Online Appendix
Can Women Have Children and a Career? IV Evidence from IVF Treatments*

Petter Lundborg, Erik Plug and Astrid Würtz Rasmussen

2017

Abstract

In this online appendix we present additional information, tables and figures to which we refer to but not include in the main text of the paper.

*Lundborg: Department of Economics, Lund University, Tycho Brahes Väg 1, 223 63 Lund, Sweden (email: petter.lundborg@nek.lu.se). Plug: Amsterdam School of Economics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands (email: e.j.s.plug@uva.nl). Rasmussen: Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus, Denmark (email: awr@econ.au.dk). We thank three referees, seminar and conference participants for their comments and suggestions. Data in this article are, in part, taken from IVF registers, held by the Danish National Board of Health (DNBH). DNBH is not responsible for the analysis or interpretation of the data presented. Financial support from the Institute for Evaluation of Labour Market and Education Policy (IFAU) and Swedish Research Council for Health, Working Life and Welfare (FORTE) is gratefully acknowledged. The authors declare that they have no relevant or material financial interest that relate to the research described in this paper.

Appendix A: Treatment exogeneity

Table 1 contains estimates from regressions linking treatment success (at the first IVF treatment with embryo implant) to pre-treatment labor market outcomes, controlling for year-of-treatment and age-at-first-treatment fixed effects. We find that the estimates attached to the pre-treatment labor market characteristics are all small and, apart from participation, far from statistically significant. The F statistics, reported at the bottom, further indicate that the pre-treatment labor market outcomes together do not predict treatment success.

Table 1: Pre-treatment characteristics and treatment success

	(1)	(2)
Pre-treatment characteristics	Success	Success
Years of schooling	0.002	0.001
	(0.002)	(0.002)
Annual earnings (in 100,000 DKK)	-0.002	-0.000
	(0.003)	(0.003)
Sickness benefits	-0.010	-0.007
	(0.009)	(0.010)
Marriage	0.002	-0.003
	(0.007)	(0.007)
Positive earnings	0.027	0.029
	(0.014)	(0.015)
Partner's years of schooling		0.002
		(0.002)
Partner's annual earnings (in 100,000 DKK)		-0.002
		(0.002)
F-test, pre-treatment labor market characteristics	1.81	1.26
R-squared	0.02	0.02
Observations	18,538	16,689

Notes: The table shows regressions on the probability of success at first IVF treatment. Column 1 controls for age at first treatment, year of first treatment, and various labor market characteristics, measured the year before IVF treatment. Column 2 adds controls for the partner's pre-treatment earnings and years of schooling. The F-test statistic is a test statistic for joint significance of labor market variables. Robust standard errors are in parentheses.

Appendix B: Some Sensitivity Checks

In Section IV.D, we perform several sensitivity checks to see whether our IVF treatment effect estimates are subject to a number of biases: omitted variable bias related to female health factors, sample selection bias related to wages being observed only for women who work, and sample selection bias related to the unbalanced nature of our short, medium and long-run samples. Full results are reported in Table 2 below.

Table 2: Sensitivity Checks

Independent variable	Earnings controlling for medical factors	Earnings	Pre-treatment education	Pre-treatment wages	Earnings
Sample	*Full sample*	*Healthy women*	*Worker sample*	*Worker sample*	*Balanced sample*
	(1)	(2)	(3)	(4)	(5)
			Panel A: Years 0-1		
IVF success	-47,196	-46,211	0.009	0.002	-49,114
	(1,464)	(2,200)	(0.041)	(1.187)	(1,644)
Observations	18,538	7,378	14,022	14,022	13,779
			Panel B: Years 2-5		
IVF success	-8,088	-7,379	0.028	0.363	-10,266
	(1,728)	(2,667)	(0.042)	(1.940)	(1,954)
Observations	18,436	7,330	12,332	12,332	13,779
			Panel C: Years 6-10		
IVF success	-5,524	-7,953	0.039	2.675	-6,960
	(2,434)	(3,847)	(0.049)	(2.442)	(2,397)
Observations	13,779	5,228	9,627	9,627	13,779
Baseline mean	223,042	222,041	-	-	220,508
Pre-treatment	1,775	-1,270	-	-	1,275
effect	(1,813)	(2,823)	-	-	(2,055)

Notes: Column 1 reruns the reduced form earnings specification (column 2 in Table 2) adding controls for the following medical factors: number of eggs retrieved, number of embryos implanted, diagnoses, causes of infertility, type of IVF treatment, and clinic indicators. Column 2 shows the effect of IVF treatment success on labor earnings for a sample of women who are *not* diagnosed with any fertility problem. Columns 3 and 4 show the effect of IVF treatment success on pre-treatment schooling and pre-treatment earnings (baseline mean). Column 5 reruns the reduced form earnings specification on a balanced panel of women. Time period t=0 refers to the year of the (potential) child birth. All regressions control for age at first IVF treatment, year of first IVF treatment, and pre-treatment education and earnings (except for the regressions in columns 3 and 4). The baseline mean refers to the mean of the outcome variable taken over years 1-4 before the year of the first IVF treatment. The pre-treatment effect refers to the reduced form effect of success at first IVF treatment on the pre-treatment baseline mean (only shown in the relevant regressions). Standard errors are in parentheses.

To be sure that our are wage results are correct and not driven by selective labor market withdrawal, we have also run our wage regressions on extended samples including non-working women with zero wages. We run reduced-form OLS regressions as well as reduced-form Median Quantile regressions. These results are tabulated in Table 3 and show (not unexpectedly) that the short-run wage responses are very different from the ones we report in the paper. We find that the short-run effect of having children on wages turns significantly negative, which is consistent with the short-run effect we find on labor market withdrawal. In the medium and long run, however, the wage responses are remarkably similar. Because of this similarity, we feel quite certain that medium and long-run wage results are not driven by sample selection.

Table 3: Reduced-Form Wage Regressions using All Women

Independent variable	Wages (wages>0)	Wages (zeros for missings)	Wages (median regression)
		Panel A: Years 0-1	
Success	2.899	-12.277	-5.286
	(2.212)	(2.185)	(0.952)
Observations	14,022	18,538	18,538
		Panel B: Years 2-5	
Success	-8.690	-5.845	-4.428
	(1.437)	(1.800)	(1.252)
Observations	12,332	18,435	18,435
		Panel C: Years 6-10	
Success	-5.348	-5.105	-4.341
	(1.861)	(2.227)	(1.355)
Observations	9,627	13,779	13,779

Notes: Column 1 replicates our main (reduced form) wage regression, where inclusion in the sample is conditional on observing positive wages. Column 2 shows the reduced form effect of IVF success on wages where missing wages take the value 0. Column 3 shows median regression results (with zeros for missing). Time period t=0 refers to the year of the (potential) child birth. All regressions control for age at first IVF treatment, year of first IVF treatment, pre-treatment education, and average wages taken over 1 to 4 years before IVF treatment. Standard errors are in parentheses.

Appendix C: Delayed Fertility of Always Takers

Most IVF treated women end up having children despite a first failed IVF attempt. We call this delayed fertility. If the impact of having children on female labor earnings is larger when children are young, delayed fertility will

affect our fertility effect estimates. To see how this works, we introduce a simplified three-period version of our fertility model and two groups of women with different demands for children (or different command over resources). Let women from group a exert a weaker demand for children; these are women who would remain childless after a first failed IVF attempt. Such women always earn Y_a in the absence of children. After a successful first treatment, these women experience an earnings loss C_a in period 1 when children are young, and c_a in periods 2 and 3 when children are older. Let women from group b exert a stronger demand for children; these are women who would end up having children in period 2, despite a first failed IVF attempt. Such women always earn Y_b in the absence of children. After a successful first treatment, they experience an earnings loss C_b in period 1, and c_b in periods 2 and 3 when children are older. After a failed first treatment, such women have a child in period 2 and experience an earnings loss C_b in period 2 when children are young, and c_b in period 3 when children are older. Let β_a and β_b represent the shares of women from group a and b who receive a successful first treatment. In the context of IV estimation, these two groups represent compliers and always takers (Imbens and Angrist 1994). If we compare fertility and earnings differences between successfully and unsuccessfully treated women in period 1, we obtain the first-stage and reduced-form effects in a simplified fertility model without covariates

$$E(F_1|Z=1) - E(F_1|Z=0) = \beta_a + \beta_b,$$

$$E(Y_1|Z=1) - E(Y_1|Z=0) = -\beta_a C_a - \beta_b C_b.$$

Combining these two equations gives us the IVF-based IV estimator

$$\lambda_1^{IV} = -\frac{\beta_a}{\beta_a + \beta_b} C_a - \frac{\beta_b}{\beta_a + \beta_b} C_b = -C_a + \frac{\beta_b}{\beta_a + \beta_b}[C_a - C_b]$$

The IV estimator for period 1 captures the effect of having any children on labor earnings, which is the weighted average of the causal effects in groups a and b. In a similar spirit, we can construct the IV estimator for periods 2 and 3. Combining the fertility and earnings differences in periods 2 and 3 between women who were successfully and unsuccessfully treated in period 1, we get the IVF-based IV estimators

$$\lambda_2^{IV} = -c_a + \tfrac{\beta_b}{\beta_a}[C_b - c_b],$$

$$\lambda_3^{IV} = -c_a.$$

The IV estimator for period 2 captures the causal effect of interest and a nuisance parameter, which is the delayed fertility effect of women in group b who experienced a failed treatment in period 1. Note that the delayed fertility effect will reduce the earnings of these women relative to those who were successful in period 1, so the estimated effect of interest is biased towards zero. The IV estimator in period 3 solves this problem.

More generally, the fertility effect of interest represents a mixture of fertility and delayed fertility effects on labor earnings. While delayed fertility is a concern, it will give us conservative fertility estimates that are still informative about the impact children have on their mothers' labor earnings.

Appendix D: All IVF and Twin Samples

In this paper we use four different samples in estimation: (i) a sample of IVF treated women without children (baseline sample); (ii) a sample of IVF treated women with children; (iii) a sample of successfully treated IVF women (with at least two embryo implants for the twin strategy); and (iv) a sample of representative women. Table 4 provides sample means and standard deviations for a selected number of variables.

Appendix E: Other External Validity Results

Another way to investigate the external validity of our IVF results is to compare labor earnings of women in our primary IVF sample to those of women in a more representative sample drawn from the full population of women who had their first born child around the same time as IVF treated women had their first IVF attempt. Figures 1a and 1b show simple descriptive graphs for average labor earnings for these two groups of mothers, six years preceding the year of childbirth, for the year of childbirth, and for ten years following childbirth. We plot two measures of labor earnings: conditional and unconditional labor earnings. The conditional earnings measure refers to the year-to-year dummy parameters in regressions that control for pre-treatment education and age at having the first child (Figure 1b). The unconditional earnings measure refers to year-to-year average earnings, regardless of the individual characteristics (Figure 1a). In terms of unconditional earnings, one can see that before as well as after childbirth the labor earnings of IVF treated women are consistently higher than those of the representative sample. In terms of conditional earnings, one can see that the decline in earnings is somewhat sharper among

Table 4: Descriptive Statistics of Selected Variables from All Samples.

	IVF sample (main sample) (1)	IVF sample (previous children) (2)	IVF sample (twin trick) (3)	Representative sample (4)
Pre-treatment outcomes:				
Age at first IVF treatment	32.18	34.97	31.53	28.27
	(4.32)	(4.08)	(3.91)	(4.30)
Year at first IVF treatment	2000.20	2000.50	2000.00	2001.45
	(3.11)	(3.11)	(2.99)	(4.07)
Annual earnings (1000 DKK)	244.9	226.1	244.4	201.7
	(140.1)	(150.7)	(131.0)	(136.4)
Schooling	12.83	12.50	12.81	12.55
	(2.34)	(2.47)	(2.30)	(2.32)
Sickness benefits	0.17	0.19	0.17	0.14
	(0.37)	(0.39)	(0.38)	(0.35)
Married	0.52	0.65	0.53	0.31
	(0.50)	(0.48)	(0.500)	(0.46)
Positive earnings	0.91	0.87	0.92	0.90
	(0.28)	(0.33)	(0.27)	(0.30)
Post-treatment outcomes:				
Annual earnings (1000 DK)	233.04	254.69	234.48	178.91
	(141.11)	(138.39)	(117.95)	(127.71)
Positive earnings	0.88	0.95	0.94	0.85
	(0.27)	(0.16)	(0.16)	(0.35)
Observations	18,538	4,598	4,557	103,826

Notes: This table shows descriptive statistics for the following samples and subsamples. Column (1): childless women entering their first IVF treatment. Column (2): women with children entering their first IVF treatment. Column (3): women with a successful first IVF treatment who had at least two embryos inserted. Column (4): a representative sample of women who had at least one child during the study period. Standard deviations within parentheses. Annual earnings are reported in 2008 Danish Kroner (DKK 100 corresponds to USD 20 as of August 2008).

the IVF women and the increase in earnings is somewhat steeper in the representative sample of women. Nevertheless, both measures of labor earnings in the IVF and representative samples follow roughly the same pattern before and after childbirth, which is suggestive that any inherent differences between these two groups of mothers are probably not the leading cause of any differences in how they respond to first born children.

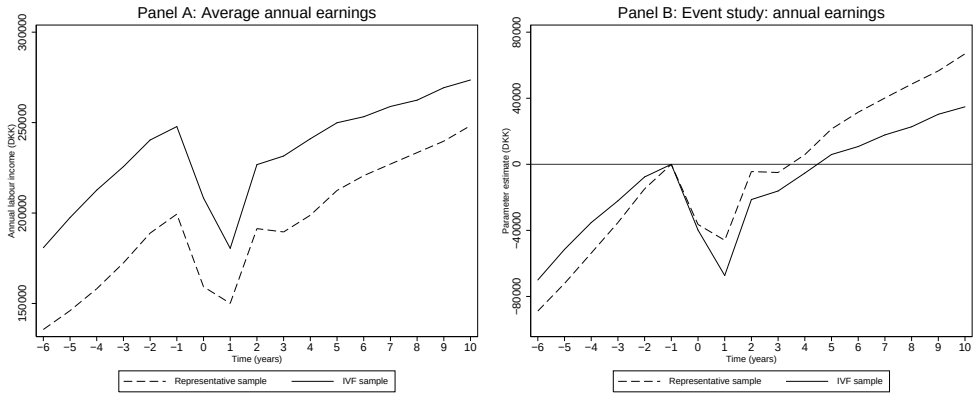

Figure 1: Relationships between (First) Child Birth and Annual Earnings.

Notes: The figures represent unconditional and conditional labor earnings. The unconditional earnings measure refers to year-to-year averages (Figure 1a). The conditional earnings measure refers to coefficients from (before and after) OLS regressions on the relationship between child birth and annual earnings controlling for the year of child birth and age of having a first child (Figure 1b). We have normalized coefficients to 0 in the year before child birth. We use the sample of IVF treated women and the sample of representative women who had their first child between 1995-2005. See text for details.

As an additional check, we run naive OLS regressions of annual labor earnings on having children, age and year fixed effects and years of education using the two samples of women (ignoring that fertility is endogenous). Results are reported in Table 5. If any inherent differences between the women with and without IVF treatment would lead to different labor supply responses to childbearing, we expect to find different associations between annual labor earnings and fertility, measured at the extensive margin. This is not what we observe. The fertility estimates in the IVF treated sample of women, which are all negative and statistically significant in the short, medium, and long run, do not differ much from those found in the sample of representative women. Also the time series patterns of the fertility estimates are roughly the same. We therefore believe that unobserved differences between women with and without IVF treatment are an unlikely reason for differences in labor supply responses.

Table 5: External Validity Check: Results from Naive Least Squares Regressions of Annual Labor Earnings on Having Children.

Independent variable	Earnings	Earnings
Sample	*IVF sample*	*Representative sample*
	(1)	(2)
	Panel A: Years 0-1	
Fertility	-54,176	-51,226
	(1,252)	(472)
Observations	184,976	529,017
	Panel B: Years 2-5	
Fertility	-18,475	-23,431
	(1351)	(498)
Observations	184,354	528,428
	Panel C: Years 6-10	
Fertility	-7,260	-12,157
	(1,647)	(565)
Observations	170,563	524,511

Notes: Column 1 shows OLS (panel) estimates of the relationship between having a first child and annual earnings for the sample of IVF treated women. Column 2 shows the corresponding estimates for the sample of representative women. Time period t=0 refers to the year of the (potential) child birth. All regressions control for age at first child birth, year of child birth, and education. Standard errors are clustered at the individual level and are shown in parentheses.

Appendix F: Other US Twin Studies

We have surveyed three US studies on the impact of children on female labor market outcomes using twins at first birth as an instrumental variable. We take annual labor earnings as the main labor market outcome, focus on impact estimates taken from the 1980 (or more recent) Census Public Use Micro Samples (PUMS), and convert each estimate into percentage point changes in average annual earnings to facilitate the comparison. Bronars and Grogger (1994) study the effect of having twins at first birth on the labor earnings of unwed mothers with children under 18 taken from the 1980 PUMS. On the full sample, they find that earnings decline 15 percentage points for having a second child. Using a sample of mothers with children between 0 to 3 years old, they find that earnings decline 22 percentage points. Bronars and Grogger also report estimates when children are older, but only the short-run estimate is large enough to be statistically significant. Jacobsen, Pearce and Rosenbloom (1999) also examine the effect of having twins at first birth on the labor earn-

ings of all mothers (including unwed mothers) with children under 18 in the 1980 PUMS. They report overall impacts as well as short, medium and long-run impacts. They find an overall decline in average annual labor earnings of 5 percentage points. When they take the children's age into account, they find that annual labor earnings decline 22 percentage points when children are 0 to 2 years old, 8 percentage points when children are 3 to 5 years old, and 4 percentage points when children are 6 to 10 years old. Again, only their short-run estimate is statistically significantly different from zero. Since they report only the reduced-form impact of having twins on female labor earnings, we have rescaled their estimates with the short, medium and long-run first-stage estimates (which are approximately 0.96, 0.67 and 0.60). Vere (2011) uses twins at first birth to examine the effect of having two or more children on the labor earnings of all women who are between 21 and 35 years at the time of sampling in the 1980, 1990 and 2000 PUMS. He finds that having an additional second child reduces earnings by 38, 24 and 20 percentage points using the 1980, 1990 and 2000 PUMS respectively. Because the children in these samples are younger, on average, Vere finds fertility effect estimates that are larger than the ones reported in the other two studies.

When we use twins at first birth as instrument for having additional children on a representative sample of all Danish women, we find that having additional children reduces earnings, on average, by 3 percentage points, regardless of the age of the children (not reported). When we take the children's age into account, we find that having additional children reduces earnings by 8 percentage points in the short run, 2 percentage points in the medium run, and 1 percentage point in the long run. Also we find statistically significant negative fertility effects in the short run, but not in the medium and long run. Overall, we find that all the fertility effect estimates reported in these US studies are larger than those obtained in Denmark.

References

Bronars, Stephen and Jeff Grogger, "The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment," *American Economic Review*, December 1994, *84* (5), 1141–56.

Imbens, Guido and Joshua Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, March 1994, *62* (2), 467–

75.

Jacobsen, Joyce, James Wishart Pearce, and Joshua Rosenbloom, "The Effects of Childbearing on Married Women's Labor Supply and Earnings: Using Twin Births as a Natural Experiment," *Journal of Human Resources*, 1999, *34* (3), 449–474.

Vere, James P, "Fertility and Parents Labour Supply: New Evidence from US Census Data," *Oxford Economic Papers*, 2011, *63* (2), 211–231.