The Effect of Female Education on Fertility and Infant Health: Evidence From School Entry Policies Using Exact Date of Birth: Web Appendix

Justin McCrary and Heather Royer

1 Estimates of the Probability of Motherhood by Age and Cohort

Here we expand on the description of Web Appendix Table 1. The table consists of 4 separate panels: 1A, 1B, 1C, and 1D. Panels 1A and 1B pertain to California; panels 1C and 1D pertain to Texas. For each panel, a row represents a cohort, and each column represents a specific age at which a mother could be observed in our administrative data.[1] The entries of the table are discontinuity estimates, with standard errors in parentheses.

The marginal rows at the bottom of the table present joint tests of no effect across cohorts, pooled estimates of the cohort-specific discontinuities, tests of the cross-cohort restrictions, and sample means.

The last column of panels 1B and 1D gives cohort-specific discontinuities corresponding to Figure 1 for California and Texas, respectively. None (one) of the cohort-specific discontinuities for California (Texas) are significantly different from zero. The lower right hand corner of panels 2B (2D) displays the pooled estimated discontinuity in the fraction observed at any age of -0.0019 (-0.0072). Compared to the overall fraction observed of 0.23 (0.29), these are very small discontinuities. Neither is statistically significant at conventional levels of significance.

The second-to-last columns of panels 2B and 2D present similarly small and insignificant estimates for the fraction observed at 23 or younger (i.e., selection into our estimation sample). The pooled discontinuity estimate for the probability of giving birth at 23 or younger is -0.0018 (-0.0048) for California (Texas), which is small relative to the mean of 0.18 (0.24).

The other columns of the table present age-specific probabilities of observation. These are useful to consider, because it is possible that there could be no aggregate change in the fraction observed, but individuals could be observed at different ages. The age-specific discontinuity estimates do not support this hypothesis, however. The estimated discontinuities are generally small, of fluctuating sign, and statistically indistinguishable from zero. For example, 2 (3) of the 162 point estimates for California are positive (negative) and significant, and 2 (4) of the 143 point estimates for Texas are positive (negative) and significant.

The third-to-last columns of panels 2B and 2D present tests of the hypothesis that the discontinuity estimates are jointly zero across all ages, for a given cohort ("test of no effect").[2] For each state and for each cohort, we fail to reject the hypothesis.

For each panel, the third-to-last block of rows presents tests of the hypothesis that the discontinuity estimates are jointly zero across all cohorts, for a given age ("test of no effect"). For California, we reject one of these hypotheses (age 28) and for Texas, we reject two of these hypotheses (ages 22 and 25). These rejections may be spurious, since we are testing many hypotheses. To test this idea, we also present a test of the hypothesis that the discontinuity estimates are jointly zero across all cohorts and all ages (bold entry in lower right of panels 2B and 2D). We fail to reject this stringent null hypothesis for both California and Texas.

[1] Beneath the cohort label, we present the number of women in these cohorts who are born within 50 days of the school entry date.

[2] Estimates and tests pooled across ages ignore the mechanical negative correlation between the indicator for being age a and the indicator for being age $a' \neq a$. Because we have many age categories, this negative correlation is negligibly small (cf., McCrary 2008). The simulation evidence we have examined shows that ignoring the negative correlation leads to extremely minor size distortions.

The second-to-last block of rows presents pooled estimates of the discontinuity in the probability of being observed, for a given age. These pooled estimates impose potentially false cross-cohort restrictions. We present the p-value for the test of the hypothesis that the discontinuity is equal across cohorts (brackets). For all but one age in California and one age in Texas, we fail to reject the cross-cohort restrictions.

2 Power

In this section, we discuss ex-ante estimates of the sample sizes required to reject point hypotheses our own study is not able to rule out. For example, our data do not rule out the possibility that a one-year increment to schooling reduces infant mortality by -0.001.

We next explain our approach to computing the requisite sample sizes. Because they are nonparametric in nature, our instrumental variables estimates have variances of the form

$$V\left[\widehat{\theta}\right] = \frac{c}{nh} \tag{A.1}$$

where n is our existing sample size, h is the bandwidth and c is a complicated function of the design matrix and prediction errors.

The estimated standard errors from our data provide a preliminary estimate of c that can be used to forecast what kinds of magnitudes of standard errors would be associated with point estimates from larger sample sizes. For example, doubling the sample size is expected to yield a standard error 70 percent as large as our estimated standard errors, holding the bandwidth fixed at $h = 50$.

An alternative approach to forecasting standard errors shrinks the bandwidth with the sample size. The theoretical econometrics literature suggests that the bandwidth should be of the order $n^{-1/5}$ (Porter 2003, Theorem 3(b)). Write $h = kn^{-1/5}$ for some k and note that this implies

$$V\left[\widehat{\theta}\right] = \frac{\tilde{c}}{n^{4/5}} \tag{A.2}$$

where $\tilde{c} = c/k$. The estimated standard errors from our data and our chosen bandwidth of $h = 50$ together furnish an estimate for $\tilde{c}$.

We thus have 2 approaches to power calculations for the regression discontinuity context. The first approach holds the bandwidth fixed (e.g., $h = 50$ regardless of sample size). The second approach shrinks the bandwidth at the theoretically prescribed rate. We turn now to calculating the sample size needed to reject a specific point hypothesis $\theta = \theta_0$ under both approaches, with a focus on tests of 5 percent size.

Under the first approach, for a two-sided test, we write

$$\frac{|\widehat{\theta}_N - \theta_0|}{\mathrm{se}_N} > 1.96 \Longleftrightarrow \frac{N-n}{n} > \frac{1.96^2}{(\widehat{\theta}_N - \theta_0)^2}\widehat{\mathrm{se}}^2 - 1 \tag{A.3}$$

where θ_0 is the point hypothesis to be tested, se_N is the standard error forecast for the new, larger sample size N, $\widehat{\mathrm{se}}$ is the estimated standard error from our data, and $\widehat{\theta}_N$ is the point estimate we expect to obtain in the larger sample size.[3] (For example, we might choose to set $\widehat{\theta}_N$ equal to the estimate based on our data (i.e., $\widehat{\theta}$), or we might choose to set $\widehat{\theta}_N$ to zero.)

The right-hand side of the second inequality gives the predicted smallest percent increase in the sample size that will allow rejection of the point null hypothesis $H_0 : \theta = \theta_0$ in favor of the alternative $H_a : \theta \neq \theta_0$.

The second approach to power calculations for the regression discontinuity context shrinks the bandwidth with the sample size. Under this approach, for a two-sided test, we write

$$\frac{|\widehat{\theta}_N - \theta_0|}{\mathrm{se}_N} > 1.96 \Longleftrightarrow \frac{N-n}{n} > \left(\frac{1.96^2}{(\widehat{\theta}_N - \theta_0)^2}\widehat{\mathrm{se}}^2\right)^{5/4} - 1 \tag{A.4}$$

[3]Note c in Equation A.1 will be the same in the old and new sample. Also, we have subtracted 1 from both sides of the inequality to make this expression in percentage terms.

Intuitively, if the fixed bandwidth approach suggests that twice as much data (i.e., 100 percent more) is required to reject a particular hypothesis of interest, then the shrinking bandwidth approach suggests that 138 percent as much data is required ($2^{5/4} - 1 \approx 1.38$).[4]

Appendix Table 4 presents these calculations for selected outcomes, point hypotheses of interest, and hypothetical point estimates, $\widehat{\theta}_N$, that would obtain in the larger sample. For selected outcomes of interest—maternal smoking, prenatal care in the first trimester, low birth weight, prematurity, and infant death—we present typical point estimates from the literature (column 1) alongside IV estimates using our data (column 2). The IV estimates are pooled estimates for California and Texas.

The remaining columns of the table present our power calculations. Column 3 reports the point estimate we might expect to obtain in the larger sample. Column 4 gives potential point hypotheses of interest, and column 5 gives the percent increase in sample size needed to reject that point hypothesis, using a two-sided test and assuming that the bandwidth is held fixed at $h = 50$ (cf., A.3). Column 6 mimics column 5, but reports the percent increase in the sample size required if the bandwidth were to be smaller than that we use here (cf., A.4). Point hypotheses from Currie and Moretti (2003) and Chou, Liu, Grossman and Joyce (2007) are presented for comparison purposes.

Rows 1 through 4 pertain to maternal smoking. For this outcome, our IV estimate is the same sign as that in the literature and statistically significant, but smaller in magnitude. We are interested in knowing what kinds of point hypotheses could be ruled out in larger samples, assuming that the point estimate in the larger sample was the same as that we obtain (-0.016). Columns 5 and 6 report that the point hypothesis -0.06 is rejected by our data, suggesting a smaller impact of schooling on smoking than in the literature. On the other hand, our data also rule out 0 as a plausible hypothesis. Our power calculations suggest that only 20 percent more data would be required to rule out a point hypothesis of -0.03. However, hypotheses such as -0.01 which are close to our point estimate, would be difficult to rule out even with very large samples.

Rows 5 through 8 pertain to prenatal care in the first trimester. Our IV estimate is again of the same sign as that in the literature. However, for this outcome, the estimate is not quite statistically significantly different from zero. We are thus interested in knowing what kind of a sample size would be required to rule out zero. Columns 5 and 6 report that even a 30-40 percent increase in sample size would be sufficient to rule out zero, assuming that the point estimate in the larger sample was the same as that we obtain (0.031). Ruling out the Currie and Moretti (2003) estimate of 0.02 seems infeasible, as does ruling out point hypotheses in the neighborhood of 0.04. On the other hand, ruling out a large point hypothesis such as 0.06 would be feasible with 50-70 percent more data.

Rows 9 through 16 pertain to low birth weight. For this outcome, our data provide substantial evidence against the Currie and Moretti (2003) estimate. In particular, our analysis rejects the point hypothesis of -0.01 and the much smaller point hypothesis of -0.005. Indeed, if a 70 percent larger sample were collected and the point estimate was equal to what it is in our data (0.014), we could rule out the extremely small point hypothesis of -0.001.

Rows 13 through 16 consider a similar set of thought experiments regarding power, but change the assumptions. In particular, in these 4 rows, we assume that in the larger sample size, we would obtain a point estimate of zero. With such a point estimate, it would be possible to reject the Currie and Moretti (2003) estimate only with much more data, and rejecting smaller point hypotheses such as -0.005 and -0.001 is likely not feasible.

Rows 17 through 24 pertain to prematurity. For this outcome, the typical point hypothesis from the literature is on the edge of the confidence region based on our data. For example, the two-sided confidence region is $[-0.014, 0.028]$ and the one-sided confidence region is $[-0.011, 0.009]$. In a 50-70 larger sample with the same point estimate, we would reject the Currie and Moretti (2003) estimate of -0.01. To rule out a hypothesis of -0.005 would require much more data than we have.

Rows 21 through 24 consider a similar set of thought experiments, but assume that in the larger sample the point estimate would be zero rather than 0.009. Ruling out point hypotheses of -0.02 and -0.01 would be feasible under such a scenario, but ruling out smaller point hypotheses would be unlikely.

Rows 25 through 32 pertain to infant death. Currie and Moretti (2003) do not estimate the effect of

[4] Under either approach, these results can be adapted to suit power calculations for one-sided tests. One-sided tests have alternative hypotheses of the form $H_a : \theta < \theta_0$ or $H_a : \theta > \theta_0$. The first type of alternative hypothesis is interesting when $\widehat{\theta}_N - \theta_0$ is positive, and the second type is interesting when $\widehat{\theta}_N - \theta_0$ is negative. One can show that the predicted smallest required percent increase in the sample size continues to have the form given by these inequalities, but with 1.64 replacing 1.96.

education on this outcome. While the country context may be quite different, Chou et al. (2007) report an estimate of -0.005. The table shows that, using our research design, very large sample sizes are necessary to make precise statements about the effect of schooling on infant death.

References

Chou, Shin-Yi, Jin-Tan Liu, Michael Grossman, and Theodore Joyce, "Parental Education and Child Health: Evidence from a Natural Experiment in Taiwan," *NBER Working Paper #13466*, October 2007.

Currie, Janet and Enrico Moretti, "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence From College Openings," *Quarterly Journal of Economics*, November 2003, *118* (4), 1495–1532.

McCrary, Justin, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, February 2008, *142* (2).

Porter, Jack, "Estimation in the Regression Discontinuity Model," 2003. Unpublished manuscript, Harvard University.

Web Appendix Table 1A. Discontinuity in Fraction Giving Birth at Specific Ages, by Cohort: California, Ages 13 through 23

	13	14	15	16	17	18	19	20	21	22	23
1969 cohort								-0.003	-0.007	0.002	0.001
46,190								(0.009)	(0.004)	(0.008)	(0.003)
1970 cohort							0.004	-0.005	-0.005	-0.005	-0.003
46,190							(0.006)	(0.005)	(0.007)	(0.003)	(0.005)
1971 cohort						0.000	-0.002	-0.008	0.001	0.001	0.000
40,464						(0.004)	(0.009)	(0.009)	(0.004)	(0.004)	(0.003)
1972 cohort					0.008	-0.005	-0.005	0.007	0.001	0.001	0.002
39,762					(0.004)	(0.011)	(0.004)	(0.009)	(0.004)	(0.007)	(0.007)
1973 cohort				0.006	0.000	0.010	0.000	-0.002	0.002	-0.003	0.001
38,594				(0.003)	(0.008)	(0.006)	(0.005)	(0.004)	(0.004)	(0.006)	(0.003)
1974 cohort			-0.001	0.001	0.000	-0.002	0.011	-0.002	-0.006	-0.003	0.003
41,154			(0.002)	(0.006)	(0.004)	(0.006)	(0.005)	(0.004)	(0.005)	(0.008)	(0.003)
1975 cohort		-0.002	0.006	0.001	-0.002	0.003	-0.004	0.004	0.003	-0.002	0.001
41,666		(0.002)	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)	(0.010)	(0.005)	(0.003)	(0.003)
1976 cohort	0.000	-0.001	-0.003	-0.002	0.003	0.002	0.001	0.004	-0.001	0.002	-0.001
44,132	(0.000)	(0.003)	(0.002)	(0.007)	(0.004)	(0.005)	(0.004)	(0.004)	(0.006)	(0.004)	(0.007)
1977 cohort	0.000	0.001	0.000	0.001	0.005	-0.002	-0.002	0.000	-0.005	0.002	0.002
45,410	(0.001)	(0.002)	(0.002)	(0.004)	(0.006)	(0.007)	(0.007)	(0.005)	(0.004)	(0.004)	(0.003)
1978 cohort	-0.001	0.001	-0.004	-0.002	-0.001	0.002	-0.002	0.000	-0.006	0.000	
46,940	(0.001)	(0.001)	(0.003)	(0.004)	(0.004)	(0.004)	(0.010)	(0.003)	(0.008)	(0.004)	
1979 cohort	-0.001	0.001	0.001	-0.002	-0.001	-0.007	-0.002	-0.002	-0.005		
50,968	(0.001)	(0.001)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)	(0.006)	(0.005)		
1980 cohort	0.000	-0.001	-0.002	-0.004	-0.006	0.001	0.002	-0.004			
52,566	(0.001)	(0.001)	(0.002)	(0.003)	(0.006)	(0.005)	(0.004)	(0.004)			
1981 cohort	-0.001	0.001	0.001	0.002	0.000	-0.006	0.001				
54,956	(0.000)	(0.001)	(0.002)	(0.002)	(0.006)	(0.007)	(0.004)				
1982 cohort	0.000	-0.001	-0.001	0.002	0.003	0.001					
55,898	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)					
1983 cohort	-0.001	-0.001	0.001	0.002	-0.001						
56,090	(0.000)	(0.001)	(0.004)	(0.003)	(0.002)						
1984 cohort	0.000	-0.001	0.000	0.001							
58,884	(0.001)	(0.001)	(0.003)	(0.002)							
1985 cohort	0.000	-0.002	0.000								
61,919	(0.000)	(0.001)	(0.001)								
1986 cohort	0.000	0.001									
62,574	(0.000)	(0.001)									
1987 cohort	0.000										
65,915	(0.000)										
Test of no effect	9.88	13.02	10.05	8.98	7.69	7.74	9.20	4.71	9.74	4.32	2.40
p-value, no effect	[0.63]	[0.37]	[0.61]	[0.70]	[0.81]	[0.80]	[0.69]	[0.97]	[0.55]	[0.93]	[0.98]
Pooled Estimates	-0.0003	-0.0004	-0.0001	0.0004	0.0004	-0.0006	0.0002	-0.0009	-0.0027	-0.0003	0.0005
Standard Errors	(0.0002)	(0.0004)	(0.0007)	(0.0010)	(0.0013)	(0.0017)	(0.0017)	(0.0019)	(0.0016)	(0.0017)	(0.0015)
p-value, cohort restrictions	[0.96]	[0.44]	[0.54]	[0.68]	[0.76]	[0.74]	[0.61]	[0.96]	[0.69]	[0.90]	[0.99]
Means	{0.0004}	{0.0027}	{0.0083}	{0.0164}	{0.0228}	{0.0286}	{0.0318}	{0.0263}	{0.0200}	{0.0154}	{0.0121}

Web Appendix Table 1B. Discontinuity in Fraction Giving Birth at Specific Ages, by Cohort: California, Ages 24 through 31

	24	25	26	27	28	29	30	31	Test of No Effect [p-value]	Observed at 23 or Younger	Observed at Any Age
1969 cohort	-0.006	0.001	0.001	0.000	0.000	0.001	0.004	0.004	10.28	-0.003	0.004
46,190	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.004)	(0.004)	[0.59]	(0.026)	(0.038)
1970 cohort	-0.001	-0.004	-0.003	0.001	0.000	0.004	0.001		8.42	-0.015	-0.012
46,190	(0.003)	(0.006)	(0.003)	(0.004)	(0.003)	(0.003)	(0.002)		[0.75]	(0.007)	(0.013)
1971 cohort	-0.005	0.005	0.004	0.004	-0.004	-0.005			9.38	-0.005	-0.006
40,464	(0.005)	(0.003)	(0.005)	(0.003)	(0.003)	(0.007)			[0.67]	(0.027)	(0.040)
1972 cohort	0.006	0.001	-0.002	-0.007	-0.003				16.62	0.010	0.005
39,762	(0.003)	(0.004)	(0.005)	(0.003)	(0.004)				[0.16]	(0.034)	(0.045)
1973 cohort	-0.004	0.002	0.001	0.002					8.65	0.017	0.019
38,594	(0.008)	(0.003)	(0.003)	(0.003)					[0.73]	(0.019)	(0.030)
1974 cohort	-0.002	-0.004	0.003						12.93	0.000	-0.005
41,154	(0.005)	(0.004)	(0.002)						[0.37]	(0.025)	(0.034)
1975 cohort	-0.002	-0.002							8.17	0.008	0.006
41,666	(0.009)	(0.003)							[0.77]	(0.016)	(0.024)
1976 cohort	0.001								4.38	0.004	0.001
44,132	(0.004)								[0.98]	(0.028)	(0.028)
1977 cohort									3.95	0.002	0.005
45,410									[0.97]	(0.021)	(0.020)
1978 cohort									3.80	-0.014	-0.015
46,940									[0.96]	(0.021)	(0.021)
1979 cohort									9.46	-0.015	-0.015
50,968									[0.40]	(0.017)	(0.017)
1980 cohort									6.16	-0.011	-0.011
52,566									[0.63]	(0.023)	(0.023)
1981 cohort									4.52	-0.001	-0.001
54,956									[0.72]	(0.016)	(0.016)
1982 cohort									2.59	0.003	0.003
55,898									[0.86]	(0.010)	(0.010)
1983 cohort									2.02	-0.002	-0.002
56,090									[0.85]	(0.011)	(0.011)
1984 cohort									0.99	-0.002	-0.002
58,884									[0.91]	(0.006)	(0.006)
1985 cohort									7.29	-0.003	-0.003
61,919									[0.06]	(0.003)	(0.003)
1986 cohort									2.96	-0.001	-0.001
62,574									[0.23]	(0.002)	(0.002)
1987 cohort									0.00	-0.001	-0.001
65,915									[0.98]	(0.001)	(0.001)
Test of no effect	2.40	9.89	4.87	3.72	9.97	2.50	2.14	1.05	**122.57**	10.38	6.24
p-value, no effect	[0.97]	[0.19]	[0.56]	[0.59]	[0.04]	[0.47]	[0.34]	[0.30]	**[0.99]**	[0.94]	[0.99]
Pooled Estimates	-0.0017	-0.0003	0.0008	0.0003	-0.0015	0.0002	0.0024	0.0035		-0.0018	-0.0019
Standard Errors	(0.0019)	(0.0015)	(0.0015)	(0.0014)	(0.0015)	(0.0026)	(0.0023)	(0.0042)		(0.0039)	(0.0048)
p-value, cohort restrictions	[0.26]	[0.57]	[0.67]	[0.04]	[0.66]	[0.50]	[0.61]	NA		[0.99]	[0.99]
Means	{0.0097}	{0.0081}	{0.0070}	{0.0059}	{0.0051}	{0.0040}	{0.0031}	{0.0020}		{0.1848}	{0.2307}

Web Appendix Table 1C. Discontinuity in Fraction Giving Birth at Specific Ages, by Cohort: Texas, Ages 13 through 23

	13	14	15	16	17	18	19	20	21	22	23
1969 cohort								0.006	-0.006	-0.005	0.001
32,522								(0.007)	(0.009)	(0.008)	(0.004)
1970 cohort							-0.005	-0.014	0.005	0.001	-0.004
33,938							(0.013)	(0.012)	(0.008)	(0.010)	(0.006)
1971 cohort						-0.009	-0.005	-0.006	0.003	0.011	-0.002
33,258						(0.008)	(0.005)	(0.005)	(0.010)	(0.004)	(0.005)
1972 cohort					0.000	-0.007	-0.004	-0.006	0.002	0.010	-0.003
31,286					(0.005)	(0.006)	(0.011)	(0.009)	(0.006)	(0.005)	(0.004)
1973 cohort				0.000	0.002	0.005	-0.003	-0.002	0.004	-0.007	0.008
30,924				(0.004)	(0.007)	(0.004)	(0.009)	(0.011)	(0.005)	(0.004)	(0.005)
1974 cohort			-0.001	0.005	-0.003	-0.006	-0.001	-0.010	0.001	0.006	0.004
31,806			(0.003)	(0.004)	(0.006)	(0.009)	(0.005)	(0.006)	(0.005)	(0.004)	(0.004)
1975 cohort		-0.002	-0.006	0.002	0.000	-0.001	-0.003	0.004	-0.002	0.000	-0.002
31,794		(0.005)	(0.005)	(0.004)	(0.007)	(0.006)	(0.010)	(0.004)	(0.005)	(0.005)	(0.004)
1976 cohort	-0.002	-0.003	-0.001	0.002	-0.010	0.002	0.015	0.003	0.004	0.006	-0.001
32,776	(0.002)	(0.003)	(0.003)	(0.006)	(0.006)	(0.005)	(0.005)	(0.010)	(0.006)	(0.005)	(0.007)
1977 cohort	0.000	-0.001	0.000	0.005	-0.001	-0.007	-0.002	0.009	0.008	0.004	
33,298	(0.001)	(0.003)	(0.004)	(0.003)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)	
1978 cohort	0.001	-0.002	-0.004	0.004	0.010	-0.003	-0.002	-0.005	0.006		
35,035	(0.001)	(0.002)	(0.003)	(0.004)	(0.005)	(0.008)	(0.005)	(0.007)	(0.004)		
1979 cohort	0.001	0.000	-0.003	-0.002	0.009	-0.008	-0.001	0.000			
37,521	(0.001)	(0.002)	(0.003)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)			
1980 cohort	0.001	0.001	-0.002	0.004	-0.006	0.000	-0.013				
39,878	(0.001)	(0.003)	(0.003)	(0.003)	(0.005)	(0.006)	(0.005)				
1981 cohort	-0.001	-0.003	0.002	-0.004	0.001	-0.003					
41,689	(0.001)	(0.001)	(0.003)	(0.003)	(0.005)	(0.009)					
1982 cohort	0.001	-0.003	-0.001	0.001	-0.004						
43,526	(0.001)	(0.002)	(0.002)	(0.004)	(0.003)						
1983 cohort	-0.001	-0.003	-0.002	-0.002							
42,507	(0.002)	(0.002)	(0.002)	(0.004)							
1984 cohort	0.000	-0.002	-0.002								
44,303	(0.001)	(0.002)	(0.002)								
1985 cohort	0.000	-0.002									
44,143	(0.001)	(0.002)									
1986 cohort	0.000										
43,854	(0.001)										
Test of no effect	7.77	14.72	7.16	9.33	13.22	10.54	17.95	10.73	6.41	18.88	5.92
p-value, no effect	[0.73]	[0.20]	[0.79]	[0.59]	[0.28]	[0.48]	[0.08]	[0.47]	[0.78]	[0.03]	[0.66]
Pooled Estimates	-0.0001	-0.0018	-0.0017	0.0012	-0.0003	-0.0036	-0.0025	-0.0019	0.0027	0.0030	0.0001
Standard Errors	(0.0003)	(0.0008)	(0.0009)	(0.0012)	(0.0017)	(0.0021)	(0.0023)	(0.0023)	(0.0021)	(0.0020)	(0.0018)
p-value, cohort restrictions	[0.67]	[0.98]	[0.96]	[0.64]	[0.22]	[0.65]	[0.08]	[0.39]	[0.93]	[0.07]	[0.55]
Means	{0.0008}	{0.0044}	{0.0117}	{0.0224}	{0.0320}	{0.0380}	{0.0390}	{0.0332}	{0.0253}	{0.0193}	{0.0148}

Web Appendix Table 1D. Discontinuity in Fraction Giving Birth at Specific Ages, by Cohort: Texas, Ages 24 through 30

	24	25	26	27	28	29	30	Test of No Effect [p-value]	Observed at 23 or Younger	Observed at Any Age
1969 cohort	-0.001	-0.003	-0.002	0.005	-0.001	-0.002	0.001	3.87	-0.008	-0.009
32,522	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	[0.97]	(0.016)	(0.038)
1970 cohort	0.005	0.002	0.000	-0.001	-0.003	-0.003		5.93	-0.017	-0.018
33,938	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)		[0.88]	(0.031)	(0.043)
1971 cohort	0.009	-0.006	0.004	-0.005	-0.002			18.61	-0.007	-0.009
33,258	(0.005)	(0.004)	(0.005)	(0.009)	(0.006)			[0.07]	(0.008)	(0.022)
1972 cohort	-0.003	-0.006	0.002	-0.001				9.74	-0.008	-0.023
31,286	(0.005)	(0.004)	(0.008)	(0.004)				[0.55]	(0.044)	(0.047)
1973 cohort	-0.001	-0.005	-0.002					9.29	0.010	-0.006
30,924	(0.003)	(0.007)	(0.004)					[0.60]	(0.038)	(0.040)
1974 cohort	-0.003	-0.001						9.24	-0.005	-0.010
31,806	(0.004)	(0.003)						[0.60]	(0.035)	(0.042)
1975 cohort	-0.010							16.32	-0.011	-0.017
31,794	(0.003)							[0.13]	(0.034)	(0.037)
1976 cohort								14.72	0.016	0.019
32,776								[0.20]	(0.024)	(0.030)
1977 cohort								9.27	0.021	0.014
33,298								[0.51]	(0.031)	(0.033)
1978 cohort								11.24	0.000	0.000
35,035								[0.26]	(0.021)	(0.021)
1979 cohort								9.28	0.001	0.001
37,521								[0.32]	(0.021)	(0.021)
1980 cohort								12.70	-0.025	-0.025
39,878								[0.08]	(0.014)	(0.014)
1981 cohort								10.66	-0.011	-0.011
41,689								[0.10]	(0.024)	(0.024)
1982 cohort								5.55	-0.020	-0.020
43,526								[0.35]	(0.009)	(0.009)
1983 cohort								3.50	-0.003	-0.003
42,507								[0.48]	(0.008)	(0.008)
1984 cohort								1.90	-0.005	-0.005
44,303								[0.59]	(0.005)	(0.005)
1985 cohort								1.35	-0.005	-0.005
44,143								[0.51]	(0.008)	(0.008)
1986 cohort								0.11	0.000	0.000
43,854								[0.74]	(0.002)	(0.002)
Test of no effect	5.92	18.91	6.14	1.37	1.76	1.20	1.14	**153.29**	12.85	11.90
p-value, no effect	[0.66]	[0.01]	[0.41]	[0.93]	[0.78]	[0.75]	[0.57]	**[0.26]**	[0.80]	[0.99]
Pooled Estimates	-0.0005	-0.0030	0.0005	-0.0007	-0.0021	-0.0024	0.0015		-0.0048	-0.0072
Standard Errors	(0.0015)	(0.0019)	(0.0023)	(0.0028)	(0.0026)	(0.0025)	(0.0038)		(0.0051)	(0.0060)
p-value, cohort restrictions	[0.01]	[0.65]	[0.85]	[0.63]	[0.92]	[0.74]	NA		[0.88]	[0.99]
Means	{0.0120}	{0.0102}	{0.0083}	{0.0066}	{0.0051}	{0.0035}	{0.0022}		{0.2411}	{0.2903}

Web Appendix Table 2. Details on Bandwidth Selection

	California							Texas						
	Imbens-Lemieux	Rule-of-Thumb	Percent increase in RMSE compared to IL choice					Imbens-Lemieux	Rule-of-Thumb	Percent increase in RMSE compared to IL choice				
			ROT	30	40	50	60			ROT	30	40	50	60
Fraction Observed	180	73	0.4	4.0	2.2	1.3	0.8	56	57	0.0	2.3	0.8	0.1	0.1
Maternal Age	142	91	0.3	4.6	2.8	2.0	1.6	126	70	0.4	3.4	2.4	1.3	0.8
Maternal Education	92	95	0.0	3.1	1.8	1.2	0.7	180	68	1.2	4.3	3.1	2.3	1.5
Low Birthweight	180	90	0.7	5.1	3.5	2.4	1.6	180	81	0.7	4.6	3.2	2.3	1.7
Prematurity	180	80	0.8	4.8	3.4	2.3	1.6	180	89	0.6	5.1	3.5	2.4	1.6
Infant Death								180	82	0.7	4.5	2.8	1.8	1.3
Mother Smokes	166	62	0.7	4.0	2.3	1.3	0.8	180	135	0.2	5.5	3.8	2.7	1.8
Mother Drinks								180	69	0.5	3.7	1.9	1.0	0.7
Mother Has STDs	180	73	0.5	3.4	1.6	1.0	0.7	180	66	1.1	4.5	2.8	1.9	1.3
Any Care	180	92	0.6	4.5	2.9	1.9	1.5	180	105	0.6	4.2	2.6	1.8	1.4
Care in First Trimester	180	71	0.9	4.7	2.9	1.9	1.3	180	74	0.7	4.0	2.5	1.4	0.9
Number of Visits	180	64	1.4	3.4	2.2	1.7	1.4	180	77	1.1	5.6	3.6	2.5	1.7
Father Present	180	122	0.4	3.5	2.6	2.0	1.7	156	129	0.1	3.3	2.1	1.7	1.2
Father's Age	180	86	0.6	3.5	2.5	1.9	1.6	180	74	1.2	5.1	3.5	2.4	1.8
Father's Education	180	69	0.6	3.3	1.9	1.3	0.9	180	75	0.9	5.0	3.4	2.5	1.6

Notes: First 2 columns present the results of different model selection methods. The Imbens-Lemieux ("IL") method estimates the RMSE directly and chooses the bandwidth that minimizes the estimated RMSE. The Rule-of-Thumb ("ROT") method is based on global polynomial approximations. The final 5 columns present the log difference between the estimated RMSE at the minimizer (IL) and RMSE at the stated bandwidth. See text for details.

Web Appendix Table 3. Tests of Overidentification: Continuity of Baseline Characteristics

California

Maternal Characteristics

Hispanic	Black	Black*	Low Birthweight*	First Month Prenatal Care*
-0.0031	-0.0025	-0.0017	-0.0012	-0.0036
(0.0069)	(0.0044)	(0.0025)	(0.0023)	(0.0160)
[0.25]	[0.52]	[0.96]	[0.55]	[0.65]
{0.42}	{0.13}	{0.10}	{0.07}	{2.76}
214,608	214,608	576,421	575,213	576,421

Grandparental Characteristics*

Native Born	Parity	Child Mortality	Age at Childbirth: Mother	Age at Childbirth: Father
0.0049	0.0188	0.0003	0.0255	0.0607
(0.0044)	(0.0127)	(0.0004)	(0.0546)	(0.0534)
[0.14]	[0.22]	[0.59]	[0.62]	[0.87]
{0.70}	{2.08}	{0.01}	{25.68}	{28.60}
513,213	575,379	573,466	576,421	558,994

Texas

Maternal Characteristics

Hispanic	Black	Black*	Low Birthweight*	First Month Prenatal Care*
0.0051	-0.0020	0.0030	-0.0001	-0.0300
(0.0077)	(0.0056)	(0.0035)	(0.0023)	(0.0188)
[0.93]	[0.68]	[0.77]	[0.71]	[0.91]
{0.39}	{0.18}	{0.15}	{0.08}	{2.95}
188,692	188,692	551,294	550,760	551,294

Grandparental Characteristics*

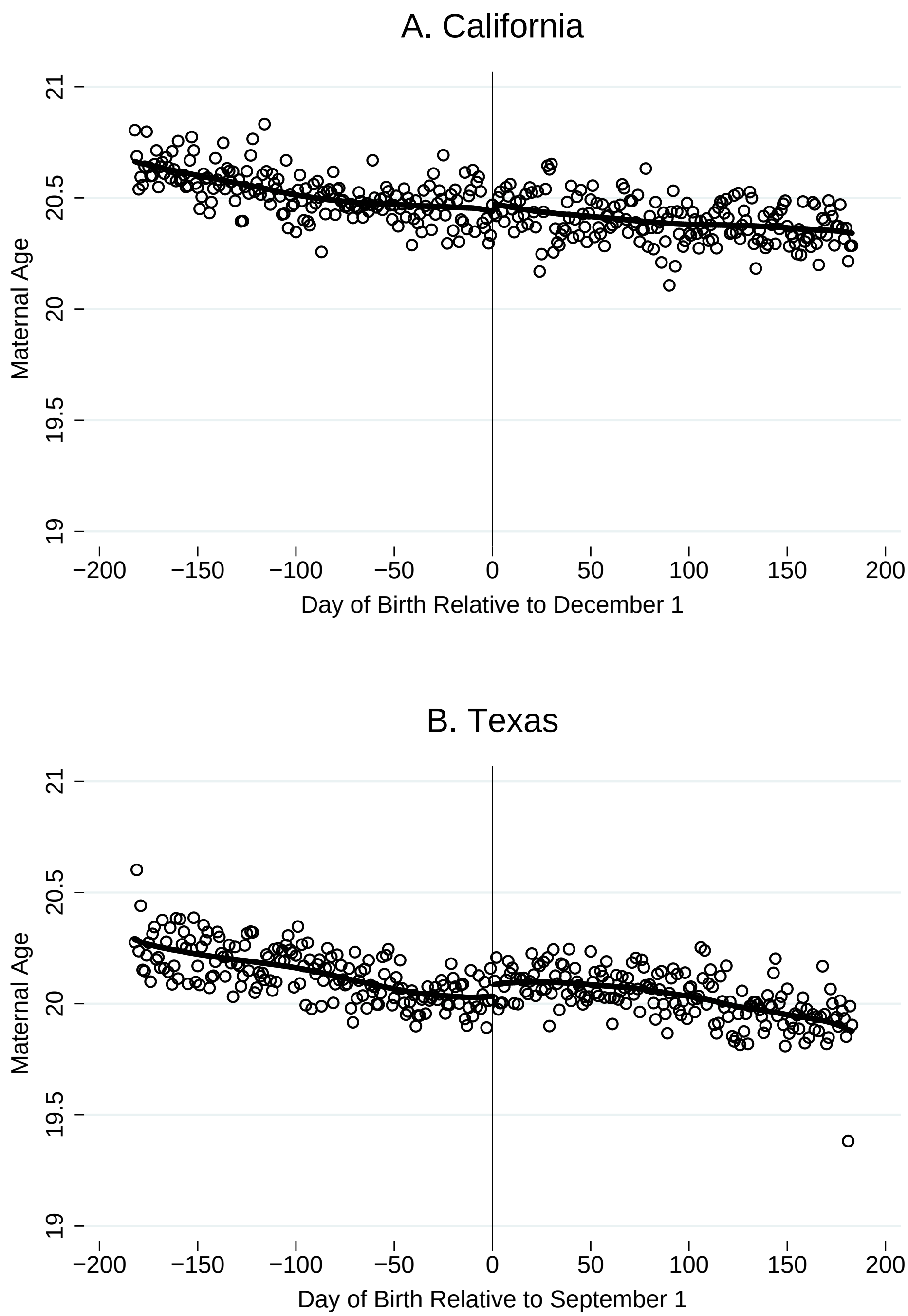
Native Born	Parity	Child Mortality	Age at Childbirth: Mother	Age at Childbirth: Father
0.0023	0.0015	0.0003	-0.0248	-0.0481
(0.0040)	(0.0109)	(0.0004)	(0.0635)	(0.0664)
[0.99]	[0.57]	[0.08]	[1.00]	[0.96]
{0.85}	{2.15}	{0.01}	{24.61}	{27.86}
502,761	551,074	551,069	551,294	475,450

Notes: Standard errors in parentheses. P-values for null hypothesis that discontinuity equal across cohorts in brackets beneath standard errors. Sample means in braces below p-values. Sample sizes below sample means. Stars indicate that estimates are based on public-use files.

Web Appendix Table 4. Power Calculations

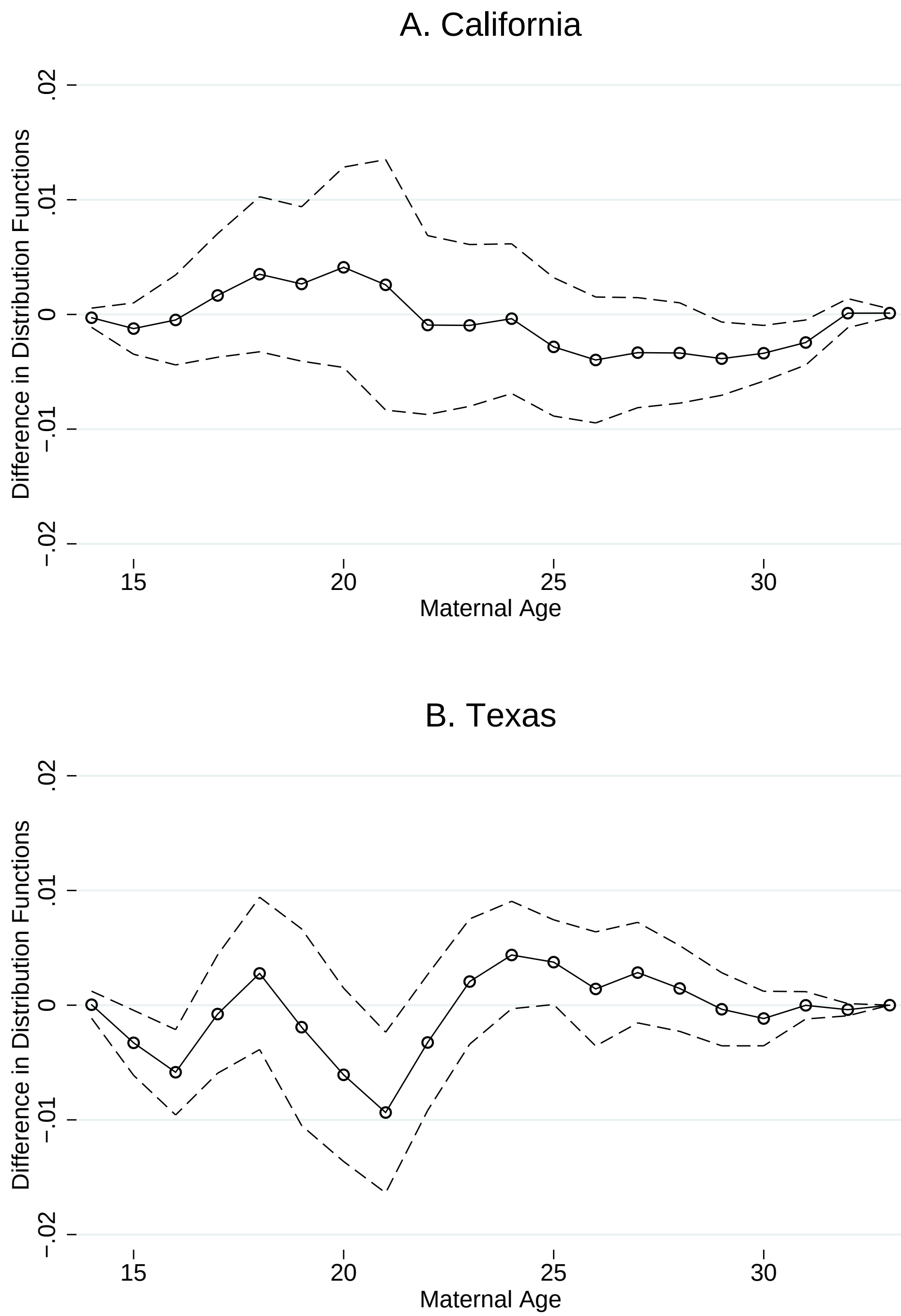
Outcome	Row	Currie and Moretti	Estimate (Std. Err.)	Presumptive Point Estimate (θ_N)	Point Hypothesis (θ_0)	Needed Increase in Sample Size: Fixed Bandwidth Approach	Shrinking Bandwidth Approach
Maternal Smoking	1	-0.06	-0.016	-0.016	-0.060	Rejected	Rejected
	2		(0.008)		-0.030	16%	21%
	3				-0.010	660%	1162%
	4				0	Rejected	Rejected
Prenatal Care in 1st Trimester	5	0.02	0.031	0.031	0	29%	37%
	6		(0.018)		0.020	910%	1702%
	7				0.040	1471%	3027%
	8				0.060	49%	65%
Low Birth Weight	9	-0.01	0.014	0.014	-0.010	Rejected	Rejected
	10		(0.010)		-0.005	Rejected	Rejected
	11				-0.001	50%	65%
	12				0	71%	96%
	13			0	-0.010	247%	373%
	14				-0.005	1287%	2577%
	15				-0.001	34577%	149540%
	16				0	NA	NA
Prematurity	17	-0.01	0.009	0.009	-0.020	Rejected	Rejected
	18		(0.012)		-0.010	52%	68%
	19				-0.005	175%	255%
	20				0	548%	933%
	21			0	-0.020	42%	55%
	22				-0.010	469%	779%
	23				-0.005	2177%	4875%
	24				0	NA	NA
Infant Death	25	NA	-0.006	-0.006	-0.010	294%	455%
	26		(0.004)		-0.005	22653%	88270%
	27				-0.001	266%	407%
	28				0	147%	209%
	29			0	-0.010	Rejected	Rejected
	30				-0.005	208%	307%
	31				-0.001	7589%	22667%
	32				0	NA	NA

Notes: Table presents estimated percent increases in sample size necessary to reject specified hypotheses, under various assumptions on the point estimate that would be obtained in a larger sample. For details on calculations, see text.



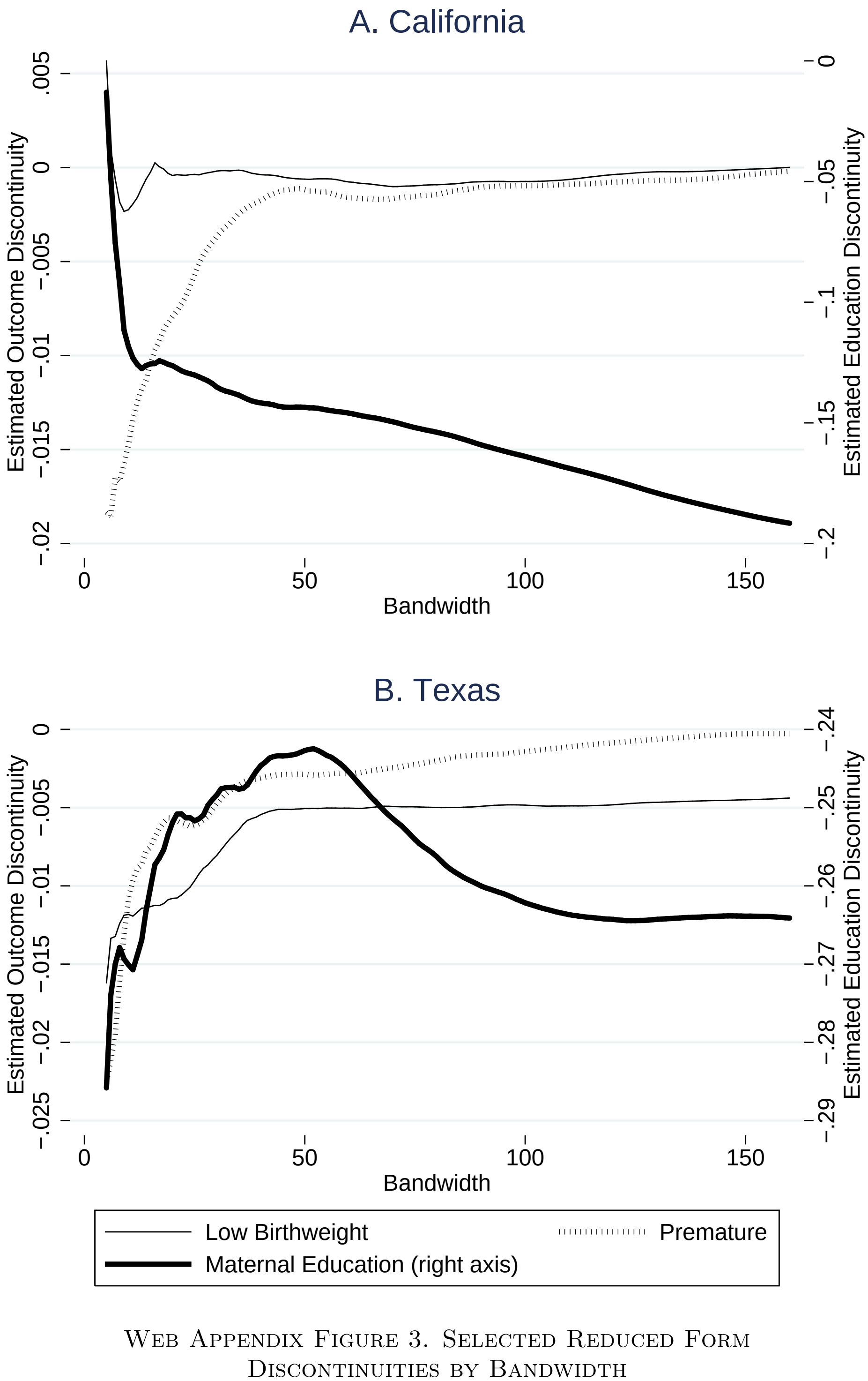
Web Appendix Figure 1. Age at First Birth

Note: Open circles are unconditional averages. Solid curve is a local linear smoother ($h = 50$). Estimates based on post-1969 cohorts. See text for details.



Web Appendix Figure 2. Distribution of Age Effects

Note: Open circles represent differences in distribution functions for age for those born before and after the school entry date. Estimates based on post-1969 cohorts. Dashed lines indicate pointwise confidence regions. See text for details.



Web Appendix Figure 3. Selected Reduced Form Discontinuities by Bandwidth

Note: Estimates based on young women from post-1969 cohorts. See text for details.