Negative socioeconomic selection and positive health selection in U.S. teenage fertility: Evidence from North Carolina linked birth certificate data

Poh Lin Tan and Philip J. Cook[1]

Abstract

This paper contributes to the literature on selection into teenage fertility in three ways. First, it is based on a population-level administrative dataset rather than survey data. Second, it provides evidence that paternal characteristics at birth matter independently of maternal characteristics for selection into teenage fertility. In particular, we find that paternal education have significant effects independently of maternal characteristics. We also show that relative parental characteristics matter – *cet. par.*, daughters born to white families in which fathers are significantly older or more educated than mothers are more likely to have a teenage birth. Third, it presents evidence of positive health selection into teenage fertility, in the sense that low-weight infants are less likely to give birth as teenagers.

[1] Sanford School of Public Policy, Duke University.

Introduction

There is strong evidence that women born to mothers who were younger at first birth are more likely to give birth as teenagers, especially among whites. Accounting for maternal education and marital status reduces but cannot fully explain the predictive power of maternal age. On the other hand, there is little evidence that biological characteristics play any role; for example, numerous studies using a diverse range of datasets find no association between teenage fertility and age at menarche (Hardy et al. 1998; Kahn and Anderson 1992; Manlove 1997; Meade, Kershaw and Ickovics 2008).

In this paper, we explore the importance of socioeconomic and biological characteristics in determining selection into teenage childbearing. Our paper contributes to the existing literature in three ways. First, it is based on a large population-level administrative dataset rather than longitudinal or retrospective survey data. Second, it provides evidence that paternal characteristics at birth and relative parental characteristics matter for selection into teenage fertility. Third, we present evidence that infant health as measured by birth weight is highly predictive of teenage childbearing, thereby providing fresh insight into the role of biology. Our findings are consistent with studies in the epidemiological literature, which find that poor infant health may reduce future fertility due to lower levels of organ development and higher childhood mortality.

Why is it important to understand selection into teenage childbearing? One of the most important reasons is that teenage childbearing is associated with negative life outcomes for both mother and child, and is hence of policy concern. A greater understanding of selection into teenage childbearing could inform policymakers about the socioeconomic groups at high risk and increase the effectiveness of policy interventions. In this study, we provide new evidence on socioeconomic selection into teenage childbearing. We find, consistent with the existing literature, that teenagers are negatively selected into childbearing in terms of parental socioeconomic characteristics. Unlike previous work (Kahn and Anderson 1992; Manlove 1998), however, we find much weaker effects of maternal age at first birth for blacks, with black teenagers born to older mothers just as likely to give birth as black teenagers born to teenage mothers.

Also unlike previous work, which has focused almost exclusively on maternal characteristics and found little evidence that paternal characteristics mattered to teenage

childbearing (Harris, Furstenberg and Marmer 1998), we provide strong evidence that paternal education at birth does matter, especially for whites. Our new results are compatible with recent findings that the quantity and quality of father-child interactions are predictive of teenage fertility behavior (Guilamo-Ramos et al. 2012; Ellis et al. 2003). We also find that relative parental characteristics matter – daughters born to white families in which fathers are older or more educated than mothers are more likely to have a teenage birth, *cet. par*.

Finally, we provide evidence of health selection into teenage childbearing. In particular, we show that very unhealthy infants as indicated by very low birth weight (less than 1500 grams) are much less likely to give birth as teenagers, and that those with low birth weight (1500-2500 grams) are somewhat less likely to have a teen birth, *cet. par*. Since poor infant health is associated with negative consequences for future labor market and other outcomes (Black, Devereux and Salvanes 2007), our results suggest that selection into teenage childbearing may be less negative than implied by the literature on negative socioeconomic selection.

Data

Our dataset was obtained from the North Carolina Education Research Data Center with permission from the NC Department of Health and Human Services. The dataset contains birth certificate data on six birth cohorts of females born in North Carolina between 1987 and 1992. (For simplicity, the six cohorts are hereafter referred to as "Gen2"; the parents of Gen2 are referred to as "Gen1"; and the babies born to Gen2 are referred to as "Gen3".) The birth certificate data include information on Gen1 characteristics and maternal health behavior, as well as Gen2 infant health, discussed in more detail below. (No information on maternal smoking or drinking is available for the cohort born in 1987.) Using randomized individual identifiers for Gen2, the birth certificate data are then linked to their public school system educational records as well as Gen3 birth certificate data. Unfortunately, members of Gen2 who are not in the public educational system are not assigned identifiers. Since these members are disproportionately born to white and highly educated women, the final sample over-represents individuals born to black and less educated women.

One drawback of this dataset is that age at first birth is not reported on birth certificates unless, of course, the current birth is the first. Since the important role of age at first birth has been established in the literature, we restrict our sample to members of Gen2 who were first

births. We also restrict our sample to births to non-Hispanic white and black women, who form the vast majority of the observations. (Our 1987 cohort includes a small number of Hispanics, since no information on ethnicity was collected for this cohort.) Around 20% of birth certificates do not contain any information about father's age, education or race; we include these observations in our analyses by constructing an indicator for observations with missing data. Our final sample consists of 74,996 female singleton Gen2 individuals born to residents of North Carolina who attended public school in NC from grade 3 to age 15, as well as an additional 957 individuals with infant death certificates. We include the latter to assess the importance of increased mortality risk as one of the mechanisms in positive health selection.

We use Vital Statistics (1995) guidelines to correct for implausible birth weights. Only a few observations were changed, reflecting the high quality of the birth certificate data. We compare our summary statistics to those published by Vital Statistics (which are based on births born to NC residents rather than births in the state) and find very similar estimates in terms of the racial, age and educational composition of mothers and proportion of babies with low birth weight. Our preterm rates based on estimates of gestational age, however, are much lower (approximately 10% compared to 13% according to Vital Statistics), and the discrepancies persist whether gestational age is calculated based on clinical estimates, last menstrual period or a combination of both methods. Hence, our analysis uses only birth weight to represent infant health using four commonly used categories: a) high birth weight (>4500g), b) normal birth weight (2500-4500g), c) low birth weight (1500-2499g) and d) very low birth weight (<1500g). Birth weight is one of the most widely used international indicators of infant health due to its strong associations with infant mortality, morbidity and fetal growth (OECD 2011). There are two additional technical advantages to using birth weight: it is likely to be a more reliable indicator of infant health than gestational age since it is directly measured at the time of birth rather than estimated based on reported data, and very few observations have missing values.

Table 1 provides the summary statistics. Since the last year of available data was 2009, we have the complete teenage fertility history of only four birth cohorts (1987 to 1990), and only the early teenage fertility history (i.e. births before age 18) of the last two cohorts. Gen2 members who give birth before the age of 20 are much more likely to have unmarried parents with less than high school education, and much less likely to have parents with college education

Table 1: Summary statistics

	% of sample		
	All	**Gen2 first birth before age 20 (cohorts 1987-1990)**	**Gen2 first birth before age 18 (All cohorts)**
Gen1 socioeconomic characteristics			
Mother's age and education			
< 19	20.43	37.57	41.72
19-22 and less than high school	5.63	10.03	10.53
19-22 and high school	17.80	22.59	20.80
19-22 and some college	5.25	3.64	3.59
> 22 and less than high school	2.50	3.80	4.30
> 22 and high school	18.98	14.49	12.72
> 22 and some college	13.96	5.82	4.51
> 22 and college	15.15	1.90	1.53
Mother's age			
23-24	11.57	9.71	8.18
25-29	25.38	12.41	11.31
30-34	10.78	3.26	3.09
> 34	2.85	0.63	0.50
Father's age and education			
< 19	3.74	6.12	6.91
19-22 and less than high school	5.31	9.47	10.04
19-22 and high school	9.16	10.23	9.04
19-22 and some college	1.70	1.12	0.90
> 22 and less than high school	7.52	10.93	10.38
> 22 and high school	25.30	18.31	15.89
> 22 and some college	12.18	4.59	3.64
> 22 and college	13.00	1.65	1.24
Father's age			
23-24	10.01	9.77	8.75
25-29	25.56	16.33	13.86
30-34	14.87	6.58	5.85
> 34	8.17	3.22	3.19
Gen1 relationship characteristics			
Mother's marital status			
Married	63.64	43.85	37.27
Unmarried and father known	16.90	22.08	24.70
Unmarried and father unknown	19.46	34.07	38.03
Relative characteristics			
Different race	1.68	1.92	2.42
Father older by > 1 year	45.08	40.56	77.45
Parents' age differ by < 3 years	27.43	19.52	18.20
Father younger by > 1 year	7.20	4.45	4.35
Father more educated by > 1 year	11.65	10.77	53.05
Parents' education differ by < 3 years	49.43	39.05	36.09
Father less educated by > 1 year	16.49	11.87	10.87

Gen2 health			
Birth weight			
< 1500g	1.74	0.71	0.53
1500g-2499g	7.15	7.77	7.80
2500g-4500g	90.26	90.95	91.04
> 4500g	0.86	0.58	0.63
Gen2 academic performance			
3rd grade			
Verbal Z-score < -1	11.23	19.70	20.48
Verbal Z-score between -1 and 1	68.71	73.18	73.81
Verbal Z-score > 1	20.06	7.12	5.71
Verbal Z-score unknown	1.52	1.94	2.08
Math Z-score < -1	13.40	22.79	23.46
Math Z-score between -1 and 1	69.87	71.12	71.82
Math Z-score > 1	16.72	6.08	4.73
Math Z-score unknown	1.44	1.78	1.90
Gifted	7.00	1.60	1.48
Handicapped	7.27	8.61	8.55
8th grade			
Verbal Z-score < -1	10.63	19.17	21.44
Verbal Z-score between -1 and 1	71.30	75.43	74.50
Verbal Z-score > 1	18.07	5.39	4.06
Verbal Z-score unknown	1.21	1.58	1.90
Math Z-score < -1	13.03	23.86	27.02
Math Z-score between -1 and 1	68.87	71.38	69.05
Math Z-score > 1	18.10	4.76	3.93
Math Z-score unknown	1.23	1.61	2.06
Gifted	19.52	5.40	3.95
Handicapped	3.56	4.76	4.51
Skipped 1 year	0.60	0.29	0.37
On time	85.25	78.29	74.75
Held back by 1 year	12.90	18.61	22.17
Held back by 2 years	1.25	2.81	2.71
Number of observations	75,953	7,647	3,794

Table 2: Proportion of Gen2 who had a teenage birth

A. Mother's age and marital status

	Mother's age at first birth				
Mother's marital status	> 19	19-22	23-29	30-34	> 34
Married	0.248	0.161	0.071	0.037	0.033
Unmarried and father known	0.273	0.212	0.172	0.147	0.082
Unmarried and father unknown	0.300	0.246	0.200	0.165	0.073
Average	0.279	0.191	0.089	0.048	0.038

B. Father's education and mother's marital status
(Mother gave birth as a teen)

	Father's education			
Mother's marital status	Less than high school	High school	Some college	College
Married	0.292	0.184	0.125	0.129
Unmarried and father known	0.303	0.231	0.157	0.059
Unmarried and father unknown	-	-	-	-
Average	0.297	0.203	0.134	0.104

C. Father's education and mother's marital status
(Mother gave birth at ages 20-24)

	Father's education			
Mother's marital status	Less than high school	High school	Some college	College
Married	0.224	0.125	0.081	0.041
Unmarried and father known	0.235	0.189	0.131	0.113
Unmarried and father unknown	-	-	-	-
Average	0.226	0.135	0.088	0.048

or to know their biological fathers. (We define fathers as unknown if mothers are unmarried and no information about father's age, race and education are recorded in the birth certificate.) They are also much less likely to have very low birth weight, even though (as we show below) births to younger and less educated women are generally at higher risk of having very low birth weight. Despite their relatively strong infant health, Gen2 teenage mothers have much lower test scores in the third grade, with little evidence of catching up by the eighth grade (the last year in which test scores are highly comparable across students in the same grade). The persistence of lower test scores at the eighth grade is particularly remarkable given that Gen2 teenage mothers are much more likely to have been held back a year or two and hence to be taking the tests at an older age than other students.

Table 2 shows that some Gen1 socioeconomic characteristics may be particularly predictive of selection into teenage childbearing. Part A of the table shows, as expected, that women born to younger and unmarried mothers are more likely to have a teenage birth, but that girls born to unmarried women in their 20s are still less likely to have teenage births than those born to married teenage mothers. Parts B and C show that father's education matters among both women born to teenage mothers and women in their early 20s – in both groups, girls with fathers with more than high school education are less likely to have teenage births than girls born to women who first gave birth in their late 20s. In the next section, we describe a more formal way to assess the impacts of Gen1 and Gen2 characteristics on selection into teenage childbearing.

Methods

We use a logistic regression model where the outcome (giving birth before age 18 or 20) and predictors (Gen1 characteristics and Gen2 infant health) are dichotomous variables. We begin by testing for socioeconomic selection and health selection separately before presenting the results of the full model. We also test whether Gen2 early and later educational outcomes are important mediators of socioeconomic and health selection by controlling for individuals' verbal and math Z-scores. In all of the models, we control for cohort fixed effects. All estimates are presented as odds ratios relative to the omitted class.

To explore the relationship between Gen1 socioeconomic characteristics and Gen2 infant health, we use an ordered logistic model where infant health is an ordered rather than dichotomous variable (normal birth weight = 0, low birth weight = 1 and very low birth weight =

2). We also use this ordered logistic model to examine whether socioeconomic and health selection is associated with Gen3 infant health, conditional on Gen2 having a teenage birth. All estimates are presented as proportional odds ratios, i.e. they give the ratio of a) the odds that an individual born to a mother with age at first birth = *x* will have very low birth weight and b) the odds that she will have low or higher birth weight. By the proportional odds assumption, this ratio is also equal to the ratio of a) the odds that she will have low birth weight and b) the odds that she will have normal or higher birth weight.

Results

Table 3 shows that there is negative socioeconomic selection into teenage childbearing in terms of maternal age and parental education, consistent with previous literature. As previously mentioned, these results are based on data for the first four cohorts only, since only these cohorts have completed their teenage childbearing by the last year of available data. The results are qualitatively similar for teenage childbearing under 18 (not shown). Unlike previous literature (Kahn and Anderson 1992; Manlove 1998), however, we find a much stronger moderating effect of race on the intergenerational patterns: black teenagers born to older mothers are no less likely to give births than black teenagers born to teenage mothers (after adjusting for maternal education). Interestingly, we also find that maternal age appears to have less of a protective effect for women born to the least educated mothers for both races.

Another interesting result from Table 3 is that fathers' characteristics matter: Gen1 paternal education in Gen 1 is predictive of Gen2 teenage childbearing in Gen 2 even after controlling for maternal education, especially for whites. On the other hand, Gen1 paternal age is much less predictive of Gen2 teenage fertility than Gen1 maternal age, possibly because paternal age may not reflect age at first birth. Moreover, Tables 3 and 4 suggest that parental relationships may matter for teenage fertility outcomes. From Table 3, Gen2 women born to unmarried mothers are significantly more likely to have a teenage birth. In addition, controlling for parental age and education, white females born to families in which the father is at least two years older or has at least two more years of education than the mother are more likely to give birth as teenagers. Finally, results in Table 4 indicate that, as expected, the protective effects of paternal age and education are stronger when parents are married – what may be somewhat surprising is that maternal age and education also matter more when parents are married.

Table 3: Impact of relative parental race, age and education on socioeconomic selection into teenage fertility by race
(dependent variable = giving birth before age 20)

	% of sample	All	Whites	Blacks
Gen1 maternal characteristics				
Mother's age and education				
< 19	37.57	-	-	-
19-22 and less than high school	10.03	1.059	1.127*	0.874
19-22 and high school	22.59	0.775***	0.761***	0.757***
19-22 and some college	3.64	0.508***	0.478***	0.490***
> 22 and less than high school	3.80	1.033	0.977	1.171
> 22 and high school	14.49	0.636***	0.571***	0.712***
> 22 and some college	5.82	0.463***	0.437***	0.441***
> 22 and college	1.90	0.228***	0.211***	0.214***
Mother's age				
23-24	9.71	-	-	-
25-29	25.75	0.819***	0.789***	0.887
30-34	10.35	0.653***	0.586***	0.769*
> 34	2.51	0.504***	0.531***	0.410***
Gen1 paternal characteristics				
Father's age and education				
< 19	6.12	-	-	-
19-22 and less than high school	9.47	1.254***	1.199**	1.286*
19-22 and high school	10.23	0.758***	0.666***	0.880
19-22 and some college	1.12	0.478***	0.375***	0.664**
> 22 and less than high school	10.93	1.231***	1.196**	1.095
> 22 and high school	18.31	0.727***	0.636***	0.903
> 22 and some college	4.59	0.489***	0.421***	0.696**
> 22 and college	1.65	0.248***	0.205***	0.470***
Father's age				
23-24	9.77	-	-	-
25-29	26.15	0.906*	0.886**	1.077
30-34	14.67	0.910	0.901	1.135
> 34	7.74	0.883	0.919	0.966
Gen1 relationship characteristics				
Mother's marital status				
Married	65.15	-	-	-
Unmarried and father known	15.06	1.380***	1.155***	1.279***
Unmarried and father unknown	19.79	1.548***	1.490***	1.301***
Relative characteristics				
Different race	1.46	1.194*	1.276**	1.293
Father older by > 1 year	45.08	1.122***	1.207***	0.910
Father younger by > 1 year	7.20	1.128	1.124	1.239*
Father more educated by > 1 year	11.65	1.191***	1.236***	1.100
Father less educated by > 1 year	16.49	0.993	0.999	1.031
Cohort fixed effects controls	-	Yes	Yes	Yes
Pseudo R^2	-	0.105	0.125	0.046
Number of observations	50,297	50,015	34,565	15,450

*Significant at 10% level. **Significant at 5% level. ***Significant at 1% level.

Table 4: Impact of relative parental race, age and education on socioeconomic selection into teenage fertility by marital status
(dependent variable = giving birth before age 20)

	% of sample	Married	Unmarried and father known	Unmarried and father unknown
Gen1 maternal characteristics				
Mother's race				
White	0.681	-	-	-
Black	0.309	1.713***	1.210***	1.007
Mother's age and education				
< 19	37.57	-	-	-
19-22 and less than high school	10.03	1.124	0.874	1.028
19-22 and high school	22.59	0.705***	0.768***	0.763***
19-22 and some college	3.64	0.439***	0.541***	0.466***
> 22 and less than high school	3.80	1.011	0.845	0.937
> 22 and high school	14.49	0.554***	0.645***	0.695***
> 22 and some college	5.82	0.422***	0.442***	0.426***
> 22 and college	1.90	0.208***	0.256***	0.198***
Mother's age				
23-24	9.71	-	-	-
25-29	25.75	0.805***	0.910	0.891
30-34	10.35	0.618***	0.797	0.834
> 34	2.51	0.541***	0.442*	0.368**
Gen1 paternal characteristics				
Father's age and education				
< 19	6.12	-	-	
19-22 and less than high school	9.47	1.272**	1.185*	-
19-22 and high school	10.23	0.651***	0.859*	-
19-22 and some college	1.12	0.380***	0.606**	-
> 22 and less than high school	10.93	1.171	1.233*	-
> 22 and high school	18.31	0.657***	0.824**	-
> 22 and some college	4.59	0.451***	0.600***	-
> 22 and college	1.65	0.234***	0.385***	-
Father's age				
23-24	9.77	-	-	-
25-29	26.15	0.880**	1.065	-
30-34	14.67	0.838**	1.349**	-
> 34	7.74	0.814*	1.144	-
Gen1 relationship characteristics				
Relative characteristics				
Different race	1.46	1.525***	1.036	-
Father older by > 1 year	45.08	1.171***	0.981	-
Father younger by > 1 year	7.20	1.103	1.234*	-
Father more educated by > 1 year	11.65	1.213***	1.149	-
Father less educated by > 1 year	16.49	1.011	0.956	-
Cohort fixed effects controls	-	Yes	Yes	Yes
Pseudo R^2	-	0.116	0.026	0.017
Number of observations	-	32,628	7,526	9,861

*Significant at 10% level. **Significant at 5% level. ***Significant at 1% level.

We now turn to the evidence for health selection, displayed in Table 5. Results in the second column show that infants who have extremely high or low weight at birth are less likely to give birth as teenagers, while infants who are moderately underweight are more likely to do so. If we exclude Gen2 individuals who received infant death certificates (and therefore did not give birth as teenagers), the odds ratio for infants who have very low birth weight compared to infants who have normal birth weight jumps up from 0.348 to 0.632, so that infant mortality accounts for around 45% of the total differential. Since other very low-birth weight individuals may have left the sample due to early childhood mortality, the true coefficient may be lower than 0.348. Hence, mortality may account for around half of the observed health selection.

After controlling for socioeconomic characteristics, infants who have very high weight are not significantly less likely to have a teenage birth than normal-weight infants, while low birth weight is now associated with lower probability of having a teenage birth. The evidence suggests that, *cet. par.*, underweight infants are less likely to become teenage mothers. While moderately underweight infants are more likely than healthy infants to give birth as teenagers due to their disadvantaged family backgrounds, very underweight infants are less likely than healthy infants to give birth as teenagers *despite* their disadvantaged family backgrounds. Table 5 shows that while socioeconomic selection differs between whites and blacks, positive health selection seems to be fairly similar, with somewhat larger effects for blacks.

Why do we find larger effects for health selection after controlling for socioeconomic characteristics? If babies with low infant health are more likely to be born to families with lower socioeconomic status, and if daughters from these families are at higher risk of teenage birth, we might expect that not controlling for socioeconomic characteristics would result in lower rather than higher odds ratios for individuals who have low birth weight. Indeed, Table 7 shows that low infant health (where low infant health is an ordered categorical variable such that normal birth weight = 0, low birth weight = 1 and very low birth weight = 2) is negatively associated with paternal education. One possible explanation is that low infant health is also positively associated with maternal age, with larger effects for age and smaller effects of parental education for blacks. Our results are consistent with those of Geronimus (1996) and Geronimus and Korenman (1993), who show using a) linked birth certificate and census data for first births in Michigan and b) data from the Longitudinal Study of Youth on sisters respectively that the relationship between maternal age and infant health is negative and especially strong for blacks.

Table 5: Negative socioeconomic selection and positive health selection into teenage fertility (dependent variable = giving birth before age 20)

	% of sample	SE selection	Health selection	Full Model
Gen1 socioeconomic characteristics				
Mother's race				
White	0.681	-		-
Black	0.309	1.299***		1.323***
Mother's age and education				
< 19	37.57	-		-
19-22 and less than high school	10.03	1.088*	-	1.093*
19-22 and high school	22.59	0.766***	-	0.764***
19-22 and some college	3.64	0.492***	-	0.488***
> 22 and less than high school	3.80	1.049	-	1.059
> 22 and high school	14.49	0.631***	-	0.631***
> 22 and some college	5.82	0.453***	-	0.453***
> 22 and college	1.90	0.222***	-	0.220***
Mother's age				
23-24	9.71	-	-	-
25-29	25.75	0.819***	-	0.829***
30-34	10.35	0.651***	-	0.665***
> 34	2.51	0.498***	-	0.515***
Father's age and education				
< 19	6.12	-		-
19-22 and less than high school	9.47	1.287***	-	1.284***
19-22 and high school	10.23	0.747***	-	0.749***
19-22 and some college	1.12	0.471***	-	0.467***
> 22 and less than high school	10.93	1.242***	-	1.244***
> 22 and high school	18.31	0.721***	-	0.717***
> 22 and some college	4.59	0.490***	-	0.485***
> 22 and college	1.65	0.251***	-	0.247***
Father's age				
23-24	9.77	-	-	-
25-29	26.15	0.909*	-	0.907*
30-34	14.67	0.914	-	0.910
> 34	7.74	0.886	-	0.884
Mother's marital status				
Married	65.15	-	-	-
Unmarried and father known	15.06	1.234***	-	1.236***
Unmarried and father unknown	19.79	1.307***	-	1.326***
Relative characteristics				
Different race	1.46	1.261**	-	1.272**
Father older by > 1 year	45.08	1.126***	-	1.130***
Father younger by > 1 year	7.20	1.133*	-	1.136*
Father more educated by > 1 year	11.65	1.201***	-	1.206***
Father less educated by > 1 year	16.49	1.006	-	1.006

Gen2 health				
Birth weight				
< 1500g	2.05	-	0.348***	0.262***
1500g-2499g	7.03	-	1.122**	0.874***
2500g-4500g	89.80	-	-	-
> 4500g	1.12	-	0.632***	0.827
Cohort fixed effects controls	-	Yes	Yes	Yes
Pseudo R^2	-	0.107	0.004	0.110
Number of observations	-	50,015	50,055	49,981

*Significant at 10% level. **Significant at 5% level. ***Significant at 1% level.

Table 6: Negative socioeconomic and positive health selection into teenage fertility by race (dependent variable = giving birth before age 20)

	% of sample	All	Whites	Blacks
Gen1 socioeconomic characteristics				
Mother's age and education				
< 19	37.57	-	-	-
19-22 and less than high school	10.03	1.061	1.130**	0.878
19-22 and high school	22.59	0.774***	0.760***	0.754***
19-22 and some college	3.64	0.506***	0.479***	0.484***
> 22 and less than high school	3.80	1.040	0.978	1.211
> 22 and high school	14.49	0.637***	0.571***	0.713***
> 22 and some college	5.82	0.463***	0.436***	0.441***
> 22 and college	1.90	0.228***	0.210***	0.212***
Mother's age				
23-24	9.71	-	-	-
25-29	25.75	0.828***	0.794***	0.906
30-34	10.35	0.667***	0.593***	0.800
> 34	2.51	0.521***	0.539***	0.434***
Father's age and education				
< 19	6.12	-	-	-
19-22 and less than high school	9.47	1.249***	1.198**	1.278*
19-22 and high school	10.23	0.761***	0.667***	0.886
19-22 and some college	1.12	0.474***	0.373***	0.653**
> 22 and less than high school	10.93	1.232***	1.199**	1.104
> 22 and high school	18.31	0.725***	0.634***	0.899
> 22 and some college	4.59	0.485***	0.418***	0.686**
> 22 and college	1.65	0.245***	0.203***	0.460***
Father's age				
23-24	9.77	-	-	-
25-29	26.15	0.903**	0.885**	1.076
30-34	14.67	0.906	0.899	1.136
> 34	7.74	0.880	0.920	0.966
Mother's marital status				
Married	65.15	-	-	-
Unmarried and father known	15.06	1.393***	1.158***	1.277***
Unmarried and father unknown	19.79	1.587***	1.516***	1.316***
Relative characteristics				
Different race	1.46	1.199*	1.287**	1.293
Father older by > 1 year	45.08	1.126***	1.209***	0.913
Father younger by > 1 year	7.20	1.130*	1.129	1.241*
Father more educated by > 1 year	11.65	1.195***	1.241***	1.107
Father less educated by > 1 year	16.49	0.993	1.000	1.019
Gen2 health				
Birth weight				
< 1500g	2.05	0.262***	0.300***	0.237***
1500g-2499g	7.03	0.874***	0.893	0.859**
2500g-4500g	89.80	-	-	-
> 4500g	1.12	0.827	0.842	0.844

Cohort fixed effects controls	-	Yes	Yes	Yes
Pseudo R^2	-	0.110	0.127	0.052
Number of observations	-	49,981	34,550	15,431

*Significant at 10% level. **Significant at 5% level. ***Significant at 1% level.

Table 7: Impact of relative parental race, age and education on infant health (dependent variable = low infant health)

	% of sample	All	Whites	Blacks
Gen1 maternal characteristics				
Mother's age and education				
< 19	41.72	-	-	-
19-22 and less than high school	10.53	1.267***	1.358***	1.314***
19-22 and high school	20.80	0.956	0.924	0.933
19-22 and some college	3.59	0.999	0.962	0.904
> 22 and less than high school	4.30	1.326***	1.281**	1.453***
> 22 and high school	12.72	1.064	1.003	1.032
> 22 and some college	4.51	1.027	0.955	0.953
> 22 and college	1.53	0.906	0.788*	0.910
Mother's age				
23-24	8.18	-	-	-
25-29	11.31	1.275***	1.202***	1.403***
30-34	3.09	1.579***	1.469***	1.696***
> 34	0.50	2.141***	1.825***	2.731***
Gen1 paternal characteristics				
Father's age and education				
< 19	6.91	-	-	-
19-22 and less than high school	10.04	0.834**	0.865	0.930
19-22 and high school	9.04	0.910	0.831*	0.955
19-22 and some college	0.90	0.637***	0.575***	0.688**
> 22 and less than high school	10.38	0.947	0.947	1.006
> 22 and high school	15.89	0.772***	0.723***	0.806**
> 22 and some college	3.64	0.620***	0.577***	0.725***
> 22 and college	1.24	0.539***	0.558***	0.615***
Father's age				
23-24	8.75	-	-	-
25-29	13.86	0.958	0.972	1.034
30-34	5.85	0.947	1.023	0.909
> 34	3.19	1.046	1.181	0.927
Gen1 relationship characteristics				
Mother's marital status				
Married	65.15	-	-	-
Unmarried and father known	15.06	1.724***	1.354***	1.162**
Unmarried and father unknown	19.79	1.925***	1.432***	1.220**
Relative characteristics				
Different race	1.46	1.047	1.268**	0.875
Father older by > 1 year	77.45	0.992	0.987	0.998
Father younger by > 1 year	4.35	1.073	1.141*	1.028
Father more educated by > 1 year	53.05	1.119**	1.125*	1.143
Father less educated by > 1 year	10.87	1.042	1.062	1.031
Cohort fixed effects controls	-	Yes	Yes	Yes
Pseudo R^2	-	0.022	0.014	0.008
Number of observations	-	75,785	52,360	23,425

*Significant at 10% level. **Significant at 5% level. ***Significant at 1% level.

Table 7 also shows that while white women born to families in which the father is substantially older or more educated are more likely to give birth as teenagers, there is somewhat weaker evidence that they are more likely to have low infant health. One possible explanation for our results is that the impact of having older or more educated fathers only makes a difference after a long period of exposure.

In results not shown here, we find that maternal health behavior (measured using trimester when prenatal care began, and amounts of smoking and drinking during pregnancy) can explain only part of a) the negative association between maternal age and infant health and b) the positive association between parental education and infant health.

We also explore whether negative socioeconomic and positive health selection are mediated by educational outcomes at early and older ages. Table 8 shows that, as expected, higher test scores at 3rd and 8th grades are associated with lower teenage fertility. However, females who are handicapped or have learning disabilities are also less likely to give birth as teenagers, suggesting a potential nonlinear relationship between academic performance and teenage childbearing. The predictive power of both maternal and paternal age and education weaken after controlling for educational outcomes in the 3rd grade, and weaken again after controlling for educational outcomes at older ages. Unlike Manlove (1997), we do not find that educational outcomes completely mediate the effect of maternal education. The results for Gen1 socioeconomic characteristics differ qualitatively from those for Gen2 health: while having very low birth weight is less predictive of teenage fertility outcomes after controlling for educational outcomes in the 3rd grade, additional controls for educational outcomes at the 8th grade make little difference.

To study whether socioeconomic and health selection into teenage childbearing is similar at early adult ages (ages 20 to 22), we repeat our analysis for these ages using data for the first three cohorts. We find evidence of negative socioeconomic selection for these ages, but no evidence of health selection beyond early mortality (results not shown). One potential explanation for why we find no evidence of health selection into early adult childbearing is that early adult fertility is partly boosted by delayed childbearing among less healthy teenagers, leading to comparable fertility levels despite lower levels of fecundability.

Table 8: Socioeconomic and positive health selection into teenage fertility, controlling for educational outcomes (dependent variable = giving birth before age 20)

	% of sample	Not controlling for educational outcomes	Controlling for early educational outcomes	Controlling for early and late educational outcomes
Gen1 socioeconomic characteristics				
Mother's race				
White	0.681	-	-	-
Black	0.309	1.323***	1.177***	1.121***
Mother's age and education				
< 19	37.57	-	-	-
19-22 and less than high school	10.03	1.093*	1.062	1.039
19-22 and high school	22.59	0.764***	0.773***	0.785***
19-22 and some college	3.64	0.488***	0.522***	0.543***
> 22 and less than high school	3.80	1.059	1.041	1.029
> 22 and high school	14.49	0.631***	0.641***	0.654***
> 22 and some college	5.82	0.453***	0.480***	0.504***
> 22 and college	1.90	0.220***	0.254***	0.281***
Mother's age				
23-24	9.71	-	-	-
25-29	25.75	0.829***	0.839***	0.845***
30-34	10.35	0.665***	0.682***	0.698***
> 34	2.51	0.515***	0.533***	0.542***
Father's age and education				
< 19	6.12	-	-	-
19-22 and less than high school	9.47	1.284***	1.233***	1.211***
19-22 and high school	10.23	0.749***	0.749***	0.756***
19-22 and some college	1.12	0.467***	0.486***	0.504***
> 22 and less than high school	10.93	1.244***	1.178**	1.145*
> 22 and high school	18.31	0.717***	0.722***	0.723***
> 22 and some college	4.59	0.485***	0.506***	0.526***
> 22 and college	1.65	0.247***	0.280***	0.302***
Father's age				
23-24	9.77	-	-	-
25-29	26.15	0.907*	0.911*	0.914*
30-34	14.67	0.910	0.923	0.921
> 34	7.74	0.884	0.891	0.886
Mother's marital status				
Married	65.15	-	-	-
Unmarried and father known	15.06	1.236***	1.220***	1.204***
Unmarried and father unknown	19.79	1.326***	1.302***	1.265***
Relative characteristics				
Different race	1.46	1.272**	1.261**	1.265**
Father older by > 1 year	45.08	1.130***	1.130***	1.136***
Father younger by > 1 year	7.20	1.136*	1.121*	1.103
Father more educated by > 1 year	11.65	1.206***	1.192***	1.182***
Father less educated by > 1 year	16.49	1.006	1.011	1.013

Gen2 health				
Birth weight				
< 1500g	2.05	0.262***	0.495***	0.484***
1500g-2499g	7.03	0.874***	0.899**	0.882**
2500g-4500g	89.80	-	-	-
> 4500g	1.12	0.827	0.847	0.842
Gen2 academic performance				
3rd grade				
Verbal Z-score < -1	11.50	-	1.202***	1.091**
Verbal Z-score between -1 and 1	66.74	-	-	-
Verbal Z-score > 1	20.13	-	0.642***	0.848***
Verbal Z-score unknown	1.63	-	1.359	1.219
Math Z-score < -1	13.61	-	1.155***	1.004
Math Z-score between -1 and 1	67.99	-	-	-
Math Z-score > 1	16.85	-	0.739***	1.004
Math Z-score unknown	1.55	-	0.896	0.595**
Gifted	6.61	-	0.683***	0.884
Handicapped	7.24	-	0.834***	0.803***
8th grade				
Verbal Z-score < -1	10.40	-	-	1.186***
Verbal Z-score between -1 and 1	70.55	-	-	-
Verbal Z-score > 1	17.81	-	-	0.741***
Verbal Z-score unknown	1.24	-	-	1.467
Math Z-score < -1	13.26	-	-	1.275***
Math Z-score between -1 and 1	67.59	-	-	-
Math Z-score > 1	17.94	-	-	0.623***
Math Z-score unknown	1.21	-	-	1.695
Gifted	19.54	-	-	0.656***
Handicapped	3.87	-	-	0.872**
Skipped 1 year	5.65	-	-	0.935
On time	78.15	-	-	-
Held back by 1 year	12.33	-	-	1.038
Held back by 2 years	1.22	-	-	1.395***
Cohort fixed effects controls	-	Yes	Yes	Yes
Pseudo R^2	-	0.110	0.117	0.124
Number of observations	-	49,981	49,141	49,109

*Significant at 10% level. **Significant at 5% level. ***Significant at 1% level.

Table 9: Impact of socioeconomic and health selection on infant health, conditional on a teenage birth
(dependent variable = low infant health)

	% of sample	All	Whites	Blacks
Gen1 socioeconomic characteristics				
Mother's age and education				
< 19	37.57	-	-	-
19-22 and less than high school	10.03	0.983	1.043	0.970
19-22 and high school	22.59	1.029	0.903	1.059
19-22 and some college	3.64	0.989	0.886	0.914
> 22 and less than high school	3.80	0.987	0.920	1.082
> 22 and high school	14.49	0.772*	0.521***	0.991
> 22 and some college	5.82	1.127	0.792	1.279
> 22 and college	1.90	0.926	0.555	1.023
Mother's age				
23-24	9.71	-	-	-
25-29	25.75	0.846	0.888	0.761
30-34	10.35	0.738	0.654	0.772
> 34	2.51	0.816	1.176	0.385
Father's age and education				
< 19	6.12	-	-	-
19-22 and less than high school	9.47	0.810	0.886	0.686
19-22 and high school	10.23	0.913	0.925	0.854
19-22 and some college	1.12	0.492*	0.702	0.278*
> 22 and less than high school	10.93	0.961	0.976	1.144
> 22 and high school	18.31	1.035	1.000	1.053
> 22 and some college	4.59	1.032	0.954	1.237
> 22 and college	1.65	0.957	1.128	0.815
Father's age				
23-24	9.77	-	-	-
25-29	26.15	0.905	0.870	1.003
30-34	14.67	1.203	1.432*	1.082
> 34	7.74	1.155	1.382	0.924
Mother's marital status				
Married	65.15	-	-	-
Unmarried and father known	15.06	1.313***	0.946	1.099
Unmarried and father unknown	19.79	1.152	0.812	0.932
Relative characteristics				
Different race	1.46	0.778	0.965	0.521
Father older by > 1 year	45.08	0.961	0.915	1.090
Father younger by > 1 year	7.20	1.384**	1.386	1.592*
Father more educated by > 1 year	11.65	1.063	1.013	1.141
Father less educated by > 1 year	16.49	0.884	0.935	0.980
Gen2 health				
Birth weight				
< 1500g	0.71	1.886**	4.405***	0.872
1500g-2499g	7.77	1.573***	1.642***	1.434***
2500g-4500g	90.95	-	-	-
> 4500g	0.58	0.960	1.176	0.691

Gender, plurality and parity controls	-	Yes	Yes	Yes
Cohort fixed effects controls	-	Yes	Yes	Yes
Pseudo R^2	-	0.032	0.038	0.030
Number of observations	-	10,596	6,061	4,535

*Significant at 10% level. **Significant at 5% level. ***Significant at 1% level.

Finally, we investigate whether socioeconomic and health selection is associated with Gen3 infant health, conditional on Gen2 having a teenage birth. Table 9 shows that, controlling for Gen3 gender, plurality and parity, babies born to teenage mothers with low or very low birth weight are also more likely to have low birth weight, while babies with younger or less educated grandparents are just as likely to have normal birth weight as other babies. If Gen2 maternal health behavior is controlled for (results not shown here), the effect of having a mother with low birth weight continues to persist and becomes slightly stronger, with the odds ratio for babies born to mothers with low and very low birth weight compared to those born to mothers with normal birth weight rising from 1.573 and 1.886 to 1.598 and 1.930 respectively.

Discussion

Our paper is based on a large population-level administrative dataset for six birth cohorts born between 1987 and 1992 in North Carolina. The dataset not only provides birth certificate data on the parental characteristics and health at birth of these cohorts, but also administrative data on their later educational and teenage fertility outcomes. Use of administrative data represents a departure from previous studies which usually rely on survey datasets. Our results are consistent with those from earlier work, which find that maternal age, education and marital status predict teenage fertility outcomes. Unlike earlier work, however, we find much weaker effects of maternal age at first birth on selection into teenage fertility for blacks.

The above results offer two main contributions in addition to what is currently known about selection into teenage fertility. First, the evidence suggests that paternal education at birth matters as much as maternal education for whites; for blacks, the effects of paternal education are weaker but continue to be significant. The predictive power of paternal education may reflect income effects as well as the potential impact of paternal education on teenagers' discount rates. Interestingly, the association between parental characteristics and teenage fertility outcomes may depend on parental relationships. We find that the protective effects of parental age and education are weaker when parents are unmarried, and that white families in which the father is substantially older or more educated than the mother are more likely to have daughters who give birth as teenagers. This may reflect the role of parental bargaining power, which disproportionately benefits children of the same sex, either due to preferences or technological advantages (Thomas 1994).

Second, the evidence suggests that there is positive health selection which excludes the least healthy infants from giving birth as teenagers. Specifically, we show that moderately unhealthy infants are more likely than healthy infants to give birth as teenagers due to their disadvantaged family backgrounds, while very unhealthy infants are less likely than healthy infants to give birth as teenagers *despite* their disadvantaged family backgrounds. Our findings are consistent with studies in the epidemiology literature which find a negative relationship between low infant health and later life fertility, possibly due to lower levels of organ development and functioning and higher childhood mortality (Ekholm et al. 2005; Nohr et al. 2009; and Swamy, Ostbye and Skjaerven 2008). Other studies, on the other hand, finds that low birth weight for gestational age is associated with higher fertility (Ekholm et al. 2005; Nettle, Coall and Dickins 2010). Nettle, Coall and Dickins argue that their results reflect the link between low infant health and lower age at menarche –the literature review above, however, suggests that age of menarche is not a significant predictor of teenage fertility.

In addition to documenting negative socioeconomic and positive health selection into teenage childbearing, our results imply three important differences between these two forms of selection. First, negative socioeconomic selection persists into early adulthood, whereas the effects of infant health on fertility in these ages are limited to the impact on early mortality. These results complement those in Hardy et al. (1998), who find that maternal age is negatively associated with probability of childbearing during both teenage years and early adulthood. Second, we find that socioeconomic selection is mediated by both early and late educational attainment, while health selection is mediated only by early educational attainment. We find that academic ability is generally associated with lower probability of teenage childbearing (consistent with Hardy et al. 1998 and Meade, Kershaw and Ickovics 2008), which may reflect higher opportunity costs of childbearing, lower educational aspirations among those with early childbearing intentions or simply the outcome of low parental support. However, we also note that females with learning disabilities are less rather than more likely to have a teenage birth. This points to a potentially nonlinear relationship between academic ability and teenage fertility, possibly due to the relationship between cognitive ability and social popularity (Manlove 1997). Third, we find that health selection into teenage childbearing is also associated with infant health conditional on having a teenage birth. The intergenerational transmission of low infant health persists even after controlling for maternal educational outcomes and health behavior. These

results are consistent with those in Swamy, Ostbye and Skjaerven (2008), who find that preterm Norwegian women were more likely to go on to have preterm births.

Conclusion

Teenage childbearing has been argued to have negative consequences for both mother and child, and has hence long been of policy interest. Understanding the role of socioeconomic selection into teenage childbearing can help to provide insight into the extent to which the observed correlations between teenage childbearing and later life outcomes are potentially due to non-causal rather than causal mechanisms. In this paper, we provide new evidence that paternal characteristics and infant health predict teenage childbearing outcomes. Since poor infant health has negative consequences for future labor market and other outcomes, selection into teenage childbearing may be less negative than implied by the literature on socioeconomic selection.

References

Black, Sandra E., Paul J. Devereux and Kjell G. Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." *Quarterly Journal of Economics* 122(1): 409-439.

Ekholm, K. et al. 2005. "The Probability of Giving Birth among Women Who Were Born Preterm or with Impaired Fetal Growth: A Swedish Population-based Registry Study." *American Journal of Epidemiology* 161(8): 725:733.

Ellis, Bruce J. et al. 2003. "Does Father Absence Place Daughters at Special Risk for Early Sexual Activity and Teenage Pregnancy?" *Child Development* 74(3): 801-821.

Geronimus, Arline T. 1996. "Black/White Differences in the Relationship of Maternal Age to Birthweight: A Population-Based Test of the Weathering Hypothesis." *Social Science and Medicine* 42(4): 589-597.

Geronimus, Arline T. and Sanders Korenman. 1993. "Maternal Youth or Family Background? On the Health Disadvantages of Infants with Teenage Mothers." *American Journal of Epidemiology* 137(2): 213-225.

Guilamo-Ramos, Vincent et al. 2012. "Paternal Influences on Adolescent Sexual Behaviors: A Structured Literature Review." *Pediatrics* 130(5): 1313-1325.

Hardy, Janet B. et al. 1998. "Like Mother, Like Child: Intergenerational Patterns of Age at First Birth and Associations with Childhood and Adolescent Characteristics and Adult Outcomes in the Second Generation." *Developmental Psychology* 34(6): 1220-1232.

Harris, Kathleen Mullan, Frank F. Furstenberg Jr. and Jeremy K. Marmer. 1998. "Paternal Involvement with Adolescents in Intact Families: The Influence of Fathers over the Life Course." *Demography* 35(2): 201-216.

Kahn, Joan R. and Kay E. Anderson. 1992. "Intergenerational Patterns of Teenage Fertility." *Demography* 29(1): 39-57.

Manlove, Jennifer 1998. "The Influence of High School Dropout and School Disengagement on the Risk of School-Age Pregnancy." *Journal of Research on Adolescence* 8(2): 187-220.

_____. 1997. "Early Motherhood in an Intergenerational Perspective: The Experiences of a British Cohort." *Journal of Family and Marriage* 59(2): 263-279.

Meade, Christina S., Trace S. Kershaw and Jeanette S. Ickovics. 2008. "The Intergenerational Cycle of Teenage Motherhood: An Ecological Approach." *Health Psychology* 27(4): 419-429.
Nettle, Daniel, David A. Coall and Thomas E. Dickins. 2010. "Birthweight and Paternal Involvement Predict Early Reproduction in British Women: Evidence from the National Child Development Study." *American Journal of Human Biology* 22: 172-179.

Nohr, E. A. et al. 2009. “Waiting Time to Pregnancy According to Maternal Birthweight and Prepregnancy BMI.” *Human Reproduction* 24(1): 226-232.

North Carolina Education Research Data Center. http://www.childandfamilypolicy.duke.edu/project_detail.php?id=35.

OECD. 2011. “Infant Health: Low Birth Weight.” *Health at a Glance 2011: OECD Indicators*, OECD Publishing.

Swamy, Geeta, Truls Ostbye and Rolv Skjaerven. 2008. “Association of Preterm Birth with Long-Term Survival, Reproduction and Next-Generation Preterm Birth.” *Journal of American Medical Association* 299(12): 1429-1436.

Thomas, Duncan. 1994. “Like Father, Like Son; Like Mother, Like Daughter: Parental Resources and Child Height.” *Journal of Human Resources* 29(4): 950-988.