Life Course Centre

The Australian Research Council Centre of Excellence for Children and Families over the Life Course

Phone +61 7 3346 7477 Email lcc@uq.edu.au lifecoursecentre.org.au

WORKING PAPER SERIES

DOES ANYONE SUFFER FROM TEENAGE MOTHERHOOD? MENTAL HEALTH EFFECTS OF TEEN MOTHERHOOD IN THE UK ARE SMALL AND HOMOGENEOUS

Martin O'Flaherty

The University of Queensland

Sara Kalucza

Umeå University

Joshua Bon

Queensland University of Technology

No. 2021-20

October 2021

Australian Government Australian Research Council

The University Of Queensland Australia

The University of Melbourne

The University of Western Australia

The University of Sydney

NON-TECHNICAL SUMMARY

Teen motherhood is associated with many poor long-term outcomes, including poorer mental health. Important questions remain, however, about the extent to which these relationships reflect effects of teen motherhood *per se,* as distinct from high levels of disadvantage experienced before becoming a mother. Moreover, emerging evidence suggests that economic effects of teen motherhood may be largest among the (relatively advantaged) women who are least likely to experience teen motherhood, but no previous work addresses this issue in relation to mental health outcomes.

Drawing on data from the 1970 British Cohort Study, we investigate the effects of becoming a teen mother (in comparison to motherhood aged 20-30) on the mother's mental health at ages 30, 34, and 42. We apply a flexible Bayesian estimation approach with important technical and substantive advantages. In particular, we are able to identify possible differences in the effects of teen motherhood dependent on a wide array of background characteristics that may not have been suggested by theory.

Our results show that overall effects of teen motherhood on later life mental health are consistently small across different outcome measures and ages. Furthermore, we show that effects of teen motherhood are also largely *homogeneous* – in addition to there being minimal overall effects, there are also no identifiable subgroups in the data for whom teen motherhood appears to have larger effects on mental health. This contrasts with previous findings that suggest heterogeneous effects of teen motherhood on economic outcomes.

We conclude that associations between teen motherhood and later life mental health reflect primarily selection into teen motherhood among disadvantaged adolescent girls, rather than effects of teen motherhood. This indicates that efforts to prevent teen motherhood are unlikely to yield substantial benefits to mental health. It is, however, likely valuable for services to continue targeting mental health support to teen mothers as a clearly disadvantaged group. We further suggest that our analytic method may be of broad utility to the social sciences, where analysts typically lack prior knowledge of the true relationships between variables, and heterogeneous effects are likely commonplace in many fields of study.

ABOUT THE AUTHORS

Martin O'Flaherty is a research fellow within the ARC Centre of Excellence for Children and Families over the Life Course and the Institute for Social Science Research at the University of Queensland. Martin's research focusses on understanding links between family formation processes and health- and socioeconomic-disadvantage over the life course, including work on young parents, young carers, and families of children with disabilities. He also works on children's time use and the application of machine learning methodologies to support causal inference. His work has appeared in leading journals including *Demography* and *Child Development*. Email: m.oflaherty@uq.edu.au

Sara Kalucza is a research fellow in the Department of Sociology at Umeå University. Sara's research mainly focuses on the long-term impact of young parenthood on health and socioeconomic outcomes. She is interested in how aspects such as intergenerational transmission, mental health and other resources in childhood and adolescence shapes life course pathways, and takes an extra interest in methods of sequence analysis. Her work has appeared in journals such as *Advances in Life course Research and Journal of Population Research.* Email: sara.kalucza@umu.se

Joshua Bon is a PhD candidate in the School of Mathematical Sciences at Queensland University of Technology (QUT) and is affiliated with the Centre for Data Science at QUT and the ARC Centre of Excellence in Mathematical Statistical Frontiers (ACEMS). Joshua's research focuses on computational statistics including methodological advances in Bayesian inference and designing appropriate statistical analyses for social science research. He is also interested in the provision of useful statistical software for applied researchers. His work has appeared in journals such as *Journal of the Royal Statistical Society: Series A (Statistics in Society)* and *Statistics and Computing*. Email: joshuajbon@gmail.com

Acknowledgements

The research was supported by the Australia Research Council (ARC) Centre of Excellence for Children and Families over the Life Course (CE140100027/CE20010025), and the Swedish Research Council for Health, Working Life and Welfare (FORTE, 2018-00861). Analysis uses data from the 1970 British Cohort Study, which is made available by the Centre for Longitudinal Studies (CLS), UCL Social Research Institute and the UK data service. The views reported in this study are those of the authors, and should not be attributed to the ARC, FORTE, CLS, or the UK Data Service. The authors wish to acknowledge the valuable assistance provided by Sebastian Kalucza in preparing the data for analysis.

ABSTRACT

Teen mothers experience disadvantage on a wide range of outcomes. However, previous research is equivocal with respect to possible long-term mental health consequences of teen motherhood and has not adequately considered the possibility that effects on mental health may be heterogeneous. Drawing on data from the BCS70, this paper applies a novel statistical machine-learning approach to estimate the effects of teen motherhood on mental health outcomes at ages 30, 34, and 42. We extend previous work by estimating not only sample-average effects but also individual-specific estimates. Our results show that sample-average mental health effects of teen motherhood are substantively small at all time points, apart from age 30 comparisons to women who first became mothers aged 25-30. Moreover, we show that these effects are largely homogeneous for all women in the sample – indicating that there are also no subgroups in the data that experience important detrimental consequences on their mental health. We conclude that there are no likely mental health benefits to policy and interventions aiming to prevent teen motherhood.

Keywords: Teen mothers; Mental health; Life course; Causal inference; Machine learning; Heterogeneous treatment effects; Bayesian statistics

Suggested citation: O'Flaherty, M., Kalucza, S., & Bon, J. (2021). 'Does Anyone Suffer from Teenage Motherhood? Mental Health Effects of Teen Motherhood in the UK are Small and Homogenous', Life Course Centre Working Paper Series, 2021-20. Institute for Social Science Research, The University of Queensland.

Introduction

Decades of research establishes that younger mothers are a disadvantaged group, at risk of experiencing a range of negative outcomes. A growing body of research suggests that younger mothers may be at risk of unfavourable long-term mental health outcomes, perhaps because of reduced socio-economic prospects or social stigma (Angelini & Mierau 2018; Henretta *et al* 2008; Xavier *et al* 2018). Mental health may also form an important link between age at motherhood and longer run health and labour market outcomes, where adolescent mothers has shown to have lower educational completion, lower lifetime income and increased welfare dependency (Gibb *et al* 2015; Kane *et al* 2013). Consequently, studies establishing the evolution, magnitude, and sources of worse mental health among young mothers are of considerable importance.

In this paper we analyse the relationship between early motherhood and long-run mental health outcomes using data from a nationally representative cohort of British women. In doing so, we contribute to the literature in several ways. First, previous research has generally assessed mental health outcomes at only a single time point, and it is therefore unclear how effects of early motherhood vary at different stages of the life course. We estimate the effects of young motherhood on mental health at three time points (ages 30, 34, and 42), enabling a direct comparison of effects from early adulthood through to mid-life.

Second, important concerns remain regarding the causal interpretation of differences in mental health by motherhood timing. In particular, it is clear that women who become mothers at a young age are disadvantaged before becoming parents (Mollborn & Morningstar 2009; Kalucza 2018), and as such poor outcomes reported by previous studies may reflect uncontrolled confounding, particularly because most previous studies control for only a limited set of prior confounders. Furthermore, even if the set of observed confounders is sufficient for nonparametric identification of causal effects, biases may still arise if the functional form linking confounders to outcome is misspecified (e.g. assumed linearity when the true relationship is non-linear). We address these issues by combining data on a rich set of controls

(collected prospectively from birth to adolescence) with a flexible estimation approach that allows for complex relationships between confounders, motherhood timing, and mental health outcomes.

Finally, although existing evidence is generally consistent with small negative effects of younger motherhood on mental health *on average*, it is possible that this masks subgroups in the population for whom the effects may be more substantially negative or, in some cases, potentially positive (Mollborn 2017). Indeed, recent evidence suggests that mental health consequences of younger parenthood may be less severe among earlier cohorts (Aitkin *et al* 2016; Grundy *et al* 2019; Gunes 2016) or even beneficial among teenagers with the most positive attitudes towards pregnancy (Whitworth 2017). Heterogeneity of this sort would have important implications for policy aimed at preventing youth pregnancy or supporting young parents' mental health. Our analytic approach enables us to provide *personalized* estimates of effect heterogeneity on the basis of a wide range of pre-parenthood characteristics.

Background and literature review

Theoretical perspectives on the relationship between teen motherhood and mental health

A number of theories point to possible causal effects of young motherhood, including the life course (Elder *et al* 2003; DiPrete & Eirich 2006) and stress process (Pearlin 1989) perspectives. Life course theory highlights the importance of the timing of role-transitions within normatively defined life course development. Transitions that depart from the normative ideal may attract social stigma, from one's self (as a possible violation of expectations) or other social actors. Indeed, younger mothers experience considerable social stigma (SmithBattle 2013; Yardley 2008), and stigma in turn is a major cause of health inequalities (Hatzenbeuhler *et al* 2013). Occurring at a stage of the life course when key financial and social supports have not yet been established, young motherhood may also tip the balance of parental demands relative to rewards negatively, producing role strain and potentially contributing to proliferation of other stressors in the family and across life domains (Pearlin 1989). This process of stress proliferation may in turn impact health outcomes, and the magnitude of these effects may grow over time as

consequences unfold across the life course (Pearlin *et al* 2005). For instance, if teen motherhood interferes with completion of education, this may create ongoing financial strain, leading to increasingly poor mental health.

While both life course and stress process perspectives suggest plausible mechanisms for causal effects of young motherhood, it is also clear that young motherhood is *caused by* early life disadvantage in a number of social dimensions. It is therefore possible that reported associations between early motherhood and later life mental health status may (partly or in full) represent failure to adequately control for selection into young motherhood. For instance, adverse childhood experiences (Hillis *et al* 2004), socioeconomic disadvantage (Penman-Aguilar 2013), poor adolescent mental health (Kalucza, 2018), and family instability (Fomby & Bosick 2013) are all factors that have been identified in the literature as causes of young motherhood that are also plausibly causes of later life mental health. Moreover, some (primarily qualitative) accounts of young motherhood highlight parenthood as a source of positive identity, meaning, and motivation for disadvantaged women who might otherwise not have access to normative career and family formation pathways (Brubaker & Wright 2006; Edin & Kefalas 2005). Contrary to the prevailing narrative of young motherhood as a contributory factor to poor mental health, these accounts suggest that associations between young motherhood and mental health may simply reflect background disadvantage, and young motherhood may carry positive consequences in some instances.

Long term mental health consequences of teen motherhood

Empirical evidence for long-run mental health effects of teen motherhood remains mixed. Studies that have identified detrimental effects of young motherhood on mental health in the medium and long term span a range of countries (including Australia, the UK, the US) and assess mental health outcomes at different points in the life course, from around age 30 to mid-life (Xavier *et al* 2018). Most studies use outcome measures of symptoms of depression/anxiety, however a notable exception is Kravdal *et al* (2017) who analyse purchase of antidepressant medication at ages 45-73 using Norwegian register data.

Both logistic regression and sibling fixed effects analyses of these data show substantially higher rates of antidepressant use among younger mothers. Several studies also analyse potential mediators, reporting findings that are generally in accordance with the stress process model. For instance, Grundy *et al* (2019) find a large effect of early fertility on UK women's depression in mid-life, mediated by wealth and high parity. Similarly, Falci *et al*'s (2010) analysis of parents in Minnesota showed that the effects of young motherhood on depressive symptoms in the late-twenties was largely accounted for by financial strain and low sense of personal control. Contrary examples in which no effect of young motherhood on later life mental health are also common, and tend to be those that analyse younger cohorts of women (Xavier *et al* 2017, 2018).

Heterogeneity in the effects of young motherhood

A range of direct and indirect evidence suggests that the effects of teen motherhood on mental health may be heterogeneous. First, two recent US studies investigate heterogeneity in the effects of teen motherhood on socioeconomic outcomes, finding that detrimental effects of teen motherhood may be confined to relatively advantaged segments of the population (Diaz & Fiel 2016; Gorry 2019). Diaz and Fiel (2016) apply propensity score methods to data from the child and young adult cohorts of the national longitudinal survey of youth 1979, and conclude that disadvantaged teens (who are most likely to become pregnant) experience few negative consequences, while teens from more advantaged backgrounds suffer comparatively larger reductions in education and earnings in early adulthood. Gorry (2019) examines heterogeneity in the effects of teen motherhood over socioeconomic and racial groups using an instrumental variables approach, and finds similarly that non-Hispanic whites and those from advantaged neighbourhoods are most negatively affected. While we are unaware of previous studies focussing on heterogeneity in physical health effects of teenage motherhood at the individual level, contextual evidence presented in Grundy and Foverskov (2016) indicates that countries where teen motherhood is a comparatively *normative* life course event show the weakest associations between teen parenthood and long-run physical health. Socioeconomic position and physical health may act as important mechanisms

for any effect of teen motherhood on long-run mental health, so to the extent that these effects are heterogeneous it is plausible that so too are the effects on mental health.

Second, several studies examine heterogeneity in the effects of teen motherhood on mental health directly, albeit on the basis of a limited set of factors to date. For example, in both the UK and Australia, teen motherhood has been found to have substantially larger effects on mental health among younger cohorts (Aitkin *et al* 2016; Grundy *et al* 2019; Gunes 2016), although it is unclear if this reflects age or cohort differences. In the UK context, Grundy *et al* (2019) found that the overall effect of teen parenthood on mental health status for women aged 55-64 in 2010 was five times larger than the same effect for women aged 65 or older. In Australia, Aitkin *et al* (2019) find that the largest detrimental effects of teen motherhood are concentrated among a slightly younger birth cohort (1956-65), although in relative terms the difference is not as large as reported for the UK. These differences may reflect secular changes in the nature and societal perception of adolescent fertility. In particular, as the commencement of fertility has shifted to older ages, teen fertility may be both increasingly selected with respect to background disadvantage and more subject to stigma from others.

Whitworth (2017), drawing on data from the Add Health study, investigates variability in the effects of teenage first birth on later mental health as a function of attitudes to pregnancy measured before birth. She concludes that there are minimal gaps in later mental health related to teen motherhood among those that express the most negative attitudes towards teen pregnancy, and actually better mental health among teen mothers that express the most positive attitudes. It is however, important to note that Whitworth's (2017) study relies upon a linear interaction term to estimate effects among those with positive attitudes to teen pregnancy, when the sample distribution of these attitudes is strongly skewed towards negative attitudes regardless of teen motherhood status. As discussed by Hainmueller *et al* (2019), such interactions are likely to be highly sensitive to model misspecification in many circumstances.

Methods for heterogeneous causal effects

As discussed, a small number of previous studies investigate heterogeneity in the effects of teen motherhood. However, in common with almost *all* studies of heterogeneous effects in the social sciences, these studies suffer from a number of important methodological limitations. We briefly discuss these limitations in the context of two predominant approaches to effect heterogeneity – statistical interactions on one hand, and the increasingly common propensity score approach advocated by Xie *et al* (2012) on the other. We then introduce an alternative approach drawing on recent methodological advances in the machine learning literature that overcomes many of these limitations.

Regression models including interactions between covariates and treatment represent the predominant approach to effect heterogeneity. This approach suffers from the key limitation that the analyst is presumed to know *a priori* the true functional form (including non-linearities and interactions) relating the covariates and treatment to the outcome. In practice, theory typically offers no strong guidance on these issues, and so analysts tend to fall back on linearity assumptions for both 'main effects' and interactions. These assumptions are often unreliable, and may lead to fragile and model dependent estimates of the quantities of interest – a recent reanalysis of papers using linear covariate-by-treatment interactions in leading political science journals concludes that the majority are unreliable due either to neglected non-linearity or lack of common support for the moderator between treated and untreated groups (Hainmueller *et al* 2019). While the substantive content of these papers is clearly distinct from the current context, we suspect that these kinds of issues are likely common across the social sciences. Researchers may also fit a series of models to data to try to identify an appropriate functional form – this approach will often result in better in-sample model fit, but brings a separate set of problems. In particular, once a given dataset has been used to select a model, standard errors from a final analysis conducted using that same data are no longer valid without additional corrections that are not part of standard practice (Berk *et al* 2013). Moreover, presented with a range of estimates, analysts may be

tempted to choose those which conform to prior beliefs, exceed conventional thresholds for null-hypothesis significance testing, or are otherwise more 'interesting' in some way.

An alternative approach, proposed by Xie *et al* (2012), aims to identify heterogeneity in treatment effects as a function of the propensity score, which is the probability of treatment conditional on pre-treatment covariates. Xie *et al*'s proposal incorporates a series of primarily nonparametric approaches to accomplish this task, and has proved useful in applied research (e.g. Diaz & Fiel 2016). The methods' primary limitation is that it can only identify effect heterogeneity as a function of the propensity score. In general, treatment effects may be a function of other covariates that are unrelated to the probability of treatment, or may vary in response to covariates that are both positively *and* negatively associated with treatment. Therefore, the propensity score approach to effect heterogeneity may obscure a considerable amount of variation in effects – variation that is likely of substantive interest. There is consequently a need for methods that can better identify heterogeneous causal effects as a function of multiple covariates, in situations where the functional forms relating treatment and covariates to the outcome is unknown to the analyst and may be non-linear.

In this paper, we demonstrate the application of Bayesian Additive Regression Trees (BART) (Chipman *et al* 2010; Hill 2011) to the estimation of both the familiar average treatment effect (ATE) and heterogeneous causal effects. We defer a full description of the model to the methods section below but note here that BART provides a highly flexible model fit by modelling the outcome as the sum of many small regression trees, which are regularized through the use of priors that 'shrink' the trees towards the null. The individual trees naturally incorporate non-linearities and interactions (including interactions of treatment with *any* covariate), and as a consequence the overall model inherits this property without the need for the analyst to know in advance the true functional forms. Importantly, a growing body of simulation evidence indicates that BART outperforms traditional estimation methods (e.g. linear regression, propensity score matching, or inverse probability of treatment weighting) and modern

competitors such as causal forests (Brand *et al* 2019; Wager & Athey 2018) as an estimator for both average and heterogeneous treatment effects (Dorie *et al* 2019; Hill 2011; Wendling *et al* 2018).

Data and methods

Data for the study are drawn from the 1970 British Birth Cohort Study (BCS70), an ongoing longitudinal survey with 11 waves to date. The survey takes as its subjects all those living in England, Scotland, and Wales who were born in a single week of 1970 (Elliott & Sheperd, 2006). By the 10th wave in 2016, the study participants were 46 years old. The initial sample was 17,196 individuals, and over the years survey data have been collected not only from the cohort members but also from the cohort members parents and teachers. At age 46, 8,581 individuals remained in the study which constitutes a 70% survey response rate, calculated from the 12,192 successfully traced and eligible cohort members (Morgan & Taylor, 2018: p26). The survey covers many different aspects of family circumstances, health, education, and social development as the children in the cohort move from childhood, through adolescence and into adulthood. Detailed information about the study, collection of data, response rates and more can be found in the cohort profile (Elliott & Sheperd (2006)) and the technical report (Morgan & Taylor (2018)).

Dependent and independent variables

We define 'young parenthood' as becoming a parent before the age of 20, cut-off in line with international public health statistics and a large body of previous research on young- and teenage parenthood. Age at first child was calculated by subtracting the respondents birthyear from the year of reported births, using information about live births from the BCS70 waves at ages 24 and onwards.

Two measures of mental health and wellbeing are used as outcomes. Our first measure of mental health is the Malaise inventory. The Malaise inventory was developed by Rutter *et al* (1970) and has since been applied extensively. The inventory consists of 24 'yes-no' self-completion questions which combine to measure levels of psychological distress or depression (Rutter *et al* 1970). The scale has been validated

for general population samples (Rodgers *et al* 1999) and covers emotional disturbance and associated physical symptoms, with scores ranging from 0 to 24. Persons responding 'yes' to eight or more items are considered at risk of depression (Rodgers *et al* 1999). The malaise inventory is included in the BSC70 at ages 16, 26, 30, 34 and 42, and in this study we utilize it as a dependent variable at ages 30, 34 and 42. Scores are dichotomized using cut-offs 9+ when using the full item scale included in the BSC70 at age 30, and 4+ for the 9-item scale included int the BSC70 at ages 34 and 42.

In addition to the malaise inventory, we use the Warwick-Edinburgh Mental Well-being Scale (WEMWBS) at age 42 (Tennant *et al* 2007). WEMWBS is a 14-item scale of mental well-being covering subjective well-being and psychological functioning including positive affect (feelings of optimism, cheerfulness, relaxation), satisfying interpersonal relationships and positive functioning (energy, clear thinking, self-acceptance, personal development, competence and autonomy) (Tennant *et al* 2007). The scale is scored by summing responses to each item answered on a 1 to 5 Likert scale, with a minimum score of 14 and a maximum of 70. Higher WEMWBS scores indicate better mental wellbeing. The WEMWBS has the advantage of being free of ceiling effects in population samples, and is thus more sensitive to variation in mental wellbeing among the non-clinical population. Our analytical sample had a mean WEMWBS score of 48.7 with a 95% confidence interval of 48.4 to 49.1. This is about 1 unit lower than the provisional Scottish population mean score of 50.7 with a 95% confidence interval of 50.3 to 51.1 (Stewart-Brown & Janmohamed, 2008).

Covariates

The BSC70 includes a rich set of background covariates. To select covariates for inclusion, we first identified topic domains we wished our covariates to cover based on previous research and theory. The 13 topic domains included delinquency, education attainment and aspirations, family history and family stability, health behaviour, family housing situation, mental health in childhood, orientation to the world, parent parenting strategy, peer and peer characteristics, physical health, parents' physical health, parent

relationship, and parent- and grandparent education and social class. We selected 70 covariates of interest with acceptable levels of missing data, measured from ages 0 to 16, sorted by topic domain. An overview of variables and domains can be found in appendix A1, and descriptive statistics in appendix A2.

Missing data

Missing data for background covariates was imputed using multiple imputation by chained estimation in the 'mice' R package (van Buuren 2012). Cases with missing outcomes or fertility data were excluded from analysis. A total of 100 imputed datasets were created, and our models are estimated separately for each imputed dataset. The first 1000 MCMC iterations are discarded as 'burn-in' for each model, and 1000 posterior samples are retained. Final estimates are constructed by pooling the retained posterior samples from each imputed dataset (Gelman *et al.* 2014), for a total of 100,000 posterior samples.

BART

BART (Chipman *et al* 2010) is a tree-based regression model notable for flexibility and parsimonious modelling of many variables. BART can capture high-order variable interactions and account for uncertainty by representing model parameters probabilistically in the Bayesian framework. This section will first describe the general structure of BART, in relation to standard regression, followed by a description of the tree structure employed by BART.

BART models generally have the same structure as other regression models, that is, for continuous outcome Y, with covariates $x = [x_1, x_2, \ldots, x_p]$ the overall regression model is described by:

$$Y = f(x) + \epsilon$$

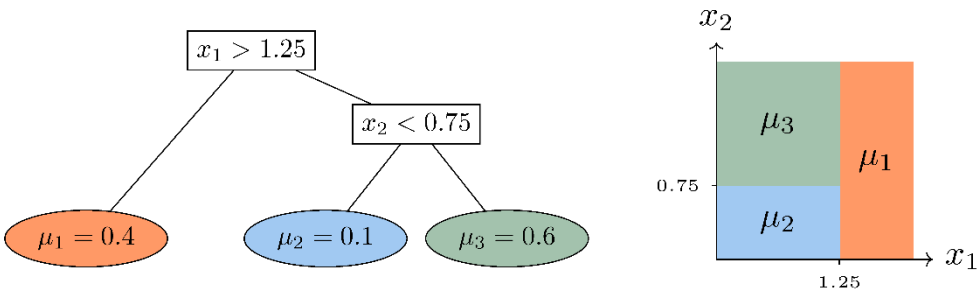
where $f(x)$ is a model of the mean relation between the outcome and covariates, and ϵ is zero-mean normally distributed noise with variance σ^2. In the case of BART, the function f is a sum of many regression trees (see below for details). In comparison, in a standard linear regression f is simply the linear function $f(x) = x\beta$. The extension of BART models to non-continuous outcomes, such as binary responses, is analogous to the extension linear models to generalised linear models (Nelder & Wedderburn, 1972) in that f is connected to the mean of the outcome through a link equation:

$$E(Y) = h^{-1}[f(x)]$$

where h is the link function. We use the probit link for h for the binary Malaise outcomes.

BART assumes a sum of trees structure for the mean model f. Each tree is a regression tree -- a binary tree T consisting of successive nodes of decision rules, and a layer of terminal nodes with mean values $M = \{\mu_1, \mu_2, \ldots, \mu_m\}$ where m is the number of terminal nodes. Each tree recursively partitions the covariates x into several subgroups with a particular mean value. This process is denoted by the function $g(x; T, M)$ which takes inputs of the covariates, tree structure with decisions, and mean values at terminal nodes. An illustrative example of a single tree is given in Figure 1.

Figure 1: Decision tree example



A single decision tree (left) with three terminal nodes, μ_1, μ_2 and μ_3. Left branches indicate the preceding nodes' condition is true, while right branches indicate they are false. The tree function $f(x_1, x_2)$ assigns values to the parameter space based on the tree. For this tree, any (x_1, x_2) with $x_1 > 1.25$ will have $f(x_1, x_2) = 0.4$. The nodes μ_2, μ_3 are interaction effects as they depend on both x_1 and x_2. A second representation of $f(x_1, x_2)$ is given on the right.

Rather than a single large tree, BART uses the sum of many smaller trees (commonly 200) to model f which provides more flexibility. Each tree is constrained by priors to be a weak learner, which retains flexibility but penalises overfitting (Chipman, *et al* 2010). Such a sum of trees structure is described mathematically as:

$$f(x) = \sum_{k=1}^{K} g(x; T_k, M_k)$$

where K is the number of trees.

BART uses the Bayesian paradigm to perform regression on a given dataset using the sum of tree structure. The model is fit using Markov chain Monte Carlo (MCMC) as described in Chipman et al. (2010). Model summaries and fitted values are defined in terms of the Bayesian posterior samples from the MCMC procedure (Gelman *et al* 2014). We used the 'BART' R package (McCulloch *et al* 2021) to estimate all models.

BART for estimation of causal effects

Causal inference with BART relies on assumptions of unconfoundedness and positivity (Hill 2011). Unconfoundedness ($Y(0), Y(1) \perp\!\!\!\perp A \mid X$) stipulates that (binary) treatment A is unrelated to the potential outcomes ($Y(0)$, $Y(1)$), conditional on the observed confounders X. Positivity ($0 < p(A = 1 \mid X) < 1$) requires that there is a non-zero probability of receiving every level of the treatment for all values of the confounders. Given these assumptions, the conditional average effect of treatment (CATE) for subjects with observed confounder values $X = x$ is $\mathrm{E}(Y \mid A = 1, X = x) - \mathrm{E}(Y \mid A = 0, X = x)$ (Hill 2011). Estimation of causal effects therefore resolves into the problem of estimating the response surfaces under treatment and no-treatment.

Sample average effects for any subgroup of interest (including the common Average Treatment Effect (ATE) or Average Treatment Effect on the Treated (ATT) may be calculated by averaging the CATE over subjects in the relevant group. For example, the ATT is calculated as the average of the CATE for women in the sample who became young mothers. Uncertainty is quantified through variation in estimates over MCMC iterations. Specifically, we report 95% credible intervals based on a normal approximation, based on evidence in Carnegie (2019) that this type of credible interval exhibits better nominal coverage. In keeping with best-practice reporting advice (Amrhein *et al* 2019), we discuss throughout the range of parameter values that are compatible with our estimates (high and low) in addition to point estimates.

Results

Table 1 shows sample descriptive statistics for age at first birth and the four outcome variables. For each of the outcomes, it is apparent that younger mothers experience notably worse outcomes. In the case of the three Malaise items, young parents are roughly ten to twelve percentage points more likely to experience elevated levels of distress than women who became mothers at an older age at each time point. For the Warwick-Edinburgh measure of mental wellbeing, young parents average 2.1 points lower,

Table 1: Sample summary statistics by outcome and young motherhood

	Normative age mothers (20-30)	Young mothers (<20)	Total	Unadjusted difference
Malaise inventory (age 30)				
Low (0-7)	1,993 (87.2%)	366 (75.9%)	2,359 (85.3%)	
High (8-24)	292 (12.8%)	116 (24.1%)	408 (14.8%)	11.3%
Total	2,285	482	2,767	
Malaise inventory (age 34)				
Low (0-3)	1,729 (81.6%)	292 (69.7%)	2,021 (79.6%)	
High (4-9)	390 (18.4%)	127 (30.3%)	517 (20.4%)	11.7%
Total	2,119	419	2,538	
Malaise inventory (age 42)				
Low (0-3)	1,837 (80.3%)	342 (70%)	2,179 (78.5%)	
High (4-9)	451 (19.7%)	147 (30%)	598 (21.5%)	10.3%
Total	2,288	489	2,777	
Warwick-Edinburgh (age 42)				
Mean (SD)	49.1 (8.5)	47.0 (9.5)	48.8 (8.7)	-2.1
Total	2,132	446	2,578	

equivalent to roughly .23 of the sample standard deviation. Sample sizes differ depending on the outcome, ranging from 2,538 to 2,777.

Figure 2: Propensity score overlap by young motherhood status

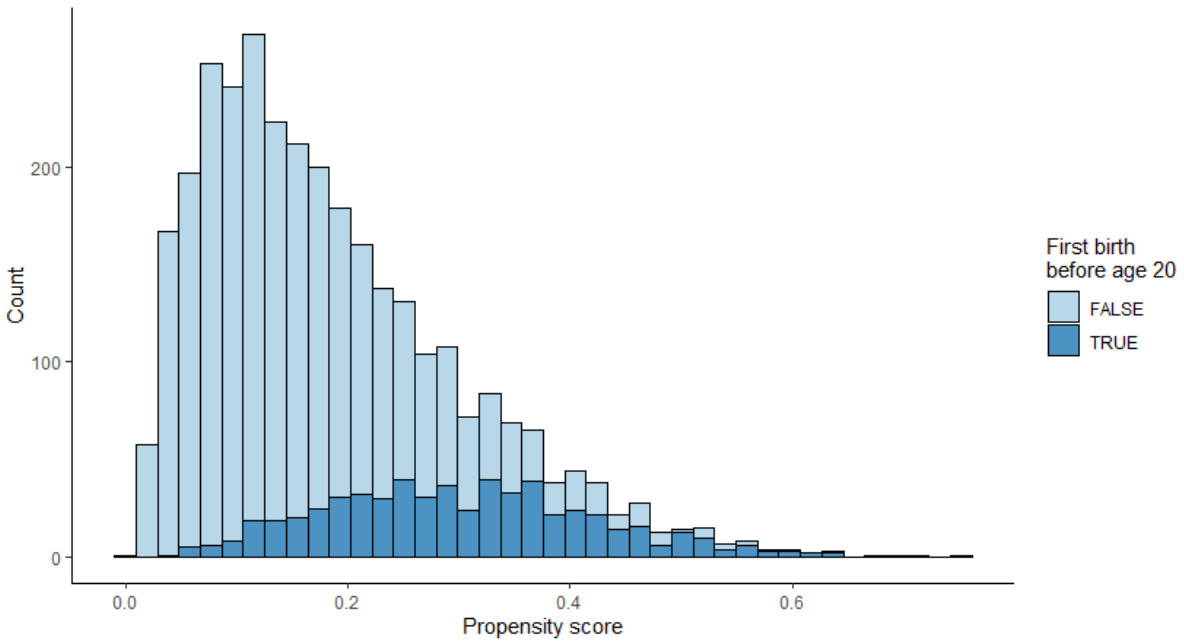


Table 2: Average effects of young motherhood by outcome

Outcome variable	ATE point estimate (95% credible interval)
Malaise (age 30)	3.6 (0.01, 7.2)
Malaise (age 34)	3.4 (-0.9, 7.7)
Malaise (age 42)	3.1 (-1.1, 7.3)
Warwick (age 42)	-0.3 (-1.2/0.5)

** denotes 95% credible intervals that exclude zero.*

To check the positivity assumption, we plot the estimated propensity score separately for young mothers and older mothers in Figure 2. The plot shows sufficient overlap, albeit with potential problems at very low propensity scores.

Table 2 presents the estimated sample average treatment effect (ATE) for each of the mental health outcome measures: dichotomous indicators of depressive symptoms based on the Malaise inventory at ages 30, 34, and 42, and the Warwick scale at age 42. Malaise outcomes are modelled using the probit link and the Warwick scale is modelled as continuous. All models control for the full set of covariates described in appendices A1 and A2. Estimates for the binary malaise outcomes are expressed as percentage point differences and estimates for the continuous Warwick outcome are expressed in raw (unstandardized) units. In all cases, our estimates show that the average treatment effects are small in magnitude. We find weak evidence of an increased risk of poor mental health for young mothers at age 30 (3.6 percentage points (CI: 0.01-7.2 percentage points)) adjusting for controls, although the lower bound of the credible interval is very close to zero. Otherwise, although the upper bounds of the credible intervals are consistent with small detrimental effects, the 95% credible interval overlaps zero for all dependent variables, indicating a slight possibility that the true effects of young motherhood on later life mental health are in fact protective. Moreover, this indicates that our analysis does not provide any strong evidence in support of deleterious average causal effects of young motherhood beyond age 30, and we note a slight decline in the magnitude of point estimates for the Malaise outcomes at older ages. With respect to the Warwick-Edinburgh scores, we note that the 95% credible interval is consistent with, *at most*, a quite small detrimental effect equal to less than one-sixth of a standard deviation (-1.2). The point

estimate for the ATE of young parenthood on the Warwick-Edinburgh scores corresponds to an effect of less than 4% of a standard deviation. In contrast, for the Malaise outcomes the upper bounds of the credible intervals correspond to an increase in the risk of psychological distress of roughly seven to eight percentage points. Given the base prevalence of these outcomes ranges from 15-22%, an increase of seven to eight percentage points would represent an important effect. Overall, the credible intervals strongly suggest very small effects of young motherhood on the Warwick-Edinburgh scale, while leaving open the (statistical) possibility of substantively meaningful detrimental effects on the Malaise outcomes.

Sensitivity analysis for average effects

The choice of cut-offs for 'young' motherhood is admittedly arbitrary, although consistent with previous literature on the topic. To address this issue, we conducted a series of sensitivity analyses. First, we varied the definition of 'young motherhood' to be either under 18 (compared to 19-30) or under 22 (compared to 23-30). Results from this analysis are shown in table 3. In all cases, we arrive at conclusions that are substantively similar to the main analysis, although the credible intervals are notably wider for the <18 vs. 19-30 comparison owing to the small number of first births occurring prior to age 18. The only notable difference in comparison to the analyses reported in table 2 is that the ATE for <18 vs. 19-30 at age 30 overlaps zero (although the point estimate is in fact larger than the corresponding value for the <22 vs 23-30 comparison).

Table 3: Average effects of young motherhood using alternative cut-points

Outcome variable	ATE point estimate (95% credible interval)	
	<18 vs. 19-30	*<22 vs. 23-30*
Malaise (age 30)	5.1 (-0.3, 10.4)	4.7 (1.8, 7.6) *
Malaise (age 34)	0.2 (-4.3, 4.6)	3.3 (-0.3, 6.9)
Malaise (age 42)	4.9 (-1.4, 11.1)	2.8 (-0.7, 6.3)
Warwick (age 42)	-0.1 (-1.3, 1.0)	-0.4 (-1.2, 0.4)

** denotes 95% credible intervals that exclude zero.*

Second, in the main analysis we assume that variation in ages at first birth within the 'young' and 'normative' groups is inconsequential for mental health. For the 'normative' group in particular, this encompasses a wide age range (20-30), during which time partnerships and human capital are typically established. We therefore conducted a series of additional analyses using more restrictive definitions of the 'normative' comparison group, as either 20-24 or 25-30. The results from these analyses are shown in table 4. In substantive terms, we find larger (roughly double) effects for the analyses of Malaise outcomes that use 25-30 as the comparison group, relative to analyses that use 20-24 as the comparison group. The credible interval for the ATE for motherhood before age 20 vs. motherhood at ages 25-30 on Malaise at age 30 excludes zero, whereas the corresponding credible interval for the <20 vs. 20-24 comparison does not. We note, however, that the substantial overlap between the credible intervals for these analyses, indicating no clear evidence that longer delays in motherhood would yield larger benefits in mental health.

Table 4: Average effects of young motherhood using alternative comparison groups

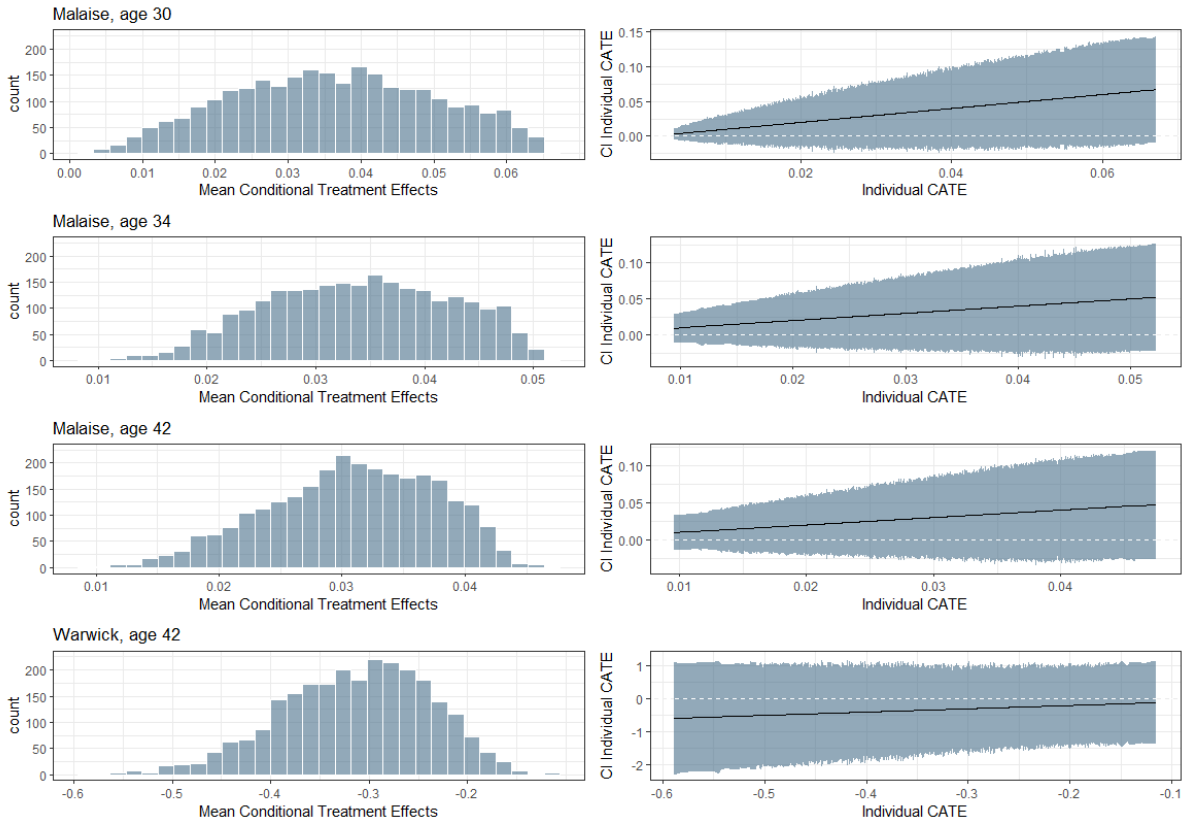
Outcome variable	ATE point estimate (95% credible interval)	
	<20 vs. 20-24	*<20 vs. 25-30*
Malaise (age 30)	2.0 (-1.9, 6.0)	5.0 (1.2, 8.8) *
Malaise (age 34)	1.9 (-2.5, 6.4)	4.5 (-0.2, 9.3)
Malaise (age 42)	2.3 (-2.2, 6.8)	4.1 (-0.6, 8.8)
Warwick (age 42)	-0.1 (-1.0, 0.7)	-0.8 (-1.9, 0.3)

** denotes 95% credible intervals that exclude zero.*

Heterogeneous effects

We next present evidence regarding possible *heterogeneity* in the effects of young motherhood across individuals. As discussed, it is possible that small average effects may conceal variation across groups, with some experiencing more detrimental consequences (and some potentially positive effects). In practice however, we find little evidence to support this contention. The left-hand panel of figure 3 plots the distribution of the individual CATE estimate on the four different outcomes across sample members. Point estimates for the Warwick-Edinburgh CATE are concentrated between roughly -0.5 and -0.1, and

Figure 3: Individual CATE estimates by outcome



Average individual CATE estimates based on 100,000 posterior samples, and 95% credible intervals (CI).

are centered close to the average treatment effect estimate of -0.3. Substantively, the largest individual CATE (-0.6) point estimate is equivalent to an effect of slightly less than seven percent of a standard deviation – still practically very small. Moreover, the variance in the individual CATEs is dwarfed by comparison with the degree of statistical uncertainty associated with the estimates – 95% credible intervals comfortably include zero for all cases. Thus, our analysis suggests that, in addition to no average effect of young motherhood (compared to the counterfactual of motherhood from age 20-30) on mental wellbeing at age 42, there are no *subgroups* in the data for whom a meaningful effect is likely to be present. We further investigated common alternative estimands (table 5), including the average treatment effect on the treated (ATT) and the average treatment effect on the controls (ATC). These estimates were substantively identical to the main analysis. Last, we calculated the ATE within quintiles of the

Table 5: Average effects of young motherhood for young mothers and normative age mothers

Outcome variable	ATT (95% credible interval)	ATC (95% credible interval)
	Young mothers	*Normative age mothers*
Malaise (age 30)	4.3 (0.2, 8.4) *	3.5 (0.03, 6.9) *
Malaise (age 34)	3.9 (-1.0, 8.7)	3.3 (-0.9, 7.5)
Malaise (age 42)	3.3 (-1.2, 8.0)	3.0 (-1.1, 7.1)
Warwick (age 42)	-0.3 (-1.1, 0.5)	-0.3 (-1.2, 0.5)

** denotes 95% credible intervals that exclude zero.*

Table 6: Average effects of young motherhood by propensity score quintiles

Outcome variable	ATE point estimate (95% credible interval)
Malaise (age 30)	
Propensity score quintile	
1(lowest)	2.2 (-0.2, 4.6)
2	3.2 (-0.1, 6.4)
3	3.7 (0.03, 7.3) *
4	4.3 (0.1, 8.4) *
5 (highest)	5.0 (0.2, 9.9) *
Malaise (age 34)	
Propensity score quintile	
1(lowest)	2.6 (-0.9, 6.1)
2	3.1 (-0.9, 7.2)
3	3.4 (-1.0, 7.9)
4	3.8 (-1.0, 8.6)
5 (highest)	4.3 (-1.1, 9.7)
Malaise (age 42)	
Propensity score quintile	
1(lowest)	2.5 (-1.0, 6.0)
2	2.9 (-1.1, 6.9)
3	3.2 (-1.1, 7.5)
4	3.3 (-1.2, 7.9)
5 (highest)	3.6 (-1.3, 8.6)
Warwick (age 42)	
Propensity score quintile	
1(lowest)	-0.4 (-1.3, 0.6)
2	-0.3 (-1.2, 0.6)
3	-0.3 (-1.2, 0.5)
4	-0.3 (-1.1, 0.5)
5 (highest)	-0.3 (-1.1, 0.6)

** denotes 95% credible intervals that exclude zero.*

propensity score, finding (table 6) no variation in the magnitude of effects of young motherhood on WEMWBS scores.

Individual CATE estimates for the Malaise outcomes vary between a 0 to 6 percentage point increase in risk of psychological distress at age 30, and 1 to 5 percentage points at ages 34 and 42. While the magnitude of this variation appears substantively important, we do not believe that this represents evidence of effect heterogeneity, for two reasons. First, the uncertainty of the estimates dwarfs the variation in the individual CATEs. Second, there is negligible heterogeneity in effects on the underlying probit scale. This means that the apparent variation in the magnitude of effects (expressed as percentage points) reflects variation in the underlying marginal probability of experiencing poor mental health, rather than the presence of interactions between young motherhood and covariates within $f(x)$. In practice, this means that young motherhood has larger effects in percentage point terms only for women who experience an elevated risk of poor mental health as a function of other background characteristics. Subgroup analyses by young motherhood status (ATT/ATC) and by quintiles of the propensity score reflect this fact, with relatively larger percentage-point effects of young motherhood among young mothers, and among women who had a higher propensity to experience young motherhood. Otherwise, there are minimal substantive differences in the findings for these subgroups compared to the main analysis of Malaise outcomes.

If analysis had identified meaningful variation in effects across individuals, we would ordinarily proceed to explore subgroups whose mental health appears to be more or less strongly affected by young motherhood. Because our analysis found no evidence of individual effect heterogeneity for the WEMWBS outcome, nor (on the probit scale) for the Malaise outcomes, we did not proceed to this analysis. Rather, the lack of effect heterogeneity suggests that any effects of young motherhood on mental health are substantially *homogeneous*, at least as a function of the (extensive) set of background covariates included in our analysis.

Discussion and conclusion

A long history of research and public commentary links young motherhood with a host of negative outcomes for both mothers and their children (Mollborn 2017). However, as data and methods have become more sophisticated, it seems increasingly likely that *causal* effects of young motherhood are small or non-existent, or in some cases confined to relatively advantaged segments of the population that are unlikely to experience young motherhood in any case (Diaz and Fiel 2016; Gorry 2019). Our results largely confirm the narrative that poorer outcomes experienced by young mothers are primarily due to the high level of background disadvantage experienced by young mothers, rather than detrimental effects of young motherhood *per se*. In our primary analyses, the point estimate for the ATE was roughly seventy percent smaller than the bivariate differences for malaise outcomes at ages 30, 34, and 42, and eighty-five percent smaller for the Warwick-Edinburgh measure of mental wellbeing at age 42. Except for the Malaise outcome at age 30, 95% credible intervals for the ATE covered zero, meaning that in comparison to those who became mothers aged 20 or older, our analyses provide only very limited evidence of harmful causal effects of young motherhood on later mental health. While our estimates indicate some support for an increased rate of mental health problems among young mothers at age 30, the effect is substantively quite small and not distinguishable from zero at older ages.

Our analysis does, however, provide some limited evidence that effects of young motherhood on mental health may depend on the counterfactual state that 'young motherhood' is compared to, and the life stage of the woman. Specifically, sensitivity analysis of malaise outcomes at age 30 in which teen mothers were compared to those who became mothers aged 25-30 (excluding mothers at 20-24) indicated a five percentage-point increase in the risk of poor mental health. With the major caveat that causal interpretation of our estimates remains subject to the strong assumption that all relevant confounders have been accounted for, this suggests that there may be some benefit to delaying motherhood into the mid to late 20s in the short term. At older ages, estimated ATEs for the same contrast, although positive and only

marginally smaller in magnitude, were not distinguishable from zero. Estimates for the contrast between teen motherhood and motherhood aged 20-24 were uniformly not distinguishable from zero.

An important limitation of many common methods is that they provide only average effects and leave open the possibility that important variation in effects may exist between subgroups in the data. Effect heterogeneity of this kind has the potential to be both illuminating theoretically and important for policy (e.g. with respect to targeting interventions), and there is therefore considerable value in methods that can properly identify heterogeneous effects. In practice, our analyses find little evidence of effect heterogeneity, as evidenced by the largely homogeneous individual CATE estimates. We note that this represents an important finding – in addition to there being minimal evidence of sample average effects of teen motherhood, our analysis further suggests that there are *no subgroups for which effects can be reliably identified*. The fact that our method allows us to arrive at a general conclusion of this nature contrasts with many alternative methods, which would permit only consideration of a limited number of pre-specified subgroups, and commonly require (unrealistic) prior knowledge of the correct functional forms.

We note that there exist alternative approaches (similar in spirit to BART) to estimation of heterogeneous causal effects – notably the 'causal forest' approach of Wager & Athey (2018; see also Brand *et al* 2019). As there are generally no reasonable grounds to believe that effects in social science are truly homogeneous, methods of this kind have the potential to furnish valuable new insights in many substantive areas, particularly with the advent of larger linked data that may support reliable inferences within smaller subgroups. To date, however, evidence suggests that BART-based methods tend to outperform alternative methods (Dorie *et al* 2019; Hill 2011; Wendling *et al* 2018). This is particularly true with respect to estimation of heterogeneous effects, but applies also to the more common sample average estimands appearing in the literature. We therefore suggest that BART has considerable potential as a tool for social science.

Our findings have several implications for policy. With respect to efforts to *delay* motherhood, our findings suggest that policy would need to achieve relatively large changes in women's birth timing (in the order of 6 years at minimum) in order to realize any meaningful benefits to mental health. As the extant literature shows small or null effects of a range of teen pregnancy prevention strategies (Baxter *et al* 2021; Marseille *et al* 2018), it seems unlikely that intervention can achieve delays of this magnitude in practice. It is consequently unclear that there exist viable teen motherhood prevention strategies that would be reasonably expected to translate into better mental health outcomes. Furthermore, after age 30 we find no effects of teen motherhood that can be reliably distinguished from zero, suggesting that prevention efforts are unlikely to achieve substantial long-term gains in mental health. We note however, that teen motherhood remains a strong (and visible) *marker* of disadvantage even in the absence of causal effects, and that motherhood brings specific needs and constraints related to caring for children. There is therefore continuing potential for younger motherhood status to be used as a mechanism for targeting services aimed at ameliorating mental health, particularly as pregnancy and early childhood is a life stage during which young mothers may be more engaged with and accessible to health and social services. Moreover, our finding of an increased rate of poor mental health among teen mothers (compared to motherhood aged 25-30) at age 30 suggests that there may be scope for intervention after birth to mitigate any potential negative short-term effects on later mental health. On this point, we stress the need for future work to consider more carefully the mechanisms associated with any detrimental effect of teen motherhood – there is a tendency in much of the literature to equate negative effects of teen motherhood with some sort of deficit in the mother herself. This ignores the evidence that teen mothers are routinely stigmatized (McArthur & Winkworth, 2018; Yardley 2008), and that the experience of stigma *per se* is strongly linked to poorer health outcomes (Hatzenbuehler *et al* 2013). In short, interventions aiming to change public *perceptions* of teen pregnancy – in lieu of targeting perceived deficits in the mother – may ameliorate any negative effects of teen pregnancy on mental health.

As with all analysis, our work is subject to limitations. First, we note that the Malaise outcomes are geared towards clinical mental health symptoms and are consequently less sensitive to non-clinical

variation in mental health. This may have limited the ability of our analyses to identify effects of young motherhood on aspects of mental health that do not amount to clinical concerns. The WEMWBS scale does, however, capture sub-clinical variation in mental wellbeing and results were largely consistent with the Malaise outcomes. Second, many covariates of potential importance to young motherhood (particularly from the age 16 data collection) were unable to be included in analysis due to very high levels of missing data. While we were able to control for an extremely rich set of potential confounders in comparison with other studies, it is nevertheless likely that some residual confounding exists. This implies also that our analysis cannot rule out the possibility that there are (unmeasured) subgroups of women for whom effects of teen motherhood are larger or smaller.

Our work builds upon a long history of studies that have investigated the consequences of teen motherhood (Mollborn 2017), and showcases the application of BART as a tool for the estimation of both common causal estimands as well as heterogeneous causal effects (Hill 2011). In most settings of interest to social scientists, there is no strong rationale to believe *a priori* that causal effects are truly homogeneous, and no strong theory to guide model specification – BART addresses both issues, and we therefore suggest that there is considerable potential for social science to benefit from BART or similar methods. Substantively, our findings indicate that causal effects of teen motherhood on later mental health are likely to be both small and homogeneous for the cohort of women we studied. We note however that the absence of effects in this cohort (now middle aged) does not rule out the possibility that such effects may arise for younger cohorts of women. In this light, we stress the need for both future research and policy aimed at preventing teen motherhood to tread lightly to avoid further reinforcing negative stereotypes or low expectations of young mothers – to the extent that we perpetuate such stigma, we risk *creating* the problem we purport to solve.

References

Aitken, Z., Hewitt, B., Keogh, L., LaMontagne, A., Bentley, R., & Kavanagh, A. (2016) Young maternal age at first birth and mental health in later life: Does the association vary by birth cohort? *Social Science and Medicine*, 157: 9-17.

Al-Sahab, B., Heifetz, M., Tamim, H., Bohr, Y., & Connolly, J. (2012). Prevalence and Characteristics of Teen Motherhood in Canada. *Maternal and Child Health Journal*, 16(1), 228–234.

Amrhein, V., Greenland, S., & McShane, B. (2019) Scientists rise up against statistical significance, *Nature*, 567(7748): 305-307.

Angelini, V., & Mierau, J. (2018). Late-life health effects of teenage motherhood. *Demographic Research*, 39(41), 1081–1104. https://doi.org/10.4054/DemRes.2018.39.41

Baxter, A., Dundas, R., Popham, F., & Craig, P. (2021) How effective was England's teenage pregnancy prevention strategy? A comparative analysis, *Social Science and Medicine*, 270: 113685

Berk, R., Brown, L., Buja, A., Zhang, K., & Zhao, L. (2013) Valid post-selection inference, *Annals of Statistics*, 41(2): 802-837.

Brand, J., Xu, J., Koch, B., Geraldo, P. (2019) Uncovering sociological effect heterogeneity using machine learning, https://arxiv.org/ftp/arxiv/papers/1909/1909.09138.pdf

Brubaker, S. J., & Wright, C. (2006). Identity transformation and family caregiving: Narratives of African American teen mothers. *Journal of Marriage and Family*, 68, 1214–1228.

Carnegie, N. (2019) Comment: Contributions of model features to BART causal inference performance using ACIC 2016 competition data, *Statistical Science*, 34(1): 90-93.

Chipman, H., George, E., & McCulloch, R. (2010) BART: Bayesian Additive Regression Trees, *The Annals of Applied Statistics*, 4(1): 266-298.

Diaz, C. & Fiel, J. (2016) The effect(s) of teen pregnancy: reconciling theory, methods, and findings, *Demography*, 53: 85-116.

DiPrete, T. A., & Eirich, G. M. (2006). Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments. Annual Review of Sociology, 32(1), 271–297. https://doi.org/10.1146/annurev.soc.32.061604.123127

Dorie, V., Hill, J., Shalit, U., Scott, M., & Cervone, D. (2019) Automated versus do-it-yourself methods for causal inference: lessons learned from a data analysis competition, *Statistical Science*, 34(1): 43-68.

Edin, K., & Kefalas, M. (2005). *Promises I can keep: Why poor women put motherhood before marriage*. Berkeley: University of California.

Elder, G., Johnson, M., & Crosnoe, R. (2003) The emergence and development of life course theory, 3-19 in Mortimer, J. & Shanahan, M. (eds.), *Handbook of the Life Course*, Springer

Elliott, J., & Shepherd, P. (2006). Cohort Profile: 1970 British Birth Cohort (BCS70). International Journal of Epidemiology, 35(4), 836–843. https://doi.org/10.1093/ije/dyl174

Falci, C., Mortimer, J., & Noel, H. (2010) Parental timing and depressive symptoms in early adulthood. *Advances in Life Course Research*, 15: 1-10.

Fomby, P. & Bosick, S. (2013) Family instability and the transition to adulthood, *Journal of Marriage and Family*, 75: 1266-1287.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., & Rubin, D. (2014). *Bayesian data analysis* (3rd ed.). Boca Raton : CRC Press.

Gibb, S. J., Fergusson, D. M., Horwood, L. J., & Boden, J. M. (2015). Early Motherhood and Long-Term Economic Outcomes: Findings From a 30-Year Longitudinal Study. *Journal of Research on Adolescence*, 25(1), 163–172. https://doi.org/10.1111/jora.12122

Gorry, D. (2019) Heterogeneous consequences of teenage childbearing, *Demography*, 56: 2147-2168.

Grundy, E., & Foverskov, E. (2016) Age at first birth and later life health in western and eastern Europe, *Population and Development Review*, 42(2): 245-269.

Grundy, E., Read, S., & Vaisanan, H. (2019) Fertility trajectories and later-life depression among parents in England, *Population Studies*, https://doi.org/10.1080/00324728.2019.1649450

Gunes, P. (2016) The effects of teenage childbearing on long-term health in the US: a twin-fixed-effects approach, *Rev Econ Household*, 14: 891-920

Hainmueller, J., Mummolo, J., & Xu, Y. (2019) How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice, *Political Analysis*, 27: 163-192

Hatzenbuehler, M., Phelan, J., & Link B. (2013) Stigma as a fundamental cause of population health inequalities, *American Journal of Public Health*, 103(5): 813-821.

Henretta, J. C., Grundy, E. M. D., Okell, L. C., & Wadsworth, M. E. J. (2008). Early motherhood and mental health in midlife: A study of British and American cohorts. *Aging & Mental Health*, 12(5), 605–614. https://doi.org/10.1080/13607860802343084.

Hill, J. (2011) Bayesian nonparametric modelling for causal inference, *Journal of Computational and Graphical Statistics*, 20(1): 217-240.

Hillis, S., Robert, F., Dube, S., Felitti, V., Marchbanks, P., & Marks, J. (2004) The association between adverse childhood experiences and adolescent pregnancy, long-term psychosocial consequences, and fetal death, *Pediatrics*, 113(2): 320-327.

Kalucza, S. (2018). Mental health problems and social disadvantages as predictors of teenage parenthood: A register-based population study of Swedish boys and girls. *Longitudinal and Life Course Studies*, 9(2), 212–225. https://doi.org/10.14301/llcs.v9i2.444

Kane, J. B., Morgan, S. P., Harris, K. M., & Guilkey, D. K. (2013). The Educational Consequences of Teen Childbearing. *Demography*, *50*(6), 2129–2150. https://doi.org/10.1007/s13524-013-0238-9

Kravdal, O., Grundy, E., & Skirbekk, V. (2017) Fertility history and use of antidepressant medication in late mid-life: a register based analysis of Norwegian women and men, *Aging and Mental Health*, 21(5): 477-486. http://dx.doi.org/10.1080/13607863.2015.1118010

McArthur, M., & Winkworth, G. (2018). Give Them a Break: How Stigma Impacts on Younger Mothers Accessing Early and Supportive Help in Australia. *British Journal of Social Work*, 48(5), 1277–1295. https://doi.org/10.1093/bjsw/bcx073

McCulloch, R., Sparapani, R., Spanbauer, C., Gramacy, R., & Pratola, M. (2021) *BART: Bayesian Additive Regression Trees*, R package version 2.9.

Mollborn, S. (2017) Teenage mothers today: what we know and how it matters, *Child Development Perspectives*, 11(1): 63-69.

Mollborn, S. & Morningstar, E. (2009) Investigating the relationship between teenage childbearing and psychological distress using longitudinal evidence, *Journal of Health and Social Behaviour*, 50: 310-326.

Morgan, Z., Brugha, T., Fryers, T., & Stewart-Brown, S. (2012) The effects of parent-child relationship status on later life mental health in two national birth cohorts, *Social Psychiatry and Psychiatric Epidemiology*, 47: 1707-1715.

Morgan, H., & Taylor, K. (2018). BCS70 technical report: Age 46 Survey. NatCen Social Research.

Pearlin, L. (1989) The sociological study of stress, *Journal of Health and Social Behaviour*, 30: 241-256.

Pearlin, L. I., Schieman, S., Fazio, E., & Meersman, S. (2005). Stress, health, and the life course: Some conceptual perspectives. Journal of Health and Social Behavior, 46, 205–219. doi:10.1177/002214650504600206

Penman-Aguilar, A., Carter, M., Snead, C., & Kourtis, A. (2013) Socioeconomic disadvantage as a social determinant of teen childbearing in the US, *Public Health Reports*, 128: 5-22

Rodgers, B., Pickles, A., Power, C., Collishaw, S., & Maughan, B. (1999). Validity of the Malaise Inventory in general population samples. *Social Psychiatry & Psychiatric Epidemiology*, 34(6), 333–341. https://doi.org/10.1007/s001270050153

Rutter, M., Tizard, J., and Whitmore, K. (1970) Education, health and behaviour, London: Longmans.

SmithBattle, L. (2013) Reducing the stigmatization of teen mothers, *MCN American Journal of Maternal/Child Nursing*, 38(4): 235-241.

Stewart-Brown, S., & Janmohamed, K. (2008). *WEMWBS User Guide, Version 1*. Retrieved May 11, 2020, from http://www.ocagingservicescollaborative.org/wp-content/uploads/2014/07/WEMWBS-User-Guide-Version-1-June-2008.pdf

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J., & Stewart-Brown, S. (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes*, 5(1), 63. https://doi.org/10.1186/1477-7525-5-63

van Buuren, S. (2012) *Flexible imputation of missing data*, Boca Raton: CRC Press

Wager, S. & Athey, S. (2018) Estimation and inference of heterogeneous treatment effects using random forests, *Journal of the American Statistical Association*, 113: 1228-1242.

Whitworth, T., (2017) Teen childbearing and depression: Do pregnancy attitudes matter? *Journal of Marriage and Family*, 79: 390-404.

Webbink, D., Martin, N. G., & Visscher, P. M. (2008). Does teenage childbearing increase smoking, drinking and body size? *Journal of Health Economics*, 27: 888-903.

Wendling, T., Jung, K., Callhan, A., Schuler, A., Shah, N., Gallego, B. (2018) Comparing methods for estimation of heterogeneous treatment effects using observational data from health care databases, *Statistics in Medicine*, 37: 3309-3324.

Yardley, E. (2008) Teenage mothers' experiences of stigma, *Journal of Youth Studies*, 11(6): 671-684.

Xavier, C., Brown, H., & Benoit, A. (2017) Teenage pregnancy and long-term mental health outcomes among Indigenous women in Canada, *Archives of Women's Mental Health*, 21: 333-340.

Xavier, C., Benoit, A., & Brown, H. (2018) Teenage pregnancy and mental health beyond the postpartum period: A systematic review, *Journal of Epidemiology and Community Health*, 72: 451-457.

Xie, Y., Brand, J., & Jann, B. (2012) Estimating Heterogeneous treatment effects with observational data, *Sociological Methodology*, 42(1): 314-347.

Appendices

Appendix A1. Overview of model covariates, with topic domains and age at survey.

Domain	Age	Measure
Delinquency	10	Fights with other children
Educational attainment and aspirations	5	Estimated reading age
	5	Attends school
	5	In need of special education treatment
	16	Feels like school is largely a waste of time
	16	Has left school
Family history/stability	0	Birth order
	0	Duration of marriage prior to birth
	0	Mother's marital status at time of birth
	0	CM mother age at first birth
	5	Both bio mother/father present
	10	parents divorced/separated
	10	Number of addresses
	16	Total number of siblings
Health behaviour	16	Is vaccinated
	16	CM age at menarche
	16	Ever been on the pill?
	16	CM contraceptive use
	16	CM have had sex
Family housing situation	5	Tenure type
	5	Interviewer description home - luxuriousness
	5	Number of household appliances present
	5	Crowding measure - number od household members per room
	5	Interviewer Social grouping of neighbourhood
Mental health in childhood	5	Frequent temper tantrums (once a week)
	5	Any sleep problems
	10	Appears miserable, depressed
	10	Diagnosed mental handicap or behavioural issue (medical)
	10	Appears worried
	16	Teen changes mood quickly / drastically
	16	Teenager displays outbursts of temper
Orientation to the world	10	Feeling shy or foolish when talking to teacher
	10	Academic self concept
Parent attitudes/beliefs	5	Attitude to child independence
Peers and peer characteristics	10	Peers smokes
	10	Relationship to peers in school
	10	Not much liked by other children (mother)
Physical health	0	Gestational age in weeks
	0	Foetal distress during delivery (heart rate > 160 / < 120)

	0	Birthweight
	5	Medical Condition: Hearing or Vision problem
	5	Ever admitted to hospital
	5	Breastfeeding duration
	10	Medical condition that affects daily life or school
	10	Any diagnosed physical condition
	10	Obesity
	10	Any disabilities
Parent relationship with cohort member	5	Any parents reads to cohort member
	10	Mothers interest in school (teacher answer)
	10	How much time do you spend talking to your parents every day?
	10	Parents expects too high standards/over concerned (teacher answer)
	10	Parents dismissive of child's potential
	10	Age at which cohort members mother think cohort member will leave school
	16	Are you and your husband completely happy with how your teenager is turning out?
	16	Weekly Activities shared with parents, n
	16	Parents have visited teens school since sept 85
	16	Would like teenager to stay in full time education after this school year, and to A levels etc.
	16	Would like teenager to continue some form of fulltime education beyond age 18.
Parental health and health behaviour (parent of cohort member)	0	Mothers height
	0	Mother's smoking history - if smoked during this pregnancy
	5	Parental Rutter Score
	5	Mother current smoking
	16	Parent total Malaise score
Parental and grandparental class/education (parent/grandparent of cohort member)	0	Mothers age at completion of full-time education
	0	Father age at completion of full-time education
	0	Father's Social Class 1970
	0	Mothers's Social Class 1970
	5	Mother is housewife
	16	Combined parental income per week
	16	Seriously troubled by financial hardship the last 12 months

Appendix. A2. Decscriptive statistics of model covariates, over age at first birth.

	Age at first birth < 20 (N=589)	Age at first birth 20 or older (N=2585)	Overall (N=3174)
Fights with other children, age 10			
Mean (SD)	20.0 (23.0)	15.5 (17.7)	16.3 (18.8)
Median [Min, Max]	13.0 [0, 100]	12.0 [0, 100]	12.0 [0, 100]
Missing	99 (16.8%)	355 (13.7%)	454 (14.3%)
Estimated reading, age 5			
Mean (SD)	4.67 (1.01)	4.98 (0.958)	4.92 (0.976)
Median [Min, Max]	4.79 [2.05, 6.80]	5.11 [2.05, 6.91]	5.11 [2.05, 6.91]
Missing	177 (30.1%)	854 (33.0%)	1031 (32.5%)
Attends school, age 5			
no	407 (69.1%)	1796 (69.5%)	2203 (69.4%)
yes	62 (10.5%)	312 (12.1%)	374 (11.8%)
Missing	120 (20.4%)	477 (18.5%)	597 (18.8%)
In need of special education treatment, age 5			
no	424 (72.0%)	1855 (71.8%)	2279 (71.8%)
yes	7 (1.2%)	27 (1.0%)	34 (1.1%)
Missing	158 (26.8%)	703 (27.2%)	861 (27.1%)
Feels like school is largely a waste of time, age 16			
no	120 (20.4%)	872 (33.7%)	992 (31.3%)
yes	114 (19.4%)	422 (16.3%)	536 (16.9%)
Missing	355 (60.3%)	1291 (49.9%)	1646 (51.9%)
Has left school, age 16			
no	232 (39.4%)	1337 (51.7%)	1569 (49.4%)
yes	14 (2.4%)	91 (3.5%)	105 (3.3%)
Missing	343 (58.2%)	1157 (44.8%)	1500 (47.3%)
Birthorder, age 0			
Mean (SD)	2.31 (1.47)	2.12 (1.33)	2.16 (1.36)
Median [Min, Max]	2.00 [1.00, 10.0]	2.00 [1.00, 11.0]	2.00 [1.00, 11.0]
Missing	49 (8.3%)	175 (6.8%)	224 (7.1%)
Duration of marriage prior to birth, age 0			
Mean (SD)	4.88 (4.23)	5.19 (4.16)	5.14 (4.17)
Median [Min, Max]	4.00 [0, 22.0]	4.00 [0, 23.0]	4.00 [0, 23.0]
Missing	103 (17.5%)	328 (12.7%)	431 (13.6%)

	Age at first birth < 20 (N=589)	Age at first birth 20 or older (N=2585)	Overall (N=3174)
Mother's marital status at time of birth, age 0			
other	59 (10.0%)	175 (6.8%)	234 (7.4%)
married	481 (81.7%)	2235 (86.5%)	2716 (85.6%)
Missing	49 (8.3%)	175 (6.8%)	224 (7.1%)
Mother age at first birth, age 0			
Mean (SD)	20.6 (3.52)	22.0 (3.76)	21.7 (3.75)
Median [Min, Max]	20.0 [13.0, 38.0]	22.0 [12.0, 42.0]	21.0 [12.0, 42.0]
Missing	53 (9.0%)	182 (7.0%)	235 (7.4%)
Both bio mother/father present, age 5			
no	73 (12.4%)	207 (8.0%)	280 (8.8%)
yes	400 (67.9%)	1921 (74.3%)	2321 (73.1%)
Missing	116 (19.7%)	457 (17.7%)	573 (18.1%)
Parents divorced/separated, age 10			
no	351 (59.6%)	1848 (71.5%)	2199 (69.3%)
yes	99 (16.8%)	252 (9.7%)	351 (11.1%)
Missing	139 (23.6%)	485 (18.8%)	624 (19.7%)
Number of adresses, age 10			
Mean (SD)	2.40 (1.20)	2.25 (1.11)	2.28 (1.13)
Median [Min, Max]	2.00 [1.00, 5.00]	2.00 [1.00, 5.00]	2.00 [1.00, 5.00]
Missing	102 (17.3%)	390 (15.1%)	492 (15.5%)
Total number of siblings, age 16			
Mean (SD)	1.80 (1.85)	1.87 (1.75)	1.86 (1.76)
Median [Min, Max]	1.00 [0, 5.00]	2.00 [0, 5.00]	2.00 [0, 5.00]
Missing	177 (30.1%)	521 (20.2%)	698 (22.0%)
Is vaccinated, age 16			
no	230 (39.0%)	1026 (39.7%)	1256 (39.6%)
yes	182 (30.9%)	1038 (40.2%)	1220 (38.4%)
Missing	177 (30.1%)	521 (20.2%)	698 (22.0%)
Age at menarche, age 16			
under 12	62 (10.5%)	257 (9.9%)	319 (10.1%)
12-13	129 (21.9%)	753 (29.1%)	882 (27.8%)
over 13	38 (6.5%)	243 (9.4%)	281 (8.9%)
not started	2 (0.3%)	11 (0.4%)	13 (0.4%)

	Age at first birth < 20 (N=589)	Age at first birth 20 or older (N=2585)	Overall (N=3174)
Missing	358 (60.8%)	1321 (51.1%)	1679 (52.9%)
Ever been on the pill?, age 16			
no	162 (27.5%)	1016 (39.3%)	1178 (37.1%)
yes	68 (11.5%)	253 (9.8%)	321 (10.1%)
Missing	359 (61.0%)	1316 (50.9%)	1675 (52.8%)
Contraceptive use, age 16			
contraception	139 (23.6%)	634 (24.5%)	773 (24.4%)
no sex	33 (5.6%)	320 (12.4%)	353 (11.1%)
none	62 (10.5%)	334 (12.9%)	396 (12.5%)
Missing	355 (60.3%)	1297 (50.2%)	1652 (52.0%)
Have had sex, age 16			
never	160 (27.2%)	1028 (39.8%)	1188 (37.4%)
once	11 (1.9%)	80 (3.1%)	91 (2.9%)
several	63 (10.7%)	180 (7.0%)	243 (7.7%)
Missing	355 (60.3%)	1297 (50.2%)	1652 (52.0%)
Tenure type, age 5			
owned	199 (33.8%)	1212 (46.9%)	1411 (44.5%)
rented	252 (42.8%)	806 (31.2%)	1058 (33.3%)
other	21 (3.6%)	101 (3.9%)	122 (3.8%)
Missing	117 (19.9%)	466 (18.0%)	583 (18.4%)
Interviewer description home - luxiouriousness, age 5			
low	41 (7.0%)	63 (2.4%)	104 (3.3%)
adequate	219 (37.2%)	753 (29.1%)	972 (30.6%)
high	203 (34.5%)	1270 (49.1%)	1473 (46.4%)
Missing	126 (21.4%)	499 (19.3%)	625 (19.7%)
Number of household appliances present, age 5			
Mean (SD)	4.62 (1.41)	5.07 (1.23)	4.99 (1.27)
Median [Min, Max]	5.00 [0, 7.00]	5.00 [0, 7.00]	5.00 [0, 7.00]
Missing	116 (19.7%)	457 (17.7%)	573 (18.1%)
Crowding measure - number resident household members / number rooms, age 5			
Mean (SD)	0.990 (0.347)	0.882 (0.297)	0.902 (0.310)
Median [Min, Max]	1.00 [0.429, 3.00]	0.800 [0.160, 3.60]	0.800 [0.160, 3.60]
Missing	124 (21.1%)	486 (18.8%)	610 (19.2%)

	Age at first birth < 20 (N=589)	Age at first birth 20 or older (N=2585)	Overall (N=3174)
Interviewer Social grouping of neighbourhood, age 5			
Mean (SD)	2.33 (0.891)	2.56 (0.883)	2.52 (0.889)
Median [Min, Max]	2.00 [1.00, 4.00]	2.00 [1.00, 4.00]	2.00 [1.00, 4.00]
Missing	134 (22.8%)	544 (21.0%)	678 (21.4%)
Freq temper tantrums (once a week), age 5			
no	388 (65.9%)	1794 (69.4%)	2182 (68.7%)
yes	62 (10.5%)	199 (7.7%)	261 (8.2%)
Missing	139 (23.6%)	592 (22.9%)	731 (23.0%)
Any sleep problems, age 5			
no	343 (58.2%)	1570 (60.7%)	1913 (60.3%)
yes	127 (21.6%)	544 (21.0%)	671 (21.1%)
Missing	119 (20.2%)	471 (18.2%)	590 (18.6%)
Appears miserable, depressed, age 10			
Mean (SD)	33.5 (31.1)	32.6 (30.1)	32.8 (30.2)
Median [Min, Max]	18.5 [0, 100]	19.0 [0, 100]	19.0 [0, 100]
Missing	101 (17.1%)	353 (13.7%)	454 (14.3%)
Diagnosed mental handicap or behavioural issue (medical), age 10			
no	523 (88.8%)	2357 (91.2%)	2880 (90.7%)
yes	18 (3.1%)	46 (1.8%)	64 (2.0%)
Missing	48 (8.1%)	182 (7.0%)	230 (7.2%)
Appears worried, age 10			
Mean (SD)	22.0 (23.8)	19.7 (21.2)	20.1 (21.7)
Median [Min, Max]	14.0 [0, 100]	14.0 [0, 99.0]	14.0 [0, 100]
Missing	97 (16.5%)	351 (13.6%)	448 (14.1%)
Teen changes mood quickly / drastically, age 16			
Mean (SD)	1.85 (0.929)	1.58 (0.823)	1.63 (0.847)
Median [Min, Max]	2.00 [1.00, 4.00]	1.00 [1.00, 4.00]	1.00 [1.00, 4.00]
Missing	268 (45.5%)	938 (36.3%)	1206 (38.0%)
Teenager displays outbursts of temper, age 16			
Mean (SD)	1.67 (0.901)	1.45 (0.739)	1.48 (0.772)
Median [Min, Max]	1.00 [1.00, 4.00]	1.00 [1.00, 4.00]	1.00 [1.00, 4.00]
Missing	269 (45.7%)	942 (36.4%)	1211 (38.2%)
Feeling shy or foolish when talking to teacher, age 10			

	Age at first birth < 20 (N=589)	Age at first birth 20 or older (N=2585)	Overall (N=3174)
no	232 (39.4%)	1025 (39.7%)	1257 (39.6%)
yes	240 (40.7%)	1040 (40.2%)	1280 (40.3%)
Missing	117 (19.9%)	520 (20.1%)	637 (20.1%)
Academic self concept, age 10			
Mean (SD)	12.4 (1.95)	12.6 (1.82)	12.6 (1.85)
Median [Min, Max]	12.0 [7.00, 18.0]	12.8 [7.00, 20.0]	12.8 [7.00, 20.0]
Missing	114 (19.4%)	502 (19.4%)	616 (19.4%)
Attitude to child independence, age 5			
Mean (SD)	1.21 (0.787)	1.34 (0.768)	1.32 (0.772)
Median [Min, Max]	1.18 [0.0274, 3.41]	1.28 [0.00218, 3.41]	1.27 [0.00218, 3.41]
Missing	223 (37.9%)	785 (30.4%)	1008 (31.8%)
Peers smokes, age 10			
most or some	83 (14.1%)	237 (9.2%)	320 (10.1%)
none	392 (66.6%)	1843 (71.3%)	2235 (70.4%)
Missing	114 (19.4%)	505 (19.5%)	619 (19.5%)
Relationship to peers in school, age 10			
Mean (SD)	9.00 (1.66)	9.14 (1.68)	9.11 (1.68)
Median [Min, Max]	9.00 [5.00, 13.0]	9.00 [5.00, 14.0]	9.00 [5.00, 14.0]
Missing	111 (18.8%)	498 (19.3%)	609 (19.2%)
Not much liked by other children (mother), age 10			
Mean (SD)	17.8 (21.6)	14.8 (17.0)	15.3 (18.0)
Median [Min, Max]	12.0 [0, 100]	12.0 [0, 99.0]	12.0 [0, 100]
Missing	101 (17.1%)	358 (13.8%)	459 (14.5%)
Gestational age in weeks, age 0			
Mean (SD)	39.8 (1.84)	39.8 (1.69)	39.8 (1.72)
Median [Min, Max]	40.0 [31.0, 47.0]	40.0 [30.0, 47.0]	40.0 [30.0, 47.0]
Missing	167 (28.4%)	652 (25.2%)	819 (25.8%)
Fetal distress during delivery (heart rate > 160 / < 120), age 0			
no	515 (87.4%)	2316 (89.6%)	2831 (89.2%)
yes	0 (0%)	7 (0.3%)	7 (0.2%)
Missing	74 (12.6%)	262 (10.1%)	336 (10.6%)
Birthweight, age 0			
Mean (SD)	3220 (519)	3250 (503)	3240 (506)

	Age at first birth < 20 (N=589)	Age at first birth 20 or older (N=2585)	Overall (N=3174)
Median [Min, Max]	3200 [1360, 5450]	3260 [1020, 5390]	3260 [1020, 5450]
Missing	50 (8.5%)	176 (6.8%)	226 (7.1%)
Medical Condition: Hearing or Vision problem, age 5			
no	432 (73.3%)	1893 (73.2%)	2325 (73.3%)
yes	41 (7.0%)	235 (9.1%)	276 (8.7%)
Missing	116 (19.7%)	457 (17.7%)	573 (18.1%)
Ever admitted to hospital, age 5			
no	337 (57.2%)	1470 (56.9%)	1807 (56.9%)
yes	65 (11.0%)	272 (10.5%)	337 (10.6%)
Missing	187 (31.7%)	843 (32.6%)	1030 (32.5%)
Breastfeeding duration, age 5			
Mean (SD)	4.03 (1.55)	3.83 (1.58)	3.87 (1.58)
Median [Min, Max]	5.00 [1.00, 5.00]	5.00 [1.00, 5.00]	5.00 [1.00, 5.00]
Missing	122 (20.7%)	471 (18.2%)	593 (18.7%)
Medical condition that affects daily life or school, age 10			
no	412 (69.9%)	1992 (77.1%)	2404 (75.7%)
yes	85 (14.4%)	242 (9.4%)	327 (10.3%)
Missing	92 (15.6%)	351 (13.6%)	443 (14.0%)
Any diagnosed physical condition, age 10			
no	409 (69.4%)	1848 (71.5%)	2257 (71.1%)
yes	132 (22.4%)	555 (21.5%)	687 (21.6%)
Missing	48 (8.1%)	182 (7.0%)	230 (7.2%)
Obesity, age 10			
underweight	63 (10.7%)	334 (12.9%)	397 (12.5%)
normal	352 (59.8%)	1574 (60.9%)	1926 (60.7%)
obese	61 (10.4%)	273 (10.6%)	334 (10.5%)
Missing	113 (19.2%)	404 (15.6%)	517 (16.3%)
Any disabilities, age 10			
yes	141 (23.9%)	517 (20.0%)	658 (20.7%)
no	330 (56.0%)	1646 (63.7%)	1976 (62.3%)
Missing	118 (20.0%)	422 (16.3%)	540 (17.0%)
Parents reads to, age 5			
no	71 (12.1%)	175 (6.8%)	246 (7.8%)

	Age at first birth < 20 (N=589)	Age at first birth 20 or older (N=2585)	Overall (N=3174)
yes	387 (65.7%)	1905 (73.7%)	2292 (72.2%)
Missing	131 (22.2%)	505 (19.5%)	636 (20.0%)
Mothers interest in school (teacher), age 10			
Mean (SD)	2.33 (1.41)	1.95 (1.30)	2.02 (1.33)
Median [Min, Max]	2.00 [1.00, 5.00]	2.00 [1.00, 5.00]	2.00 [1.00, 5.00]
Missing	128 (21.7%)	547 (21.2%)	675 (21.3%)
How much time do you spend talking to your parents every day?, age 10			
not much	209 (35.5%)	715 (27.7%)	924 (29.1%)
a lot	264 (44.8%)	1351 (52.3%)	1615 (50.9%)
Missing	116 (19.7%)	519 (20.1%)	635 (20.0%)
Parents expects too high standards/over concerend (teacher), age 10			
no	453 (76.9%)	1921 (74.3%)	2374 (74.8%)
yes	26 (4.4%)	175 (6.8%)	201 (6.3%)
Missing	110 (18.7%)	489 (18.9%)	599 (18.9%)
Parents dissmisive of childs potential, age 10			
no	458 (77.8%)	2043 (79.0%)	2501 (78.8%)
yes	21 (3.6%)	53 (2.1%)	74 (2.3%)
Missing	110 (18.7%)	489 (18.9%)	599 (18.9%)
Age mother think CM will leave school, age 10			
16 years	250 (42.4%)	834 (32.3%)	1084 (34.2%)
17 years	71 (12.1%)	385 (14.9%)	456 (14.4%)
18 years	153 (26.0%)	890 (34.4%)	1043 (32.9%)
don't know	16 (2.7%)	93 (3.6%)	109 (3.4%)
Missing	99 (16.8%)	383 (14.8%)	482 (15.2%)
Are you and your husband completely happy with how your teenager is turning out?, age 16			
no	144 (24.4%)	494 (19.1%)	638 (20.1%)
yes	182 (30.9%)	1177 (45.5%)	1359 (42.8%)
Missing	263 (44.7%)	914 (35.4%)	1177 (37.1%)
Weekly Activities shared with parents, n, age 16			
Mean (SD)	1.57 (2.33)	1.74 (2.22)	1.71 (2.24)
Median [Min, Max]	0 [0, 13.0]	1.00 [0, 11.0]	1.00 [0, 13.0]
Missing	177 (30.1%)	521 (20.2%)	698 (22.0%)
Parents have visited teens school since sept 85, age 16			

	Age at first birth < 20 (N=589)	Age at first birth 20 or older (N=2585)	Overall (N=3174)
no	170 (28.9%)	726 (28.1%)	896 (28.2%)
yes	148 (25.1%)	899 (34.8%)	1047 (33.0%)
Missing	271 (46.0%)	960 (37.1%)	1231 (38.8%)
Would like teenager to stay in full time educationa after this school year, and to A levels etc., age 16			
no	225 (38.2%)	1116 (43.2%)	1341 (42.2%)
yes	102 (17.3%)	568 (22.0%)	670 (21.1%)
Missing	262 (44.5%)	901 (34.9%)	1163 (36.6%)
Would like teenager to continue som form of fulltime education beyond age 18, age 16			
no	274 (46.5%)	1372 (53.1%)	1646 (51.9%)
yes	53 (9.0%)	312 (12.1%)	365 (11.5%)
Missing	262 (44.5%)	901 (34.9%)	1163 (36.6%)
Mothers height, age 0			
Mean (SD)	160 (6.30)	161 (6.51)	161 (6.47)
Median [Min, Max]	160 [140, 182]	161 [138, 185]	161 [138, 185]
Missing	54 (9.2%)	203 (7.9%)	257 (8.1%)
Mother's smoking history - if smoked during this pregnancy, age 0			
no	265 (45.0%)	1402 (54.2%)	1667 (52.5%)
yes	275 (46.7%)	1008 (39.0%)	1283 (40.4%)
Missing	49 (8.3%)	175 (6.8%)	224 (7.1%)
Parental Rutter Score, age 5			
Mean (SD)	8.54 (4.88)	7.37 (4.38)	7.58 (4.49)
Median [Min, Max]	8.00 [0, 27.0]	7.00 [0, 26.0]	7.00 [0, 27.0]
Missing	140 (23.8%)	558 (21.6%)	698 (22.0%)
Mother current smoking, age 5			
no	213 (36.2%)	1197 (46.3%)	1410 (44.4%)
yes	247 (41.9%)	866 (33.5%)	1113 (35.1%)
Missing	129 (21.9%)	522 (20.2%)	651 (20.5%)
1986: (CM's parent) total Malaise score, age 16			
Mean (SD)	9.19 (5.98)	7.65 (5.26)	7.90 (5.41)
Median [Min, Max]	8.00 [0, 32.0]	6.00 [0, 35.0]	7.00 [0, 35.0]
Missing	280 (47.5%)	980 (37.9%)	1260 (39.7%)
Mother age completion of full-time education, age 0			
<15	45 (7.6%)	152 (5.9%)	197 (6.2%)

	Age at first birth < 20 (N=589)	Age at first birth 20 or older (N=2585)	Overall (N=3174)
15	352 (59.8%)	1440 (55.7%)	1792 (56.5%)
16	83 (14.1%)	404 (15.6%)	487 (15.3%)
17	25 (4.2%)	162 (6.3%)	187 (5.9%)
>17	27 (4.6%)	228 (8.8%)	255 (8.0%)
Missing	57 (9.7%)	199 (7.7%)	256 (8.1%)
Father age completion of full-time education, age 0			
<15	39 (6.6%)	174 (6.7%)	213 (6.7%)
15	351 (59.6%)	1342 (51.9%)	1693 (53.3%)
16	53 (9.0%)	333 (12.9%)	386 (12.2%)
17	21 (3.6%)	153 (5.9%)	174 (5.5%)
>17	36 (6.1%)	285 (11.0%)	321 (10.1%)
Missing	89 (15.1%)	298 (11.5%)	387 (12.2%)
Father's Social Class 1970, age 0			
one	15 (2.5%)	89 (3.4%)	104 (3.3%)
two	37 (6.3%)	281 (10.9%)	318 (10.0%)
three, non manual	35 (5.9%)	281 (10.9%)	316 (10.0%)
three, manual	231 (39.2%)	1096 (42.4%)	1327 (41.8%)
four	103 (17.5%)	334 (12.9%)	437 (13.8%)
five	41 (7.0%)	140 (5.4%)	181 (5.7%)
other	27 (4.6%)	69 (2.7%)	96 (3.0%)
not_supported	43 (7.3%)	109 (4.2%)	152 (4.8%)
Missing	57 (9.7%)	186 (7.2%)	243 (7.7%)
Mothers's Social Class 1970, age 0			
one or two	28 (4.8%)	211 (8.2%)	239 (7.5%)
three, non manual	133 (22.6%)	691 (26.7%)	824 (26.0%)
three, manual	20 (3.4%)	114 (4.4%)	134 (4.2%)
four	148 (25.1%)	467 (18.1%)	615 (19.4%)
five	5 (0.8%)	21 (0.8%)	26 (0.8%)
other	6 (1.0%)	14 (0.5%)	20 (0.6%)
home duties	156 (26.5%)	677 (26.2%)	833 (26.2%)
Missing	93 (15.8%)	390 (15.1%)	483 (15.2%)
Mother housewife, age 5			
no	193 (32.8%)	928 (35.9%)	1121 (35.3%)

	Age at first birth < 20 (N=589)	Age at first birth 20 or older (N=2585)	Overall (N=3174)
yes	276 (46.9%)	1190 (46.0%)	1466 (46.2%)
Missing	120 (20.4%)	467 (18.1%)	587 (18.5%)
Combined parental income per week, age 16			
<150	157 (26.7%)	488 (18.9%)	645 (20.3%)
150-249	64 (10.9%)	417 (16.1%)	481 (15.2%)
250-399	40 (6.8%)	290 (11.2%)	330 (10.4%)
400+	12 (2.0%)	111 (4.3%)	123 (3.9%)
Missing	316 (53.7%)	1279 (49.5%)	1595 (50.3%)
Seriously troubled by financial hardship the last 12 months, age 16			
no	251 (42.6%)	1474 (57.0%)	1725 (54.3%)
yes	85 (14.4%)	243 (9.4%)	328 (10.3%)
Missing	253 (43.0%)	868 (33.6%)	1121 (35.3%)