The impact of young motherhood on education, employment and marriage

Bruce Bradbury

University of New South Wales

September 2006

Social Policy Research Centre

THE IMPACT OF YOUNG MOTHERHOOD ON EDUCATION, EMPLOYMENT AND MARRIAGE

By Bruce Bradbury

SPRC Discussion Paper No. 148
September 2006

Published by
The Social Policy Research Centre
University of New South Wales
Sydney NSW 2052
Australia

ISSN: 1447-8978
ISBN: 978 0 7334 2408 3

Author's Note:
This research has been funded by the Commonwealth Department of Families, Community Services and Indigenous Affairs (FaCSIA). The author would like to thank Bob Gregory, Christina Lee, Deborah Loxton, FaCSIA staff and seminar participants in Syracuse, Canberra and Adelaide for comments on earlier versions of the paper. The paper relies on data collected as part of the Australian Longitudinal Survey on Women's Health. This survey is funded by the Australian Government Department of Health and Ageing and conducted by a team of researchers at the Universities of Newcastle and Queensland. I am grateful to the research team for permission to use the data, and to the anonymous women who participated in the study.

Correspondence to:
b.bradbury@unsw.edu.au

Abstract

The poor socio-economic outcomes of women who have their first child when young are well documented. However, the policy implications of this association depend upon the causal mechanisms that underlie it. Recent studies in the US and UK have used miscarriage as an instrument to identify the direct causal impact of young childbearing – with US research suggesting that early child-bearing may even have a beneficial impact upon mother's outcomes. This paper uses this method to examine this issue for a new Australian panel of young women.

No evidence is found for an adverse impact of young childbirth on education, labour market, income or location. Instead these outcomes follow the patterns that might be expected on the basis of selection effects. On the other hand, young motherhood does have an impact on partnering outcomes. Being a young mother reduces the likelihood of being legally married (instead of defacto partnered) when aged in the late 20s. Also, having a child in the early rather than late 20s leads to a greater likelihood of being a lone parent at around age 30.

1 Introduction

As in many other countries, Australian women who have their first childbirth when young have significantly poorer socio-economic outcomes than women who delay child-rearing. They are more likely to be receiving income support payments, have poorer educational outcomes, are less likely to be partnered once they reach their 30s, and if partnered, their partner has a lower income (Bradbury, 2006).

Though Australian teenage birth rates are now lower than in other English-speaking countries, most non-English speaking OECD countries have even lower rates.[1] Moreover, the correlations with socio-economic disadvantage described above also occur for first mothers aged in their early 20s.

However, these associations can arise either because having a child when young directly causes socio-economic disadvantage (a direct effect) or because women from a disadvantaged background are more likely to have their children when young and are also more likely to be disadvantaged later on (a selection effect). Plausible explanations for both patterns can be found. Women from disadvantaged families do tend to have their first birth earlier (eg Stewart, 2003; Wolfe, Wilson and Haveman, 2001). In the literature, this has been described as arising from either rational decision-making – based on their relatively poor expected returns in the labour market, or in terms of the constraints on their ability to control their fertility – eg unwanted sex and access to contraception and abortion services.

For policy purposes it is important to identify the relative importance of these direct and selection effects. In general, there are three broad categories of policy intervention that might be used to ameliorate the poor socio-economic outcomes of young mothers and their children.

- Measures to discourage women from having their children when young. Examples include sex and (anti-)parenting education (of both young men and women), and increased access to contraception and termination services in rural areas.

- Providing additional support to women who do have their children when young. Eg income support, parenting training, specialised childcare services etc.

- Providing additional support to women who are likely to have their child when young. That is, increasing the opportunities for all disadvantaged young women.

If the selection effect is important but the direct effect is not, then the first of these policy interventions will have little impact on outcomes for mothers. It will simply lead to women having their children a little later but with their socio-economic outcomes unchanged.

On the other hand, if the direct effect is important, then providing additional support to young mothers may actually lead to worse outcomes for mothers if it encourages more of them to have their children when young. (The third policy option can be justified under both causal hypotheses). Hence it is important to know whether there is actually a direct effect of early child childbearing on outcomes. This paper looks at outcomes for mothers.

1 The Australian rate is 18 per 1,000 compared to 52 in the US, 31 in the UK and 20 in Canada. Rates in other countries are Germany 13, France 9, Sweden 7, Japan 5 (UNICEF, 2001, data for 1998).

From a physiological perspective, having a child at 18 or 19 may be better than having one at 35. To the extent to which disadvantages flow from the event of being a young mother, these are most likely to arise in the socio-economic sphere. As UNICEF notes,

> Teenage parenthood has come to be regarded as a significant disadvantage in a world which increasingly demands an extended education, and in which delayed childbearing, smaller families, two income households, and careers for women are increasingly becoming the norm (UNICEF, 2001, pp5-6).

To estimate the impact of young motherhood (as opposed to its correlates), it is necessary to identify a group of women who are similar to the young mothers in all relevant respects except for their motherhood status.

One method is to use a sample of women who experience a miscarriage within the same age window as the comparison group. Miscarriage is (largely) random with respect to socio-economic outcome variables and so, even though not all non-miscarriages lead to births (because of abortions), this is a suitable instrument for identifying the impact of childbirth.[2]

This technique has recently been employed by Hotz et al (1997 and 2005) for the US and Ermisch and Pevalin (2003, 2005) for the UK. The US studies find that becoming a mother at a young age (under 18) actually increases the later earnings of mothers compared to what they would have been if they had not had a child. The hours of work of young mothers are higher and their wage rates are the same. Part of the hours of work effect might be due to the fact that women with earlier first childbirths tend to have older children in their late 20s. Nonetheless, the lack of negative outcomes lead Hotz et al (2005, p25) to conclude that this research 'casts doubt on the view that postponing childbearing will improve the socioeconomic attainment of teen mothers in any substantial way'.

Ermisch and Pevalin in the UK do not find any negative education or labour outcomes of having a teenage birth, but do find a significant negative impact upon the mothers' success in the 'marriage market'. Women who were teenage mothers are more likely to have unemployed or low-income husbands at age 30.

This paper employs similar methods to assess the impact of young motherhood in the Australian environment. The paper uses data from the Australian Longitudinal Study of Women's Health. The survey is introduced in the next section, with initial descriptions of the association between pregnancy experiences and outcomes shown in Section 3. Section 4 discusses the use of miscarriage as an identification device, Section 5 the empirical strategy and Section 6 describes the results.

No evidence is found for an adverse impact of young childbirth on education, labour market, income or locational outcomes. Instead these outcomes largely follow the patterns that might be expected on the basis of selection effects. On the other hand, young motherhood does have an impact on partnering outcomes. Being a young mother reduces the likelihood of being legally (rather than defacto) married when aged in the late 20s. Also, having a child in the early rather than late 20s leads to a greater likelihood of being a lone parent at around age 30.

---

2 Other approaches include comparisons of siblings and the use of cross-state policy variation as instruments (see Hoffman, 1998 and Holmlund, 2005, for recent surveys).

2 The Australian Longitudinal Survey of Women's Health

The Australian Longitudinal Study on Women's Health (ALSWH) is a large longitudinal postal survey of a nationally representative sample of Australian women. An overview of the study can be found in Lee et al (2005) and more information can be found at www.newcastle.edu.au/centre/wha. The study includes three samples covering different age groups. This report only uses data from the Young sample, women aged 18 to 23 years in April 1996.

The initial sample was drawn randomly from the Australian Medicare[3] database, with women in rural and remote areas having twice the probability of selection. The first questionnaire was sent out in mid-July 1996. Just under 15,000 of the young women responded.[4] This represents around 41 to 42 per cent of the initial sample of women to whom letters were sent out. Though a good response for this type of survey, this low response rate needs to be considered when interpreting results.[5]

Wave 2 of the survey was administered in 2000 and Wave 3 in 2003 (when the women were aged 25 to 30 years old). By Wave 3, the sample size was only 9,074 – about 25 per cent of the original randomly selected sample. A comparison with Census data in Waves 1 and 2 suggests an under-representation of women with lower education levels and of immigrant women – as might be expected from a self-completion postal survey.

Since the main variables used in this study are those from the first and third waves, the starting population for analysis is women who responded to these two surveys and who responded to the main fertility questions – a sample size of 8,862. Two populations are considered. The first is all women in this overall sample (*All Women* sample). The second population is women who experienced at least one pregnancy prior to wave 1 – the *Pregnant Women* sample.

2.1 Fertility Information

In wave 1 the respondents were asked how many times they had had a miscarriage, had a termination or given birth to a child. The number of pregnancies completed prior to wave 1 is defined as the sum of these three variables. Pregnancies that were still underway (ie if the woman was pregnant at the time of survey completion) are not counted.[6]

These data are self-reported rather than based on medical records and are subject to an unknown degree of reporting error. It is possible, for example that terminations might not be reported and early miscarriages might not have been recognised as such. There was no explicit question in wave 1 about still-births and so it is possible that these might be omitted

---

3 Medicare is the Australian universal health insurance system.

4 For details of the sample, see http://www.newcastle.edu.au/centre/wha/Project/sample.html

5 Some of the apparent non-response was due to address errors in the Medicare database. Due to privacy restrictions it was not possible to use common methods such as telephone follow-ups to increase the response rate.

6 The survey also directly asked how many times the woman had been pregnant. In 3 per cent of cases, this is not consistent with the measure used here. An analysis of these discrepancies is available from the author, where it is concluded that the calculated variable used here is probably the better measure.

from all three categories. Other than this, it is reasonable to expect that births would be well reported.[7]

In the All Women sample, 87 per cent had no pregnancies prior to wave 1 (when they were aged between 18 and 24), 8.9 per cent had one pregnancy, 2.6 per cent had two pregnancies, 0.9 per cent had three, and 0.5 per cent had four or more.

Some 5.2 per cent of women had one birth, and 1.7 per cent had two or more. Of all the pregnancies, half ended in births, one-third in terminations and the remainder as miscarriages. These proportions are based on unweighted data and should not be considered accurate estimates of population rates. (The design over-representation of rural women biases the birth rate up, but the non-response over-representation of high-education women biases it down).

Table 1 summarises the combinations of pregnancy outcomes found in the sample. Just over 600 had experienced at least one childbirth (plus possibly a miscarriage or termination), 361 reported having a termination only, and 125 a miscarriage only. A small number (28) reported having both miscarriage and termination (but not childbirth).

Table 1 also shows the mean age of each group (as at end 1996). The no pregnancy and miscarriage only group were slightly younger than the other groups.

Table 1 Pregnancy experiences prior to wave 1

Number of			Summary title	Unweighted N	Mean age as at end 1996 (weighted)
Child-births	Mis-carriages	Terminat-ions			
0	0	0	No completed pregnancy	7,723	20.7
0	0	1+	Termination(s) only	369	21.3
0	1+	0	Miscarriage(s) only	125	20.8
0	1+	1+	Miscarriage(s) and termination(s)	29	21.8
1+	0	0	Childbirth(s) only	447	21.5
1+	0	1+	Childbirth(s) plus other	63	21.6
1+	1+	0	Childbirth(s) plus other	82	21.8
1+	1+	1+	Childbirth(s) plus other	24	22.1

Source: Author's calculations from the Australian Longitudinal Survey of Women's Health.

7 Giving children up for adoption is rare in Australia and so a wish to conceal this is unlikely to be a major source of measurement error.

Table 2 Pregnancies, births, miscarriages and terminations by age

	Age as at end 1996							
	17-20				21-24			
	Number of completed pregnancies(a)				Number of completed pregnancies(a)			
	0	1	2+	All	0	1	2+	All
Number of births								
0	3,682	152	20	3,854	4,041	295	56	4,392
1	-	94	28	122	-	250	90	340
2+	-	-	24	24	-	-	130	130
Number of miscarriages								
0	3,682	202	39	3,923	4,041	479	159	4,679
1	-	44	19	63	-	66	92	158
2+	-	-	14	14	-	-	25	25
Number of terminations								
0	3,682	138	44	3,864	4,041	316	156	4,513
1	-	108	21	129	-	229	74	303
2+	-	-	7	7	-	-	46	46
All	3,682	246	72	4,000	4,041	545	276	4,862
All (%)	92.1	6.2	1.8	100.0	83.1	11.2	5.7	100.0

Note: (a) Prior to the questionnaire completion in mid-1996.
Source: Author's calculations from the Australian Longitudinal Survey of Women's Health. Unweighted data

Table 2 shows the numbers of pregnancy outcomes by age as at end 1996.[8] Almost all the births occurring in the sub-sample aged 17-20 can be classed as 'teenage births' since the survey was mailed out in mid-1996.

In the younger group, 6.2 per cent had one completed pregnancy prior to wave 1, and 1.8 per cent had two or more completed pregnancies. The corresponding figures for the older half of the sample are 11.2 and 5.7 per cent. In the younger group there were slightly more births than terminations, while in the older group there were substantially more births. The miscarriage rate was higher for the younger group (approx 23 per cent of all pregnancies compared to 18 per cent for the older group).

2.2 The impact of miscarriage

What is the impact of miscarriage on short and longer-term fertility outcomes? Is a miscarriage simply followed by another pregnancy, or does having a miscarriage have a significant long-term impact on fertility?

The relationship between miscarriage and subsequent births for the pregnant women sample is summarised in the logistic regression results shown in Table 3. This regression estimates the influence of miscarriage on the probability of having a birth prior to the first wave of the survey. Because 31 per cent of the pregnant women sample had more than one pregnancy prior to wave 1, miscarriage is summarised as the fraction of pregnancies that ended in miscarriage. The number of pregnancies and the woman's age (at end 1996) are included as controlling variables. Interactions between these variables (and age-squared) were tested,

8 For confidentiality reasons, the survey data file only includes year of birth, and so it is not possible to calculate age as at the survey completion date.

but were not significant. The second panel of the table repeats the regression while including a number of background characteristic variables as controls. The results are very similar to the simpler estimates in the top panel.

Overall, 54 per cent of the pregnant women sample had at least one childbirth prior to wave 1. As expected, the fraction of pregnancies ending is miscarriage is strongly associated with not having a child, with an odds-ratio of under 5 per cent (ie exp(–3.05)). By definition, if all the pregnancies of a particular woman ended in miscarriage, she would not record any childbirths (though in fact the fitting process of the logistic regression actually predicts a 5 per cent childbirth rate among those with one pregnancy and aged 21). On the other hand, among women who did not have a miscarriage (and were aged 21 with only one pregnancy), 51 per cent had a child. The remainder had terminations.

In addition however, this summary model predicts that of those 21-year-old women with two pregnancies, but only one miscarriage, some 45 per cent will have at least one childbirth prior to wave 1. The actual fraction across all ages is even higher at 72 per cent (compared to 85 per cent among those with no miscarriages). This points to the fact that, for most women, having a miscarriage implies a delay in first childbirth, and this delay might not be that long.

This is addressed more explicitly in Table 4, which shows the impact of miscarriage prior to wave 1 on age at first birth – taking account of births recorded in all three waves. Again, the population is the pregnant women sample. The first two data columns in the table show the results of an OLS regression of age at first birth as a function of age at end 1996, number of pregnancies prior to wave 1 and the fraction of those pregnancies that ended in miscarriage. Women who did not have any births recorded prior to wave 3 are not included in this estimation. These women are included in the second two columns, which show a similar estimation using the Tobit estimation procedure. This takes account of the fact that age at first childbirth is censored at the woman's age in wave 3. It does, however, impose strong distributional assumptions on age at first childbirth for those women who are censored.[9]

The impact of miscarriage on age at first childbirth is much the same under both estimation approaches. Controlling for other background characteristics also doesn't make much difference. Compared to not having any miscarriages, having all her pregnancies prior to wave 1 end in miscarriage implies that a woman's age at first childbirth will increase by around 2.8 years. The effect is very similar if the estimation is repeated for just those with only one pregnancy (not shown).

[9] The SAS QLIM procedure was used (version 9.1.2). The censoring is allowed to vary between women depending on their age in wave 3.

Table 3 The impact of miscarriage on probability of having a child prior to wave 1 (Women with at least one pregnancy prior to wave 1)

Variable	Logistic regression estimates Parameter estimate	 Standard error
Intercept	0.05	0.08
Age at end 1996 (minus 21)	0.05	0.05
Number of pregnancies prior to wave 1 interview (minus 1)	1.27	0.13
Fraction of pregnancies ending in miscarriage	-3.05	0.34
N	1139	
Mean proportion with birth	0.54	
Predicted proportion when age = 21, number of pregnancies = 1 and no pregnancies end in miscarriage	0.51	
Corresponding predicted proportion when all pregnancies end in miscarriage	0.05	
Predicted proportion when age =21, number of pregnancies = 2 and half these end in miscarriage	0.45	

Estimates controlling for smoking, ATSI status, English speaking, location and father's occupation

Intercept	-0.35	0.22
Age at end 1996 (minus 21)	0.07	0.05
Number of pregnancies prior to wave 1 interview (minus 1)	1.30	0.13
Fraction of pregnancies ending in miscarriage	-3.17	0.34
N	1139	
Mean proportion with birth	0.54	
Predicted proportion when age = 21, number of pregnancies = 1, no pregnancies end in miscarriage and with reference background characteristics	0.41	
Corresponding predicted proportion when all pregnancies end in miscarriage	0.03	
Predicted proportion when age =21, number of pregnancies = 2, half these end in miscarriage and with reference background characteristics	0.35	

Source: Author's calculations from the Australian Longitudinal Survey of Women's Health (Unweighted data).

Table 4 The impact of miscarriage on age at first childbirth (Women with at least one pregnancy prior to wave 1)

	OLS (for women with a birth prior to wave 3)			Tobit (includes those with no birth)	
Variable	Parameter estimate	Standard error		Parameter estimate	Standard error
Intercept	20.59	0.11		23.31	0.18
Age at end 1996 (minus 21)	0.73	0.07		0.65	0.11
Number of pregnancies prior to wave 1 interview (minus 1)	-1.23	0.10		-2.15	0.18
Fraction of pregnancies ending in miscarriage	2.67	0.29		2.87	0.46
N	852			1115	
R^2	0.24				
s				4.79	
Estimates controlling for smoking, ATSI status, English speaking, location and father's occupation					
Age at end 1996 (minus 21)	0.75	0.07		0.57	0.11
Number of pregnancies prior to wave 1 interview (minus 1)	-1.19	0.10		-2.01	0.17
Fraction of pregnancies ending in miscarriage	2.71	0.29		2.90	0.44
N	852			1115	
R^2	0.27				
σ				4.56	

Source: Author’s calculations from the Australian Longitudinal Survey of Women’s Health (Unweighted data).

3 Outcomes for Women with Different Pregnancy Experiences prior to Wave 1

The selection effect for young motherhood is likely to work at two levels. First, young women and their partners may or may not engage in sexual behaviour that will be likely to lead to pregnancy. Second, pregnant women make decisions about terminations.

Selection on the basis of characteristics that are associated with socio-economic outcomes is possible at both stages. That is, women expecting poor educational or labour market outcomes may be less concerned about becoming pregnant, and also less likely to have a termination if they do become pregnant. If selection effects operate at both these decision-making levels we might expect that socio-economic outcomes might be ranked in the following order (better outcomes first). This ranking is solely based upon the anticipated selection effect and takes no account of any direct effect of childbirth.

1. Women who do not become pregnant at a young age.

2. Women who become pregnant, but have a termination.

3. Women who have a miscarriage.

4. Women who have a child at a young age.

Women who have a miscarriage will come from both groups 2 and 4. That is, some might be intending to have a termination while others might be intending to continue to childbirth. Hence, they might be expected to have outcomes between these two groups.

If, on the other hand, there were no selection effect, there should be no difference between groups 1, 2 and 3. If having a child has a direct effect on outcomes then group 4 will be different from the first three groups. Comparing outcomes among these four groups is thus a useful starting point in examining whether selection or direct effects predominate.

This comparison is complicated by the fact that many women in the ALSWH survey had experienced more than one pregnancy, sometimes with different outcomes. To summarise this, the groupings shown in Table 1 are used (with the last three categories grouped together).

Table 5 shows the wave 3 demographic characteristics of each of these pregnancy experience groups. If selection effects are the main link between socio-economic characteristics and fertility experience, we would expect that socio-economic outcomes should generally become poorer as we move from left to right across the table (though the placement of the multiple outcome categories might be less clear). If the direct effect of childbirth were the only (negative) effect, we would expect to find the first four columns having similar patterns while the last two columns had poorer outcomes.

It is possible that marital status is both an outcome of childbirth and also associated with selection effects. If separation is considered as an adverse effect, then the observed pattern does seem to be consistent with a direct effect. Those who had a childbirth prior to wave 1, were much more likely to be classed as separated, divorced or widowed some seven years later. However, on the other hand, the proportion of the sample never married varies in line with a selection effect; those 'closer' to having children are less likely to be single. In this case the selection might be operating via preferences for family formation rather than via differences in labour market opportunities.

The household type patterns follow those for marital status. Interestingly, the proportion living with their parents when aged 25-30 falls steadily, in line with a selection effect. Either because of poorer labour market opportunities or preferences for independence or family formation (and hence possible childbirth after wave 1), those with pregnancies prior to wave 1 (but no child at wave 1) were less likely to be living at home than those with no early pregnancy. Note also that a large fraction of the pregnant but no child group were living with children in wave 3, with terminations or miscarriages simply delaying the first child.

Table 5 Demographic Characteristics by Pregnancy Experience

	Pregnancy Experience in Wave 1 (when aged 18-23)					
Characteristics in wave 3 (when aged 25-30)	No previous pregnancy	Termination only	Termination and miscarriage	Miscarriage only	Birth only	Birth plus other
Marital Status (%)						
Never married	38.8	30.0	31.6	21.6	17.2	24.2
Married	39.3	36.4	34.2	52.8	50.5	37.3
De facto	19.2	28.9	24.1	23.3	20.6	24.8
Married plus de facto	*58.5*	*65.3*	*58.3*	*76.1*	*71.2*	*62.1*
Separated/Divorced/Widowed	2.7	4.7	10.1	2.3	11.7	13.7
Total	100.0	100.0	100.0	100.0	100.0	100.0
Household Type						
Respondent only	8.0	7.9	8.9	4.5	0.0	1.4
Respondent plus spouse only	38.8	29.4	13.9	22.7	3.8	0.5
With spouse and children	20.3	33.2	46.3	47.9	67.3	63.5
With children only	1.7	5.3	10.1	6.4	21.8	26.0
as % of all with children	*7.8*	*13.9*	*18.0*	*11.7*	*24.5*	*29.1*
With respondent's parent(s)	12.4	10.0	7.0	2.5	1.2	1.4
With other adults (no children)	14.8	8.6	13.9	11.7	0.9	0.0
Other	4.0	5.7	0.0	4.4	5.1	7.2
Total	100.0	100.0	100.0	100.0	100.0	100.0

Source: Author's calculations from the Australian Longitudinal Survey of Women's Health. Weighted data.

Finally, the 'lone parenthood rate' (those living with children only as a percentage of all those with children in wave 3), generally increases across the table. With the exception of the miscarriage only group, this is consistent with a selection effect whereby women who are more likely to have a young childbirth are less likely to have a partner in wave 3. This could be due to poorer partnering opportunities in the regions in which these women live. On the other hand, the much higher rate of partnering (among those with children) of those with miscarriages compared to those with births is suggestive of a direct impact of having a young birth. This is probably associated with the higher rate of separation among those with births.

Table 6 summarises the education and labour market characteristics of the different fertility groups. For the very lowest education level, there is some evidence of a direct effect, as there are markedly larger fractions in the two birth groups. Otherwise, the large differences in educational levels generally follow the pattern that might be expected on the basis of a selection effect. Of those with no pregnancy prior to wave 1, only 6 per cent had not progressed beyond a year 10 qualification. Among those with a childbirth, 29 per cent fell into this category (45 per cent of the birth plus other group). The termination and miscarriage groups had intermediate rates of low qualifications in line with the expected selection effect pattern.

Table 6 Labour Market Characteristics by Pregnancy Experience

Characteristics in wave 3 (when aged 25-30)	Pregnancy Experience in Wave 1 (when aged 18-23) No previous pregnancy	Termination only	Termination and miscarriage	Miscarriage only	Birth only	Birth plus other
Highest Educational Qualification (%)						
No formal qualifications	0.6	0.9	0.0	1.0	5.2	10.5
Year 10 or equivalent	5.4	12.4	17.2	26.4	23.7	34.0
Year 12 or equivalent	16.2	22.7	25.3	25.9	28.0	27.9
Trade/apprenticeship	2.7	3.4	0.0	2.7	1.9	2.2
Certificate/diploma	21.5	26.7	18.3	19.4	25.8	22.4
University degree	40.7	23.3	27.2	19.4	12.4	3.1
Higher university degree	12.9	10.5	12.0	5.2	3.0	0.0
Total	100.0	100.0	100.0	100.0	100.0	100.0
Employment Status (as % of population)						
Employed	84.9	79.6	76.6	70.7	55.5	55.8
Unemployed	6.2	7.6	11.0	10.1	15.7	17.5
% > 6 months	*48.5*	*66.9*	*81.4*	*77.4*	*78.5*	*86.2*
Not in Labour Force	9.0	12.9	12.4	19.3	28.8	26.7
Total	100.0	100.0	100.0	100.0	100.0	101.0
Occupation						
Manager or administrator	7.9	11.0	5.2	13.1	8.8	13.2
Professional	43.4	24.7	26.1	21.4	20.1	11.2
Associate professional	9.1	10.5	7.0	6.3	7.4	2.9
Tradesperson or related worker	3.1	5.9	2.6	7.4	6.2	3.8
Advanced clerical or service worker	14.6	20.4	27.8	19.2	11.3	11.5
Intermediate clerical, sales/service worker	14.5	16.4	19.1	13.8	24.0	27.7
Intermediate production or transport worker	0.5	1.1	0.0	1.8	0.3	4.8
Elementary clerical, sales or service worker	4.4	5.2	9.6	6.8	9.0	16.2
Labourer or related worker	2.5	4.8	2.6	10.2	12.9	8.6
Total with jobs	100.0	100.0	100.0	100.0	100.0	100.0

Source: Author's calculations from the Australian Longitudinal Survey of Women's Health. Weighted data.

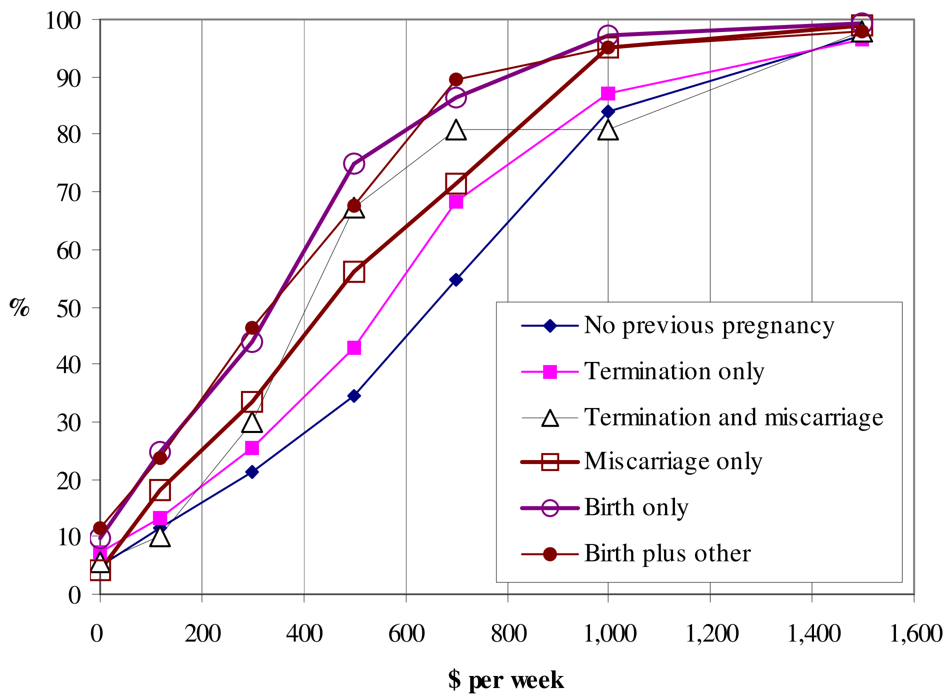
It is not surprising that education patterns at the lower end should follow the patterns of selection rather than direct effects. Very few of the pre-wave 1 childbirths would have occurred prior to the end of year 10. It is more likely that there may be direct impact of childbirth on the achievement of higher levels of education. However, patterns of university degree attainment also follow a selection pattern. Of those who had never been pregnant prior to wave 1, over half had university degrees by wave 3. Among those who had an early childbirth (only), about 15 per cent had a university degree. The termination and miscarriage groups have intermediate rates of higher education as expected.

The story with respect to employment status is more complicated because the presence of (particularly young) children in the household also has an impact on employment status. The two groups of women who had early childbirths have the lowest probability of employment (around 56 per cent), compared to over 70 per cent for the other groups. However, this may well reflect their higher probability of having children in their household at wave 3, and is not necessarily indicative of poorer employment prospects in the longer term. Unemployment is a better indicator of women not being able to achieve their desired employment status. The pattern here appears to be generally that expected via a selection effect, though there is a higher proportion of the population unemployed among those with a childbirth.

Of those employed, professional employment tends to follow the patterns for higher education, though there is no such pattern for managerial/administrative employment.

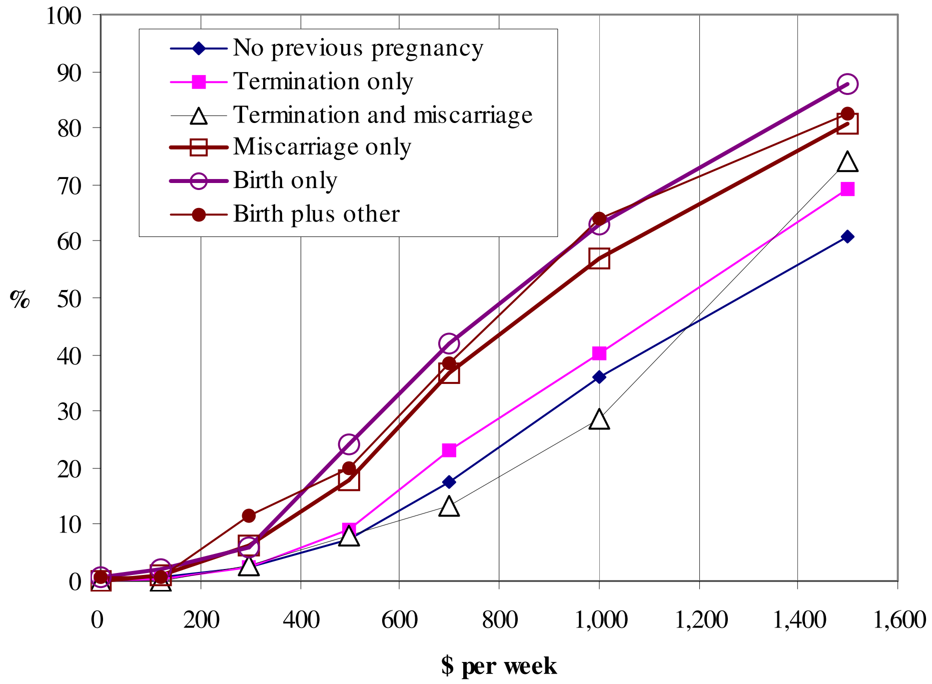
The incomes of these women are shown in Figure 1. In this figure lines lying to further to the right represent higher levels of income. The incomes of those women with no pregnancy in wave 1 are clearly higher, with fewer having incomes below $700 per week. Those women with a miscarriage generally have incomes between the termination and birth groups.

Figure 1 Cumulative Personal Income Distribution by Pregnancy Experience



For household income (Figure 2), there are two clear groupings. Women with no pregnancy or a termination have a high income, whereas women with a childbirth or a miscarriage have a lower income level (higher lines). The low income for women with miscarriage suggests a selection effect, as these women did not have a child (prior to wave 1).

Figure 2 Cumulative Household Income Distribution by Pregnancy Experience



4 Identifying the Impact of a Young Childbirth

The results presented above suggest that selection effects are the main reason for the association between young childbirth and later socio-economic outcomes. The remainder of the paper seeks to more directly identify whether there is any direct effect on outcomes.

The classic experimental approach to causal inference would identify this relationship by randomly assigning women to have their first child at different ages (eg a 'young' and an 'old' age). The impact of young motherhood is then calculated as the difference in average outcomes between the young and old childbirth groups. Such a random assignment would be applied to a representative sample of women and thus be an estimate of the average 'treatment' effect across all women. Of course, such a random assignment is neither ethically or biologically feasible. What can we estimate using observational data?

4.1 The bias in simple difference estimators

One approach is to simply compare the outcomes of women who have children when young, with women who do not. As noted above, this difference in outcomes is likely to be a biased estimate of the impact of young childbirth. This is because there are likely to be unobserved 'third variables' (eg labour market skills) which influence both outcomes and fertility-related behaviour.

In fact, there are two possible sources of selection bias. The first arises because, even if they did not have a child, the women who have a young childbirth may have differed in their average outcomes. For example, women whose labour market prospects are poorer may consequently decide to have a child earlier than women who are likely to have higher earnings. The difference between the average later life labour market outcomes for women with and without young births will thus be a biased estimator of the causal impact of a young birth. It will suggest that a young birth decreases expected outcomes more than is actually the case.

A second bias operates in the opposite direction. It is possible that those women who think that there will only be a small *change* in their outcomes because of a young childbirth will be more likely to have a birth. If this is the case, then this will reduce the observed difference between the two groups.

This second bias arises from the fact that the impact of a young childbirth might differ between those who do and who don't have a child (Ermisch and Pevalin, 2003). Estimation becomes a little simpler if we restrict attention to just one of these groups. In this case, we can examine the effect of a young childbirth on the population of women who do actually have a young childbirth.

In the evaluation literature, this is known as the effect of *treatment on the treated*. In the present example this is a policy-relevant focus. The most likely policy intervention in this area is to intervene to reduce childbirths among young women. Such a policy would thus impact upon the population of women who would have had a young childbirth, and so this is an appropriate group on which to focus attention.

For this group, the above discussion suggests that we would expect the simple difference between women with and without young childbirths to be an overestimate of the negative education and labour market effects of young childbirth.

4.2 Miscarriage as a randomisation device

One method for addressing these identification biases is to utilise information on miscarriages. To use this information, we restrict our attention to women who become pregnant within the age window of interest (eg teenage years). Women whose pregnancies miscarry can then be compared with women who do not have a miscarriage. The miscarriage thus serves as a 'random assignment' with one group of women having childbirths and another not.

However, there are two reasons why this is not the same as a randomised experiment. First, miscarriages are not entirely random. Some women are more likely than others to miscarry and the characteristics that influence miscarriage might be associated with later life outcomes. Some of these characteristics are known. For example, women who smoke are more likely to miscarry. If these characteristics are observed then we can control for them and estimate the impact of miscarriage within each group defined by these characteristics. In the ALSWH survey, extensive information on smoking habits is collected in the first wave, and so this and other characteristics that might be related to miscarriage propensities can be included as controlling variables. This is the approach used by Hotz, McElroy and Sanders (2005).

However, it is possible that there might be some other, unobserved, aspects of health status that both influence miscarriage propensities and later life outcomes. For example, women who are generally more healthy may be less likely to miscarry, and more likely to have favourable labour market outcomes. If this is the case, then the use of miscarriage as a randomisation device will lead to biased estimates. Hotz, Mullin and Sanders (1997) develop a method for placing bounds on the likely impact of this selection, based on assumptions about the proportions of miscarriages that are truly random. This method is also employed by Ermisch and Pevalin (2003). Unfortunately this approach leads to quite wide bounds for the likely effect and so is not employed here. Hotz, McElroy and Sanders (2005) however, argue that the assumption of randomness of miscarriages (controlling for characteristics such as smoking) is not rejected by these more comprehensive tests, and so estimates using this assumption are still valid.

A second, complication is that many pregnancies end in abortion rather than miscarriages or live births. While miscarriage might be considered close to random when controlling for suitable health status variables, this is not the case for abortion, which requires a deliberate decision. It is likely that some of the factors associated with the choice to have an abortion may be directly related to later life outcomes – such as the likely labour market implications of having a child.

With respect to a given pregnancy, women who have a miscarriage do not have an associated childbirth. Women who do not have a miscarriage may have a birth or not, depending upon whether they have an abortion. If we think of the 'treatment' as being the absence of a miscarriage, then this situation is the same as the situation in many randomised experiments, where some subjects who are selected to receive the treatment do not in fact receive it. In this case, it is common to compare groups based upon the 'intention to treat' rather than the actual treatment.

In the present example, the analogous 'intention to treat' comparison is to select women who have a pregnancy within the relevant age window and then compare the later life outcomes of those for whom the pregnancy ends in a miscarriage with those where this is not the case. Note that the latter group includes those who have an abortion. This is described here as the 'miscarriage effect'.

Using similar notation to Ermisch and Pevalin (2003), let y indicate an outcome variable, $p = 1$ if the woman became pregnant at a young age (0 otherwise) and $M = 1$ if she has a miscarriage (0 otherwise). This miscarriage effect (or 'intention to treat') estimator is then defined as

$$I = \mathrm{E}(y|\, M = 0, p = 1) - \mathrm{E}(y|\, M = 1, p = 1) \qquad (1)$$

The outcome variable can be considered as also being conditional on some set of exogenous explanatory variables, x, but for clarity this is not included here. This is estimated as the average difference in outcomes between women who become pregnant and do not have a miscarriage and the outcomes for women who do have a miscarriage (while possibly controlling for x). The first group includes both women who go on to childbirth as well as women who have an abortion.

This comparison shows the impact of having a miscarriage on later life outcomes. In itself, this result is not of much direct interest. Any policy intervention to reduce young childbirths is not likely to be implemented via an increase in miscarriages. Instead, policy interventions will focus on reducing the number of pregnancies or possibly increasing the number of abortions (eg in places where abortion is currently illegal or unavailable). Nonetheless, if the miscarriage effect is zero the young birth estimator described below will also be zero. Since the statistical properties of the miscarriage estimator are more robust, it is useful to estimate this so as to reach a qualitative conclusion about whether young birth has an identifiable impact.

Ideally, however, what is needed is an estimate of the impact of young childbirth per se, rather than just the impact of not having a miscarriage. To use the randomisation arising from miscarriage to estimate this more policy-relevant question additional, arguably quite reasonable, assumptions are needed.

- Miscarriage is assumed to be random (conditional on observed characteristics such as smoking habits).
- The only impact of miscarriage on later life outcomes acts via the fact that it precludes a live birth in this pregnancy. For example, if miscarriage played a role in triggering depression then this assumption would be invalid, as the depression might in turn impact on later life outcomes.[10]
- Having an abortion has no direct later life impact other than via the prevention of a having a childbirth. Again, if physical or emotional trauma resulting from the abortion influenced later life outcomes then this assumption would be invalid.

If these assumptions hold, then we can form a consistent instrumental variable estimator (or Wald estimator) of the impact of having a young childbirth as

$$\alpha = \frac{\begin{pmatrix} \text{Average outcome of women with pregnancies but no miscarriage} - \\ \text{Average outcome of women with a miscarriage} \end{pmatrix}}{\text{Probability that the intended pregnancy outcome was a birth rather than an abortion}}$$

or, more concisely,

$$\alpha = \frac{\text{Miscarriage effect}}{\text{1 - Abortion probability}} \qquad (2)$$

The intuition behind this estimator is as follows.

The assumptions above imply that the only impact of miscarriage is on the likelihood that the woman will have a childbirth. If there were no abortions (the denominator equal to one), then absence of a miscarriage implies a childbirth (and vice versa) and hence the miscarriage effect is the same as the impact of a childbirth. However, lack of miscarriage does not always imply a childbirth, and so the miscarriage effect represents an attenuation of the impact of childbirth on later life outcomes.

The miscarriage effect is comprised of comparisons within two population sub-groups: women who intended to have an abortion, and women who did not. The miscarriage effect is the weighted average of the miscarriage effect in these two sub-populations. For the sub-population of women who intend to have an abortion, the above assumptions imply that a miscarriage will have no impact on later life outcomes since they would have no childbirth whether or not the miscarriage occurred. That is, for this sub-group the miscarriage effect is zero.

---

[10] This also precludes general equilibrium influences. If, for example, a miscarriage for woman A, increased the availability of childcare for woman B (who did not have a miscarriage), and hence influenced the labour market participation of woman B then the estimates shown here would not be accurate estimates of the impact of miscarriages. Given that we are considering only a small sub-set of families, such an effect is likely to be negligible.

For sub-population of women who will not have an abortion, having a miscarriage means not having a child and not having a miscarriage implies a childbirth. So, for this sub-group the miscarriage effect is identical to the impact of having a child. The observed miscarriage effect is thus the effect for this group, weighted by the probability of being in this group (ie the probability of not having an abortion). Equation (2) inverts this relationship to show the impact of having a child as a function of the observed (overall) miscarriage effect.

More formally, let l = B or A indicate the (latent) pregnancy outcome that would have occurred if there had been no miscarriage, with B representing birth and A abortion. For a given pregnancy, the probability that $l = B$ is represented by P_B (with $P_A = 1 - P_B$). Among the population of women with pregnancies, the estimator above can thus be written as

$$\alpha = \{E(y \mid M = 0) - E(y \mid M = 1)\}/P_B \qquad (3)$$

This is estimated as the difference between the mean outcomes of pregnant women with no miscarriages and the outcomes for women with miscarriages, all divided by the latent birth probability.

To see how this scaling arises, we decompose the two expectations in this expression into the expected outcome among women who had (or would have had) an abortion and those who did not. These outcomes are weighted by the probabilities of abortion or birth. Because miscarriages are assumed to be random, these probabilities are the same in the miscarriage and non-miscarriage group.

$$E(y \mid M = 0) = P_B E(y \mid M = 0, l = B) + P_A E(y \mid M = 0, l = A) \qquad (4)$$

$$E(y \mid M = 1) = P_B E(y \mid M = 1, l = B) + P_A E(y \mid M = 1, l = A) \qquad (5)$$

The term $E(y \mid M = 0, l = A)$ in the first of these equations is the expected outcome among women who did not have a miscarriage, but did have an abortion. The term $E(y \mid M = 1, l = A)$ in the second equation is the expected outcome of women who would have had an abortion but in fact had a miscarriage. Again, if we assume that miscarriages are random then these two expectations are equal. Hence

$$E(y \mid M = 0) - E(y \mid M = 1) \quad = \quad P_B\{E(y \mid M = 0, l = B) - E(y \mid M = 1, l = B)\} \qquad (6)$$

The term in braces on the RHS is thus equal to α in equation (3). It represents the expected outcome difference between those women without a miscarriage who have a birth and those women who had a miscarriage but would have chosen a birth. It is thus an *estimate of the impact of having a young childbirth among the sub-population of women who were pregnant in the age window and would not have chosen an abortion*. As noted above, this is a population of policy relevance – these are the women most likely to be targeted by any policy designed to reduce fertility at young ages.

However it should be remembered that the estimate provides no direct information on the outcomes for the remaining population. For example, no direct information is provided on the likely outcomes of a reduction in availability of abortions, because this would apply to those women who presently would have abortions. Any inference for this group would need additional assumptions (such as an assumption of constant impacts across the population). We return to this issue in the conclusion.

Finally, we can show this is equivalent to conventional instrumental variable (IV) estimators. If z = 1 when a woman has a young child birth (0 otherwise) the Wald IV estimate of the impact of z on y is given by

$$\alpha = \frac{\text{Impact of } M \text{ on } y}{\text{Impact of } M \text{ on } z} = \frac{Cov(M, y)}{Cov(M, z)} = \frac{E(y \mid M = 0) - E(y \mid M = 1)}{E(z \mid M = 0) - E(z \mid M = 1)} \qquad (7)$$

But $E(z \mid M = 0) = P_B$ and $E(z \mid M = 1) = 0$ because a miscarriage totally precludes a birth from this pregnancy. The Wald estimate given in this equation is also identical to that obtained via the 2SLS estimate (Angrist, Imbens and Rubin, 1996).

All the quantities in expression (3) can be estimated from information on a random sample of pregnancies. All that is needed is the average outcome for women with a pregnancy that does not end in miscarriage, the average outcome for women with a miscarriage and the probability of choosing abortion rather than childbirth. The latter can be estimated as the proportion of those women who do not have miscarriages who have abortions. All of these terms can be estimated conditionally on observed characteristics (including those characteristics that might influence the chance of a miscarriage). The next section describes how the data available in the ALSWH survey is adapted to undertake this estimation.

5 Estimation Method

The ALSWH data does not contain detailed information on each pregnancy and its outcome. The survey includes information on the number of pregnancies and pregnancy outcomes, but does not contain information on the date of each pregnancy. We could directly apply the above approach to this data by restricting to women with only one pregnancy. However, this would both discard much relevant data and also (potentially) introduce selection biases. Consequently, the empirical analysis employed here modifies this simple model by using the fraction of pregnancies that resulted in miscarriage as the identifying instrument.

The analysis population is the pregnant women sample, women who had at least one (completed) pregnancy prior to wave 1. Table 7 shows the births and miscarriages of this sample (disaggregated by age). The birth and miscarriage variables are defined as follows.

The variable flagging births, z_i, is set equal to one if the woman had one or more births prior to wave 1.

The variable, M_i is defined as the proportion of pregnancies that end in miscarriage (ie as in Section 2.2). For 77 per cent of the pregnant women sample, no pregnancies ended in miscarriage, in 11 per cent of cases all pregnancies ended in miscarriage, 6 per cent half the pregnancies, 3 per cent one-third of pregnancies and in 1 per cent of cases, one-quarter of their pregnancies ended in miscarriage. All the estimations also control for the total number of pregnancies experienced prior to wave 1.

A key feature of the ALSWH data is that the age ranges for which we have information on the numbers of different pregnancy outcomes varies across the sample – depending upon the woman’s age at wave 1. To take account of this, the relationships in the model are introduced in interaction with age.

For each outcome variable y, several relationships are estimated. First, the relationship between the outcome and having a birth prior to wave 1.

$$y_i = \alpha z_i + \beta a_i z_i + x_i \gamma + e_i \tag{8}$$

where a_i is age (at end 1996 minus 20),[11] x_i is a vector of control variables (including the number of pregnancies prior to wave 1), and e_i an error term. This is estimated both for the full population of women and for the pregnant women sub-sample.

Then estimates for the 'reduced form' or miscarriage effect equation are presented (both linear probability and logistic estimates are presented for binary dependent variables).

$$y_i = \alpha' M_i + \beta' a_i M_i + x_i \gamma' + e'_i \tag{9}$$

This shows the direct impact of miscarriages on outcomes. Estimates of this equation can be used to identify the direction of the childbirth impact. Finally equation (8) is estimated by 2SLS using instruments M_i, $a_i M_i$ and x for the endogenous variables z_i and $a_i z_i$. Though most of the outcome variables, y, and the endogenous variable, z, are binary, 2SLS is used here following the recommendations of Angrist and Krueger (2001).[12]

Because of age definition, the parameter α will be an estimate of the impact of delaying the first childbirth past age 20 (and $\alpha + \beta$ = the effect of delaying past 21 etc). This estimate of α obtained in the 2SLS estimation thus represents *the impact on y of having one or more births prior to wave 1 compared to the alternative of having at least one pregnancy prior to wave 1 but not having any births because of miscarriage*.

[11] The survey records age as at end 1996 rather than at date of survey (mid 1996). Women with $a_i = 0$ therefore turned 20 at some time during calendar year 1996, and so on average were turning 20 at about the time of the survey.

[12] Angrist and Krueger (2001:80) argue that the use of non-linear specifications (eg logit) in the first stage of the instrumental variable estimation procedure carries a high risk of specification error, and that even if the second stage dependent variable is binary the least squares estimates still capture a relevant average treatment effect. They conclude, 'two-stage least squares is a robust estimation method that provides a natural starting point for instrumental variables applications'.

Table 7 Women with at least one completed pregnancy at wave 1: Number of births by number of miscarriages

	Number of births				
Number of miscarriages	0	1	2	3+	All
Age 18-20 at end 1996					
0	113	107	16	5	241
1	50	*11*	*1*	*1*	63
2	8	*3*	*1*	-	12
3+	1	*1*	-	-	2
Age 21-24 at end 1996					
0	256	289	75	18	638
1	89	*45*	*17*	*7*	158
2	5	*5*	*7*	*3*	20
3+	1	*1*	*2*	*1*	5
All ages					
0	369	396	91	23	879
1	139	*56*	*18*	*8*	221
2	13	*8*	*8*	*3*	32
3+	2	*2*	*2*	*1*	7

Source: Author's calculations from the Australian Longitudinal Survey of Women's Health. Unweighted data

Though this estimation procedure does not explicitly use information on abortions, this information enters indirectly via the first stage estimates of the relationship between miscarriages and births. That is, the existence of abortions means that the lack of a miscarriage does not necessarily imply a birth.

5.1 Socio-Economic outcomes

Though socio-economic outcomes are not a primary focus of the Australian Longitudinal Survey of Women's Health, there are a number of socio-economic characteristics collected. These include information on marital status and household composition, education level, labour market status, income category, feelings of stress in different life areas, job security and residential postcode. The following outcome variables are considered here. In some cases, the analysis is only undertaken for a restricted population (for non-missing data and population selections as indicated). In some cases it is appropriate to incorporate additional controlling variables for specific outcome variables. These are also indicated below.

Demographic indicators:

Partnered = 1 if married or in a defacto relationship, 0 otherwise.

Married = 1 if legally married (0 if in a defacto relationship or not partnered).

Lone parent = 1 if living with children only (Population: women living with children only or with spouse and children).

Education indicators:

Year 12 or more = 1 if highest education level is Year 12 or above.

Certificate or more = 1 if highest education level is Certificate, diploma or trade qualification or above.

Degree = 1 if highest education level is university degree.

Employment indicators:

Employed = 1 if employed.

Unemployed = 1 if unemployed ('unemployed and actively seeking work').

Occupation indicators (population: those with a paid job):

Manager = 1 if current occupation is 'Manager or administrator (eg. magistrate, farm manager, general manager, director of nursing, school principal)'.

Professional = 1 if current occupation is 'Professional (eg. scientist, doctor, registered nurse, allied health professional, teacher, artist)'.

Intermediate = 1 if current occupation is 'Advanced clerical or service worker (eg secretary, personal assistant, flight attendant, law clerk)', 'Associate professional (eg technician, manager, youth worker, police officer)' or 'Tradesperson or related worker (eg hairdresser, gardener, florist)'.

Other_Occupation = 1 if has paid job not included above.

Income measures:

Own income: Gross income recoded to the midpoints of the income ranges ('What is the average gross (before tax) income that you receive each week, including pensions, allowances and financial support from parents?' – 8 categories).

Low personal income: Gross income below $300 per week.

Household income: Gross household income recoded to midpoints ('What is the average gross (before tax) income of your household (eg you and your partner, or you and your parents sharing a house?' – 8 categories) (Population: partnered women).

Low household income: Gross household income below $1,000 per week. (Population: partnered women).

Other income: Household income – Own income (Population: partnered women).

Low other income: If the mid point of the household income category minus the midpoint of the personal income category is less than $700 per week. (Population: partnered women).

SEIFA_Change: Change in the ABS SEIFA index of disadvantage for their postcode of residence between waves 1 and 3.[13] High scores indicate a move to a more advantaged

[13] This is calculated using postcode as reported in each wave together with ABS data on mean SEIFA scores in each SLA in 2001 and ABS data describing the proportion of people in each postcode living in each SLA. Note that the SEIFA scores are all defined as at the 2001 Census. They thus do not take account of any changes in the characteristics of regions over time.

region. For this variable, additional controlling variables flagging whether the woman was living with her parents in wave 1 and in wave 3 are included.

5.2 Control Variables: Determinants of Miscarriage

For the reasons mentioned in Section 4, it is important to control for the characteristics potentially associated with miscarriage. The key variables are smoking and drinking behaviour. Ideally, these should be measured prior to the commencement of pregnancy. This information is not available, however it is possible to use a set of measures that a closely correlated with this. The following variables are included as controls:

Age: Age in whole years as at end 1996.

Years_Smoking: Years since first started smoking (0 for never-smokers and missing cases). Defined as age as at end 1996 minus age started smoking (ages younger than 13 coded to 13). This ignores the fact that some people may have stopped smoking.

Smoked_Daily =1 if ever smoked daily for 6 months or more.

Cigarettes Number of cigarettes usually smoked per day, now or in the past. (Missing set to zero).

Drinker = 1 if the woman falls into the categories of low, intermediate or high risk drinker (as opposed to non-drinker, drinks rarely, or no-risk drinker).[14] This variable was tested in early specifications, but is not included in the models here. Exploratory analysis suggested that drinking behaviour was probably directly influenced by childbirth (with mothers observed to be drinking substantially less than non-mothers). As such, this was not suitable for use as a controlling variable.

5.3 Other Control Variables

The ALSWH survey over-sampled women in rural and remote areas. If the relationship between childbirth and outcomes is the same in both urban and rural/remote areas, then this sampling scheme will not influence the results.

The variable *Remote* is included in all estimations, equal to one for women who were in rural or remote regions at the time of sample selection. Supplementary estimates using weights inversely proportional to the probability of selection were also calculated. They are generally very similar (the one possible exception is discussed below).

Other control variables that are used to control for any differences in the social origins of the different groups are

ATSI = 1 if of Aboriginal or Torres Strait Islander origin (otherwise and non-response coded to 0).

Poor_English = 1 if not self-described as being either an English speaker or speaking English very well.

---

[14] Defined as for the variable *AlcSt* in Section 2.3.1 in *Data Dictionary Supplement of the Australian Longitudinal Study on Women's Health* September 2004.

Father's occupation when growing up. A set of 8 binary variables are used to describe the main occupational groups.

In addition, estimates are also presented for some of the labour market outcome variables with additional control variables flagging the presence of children in the household in wave 3 (*W3ChildU4* and *W3Child4_12*). This controls for the fact that labour market outcomes might be directly influenced by the children currently in the household.

6 Results

Table 8 and Table 10 present estimates of the impact of having a child prior to age 20 (ie the values of α). Table 9 and Table 11 present corresponding estimates of the impact of delaying the first childbirth past age 22 ($\alpha + 2\beta$). The mean age (at end 1996) of the all women sample is 20.7 and for the pregnant women sample 21.4.

The first data column of these tables shows the apparent impact of a young childbirth – estimated across the all women sample. It mirrors the comparisons shown in Table 5 and Table 6 between women with births and all other women (except that here we are holding the control variables constant). The estimate for the dependent variable Partnered is 0.09, indicating that women who had a childbirth prior to age 20 were 9 percentage points more likely to be partnered in wave 3 than other women. Estimates, such as this one, which are significantly different from zero at the 5 per cent level, are shown in bold.[15]

The second column shows similar estimates, but here calculated across the Pregnant Women sample. This controls for selection effects that might lead women with different characteristics to become pregnant. In this population, women with births were only slightly more likely to be partnered, and this is not significantly different from zero.

The third column shows the impact of having a miscarriage on the outcome variable, with the first line of the table showing the impact of miscarriage on the probability of having a birth (strongly negative, see also Section 2.2). Miscarriage is negatively associated with birth, and so miscarriage is a negative indicator of the causal relationship between birth and outcomes. That is, if we think that having a birth causes a woman to be more likely to be partnered in wave 3, then the miscarriage effect should be negative. In this case it is positive (ie in the opposite direction to the relationship suggested by the first column) but not significantly different from zero.

Estimates in the first three columns are calculated as a linear probability models using OLS. Though this is not the preferred estimation method when dependent variables are binary this is the most comparable estimate to the 2SLS estimates in the final column. For the miscarriage effect, the tables also show corresponding estimates from a logistic regression (for variables with binary outcomes). In this case the parameter estimate shows the impact of miscarriage on the log odds-ratio of the outcome event occurring. More useful for our

15 The results shown here are unweighted estimates. Weighted OLS and Logistic regression estimates which take account of the different sampling fractions in the urban and rural areas have also been calculated (using SAS 9.1.2, proc surveyreg and surveylogistic). The results (available from the author) are very similar. For example, the logistic estimate for married is 0.53 with a 't' statistics of 2.0 rather than the 0.51 and 2.2 shown in Table 8. The one exception (in Table 10) is described in the text.

purposes here is the 't' estimate for the statistic – a measure of the statistical significance of the relationship. In most cases this is very similar to that obtained from the OLS estimates.

Table 8 The Impact of Having a Child Prior to Age 20: Demographic, Education and Labour Market Outcomes

Dependent variable (wave 3, except for birth)	**All women**	**Women with at least one pregnancy prior to wave 1**			
	Birth effect	Birth effect	Miscarriage effect		Birth effect
	OLS	OLS	OLS	Logistic	2SLS
	('t' statistics in parentheses)				
	1	2	3	4	5
Birth (1st stage regression)			**-0.49**		
			(-9.8)		
Partnered	**0.09**	0.02	0.02	0.14	-0.05
	(3.0)	(0.4)	(0.5)	(0.6)	(-0.4)
Married	**0.07**	0.07	**0.12**	**0.51**	**-0.24**
	(2.3)	(1.8)	(2.2)	(2.2)	(-2.0)
Lone parent	**0.11**	0.05	0.03	0.06	-0.06
	(5.0)	(1.1)	(0.6)	(0.2)	(-0.6)
Year 10 or more	**-0.07**	**-0.07**	0.03	0.84	-0.06
	(-10.1)	(-4.3)	(1.3)	(1.2)	(-1.3)
Year12 or more	**-0.28**	**-0.19**	0.03	0.15	-0.06
	(-15.2)	(-5.0)	(0.5)	(0.6)	(-0.6)
Certificate or more	**-0.29**	**-0.16**	0.02	0.07	-0.04
	(-10.2)	(-3.8)	(0.3)	(0.3)	(-0.3)
Degree	**-0.24**	**-0.12**	0.04	0.29	-0.07
	(-8.1)	(-3.8)	(0.9)	(0.9)	(-0.9)
Employed	**-0.25**	**-0.17**	0.01	0.06	-0.03
	(-10.2)	(-4.1)	(0.3)	(0.2)	(-0.2)
Employed (cont. for children)	**-0.08**	**-0.10**	0.02	0.13	-0.06
	(-3.5)	(-2.5)	(0.5)	(0.5)	(-0.5)
Unemployed	**0.11**	0.06	-0.07	-0.61	0.14
	(6.3)	(1.9)	(-1.7)	(1.6)	(1.7)
Manager	0.02	0.00	0.08	0.75	-0.19
	(1.0)	(0.1)	(1.9)	(1.8)	(-1.8)
Professional	**-0.15**	-0.03	0.01	0.00	-0.02
	(-3.6)	(-0.8)	(0.1)	(0.0)	(-0.1)
Intermediate	**-0.11**	**-0.13**	0.03	0.12	-0.07
	(-2.7)	(-2.6)	(0.4)	(0.4)	(-0.4)
Other occupation	**0.23**	**0.17**	-0.12	-0.55	0.31
	(6.2)	(3.3)	(-1.8)	(1.8)	(1.7)

Source: Author's calculations from the Australian Longitudinal Survey of Women's Health. Unweighted data

Notes: Columns 1, 2 and 5 show the impact on the dependent variable of having a birth prior to wave 1 (normalised to age 20), controlling for the other variables as specified in the text. Columns 3 and 4 show the impact of having a miscarriage (column 3 OLS and column 4, logistic regression). The estimate divided by its standard error is shown in parentheses underneath each estimate. Where the absolute value of this is greater than 1.96, the estimate is shown in bold.

Finally, the last column shows the 2SLS estimate of the birth effect. This is approximately equal to the OLS miscarriage effect divided by the impact of miscarriage on birth (the first line).[16] For the partnered outcome, the birth effect suggests that those with a birth prior to

[16] The impact of miscarriage on birth shown is for the population for the Partnered outcome variable. For some outcome variables the population is different and so the impact is different to that shown.

age 20 are 5 percentage points less likely to be partnered. This result (in the opposite direction to the all women estimate in the first column) flows directly from the fact that the miscarriage effect is positive. However, as for the miscarriage effect itself, this result is not significantly different from zero. These patterns reflect the results shown in Table 8, where 76 per cent of the miscarriage only group are partnered, compared to 71 per cent of birth only group and even lower percentages for the termination groups.

Table 9 The Impact of Having a Child Prior to Age 22: Demographic, Education and Labour Market Outcomes

Dependent variable (wave 3, except for birth)	**All women**	**Women with at least one pregnancy prior to wave 1**			
	Birth effect	Birth effect	Miscarriage effect		Birth effect
	OLS	OLS	OLS	Logistic	2SLS
	('t' statistics in parentheses)				
	1	2	3	4	5
Birth (1st stage regression)			**-0.48**		
			(-9.8)		
Partnered	**0.05**	0.03	0.07	0.37	-0.15
	(2.4)	(1.0)	(1.3)	(1.4)	(-1.3)
Married	0.04	**0.08**	**0.13**	**0.54**	**-0.27**
	(1.9)	(2.4)	(2.4)	(2.3)	(-2.2)
Lone parent	**0.14**	**0.11**	-0.11	**-0.93**	**0.19**
	(8.1)	(3.2)	(-1.9)	(2.1)	(2.0)
Year 10 or more	**-0.04**	**-0.03**	0.01	0.20	-0.01
	(-7.7)	(-2.5)	(0.3)	(0.3)	(-0.3)
Year12 or more	**-0.23**	**-0.14**	-0.03	-0.17	0.07
	(-16.5)	(-4.5)	(-0.6)	(0.7)	(0.6)
Certificate or more	**-0.26**	**-0.20**	-0.02	-0.07	0.04
	(-12.5)	(-5.9)	(-0.3)	(0.3)	(0.3)
Degree	**-0.24**	**-0.17**	0.03	0.27	-0.07
	(-11.1)	(-6.9)	(0.8)	(0.9)	(-0.8)
Employed	**-0.21**	**-0.16**	0.05	0.21	-0.10
	(-11.9)	(-4.7)	(0.9)	(0.9)	(-0.9)
Employed (cont. for children)	**-0.09**	**-0.11**	0.04	0.24	-0.11
	(-4.7)	(-3.0)	(0.9)	(0.9)	(-0.9)
Unemployed	**0.09**	0.05	-0.01	-0.11	0.02
	(7.6)	(1.9)	(-0.3)	(0.3)	(0.3)
Manager	-0.02	-0.03	0.04	0.42	-0.09
	(-1.3)	(-1.3)	(1.0)	(1.0)	(-0.9)
Professional	**-0.14**	**-0.07**	0.00	0.00	0.00
	(-5.0)	(-2.1)	(-0.0)	(0.0)	(0.0)
Intermediate	**-0.06**	**-0.10**	0.00	-0.01	0.01
	(-2.3)	(-2.5)	(-0.0)	(0.0)	(0.0)
Other occupation	**0.22**	**0.21**	-0.04	-0.17	0.08
	(8.7)	(5.0)	(-0.6)	(0.6)	(0.5)

Source and Notes: See Table 8.

Though there is no impact of births on the probability of being partnered in wave 3, there is an impact on the probability of being legally married (vs being single or defacto partnered). Compared to all women, women who had a birth are 7 percentage points more likely to be married (column 1). However, those who had a miscarriage are also more likely to be married than those who were pregnant but did not miscarry. This translates into a substantial 24 percentage point direct impact of birth on the marriage rate, with young birth causing a reduction in the probability of being legally married. That is, the estimate of the causal relationship is the opposite of that apparent from the simple association (both are significantly different from zero at the 5% level – though only just).

This paradoxical result stems from a combination of selection and causal effects acting in opposite directions, and can most clearly be understood with reference to Table 5 (on page 10). Ignoring the mixed categories, that table showed the miscarriage and birth groups to have higher rates of marriage than the other groups (no pregnancy and termination). Following the reasoning introduced in Section 2, if women who are more interested in marrying by wave 3 are also more interested in becoming pregnant (possibly because they are already married or engaged in wave 1), then this selection effect would lead to marriage rates increasing across the groups as

S: Prob(marriage | no pregnancy) <
Prob(marriage | termination) <
Prob(marriage | miscarriage) <
Prob(marriage | birth)

(The first term should be read as the probability of marriage given no pregnancy, and so on). On the other hand, if having a birth rather than miscarriage directly reduces the probability of getting married (everything else held constant), and there are no selection effects, then we would expect to find marriage rates related as

D: Prob(marriage | no pregnancy) =
Prob(marriage | termination) =
Prob(marriage | miscarriage) >
Prob(marriage | birth)

The combination of these two effects can thus lead to a marriage rate pattern of

C: Prob(marriage | no pregnancy) <
Prob(marriage | termination) <
Prob(marriage | miscarriage) =
Prob(marriage | birth)

This is not quite what we find in Table 5. There, the no pregnancy group has a similar marriage rate to the termination groups (actually slightly smaller), suggesting that the first stage of the selection effect, S, is weak. Nonetheless, the combined effect, C, does imply that marriage may be more common among those who give birth than the average of those who do not, while at the same time the direct effect of birth (D) is to make mothers less likely to get married.

As noted above, there is no significant corresponding impact of birth on partnership rates (ie including both de facto and de jure marriages). The effect there is in the same direction as for marriage, but much smaller. This implies that the reduction in formal marriage is largely offset by an increase in de facto marriage.

The third marriage status variable measures the impact of a young birth on lone parenthood status at wave 3. The population for this calculation is women with children in wave 3. The all women estimate in column 1 shows mothers with young first births to be 11 percentage points more likely to be lone parents in their late 20s. However, this is much smaller when we restrict attention to the pregnant women sample, and disappears in the miscarriage and

2SLS estimates – clearly fitting the pattern of a relationship that is primarily a selection effect.

However, this result does vary according to the woman's age in wave 1. For those aged 22 at the end of 1996, Table 9 shows that having a child prior to wave 1 does lead to a higher probability of being a lone parent in wave 3. The OLS miscarriage effect is not quite significant at the 5 per cent level, but the logistic regression (a more appropriate test) is.

The lone parent outcome variable here is only defined for women who have children in wave 3. So this result should be interpreted as suggesting that having a child in your early rather than late 20s makes you more likely to be a lone parent around age 30.

With respect to education, there is evidence of strong selection effects, but no significant miscarriage effect. Women with higher levels of education (by wave 3) are much more likely to give birth, even within the sub-sample of women who had a pregnancy before wave 1. Having a miscarriage tends to increase educational outcomes by a small amount, but this is not significantly different from zero in any cases. The largest miscarriage effect is for year 10 completion, where year 10 completion is 3 percentage points higher, but the t statistic for this is only 1.3 (or 1.2 in the logistic regression model).[17]

As with all the insignificant results here, it is possible that there is a direct impact of young childbirth, but our sample size here is insufficient to identify it. For example, the most likely educational impact might be expected to be on year 12 completion. Here the birth effect is estimated as reducing year 12 completion by 6 percentage points, but with a standard error of around 10 percentage points (the estimate divided by the t statistic). If this precision were to be maintained with a larger effect, we would need to see a direct effect of around a 20 percentage point reduction before it would appear as significant here. This is not implausible and indeed it is similar to the observed association between birth and education shown in the first two columns of Table 8. So, in this case, if the birth effects shown in the first two columns were in fact entirely due to a direct effect then they would probably appear as statistically significant in the last three columns.

Women giving birth prior to wave 1 are much less likely to be employed in wave 3, but there is no miscarriage association. Indeed, much of this association simply reflects the correlation of birth with the presence of children in the household in wave 3. When this is controlled for the associations in columns 1 and 2 are much reduced but still significant, probably reflecting the selection effects apparent for education.

The patterns for occupation reflect those in the earlier Table 6. Among those with jobs, women with young births are much less likely to have jobs requiring educational qualifications, but are just as likely to have managerial jobs – reflecting the education selection effects described above. There is some suggestion that having a miscarriage might increase the chance of having a managerial job (almost significant) vs a job in the 'other' categories.

[17] Because of the cumulative definitions of income, the year 10 effect is also included in the other variables. The origins of this pattern can be seen in the first line of Table 6, where the birth groups have many more women in the lowest educational level than the miscarriage group. The assumptions introduced in Section 4 imply that, in the absence of a direct birth effect, the miscarriage group should be an average of the termination and birth groups. However, the birth group has a higher no-qualification rate than expected on the basis of this – implying a direct effect. However we cannot discount the possibility that this result has arisen from sampling variation.

Table 10 presents similar results based on the income variables collected in the wave 3 survey. Three different income concepts are defined: own income, household income, and other income, calculated as the difference between the first two variables. The household income and other income variables are only defined for partnered women. Results are also presented where the dependent variable is a binary variable flagging low incomes (as defined in Section 5.1) and also where the number and age of the children in the household in wave 3 are controlled for.

Table 10 The Impact of Having a Child Prior to Age 20: Income Effects

Dependent variable (wave 3, except for birth)	**All women**	**Women with at least one pregnancy prior to wave 1**			
	Birth effect	Birth effect	Miscarriage effect		Birth effect
	OLS	OLS	OLS	Logistic	2SLS
	('t' statistics in parentheses)				
	1	2	3	4	5
Own income	**-170**	**-126**	10.7		-22.7
	(-6.6)	(-4.1)	(0.3)		(-0.3)
Own income (cont. for children)	46.6	-8.1	6.8		-7.9
	(1.9)	(-0.3)	(0.2)		(-0.1)
Low personal income	**0.16**	**0.09**	0.03	0.14	-0.07
	(5.4)	(2.1)	(0.6)	(0.6)	(-0.6)
Low p. income (cont. for children)	**-0.06**	-0.01	0.02	0.11	-0.05
	(-2.2)	(-0.1)	(0.4)	(0.4)	(-0.4)
Household income (partnered only)	**-174**	**-144**	-66.0		125
	(-5.0)	(-3.0)	(-1.0)		(1.0)
Household income (cont. for children)	-16	-58	-94.41		189
	(-0.5)	(-1.1)	(-1.5)		(1.4)
Low household income	**0.18**	**0.16**	0.04	0.17	-0.07
	(4.7)	(3.0)	(0.5)	(0.6)	(-0.5)
Low h. income (cont. for children)	0.03	0.09	0.06	0.29	-0.13
	(0.8)	(1.5)	(0.9)	(0.9)	(-0.9)
Other income (partnered only)	9	8	-93.95		183
	(0.3)	(0.2)	(-1.6)		(1.6)
Low other income	0.02	0.02	0.03	0.14	-0.06
	(0.5)	(0.4)	(0.4)	(0.4)	(-0.4)
SEIFA change	0	2	3.00		-6
	(-0.0)	(0.4)	(0.6)		(-0.6)

Source and Notes: See Table 8.

In both the population of all women (first column) and of women who experienced a pregnancy (column 2), women who had a birth prior to wave 1 have a much lower income than women who did not. However, much of this association simply reflects the fact that these women are more likely to have children in the household in wave 3. Controlling for this, women with a birth have a higher income (all women, not quite significant) or no difference (women with pregnancies). The miscarriage effect (third column) is small and not significantly different from zero.

Household income (for partnered women) is also significantly lower, but not if we control for the presence of children. The miscarriage effects are not significantly different from zero.

The last income panel ('other income') attempts to focus on the income of the women's partners, though the indirect method used to calculate partner income no doubt introduces noise into the data. Here we find essentially no relationship in any of the models. Note, however, that 'other income' is almost significant. In the weighted estimates (see footnote 15) this parameter is in fact significant at the 5 per cent level. Having a miscarriage leads to a lower partner income (among those with partners). This implies that a young birth actually increases partner incomes (among those with partners).[18] This is the opposite of the result found by Ermisch and Pevalin in the UK.

Finally, the last panel looks at the change in the SEIFA scores of the post code regions in which the women live (for women who do not move, this variable is zero). No significant effects are found.

Table 11 The Impact of Having a Child Prior to Age 22: Income Effects

Dependent variable (wave 3, except for birth)	**All women**	**Women with at least one pregnancy prior to wave 1**			
	Birth effect	Birth effect	Miscarriage effect		Birth effect
	OLS	OLS	OLS	Logistic	2SLS
	('t' statistics in parentheses)				
	1	2	3	4	5
Own income	**-197**	**-170**	0.5		-0.2
	(-10.4)	(-6.7)	(0.0)		(-0.0)
Own income (cont. for children)	-14.3	**-61.5**	-25.4		67.3
	(-0.7)	(-2.2)	(-0.7)		(0.7)
Low p. income	**0.15**	**0.14**	0.00	-0.03	0.01
	(7.0)	(4.0)	(-0.1)	(0.1)	(0.1)
Low p. income (cont. for children)	-0.02	0.06	0.01	0.04	-0.03
	(-0.9)	(1.6)	(0.2)	(0.2)	(-0.2)
Household income	**-219**	**-192**	4.19		-13
	(-8.9)	(-5.1)	(0.1)		(-0.1)
Household income (cont. for children)	**-90**	**-122**	-9.16		18
	(-3.4)	(-2.9)	(-0.2)		(0.1)
Low h. income	**0.21**	**0.20**	-0.01	-0.03	0.02
	(7.9)	(4.7)	(-0.1)	(0.1)	(0.1)
Low h. income (cont. for children)	**0.10**	**0.14**	0.00	0.02	-0.01
	(3.3)	(3.0)	(0.1)	(0.1)	(-0.0)
Other income	-31	-30	-24.48		47
	(-1.4)	(-0.8)	(-0.5)		(0.4)
Low other income	**0.08**	0.07	0.00	-0.01	0.01
	(2.6)	(1.6)	(-0.0)	(0.0)	(0.1)
SEIFA change	1	0	0.74		-1
	(0.4)	(-0.1)	(0.1)		(-0.1)

7 Caveats

7.1 Non-randomness of miscarriages

These estimates assume that, controlling for a small number of characteristics, miscarriages are random with respect to the social background factors that might also influence socio-economic outcomes. In their survey, Kline, Stein and Susser (1989) find evidence that smoking and alcohol consumption increase the risk of miscarriage. With respect to socio-economic factors more generally, they do not find strong evidence of an association, but find no strong evidence to disprove such a link either.

[18] The difference between the weighted and unweighted results, implies that this relationship primarily holds in urban areas.

If social disadvantage is associated with a higher probability of miscarriage, then the impacts estimated above are most likely to be too small (in absolute terms). The impact of young childbirth on marriage, for example, is proportional to the probability of being married for those with no miscarriages minus the probability for those with a miscarriage (see equation (1) on page 15). In the data presented here, this difference is negative, implying that young childbirth decreases the probability of marriage. If background disadvantage implied both miscarriage and non-marriage this would act in opposite direction, implying that the true effect would be more negative than that estimated here.[19]

Some exploratory analysis has been undertaken to test whether miscarriages are associated with the socio-economic background factors recorded in the survey. However, even with respect to the observed background factors, random occurrence is difficult to test directly. This is because observed miscarriages are influenced by the probability of abortion – which we have strong reasons to believe will have strong associations with social background (see the discussion in Section 3). A pregnancy which is going to end as a mid or late-term miscarriage might instead be terminated before this, in which case it will be recorded in our data as an abortion rather than a miscarriage. The probability of observing a miscarriage (vs a birth or termination) will thus be negatively related to the factors that cause terminations.[20] Nonetheless, most miscarriages occur earlier in the pregnancy than abortions, and so this bias might thus be only small.

Using the assumption that this is the case, the probability of miscarriage as a function of observable characteristics was examined. The observable background factors included: age at end 1996, whether of Aboriginal or Torres Strait Islander origin, whether they described how they spoke English as other than 'very well', their father's occupation when they were aged 14 (8 categories), whether they lived in a remote or very remote region and the ABS SEIFA index of disadvantage of the region. Also included were indicators of smoking and alcohol consumption (years smoking, whether smoked daily and how many cigarettes smoked, and whether their drinking behaviour scored at above the negligible risk level). These latter variables are not necessarily exogenous and might be influenced by fertility outcomes. (Indeed a separate analysis shows the drinking indicator to decline significantly among women with childbirths).

In any event, however, these variables as a group do not have a statistically significant relationship with miscarriage. A logistic regression on the population of women with only one pregnancy prior to wave 1 cannot reject the hypothesis that all parameters other than the intercept are equal to zero (likelihood ratio χ^2=16.5, df=16, p=0.41). Variables significant at the individual 5 per cent level are age at end 1996 and English proficiency. The latter arises because 3 of the 7 women describing themselves as having not very good English also had a miscarriage. With such a small group, the asymptotic hypothesis test might not be accurate, and indeed an exact test of the bivariate relationship does not reject the null hypothesis of no association at the 5 per cent level (an additional analysis including country of birth found no significant relationship). The age relationship is perhaps more puzzling as it implies that

---

19 Hotz et al (1997) show how these estimates can be bounded using assumptions about the proportion of miscarriages likely to be random (and taking account of misreporting). Such bounds are likely to be uninformative for the estimates here (which are only just statistically significant).

20 An alternative approach might be to observe the probability of miscarriage conditional on the probability of not having a termination. However this will be biased in the opposite direction, as women with lower propensities for terminations will predominate in the sample, and these have different characteristics to the average woman.

older women are less likely to have a miscarriage – which is contrary to the epidemiological patterns identified in Kline et al. Note however, that age here is age at interview rather than age at miscarriage. Given the acceptance of the overall test however, the most likely explanation is that this reflects sampling variation.

Though there does not seem to be strong evidence of non-randomness of miscarriages in this sample, these background factors are used as controlling variables in the results shown above.

7.2 Reporting error

There are three aspects of the survey collection methodology that may bias the results presented above.

- Survey non response,
- non-reporting of miscarriages and
- non-reporting of terminations.

Like many postal surveys, the response rate was relatively low, with people changing address and those with less education less likely to complete the survey. The results, therefore, should be considered to reflect the impact of young childbirth in a population more heavily weighted towards those who had not moved address and with higher education levels (though recall that the data also weights towards the non-urban population). It is not clear, however, in which direction this result will differ from a relationship that was estimated across the whole population.

Among those responding, the two most likely sources of bias are the non-reporting of terminations and the non-reporting of miscarriages. These might bias the results if the average outcomes for the non-reporters differed from that of the reporters.

The key results all stem from the estimate of the miscarriage effect on each outcome variable (equation (1)). In the tables above, this is estimated the average outcome for women who had a miscarriage minus the average outcome for women who were pregnant but did not have a miscarriage (while controlling for other variables).

Women who had a single pregnancy that miscarried, but did not report it in the survey, would be recorded in neither of these comparison groups – though they should be included in the first. Taking the example of the marriage outcome variable, it is possible that women with less education might be less likely to be aware of a miscarriage and also less likely to be legally married. This means that the average marriage rate for women with a miscarriage would be lower if these women were in fact included in this group. The results above show that women with miscarriages are more likely to be married, so this pattern of missing data could mean that this is an overestimate of the impact of miscarriage. However, this bias is likely to be small. Its size depends upon the product of two factors: the fraction of women who had a miscarriage but did not report it, and the difference of their marriage rate from that of the women who did report their miscarriage. If either of these factors is small, then the bias will be small, but we don't have any direct evidence on this.

Non-reporting of terminations affects the second of these comparison groups, women who were pregnant but did not have a miscarriage (ignoring multiple pregnancies for simplicity). If the type of women who were not likely to report a termination were also more likely to be married, then the average marriage rate for this comparison group would be higher if they were to be included. Again, this could possibly be an explanation for the relationship found. However, here the same number of non-reportings will have a smaller impact, because it would have an impact in proportion to its fraction of all women with pregnancies that don't miscarry (a larger number than the number of miscarriages).

If non-reporting is small then neither of these two biases will have much impact. If non-reporting is substantial, then they will only have an impact if the outcome variables differ substantially between the reporters and non-reporters. I speculate that this is probably not an important problem in this data but the possibility of these biases must be included as a caveat on the results.

8 Discussion

Being a young mother is strongly associated with poorer socio-economic outcomes in later life. However, the results in this report suggest that, for economic outcomes at least, this is primarily a selection effect. Women with characteristics that mean they are less likely to do well in education or the labour market are more likely to undertake (or be subjected to) behaviour that might lead to pregnancy and less likely to have a termination if they do get pregnant. The reasons for this reflect their capabilities, their preferences and the opportunities available to them.

To the extent to which this selection effect is the main reason for poor later life outcomes of young mothers, there is little point in discouraging these women from having children.[21] Even if they did not have children the evidence here suggests that they would have had similar education and labour market outcomes. Though the sample size of the ALSWH means that the estimates here are not very precise, they do confirm similar results found the US and UK.

However, this conclusion that young motherhood does not have a causal impact upon education and labour market outcomes might seem paradoxical. In this data, approximately one third of pregnancies prior to age 23 ended in an abortion. If births do not have an impact on later life outcomes, why do so many women go through what can at best be described as a very unpleasant experience?

This contradiction is resolved if we realise that the estimate of no impact applies to a particular population: women who get pregnant and do not intend to have an abortion. This is a population of particular policy relevance; these are the women who would be targeted by any policy designed to reduce fertility at young ages. Nonetheless, most women do not fall into this population, and there is reason to believe that the impact of having a child when young will be greater for other women. Many of the women who do not want to get pregnant, or who do not wish to have a child if they do, have this intention precisely because they believe motherhood will have an impact on their educational or labour market

[21] Unless we wish to support eugenicist arguments that we wish to discourage disadvantaged women from having too many children.

outcomes. Those who don't think there will be an impact are more likely to want to have a child. So the result here might be seen as simply confirming this expectation.[22]

In any event, the results here do point to some impacts in areas other than education and employment. In particular, some evidence of an impact on partnering outcomes is found. Being a young mother reduces the likelihood of being legally married when aged in the late 20s, and having a child in the early rather than late 20s leads to a greater likelihood of being a lone parent at around age 30.

The marital status result may reflect the fact that young motherhood reduces the likelihood of being in a relationship with the father of their child – because the relationships existing in the late 20s are more likely to have been formed after the first childbirth. If this is the case, it might have implications for child outcomes. Unfortunately this cannot be directly tested in our dataset, and other data sources that do contain this information do not include the information on miscarriage that would enable the identification of causal links.

Though they didn't find this particular link, Ermisch and Pevalin in the UK also found impacts on partnering patterns. Women who were young mothers ended up with lower income husbands by age 30. We don't find this result here, and in some specifications find that the young mothers actually have higher income partners.

Several potential mechanisms could explain the association with lone parenthood. These include an argument that women with children are generally less successful in the marriage market – consistent with Ermisch and Pevalin's results. However, this result might also be due to interactions between marital dynamics and child ages. Those who have their children in their late 20s will have young children when aged 30. Their relatively low lone parenthood rates might simply be due to lower rates of separation in families with young children (see Bradbury and Norris, 2005). This is an issue that might be resolved when future waves of the ALSWH data become available.

The most obvious extension to the work included here would thus be a reanalysis when future waves are undertaken. For many of the outcomes (eg labour market status) wave 3 is still too early to identify long-term impacts. Another extension would be the use of subsidiary data sources to try to identify the extent of non-reporting of miscarriages and terminations and the impact of these on the estimated results.

Finally, it might be possible to use a similar approach to examine the characteristics of the *children* of young mothers. That is, one might compare the outcomes of children whose mother had previously miscarried at a particular age with the outcomes of children who were born when their mother was this age. No existing Australian dataset contains sufficient information on both children and mothers, but it might be possible to collect this by supplementing one of the existing longitudinal surveys with retrospective information on pregnancies and miscarriages.

---

[22] This does not necessarily imply that pregnancies and births are chosen. If constraints on choice (eg lack of access to fertility control services and/or religious prohibitions) were more likely among more disadvantaged women then this could lead to the same patterns found here.

References

Angrist, J. D., G. W. Imbens and D. B. Rubin (1996), 'Identification of causal effects using instrumental variables' *Journal of the American Statistical Association* 91(434, June): pp. 444-55.

Angrist, Joshua D. and Alan B. Krueger (2001), 'Instrumental variables and the search for identification: From supply and demand to natural experiments' *Journal of Economic Perspectives* 15: pp. 69-85.

Bradbury, B. (2006, forthcoming), 'Disadvantage among Australian young mothers', *Australian Journal of Labour Economics* 9(2).

Bradbury, B. and K. Norris (2005), *Family Dynamics in Australia* Final Report of the Fluid Families Project, Report for the Department of Family and Community Services under the Deed of Agreement for the Provision of Social Policy Research and Services, Social Policy Research Centre, June.

Ermisch, J. and D. J. Pevalin (2003), 'Does a "teen-birth" have longer-term impacts on the mother? Evidence from the 1970 British Cohort Study', *ISER Working Paper* No. 2003-28.

Ermisch, J. and D. J. Pevalin (2005), 'Early motherhood and later partnerships', *Journal of Population Economics,* 18: pp. 469-89.

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, pp. 236-9

Holmlund, H. (2005), 'Estimating long-term consequences of teenage childbearing; An examination of the siblings approach', *Journal of Human Resources,* 40(3): pp. 716-743.

Hotz, V. J., S. McElroy and S. Sanders (2005), 'Teenage childbearing and its life cycle consequences: exploiting a natural experiment', *Journal of Human Resources,* XL (3): pp. 683-715.

Hotz, V. Joseph, Charles H. Mullin and Seth G. Sanders (1997), 'Bounding causal effects using data from a contaminated natural experiment: analyzing the effects of teenage childbearing, *Review of Economic Studies,* 64: pp. 575-603.

Kline, Jennie, J. Stein and M. Susser (1989), *Conception to Birth: Epidemiology of Prenatal Development,* Monographs in Epidemiology and Biostatistics, Volume 14, OUP, New York.

Lee, C., A. J. Dobson, W. J. Brown, L. Bryson, J. Byles, P. Warner-Smith and A. F. Young (2005), 'Cohort profile: the Australian longitudinal study on women's health', *International Journal of Epidemiology*, (published May 13).

Stewart, J. (2003), 'The mommy track: the consequences of gender ideology and aspirations on age at first motherhood', *Journal of Sociology and Social Welfare,* 30(2): pp. 3-30.

UNICEF (2001), *A League Table of Teenage Births in Rich Nations,* Innocenti Report Card, Issue No. 3, July. Florence: UNICEF Innocenti Research Centre http://www.unicef-icdc.org/publications/index.html

Wolfe, B., K. Wilson and R. Haveman (2001), 'The role of economic incentives in teenage nonmarital childbearing choices', *Journal of Public Economics,* 81, pp. 473–511.