The black-white gap in non marital fertility education and mates in segmented marriage markets

Stone, Joe A.

2 January 2012

The Black-White Gap in Non Marital Fertility

Education and Mates in Segmented Marriage Markets

Joe A. Stone
Department of Economics
University of Oregon

January 2, 2012

Abstract

This study is the first to find that mate availability explains much of the race gap in non marital fertility in the United States. Both a general and an education-based metric have strong effects. The novel statistical power arises from difference-in-differences for blacks and whites, multiple cohorts, periods, and coefficient restrictions consistent with both the data and models in which differences in mate availability can induce blacks and whites to respond in opposite directions to *changes* in mate availability. Results are robust to several alternative specifications and tests and appear relevant where marriages are segmented along racial, religious, or other lines.

JEL Categories A10 J12 J13.

Keywords race marriage, fertility, education

Overview

Non marital birth shares, the shares of non marital births in total births, are much higher for blacks than for whites in the U. S., but have increased for both. Many explanations—economic, political, legal, biological, and cultural have been proposed, generating voluminous evidence for the effects of various public policies on female headship and non marital fertility.[1] There is much less direct evidence to explain the pronounced race differential in non marital fertility. Wilson (1987), Willis (1999), Brien (1997), and others emphasize mate availability, relative resources, education, and other factors, but our evidence is the first to confirm that metrics of mate availability explain much of the race difference in non marital birth shares (NBS). Brien (1997) suggests that prior weak results, as in South and Lloyd (1992), arise from measurement error in locale-specific data, as compared to use of state-aggregate data, but state-aggregate estimates can also have weak power, perhaps because variations tend to be commonly shared, leaving little between-state variation and low power in estimation. The strongest evidence to date for the role of mate availability in non marital fertility does not apply directly to the black-white gap and is drawn from idiosyncratic sources of variation, such as prison incarceration rates (Charles and Luoh, 2006), WWI military deaths (France Abramitzky et. al., 2010), and sex ratios among second- and third-generation immigrants in the United States (Angrist, 2001).

Here, we focus directly on the race gap in non marital fertility. The novel strength of our estimates is largely due to a difference-in-differences specification for blacks and whites, multiple birth cohorts and periods, and coefficient restrictions consistent with both the data and theory. We assess robustness in several ways, e.g. by using alternative

[1] For example, Gray et al (2006) and Moffitt (2000)

age ranges for cohort variables, inclusion of controls for both fixed and time-varying age effects, and use of alternative tests, including tests of Granger causality, which fail to reject the key regression results. Moreover, our null hypotheses require a specific set of restrictions in the pattern of effects for changes in the metrics of mate availability among blacks and whites—restrictions not rejected by the data and grounded in models of fertility and marriage in which blacks and whites respond in opposite directions to changes in mate availability due to their initial differences in mate availability. Hence, odds that potential sources of bias would yield spurious results simultaneously consistent with all null hypotheses in multiple estimation methods appear remote.

Theoretical Context

Wilson (1987), Willis (1999), and others argue that blacks and whites are likely to respond differently to changes in mate availability. Where eligible women exceed men in a cohort, as for blacks, then with positive assortative mating, children tend to be born outside marriage to low-income/education fathers and mothers, at least when the incomes of the mothers exceed a value critical to their decision to bear a child, so that increases in the ratio of men's to women's incomes/education or in the ratio of men to women decrease non marital birth rates when eligible women outnumber men, but where the ratio of eligible men to women is above parity, as it is for whites, children tend to be born within marriage to high-income/education men and women in positively assortative matches.

Model intuition

In terms of marginal effects, the Wilson-Willis approach yields the intuitive result that a relative increase in the short (long) side of mate availability decreases (increases) NBS. The overall ratio of men to women has been below parity for blacks in recent decades but above parity for whites, so that males are the short (long) side of mate availability for blacks (whites), a divergence due largely to higher death, incarceration, military enlistment, and institutionalization rates among black men. Temporal variations in these rates tend to be widely shared, with standard deviations well below the corresponding means. Hence, between-locale variation typically yields weak estimates.

Data and Variables

All data are publicly available and refer to the civilian, noninstitutionalized population, ages 20-44.[2] We limit the analysis to birth cohorts for which full data are available, from 1972 through 2002. The cohorts, their birth years, and the years at which they are ages 20-24 and 40-44 are presented in Table 1. As with many cohort analyses, we employ five-year age ranges, a range wide enough to provide reliable measures, but narrow enough to limit time- and age-varying heterogeneity within the cohort. Robustness checks using alternative age groupings suggest that the cohort data are aligned appropriately for gender differences in age at first marriage and temporal differences, as emphasized by Neelakantan and Tertilt (2008).

Table 2 reports summary statistics for all variables used in estimation.[3] The dependent variable is the non marital birth share (NBS), the share of non marital births among all births. As expected, NBSB, the NBS for blacks, (at 43.4) is greater than

[2] Non marital births are from *National Vital Statistics Reports* (2000, 2002); Total births from *Vital Statistics of the United States*

[3] Data for sex ratio are from U.S. Bureau of the Census

NBSW, the NBS for whites (12.3), though both have risen over time, as reflected in the minimum and maximum values.

Our primary, most general measure of mate availability is SEXRATIO, the ratio of eligible men to women when each cohort was 20-24. This variable is calculated by dividing the number of men in this age group for a given birth cohort by the number of women and multiplying by 100 to express the result in percentage points. The average sex ratio for whites, SEXRATIOW, is 102.5, ranging narrowly from 100.9 for the oldest cohort to 107.0 for the most recent cohort. By contrast, the average sex ratio for blacks SEXRATIOB is only 93.6, ranging from 88.6 for the oldest cohort to 97.5 for the most recent cohort and uniformly below unity.

SEXRATIO values for whites and blacks do not overlap, hence, estimation by race or by whether the sex ratio is above or below parity are equivalent. As anticipated, standard deviations for SEXRATIO relative to the corresponding means are larger than the relative variation typical for the locale-specific data shown to be subject to substantial measurement error.

Our qualitative metric of mate availability at the upper end of potential marriage pairs is POSTS, the ratio of male to female school enrollments at ages 20-24, which primarily reflect post-secondary school enrollments. Use of relative wages in prior studies as an alternative qualitative or resource metric yields weak results, perhaps either because differences between average earnings for men and women, especially among blacks are much smaller than differences for post-secondary schooling or because measurement error dominates in locale-specific data.

Use of the broader spatial data for POSTS is consistent with the greater spatial mobility for men and women at higher levels of education and skill. POSTS is favorable on average for women of both races early in the sample period, but has become increasingly unfavorable, especially for black women. Overall, the average for POSTSW is 1.64, ranging from over 4.0 for the oldest cohort to only .89 for the most recent cohort. The average ratio for POSTSB is 1.25, ranging from 1.78 for the oldest cohort down to only 0.71 for the most recent cohort. These shifts exhibit the now-familiar dominance of school enrollments for women over men during these ages–particularly so for blacks, for whom incarceration rates of young males have increased relative female-male enrollments for black women (Mechoulan, 2006).

We find similar results using alternative age ranges for school enrollment, e.g., either 18-21 or 18-24 years old. Consistent with the Wilson-Willis model, we expect the effect of an increase in SEXRATIOB on NBSB to be negative because males are 'short' among blacks, and the effect of an increase in SEXRATIOW on NBSW to be positive because women are short among whites, if only slightly.

By contrast, we expect the effect of POSTS on NBS to be positive for both whites and blacks because until only very recently, males outnumber females in both cases. Based on familiar racial differences in the timing of nonmarital births, the black-white difference in NBS should decline with age, so we control for cohort age and subsequently, also for age-year interactions to account for temporal changes in racial differences in the timing of fertility.

Estimation Issues

Specification

Our focus is directly on the black-white gap in NBS, and our estimation strategy is to use differences for blacks and whites across birth cohorts, time and age, to identify the effects of SEXRATIO and POSTS. Of course, age, year, and birth cohort are linearly dependent, so only two of these three effects are identified in linear form. Because values for SEXRATIO for blacks and whites do not overlap, estimation separately by race and separately by whether SEXRATIO is above or below unity are equivalent, so we begin by expressing the black-white difference in NBS as a linear function of fixed year intercepts and black-white differences in SEXRATIO and POSTS, with all variables in log form, which mitigates differences in scale for blacks and whites and yields more robust results for cohort i, age j, and year t:

$[NBSW_{ijt} - NBSB_{ijt}] =$

$$c_t + b^{w}_1 \, SEXRATIOW_{ijt} - b^{B}_1 \, SEXRATIOB_{ijt} + b^{W}_2 \, POSTSW_{ijt} - b^{B}_2 \, POSTSB_{ijt} + b_3 \, AGE + e_{ijt} \quad (1)$$

Where cohort i, year t, age j, and other effects common to blacks and whites are eliminated by differencing by race, so that the age and fixed year effects capture residual race differences.

For the two metrics for mate availability, SEXRATIO and POSTS, the Wilson Willis models predict opposite, possibly equal, effects for blacks and whites. That is, for equal but opposite effects to hold :

$$b^{w}_1 + b^{B}_1 = b^{B}_2 + b^{W}_2 = 0 \quad (2)$$

in which case, eq (1) reduces to eq (3) below:

$[NBSW_{ijt} - NBSB_{ijt}] =$

$$c_t + b_1 \, (SEXRATIOW_{ijt} - SEXRATIOB_{ijt}) + b_2 (POSTSW_{ijt} - POSTSB_{ijt}) + b_3 \, AGE + e_{ijt} \quad (3)$$

Where we expect $b_3 < 0$, and $b_1, b_2 > 0$

Endogeneity

There is little reason to suspect endogeneity bias for SEXRATIO because the cohort sex ratio at age 20-24 is predetermined and independent of NBS. However, POSTS is predetermined only for *subsequent* ages within a cohort and strictly exogenous only across cohorts. Even so, we can assess the importance of endogeneity bias for POSTS in several ways, including tests of Granger causality and examining the sensitivity of the estimated effect of POSTS to using alternative age ranges for school enrollments, Moreover, we expect the specific signs and restrictions in eq. (2), so that the likelihood of any bias *simultaneously* consistent with all the predictions for blacks and whites appears remote. In addition, we are also able to replicate key results with tests of Granger causality.

Results

Column 1 of Table 3 presents the baseline regression for eq. (3) The regression includes fixed year effects, which are jointly significant, but omitted for brevity. The difference-in-differences specification in eq. (3) pushes the data hard, leaving only 23 degrees of freedom, roughly the minimum needed to rely on the small-sample properties of ordinary least squares. Hence, the power of the estimates in column 1 rests not on large numbers, but on the extent to which cohort variations in the data are sufficiently large and common enough across locales to identify the effects of mate availability, given the restrictions of eq. (2). Otherwise, with such limited degrees of freedom, standard errors will likely be large and yield insignificant coefficients even where significance is expected. In fact, however, power does not appear to be a problem for the estimates. The equation fit is strong, with an R-squared of 0.766, and all coefficients are significant in the hypothesized direction. In addition, the coefficient for AGE is significantly negative, indicating that nonmarital fertility falls relative to total fertility more rapidly with age for blacks as compared to whites.[4] The coefficients for

[4] Gray and Stone (2010) examine factors determining black-white differences in the timing of both marital and nonmarital fertility.

SEXRATIO and POSTS are significantly positive, as hypothesized and indicate in each case that an increase in the short side of mate availability for blacks or whites decreases the corresponding NBS.

Robustness and Specification Tests

Age-year interactions

Addition of an age-year interaction in Column 2 demonstrates that estimates for the metrics of mate availability are not sensitive to the significant interaction between cohort age and year, and indicates that the age-NBS profile for the race gap in NBS became even more negative later in the later years of the period.

Residual diagnostics

The Jarque-Bera test of normality does not reject the null hypothesis that the regression residuals follow a normal distribution (p=.0383), and The Q statistic for first-order autocorrelation fails to reject the null of zero autocorrelation, (p =0.327), and the Q-statistic for first-order autocorrelation fails to reject the null of zero autocorrelation, (p =0.327). The absence of significant autocorrelation is consistent with the absence of significant specification error and lends credence to reliance on the recursive inertia of predetermined data at earlier ages of a cohort for identification. Results are not sensitive to use of different, neighboring age ranges for either SEXRATIO or POSTS, and tests for Granger causality based on the regression residuals reject the null hypothesis of no Granger causality for each of the hypothesized links in eq. (3), and yet also fail to reject the null of no reverse Granger causality from the race gap in NBS to either SEXRATIO or POSTS, so the tests raise no significant concerns regarding reverse causality or endogeneity bias, at least in the Granger sense.

Coefficient restrictions

Despite the strong significance and power of the individual estimates, a Wald test of the joint linear restrictions set out in eq. (2) and imposed in column 1 (for equal, but opposite effects for blacks and whites) fails to reject the restrictions ($p = 0.712$).

Unrestricted estimates of eq. (3) are presented in Table 4. The separate variables by race for SEXRATIO and POSTS enter significantly as a set, but POSTSW is the only metric for mate availability to enter significantly.

One can see from the larger standard errors and weaker power for the other individual estimates that the restrictions play an important role in the power of the individual estimates in column 1 of Table 3, despite the fact that the joint test of the restrictions fails to reject them. Note that the significantly positive effect of POSTSW is consistent with the Wilson-Willis approach, where a relative increase (decrease) in the long side of mate availability , in this case white males enrolled in school at ages 20-24, increases (decreases) NBS. The weak individual estimates in Table 4 are in line with other similarly weak prior estimates, suggesting that prior evidence may mask the significant effects found here, perhaps because either the opposing effects for blacks and whites cancel to zero in pooled specifications for blacks and whites or because the identifying power of differencing by race or cross-race restrictions is not exploited.

Explaining the Racial Gap

Regression variance

The model expressed by eq. (3) explains about 77% of the sample variance of the (log) race gap in nonmarital birth shares for blacks and whites. Because this explanatory power is not due solely to our two metrics of mate availability, we isolate their effects using the estimates in Table 3 to simulate changes in the race gap.

Model simulation

To use the estimated model to simulate the proportion of the maximum change in the (log) race gap explained by SEXRATIO and POSTS, we 1) calculate the maximum change for the race gap in the sample (i.e. from min to max), and 2) use the model coefficients along with the maximum change in each explanatory variable to simulate the change in the race gap attributable to the two metrics of mate availability. The maximum change in the log race gap is 2.81, of which SEXRATIO explains 0.85, or 30.2%, and POSTS explains .60, or 21.4%. Hence the two metrics jointly account for 51.6%, more than half of the maximum change in the race gap in NBS. AGE differences explain- .544, -19.4%, so the proportion of the change in the race gap (net of age effects) explained by metrics of mate availability is higher, roughly 64%.

Concluding Remarks

Our two metrics of mate availability reflect and are heavily influenced by education and other public policies, especially the decades-long 'war on drugs' in the U.S, possibly at the loss of a generation of black men to families they might have fathered with their children's mother. Results here complement other recent evidence on effects of incarceration rates in the U. S. and WWI military deaths in France on nonmarital fertility, but indicate broader effects for both a general and an education-based metric for mate availability directly on the black-white gap in non marital fertility. General mate availability is less a factor among whites, though the relative pools of college educated men and women are now rapidly changing among both whites and blacks. Our results may be relevant where marriages are segmented along racial, ethnic, or religious lines, as for example, among castes in India, ethnic Malays and Chinese in Malaysia; or whites, South Asians, and blacks in the U. K. (Barthoud 2011). Finally, the results here confirm theoretical conjectures that the direction of the effect of a change in mate availability depends on both the direction of the change and whether the change comes from the long or short side of the marriage market.

References

Abramitzky, R., Delavando, A. and Luis Vasconcelos. 2010. “Marrying Up: The Role of Sex Ratio in Assortative Matching.” Working paper Stanford U.

Berthoud. 2011. “*Family Formation in Multicultural Britain: Three patterns of diversity*.” working paper Institute for Social and Economic Research, U. Essex

Brien, Michael J. 1997. "Racial Differences in Marriage and the Role of Marriage Markets." *Journal of Human Resources* 32(4):741-78

Census. 2005. *Fertility of American Women: 2004.* Washington: Commerce Dept.

Charles, Kerwin K. and Ming Ching Luoh. 2006. “Male Incarceration, the Marriage Market, and Female outcomes,” NBER working paper

Gray, Jo A., Stockard, Jean and Joe A. Stone. 2006. “The Rising Share of Non marital Births: Fertility Choice or Marriage Behavior?” *Demography* 43:2, 241-53

Mechoulan Stéphane. 2006 “External Effects of Black Male Incarceration on Black Females.” U. Toronto.

Moffitt, Robert. 2000. “Welfare and Female Headship in U.S. Time Series.” *American Economic Review* 90:373-77

Neelakantan, U. and M. Tertilt. 2008. “A note on marriage market clearing.” *Economics Letters* 101 (2): 103-105.

South, Scott J. and Kim M. Lloyd. 1992. “Marriage Markets and Nonmarital Fertility in the United States.” *Demography* 29: 247-64

Willis, Robert J. 1999. “A Theory of Out-of-Wedlock Childbearing.” *Journal of Political Economy* 107(6): S33-64.

Wilson, W. 1987. *The Truly disadvantaged:* Chicago: U Chicago Press

Table 1. Cohorts and Birth Years (1972-2007), (Ages 20-24)

Cohort #	Age 20-24 in	Age 40-44 in
1	1952	1972
2	1957	1977
3	1962	1982
4	1967	1987
5	1972	1992
6	1977	1997
7	1982	2002
8	1987	
9	1992	
10	1997	
11	2002	

Table 2. Summary statistics (1972-2007), (Ages 20-44)

	NBSB	NBSW	SEXRATIOB	SEXRATIOW	POSTSB	POSTSW
Mean	43.42	12.28	93.64	102.49	124	165
Median	41.60	10.54	93.42	102.11	125	130
Max	81.30	44.60	97.47	106.998	178	414
Min	23.08	2.71	88.57	100.94	71	89
Std. Dev.	15.65	9.49	2.07	1.41	34	75
Obs.	35					

Notes:
See text for sources of data.
All ratios in percentage points.

Bsuffix	Blacks
Wsuffix	Whites
NBS	Non marital births as a share of total births
SEXRATIO	Ratio of males to females
POSTS	Ratio of males to females enrolled in school at ages 20-24
AGE	Average cohort age (same for blacks and whites)

Table 3. Race Gap in Non marital Birth Shares ln (white/black), (1972-2007), (ages 20-44)

(Robust std. errors below coefficients)

	(1)	(2)
Constant	1.4682	46.802*
	1.5473	5.445
lnSEXRATIO	14.2381	7.686*
	5.5743	1.775
lnPOSTS	0.6412*	0.5473*
	0.2068	0.1509
lnAGE	-0.544*	-7.129*
	0.221	0.774
Year effects	Yes*	Yes*
AGE x YEAR/k	No	0.002*
R-squared	0.766*	0.937*

See notes, Table 2.

* significant .05

Table 4. Unrestricted Estimates, Non marital Birth Share ln (white/black), (1972-2007), (Ages 20-24)

(Robust std. errors below coefficients)

Constant	-0.469
	34.610
lnSEXRATIOW	3.268
	5.857
lnSEXRATIOB	-3.149
	7.261
lnPOSTSW	1.564*
	0.408
lnPOSTSB	-0.381
	0.527
lnAGE	-1.203*
	0.374
Year Effects	Yes
AGE x YEAR	No
R-Squared	0.809*

See notes to Table 2
* significant .05