qualitative research on adolescent pregnancy: A descriptive review and analysis', Journal of Pediatric Nursing, 18(6), pp. 397–408. doi: 10.1016/S0882-5963(03)00160-X.
Published Year	2003
Review quality score	2 - low
Pathways reviewed	Educational/career aspiration/planning

Reference	Steenland, M. W. et al. (2013) 'The effect of follow-up visits or contacts after contraceptive initiation on method continuation and correct use', Contraception, 87(5), pp. 625–630. doi: 10.1016/j.contraception.2012.09.018.
Published Year	2013
Review quality score	4 - high
Pathways reviewed	Contraception initiation follow-up

Reference	Strunk, J. A. (2008) 'The effect of school-based health clinics on teenage pregnancy and parenting outcomes: an integrated literature review', The Journal of school nursing, 24(1), pp. 13–20. doi: 10.1177/10598405080240010301.
Published Year	2008
Review quality score	2 - low
Pathways reviewed	Targeting rapid-repeat pregnancies, School-based SH clinic

Reference	Sutton, M. Y. et al. (2014) 'Impact of parent-child communication interventions on sex behaviors and cognitive outcomes for Black/African-American and Hispanic/Latino youth: A systematic review, 1988-2012', Journal of Adolescent Health, 54(4), pp. 369–384. doi: https://doi.org/10.1016/j.jadohealth.2013.11.004.
Published Year	2014
Review quality score	4 - high
Pathways reviewed	Family/Community engagement

Reference	Tang, J. H. et al. (2012) 'Hormonal and intrauterine methods for contraception for women aged 25 years and younger', Cochrane database of systematic reviews, 11(CD009805). doi: 10.1002/14651858.CD009805.pub2.
Published Year	2012
Review quality score	4 - high
Pathways reviewed	Changing contraceptive technologies

Reference	Thomas, M. H. (2000) 'Abstinence-based programs for prevention of adolescent pregnancies: A review', Journal of Adolescent Health, 26(1), pp. 5–17. doi: 10.1016/s1054-139x(98)00126-8.
Published Year	2000
Review quality score	0 - low
Pathways reviewed	Abstinence-based education

Reference	Tolli, M. V (2012) 'Effectiveness of peer education interventions for HIV prevention, adolescent pregnancy prevention and sexual health promotion for young people: A systematic review of European studies', Health Education Research, 27(5), pp. 904–913. doi: 10.1093/her/cys055.
Published Year	2012
Review quality score	4 - high
Pathways reviewed	Peer-contact sexual health intervention

Reference	Trivedi, D. et al. (2009) 'Early fatherhood: a mapping of the evidence base relating to pregnancy prevention and parenting support', Health Education Research, 24(6), pp. 999–1028. doi: 10.1093/her/cyp025.
Published Year	2009
Review quality score	2 - low
Pathways reviewed	School-based pregnancy education, Community-based pregnancy education

Reference	Underhill, K., Montgomery, P. and Operario, D. (2008) 'Abstinence-plus programs for HIV infection prevention in high-income countries', Cochrane Database of Systematic Reviews, 1(CD007006). doi: 10.1002/14651858.CD007006.
Published Year	2008
Review quality score	4 - high
Pathways reviewed	Abstinence-based education

Reference	Underhill, K., Operario, D. and Montgomery, P. (2007) 'Abstinence-only programs for HIV infection prevention in high-income countries', Cochrane Database of Systematic Reviews, 4(CD005421). doi: 10.1002/14651858.CD005421.pub2.
Published Year	2007
Review quality score	4 - high
Pathways reviewed	Abstinence-based education

Reference	Usinger, K. M. et al. (2016) 'Intrauterine contraception continuation in adolescents and young women: A systematic review', Journal of Pediatric and Adolescent Gynecology, 29(6), pp. 659–667. doi: 10.1016/j.jpag.2016.06.007.
Published Year	2016
Review quality score	3 - high
Pathways reviewed	Changing contraceptive technologies

Reference	Wakhisi, A. et al. (2011) 'The effectiveness of social marketing in reduction of teenage pregnancies: a review of studies in developed countries (Structured abstract)', Social Marketing Quarterly, pp. 56–90. doi: 10.1080/15245004.2010.546941.
Published Year	2011
Review quality score	4 - high
Pathways reviewed	School-based SH clinic, Peer-contact sexual health intervention, Adolescent SH clinic access and use, School-based pregnancy education, Community-based pregnancy education, Family/Community engagement

Reference	Widman, L. et al. (2018) 'Technology-based interventions to reduce sexually transmitted infections and unintended pregnancy among youth', Journal of Adolescent Health, 62(6), pp. 651–660. doi: 10.1016/j.jadohealth.2018.02.007.
Published Year	2018
Review quality score	4 - high
Pathways reviewed	Digital media-based SH intervention

Reference	Wight, D. and Fullerton, D. (2013) 'A review of interventions with parents to promote the sexual health of their children', Journal of Adolescent Health, 52(1), pp. 4–27. doi: 10.1016/j.jadohealth.2012.04.014.
Published Year	2013
Review quality score	4 - high
Pathways reviewed	Family/Community engagement

Reference	Wilson, A. et al. (2015) 'Motivational interviews to improve contraceptive use in populations at high risk of unintended pregnancy: A systematic review and meta-analysis', European Journal of Obstetrics Gynecology and Reproductive Biology, 191, pp. 72–79. doi: 10.1016/j.ejogrb.2015.05.010.
Published Year	2015
Review quality score	4 - high
Pathways reviewed	Counselling or medical staff one-to-one

Reference	Zapata, L. B. et al. (2015) 'Impact of contraceptive counseling in clinical settings: A systematic review', American Journal of Preventive Medicine, 49(2), pp. S31–S45. doi: 10.1016/j.amepre.2015.03.023.
Published Year	2015
Review quality score	4 - high
Pathways reviewed	Counselling or medical staff one-to-one

Reference	Zoritch, B., Roberts, I. and Oakley, A. (1998) 'The health and welfare effects of day-care: A systematic review of randomised controlled trials', Social Science and Medicine, 47(3), pp. 317–327.
Published Year	1998
Review quality score	4 - high
Pathways reviewed	Early-years intervention

Appendix E Review quality by domain of risk of bias

Reference	1 – Was it systematic: defining objectives and setting appropriate study eligibility criteria?	2 – Were search and screening methods appropriate/clear and not likely to have missed key studies?	3 – Were findings synthesised systematically: presenting all available evidence?	4 – Was risk of bias and confounding adequately assessed and were results presented to take account of this?	score	rating
Aslam et al., 2017	Yes	Yes	Yes	Yes	4	high
Bennett and Assefi, 2005	Yes	Yes	Yes	Yes	4	high
Blank et al., 2010	Yes	Yes	Yes	Yes	4	high
Blank et al., 2012	Yes	Yes	Yes	Yes	4	high
Charania et al., 2011	Yes	Yes	Yes	Yes	4	high
Chin et al., 2012	Yes	Yes	Yes	Yes	4	high
DiCenso, Guyatt and Willan, 1999	Yes	Yes	Yes	Yes	4	high
DiCenso et al., 2002	Yes	Yes	Yes	Yes	4	high
Fullerton et al., 1997	Yes	Yes	Yes	Yes	4	high
Gavin et al., 2010	Yes	Yes	Yes	Yes	4	high
Goesling et al., 2014	Yes	Yes	Yes	Yes	4	high
Harden et al., 2006	Yes	Yes	Yes	Yes	4	high
Hieftje et al., 2013	Yes	Yes	Yes	Yes	4	high
Jackson et al., 2012	Yes	Yes	Yes	Yes	4	high
Johnson et al., 2011	Yes	Yes	Yes	Yes	4	high
Kim and Free 2008	Yes	Yes	Yes	Yes	4	high
Lin et al., 2008	Yes	Yes	Yes	Yes	4	high
Lopez et al., 2009	Yes	Yes	Yes	Yes	4	high
Lopez, Grey, et al., 2016	Yes	Yes	Yes	Yes	4	high

Reference	1 – Was it systematic: defining objectives and setting appropriate study eligibility criteria?	2 – Were search and screening methods appropriate/clear and not likely to have missed key studies?	3 – Were findings synthesised systematically: presenting all available evidence?	4 – Was risk of bias and confounding adequately assessed and were results presented to take account of this?	score	rating
Lopez, Bernholc, et al., 2016	Yes	Yes	Yes	Yes	4	high
Maravilla et al., 2016	Yes	Yes	Yes	Yes	4	high
Marseille et al., 2018	Yes	Yes	Yes	Yes	4	high
Matthias, 2002	Yes	Yes	Yes	Yes	4	high
Meyer et al., 2011	Yes	Yes	Yes	Yes	4	high
O'Connor et al., 2014	Yes	Yes	Yes	Yes	4	high
Rodriguez et al., 2013	Yes	Yes	Yes	Yes	4	high
SmithBattle et al., 2017	Yes	Yes	Yes	Yes	4	high
Steenland et al., 2013	Yes	Yes	Yes	Yes	4	high
Sutton et al., 2014	Yes	Yes	Yes	Yes	4	high
Tang et al., 2012	Yes	Yes	Yes	Yes	4	high
Tolli, 2012	Yes	Yes	Yes	Yes	4	high
Underhill, Operario and Montgomery, 2007	Yes	Yes	Yes	Yes	4	high
Underhill, Montgomery and Operario, 2008	Yes	Yes	Yes	Yes	4	high
Wakhisi et al., 2011	Yes	Yes	Yes	Yes	4	high
Widman et al., 2018	Yes	Yes	Yes	Yes	4	high
Wight and Fullerton, 2013	Yes	Yes	Yes	Yes	4	high
Wilson et al., 2015	Yes	Yes	Yes	Yes	4	high
Zapata et al., 2015	Yes	Yes	Yes	Yes	4	high
Zoritch et al., 1998	Yes	Yes	Yes	Yes	4	high

Reference	1 – Was it systematic: defining objectives and setting appropriate study eligibility criteria?	2 – Were search and screening methods appropriate/clear and not likely to have missed key studies?	3 – Were findings synthesised systematically: presenting all available evidence?	4 – Was risk of bias and confounding adequately assessed and were results presented to take account of this?	score	rating
Andrzejewski, Liddon and Leonard, 2019	Yes	Yes	Yes	No	3	high
Baxter et al., 2011b	Yes	Yes	Yes	No	3	high
Brittain et al., 2015b	Yes	Yes	Yes	No	3	high
Cardoza et al., 2012	Yes	Yes	Yes	No	3	high
Cooper et al., 2014	Yes	Yes	Yes	No	3	high
Dean et al., 2014	Yes	Yes	No	Yes	3	high
Denno, Chandra-Mouli and Osman, 2012	Yes	Yes	Yes	No	3	high
DeSmet et al., 2015	Yes	Yes	Yes	No	3	high
Downing et al., 2011	Yes	Yes	Yes	No	3	high
Franklin et al., 1997	Yes	Yes	Yes	No	3	high
Gavin et al., 2015	Yes	No	Yes	Yes	3	high
Guse et al., 2012	Yes	Yes	Yes	No	3	high
Jones et al., 2014	Yes	Yes	Yes	No	3	high
Kirby et al., 1994	Yes	Yes	Yes	No	3	high
Kirby 2001	Yes	Yes	No	Yes	3	high
Kirby, 2002a	Yes	Yes	No	Yes	3	high
L'Engle et al., 2016	Yes	Yes	Yes	No	3	high
Lazarus et al., 2010	Yes	Yes	Yes	No	3	high
Manlove, Fish and Moore, 2015	Yes	Yes	Yes	No	3	high
Mason-Jones et al., 2012	Yes	Yes	Yes	No	3	high

Reference	1 – Was it systematic: defining objectives and setting appropriate study eligibility criteria?	2 – Were search and screening methods appropriate/clear and not likely to have missed key studies?	3 – Were findings synthesised systematically: presenting all available evidence?	4 – Was risk of bias and confounding adequately assessed and were results presented to take account of this?	score	rating
McLellan and Dale, 2013	Yes	Yes	Yes	No	3	high
Usinger et al., 2016	Yes	Yes	Yes	No	3	high
Allen-Meares, Montgomery and Kim, 2013	Yes	No	Yes	No	2	low
Beltz et al., 2015	Yes	No	Yes	No	2	low
Fleming et al. 2015	No	Yes	No	Yes	2	low
Robin et al., 2004	No	No	Yes	Yes	2	low
Spear and Lock, 2003	Yes	Yes	No	No	2	low
Strunk, 2008	No	Yes	Yes	No	2	low
Trivedi et al., 2009	No	Yes	Yes	No	2	low
Grunseit et al., 1997	No	No	No	Yes	1	low
Kao et al., 2012	No	No	Yes	No	1	low
Kerr and Matlak, 1998	No	No	Yes	No	1	low
Kirby and Coyle, 1997	No	No	No	Yes	1	low
American Academy of Pediatrics, 2001	No	No	No	No	0	low
Baldwin and Edelman, 2013	No	No	No	No	0	low
Cheyne, 1999	No	No	No	No	0	low
Coyne and D'Onofrio, 2012	No	No	No	No	0	low
Eshre Capri Workshop Group, 2015	No	No	No	No	0	low
Fletcher et al., 2007	No	No	No	No	0	low
Franklin and Corcoran, 2000	No	No	No	No	0	low

Reference	1 – Was it systematic: defining objectives and setting appropriate study eligibility criteria?	2 – Were search and screening methods appropriate/clear and not likely to have missed key studies?	3 – Were findings synthesised systematically: presenting all available evidence?	4 – Was risk of bias and confounding adequately assessed and were results presented to take account of this?	score	rating
Hawes, Wellings and Stephenson, 2010	No	No	No	No	0	low
Hoyt and Broom, 2002	No	No	No	No	0	low
Card, 1999	No	No	No	No	0	low
Kirby, 2002a	No	No	No	No	0	low
Kirchengast 2012	No	No	No	No	0	low
Nitz, 1999	No	No	No	No	0	low
Ott and Santelli, 2007	No	No	No	No	0	low
Pierre and Cox, 1997	No	No	No	No	0	low
Sedlecky and Stancovic, 2016	No	No	No	No	0	low
Thomas 2000	No	No	No	No	0	low

See Appendix D for full references of all included studies

References

Abadie, A. and Cattaneo, M. D. (2018) 'Econometric methods for program evaluation', *Annual Review of Economics*, 10(1). doi: 10.1146/annurev-economics-080217-053402.

Abadie, A., Diamond, A. and Hainmueller, A. J. (2010) 'Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program', *Journal of the American Statistical Association*. doi: 10.1198/jasa.2009.ap08746.

Abadie, A., Diamond, A. and Hainmueller, J. (2011) 'Synth: An R package for synthetic control methods in comparative case studies', *Journal of Statistical Software*, 42(13), pp. 1–17. Available at: http://www.jstatsoft.org/v42/i13/.

Abadie, A., Diamond, A. and Hainmueller, J. (2015) 'Comparative politics and the synthetic control method', *American Journal of Political Science*, 59(2), pp. 495–510. doi: 10.1111/ajps.12116.

Adler, M. W. (1997) 'Sexual health–a Health of the Nation failure', *BMJ*, 314(7096), p. 1743. doi: 10.1136/bmj.314.7096.1743.

Andrzejewski, J., Liddon, N. and Leonard, S. (2019) 'Condom availability programs in schools: A review of the literature', *American Journal of Health Promotion*, 33(3), pp. 457–467. doi: 10.1177/0890117118791118.

Arai, L. (2009) *Teenage pregnancy: The making and unmaking of a problem*. The Policy Press. doi: 10.1080/14681811.2011.609056.

Aslam, R. W. *et al.* (2017) 'Intervention Now to Eliminate Repeat Unintended Pregnancy in Teenagers (INTERUPT): A systematic review of intervention effectiveness and cost-effectiveness, and qualitative and realist synthesis of implementation factors and user engagement', *BMC Medicine*, 15(1), p. 155. doi: 10.1186/s12916-017-0904-7.

Bache, S. M. and Wickham, H. (2014) 'magrittr: A Forward-Pipe Operator for R', R package version 1.5. Available at: https://cran.r-project.org/package=magrittr.

Baggott, R. (2012) 'Policy success and public health: The case of public health in England', *Journal of Social Policy*, 41(2), pp. 391–408. doi:

10.1017/S0047279411000985.

Balshem, H. *et al.* (2011) 'GRADE guidelines: 3. Rating the quality of evidence.', *Journal of clinical epidemiology*, 64(4), pp. 401–6. doi: 10.1016/j.jclinepi.2010.07.015.

Barnet, B. *et al.* (2010) 'Cost-effectiveness of a motivational intervention to reduce rapid repeated childbearing in high-risk adolescent mothers a rebirth of economic and policy considerations', *Archives of Pediatrics and Adolescent Medicine*, 164(4), pp. 370–376. doi: 10.1001/archpediatrics.2010.16.

Barrett, G., Smith, S. and Wellings, K. (2004) 'Conceptualisation, development, and evaluation of a measure of unplanned pregnancy', *Journal of Epidemiology and Community Health*, 58(5), pp. 426–433. doi: 10.1136/jech.2003.014787.

Barrett, G. and Wellings, K. (2000) 'Understanding pregnancy intentions: A problem in evidence everywhere', *Family Planning Perspectives*, 32(4), pp. 194–194. Available at: https://www.jstor.org/stable/2648238 (Accessed: 19 October 2020).

Baxter, A. J. (2019) 'SPHSUgraphs: Select SPHSU brand colours For ggplot2 graphs', R package version 0.1.0. Available at: https://github.com/andrewbaxter439/SPHSUgraphs.

Baxter, S. *et al.* (2011) 'Views of contraceptive service delivery to young people in the UK: A systematic review and thematic synthesis', *Journal of Family Planning and Reproductive Health Care*, 37(2), pp. 71–84. doi: 10.1136/jfprhc.2010.0014.

Bell, J. *et al.* (2004) *Living on the edge: Sexual behaviour and young parenthood in seaside and rural areas*. Hull: University of Hull. Available at: https://research.brighton.ac.uk/en/publications/living-on-the-edge-sexual-behaviour-and-young-parenthood-in-rural (Accessed: 1 September 2020).

Beltz, M. A. *et al.* (2015) 'State policy and teen childbearing: A review of research studies', *Journal of Adolescent Health*, 56(2), pp. 130–138. doi: 10.1016/j.jadohealth.2014.11.001.

Bennett, S. E. and Assefi, N. P. (2005) 'School-based teenage pregnancy prevention programs: A systematic review of randomized controlled trials', *Journal of Adolescent Health*, 36(1), pp. 72–81. doi: 10.1016/j.jadohealth.2003.11.097.

Billingsley, M. (2011) ‘Cuts threaten to undermine progress in reducing teenage pregnancies, charities claim’, *BMJ*, 342(d1264). doi: 10.1136/bmj.d1264.

Black, N. (1996) ‘Why we need observational studies to evaluate the effectiveness of health care’, *BMJ*, 312, pp. 1215–1218. doi: 10.1136/bmj.312.7040.1215.

Blair, T. (1997a) *Blair: Why we must help those excluded from society*, *The Independent*. Available at: https://www.independent.co.uk/news/blair-why-we-must-help-those-excluded-from-society-1287576.html (Accessed: 6 October 2020).

Blair, T. (1997b) *‘Bringing Britain Together’; Speech Archive*, *British Political Speech*. Available at: http://www.britishpoliticalspeech.org/speech-archive.htm?speech=320 (Accessed: 6 October 2020).

Blank, L. *et al.* (2010a) ‘Systematic review and narrative synthesis of the effectiveness of contraceptive service interventions for young people, delivered in educational settings’, *Journal of Pediatric and Adolescent Gynecology*, 23(6), pp. 341–351. doi: 10.1016/j.jpag.2010.03.007.

Blank, L. *et al.* (2010b) ‘Systematic Review and Narrative Synthesis of the Effectiveness of Contraceptive Service Interventions for Young People, Delivered in Educational Settings’, *Journal of Pediatric and Adolescent Gynecology*, 23(6), pp. 341–351. doi: 10.1016/j.jpag.2010.03.007.

Blank, L. *et al.* (2012) ‘Systematic review and narrative synthesis of the effectiveness of contraceptive service interventions for young people, delivered in health care settings’, *Health Education Research*, 27(6), pp. 1102–1119. doi: 10.1093/her/cys056.

Borzekowski, D. L. G. and Rickert, V. I. (2001) ‘Adolescents, the Internet, and health: Issues of access and content’, *Journal of Applied Developmental Psychology*, 22(1), pp. 49–59. doi: 10.1016/S0193-3973(00)00065-4.

Bouttell, J. *et al.* (2018) ‘Synthetic control methodology as a tool for evaluating population-level health interventions’, *Journal of Epidemiology and Community Health*, 72(8), pp. 673–678. doi: 10.1136/JECH-2017-210106.

Bowring, A. L. *et al.* (2018) ‘Features of successful sexual health promotion programs for

young people: findings from a review of systematic reviews', *Health Promotion Journal of Australia*, 29(1), pp. 46–57. doi: 10.1002/hpja.3.

Bowring, A. L., Lim, M. C. and Hider, K. (2016) *Interventions to improve the sexual and reproductive health of young people: a systematic review of reviews.* Melbourne, Australia: Family Planning Victoria. Available at: http://www.fpv.org.au/assets/resources/SRH_Interventions_SRoR_Final-18082016.pdf.

Bradley, T., Cupples, M. E. and Irvine, H. (2002) 'A case control study of a deprivation triangle: Teenage motherhood, poor educational achievement and unemployment', *International Journal of Adolescent Medicine and Health*, 14(2), pp. 117–123. doi: 10.1515/IJAMH.2002.14.2.117.

Brinkman, S. A. *et al.* (2016) 'Efficacy of infant simulator programmes to prevent teenage pregnancy: a school-based cluster randomised controlled trial in Western Australia', *The Lancet*, 388, pp. 2264–2271. doi: 10.1016/S0140-6736(16)30384-1.

Brittain, A. W. *et al.* (2015) 'Youth-friendly family planning services for young people: A systematic review', *American Journal of Preventive Medicine*, 49(2), pp. S73–S84. doi: 10.1016/j.amepre.2015.03.019.

Carabine, J. (2007) 'New Labour's teenage pregnancy policy: Constituting knowing responsible citizens?', *Cultural Studies*, 21(6), pp. 952–973. doi: 10.1080/09502380701470635.

Cardoza, V. J. *et al.* (2012) 'Sexual health behavior interventions for U.S. Latino adolescents: A systematic review of the literature', *Journal of Pediatric and Adolescent Gynecology*, 25(2), pp. 136–149. doi: 10.1016/j.jpag.2011.09.011.

Chang, W. *et al.* (2019) 'Shiny: Web application framework for R', R package version 1.3.2. Available at: https://cran.r-project.org/package=shiny.

Charania, M. R. *et al.* (2011) 'Efficacy of structural-level condom distribution interventions: A meta-analysis of U.S. and international studies, 1998–2007', *AIDS and Behavior*, 15(7), pp. 1283–1297. doi: 10.1007/s10461-010-9812-y.

Cheesbrough, S., Ingham, R. and Massey, D. (2002) *Reducing the rate of teenage*

conceptions: A review of the evidence (US, Canada, Australia and New Zealand). London: Health Development Agency. Available at: https://eprints.soton.ac.uk/163461/1/teenconcep_revise_v3.pdf (Accessed: 24 August 2020).

Chin, H. B. *et al.* (2012) 'The effectiveness of group-based comprehensive risk-reduction and abstinence education interventions to prevent or reduce the risk of adolescent pregnancy, human immunodeficiency virus, and sexually transmitted infections', *American Journal of Preventive Medicine*, 42(3), pp. 272–294. doi: 10.1016/j.amepre.2011.11.006.

Collumbien, M. and Macdowall, W. (2012) 'Promotion of Sexual Health', in Wellings, K., Mitchell, K., and Collumbien, M. (eds) *Sexual Health: a public health perspective*. Maidenhead, Berks: Open University Press, pp. 139–147.

Connolly, A. *et al.* (2014) 'Association between long-acting reversible contraceptive use, teenage pregnancy, and abortion rates in England', *International Journal of Women's Health*, 6, pp. 961–74. doi: 10.2147/IJWH.S64431.

Cooper, B. *et al.* (2014) 'Brief sexuality communication - A behavioural intervention to advance sexually transmitted infection/HIV prevention: A systematic review', *BJOG: An International Journal of Obstetrics and Gynaecology*, 121(s5), pp. 92–103. doi: 10.1111/1471-0528.12877.

Corcoran, J. (1999) 'Ecological factors associated with adolescent pregnancy: a review of the literature', *Adolescence*, 34(135), pp. 603–619.

Corporation, M. and Weston, S. (2019) 'doParallel: Foreach parallel adaptor for the "parallel" package', R package version 1.0.15. Available at: https://cran.r-project.org/package=doParallel.

Craig, P. *et al.* (2012) 'Using natural experiments to evaluate population health interventions: New medical research council guidance', *Journal of Epidemiology and Community Health*, 66(12), pp. 1182–1186. doi: 10.1136/jech-2011-200375.

Craig, P., Dundas, R., *et al.* (2016) 'How successful was the English teenage pregnancy strategy?', *The Lancet*, 388, p. 2604. doi: 10.1016/S0140-6736(16)32223-1.

Craig, P., Leyland, A., *et al.* (2016) 'Reply to Wellings [unpublished manuscript]', *The Lancet*, (November). Available at: https://osf.io/27c3r/.

Craig, P. *et al.* (2017) 'Natural Experiments: An overview of methods, approaches, and contributions to public health intervention research.', *Annual review of public health*, 38(1), pp. 39–56. doi: 10.1146/annurev-publhealth-031816-044327.

Craig, P. *et al.* (2019) *Developing and evaluating complex interventions: new guidance*. Medical Research Council. Available at: www.mrc.ac.uk/complexinterventionsguidance.

Dagan, N. *et al.* (2021) 'BNT162b2 mRNA Covid-19 vaccine in a nationwide mass vaccination setting', *New England Journal of Medicine*. doi: 10.1056/NEJMoa2101765.

Dahlgren, G. and Whitehead, M. (1991) *Policies and strategies to promote social equity in health*. Stockholm, Sweden: Institute for Futures Studies.

Dean, S. V. *et al.* (2014) 'Preconception care: Promoting reproductive planning', *Reproductive Health*, 11(Suppl 3). doi: 10.1186/1742-4755-11-S3-S2.

Denno, D. M., Chandra-Mouli, V. and Osman, M. (2012) 'Reaching youth with out-of-facility HIV and reproductive health services: A systematic review', *Journal of Adolescent Health*, 51(2), pp. 106–121. doi: 10.1016/j.jadohealth.2012.01.004.

DeSmet, A. *et al.* (2015) 'A systematic review and meta-analysis of interventions for sexual health promotion involving serious digital games', *Games for health journal*, 4(2), pp. 78–90. doi: 10.1089/g4h.2014.0110.

DiCenso, A. *et al.* (2002) 'Interventions to reduce unintended pregnancies among adolescents: systematic review of randomised controlled trials.', *BMJ*, 324(7351), p. 1426. doi: 10.1136/BMJ.324.7351.1426.

DiCenso, A., Guyatt, G. and Willan, A. (1999) 'A systematic review of the effectiveness of adolescent pregnancy primary prevention programs'. Ontario: Public Health Branch Ontario Ministry of Health.

Downing, J. *et al.* (2011) 'A systematic review of parent and family-based intervention effectiveness on sexual outcomes in young people', *Health Education Research*, 26(5), pp.

808–833. doi: 10.1093/her/cyr019.

Dryfoos, J. G. (1990) ‘A review of interventions to prevent pregnancy.’, *Advances in adolescent mental health*, 4, pp. 121–35. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12317625 (Accessed: 3 November 2017).

Elliott, L. *et al.* (2010) *Evaluation of healthy respect phase two: final report*. NHS Health Scotland. Available at: http://www.healthscotland.com/scotlands-health/evaluation/programme/evaluation-respect-2.aspx#Phase 2 (Accessed: 24 October 2017).

Elliott, L. *et al.* (2013) ‘Has untargeted sexual health promotion for young people reached its limit? A quasi-experimental study.’, *Journal of Epidemiology and Community Health*, 67, pp. 398–404. doi: 10.1136/jech-2012-201034.

Euro-Peristat (2018) *Country Teams - Euro-Peristat*. Available at: https://www.europeristat.com/index.php/our-network/country-teams.html (Accessed: 3 April 2019).

Falk, G. *et al.* (2006) ‘Teenage mothers - a high-risk group for new unintended pregnancies’, *Contraception*, 74(6), pp. 471–475. doi: 10.1016/j.contraception.2006.07.014.

Fleming, N. *et al.* (2015) ‘Adolescent Pregnancy Guidelines’, *Journal of Obstetrics and Gynaecology Canada*, 37(8), pp. 740–756. doi: 10.1016/S1701-2163(15)30180-8.

Fletcher, A. *et al.* (2007) ‘Interventions addressing the social determinants of teenage pregnancy’, *Health Education*, 108(1), pp. 29–39. doi: 10.1108/09654280810842111.

Forrest, S. and Wellings, K. (2012) ‘Sex education: theory and practice’, in Wellings, K., Mitchell, K., and Collumbien, M. (eds) *Sexual Health: a public health perspective*. Maidenhead, Berks: Open University Press, pp. 148–160.

Fox, J. and Weisberg, S. (2019) *car: An R Companion to Applied Regression*. 3rd edn. Thousand Oaks, CA: Sage. Available at: https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Franklin, C. *et al.* (1997) ‘Effectiveness of prevention programs for adolescent pregnancy: A meta-analysis’, *Journal of Marriage and the Family*, 59(32488), pp. 551–567. doi: 10.2307/353945.

French, R. and Mitchell, K. (2012) ‘Sexual health services’, in Wellings, K., Mitchell, K., and Collumbien, M. (eds) *Sexual Health: a public health perspective*. Maidenhead, Berks: Open University Press, pp. 161–172.

Frieden, T. R. (2010) ‘A framework for public health action: the health impact pyramid.’, *American journal of public health*, 100(4), pp. 590–595. doi: 10.2105/AJPH.2009.185652.

Fullerton, D. *et al.* (1997) ‘Preventing unintended teenage pregnancies and reducing their adverse effects’, *Quality in Health Care*, 6, pp. 102–108. doi: 10.1136/qshc.6.2.102.

Furedi, A. (1999) ‘The public health implications of the 1995 “pill scare”’, *Human Reproduction Update*, 5(6), pp. 621–626. doi: 10.1093/humupd/5.6.621.

Gavin, L. E. *et al.* (2010) ‘A review of positive youth development programs that promote adolescent sexual and reproductive health’, *Journal of Adolescent Health*, 46(3), pp. S75–S91. doi: 10.1016/j.jadohealth.2009.11.215.

Gavin, L. E. *et al.* (2015) ‘Programs to strengthen parent-adolescent communication about reproductive health: A systematic review’, *American Journal of Preventive Medicine*, 49(2), pp. S65–S72. doi: 10.1016/j.amepre.2015.03.022.

Geronimus, A. T. (1997) ‘Teenage childbearing and personal responsibility: An alternative view’, *Politial Science Quarterly*, 112(3), pp. 405–430. doi: 10.2307/2657564.

Girma, S. and Paton, D. (2015) ‘Is education the best contraception: The case of teenage pregnancy in England?’, *Social Science and Medicine*, 131, pp. 1–9. doi: 10.1016/j.socscimed.2015.02.040.

Glasgow, R. E., Vogt, T. M. and Boles, S. M. (1999) ‘Evaluating the public health impact of health promotion interventions: the RE-AIM framework’, *American journal of public health*, 89(9), pp. 1322–7. doi: 10.2105/ajph.89.9.1322.

Glasier, A. and Wellings, K. (2012) ‘Unplanned pregnancy’, in Wellings, K., Mitchell, K.,

and Collumbien, M. (eds) *Sexual Health: a public health perspective*. Maidenhead, Berks: Open University Press, pp. 33–46.

Glasziou, P. *et al.* (2007) 'When are randomised trials unnecessary? Picking signal from noise', *BMJ*, 334(7589), pp. 349–351. doi: 10.1136/bmj.39070.527986.68.

Goesling, B. *et al.* (2014) 'Programs to reduce teen pregnancy, sexually transmitted infections, and associated sexual risk behaviors: A systematic review', *Journal of Adolescent Health*, 54(5), pp. 499–507. doi: 10.1016/j.jadohealth.2013.12.004.

Gohel, D. (2020a) 'officer: Manipulation of Microsoft Word and PowerPoint documents', R package version 0.3.8. Available at: https://cran.r-project.org/package=officer.

Gohel, D. (2020b) 'rvg: R graphics devices for vector graphics output', R package version 0.2.4. Available at: https://cran.r-project.org/package=rvg.

Grunseit, A. *et al.* (1997) 'Sexuality education and young people's sexual behavior: a review of studies', *Journal of Adolescent Research*, 12(4), pp. 421–453. doi: 10.1177/0743554897124002.

Guldi, M. and Herbst, C. M. (2017) 'Offline effects of online connecting: the impact of broadband diffusion on teen fertility decisions', *Journal of Population Economics*, 30(1), pp. 69–91. doi: 10.1007/s00148-016-0605-0.

Guse, K. *et al.* (2012) 'Interventions using new digital media to improve adolescent sexual health: A systematic review', *Journal of Adolescent Health*, 51(6), pp. 535–543. doi: 10.1016/j.jadohealth.2012.03.014.

Guyatt, G. H. *et al.* (2000) 'Randomized trials versus observational studies in adolescent pregnancy prevention.', *Journal of clinical epidemiology*, 53(2), pp. 167–74. doi: 10.1016/S0895-4356(99)00160-2.

Guyatt, G. H., Oxman, A. D., Vist, G. E., *et al.* (2008) 'GRADE: an emerging consensus on rating quality of evidence and strength of recommendations', *BMJ*, 336, pp. 924–926.

Guyatt, G. H., Oxman, A. D., Kunz, R., *et al.* (2008) 'What is "quality of evidence" and why is it important to clinicians?', *BMJ*, 336, pp. 995–998.

Hadley, A. (2014) ‘Reducing teenage pregnancy: Lessons learned from the UK Government’s Teenage Pregnancy Strategy for England’, *Journal of Adolescent Health*, 54(2 SUPPL. 1), p. S35. doi: 10.1016/j.jadohealth.2013.10.083.

Hadley, A., Chandra-Mouli, V. and Ingham, R. (2016) ‘Implementing the United Kingdom Government’s 10-year Teenage Pregnancy Strategy for England (1999–2010): Applicable lessons for other countries’, *Journal of Adolescent Health*, 59, pp. 68–74. doi: 10.1016/j.jadohealth.2016.03.023.

Hadley, A., Ingham, R. and Chandra-Mouli, V. (2016) ‘Implementing the United Kingdom’s ten-year Teenage Pregnancy Strategy for England (1999-2010): How was this done and what did it achieve?’, *Reproductive Health*, 13(1), p. 139. doi: 10.1186/s12978-016-0255-4.

Harden, A. *et al.* (2006) *Young people, pregnancy and social exclusion: a systematic synthesis of research evidence to identify effective, appropriate and promising approaches for prevention and support, EPPI-Centre, Social Science Research Unit, Institute of Education, University of London*. London: The Evidence for Policy and Practice Information and Co-ordinating Centre. Available at: https://eppi.ioe.ac.uk/cms/Portals/0/PDF reviews and summaries/pregnancy_social_exclusion.pdf.

Healthcare Improvement Scotland (2021) *Draft Sexual Health Standards: March 2021*. Available at: http://www.healthcareimprovementscotland.org/our_work/standards_and_guidelines/stnds/sexual_health_standards.aspx.

Heap, K. L., Berrington, A. and Ingham, R. (2020) ‘Understanding the decline in under-18 conception rates throughout England’s local authorities between 1998 and 2017’, *Health and Place*, 66, pp. 1353–8292. doi: 10.1016/j.healthplace.2020.102467.

Henderson, M. *et al.* (2007) ‘Impact of a theoretically based sex education programme (SHARE) delivered by teachers on NHS registered conceptions and terminations: final results of cluster randomised trial.’, *BMJ (Clinical research ed.)*, 334(7585), p. 133. doi: 10.1136/bmj.39014.503692.55.

Hernán, M. (2021) *The trial's estimates for severe disease and specific age groups are imprecise. Second, an observational analysis emulates a #targettrial (an order of magnitude greater) and confirms the vaccine's effectiveness on severe disease and in different age..., [Twitter]*. Available at: https://twitter.com/_MiguelHernan/status/1364700332748640268 (Accessed: 26 February 2021).

Hernán, M. A., Hernández-Díaz, S. and Robins, J. M. (2004) 'A Structural Approach to Selection Bias', *Epidemiology*, 15(5), pp. 615–625. doi: 10.1097/01.ede.0000135174.63482.43.

Hernán, M. A. and Robins, J. M. (2020) *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC. Available at: https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/ (Accessed: 8 December 2017).

Hieftje, K. *et al.* (2013) 'Electronic media-based health interventions promoting behavior change in youth: A systematic review', *JAMA Pediatrics*, 167(6), pp. 574–580. doi: 10.1001/jamapediatrics.2013.1095.

Hill, A. B. (1965) 'The environment and disease: Association or causation', *Proceedings of the Royal Society of Medicine*, 58(5), pp. 295–300. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14283879 (Accessed: 22 January 2018).

Hodgkinson, S. *et al.* (2014) 'Addressing the mental health needs of pregnant and parenting adolescents', *Pediatrics*, 133(1), pp. 114–122. doi: 10.1542/peds.2013-0927.

Hoffman, S. D. (1998) 'Teenage childbearing is not so bad after all... Or is it? A review of the new literature', *Family Planning Perspectives*, 30(5), pp. 236–239.

House of Commons Health Committee (2009) *Health Inequalities: Third Report of Session 2008–09*. London: House of Commons.

Howick, J., Glasziou, P. and Aronson, J. K. (2009) 'The evolution of evidence hierarchies: what can Bradford Hill's "guidelines for causation" contribute?', *Journal of the Royal Society of Medicine*, 102(5), pp. 186–194. doi: 10.1258/jrsm.2009.090020.

Hoyt, H. H. and Broom, B. L. (2002) 'School-based teen pregnancy prevention programs:

a review of the literature.', *The Journal of School Nursing*, 18(1), pp. 11–17. doi: 10.1177/10598405020180010401.

Iannone, R., Cheng, J. and Schloerke, B. (2020) 'gt: Easily Create Presentation-Ready Display Tables', R package version 0.3.0. Available at: https://cran.r-project.org/package=gt.

Information Services Division Scotland (2018a) *Teenage Pregnancy: Year of Conception ending 31 December 2016*. Available at: https://www.isdscotland.org/Health-Topics/Sexual-Health/Publications/2018-07-03/2018-07-03-TeenPreg-Report.pdf.

Information Services Division Scotland (2018b) *Termination of Pregnancy - Year ending December 2017*. Available at: https://www.isdscotland.org/Health-Topics/Sexual-Health/Publications/2018-05-29/2018-05-29-Terminations-2017-Report.pdf (Accessed: 3 April 2019).

Information Services Division Scotland (2019) *Teenage Pregnancy: Year of Conception, ending 31 December 2017*. Available at: https://www.isdscotland.org/Health-Topics/Sexual-Health/Publications/2019-07-02/2019-07-02-TeenPreg-Report.pdf.

Ingham, R. (2012) 'Teenage pregnancy: the making and unmaking of a problem - Review', *Sex Education*, 12(2), pp. 254–257. doi: 10.1080/14681811.2011.609056.

Irvine, H. *et al.* (1997) 'The implications of teenage pregnancy and motherhood for primary health care: unresolved issues.', *The British Journal of General Practice*, 47(418), pp. 323–326. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9219414 (Accessed: 3 October 2019).

Johnson, B. T. *et al.* (2011) 'Interventions to reduce sexual risk for human immunodeficiency virus in adolescents: A meta-analysis of trials, 1985-2008', *Archives of Pediatrics & Adolescent Medicine*, 165(1), pp. 77–84. doi: 10.1001/archpediatrics.2010.251.

Jones, K. *et al.* (2014) 'The impact of health education transmitted via social media or text messaging on adolescent and young adult risky sexual behavior: A systematic review of the literature', *Sexually Transmitted Diseases*, 41(7). doi: 10.1097/OLQ.0000000000000146.

Kane, R. and Wellings, K. (1999) *Reducing the rate of teenage conceptions – an international review of the evidence: data from Europe*. London. Available at: https://core.ac.uk/download/pdf/17171096.pdf (Accessed: 24 August 2020).

Kassambara, A. (2020) 'ggpubr: "ggplot2" Based Publication Ready Plots', R package version 0.2.5. Available at: https://cran.r-project.org/package=ggpubr.

Kearney, M. S. and Levine, P. B. (2015a) 'Investigating recent trends in the U.S. teen birth rate', *Journal of Health Economics*, 41(2012), pp. 15–29. doi: 10.1016/j.jhealeco.2015.01.003.

Kearney, M. S. and Levine, P. B. (2015b) 'Media influences on social outcomes: The impact of MTV's 16 and pregnant on teen childbearing', *American Economic Review*, 105(12), pp. 3597–3632. doi: 10.1257/aer.20140012.

Kerr, D. L. and Matlak, K. A. (1998) 'Alcohol use and sexual risk-taking among adolescents: A review of recent literature', *Journal of HIV/AIDS Prevention and Education for Adolescents and Children*, 2(2), pp. 67–88. doi: 10.1300/J129v02n02_05.

Kim, C. R. and Free, C. (2008) 'Recent evaluations of the peer-led approach in adolescent sexual health education: a systematic review', *International Family Planning Perspectives*. 2008/07/23, 34(2), pp. 89–96. doi: 10.1363/ifpp.34.0089.08.

Kirby, D. *et al.* (1994) 'School-based programs to reduce sexual risk behaviors: A review of effectiveness', *Public Health Reports*, 109(3), pp. 339–360. Available at: https://www.jstor.org/stable/4597597.

Kirby, D. (2001) 'Emerging Answers: Research findings on programs to reduce teen pregnancy (Summary)', *American Journal of Health Education*, 32(6), pp. 348–355. doi: 10.1080/19325037.2001.10603497.

Kirby, D. (2002a) 'Effective approaches to reducing adolescent unprotected sex, pregnancy, and childbearing', *Journal of Sex Research*, 39(1), pp. 51–57. doi: 10.1080/00224490209552120.

Kirby, D. (2002b) 'The impact of schools and school programs upon adolescent sexual behavior', *Journal of Sex Research*. 2002/12/12, 39(1), pp. 27–33. doi:

10.1080/00224490209552116.

Kirby, D. (2003) ‘Primary prevention programmes are not effective in preventing adolescent pregnancies – meta-analysis; Commentary’, *Evidence-based Obstetrics and Gynecology*, 5, pp. 14–15. doi: 10.1016/S1361-259X(03)00036.

Kirby, D. and Coyle, K. (1997) ‘Youth Development Programs’, *Children and Youth Services Review*, 19(5), pp. 437–454. doi: 10.1016/S0190-7409(97)00026-1.

Kirchengast, S. (2012) ‘Teenage-pregnancies from a human life history viewpoint - an updated review with special respect to prevention strategies’, *Current Women’s Health Reviews*, 8(3), pp. 248–255. doi: 10.2174/157340412803760685.

Kontopantelis, E. *et al.* (2015) ‘Regression based quasi-experimental approach when randomisation is not an option: Interrupted time series analysis’, *BMJ (Online)*, 350. doi: 10.1136/bmj.h2750.

Kost, K., Maddow-Zimet, I. and Arpaia, A. (2017) *Pregnancies, births and abortions among adolescents and young women in the United States, 2013: National and state trends by age, race and ethnicity*. New York. Available at: https://www.guttmacher.org/report/us-adolescent-pregnancy-trends-2013 (Accessed: 17 January 2018).

L’Engle, K. L. *et al.* (2016) ‘Mobile phone interventions for adolescent sexual and reproductive health: A systematic review’, *Pediatrics*, 138(3). doi: 10.1542/peds.2016-0884.

Lawlor, D. A. and Shaw, M. (2002) ‘Too much too young? Teenage pregnancy is not a public health problem’, *International Journal of Epidemiology*, 31(3), pp. 552–554. doi: 10.1093/ije/31.3.552.

Lawlor, D. A., Shaw, M. and Johns, S. (2001) ‘Teenage pregnancy is not a public health problem.’, *BMJ (Clinical research ed.)*, 323(7326), pp. 1428–9. doi: 10.1136/BMJ.323.7326.1428.

Lazarus, J. V *et al.* (2010) ‘Systematic review of interventions to prevent the spread of sexually transmitted infections, including HIV, among young people in Europe’, *Croatian Medical Journal*, 51(1), pp. 74–84. doi: 10.3325/cmj.2010.51.74.

Letourneau, N. L., Stewart, M. J. and Barnfather, A. K. (2004) 'Adolescent mothers: Support needs, resources, and support-education interventions', *Journal of Adolescent Health*, 35(6), pp. 509–525. doi: 10.1016/S1054-139X(04)00069-2.

Lin, J. S. *et al.* (2008) 'Behavioral counseling to prevent sexually transmitted infections: A systematic review for the U.S. Preventive Services Task Force', *Annals of Internal Medicine*, 149(7), pp. 497–508. doi: 10.7326/0003-4819-149-7-200810070-00011.

Lopez Bernal, J., Cummins, S. and Gasparrini, A. (2017) 'Interrupted time series regression for the evaluation of public health interventions: a tutorial', *International Journal of Epidemiology*, 46(1), pp. 348–355. doi: 10.1093/ije/dyw098.

Lopez Bernal, J., Cummins, S. and Gasparrini, A. (2018) 'The use of controls in interrupted time series studies of public health interventions', *International Journal of Epidemiology*, 47(6), pp. 2082–2093. doi: 10.1093/ije/dyy135.

Lopez, L. M. *et al.* (2009) 'Theory-based strategies for improving contraceptive use: a systematic review', *Contraception*, 79(6), pp. 411–417. doi: 10.1016/j.contraception.2008.12.006.

Lopez, L. M., Bernholc, A., *et al.* (2016) 'School-based interventions for improving contraceptive use in adolescents', *Cochrane Database of Systematic Reviews*, 6(CD012249). doi: 10.1002/14651858.CD012249.

Lopez, L. M., Grey, T. W., *et al.* (2016) 'Theory-based interventions for contraception', *Cochrane Database of Systematic Reviews*, 11(CD007249). doi: 10.1002/14651858.CD007249.pub5.

Ma, R. (2016) 'England's teenage pregnancy strategy has been a success: now let's work on the rest', *The Lancet*, 388(10055), p. 1981. doi: 10.1016/S0140-6736(16)31856-6.

Manlove, J., Fish, H. and Moore, K. A. (2015) 'Programs to improve adolescent sexual and reproductive health in the US: a review of the evidence', *Adolescent Health, Medicine and Therapeutics*, p. 47. doi: 10.2147/AHMT.S48054.

Marseille, E. *et al.* (2018) 'Effectiveness of school-based teen pregnancy prevention programs in the USA: A systematic review and meta-analysis', *Prevention Science*, 19(4),

pp. 468–489. doi: 10.1007/s11121-017-0861-6.

Mason-Jones, A. J. *et al.* (2012) ‘A systematic review of the role of school-based healthcare in adolescent sexual, reproductive, and mental health’, *Systematic Reviews*, 1(49). doi: 10.1186/2046-4053-1-49.

Mason-Jones, A. J. *et al.* (2016) ‘School-based interventions for preventing HIV, sexually transmitted infections, and pregnancy in adolescents (Review)’, *Cochrane Database of Systematic Reviews*, 11(CD006417). doi: 10.1002/14651858.CD006417.pub3.

Matthias, K. (2002) *Youth-specific primary health care – access, utilisation and health outcomes: a critical appraisal of the literature*. Christchurch, New Zealand: New Zealand Health Technology Assessment.

Max Planck Institute for Demographic Research (2019) *Human Fertility Database*. Available at: https://www.humanfertility.org/cgi-bin/main.php (Accessed: 25 March 2019).

Mcdermott, E., Graham, H. and Hamilton, V. (2004) *Experiences of being a teenage mother in the UK: A report of a systematic review of qualitative studies*.

McLellan, J. and Dale, H. (2013) ‘Can technology be effective in interventions targeting sexual health and substance use in young people; a systematic review’, *Health and Technology*, 3(3), pp. 195–203. doi: 10.1007/s12553-013-0059-2.

McLeod, A. I. and Vingilis, E. R. (2005) ‘Power Computations for Intervention Analysis’, *Technometrics*, 47(2), pp. 174–181. doi: 10.1198/004017005000000094.

McLeod, A. I. and Vingilis, E. R. (2008) ‘Power Computations in Time Series Analyses for Traffic Safety Interventions’, *Accident; analysis and prevention*, 40(3), p. 1244. doi: 10.1016/J.AAP.2007.10.007.

Meyer, J. L., Gold, M. A. and Haggerty, C. L. (2011) ‘Advance provision of emergency contraception among adolescent and young adult women: A systematic review of literature’, *Journal of Pediatric and Adolescent Gynecology*, 24(1), pp. 2–9. doi: 10.1016/j.jpag.2010.06.002.

Michie, S., Atkins, L. and West, R. (2014) *The behaviour change wheel book - a guide to*

designing interventions. 1st edn. Silverback Publishing. Available at: http://www.behaviourchangewheel.com/online-book#1 (Accessed: 31 August 2020).

Michie, S., van Stralen, M. M. and West, R. (2011) 'The behaviour change wheel: a new method for characterising and designing behaviour change interventions.', *Implementation Science*, 6, p. 42. doi: 10.1186/1748-5908-6-42.

Microsoft and Weston, S. (2020) 'foreach: Provides foreach looping construct', R package version 1.5.0. Available at: https://cran.r-project.org/package=foreach.

Mirai Solutions GmbH (2018) 'XLConnect: Excel connector for R', R package version 0.2-15. Available at: https://cran.r-project.org/package=XLConnect.

Mitchell, K. R. *et al.* (2021) 'What is sexual wellbeing and why does it matter for public health?', *The Lancet Public Health*, 0(0). doi: 10.1016/S2468-2667(21)00099-2.

Mitchell, K., Wellings, K. and Zuurmond, M. (2012) 'Young People', in Wellings, K., Mitchell, K., and Collumbien, M. (eds) *Sexual Health: a public health perspective*. Maidenhead, Berks: Open University Press, pp. 73–84.

Munafò, M. R. and Davey Smith, G. (2018) 'Robust research needs many lines of evidence', *Nature*, 553(7689), pp. 399–401. doi: 10.1038/d41586-018-01023-3.

Murad, M. H. *et al.* (2016) 'New evidence pyramid', *Evidence-Based Medicine*, pp. 125–127. doi: 10.1136/ebmed-2016-110401.

National Records of Scotland (2018) *Births Time Series Data*. National Records of Scotland. Available at: https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/vital-events/births/births-time-series-data (Accessed: 3 April 2019).

Nelson, B. K. (1998) 'Time Series Analysis Using Autoregressive Integrated Moving Average (ARIMA) Models', *Academic Emergency Medicine*, 5(7), pp. 739–744. doi: 10.1111/J.1553-2712.1998.TB02493.X.

Nettle, D. (2011) 'Flexibility in reproductive timing in human females: integrating ultimate and proximate explanations', *Philosophical Transactions of the Royal Society*, 366, pp.

357–365. doi: 10.1098/rstb.2010.0073.

Nettle, D., Coall, D. A. and Dickins, T. E. (2009) ‘Birthweight and paternal involvement predict early reproduction in British women: Evidence from the National Child Development Study’, *American Journal of Human Biology*, 22(2). doi: 10.1002/ajhb.20970.

NHS (2018) *Miscarriage - NHS*. Available at: https://www.nhs.uk/conditions/miscarriage/ (Accessed: 26 March 2021).

Nitz, K. (1999) ‘Adolescent pregnancy prevention: A review of interventions and programs’, *Clinical Psychology Review*, 19(4), pp. 457–471. doi: 10.1016/S0272-7358(98)00087-7.

Nixon, C., Wight, D. and Henderson, M. (no date) ‘The consequences of reducing sexual health and relationships education in schools: a natural experiment’, *[submitted for publication]*.

O’Connor, E. A. *et al.* (2014) ‘Behavioral sexual risk-reduction counseling in primary care to prevent sexually transmitted infections: A systematic review for the U.S. Preventive Services task force’, *Annals of Internal Medicine*, 161(12), pp. 874–883. doi: 10.7326/M14-0475.

Oettinger, G. S. (1999) ‘The effects of sex education on teen sexual activity and teen pregnancy’, *Journal of Political Economy*, 107(3), pp. 606–644. doi: 10.1086/250073.

Ofcom (2018) *Children and parents: Children and Parents: Media Use and Attitudes Report*. London.

Ofcom (2020) *Adults’ Media Use & Attitudes report 2020*. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0031/196375/adults-media-use-and-attitudes-2020-report.pdf (Accessed: 31 August 2020).

Office for National Statistics (2017) *Conception Statistics QMI*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/conceptionandfertilityrates/methodologies/conceptionstatisticsqmi (Accessed: 5 January 2018).

Office for National Statistics (2019) *Conceptions in England and Wales: 2017*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/conceptionandfertilityrates/bulletins/conceptionstatistics/2017 (Accessed: 23 April 2019).

Organisation for Economic Co-operation and Development (2018) *Education resources - Public spending on education - OECD Data*. Available at: https://data.oecd.org/eduresource/public-spending-on-education.htm (Accessed: 21 October 2019).

Ott, M. A. and Santelli, J. S. (2007) 'Abstinence and abstinence-only education', *Current Opinion in Obstetrics and Gynecology*, 19(5), pp. 446–452. doi: 10.1097/GCO.0b013e3282efdc0b.

Paton, D. (2020) 'The mysterious fall of the teenage pregnancy', *The Spectator*. Available at: https://www.spectator.co.uk/article/the-mysterious-fall-of-the-teenage-pregnancy (Accessed: 30 October 2020).

Paton, D., Bullivant, S. and Soto, J. (2020) 'The impact of sex education mandates on teenage pregnancy: International evidence', *Health Economics*, 29(7), pp. 790–807. doi: 10.1002/hec.4021.

Paton, D. and Wright, L. (2017) 'The effect of spending cuts on teen pregnancy', *Journal of Health Economics*, 54, pp. 135–146. doi: 10.1016/j.jhealeco.2017.05.002.

Patterson, S. P., McDaid, L., *et al.* (2019) 'How men and women learn about sex: multi-generational perspectives on insufficient preparedness and prevailing gender norms in Scotland', *Sex Education*, pp. 1–16. doi: 10.1080/14681811.2019.1683534.

Patterson, S. P., Hilton, S., *et al.* (2019) 'What are the barriers and challenges faced by adolescents when searching for sexual health information on the internet? Implications for policy and practice from a qualitative study', *Sexually Transmitted Infections*, 95(6), pp. 462–467. doi: 10.1136/sextrans-2018-053710.

Pedersen, T. L. (2019) 'patchwork: The Composer of Plots', R package version 1.0.0. Available at: https://cran.r-project.org/package=patchwork.

Pedersen, T. L. and Robinson, D. (2019) ‘gganimate: A Grammar of Animated Graphics’, R package version 1.0.3. Available at: https://cran.r-project.org/package=gganimate.

Penfold, R. B. and Zhang, F. (2013) ‘Use of interrupted time series analysis in evaluating health care quality improvements’, *Academic Pediatrics*, 13(6 SUPPL.). doi: 10.1016/j.acap.2013.08.002.

Pinheiro, J. *et al.* (2019) ‘nlme: Linear and Nonlinear Mixed Effects Models’, R package version 3.1-140. Available at: https://cran.r-project.org/package=nlme.

Public Health England and Local Government Association (2018) *Teenage pregnancy prevention framework: supporting young people to prevent unplanned pregnancy and develop healthy relationships*. Available at: https://www.gov.uk/government/publications/teenage-pregnancy-prevention-framework (Accessed: 22 October 2019).

R Core Team (2020) ‘R: A Language and Environment for Statistical Computing’. Vienna, Austria. Available at: https://www.r-project.org/.

Reeves, B. C. *et al.* (2020) ‘Chapter 24: Including non-randomized studies on intervention effects’, in Higgins, J. P. T. et al. (eds) *Cochrane Handbook for Systematic Reviews of Interventions*. 6.1. Cochrane. Available at: https://training.cochrane.org/handbook/current/chapter-24 (Accessed: 24 February 2021).

Reyes, J. W. (2015) ‘Lead exposure and behavior: Effects on antisocial and risky behavior among children and adolescents’, *Economic Inquiry*, 53(3), pp. 1580–1605. doi: 10.1111/ecin.12202.

Robinson, D. and Hayes, A. (2020) ‘Broom: Convert statistical analysis objects into tidy tibbles’, R package version 0.5.5. Available at: https://cran.r-project.org/package=broom.

Rodriguez, M. I. *et al.* (2013) ‘Advance supply of emergency contraception: A systematic review’, *Contraception*, 87(5), pp. 590–601. doi: 10.1016/j.contraception.2012.09.011.

Royal Society for Public Health (2017) *Status of mind-social media and young people’s mental health and wellbeing*. London. Available at: https://www.rsph.org.uk/static/uploaded/d125b27c-0b62-41c5-a2c0155a8887cd01.pdf

(Accessed: 20 May 2021).

RSHP (2019) *Relationships, Sexual Health and Parenthood (RSHP) education: A new national teaching and learning resource for Scotland content of the new resource - Communications brief*. Available at: https://blogs.glowscotland.org.uk/re/public/renfrewshirehwb/uploads/sites/2905/2019/11/RSHP-Resource-Communications-Brief-Screen-FINAL.pdf (Accessed: 17 July 2020).

RStudio Team (2015) 'RStudio: Integrated development environment for R'. Boston, MA. Available at: http://www.rstudio.com/.

Santelli, J. *et al.* (2006) 'Abstinence and abstinence-only education: A review of U.S. policies and programs', *Journal of Adolescent Health*. doi: 10.1016/j.jadohealth.2005.10.006.

Scottish Government (2016) *Pregnancy and parenthood in young people strategy*. Available at: https://www.gov.scot/publications/pregnancy-parenthood-young-people-strategy/ (Accessed: 6 August 2020).

Sedgh, G. *et al.* (2015) 'Adolescent pregnancy, birth, and abortion rates across countries: Levels and recent trends', *Journal of Adolescent Health*, 56(2), pp. 223–230. doi: 10.1016/j.jadohealth.2014.09.007.

Shackleton, N. *et al.* (2016) 'School-based interventions going beyond health education to promote adolescent health: Systematic review of reviews', *Journal of Adolescent Health*, 58(4), pp. 382–396. doi: 10.1016/j.jadohealth.2015.12.017.

Shea, B. J. *et al.* (2017) 'AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both', *BMJ (Online)*, 358, p. 4008. doi: 10.1136/bmj.j4008.

Sievert, C. (2018) 'plotly: Plotly for R'. Available at: https://plotly-r.com.

Sipsma, H. L. *et al.* (2017) 'Impact of social service and public health spending on teenage birth rates across the USA: an ecological study', *BMJ Open*, 7(5), p. e013601. doi: 10.1136/bmjopen-2016-013601.

Skinner, S. R. and Marino, J. L. (2016) ‘England’s Teenage Pregnancy Strategy: a hard-won success’, *The Lancet*, 388(10044), pp. 538–540. doi: 10.1016/S0140-6736(16)30589-X.

Slowikowski, K. (2020) ‘ggrepel: Automatically position non-overlapping text labels with “ggplot2”’, R package version 0.8.2. Available at: https://cran.r-project.org/package=ggrepel.

Smith, G. D. and Ebrahim, S. (2001) ‘Epidemiology—is it time to call it a day?’, *International Journal of Epidemiology*, 30(1), pp. 1–11. doi: 10.1093/ije/30.1.1.

SmithBattle, L. *et al.* (2017) ‘An umbrella review of meta-analyses of interventions to improve maternal outcomes for teen mothers’, *Journal of Adolescence*, 59, pp. 97–111. doi: 10.1016/j.adolescence.2017.05.022.

Social Exclusion Unit (1999) *Teenage pregnancy: report by the Social Exclusion Unit presented to Parliament by the Prime Minister by command of Her Majesty, June 1999.* Social Exclusion Unit.

Spear, H. J. and Lock, S. (2003) ‘Qualitative research on adolescent pregnancy: A descriptive review and analysis’, *Journal of Pediatric Nursing*, 18(6), pp. 397–408. doi: 10.1016/S0882-5963(03)00160-X.

Statistics New Zealand (2019a) *Abortions by age of mother (Annual-Dec) - Table reference ABN005AA*. Available at: http://stats.gov.nz (Accessed: 19 February 2019).

Statistics New Zealand (2019b) *Live births by age of mother (Annual-Dec) - Table Reference VSB004AA*. Available at: http://stats.govt.nz (Accessed: 19 February 2019).

Steenland, M. W. *et al.* (2013) ‘The effect of follow-up visits or contacts after contraceptive initiation on method continuation and correct use’, *Contraception*, 87(5), pp. 625–630. doi: 10.1016/j.contraception.2012.09.018.

Stock, J. H. and Watson, M. W. (2020) *Introduction to Econometrics*. 4th edn. Harlow: Pearson.

Sutton, M. Y. *et al.* (2014) ‘Impact of parent-child communication interventions on sex

behaviors and cognitive outcomes for Black/African-American and Hispanic/Latino youth: A systematic review, 1988-2012', *Journal of Adolescent Health*, 54(4), pp. 369–384. doi: https://doi.org/10.1016/j.jadohealth.2013.11.004.

Swann, C. *et al.* (2003) *Teenage pregnancy and parenthood: a review of reviews: Evidence briefing*. Health Development Agency. Available at: https://www.nice.org.uk/niceMedia/documents/teenpreg_evidence_briefing_summary.pdf.

Tang, J. H. *et al.* (2012) 'Hormonal and intrauterine methods for contraception for women aged 25 years and younger', *Cochrane Database of Systematic Reviews*, 11(CD009805). doi: 10.1002/14651858.CD009805.pub2.

Teenage Pregnancy Strategy Evaluation (2005) *Final Report, Synthesis*. Available at: http://eprints.lincoln.ac.uk/id/eprint/24203/1/TPSE.pdf.

The Family Planning Association (2015) *Teenage Pregnancy - Northern Ireland Factsheet - FPA*. Available at: https://silo.tips/download/1-teenage-pregnancy-key-facts (Accessed: 13 January 2020).

The Harris Poll (2019) *LEGO Group kicks off global program to inspire the next generation of space explorers as NASA celebrates 50 years of moon landing*. Available at: https://theharrispoll.com/lego-group-kicks-off-global-program-to-inspire-the-next-generation-of-space-explorers-as-nasa-celebrates-50-years-of-moon-landing/ (Accessed: 19 October 2020).

The National Assembly for Wales (2000) *A strategic framework for promoting sexual health in Wales*. Available at: http://www.wales.nhs.uk/sitesplus/documents/888/WAGStrategicFramework.pdf (Accessed: 9 August 2021).

The World Bank Group (2019) *World Bank Open Data*. Available at: https://data.worldbank.org/ (Accessed: 23 April 2019).

Theis, D. R. Z. and White, M. (2021) 'Is obesity policy in england fit for purpose? Analysis of government strategies and policies, 1992–2020', *The Milbank Quarterly*. doi: 10.1111/1468-0009.12498.

Tolli, M. V (2012) ‘Effectiveness of peer education interventions for HIV prevention, adolescent pregnancy prevention and sexual health promotion for young people: A systematic review of European studies’, *Health Education Research*, 27(5), pp. 904–913. doi: 10.1093/her/cys055.

Toseeb, U. and Inkster, B. (2015) ‘Online social networking sites and mental health research’, *Frontiers in Psychiatry*, 6. doi: 10.3389/fpsyt.2015.00036.

Toynbee, P. (2013) *The drop in teenage pregnancies is the success story of our time, The Guardian*. Available at: https://www.theguardian.com/commentisfree/2013/dec/13/drop-teenage-pregnancies-success-story-children (Accessed: 17 October 2019).

Triggle, N. (2018) *Is social media to thank for low teen pregnancy rates? - BBC News, BBC News*. Available at: http://www.bbc.co.uk/news/health-43506784 (Accessed: 28 March 2018).

Trivedi, D. *et al.* (2007) *Update on review of reviews on teenage pregnancy and parenthood: Submitted as an Addendum to the first evidence briefing 2003*. Centre for Research in Primary and Community Care, University of Hertfordshire. Available at: https://www.nice.org.uk/guidance/ph3/evidence/update-on-review-of-teenage-pregnancy-and-parenthood-65843245.

Twenge, J. M. (2017) *Have Smartphones Destroyed a Generation?, The Atlantic*. Available at: https://www.theatlantic.com/magazine/archive/2017/09/has-the-smartphone-destroyed-a-generation/534198/ (Accessed: 1 November 2019).

UKRI (2020) *Open research*. Available at: https://www.ukri.org/our-work/supporting-healthy-research-and-innovation-culture/open-research/ (Accessed: 12 February 2021).

Underhill, K., Montgomery, P. and Operario, D. (2007) ‘Sexual abstinence only programmes to prevent HIV infection in high income countries: Systematic review’, *British Medical Journal*, 335(7613), pp. 248–252. doi: 10.1136/bmj.39245.446586.BE.

Underhill, K., Montgomery, P. and Operario, D. (2008) ‘Abstinence-plus programs for HIV infection prevention in high-income countries’, *Cochrane Database of Systematic Reviews*, 1(CD007006). doi: 10.1002/14651858.CD007006.

Underhill, K., Operario, D. and Montgomery, P. (2007a) ‘Abstinence-only programs for HIV infection prevention in high-income countries’, *Cochrane Database of Systematic Reviews*, 4(CD005421). doi: 10.1002/14651858.CD005421.pub2.

Underhill, K., Operario, D. and Montgomery, P. (2007b) ‘Systematic review of abstinence-plus HIV prevention programs in high-income countries’, *PLoS Medicine*, 4(9), pp. 1471–1485. doi: 10.1371/journal.pmed.0040275.

UNESCO Education Sector (2017) *Early and unintended pregnancy & the education sector - Evidence review and recommendations*. Paris.

University of California and Max Planck Institute for Demographic Research (2019) *The Human Mortality Database*. Available at: https://www.mortality.org/ (Accessed: 9 April 2019).

Usinger, K. M. *et al.* (2016) ‘Intrauterine contraception continuation in adolescents and young women: A systematic review’, *Journal of Pediatric and Adolescent Gynecology*, 29(6), pp. 659–667. doi: 10.1016/j.jpag.2016.06.007.

Wakhisi, A. *et al.* (2011) ‘The effectiveness of social marketing in reduction of teenage pregnancies: a review of studies in developed countries (Structured abstract)’, *Social Marketing Quarterly*, pp. 56–90. doi: 10.1080/15245004.2010.546941.

Warnes, G. R., Bolker, B. and Lumley, T. (2020) ‘gtools: Various R Programming Tools’, R package version 3.8.2. Available at: https://cran.r-project.org/package=gtools.

Wellings, K. *et al.* (2013) ‘The prevalence of unplanned pregnancy and associated factors in Britain: Findings from the third National Survey of Sexual Attitudes and Lifestyles (Natsal-3)’, *The Lancet*, 382(9907), pp. 1807–1816. doi: 10.1016/S0140-6736(13)62071-1.

Wellings, K., Palmer, M. J., Geary, R. S., *et al.* (2016) ‘Changes in conceptions in women younger than 18 years and the circumstances of young mothers in England in 2000–12: an observational study’, *The Lancet*, 388(10044), pp. 586–595. doi: 10.1016/S0140-6736(16)30449-4.

Wellings, K., Palmer, M. J., Wilkinson, P., *et al.* (2016) ‘How successful was the English teenage pregnancy strategy? – Authors’ reply’, *The Lancet*, 388(10060), pp. 2604–2605.

doi: 10.1016/S0140-6736(16)32224-3.

Wellings, K., Mitchell, K. and Collumbien, M. (2012) *Sexual health: a public health perspective, Open University Press*. Edited by K. Wellings, K. Mitchell, and M. Collumbien. Maidenhead, Berks: Open University Press.

Wenseleers, T. and Vanderaa, C. (2018) 'export: Streamlined Export of Graphs and Data Tables', R package version 0.2.2. Available at: https://cran.r-project.org/package=export.

Whitaker, R. *et al.* (2016) 'Intervention Now to Eliminate Repeat Unintended Pregnancy in Teenagers (INTERUPT): a systematic review of intervention effectiveness and cost-effectiveness, and qualitative and realist synthesis of implementation factors and user engagement', *Health Technology Assessment*, 20(16), pp. 1–214. doi: 10.3310/hta20160.

Whiting, P. *et al.* (2016) 'ROBIS: A new tool to assess risk of bias in systematic reviews was developed', *Journal of Clinical Epidemiology*, 69, pp. 225–234. doi: 10.1016/j.jclinepi.2015.06.005.

WHO Regional Office for Europe (2004) *Evidence-informed policy-making - What are the most effective strategies for reducing the rate of teenage pregnancies?* Available at: https://www.euro.who.int/en/data-and-evidence/evidence-informed-policy-making/publications/hen-summaries-of-network-members-reports/what-are-the-most-effective-strategies-for-reducing-the-rate-of-teenage-pregnancies (Accessed: 24 August 2020).

WHO Regional Office for Europe (2019) *Health for All explorer - European Health Information Gateway*. Available at: https://gateway.euro.who.int/en/hfa-explorer/ (Accessed: 9 September 2019).

Wickham, H. (2017) 'tidyverse: Easily Install and Load the "Tidyverse"', R package version 1.2.1. Available at: https://cran.r-project.org/package=tidyverse.

Wickham, H. *et al.* (2019) 'svglite: An "SVG" Graphics Device', R package version 1.2.2. Available at: https://cran.r-project.org/package=svglite.

Wickham, H. and Bryan, J. (2019) 'readxl: Read Excel Files', R package version 1.3.1. Available at: https://cran.r-project.org/package=readxl.

Widman, L. *et al.* (2018) 'Technology-based interventions to reduce sexually transmitted infections and unintended pregnancy among youth', *Journal of Adolescent Health*, 62(6), pp. 651–660. doi: 10.1016/j.jadohealth.2018.02.007.

Wight, D. *et al.* (2002) 'Limits of teacher delivered sex education: interim behavioural outcomes from randomised trial.', *BMJ (Clinical research ed.)*, 324(7351), p. 1430. doi: 10.1136/BMJ.324.7351.1430.

Wight, D. *et al.* (2015) 'Six steps in quality intervention development (6SQuID)', *Journal of Epidemiology and Community Health*, 70(5), pp. 520–525. doi: 10.1136/jech-2015-205952.

Wight, D. and Fullerton, D. (2013) 'A Review of Interventions With Parents to Promote the Sexual Health of Their Children', *Journal of Adolescent Health*, 52(1), pp. 4–27. doi: 10.1016/j.jadohealth.2012.04.014.

Wilkinson, P. *et al.* (2006) 'Teenage conceptions, abortions, and births in England, 1994-2003, and the national teenage pregnancy strategy', *The Lancet*, 368(9550), pp. 1879–1886. doi: 10.1016/S0140-6736(06)69777-8.

Williams, O. E. (2006) 'Sexual health in Wales', *Sexually Transmitted Infections*, 82(1), p. 68. doi: 10.1136/STI.2005.019331.

Wilson, T. (2014) 'Compulsory Education and Teenage Motherhood', (November). Available at: http://www.sheffield.ac.uk/polopoly_fs/1.247219!/file/E3_wilson.pdf.

World Health Organization (2020) *Adolescent pregnancy*. Available at: https://www.who.int/news-room/fact-sheets/detail/adolescent-pregnancy (Accessed: 3 December 2020).

Zapata, L. B. *et al.* (2015) 'Impact of contraceptive counseling in clinical settings: A systematic review', *American Journal of Preventive Medicine*, 49(2), pp. S31–S45. doi: 10.1016/j.amepre.2015.03.023.

Zhang, F., Wagner, A. K. and Ross-Degnan, D. (2011) 'Simulation-based power calculation for designing interrupted time series analyses of health policy interventions', *Journal of Clinical Epidemiology*, 64(11), pp. 1252–1261. doi:

10.1016/j.jclinepi.2011.02.007.

Zoritch, B., Roberts, I. and Oakley, A. (1998) ‘The health and welfare effects of day-care: A systematic review of randomised controlled trials’, *Social Science and Medicine*, 47(3), pp. 317–327.