Understanding the Effects of Early Motherhood in Britain: the Effects on Mothers*

Alissa Goodman† Greg Kaplan‡ Ian Walker§

September 13, 2004

Abstract

This paper examines the socio-economic consequences of teenage motherhood for a cohort of British women born in 1970. We apply a number of different methodologies to the same dataset, including OLS, a propensity score matching estimator, and an instrumental variables estimator, using miscarriages as an instrument. We bound the biases introduced through IV due to non-randomness, and misreporting of the instrument. Our results are sensitive to the methodologies used. Taking only observed characteristics into account, the effects of teenage motherhood appear large and negative. The pathways are through bigger family size, and negative labour market outcomes for the mother and her partner, and are mitigated by transfers from the state through the British benefit system. Our IV estimates show that almost all these effects are reduced to zero once unobserved heterogeneity is taken into account. However our IV bounds show that biases introduced by non-randomness and misreporting of our instrument could be responsible for all of this apparent reduction in effects.

JEL Codes J31

Keywords teenage pregnancy, miscarriage, instrumental variables

---

*Funding for this research comes from HM-Treasury's Evidence Based Policy Fund with co-funding from the Department for Education and Skills, the Department for Work and Pensions, Inland Revenue and the Department for Culture, Media and Sport. We are grateful to John Ermisch and Gauthier Lanot for comments on earlier results presented at a DWP workshop in 2002, and to Erich Battistin, Laura Blow, and Frank Windmeijer, Barbara Sianesi and other colleagues at IFS, for much discussion and advice. We also benefited from comments from Jeffrey Smith and an anonymous referee. The British Cohort Study data was supplied by the UK ESRC Data Archive at the University of Essex and is used with permission. The usual disclaimer applies.

†Insititute for Fiscal Studies, alissa_g@ifs.org.uk

‡New York University and Institute for Fiscal Studies, gregkaplan@nyu.edu

§University of Warwick and Institute for Fiscal Studies, i.walker@warwick.ac.uk

1 Introduction

The worsening record on teenage pregnancies of both Britain and the USA relative to other countries motivates a continued interest in estimating the long-term socioeconomic consequences of teenage motherhood. UK teenage birth rates are the highest in Western Europe and yet are still significantly less than in the USA.[1] Britain is the only country in Western Europe which has not experienced a significant decline in teenage fertility rates in the last thirty years.[2] This paper is concerned with estimating the effects of early motherhood for a cohort of British women born in 1970, calculating how much of the well documented association of early motherhood and negative later-life economic outcomes can be given a causal interpretation. In addition we explore the extent to which the state insures teenagers against any economic loss associated with early motherhood through income transfers later in life.

The question of whether early motherhood is an indicator of prior disadvantage, a pathway to future disadvantage, or both, is one that has been extensively debated in recent literature. This question has important policy implications for the nature, timing and targeting of interventions to assist young mothers. It has also challenged researchers to find appropriate econometric techniques to distinguish between these conflicting stories. Existing data and methodologies have led to disparate evidence. Conventional estimates have indicated large negative socioeconomic effects of early motherhood and support interventions aimed at reducing the incidence of teenage conceptions. More recent evidence, that allows for the impact of prior disadvantage, has indicated smaller (and in some cases even zero or positive) effects, suggesting that the pathway to disadvantage started much earlier in the young woman's life and cannot be entirely attributed to early motherhood. If this is the case, policies which are aimed simply at preventing teenage conceptions or births will be less effective in ameliorating the negative outcomes of concern than the raw data would otherwise suggest.

Choosing between these two stories is an empirical question and our paper develops and applies the methodologies to distinguish between them. We start by comparing linear regression estimates with semi-parametric propensity score matching estimates because of fears that the regression estimates may be sensitive to functional form and a potential lack of common support. A recent example of matching estimates used in this context is provided by Levine and Painter (2003) who suggest that when matching can be performed within schools, the estimated effects of teen motherhood on educational outcomes are approximately half those obtained with conventional regression models. Our focus is on economic, rather than educational outcomes and, in this respect, our work can be thought of as complimentary to the earlier research.

We then move on to addressing the problem of unobserved heterogeneity. To this end, a number of techniques have been attempted in the recent liter-

[1]For example, UNICEF (2001) reports the number of teenage births per 1000 population as 52.1 for the USA, 30.8 for the UK, 13.2 for Germany, 9.2 for France and 7.9 for Spain

[2]See Social Exclusion Unit (1999)

ature. These have included family fixed effects (siblings and cousins)[3], twins studies[4] and instrumental variables[5]. In this paper, we follow Hotz, McElroy and Sanders (1997) and exploit data on miscarriages to form an instrumental variable that, under certain assumptions, can yield consistent estimates of the effects of early motherhood on those that experienced early motherhood - that is, the effect of the treatment on the treated. The approach is akin to a natural experiment, where the experience of miscarriage exogenously delays age at first birth, allowing the construction of a counterfactual for the outcomes of teenage mothers, had they not given birth as a teenager[6]. Attempts to use this method, such as the paper cited above, have been controversial because they have resulted in much smaller effects than traditional estimates. For example, Hotz, McElroy and Sanders find that early motherhood tends to raise levels of labour supply, accumulated work experience and labour market earnings by the time a teen mother reaches her late twenties. The use of this method is also controversial because estimates based on this methodology are potentially biased for a number of reasons[7], most notably non-random occurence and misreporting of miscarriages. We are interested in whether this bias can account for all of the difference between the IV estimates and conventional estimates. Hotz, Mullins and Sanders (1997) calculate a bound on the maximum amount of bias introduced thorugh non-random occurence of miscarriages. We show how, under plausible assumptions, it is possible to further bound the IV estimates for the effects of misreporting of miscarriages by exploiting aggregate administrative data from the Office for National Statistics (ONS) on births, miscarriages and abortions.

Our results suggest a number of points of interest. First, that the biases inherent in the miscarriage estimates may account for their small size relative to OLS and matching estimates. Second, when we examine the constituent parts of equivalised family income we find that whereas there are large effects on equivalised family income at age 30, almost all of this effect is through household size and composition, with little or no effect on household income. Third, teenage motherhood results in signifcantly higher benefit income from the state, fully compensating for the negative effect of teenage motherhood on own wages and partner's wages. Fourth, all of the negtiave socio-economic effects of teenage motherhood at age 30 are uniformly larger for teenagers who gave birth between the ages 18 and 20 than those whose first birth was before age 18, suggesting that some of these effects may be temporary.

The contribution of this paper is twofold. Our first contribution is methodological: by presenting a number of estimates of the impact of early motherhood

[3]See for example Ribar (1999), Hoffman, Foster and Furstenberg Jr (1993a) and Geronimus and Korenman (1992).

[4]See for example Bronars and Groggar (1994).

[5]See for example Klepinger, Lundberg and Plotnick (1998) and Chevalier and Viitanen (2002).

[6]Recent work by Ermisch (2004) also applies this method to UK data, however that paper restricts itself to examining the consequences of teenage motherhood for outcomes in the marriage market

[7]These are discussed in Sections 3 and 4.

using different techniques on same data set, we are able to compare parameter estimates under a wide range of assumptions. Here we add to the literature both by extending the work of Hotz, Mullins and Sanders (1997) to take into account a broader range of potential biases arising from the mis-reporting of miscarriages, and also by applying a propensity score matching estimator, complementing the work in Levine and Painter (2003). Our second contribution is empirical. By decomposing the effect on family equivalised income of the mother at age 30 into its constituent parts, we are able to assess the likely pathways contributing to the effects we find. We highlight two particular contributors which have not been much focussed on in the literature to date: the first is the impact of teenage motherhood on family size and composition, and its importance in determining the socioeconomic consequences for mothers at age 30; second we highlight the importance of the British benefit system in insuring teenage mothers against any long term negative economic effects.

The paper proceeds as follows. In section 2 we examine the various approaches that have been used to estimate the effects of early motherhood in the existing literature. Section 3 provides a discussion of the use of miscarriages as an instrumental variable and propensity score matching. In section 4 we outline the formal econometric framework and derive bounds for our IV estimates. In section 5 we discuss the data and in section 6 we present the results of the econometric analyses. Finally, section 7 concludes.

2 Approaches and Findings in the Existing Literature

In the last decade a number of new studies have used a variety of innovative methods to control for unobserved characteristics influencing selection into teenage motherhood. Whereas earlier studies were based on linear models, controlling for observed characteristics only[8], the newer literature has treated this as an evaluation problem, with an explicit emphasis on the estimation of a treatment parameter for early motherhood. The various approaches have differed primarily in the comparison group that has been used to construct the counterfactual outcome for teen mothers.

These new approaches have generated a debate in the literature as to whether once these unobserved characteristics are controlled for, any negative effects of early childbearing remain. However, drawing any robust conclusions from this debate has been difficult because of the sensitivity of the results to the empirical methodology chosen and the data set being used.[9]

One group of studies exploit family fixed-effects to compare the outcomes for teenage mothers with those of their sisters. Geronimus and Korenman (1992) used samples drawn from the National Longitudinal Survey of Young Women

[8]See for example Hofferth and Moore (1979) for the USA, and Hobcraft and Kiernan (2001) for the UK.

[9]Hoffman (1998) provides a good synthesis of this debate.

(NLSYW), National Longitudinal Survey of Youth (NLSY) and Panel Study of Income Dynamics (PSID) and found that fixed-effects estimates were smaller than conventional estimates. In the case of the NLSYW results, the effects were not statistically different from zero, implying that once family-level unobserved characteristics are controlled for, there remains little or no effect on subsequent socioeconomic outcomes. However, Hoffman, Foster and Furstenberg Jr (1993b) noted that the NLSYW results are somewhat of an outlier, with the PSID and NLSY results indicating that, while substantially smaller than conventional estimates, the effects of early childbearing are still negative and significant, even in the fixed-effects models. This conclusion was supported by further analysis of the PSID data in Hoffman, Foster and Furstenberg Jr (1993a). One possible explanation for the surprising results in the NLSWY data is the older age at which outcomes are measured (28-31 compared with 21-33 in the PSID and NLSY data), suggesting that there could be a significant temporary effect of early motherhood, but that this effect disappears over time.

However, even if one were to believe the PSID and NLSY results, it is unlikely that family fixed-effects are able to appropriately control for unobserved characteristics influencing selection into teenage motherhood. Maintaining that these characteristics differ only at the family and not the individual level, so that sisters are identical in all unobserved aspects that would influence both the decision to give birth at a young age and later socioeconomic outcomes (such as career motivation) is perhaps an unrealistically strong assumption.

More recently, Ribar (1999) developed a simultaneous equation model for sisters' outcomes to calculate the effects of teenage motherhood under different assumptions about the correlation of siblings unobserved characteristics. Maintaining the assumption that is equivalent to a family fixed-effects model results in estimates for family income-to-needs ratio[10] and years of education from the NLSY that are significantly negative, and comparable to those in Geronimus and Korenman (1992). However, estimates of effects for family income are not statistically different from zero. Under a different set of assumptions, which are equivalent to allowing each sister's fertility to instrument for the other's childbearing behaviour, he finds implausibly large, negative effects of early childbearing.[11]

A different form of fixed-effects analysis is explored in Brien, Loya and Pepper (2002) who control for individual unobserved heterogeneity by looking at changes in mothers' cognitive development over time. Because the authors observe two test scores before a teenager gives birth and one test score after, they are able to control for unobserved factors that influence the level and growth of test scores. Their differences-in-differences analysis indicates that while teenage mothers have lower test scores than teenagers who did not give birth, the direct effects of giving birth on test scores are negligible.

A particularly innovative idea implemented by Bronars and Groggar (1994)

[10]The income-to-needs ratio is income divided by the poverty level for the woman's reported family size.

[11]One possible explanation for the unusual IV results is that sisters' fertilities are not strongly correlated, so effectively this is a weak instrument problem.

was to exploit the random nature of giving birth to twins, conditional on becoming pregnant, to create a natural experiment. The idea rests on the assumption that the effect of giving birth to twins as a teenager on later socioeconomic outcomes is twice that of giving birth to a singleton as a teenager. If this is the case then one can compare outcomes for teenagers who gave birth to twins with outcomes for teenagers who bore singletons to get consistent estimates of the effects of teenage motherhood. The assumed randomness of giving birth to twins accounts for unobserved heterogeneity. They find that there are substantial effects on the short-run labour force participation for all teenage mothers, but lasting effects on the probability of eventual marriage and family earnings only for blacks. However it is unlikely that the necessary assumption for identification holds. Rather, it is probably the case that if effects of teenage motherhood exist, most of the effect is captured by the presence of any children (compared to none), so that the effect on teenagers bearing twins is likely to be less than twice that for teenagers bearing singletons.

Other researchers have searched for appropriate instrumental variables that can explain teenage fertility but are not related to unobserved characteristics that influence later socioeconomic outcomes. The most commonly used instruments have been age at menarche, and regional indicators of sexual awareness and access to contraception. For example, Chevalier and Viitanen (2003) use age of menarche as an instrument, whilst Klepinger, Lundberg and Plotnick (1998) used menarche and state/county level information. However, studies which use age of menarche as an instrument for uncovering the effects of teenage motherhood need to be interpreted with caution. Although age at menarche may exogenously alter the timing of pregnancy, it seems unlikely that it would affect whether or not a young woman gives birth, conditional on becoming pregnant. As such, these studies estimate a different treatment parameter to the one that is of interest in this paper. In section 4, we clearly outline our parameter and population of interest.

Finally, a controversial, but potentially helpful methodology has been to exploit the random nature of miscarriages as a mechanism for exogenously delaying age at first birth. This methodology, and the consequences of violations of the assumptions underlying this technique, are discussed in detail in the following section.

Britain and the USA have acute problems with teenage pregnancy[12], and while the studies cited above examine the USA, there is little British evidence on which to base policy prescriptions. The existence of full, retrospective pregnancy histories in the 30 year old sweep of the British Cohort Study (BCS) makes it possible to apply some of the aforementioned techniques to examine the pattern of results for a newer cohort than has previously been analysed in Britain. To our knowledge, Chevalier and Viitanen (2003) is the first UK example and uses age at menarche as an instrument to control for unobserved heterogeneity in an earlier cohort of children born in 1958 - the NCDS. A further analysis is very recent work by Ermisch (2004) which uses the same BCS dataset and the

[12] See Social Exclusion Unit (1999)

same instrument as we use here to look at the effects of teenage motherhood on the quality of the cohort members' partners. We compliment and extend that work by using a propensity score matching method, by accounting for possible misreporting of miscarriages and by considering a broader range of outcomes, including the disaggregation of equivalised family income into its constituent parts. Finally, Robinson (2002) constructs synthetic cohorts from cross-section surveys pooled over time to estimate the lifecycle evolution of the wage penalty associated with teen motherhood. Her results show that the wage gap between teen mothers and others is largest in the late 20's and early 30's and closes only slowly thereafter.[13] She further shows that the wage penalty appears to be larger for recent cohorts. Our data correspond to the age where the wage difference is estimated to be at its maximum.

3 Miscarriages as an Instrumental Variable and Propensity Score Matching

3.1 Miscarriages as an Instrumental Variable

The idea of exploiting miscarriages as a natural experiment to estimate the effects of teenage childbearing was first attempted by Hotz, McElroy and Sanders (1997). The idea is that, if miscarriages occur randomly and are reported correctly, then they represent situations where age at first birth has been exogenously delayed. By comparing outcomes for young women whose first pregnancy ended in a miscarriage with those who gave birth, it is possible to control for all unobserved factors that simultaneously influence the decision to become pregnant as a teenager, the decision to not terminate the pregnancy and the outcome being considered.

However, this methodology has been criticized on various grounds. Importantly, most of the problems with using miscarriages tend, under plausible assumptions, to induce an upwards bias in the estimates, towards zero.[14] This means that it is unclear whether the small effects estimated in Hotz, Mullins and Sanders (1997) and Hotz, McElroy and Sanders (1997) are indicating downward bias in conventional estimates or are being driven by the upward biases inherent in the miscarriage method. It is hence useful to specify the conditions required for miscarriages to provide consistent estimates of the true effects, so that we can get a firm grasp on whether violation of these conditions can explain the discrepancy in results.

[13] While her paper does not address causality, it does examine the results for sensitivity to the inclusion of parental class and country of origin and finds the results to be insensitive to the inclusion of these pre-existing conditions. However this does not, of course, preclude sensitivity to other possible controls or for selection on unobserved variables.

[14] The socio-economic outcomes being considered are all defined such that a more negative co-efficient represents a stronger negative effect of early motherhood. Hence, when we use the term 'upward bias', we refer to an under-estimate of the effect, whilst a 'downward bias' refers to an over-estimate of the negative effect.

Condition 1 The occurrence of a miscarriage for a pregnant teenager is random with respect to any existing unobserved characteristics that are correlated with the outcome of interest.

Condition 2 All pregnancies and their outcomes are reported correctly.

Condition 3 The occurrence of a miscarriage has no independent effect on the outcome of interest

Numerous researchers have observed that Condition 1 may not be satisfied. For example, there is some evidence that drinking and smoking while pregnant may increase the probability of a young woman experiencing a miscarriage. If the decision to smoke and/or drink while pregnant is correlated with other unobserved factors that impact on future socioeconomic outcomes, then this will lead to biased and inconsistent estimates. Another potential source of non-randomness is domestic abuse that results in a miscarriage.

However, the epidemiological literature seems to indicate that the vast majority of miscarriages are random, particularly with respect to future socioeconomic outcomes. Regan (2001) notes that approximately 50% of reported miscarriages are due to a variety of foetal chromosomal abnormalities and the remainder are largely due to neural tube defects, viral and bacterial infections in the mother and other foetal genetic defects. All of these causes can be considered as random with respect to future socioeconomic outcomes, conditional on observed characteristics. Moreover, Regan (2001) also notes that the remaining non-random causes of miscarriages are primarily pre-existing complicating factors, such as type-1 diabetes, the occurrence of which one would not expect to be correlated with economic and educational outcomes, after controlling for other background factors.

Hotz, Mullins and Sanders (1997) are able to calculate bounds for the true causal effect of early motherhood, accounting for the extent of violations of Condition 1. For most of their samples and outcomes, they are unable to reject conventional point estimates of the effects, based on the bounds. Two different figures were used for the proportion of miscarriages that occur randomly - an extremely conservative estimate of 38%, and a more realistic estimate of 84%, although the conclusions are not overly sensitive to the estimate used. In this paper, we use a variant on this method to account for violation of Condition 1, and as we will show, to similar effect.

It is important to note that for violation of Condition 1 to induce upward bias in the estimates it is necessary that the correlation between unobserved characteristics and a miscarriage being non-random is negative. In other words, those teenagers experiencing a non-random miscarriage must realise worse outcomes than the teenagers whose miscarriages are random.

Condition 2 may be violated in a number of ways. We go beyond the methodology of Hotz, Mullins, and Sanders (1997), by considering the two most likely possibilities. The first type of misreporting we consider is non-reporting of pregnancies, that may have occurred up to 15 years ago. This is a problem in all studies that use retrospective pregnancy history information such as we use here. In particular if the sample of teenagers who report their pregnancies is not representative of the total population of teenagers who experienced a pregnancy, then this may affect estimates of the effect of early motherhood. In particular, if females who became pregnant as a teenager and experienced a miscarriage but did not report the pregnancy went on to achieve better outcomes on average than teenagers who did report the miscarriage, then this will induce upwards bias in the IV estimates. Second, young women may be especially reluctant to report an abortion that they may have had. There is thus the possibility that whilst the teenage pregnancy is correctly reported, the outcome of the pregnancy is misclassified as a miscarriage. This type of misreporting will lead to an understatement of the effects of early motherhood if those women who report abortions as miscarriages are more disadvantaged than the general population of teenage mothers. A key contribution of this paper is to demonstrate how, under relatively weak assumptions, we can use aggregate administrative ONS data on births, miscarriages and abortions to bound the biases introduced into the IV estimates from both these types of misreporting. The reliability of our bounds depend on the assumption that the ONS data represents the true population proportion of births, miscarriages and abortions among pregnant teenagers. To the extent that the ONS data is biased, so will our bounds. The derivation of these bounds is set out in section 4.

Finally, violation of condition 3 may also affect our results. This condition is equivalent to the absence of a placebo effect in a controlled laboratory experiment. It states that the only way in which a miscarriage can affect the outcome under consideration is by preventing a birth (and the effects associated with a birth) from having occurred. However, the experience of a miscarriage for a pregnant teenager may be accompanied by feelings from elation to depression. It is conceivable that the loss of a wanted child could have important lasting effects on the young woman, while it is also possible that the loss of a pregnancy that was likely to be terminated by abortion has a positive impact on the teenager. We do not calculate bounds for these possible effects. The question we must ask is whether we think that these effects are important and long-lasting enough to explain differences in socioeconomic outcomes ten to fifteen years on. For example, if post-miscarriage depression and other physical effects have an immediate impact on schooling, then this could have longer-term effects on other economic indicators.

3.2 A propensity score matching estimator

As well as using miscarriages as an instrumental variable to find the impact of teenage motherhood, we also present results from a propensity score matching estimator. This technique is quite different from the instrumental variables estimator, since it does not allow us to control for any unobserved factors that influence the decision to not terminate a pregnancy and the outcome being considered. However, by matching teen mothers to other pregnant teenagers who did not give birth, the technique does control for unobserved factors influencing selection into pregnancy. We measure the impact of early motherhood on the assumption that, conditional on the observed covariates and conditional on falling pregnant, there are no unobserved factors determining selection into early motherhood that also determine later life outcomes. In this respect it is similar to estimates derived using linear regression (also presented here), however it does not require the researcher to specify any particular functional form for the relation between early motherhood and later life outcomes, making the specification completely flexible. The specific background characteristics on which we match are outlined in section 5.

4 Empirical Framework

Following the notation of Hotz, Mullins and Sanders (1997), let Y_1 be the socioeconomic later-life outcome that would result if a young woman gave birth as a teen and let Y_0 be the outcome that would result if she did not. Further, let x be a vector of observed background characteristics that are correlated with both the outcome of interest and the decision to give birth as a teen, and let D be an indicator variable equal to 1 if the young woman gave birth as a teen. Initially, we restrict our attention to an unusual sub-population of females that became pregnant as a teenager - those that reported their pregnancies plus those whose pregnancies ended in miscarriage and did not report the pregnancy. [15] This is neither the full population of pregnant teenagers, nor the observed population of teenagers who report pregnancies. It is however a useful population for estimating and bounding our parameter of interest. Further below, we relate our estimators to those estimated from the sub-population of reported pregnancies only.

Our choice of population means that D is equal to 0 for females that reported a pregnancy as a teenager but did not give birth and those who had a miscarriage as a teenager but did not report the pregnancy. Conditioning on this sub-population is left implicit in what follows. The outcome for an individual female, as a function of the observed characteristics, x, can then be written as:

$$Y(x) = Y_0(x) + \beta(x)D$$

where the effect of early motherhood on the outcome of interest is

$$\beta(x) = Y_1(x) - Y_0(x).$$

We are interested in identifying the average effect of giving birth as a teen for those females who did give birth as a teen. Because we are implicitly conditioning on falling pregnant as a teenager, we are separating out the effects on future outcomes of teenage pregnancy and teenage motherhood. The parameter

[15] Our focus on this sub-population is motivated by our interest in bounding the bias introduced by non-reporting of pregnancies that ended in miscarriages. We do not attempt to bound the effects of non-reporting of pregnancies that ended in births or abortions. We are hence implicitly assuming that any non-reporting births and abortions is random with respect to the outcomes being considered.

that we are identifying is the effect of the birth itself, over and above any effect of an early conception. This is known as the Average Treatment on the Treated (ATT) in the evaluation literature and can be expressed as

$$E(\beta|D=1,x) = E(Y_1|D=1,x) - E(Y_0|D=1,x). \tag{1}$$

The first term in (1) is readily identified from data on age at first birth, but the second term, commonly referred to as the counterfactual, is not. Herein lies the identification problem inherent in studies that estimate the effects of early motherhood using simple OLS regressions. These studies effectively replace the second term in (1) with $E(Y_0|D=0,x)$. To the extent that $E(Y_0|D=0,x)$ differs from $E(Y_0|D=1,x)$, these estimates will be biased and inconsistent. Without some additional data and/or assumptions the ATT is not identified.

If data on miscarriages is available, then under certain assumptions the ATT is identified. Let Z^* be an indicator for the occurrence of a random miscarriage, equal to 1 if the young woman became pregnant as a teen with the pregnancy ending in a randomly occurring miscarriage, and 0 if the pregnancy did not end in a randomly occurring miscarriage (either because no miscarriage occurred or because a non-random miscarriage occurred). By *random* miscarriage, we mean miscarriages that satisfy Condition 1 in section 3. Miscarriages that do not satisfy this condition are referred to as *non-random* miscarriages. As with D, we define Z^* over the sub-sample of women became pregnant as a teenager and reported the pregnancy as well as teenagers who experienced pregnancies that ended in miscarriage and did not report the pregnancy. Due to the nature of pregnancy resolution, it is clear that a pregnancy can not end in both a birth and a miscarriage and so $\Pr(D=1 \cap Z^*=1)=0$.

Regan (2001) provides evidence that as many as 99% of miscarriages occur within the first 13 weeks of pregnancy. However, aggregate statistics of abortion by gestation age suggest that one third of abortions occur at less than nine weeks gestation and a further 54% from nine to twelve weeks[16]. We make the (reasonable) assumption that abortions are not delayed in anticipation of a possible miscarriage.[17] Let the latent preference for births and abortions be denoted by D^L for teenagers experiencing miscarriages, equal to 1 for females with a latent preference for births and 0 for those who would have chosen an abortion. This sequence of events can be described in Figure 1.

[Figure 1 near here]

In order for the ATT to be identified, we make the following further assumptions:

Assumption 1. $\Pr(D^L=0|Z^*=1,x) = Pr(D=0|Z^*=0,x)$,and $\Pr(D^L=1|Z^*=1,x) = Pr(D=1|Z^*=0,x)$

[16]See http://www.statistics.gov.uk/STATBASE/Expodata/Spreadsheets/D8492.xls

[17]Our data pre-dates the introduction of the morning-after pill in the UK, so that any complications that may be introduced are avoided.

The proportion of females with a latent preference for births over abortions among those females experiencing a miscarriage is the same as the proportion of females with a preference for births among those females not experiencing miscarriages, conditional on observed characteristics.

Assumption 2. $(Z^* \perp Y_0)|x$

The occurrence of a random miscarriage is independent of the outcome that would have been realized in the absence of a birth, conditional on observed characteristics.

Consider the expression, $E(Y|Z^* = 1, x) - E(Y|Z^* = 0, x)$. Exploiting the sequence of events in Figure 1, this can be written, under Assumptions 1 and 2, as

$$\begin{aligned} E(Y|Z^* = 1, x) - E(Y|Z^* = 0, x) = \\ \{E(Y_0|Z^* = 1, D^L = 1, x) - E(Y_1|Z^* = 0, D = 1, x)\} \\ * Pr(D = 1|Z^* = 0, x). \end{aligned} \tag{2}$$

Noting that under Assumption 2, the term in braces in (2) is equivalent to $-[E(Y_1|D = 1) - E(Y_0|D = 1)]$, and re-arranging, gives an unbiased estimator for the ATT defined in (1)

$$E(\beta|D = 1, x) = \frac{E(Y|Z^* = 1, x) - E(Y|Z^* = 0, x)}{-\Pr(D = 1|Z^* = 0, x)}. \tag{3}$$

This parameter can be consistently estimated by using Z^* as an instrument for D in the following model. Conditioning on observed characteristics, x, is left implicit, and is implemented by regressing all variables on the vector of observed characteristics, x, and a constant, and working with residuals. Below, Y, D, Z and Z^* refer to these residuals.

$$\begin{aligned} Y &= \beta D + u \\ D &= \gamma Z^* + v; \quad v = -\gamma \text{ for } Z^* = 1 \end{aligned}$$

in which case the IV estimator for β is given by

$$\beta^*_{IV} = \frac{cov(Z^*, Y)}{cov(Z^*, D)}, \tag{4}$$

and is equivalent to the expression in (1). This estimator is unbiased and consistent for β.

However, Z^* is not observed. Let us assume that instead we observe Z[18], an indicator variable equal to 1 where a woman became pregnant as a teen and *re-*

[18]Actually we do not observe Z because Z is defined over the same population as Z^* and D. Z can be thought of as those reported miscarriages within the set of all reported pregnancies and unreported miscarriages. Instead,we observe a variable, $\tilde{Z}$, that is slightly different from Z. In section 4.6 we show how the bound derived here in terms of Z can be adjusted to reflect $\tilde{Z}$.

ported a miscarriage (either correctly or incorrectly) and where that miscarriage, if one did in fact occur, could have occurred either *randomly* or *non-randomly.*

Replacing Z^* with Z in the model above, and using the IV estimator

$$\hat{\beta}_{IV} = \frac{cov(Z,Y)}{cov(Z,D)} \tag{5}$$

will result in biased and inconsistent estimates for β. This is the estimator that has been the source of objections to the use of miscarriage as an instrument for births.[19]

In general, existing estimates of β that do not control for unobserved heterogeneity, have tended to indicate strong negative effects of early motherhood.[20] For the purposes of comparison, we will refer to these estimators as OLS estimators for β,

$$\hat{\beta}_{OLS} = \frac{cov(D,Y)}{var(D)}. \tag{6}$$

However estimates based on (5) have tended to be zero or even positive.[21] Our task is to estimate the extent to which the difference between $\hat{\beta}_{IV}$ and $\hat{\beta}_{OLS}$ reflects upward bias in $\hat{\beta}_{IV}$ (the difference between $\hat{\beta}_{IV}$ and β^*_{IV}) and the extent to which it signifies overstatement of the effects of early motherhood in earlier estimates (the difference between β^*_{IV} and $\hat{\beta}_{OLS}$).

Thinking again of D, Z and Z^* as events rather than residuals, we can define the true outcomes for pregnancy resolution as either birth (B), abortion (A), random miscarriage which was reported (M), a non-random miscarriage which was reported (NR) or a miscarriage for which the pregnancy went unreported (UM). The relationship between Z and Z^* can then be expressed as

$$Z = Z^* + \epsilon$$

and the joint distribution of (D, Z^*, Z) can be conceptualized as shown below, with the triple (D, Z^*, Z) taking one of only 5 possible values.

$$(D, Z^*, Z) = \begin{cases} (1,0,0) & \rightarrow & B & \rightarrow & \epsilon = 0 \\ (0,1,1) & \rightarrow & M & \rightarrow & \epsilon = 0 \\ (0,0,1) & \rightarrow & A, NR & \rightarrow & \epsilon = 1 \\ (0,1,0) & \rightarrow & UM & \rightarrow & \epsilon = -1 \\ (0,0,0) & \rightarrow & A & \rightarrow & \epsilon = 0 \end{cases}$$

Under Assumptions 1 and 2, β^*_{IV} is consistent for β. If $cov(Z,D) \approx cov(Z^*,D)$, then the difference between this consistent estimator and the identifiable one, $\hat{\beta}_{IV}$, can be expressed as

$$\hat{\beta}_{IV} - \beta^*_{IV} = \frac{cov(Z,Y)}{cov(Z,D)} - \frac{cov(Z^*,Y)}{cov(Z^*,D)} \approx \frac{cov(\epsilon,Y)}{cov(Z,D)}. \tag{7}$$

[19] See for example the criticism of Hotz, McElroy and Sanders (1996) in Hoffman (1998).
[20] For example, see Hobcraft and Kiernan (2001)
[21] For example Hotz McElroy Sanders (1996) and Hotz, Mullins and Sanders (1997)

The assumption that $cov(Z,D) \approx cov(Z^*,D)$ is necessary to ensure that this difference can be evaluated. This is a testable assumption, and is one that holds approximately in our data, as will be seen in section 5. Below we discuss the implications of non-equality of $cov(Z,D)$ and $cov(Z^*,D)$.

The mutually exclusive nature of births and miscarriages ensures that $cov(Z,D)$ is negative and hence the maximum value of (7), representing the maximum upward difference between $\hat{\beta}_{IV}$ and β^*_{IV}, is realized at the lower bound for $cov(\epsilon,Y)$. We can write $cov(\epsilon,Y)$ as

$$cov(\epsilon,Y) = Pr(\epsilon=1)\Big[E(Y|\epsilon=1) - E(Y)\Big] \\ - Pr(\epsilon=-1)\Big[E(Y|\epsilon=-1) - E(Y)\Big]. \quad (8)$$

Assuming that we have consistent estimates for $Pr(\epsilon=1)$ and $Pr(\epsilon=-1)$, finding a lower bound for $cov(\epsilon,Y)$ relies on finding a lower bound for $E(Y|\epsilon=1)$ and an upper bound for $E(Y|\epsilon=-1)$. $E(Y|\epsilon=1)$ can be further written as

$$E(Y|\epsilon=1) = E(Y|\epsilon=1,A)P_A + E(Y|\epsilon=1,NR)P_{NR}, \quad (9)$$

where P_A and P_{NR} refer to $Pr(A|\epsilon=1)$ and $Pr(NR|\epsilon=1)$ respectively.

4.1 Bounding $E(Y|\epsilon=1,NR)$

Employing the techniques outlined in Horowitz and Manski(1995), we can write this conditional expectation as

$$E(Y|\epsilon=1,NR) = E(Y|Y<Y_{(k_{NR})},Z=1)Pr(Y<Y_{(k_{NR})}|\epsilon=1,NR) \\ + E(Y|Y>Y_{(k_{NR})},Z=1)Pr(Y>Y_{(k_{NR})}|\epsilon=1,NR) \quad (10)$$

where $k_{NR} = Pr(\epsilon=1,NR|Z=1) = Pr(Z^*=0,NR|Z=1)$, represents the proportion of non-random miscarriages among the set of reported miscarriages, and $Y_{(\alpha)}$ represents the α-th percentile of the distribution of Y in the relevant conditioning sub-population. It is clear that a lower bound on this expectation is formed when

$$Pr(Y<Y_{(k_{NR})}|\epsilon=1,NR) = 1$$

and

$$Pr(Y>Y_{(k_{NR})}|\epsilon=1,NR) = 0,$$

so that a lower bound for $E(Y|\epsilon=1,NR)$ is given by

$$lowerbound[E(Y|\epsilon=1,NR)] = E(Y|Y\le Y_{(k_{NR})},Z=1). \quad (11)$$

4.2 Estimating $E(Y|\epsilon = 1, A)$

$E(Y|\epsilon = 1, A)$ refers to the mean outcome for those females who became pregnant as teenagers and terminated the pregnancy by means of an abortion, but reported the outcome of the pregnancy as a miscarriage. Since this group of women is not readily distinguishable from those who had miscarriages (and correctly reported them) as teenagers, a further assumption is needed to estimate this conditional expectation. Note that for this type of misreporting to induce an upward bias in the IV estimates, it is necessary that those who misreport abortions as miscarriages experience, on average, worse outcomes than would have been experienced had they not given birth, i.e. $E(Y|\epsilon = 1, A) < E(Y_0|D = 1)$. Moreover, if one also believed that the ATT is constrained to be negative or zero, so that giving birth as a teenager could not have a positive effect on the outcomes being considered i.e. $E(Y_0|D = 1) \geq E(Y_1|D = 1)$, then a plausible lower bound for $E(Y|\epsilon = 1, A)$ is $E(Y|D = 1)$. However, under a relatively weak assumption, it is possible to tighten this bound.

Assumption 3. $(Y \perp Z)|Z^* = 0, D = 0, A, x$

Given that a young woman had an abortion and reported the pregnancy, whether it was correctly reported as an abortion or incorrectly reported as a miscarriage is independent of the outcome, conditional on x.[22]

Under Assumption 3, $E(Y|\epsilon = 1, A) = E(Y|Z = 0, D = 0)$ which is observed in the data. We use the mean outcome for females who report abortions as a lower bound for the mean outcome of females who had an abortion but misreported it as a miscarriage. This will result in a tighter lower bound whenever those females who report an abortion as a teenager experience better outcomes than those who gave birth as a teenager, i.e. whenever $E(Y|Z = 0, D = 0) > E(Y|D = 1)$.

4.3 Estimating $E(Y|\epsilon = -1)$

This sub-population refers to those females who became pregnant as a teenager and experienced a miscarriage, but failed to report the pregnancy. For this type of misreporting to induce upward bias in the IV estimates, it is necessary that these women went on to achieve better outcomes than the true counterfactual for teenage mothers, i.e. $E(Y|\epsilon = -1) > E(Y_0|D = 1)$. One possible explanation would be that these teenagers regretted their early pregnancy and used the incident as motivation to succeed in later life. These females could feasibly not want to recollect an experience that they view as a negative aspect to their past. An alternative explanation is that these are the type of individuals who find it easier to put adverse events behind them and move on in life.It is possible that these are latent-abortion type individuals, who would have chosen to abort their pregnancy. Whichever stance one takes, it seems unlikely that these teenagers

[22] This is sometimes known as a non-differential measurement error assumption , in this case with regards the Abortion subset only.

have unobserved characteristics that are better for their future outcomes that those who reported an abortion. Hence a plausible upper bound is the mean outcome for teenagers who reported an abortion, $E(Y|Z=0,D=0)$.

Under Assumptions 1 to 3, a lower bound for $cov(\epsilon,Y)$ is given by

$$lowerbound[cov(\epsilon,Y)] = Pr(\epsilon=1)\Bigg[P_A * E(Y|Z=0,D=0)$$

$$+P_{NR}E(Y|Y \leq Y_{k_{NR}}, Z=1) - E(Y)\Bigg] - Pr(\epsilon=-1)\{E(Y|Z=0,D=0) - E(Y)\}. \tag{12}$$

4.4 Estimating Probabilities using External Data

Recall the testable assumption required for equation (7) that $cov(Z,D) \approx cov(Z^*,D)$. This assumption effectively says that the number of reported miscarriages observed in our data (including random miscarriages, non-random miscarriages and abortions reported as miscarriages) is the same as the actual number of random miscarriages that should have occurred, given the number of pregnancies that were reported (including actual random miscarriages and actual random miscarriages where the pregnancy was not reported). In other words, the number of non-random miscarriages and abortions reported as miscarriages ($D=0, Z^*=0, Z=1$) is the same as the number of non-reported random miscarriages ($D=0, Z^*=1, Z=0$). Hence this assumption also implies $Pr(\epsilon=1) = Pr(\epsilon=-1)$.[23]

Using this result along with the identity, $P_A + P_{NR} = 1$, to simplify the expression in (12) gives

$$lowerbound[cov(\epsilon,Y)] = \\ Pr(\epsilon=1)P_{NR}\Bigg[E(Y|Y \leq Y_{(k_{NR})}, Z=1) - E(Y|Z=0,D=0)\Bigg]. \tag{13}$$

This in turn can be re-written as[24]

$$lowerbound[cov(\epsilon,Y)] = \\ E(Z)k_{NR}\Bigg[E(Y|Y \leq Y_{(k_{NR})}, Z=1) - E(Y|Z=0,D=0)\Bigg]. \tag{14}$$

[23] To see this formally, note that $cov(D,Z) = Pr(Z=1)Pr(D=1)$ and $cov(D,Z^*) = Pr(Z^*=1)Pr(D=1)$, implying that $Pr(Z^*=1) = Pr(Z=1)$. Furthermore $Pr(\epsilon=1) = Pr(Z^*=0 \cap Z=1) = Pr(Z=1) - Pr(Z^*=1 \cap Z=1)$ and $Pr(\epsilon=-1) = Pr(Z^*=1 \cap Z=0) = Pr(Z^*=1) - Pr(Z^*=1 \cap Z=1)$. Hence $Pr(\epsilon=1) = Pr(\epsilon=-1)$.

[24] To see the equality of $Pr(\epsilon=1)P_{NR}$ and $E(Z)k_{NR}$, note that $Pr(\epsilon=1)P_{NR} = Pr(Z^*=0 \cap Z=1) * Pr(NR|Z^*=0, Z=1) = Pr(Z^*=0, Z=1, NR)$ and $E(Z)k_{NR} = Pr(Z=1) * \frac{Pr(Z^*=0,Z=1,NR)}{Pr(Z=1)} = Pr(Z^*=0, Z=1, NR)$.

Hence from (7) a lower bound for ${\beta^*}_{IV}$ is given by

$$lowerbound[{\beta^*}_{IV}] = \hat{\beta}_{IV} + \frac{k_{NR}\left[E(Y|Y \leq Y_{k_{NR}}, Z=1) - E(Y|Z=0, D=0)\right]}{Pr(D=1)}. \quad (15)$$

To evaluate this bound, the only probability that is required is k_{NR}.

Assumption 4 *$Pr[Z^* = 0, NR|(Z^* = 0, NR) \cup (Z^* = 1)]$ is known or can be estimated. This is the proportion of reported miscarriages that can be classified as having occurred non-randomly with respect to Y_0.*

To evaluate Assumption 4, we use epidemiologic evidence on the proportion of clinically recognized pregnancies ending in miscarriage that occur randomly. We present results for an estimate of this probability of 15%. This is consistent with the evidence in Regan(2001) and the estimate of 16% used in Hotz, Mullins and Sanders (1997). An implicit assumption that we make in constructing this probability is that the proportion of miscarriages that occur randomly is the same amongst reported and unreported miscarriages.

It is important to note that this method does not account for under-reporting of births and abortions in our data, but only misclassification of pregnancies, given that a pregnancy was in fact reported, and non-reported miscarriages.[25]. Furthermore, to the extent that one believes that the ONS birth and abortion statistics suffer from measurement error themselves, these results will be biased. We assume that the ONS statistics represent the true distribution of (D, Z^*).

4.5 The assumption that $cov(Z, D) \approx cov(Z^*, D)$

The assumption that $cov(Z, D) \approx cov(Z^*, D)$ is a testable assumption and holds if $P(Z = 1) \approx P(Z^* = 1)$. The interpretation of this condition is that the proportion of reported pregnancies that are reported as miscarriages is equal to the proportion of miscarriages which would be observed amongst the reported pregnancies if all miscarriages were reported correctly. It holds approximately in our data because there are two offsetting effects. On the one hand $P(Z = 1)$ may be greater than $P(Z^* = 1)$ because some abortions are erroneously reported as miscarriages and because some reported miscarriages cannot be considered as having occurred randomly. On the other hand, $P(Z = 1)$ may be less than $P(Z^* = 1)$ because some miscarriages go unreported. Table 3 in section 5 presents estimates of these probabilities from the British Cohort Study (Z) and the Office of National Statistics (Z^*)[26]. The numbers in Table 3 are

[25]Whether underreporting of births and abortions would significantly bias the IV estimates would depend on difference in the make up of the outcomes for reported and unreported pregnancies. I suspect that this bias would be quantitatively similar for OLS and IV estimates and so could not explain the difference between the two estimates. This is something I want to show . . .

[26]See the footnote to Table 3 for details of the sources for these statistics.

consistent this assumption, however this does not rule out other explanations for the difference between the BCS and ONS data.

To see explicitly the effect of this assumption for our bound, equation (16) shows the formula for the bound, maintaining all assumptions needed for the bound in equation (15) but relaxing the assumption that $cov(Z, D) \approx cov(Z^*, D)$.

$$
\begin{aligned}
&lowerbound[\beta^*{}_{IV}] = \hat{\beta}_{IV} \\
&\quad + \frac{Pr(Z=1|Z^*=1)}{Pr(Z^*=1|Z=1)} \frac{k_{NR}\Big[E(Y|Y \leq Y_{k_{NR}}, Z=1) - E(Y|Z=0, D=0)\Big]}{Pr(D=1)} \\
&+ \frac{Pr(Z^*=1|Z=1) - Pr(Z=1|Z^*=1)}{Pr(Z^*=1|Z=1)Pr(D=1)} \Big[E(Y|Z=1) - E(Y|Z=0, D=0)\Big]
\end{aligned} \tag{16}
$$

When $Pr(Z^* = 1|Z = 1) = Pr(Z = 1|Z^* = 1)$[27], the expression above simplifies to the formula in equation (15), showing clearly the simplifying effect that this assumption has for identification of the bound. This expression also makes clear the extra information that would be needed in order to identify the bound without making the assumption that $cov(Z, D) \approx cov(Z^*, D)$. We would need to be able to estimate $Pr(Z^* = 1|Z = 1)$ and $Pr(Z = 1|Z^* = 1)$, that is, the joint distribution of Z and Z^*.

What would be the implications for our estimates if this assumption did not hold? If $P(Z = 1) < P(Z^* = 1)$, that is if the proportion of observed miscarriages is less than the proportion of actual miscarriages, then this would imply that $cov(Z, D) > cov(Z^*, D)$ and the estimated lower bound for β^* would be greater than the actual lower bound. In other words the true bound would be tighter than our estimate. Conversely, if $P(Z = 1) > P(Z^* = 1)$, the true bound would be looser (more negative) than our estimate.

4.6 Further adjustments for unreported miscarriages

Recall that the bound in equation (15) is derived assuming that data is available for the population of females reporting a pregnancy as a teenager as well as those females who experienced a miscarriage as a teenager but did not report the pregnancy. In this section we adjust the bound to account for the fact that teenagers not reporting miscarriages are not observed in our data. Let $\tilde{D}$, $\tilde{Z}^*$ and $\tilde{Z}$ be analogous variables to D, Z^* and Z but defined only over the population of females reporting pregnancies as teenagers. The joint distribution of $(\tilde{D}, \tilde{Z}^*, \tilde{Z})$ can then be conceptualized as shown below, with the triple $(\tilde{D}, \tilde{Z}^*, \tilde{Z})$ now taking one of only 4 possible values.

$$
(\tilde{D}, \tilde{Z}^*, \tilde{Z}) = \begin{cases} (1,0,0) & \rightarrow \quad B & \rightarrow \quad \epsilon = 0 \\ (0,1,1) & \rightarrow \quad M & \rightarrow \quad \epsilon = 0 \\ (0,0,1) & \rightarrow \quad A, NR & \rightarrow \quad \epsilon = 1 \\ (0,0,0) & \rightarrow \quad A & \rightarrow \quad \epsilon = 0 \end{cases}
$$

[27] $cov(Z, D) = cov(Z^*, D)$ implies $Pr(Z^* = 1|Z = 1) = Pr(Z = 1|Z^* = 1)$

If $\hat{\beta}^{\tilde{Z}}_{IV}$ is the IV estimator calculated using a sample from this population, then the difference between $\hat{\beta}^{\tilde{Z}}_{IV}$ and $\hat{\beta}_{IV}$ can be written as

$$\begin{aligned}\hat{\beta}^{\tilde{Z}}_{IV} - \hat{\beta}_{IV} &= \frac{cov(\tilde{Z},Y)}{cov(\tilde{Z},\tilde{D})} - \frac{cov(Z,Y)}{cov(Z,D)} \\ &= \frac{E(Y|\tilde{Z}=1) - E(Y)}{-Pr(\tilde{D}=1)} - \frac{E(Y|Z=1) - E(Y)}{-Pr(D=1)}.\end{aligned}$$

Using the fact that $E(Y|Z=1) = E(Y|\tilde{Z}=1)$ and denoting the number of unreported miscarriages as UM and the number of reported births as B, this becomes

$$\hat{\beta}^{\tilde{Z}}_{IV} - \hat{\beta}_{IV} = \frac{UM}{B}\Big[E(Y|\tilde{Z}=1) - E(Y)\Big]. \tag{17}$$

Substituting into equation (15), we get the corrected bound.

$$\begin{aligned} lowerbound[\beta^{*}{}_{IV}] = {} & \hat{\beta}^{\tilde{Z}}_{IV} \\ & + \frac{k_{NR}}{Pr(D=1)}\Big[E(Y|Y \le Y_{(k_{NR})}, \tilde{Z}=1) - E(Y|\tilde{Z}=0,\tilde{D}=0)\Big] \\ & - \frac{UM}{B}\Big[E(Y|\tilde{Z}=1) - E(Y)\Big] \end{aligned} \tag{18}$$

In order to calculate this bound we need an estimate of the ratio of unreported teenage miscarriages to teenage births in our sample. We can estimate the number of unreported pregnancies by comparing the number of reported pregnancies with what we would expect based on ONS data. The number of unreported miscarriages can then be calculated by multiplying the number of unreported pregnancies by the proportion of pregnancies that end in miscarriage in the ONS statistics. Assuming that 12% of teenage pregnancies end in miscarriage, we get an estimate of UM/B of approximately 8%. This is the percentage that we use in our estimates.[28]

5 Data and Methodology

Our data comes from the British Cohort Study (BCS), a longitudinal study of a cohort of approximately 17,000 children born in Britain in the week 5-11 April 1970. Surviving members of the cohort have been followed up at ages 5, 10, 16 and 26, and most recently at age 29/30 in 1999/2000. The starting point for our sample is those females who responded to a questionnaire about their past fertility history as part of the age 30 interviews. This provides us with a sample of 5771 females.

We use two definitions of 'teenager' in all of our analysis – those aged up to (but not including) 18 years, and up to 20 years. Ideally, we would like to classify

[28]Because $E(Y|\tilde{Z}=1) - E(Y)$ is close to zero for most outcomes, the estimates are not sensitive to the assumption for UM/B

females based on age at first conception. Unfortunately, date of conception is not available in the BCS data. Instead, we classify based on age at the outcome of the first pregnancy.[29] Although the 18 year definition of a teenager may be considered preferable on theoretical grounds - it more closely reflects the time at which a pregnancy may trigger the mechanisms implicated in worsening later life socioeconomic outcomes - we focus on the 20 year definition because it provides larger sample sizes, allowing more robust inference. Moreover, this is the definition that has been more commonly adopted in the existing US literature.[30] Where a female became pregnant only once before the relevant cut-off age, we classify the outcome as either a birth, abortion (induced abortion) or miscarriage (spontaneous abortion).[31][32]

Previous studies that have exploited miscarriages to estimate the effects of early motherhood have been criticized for their treatment of females experiencing multiple pregnancies as teenagers.[33] The criticism centres on the fact that in these studies, many of the females in the miscarriage sample experienced additional pregnancies as a teenager which ended in either abortions or live births. Table 1 shows the number of teenagers who have had zero, one, two, three and four pregnancies before each cut-off age. No teenagers had more than four pregnancies by age 20 in our sample. Furthermore, many females who experienced a miscarriage as a teenager also experienced an abortion or gave birth before the relevant cut-off age. Table 2 shows the number of females in each of these categories. Including females in the miscarriage sample who also gave birth or had an abortion as a teenager would have a similar effect to contaminating the control group with the treated group in an experimental design, biasing the IV estimates. Moreover, by looking at the outcome of the other pregnancies we can learn something about the teenager's latent pregnancy resolution decision, had the pregnancy not ended in a miscarriage. In other words, we have information as to whether the teenager would have chosen to abort the pregnancy. In cases where the teenager had a latent abortion preference, we can no longer claim that the miscarriage served to exogenously delay age at first birth.

[Table 1 here]

[29]We could choose to impute dates of conception based on the outcome of the pregnancy and the date of the outcome. While this would give us a slightly larger sample of teenagers who became pregnant, it is unlikely that this would significantly affect our results. It also should be noted that aborted and miscarried pregnancies predate births by about 6 months. For this reason our age-cut-offs mean that we could very slightly undercount teenage abortions and miscarriages relative to the number of pregnancies.

[30]For example Ribar (1999).

[31]For the purposes of this paper we refer to induced abortions as "abortions" and spontaneous abortions as "miscarriages".

[32]The BCS data draws a distinction between pregnancies ending in miscarriage and those ending with a stillbirth. There is an argument for reclassifying stillbirths as miscarriages because stillbirths represent situations in which age at first birth has been exogenously delayed, however we exclude observations where the female had a stillbirth but no live birth or abortion by the age cut-off. This is done because Condition 3, discussed in Section 3, is much less likely to hold for stillbirths than for miscarriages. Only 3 females fall into this category and inclusion of these observations in the miscarriage sample does not significantly affect the results.

[33]For example, Hoffman (1998) makes this criticism about Hotz, Mullins and Sanders (1997).

[Table 2 here]

To overcome this problem, we define our non-pregnant, pregnant, birth, abortion and miscarriage samples as follows:

Non-pregnant Sample Females who did not report any pregnancy prior to the relevant cut-off age.

Pregnant Sample Females who reported at least one pregnancy prior to the relevant cut-off age.

Birth Sample $\left(\widetilde{D} = 1, \widetilde{Z} = 0\right)$ Females who had at least one birth prior to the relevant cut-off age.

Abortion Sample $\left(\widetilde{D} = 0, \widetilde{Z} = 0\right)$ Females who had at least one abortion and no births prior to the relevant cut-off age.

Miscarriage Sample $\left(\widetilde{D} = 0, \widetilde{Z} = 1\right)$ Females who had at least one birth prior to the relevant cut-off age.

Adopting these sample definitions has the effect of ensuring that the birth, abortion and miscarriage samples are mutually exclusive and together comprise the pregnant sample. Figure 1 shows the number of females in each of the samples for the two definitions of teenagers. Although the number of miscarriages is smaller than one would like for statistical purposes, the samples sizes are broadly consistent with those in Hotz, Mullins and Sanders (1997).

[Figure 2 here]

To give an indication of the possible extent of under-reporting and misreporting of pregnancies in the BCS data, Table 3 compares information from our data to the total number of conceptions, births, abortions and miscarriages per 1000 women aged 15-19 for the cohort born in 1970 based on official ONS population statistics (where available). It also shows – based on these figures - the proportion of all conceptions ending in birth, abortion or miscarriage. Official statistics are not available for miscarriage rates, however it is commonly accepted[34] that between 10% and 15% of clinically recognised pregnancies (births, abortions and miscarriages) end in miscarriage, with this proportion increasing with age. Thus a reasonable estimate for 15-19 year old females is somewhere in the vicinity of 10% to 12%.

[Table 3 here]

There is clearly a substantial amount of under-reporting of pregnancies amongst the BCS sample, with around 75-80 pregnancies per 1000 women going

[34] See Regan (1997).

unreported. As would be expected, a disproportionate amount of this under-reporting is among those pregnancies ending in abortions, with only 21% of pregnancies being reported as ending in abortion for the BCS sample, compared to 29-30% for the ONS statistics. Moreover, the fact that the BCS data show 12% of pregnancies ending in miscarriage, combined with the apparent higher proportion of pregnancies ending in births in the BCS sample (67% compared with 59%), suggest that indeed some abortions could be erroneously reported as miscarriages in the BCS data. These figures support both the notion that the unreported pregnancies are more likely to end in abortions or miscarriages than reported pregnancies, and the belief that some abortions are being misreported as miscarriages.

The outcomes we investigate cover a range of economic outcomes, all measured at age 29 or 30. Our primary outcome of interest is the natural logarithm of equivalised family income[35], however, in order to more fully understand what is driving the effects of this broad outcome, we break it down into its component parts. First, we investigate the cohort members' family size as measured by the equivalence scale[36] and the number of children that the individual has had by age 30. We then examine the cohort members' own labour market outcomes, including the natural logarithm of their hourly and weekly net wages[37] and their total hours worked. Next, we look at the natural logarithm of the cohort member's partner's weekly wages. We also examine two outcomes related to the dependency of the female on Government benefits - the logarithm of real benefits received per week and an indicator variable for whether the cohort member was in receipt of means-tested benefits. Finally, we present results for two educational outcomes - age left full-time education and whether or not the female continued in post-compulsory schooling. Because we believe the schooling decision to be highly endogenous to the pregnancy and birth decisions, it is difficult to give any robust interperetation to these results. There are arguments for both teenage pregnancy leading to an early termination of formal education and for being out of full-time education leading to an increased likelihood of becoming pregnant or proceeding with an unplanned pregnancy. We present the results here to provide a UK compliment to work in the US such as Levine and Painter(2003), rather than as a description of any causal effects. Results for the impact of early motherhood on these outcomes are presented in section 6.

Table 4 displays summary statistics for each of these outcomes for the various samples defined above. To conserve space, summary statistics are only shown for the 20-year definition of a teenager. The descriptive results are qualitatively similar for the other age groups in that for all outcomes and age definitions,

[35]Equivalised family income comprises cohort member's real net weekly income, partner's real net weekly income, real benefits received per week and real net weekly income from other sources (interest payments etc), adjusted to take account of household composition and size.

[36]We use McClement's (1977) equivalence scale. This is the most commonly used equivalence sacle in the UK. It does not take into account regional differences in living costs because there is no official data on this in the UK. Details of the scale are given in the footnote to Table 4.

[37]Where net wage data was missing, net wages were imputed from gross wages.

the birth sample has substantially lower (worse) average outcomes than for the not-pregnant, abortion and miscarriage samples.

[Table 4 here]

All regressions we report control for a range of background characteristics. The controls included are: age mother and father left full-time education; maths, reading and ability test scores at age 10; mother's age at birth; father's social class; banded family income at age 10 and age 16; and indicators at age 16 for whether the family had experienced financial hardship in the last year, and whether the girl's mother thinks sex education is important, whether her daughter will do A-levels[38], and whether her daughter will continue in full time education past age 18. We also include an indicator for whether the teenager had had a longstanding illness or disability[39]. The propensity score matching estimates we report are based on this same vector of observed characteristics.[40] Table 5 displays descriptive statistics for the background variables for the various samples.

We explicitly choose not to control for marital status of teenagers around the time of the birth. Our reasoning is that the marriage decision is endogenous to the birth decision. We want to estimate the full effect of teenage motherhood on human capital accumulation and later life outcomes, including any effects that are compounded by a correlated decision to enter into a young marriage.

[Table 5 here]

Three points are immediately clear from Table 5, emphasizing the selection problem that we are faced with when trying to estimate the causal effect of early motherhood. First, the birth sample comes from substantially more disadvantaged backgrounds on average than both the full sample and the not-pregnant sample. Those individuals who gave birth as a teenager have test scores at age 10 that are on average between 7 and 11 percentage points lower than teenagers who did not become pregnant. For each dimension, there is evidence that teenage birth is to some extent an indicator of prior disadvantage. Second, there is a remarkable similarity between the background characteristics for teenagers in the abortion sample and those in the not-pregnant sample. The income distributions at age 10 and 16, and the distribution of father's social class are almost identical for the two groups. This point, and the one noted above, provide a further warning against simply comparing the outcomes for teen mothers with non-teen mothers in order to assess the impact of teenage

[38]In the UK, A-levels are qualifications typically taken at the end of two years of post-compulsory schooling around the age of 18. A-level performance is a criterion for university admission.

[39]This was defined as a disability that severely affected ordinary life or missing more than 3 months of school in the last 12 months for health reasons at age 10.

[40]We use the stata routine psmatch2 to calculate matching estimates. See Leuven and Sianesi (2003) for details.

motherhood, even after conditioning on becoming pregnant as a teenager. Moreover, the vastly different background characteristics between the birth sample and the abortion and not-pregnant samples, suggest that simply controlling for these characteristics in a linear model may not be sufficient to identify effects for the birth sample. The problem of there being only a narrow region of common support among these background characteristics suggest that a more flexible framework, such as propensity score matching, may be more appropriate. We impose common support through the use of an Epanechnikov kernel matching estimator, as explained below.

Finally, Table 5 indicates that the characteristics of the miscarriage sample lie somewhere between the birth and not-pregnant samples, but closer to the birth sample. This supports the idea that the miscarriage sample comprises a mixture of latent birth type women and latent abortion type women, with a higher proportion of the miscarriage type having a latent-preference for birth.

6 Results

The results from our analysis cover four broad areas – family income, receipt of means-tested benefits, employment and wages and the cohort member's partnership. We also present some results regarding educational attainment, though with some caveats (see below). The aim is to understand how early motherhood affects the mother's socioeconomic status and living standard at age 30 (captured by equivalised family income), the pathways that lead from teenage motherhood to these living standards as well as the extent to which state-provided benefits compensate for any negative effects on wages and partner's income.

Recall too that we are also interested in how the estimates vary across the different methodologies. Specifically, we ask the question of whether imposing common support through propensity score matching results in smaller estimates than conventional linear models and whether the bias in the IV estimates from misreporting and non-randomness can explain their much smaller size. For each outcome we present six sets of estimates. First, we show OLS estimates of the effects of early motherhood for the full sample of females and for the sample of those who became pregnant as a teenager (columns 1 and 2). These are the 'conventional' linear models that control for observed characteristics only.

Next we use propensity score matching to compare outcomes for teenage mothers with similar pregnant teenagers who did not give birth (columns 3 and 4). This also controls for observed characteristics only, but within a more flexible framework that also allows us to impose common support, by restricting the individuals to whom we compare teenage mothers to those with similar background characteristics. We present estimates from Epanechnikov Kernel density matching with two bandwidths, one imposing common support within a propensity score bandwidth of 0.01, the other within a bandwidth of 0.001. The larger bandwidth imposes common support less strictly and so results in more efficient estimates at the cost of potentially more biased estimates.

To give an indication of the success of our attempt to match teenage mothers with similar non-teenage mothers in our sample, Table 6 shows the result from our baseline probit estimation. To conserve on space we only present results for the 20-year definition of a teenager. Other samples give quantitatively similar results. Table 7 gives an indication of the number of observations lost to common support for the same sample and a summary of the balancing tests performed. The median bias amongst the explanatory variables decreases from about 12.5% to 3.2% for the 0.01 bandwidth after matching, and from about 12.5% to 6.5% for the 0.001 bandwidth after matching. The joint test for non-significance of the explanatory variables can also no longer be rejected after matching.

[Table 6 here]

[Table 7 here]

Next, we present the first set of estimates that control for unobserved heterogeneity (column 5). These are the IV estimates, $\beta_{IV}^{\tilde{Z}}$, using miscarriages as an instrument for teenage births. However, we remind the reader that these may be biased towards zero. Accordingly, in column 6, we give estimates for a lower bound of β_{IV}^{*}, accounting for non-randomness and misreporting of miscarriages as per (18). Results are shown for estimates of the bounds where the proportion of non-random miscarriages among the set of reported miscarriages (k_{NR}) is assumed to be 0.15. All regressions include the background variables discussed above as controls. We implement IV estimation by first regressing all variables on our control variables and then working with residuals. Results from the first stage are available from the authors upon request. As would be expected from their binary nature, the correlation of miscarriages with births is significant and negative.

We present results for two definitions of teenagers: Table 8 shows estimates of the effects of motherhood before age 20, and Table 9, before age 18. The outcomes we consider are set out in 5 different panels, covering family income (panel A), state benefits (panel B), labour market outcomes (panel C), partnership status (panel D), and education (panel E).

6.1 Pathways to Disadvantage

The results shown in Table 8 provide an interesting picture of the effects of teenage motherhood on socio-eonomic status of the mother by age 30. Our first observations relate to the insights we can gain about the pathways to possible disadvantage by looking at a range of different mothers' oucomes. Here we focus mostly on what we learn from our OLS and PSM results - we go on to discuss what we learn from adopting the IV methodology in the next section below.

[Table 8 here]

Two factors stand out. First, many of our estimates suggest that teenage motherhood leads to lower needs-adjusted (equivalised) family income by age 30. However all our estimates suggest that this is driven by larger families, not lower total income. Second, many of our estimates (though not the central IV estimates, see below) highlight the importance of the government in insulating teenage mothers from further negative effects: effectively making up for worse labour market outcomes through benefit payments.

Starting with our baseline measure of the overall economic welfare of the teenage mother at age 30, net equivalised weekly family income, we can see that compared to women from the same age group who did not become teenage mothers, our OLS results show teenage mothers with an average 39% lower equivalised family income (column 1). Restricting the comparison group to those who became pregnant but did not give birth as a teenager, this difference is reduced to around 36%. The more flexible PSM approach suggests this difference is smaller again at around 28-31%.

Interestingly, lower living standards appear to be generated by family size differences rather than differences in income levels. This can be seen from the fact that large negative effects are not apparent, even in our OLS estimates, when we consider the impact on *un*equivalised family income. Teenage motherhood results in a cut in unadjusted family income of around just 7 per cent (column 2 OLS esimate), and this estimate is not significantly different from zero. This view is confirmed by considering the impact of teenage motherhood directly on family size and composition variables: our OLS and PSM estimates suggest that teenage motherhood results in an increase in the equivalence scale of around one third (i.e. teenage motherhood means the family requires one third as much income to reach the same standard of living as had the early motherhood not occurred). Moreover, this is not just a function of the particular equivalence scale we use. The fourth row of the same table shows that teenage mothers have an average of 1.2 -1.5 more children in the household by age 30 when compared with similar pregnant teenagers who did not give birth.

However the picture is a more subtle one than simply a bigger family resulting in lower living standards: it is also important to realise that the state plays an important role in maintaining the family incomes of teenage mothers: this can be seen by the fact that the OLS and PSM results suggest significantly worse labour market outcomes for both the mother herself, and for her partner, if present in the household. In particular the mother has a lower probability of being in work, whilst for those who do work, it significantly reduces the number of hours worked. Not surprisingly, these shorter hours mean that teenage motherhood leads to a reduction in weekly earnings. Additionally, Table 8 (Panel C) shows that on all but the central IV estimates, hourly earnings are also significantly reduced (though by less than weekly earnings).

There is little evidence in Panel D Table 8 that teenage motherhood affects the probability of having a partner at 30.[41] This means that, even on our conventional estimates, lone parenthood can be ruled out as an important contributor through which teenage motherhood confers disadvantage at this age. However teenage motherhood is associated with having a partner who is less well qualified, and who has a lower weekly wage compared to those who did

[41] Defined as a cohabitee or legal spouse.

not become mothers as a teenager, when we consider our OLS and matching estimates alone.

Putting all these results together, an interesting picture emerges. Taking our estimates that do not control for unobserved heterogeneity at face value, teenage motherhood does lead to lower living standards at age 30 but does not lower household income, unadjusted for needs. The missing link here is benefits from the state: and here we can see (Panel B of Table 8) that virtually all the negative impact of teenage motherhood on labour market outcomes and the quality of partnership is compensated by a higher likelihood of means-tested support from the state and a higher level of benefit income. Given the earlier results for family income, this evidence seems to suggest that in the UK, the state plays an important role in counteracting any negative labour market effects of teen motherhood. Relative to similar non-teenage mothers, our estimates suggest an that teenage motherhood leads to an increased reliance on the benefit system for support.

6.2 Differences Across Methodologies

Our second set of observations relate to the estimates obtained through different methodologies: in general, when we depart from conventional OLS estimation, we find considerably smaller effects of teenage motherhood on all the outcomes we consider. In particular our IV estimates, controlling for unobserved heterogeneity, suggest zero effects of teenage motherhood on almost all the considered outcomes. However the size of the estimated IV bounds mean that non-randomness and mis-reporting of our instrument could, in the extreme, be entirely responsible for this reduction in the apparent effects.

For example, considering again our baseline indicator of family socio-economic status at age 30, net weekly equivalised household income, we can see from Table 8 that the central IV results in column (5) suggest that unobserved heterogeneity may well be an important factor in driving the negative impacts we discussed in the section above. If these estimates are to be believed, then becoming a teenage mother results in a considerably smaller cut - of around 14% - in equivalised family income at age 30 on this estimate, and this is not significantly different from zero at the 5% level.[42] The estimated IV bound in column (6), however, shows that at the extreme the IV estimate could in fact indicate a cut of 44 per cent in equivalised family income resulting from teenage motherhood - an estimate larger in absolute terms than either the OLS or PSM estimates. This pattern of results across estimation techniques is consistent across almost all the outcomes we consider. On balance, therefore, we are unable to conclude that teenage motherhood does not have strongly negative effects on family income and its components, with the pathways to disadvantage discussed in the section above also consistent with the pattern of results revealed by the IV bounds. But depending upon how much faith we are willing to place in the use of miscarriage

[42]It should be noted, however, that the lack of significance of this estimate is due to the poor precision of our IV estimates, rather than a point estimate particularly close to zero.

as an instrument, our IV results provide evidence that the impact may well not be as negative as the conventional estimates suggest.

6.3 Age at First Birth

Our third set of observations relate to the apparent difference in the effect of teenage motherhood, depending on the age of the mother when she had her first child. Table 9 shows the estimated impact of teenage motherhood on outcomes at age 30 for a subsample who gave birth before the age of 18. In many cases the estimated effects are *less* negative than those found for our main, larger sample, which also contains those who became mothers at age 18 or 19. Two hypotheses could explain this phenomenon. First, this could be taken as evidence that any negative effects of teenage motherhood at age 30, are, at least in part temporary. If the effects of being a teenage mother were permanent and distinctive, we might expect to see larger effects for the under 18 sample in Table 9. Second, this phenomenon could also suggest that the youngest mothers are in general more protected by their families from the negative effects of early motherhood than those who give birth slightly later. These considerations suggest that further work is required in unravelling these possibly competing hypotheses.

[Table 9 here]

[Table 10 here]

6.4 Educational Outcomes

The final set of observations relate to the cohort members' educational attainment. This is likely to be an important mechanism through which teenage motherhood confers later life disadvantage, and one which is likely to create permanent, rather than temporary differences between teenage, and non-teenage mothers. However, establishing a causal interpretation for this is complicated by the fact that the most important decisions made by young people about their education are likely to take place around the same time, or even before the pregnancy and motherhood decisions we are considering. Hence the decision to become a young mother may in part be a direct result of leaving school young, and not the other way round. For this reason, to answer questions about the relationship between teenage motherhood and human capital formation it would be more appropriate to examine a panel data set of teenagers around the school leaving age (such as that in Levine and Painter (2003)) rather than the longitudinal data that we have here. We choose to include these estimates for the sake of completeness and to compliment the existing literature and we ask the reader to interpret the results with caution. The estimates in Panel E of Table 8 show that those who gave birth as teenagers are significantly less likely to go on to post-compulsory education than those who do not. This could be an important mechanism through which teenage childbearing leads to the negative effects that we have already seen, and some of our estimates also suggest a significantly younger age leaving full-time education.

Finally it should be noted that for educational outcomes, there are sizeable differences between the linear regression results for the pregnant sample and the propensity score matching estimates. As in Levine and Painter (2003) we find that the PSM estimates with fairly wide bandwidth (column (3)) are somewhat more modest than linear regression and close to the IV results. Levine and Painter (2003) report that their results are insensitive to the bandwidth used while we find that tightening the bandwidth considerably to 0.001 (column (4)) we estimate much larger effects than linear regression – indeed the PSM estimates are now very close to the linear regression results for the whole sample (column (1)).

7 Conclusions

This paper provides evidence on the effects of teenage motherhood on women's later life outcomes, by considering the impact of becoming a teenage mother on a cohort of British women observed at age 30. In line with the recent literature, we have employed a number of methods to account for both observed and unobserved characteristics influencing selection into teenage motherhood. Our results confirm that estimates are indeed sensitive to the methodology employed. When observed characteristics only are taken into account, the effects of teenage motherhood on a woman's socioeconomic status at age 30 appear to be large and negative. However, it is family size and composition, rather than household income, that appear to be the important drivers of disadvantage. In the UK, it appears that benefit income does a good job of compensating for the negative effects on labour market outcomes and partners' incomes. An interesting question for future research would be to compare this finding with US data.

Our analysis also suggests that once we take unobserved characteristics determining selection into teenage motherhood into account, the evidence for strong negative effects on later life outcomes becomes less clear cut. Our IV results - which exploit data on miscarriages as a source of exogenous variation in teenage motherhood - suggest that many of the negative effects may be significantly reduced or even disappear once such unobserved heterogeneity is taken into account. This is in line with the results from many of the other papers cited in section 3. As in other work the size of our treatment group is a problem that undermines precision – we have only 46 (77) miscarriages by age 18 (20) compared to 353 (794) births.

However there may be biases introduced into the IV estimates as a result of non-random occurrence and misreporting of miscarriages. We have shown how it is possible to extend the methods in Hotz, Mullins and Sanders (1997) to account for misreporting as well as non-randomness of miscarriages. We show that for most outcomes we consider, the apparent lack of strong negative effects using IV could indeed be driven by biases in our IV estimator. This is shown by the fact that our estimates of lower bounds for our IV results are again large and significantly negative, and indeed are broadly in line with the conventional

OLS estimates of the impact of teenage motherhood. This means that we are unable to conclude that the conventional estimates (i.e. those shown in columns 1, 2, and 3 of Tables 8 to 12) could not, in fact, be the true estimates of the impact of teenage motherhood. Rather, a cautious interpretation of our results would conclude that these conventional OLS estimates probably represent the worst possible effects of early motherhood, whereas the IV estimates probably represent the best.

What does all this mean for the policymaker, trying to decide if teenage motherhood is simply an indicator of prior disadvantage, or a pathway to future disadvantage? From the evidence in this paper alone, we cannot rule out that it is teenage motherhood that leads to lower socioeconomic status in later life, rather than earlier disadvantage alone. However, our own IV estimates – though potentially biased – do add to the growing body of evidence, amassed using a variety of different, and all imperfect, methods, which suggests that the importance of teenage motherhood may, in fact, be small compared to the role that prior disadvantage plays.

Our results also shed some light on some other important issues. First we have shown some of the contributing factors to the lower economic status at age 30 experienced by teenage mothers. In particular, those who become teenage mothers are less likely to be in work, work fewer hours, and earn a lower hourly wage than those who do not. There is no difference in the likelihood of having a partner, but the partners of teenage mothers have lower educational qualifications and labour market status than the partners of those who did not become teenage mothers. We have also shown that teenage mothers' families have greater needs - as determined by their family size and composition - for any given level of income.

Our analysis has also highlighted the importance of disentangling timing issues from any long-term permanent disadvantage that might be incurred by teenage motherhood. For example, we showed that for most outcomes, the effects of early motherhood at age 30 are larger for females falling pregnant between 18 and 20 years old than those falling pregnant before age 18. One explanation is that the effects of early pregnancy diminish over time and do not persist into later periods in life. This could be because teenage motherhood tends to bring forward in time some of the disadvantage incurred by most mothers when they raise children. Another explanation could be that for those who give birth at a younger age, the young mother's own family typically provides more support, and protects the teenager from some of the more negative effects of becoming a mother at a young age. More research – following individuals over a longer timespan – is required to ascertain the extent to which more permanent disadvantage also ensues.

References

[1] Brien, M.J., Loya, G.E. and Pepper, J.V. (2002), *Teenage Childbearing and Cognitive Development*, Journal of Population Economics, 15(3), pp.

391-416

[2] Bronars, S.G. and Grogger, J. (1994), *The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment*, American Economic Review, 84(5), pp. 1141-1156

[3] Chevalier, A. and Viitanen, T.K. (2003), *The Long-Run Labour Market Consequences of Teenage Motherhood in Britain*, Journal of Population Economics, 16(2), pp. 323-343

[4] Ermisch, J. (2004), *Early Motherhood and Later Partnerships*, Journal of Population Economics *forthcoming*

[5] Geronimus, A.T. and Korenman, S. (1992), *Accounting for No-Shows in Experimental Evaluation Designs*, Quarterly Journal of Economics, 107(4), pp. 1187-1214

[6] Hobcraft, J. and Kiernan, K. (2001), *Childhood Poverty, Early Motherhood and Adult Social Exclusion*, British Journal of Sociology, 52(3), pp. 495-517.

[7] Hofferth, S.L. and Moore, K.A. (1979), *Early Childbearing and Later Economic Well-Being*, American Sociological Review, 44(5), pp. 784-815

[8] Hoffman, S.D., Foster, E.M. and Furstenberg Jr, F.F. (1993a), *Reevalutaing the Costs of Teenage Childbearing*, Demography, 30(1), pp. 1-13

[9] Hoffman, S.D., Foster, E.M. and Furstenberg Jr, F.F. (1993b), *Reevalutaing the Costs of Teenage Childbearing: Response to Geronimus and Korenman*, Demography, 30(2), pp. 291-296

[10] 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), 236–239.

[11] Horowitz, J.L. and Manski, C.F.(1995), *Identification and Robustness with Contaminated and Corrupted Data*, Econometrica, 63(2), pp. 281-302

[12] Hotz, V.J., Mullin, C.H. and Sanders, S.G. (1997), *Bounding Causal Effects Using Data From a Contaminated Natural Experiment: An Analysis of the Effects of Teenage Childbearing*, Review of Economic Studies, 64(4), pp. 575-603

[13] Hotz, V.J., McElroy, S. and Sanders, S.G. (1997), *Teenage Childbearing and its Life Cycle Consequences: Exploiting a Natural Experiment*, NBER Working Paper, No. 7397.

[14] Klepinger, K., Lundberg, S. and Plotnick, R. (1998), *How Does Adolescent Fertility Affect the Human Capital and Wages of Young Women?*, Journal of Human Resources, 34(3), pp. 23-28

[15] Leuvin, E. and Sianesi (2003), *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing*, http://ideas.repec.org/c/boc/bocode/s432001.html. Version 1.2.3.

[16] Levine, D.I. and Painter, G. (2003), *The Costs of Teenage Out-of-Wedlock Childbearing: Analysis with a Within-School Propensity Score Matching Estimator*, Review of Economics and Statistics, 85(4) pp. 884–900

[17] McClements, L. (1977), *Equivalence Scales for Children,* Journal of Public Economics, 8(2), pp. 191-210

[18] Regan, L. (2001), *Miscarriage: What Every Women Needs to Know*, Orion Books Ltd, where?

[19] Ribar, D.C. (1999), *The Socioeconomic Consequences of Young Women's Childbearing: Reconciling Disparate Evidence*, Journal of Population Economics, 12(4), pp. 547-565.

[20] Robinson, H. (2002), *My Generation: The Changing Penalty of Teenage Motherhood,* Cardif Business School, Discussion Paper.

[21] Social Exclusion Unit (1999) *Teenage Pregnancy*, CM 4342, London, TSO.

[22] Trussell, T.J. (1976), *Economic Consequences of Teenage Childbearing*, Family Planning Perspectives, 8(4), pp. 184-190

[23] UNICEF (2001), *A League Table of Teenage Births in Rich Nations*, Innocenti Report Card No.3, UNICEF Innocenti Research Centre, Florence.

Figure 1 Pregancy, Misscarriage, Abortion and Birth

$$
\begin{tikzcd}
 & & & \mathit{Pregnancy} \arrow[dll] \arrow[drr] & & & \\
 & Z^* = 0 \arrow[dl] \arrow[dr] & & & & \begin{array}{c} Z^* = 1 \\ \mathit{Random Miscarriage} \end{array} \arrow[dl] \arrow[dr] & \\
\begin{array}{c} D = 0 \\ \mathit{Abortion} \end{array} \arrow[d] & & \begin{array}{c} D = 1 \\ \mathit{Birth} \end{array} \arrow[d] & & \begin{array}{c} D^L = 0 \\ (D = 0) \end{array} \arrow[d] & & \begin{array}{c} D^L = 1 \\ (D = 0) \end{array} \arrow[d] \\
Y_0 & & Y_1 & & Y_0 & & Y_0
\end{tikzcd}
$$

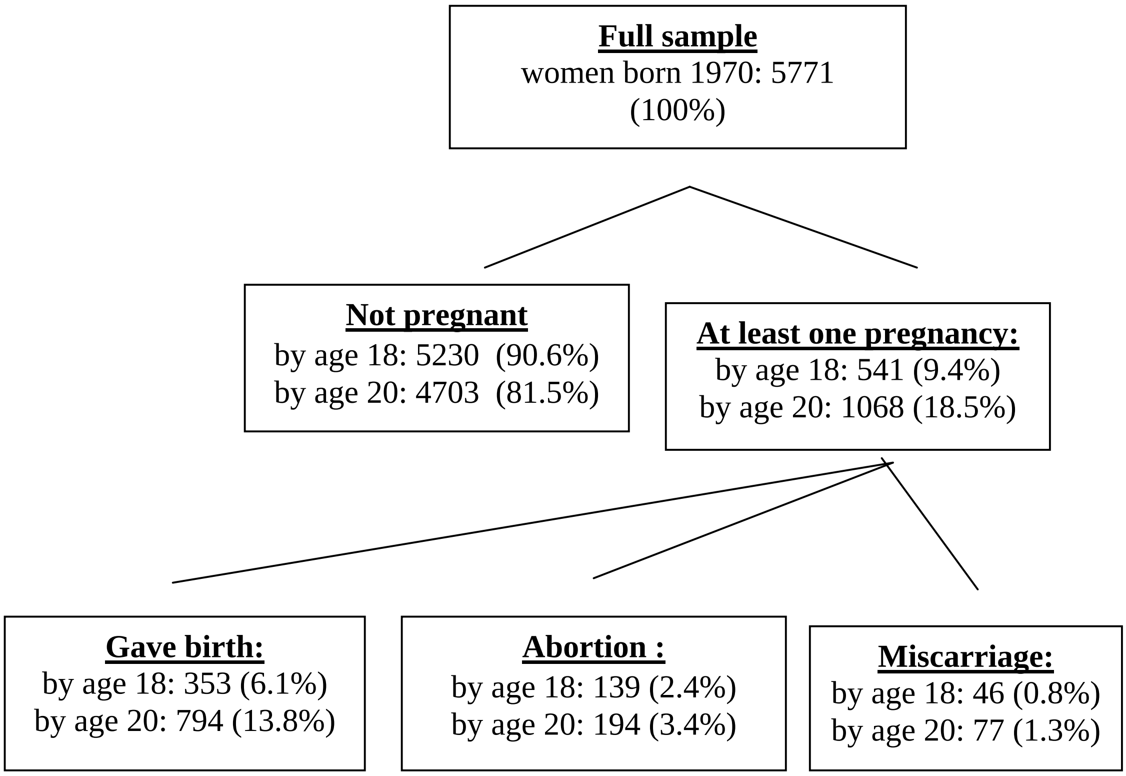
Table 1: Distribution of Number of Pregnancies

Number of Pregnancies	**by age 18**		**by age 20**	
	Number	**Percentage**	**Number**	**Percentage**
0	5230	90.6	4703	81.5
1	469	8.1	786	13.6
2	65	1.1	239	4.1
3	6	0.1	39	0.7
4	1	0.0	4	0.1
Total	5771	100	5771	100

Table 2: Other Pregnancies for Teenagers who Miscarried

	Miscarriage by 18	**%**	**Miscarriage by 20**	**%**
Also gave birth	18	26.9	63	41.7
Also had abortion	2	3.0	6	4.0
Also gave birth and had abortion	1	1.5	4	2.6
Only had miscarriages	46	68.7	77	51.0
Total	67	100	151	100

Figure 2: Sample sizes



Note: the Pregnant sample includes 3 more women in total than the birth, miscarriage, and abortion samples combined because of the stillbirths discussed in footnote 17.

Table 3: Fertility rates for 1970 Cohort, aged 15-19

	ONS per 1000	**BCS70 Per 1000**	**ONS (%)**	**BCS70 (%)**
Births	152	114	59-60	67
Abortions	74	36	29-30 [a]	21
Miscarriages	25-31	24	10-12[b]	12
	251-257	176	100	100

Notes: Birth rates refer to the number of registered live births in England and Wales. Abortion rates refer to the number of recorded abortions in England and Wales.

a: Abortion rates for the 1970 cohort are not available directly. The earliest year for which age-specific ONS abortion data is available is 1991, when the 1970 cohort would have been 21 years old. To calculate the abortion rates in this table, we use information on abortion rates of women from more recent cohorts, applying the percentage of conceptions (births and abortions) ending in abortion for each age 15-19, averaged over the years 1991-95, to the relevant birth rates for the 1970 cohort.

b: by assumption – see text above.

Sources: ONS Series FM1 no. 30 (revised) Table 10.1. and Table 12.2 and authors' calculations.

Table 4: Summary statistics – Outcome variables, 20 year definition of teenager

Outcome	Full Sample	Not Pregnant	Pregnant	Birth	Abortion	MisCarriage
Family Income						
Log Equivalised	5.76	5.84	5.41	5.30	5.84	5.58
Family Income	(0.77)	(0.77)	(0.65)	(0.60)	(0.66)	(0.74)
	5515	4489	1026	768	181	74
Log Family	5.78	5.81	5.63	5.59	5.81	5.63
Income	(0.77)	(0.79)	(0.66)	(0.64)	(0.71)	(0.75)
	5515	4489	1026	768	181	74
McClements	1.06	1.00	1.29	1.38	1.00	1.10
Equivalence	(0.30)	(0.26)	(0.34)	(0.30)	(0.29)	(0.29)
Scale[a]	5771	4703	1068	794	194	77
Number of children	0.95	0.71	1.99	2.35	0.83	1.12
	(1.07)	(0.90)	(1.16)	(0.99)	(0.91)	(1.01)
	5771	4706	1065	794	194	77
Work						
In Work?	0.68	0.72	0.51	0.46	0.66	0.69
	(0.47)	(0.45)	(0.50)	(0.50)	(0.48)	(0.47)
	5771	4703	1068	794	194	77
Log Weekly	5.14	5.21	4.73	4.56	5.21	4.78
Wage	(0.79)	(0.77)	(0.80)	(0.73)	(0.74)	(0.93)
	3907	3360	547	365	128	53
Log Hourly	1.73	1.76	1.56	1.49	1.77	1.55
Wage	(0.41)	(0.41)	(0.42)	(0.39)	(0.40)	(0.54)
	3891	3348	543	361	128	53
Hours Worked	35.15	36.10	29.25	26.53	35.66	32.32
per Week	(12.97)	(12.56)	(13.85)	(12.92)	(13.97)	(14.35)
	3938	3389	549	366	129	53
Partner						
Partner	0.70	0.71	0.69	0.70	0.63	0.69
in Household?	(0.46)	(0.45)	(0.46)	(0.46)	(0.48)	(0.47)
	5771	4703	1068	794	194	77
Log Weekly	5.64	5.66	5.51	5.45	5.65	5.66
Wage	(0.65)	(0.64)	(0.71)	(0.73)	(0.68)	(0.59)
	3372	2810	562	406	110	44
Post-	0.57	0.60	0.48	0.44	0.60	0.55
Compulsory	(0.49)	(0.49)	(0.50)	(0.50)	(0.49)	(0.50)
Schooling?	5771	4703	1068	794	194	77
Benefit variables						
Log Weekly	3.69	3.49	4.17	4.25	3.77	3.79
Benefit	(1.02)	(0.97)	(0.99)	(0.98)	(0.96)	(0.92)
Income	3266	2329	937	772	109	54
On Means-	0.79	0.85	0.50	0.42	0.77	0.69
Tested	(0.41)	(0.35)	(0.50)	(0.49)	(0.42)	(0.47)
Benefits?	5757	4689	1068	794	194	77
Education						
Age Left	17.48	17.72	16.48	16.26	17.28	16.78
Full-Time	(2.26)	(2.35)	(1.43)	(1.09)	(2.20)	(1.51)
Education	5607	4552	1055	791	185	76
Post-	0.50	0.56	0.25	0.19	0.48	0.31
Compulsory	(0.50)	(0.50)	(0.43)	(0.39)	(0.50)	(0.47)
Schooling?	5771	4703	1068	794	194	77

[a] Equivalence scales provide the means of adjusting a household's income for size and composition so that incomes can be sensibly compared across different households. Official income statistics use the McClements (1977) equivalence scale, in which an adult couple with no dependent children is taken as the benchmark with an equivalence scale of one. The equivalence scales for other types of households can be calculated by adding together the implied contributions of each household member . The scale used is: Head, 0.61; Partner/Spouse, 0.39; Other second adult, 0.46; Third adult, 0.42; Subsequent adults, 0.36; Each child aged 0-1, 0.09; Each child aged 2-4, 0.18; Each child aged 5-7, 0.21; Each child aged 8-10, 0.23; Each child aged 11-12, 0.25; Each child aged 13-15, 0.27; Each child aged 16-18, 0.36.

Table 5: Summary Statistics – Background variable, 20 year definition of a teenager

Background variables	**Full**	**Not pregnant**	**Pregnant**	**Birth**	**Abortion**	**Miscarriage**
Age father left	16.00	16.12	15.45	15.29	16.07	15.55
FT education	(2.25)	(2.35)	(1.56)	(1.32)	(2.25)	(1.39)
	5208	4270	938	695	167	73
Age mother left	15.74	15.83	15.35	15.24	15.78	15.42
FT education	(1.65)	(1.73)	(1.12)	(0.93)	(1.58)	(1.30)
	5208	4273	935	692	167	73
Maths Score	61.77	63.08	56.10	54.69	61.15	58.08
Age 10	(16.17)	(15.91)	(16.05)	(16.07)	(15.91)	(14.02)
	4327	3513	814	601	144	66
Reading Score	63.28	64.87	56.42	53.63	65.21	61.84
Age 10	(19.44)	(18.94)	(20.05)	(19.69)	(19.16)	(18.96)
	4651	3773	878	646	160	69
Ability Score	52.98	54.06	48.31	46.62	53.94	51.34
Age 10	(13.34)	(13.13)	(13.26)	(12.95)	(13.04)	(12.96)
	4563	3703	860	634	153	70
Mother's Age	25.97	26.20	24.95	24.67	25.97	25.33
at birth	(5.35)	(5.24)	(5.74)	(5.75)	(5.12)	(6.78)
Father's class:						
- I	6%	7%	3%	2%	7%	5%
- II	24%	26%	15%	12%	28%	16%
- III.manual	9%	10%	7%	6%	12%	5%
- III.nonmanual	44%	43%	51%	52%	43%	61%
- IV	12%	11%	17%	20%	7%	9%
- V	4%	3%	7%	8%	3%	4%
Income at 10:						
<£50pw	6%	5%	10%	12%	4%	10%
£50-£100	30%	27%	39%	43%	26%	31%
£100-£150	35%	35%	32%	31%	36%	34%
£150-£200	16%	17%	13%	10%	21%	16%
>£200pw	13%	15%	6%	4%	12%	9%
Income at 16:						
<£100pw	16%	12%	31%	35%	18%	19%
£100-£150	14%	14%	16%	17%	12%	21%
£150-£200	15%	15%	14%	15%	9%	17%
£200-£250	12%	13%	8%	8%	7%	7%
£250-£300	10%	10%	7%	6%	14%	7%
£300-£350	6%	6%	4%	4%	5%	2%
>£350pw	12%	13%	7%	5%	15%	7%

Table 7: Propensity Score Balancing Tests and Diagnostics

A: Balancing Tests

	Median Bias		**p-value from LR Test of No Effect of Explanatory Variables**	
Bandwidth	**(unmatched)**	**(matched)**	**(unmatched)**	**(matched)**
0.01	12.49	3.20	0.000	0.924
0.001	12.49	6.53	0.000	0.032

B: Imposition of Common Support

Bandwidth = 0.01

	On Support	**Off Support**	**Total**
No Birth	255	0	255
Birth	766	2	768
	1,021	2	1,023

Bandwidth = 0.001

	On Support	**Off Support**	**Total**
No Birth	255	0	255
Birth	406	362	768
	661	362	1,023

[a] Results above are for the subsample for which data was available for Family Income, our baseline outcome of interest. Sample sizes vary slightly for other outcomes due to missing data. Results above are for the 20-year definition of teenager. Other samples give very similar results.

Table 8. Impact of teenage motherhood - 20 yr definition

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	**OLS**	**PSM**	**PSM**	**IV**	**IV - Bound**
			bw = 0.01	bw=0.001		85% random
	Full Sample	Preg Sample	Preg Sample	Preg Sample	Preg Sample	Preg Sample
Panel A - Family Income and Composition Variables						
Log Equivalised Family Income	-0.390	-0.359	-0.275	-0.310	-0.138	-0.447
	(0.025)	(0.049)	(0.083)	(0.084)	(0.107)	(0.083)
Log Family Income	-0.090	-0.067	0.053	-0.004	0.067	-0.220
	(0.027)	(0.054)	(0.083)	(0.091)	(0.114)	(0.093)
McClements Equivalence Scale	0.345	0.338	0.352	0.326	0.240	0.359
	(0.012)	(0.022)	(0.029)	(0.040)	(0.044)	(0.043)
Number of Children	1.487	1.340	1.435	1.247	1.060	1.566
	(0.039)	(0.072)	(0.100)	(0.128)	(0.164)	(0.166)
Panel B - Benenfit Variables						
Log Weekly Benefit Income	0.356	0.257	0.225	0.249	0.200	0.645
	(0.019)	(0.034)	(0.124)	(0.141)	(0.149)	(0.150)
On Means Tested Benefits?	0.146	0.173	0.119	0.114	0.149	0.369
	(0.040)	(0.067)	(0.049)	(0.058)	(0.075)	(0.078)
Panel C - Wage Variables						
In work?	-0.200	-0.133	-0.206	-0.193	-0.173	-0.453
	(0.020)	(0.036)	(0.053)	(0.064)	(0.072)	(0.075)
Log Weekly Wage	-0.490	-0.368	-0.419	-0.343	0.008	-0.509
	(0.040)	(0.075)	(0.118)	(0.185)	(0.185)	(0.157)
Log Hourly Wage	-0.174	-0.141	-0.181	-0.202	0.071	-0.166
	(0.022)	(0.045)	(0.085)	(0.105)	(0.116)	(0.078)
Hours Worked per Week	-7.756	-5.946	-7.875	-7.372	-3.469	-10.077
	(0.717)	(1.327)	(2.075)	(3.345)	(2.857)	(2.630)
Panel D - Partner Variables						
Partner in Household?	0.000	0.068	-0.001	-0.005	0.015	-0.113
	(0.018)	(0.035)	(0.053)	(0.060)	(0.074)	(0.078)
Log Partner's Weekly Wage	-0.146	-0.173	-0.119	-0.114	-0.149	-0.365
	(0.040)	(0.067)	(0.087)	(0.163)	(0.126)	(0.111)
Partner Post-Compulsory Schooling?	-0.096	-0.138	-0.094	-0.088	-0.095	-0.221
	(0.020)	(0.036)	(0.054)	(0.064)	(0.075)	(0.079)
Panel E - Education Variables						
Age Left Full-Time Education	-0.663	-0.415	-0.201	-0.648	-0.142	-0.620
	(0.052)	(0.104)	(0.114)	(0.148)	(0.193)	(0.194)
Post-Compulsory Schooling?	-0.205	-0.115	-0.128	-0.199	-0.010	-0.112
	(0.016)	(0.031)	(0.047)	(0.057)	(0.061)	(0.064)

Table 9. Impact of teenage motherhood - 18 yr definition

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	**OLS**	**PSM**	**PSM**	**IV**	**IV - Bound**
			bw = 0.01	bw=0.001		85% random
	Full Sample	Preg Sample	Preg Sample	Preg Sample	Preg Sample	Preg Sample
Panel A - Family Income and Composition Variables						
Log Equivalised Family Income	-0.389	-0.289	-0.275	-0.277	-0.050	-0.378
	(0.032)	(0.053)	(0.081)	(0.139)	(0.141)	(0.115)
Log Family Income	-0.087	-0.029	-0.006	-0.032	0.115	-0.192
	(0.033)	(0.058)	(0.090)	(0.143)	(0.151)	(0.125)
McClements Equivalence Scale	0.353	0.296	0.312	0.314	0.177	0.340
	(0.018)	(0.032)	(0.048)	(0.078)	(0.087)	(0.081)
Number of Children	1.461	1.092	1.435	1.247	0.839	1.590
	(0.039)	(0.072)	(0.155)	(0.258)	(0.287)	(0.254)
Panel B - Benenfit Variables						
Log Weekly Benefit Income	0.339	0.182	0.157	0.101	0.075	0.553
	(0.026)	(0.046)	(0.155)	(0.284)	(0.244)	(0.229)
On Means Tested Benefits?	0.274	0.248	0.199	-0.089	0.238	0.240
	(0.067)	(0.105)	(0.066)	(0.116)	(0.114)	(0.117)
Panel C - Wage Variables						
In work?	-0.184	-0.082	-0.183	-0.194	-0.173	-0.105
	(0.027)	(0.047)	(0.069)	(0.116)	(0.105)	(0.108)
Log Weekly Wage	-0.464	-0.315	-0.277	-0.323	-0.125	-0.503
	(0.060)	(0.091)	(0.165)	(0.427)	(0.298)	(0.249)
Log Hourly Wage	-0.196	-0.099	-0.071	-0.048	-0.040	-0.180
	(0.029)	(0.051)	(0.107)	(0.221)	(0.157)	(0.152)
Hours Worked per Week	-5.820	-4.084	-5.601	-5.893	-2.999	-8.630
	(1.128)	(1.805)	(3.147)	(8.347)	(4.690)	(4.424)
Panel D - Partner Variables						
Partner in Household?	0.003	0.072	-0.004	-0.002	-0.049	-0.195
	(0.026)	(0.045)	(0.068)	(0.111)	(0.113)	(0.116)
Log Partner's Weekly Wage	-0.274	-0.248	-0.199	0.089	-0.238	-0.509
	(0.067)	(0.105)	(0.174)	(0.439)	(0.270)	(0.270)
Partner Post-Compulsory Schooling?	-0.087	-0.064	-0.085	-0.092	0.051	-0.068
	(0.028)	(0.047)	(0.070)	(0.121)	(0.114)	(0.118)
Panel E - Education Variables						
Age Left Full-Time Education	-0.685	-0.318	-0.145	-0.686	0.190	-0.239
	(0.062)	(0.115)	(0.146)	(0.255)	(0.191)	(0.187)
Post-Compulsory Schooling?	-0.210	-0.108	-0.087	-0.204	0.107	0.020
	(0.021)	(0.038)	(0.054)	(0.093)	(0.073)	(0.074)

Table 10. Impact of teenage motherhood - 18-20 yr definition

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	**OLS**	**PSM**	**PSM**	**IV**	**IV - Bound**
			bw = 0.01	bw=0.001		85% random
	Full Sample	Preg Sample	Preg Sample	Preg Sample	Preg Sample	Preg Sample
Panel A - Family Income and Composition Variables						
Log Equivalised Family Income	-0.311	-0.418	-0.340	-0.433	-0.208	-0.502
	(0.033)	(0.073)	(0.126)	(0.177)	(0.136)	(0.121)
Log Family Income	-0.074	-0.165	-0.108	-0.187	-0.077	-0.349
	(0.035)	(0.081)	(0.137)	(0.193)	(0.141)	(0.130)
McClements Equivalence Scale	0.266	0.296	0.322	0.222	0.164	0.295
	(0.015)	(0.029)	(0.048)	(0.080)	(0.053)	(0.051)
Number of Children	1.201	1.276	1.177	1.194	0.847	1.370
	(0.475)	(0.098)	(0.167)	(0.260)	(0.187)	(0.180)
Panel B - Benenfit Variables						
Log Weekly Benefit Income	0.296	0.221	0.221	0.242	0.151	0.581
	(0.025)	(0.051)	(0.231)	(0.345)	(0.201)	(0.205)
On Means Tested Benefits?	0.034	0.193	0.303	0.283	0.175	0.319
	(0.045)	(0.091)	(0.093)	(0.138)	(0.097)	(0.100)
Panel C - Wage Variables						
In work?	-0.172	-0.177	-0.165	-0.173	-0.238	-0.388
	(0.025)	(0.051)	(0.092)	(0.136)	(3.534)	(0.000)
Log Weekly Wage	-0.448	-0.385	-0.348	-0.483	0.011	-0.571
	(0.050)	(0.110)	(0.321)	(0.511)	(0.232)	(0.195)
Log Hourly Wage	-0.136	-0.093	0.047	-0.059	0.083	-0.195
	(0.030)	(0.079)	(0.341)	(0.421)	(0.166)	(0.112)
Hours Worked per Week	-8.279	-6.413	-8.668	-8.445	-0.932	-8.157
	(0.853)	(1.834)	(4.651)	(8.325)	(3.665)	(3.534)
Panel D - Partner Variables						
Partner in Household?	-0.002	0.026	-0.003	0.004	-0.008	-0.140
	(0.023)	(0.050)	(0.083)	(0.132)	(0.091)	(0.094)
Log Partner's Weekly Wage	-0.034	-0.193	-0.303	-0.283	-0.175	-0.423
	(0.045)	(0.091)	(0.148)	(0.370)	(0.143)	(0.139)
Partner Post-Compulsory Schooling?	-0.084	-0.155	-0.084	-0.085	-0.070	-0.205
	(0.025)	(0.052)	(0.097)	(0.136)	(0.102)	(0.106)
Panel E - Education Variables						
Age Left Full-Time Education	-0.506	-0.357	-0.093	-0.531	-0.160	-0.696
	(0.066)	(0.156)	(0.218)	(0.388)	(0.275)	(0.279)
Post-Compulsory Schooling?	-0.159	-0.063	-0.109	-0.158	-0.012	-0.121
	(0.020)	(0.045)	(0.075)	(0.125)	(0.083)	(0.086)