Comment on the Identification Strategy in "Family Ruptures, Stress, and the Mental Health of the Next Generation"

Matsumoto, Brett

June 2016

Comment on the Identification Strategy in "Family Ruptures, Stress, and the Mental Health of the Next Generation"

Brett Matsumoto[†]

Bureau of Labor Statistics[‡]

June 6, 2016

Abstract

The identification strategy of using expected date of birth to define treatment used by Persson and Rossin-Slater (2016) to estimate the causal effect of *in utero* exposure to stress on later outcomes has two potential flaws. First, the endogeneity of actual birth date may make little practical difference. Second, there is likely greater measurement error of gestation age in the treatment group. The conclusion that the authors are the first to obtain a causal estimates in this context would be more credible had these issues been addressed.

1 Introduction

In their paper "Family Ruptures, Stress, and the Mental Health of the Next Generation", Persson and Rossin-Slater attempt to estimate the causal effect of *in utero* exposure to stress and mental health outcomes later in life. They do this by considering mothers who experience the loss of a close relative during pregnancy as the treatment group. The control group is mothers who experience this loss shortly after birth. Let c denote conception, b

[†]Contact: matsumoto.brett@bls.gov. Prior to writing this comment, I posted an outline of the criticisms contained in this document in various online forums. I can certify that I am the author of the relevant posts and that the ideas put forth here are my own. I would particularly like to acknowledge the support of anonymous economists at econjobrumors.com who played a critical role in helping me to formulate the ideas contained within this note.

[‡]The views and opinions expressed in this paper are my own and do not reflect the views of the BLS, Department of Labor, or the Federal Government.

denote birth, and e_b denote expected date of birth ($e_b = c + 280$ days). Then the authors estimate the causal effect using the following equation (equation 5 in the paper):

$$Outcome = \beta_0 + \beta_1 \mathbf{1}[c \leq \ RelativeDeath \ < e_b] + \epsilon, \tag{1}$$

where the estimation sample is limited to mothers who experience the relative death during pregnancy or within one year of birth. The concern with defining the treatment variable as the relative death occurring prior to the actual date of birth is that the actual date of birth may be correlated with the error term. On the one hand, there is the mechanical relationship that a relative death is more likely to occur during pregnancy the longer the pregnancy lasts. Also, there is the concern that the treatment defined using actual date of birth is endogenous since relative death can affect the length of the pregnancy. The authors claim that defining the treatment group based on the expected date of birth corrects for the endogeneity. Alternatively, the causal parameter could be estimated using treatment based on expected date of birth as an instrument for treatment based on actual date of birth.

It should be noted that variations of this identification strategy have been used previously and that the objections raised apply to those papers as well (see Black et al. (2016); Currie and Rossin-Slater (2013); Rossin-Slater (2013)). What follows is an exposition of the two main concerns, and this note concludes with at brief remark.

2 Endogeneity of Actual Date of Birth Examined

As applied economists, we are trained to accept the pervasiveness of endogeneity. Any claim that X is exogenous is treated with skepticism, while a claim that X is endogenous is usually accepted without question. In this context, the authors claim that the regression of the outcome on the endogenous X (treatment based on actual birth date) does not yield a causal estimate, while the regression of the outcome on the exogenous Z (treatment based on expected birth) does. The two stage least squared estimate of the causal parameter using treatment based on expected birth date as an instrument for treatment based on actual birth date is:

$$\beta_1^{IV} = \frac{\partial Y / \partial X}{\partial X / \partial Z} \tag{2}$$

The denominator is the coefficient from the first stage regression. If the instrument perfectly predicts the endogenous variable, the causal parameter simply becomes the OLS estimate of Y on X. In this context endogeneity may not generate much bias in the OLS estimate based on the results of the first stage regression (R^2 of approximately 0.97, Table D-1) and the statement that "the instrument (relative death before expected birth date) is different

from the actual exposure variable (relative death before actual birth date) for only about 1 percent of the individuals in our data."

In a related paper Black et al. (2016), present the results using the same identification strategy on different data and outcomes. They report that the first stage R^2 is very close to one and that the OLS estimates and the 2SLS estimates are very similar when considering later in life outcomes, but they do find differences when considering birth outcomes. Other studies, such as Currie and Rossin-Slater (2013) and Rossin-Slater (2013), in different contexts but still looking at *in utero* effects of some treatment find that instrumenting for treatment based on actual date of birth using expected date of birth does yield different results. Those studies instrument for not only the endogeneity of pregnancy length but also location change. This yields a lower first stage R^2. None of the studies that use expected date of birth to instrument for actual date of birth conduct statistical tests for endogeneity of actual date of birth for the various outcomes. A coefficient that is statistically significant under OLS but not statistically significant using 2SLS is not necessarily due to a correction of bias as it could be due to the higher variance of 2SLS. Similarly, a coefficient that is statistically significant using 2SLS may not be statistically different from a non-significant OLS estimate.

For example, consider the results for the effect of a death during pregnancy on birth weight in table 2 of Black et al. (2016). The naive OLS estimate is -8.081 with a standard error of 5.857, which is not significantly different from zero. The IV estimate is -22.391 with a standard error of 6.153, which is significant at the 5% level. This yields a Hausman test statistic of over 50, which is strongly significant so the null hypothesis of treatment based on actual birth date being exogenous can be rejected. The Hausman test assumes homoskedasticity, but the authors could have easily performed a regression based test. Black et al. (2016) are able to attribute the vast majority of the difference in coefficients to the mechanical relationship between the length of pregnancy and the likelihood of a death occurring during pregnancy. However, correcting for this mechanical relationship only appears to matter for birth outcomes (although not for all birth outcomes). For the long term outcomes, a Hausman test is unable to reject the null hypothesis that treatment based on actual date of birth is exogeneous.

Authors who use this identification strategy should be aware that the exogenous treatment based on the expected date of birth almost perfectly predicts the treatment based on actual date of birth and test for the endogeneity of treatment based on actual date of birth. For some outcomes, the endogeneity of treatment based on actual date of birth does not appear to be significant enough to affect the results. The good news for researchers is that in these cases instrumentation is not necessary to generate an estimate of the causal effect.

Persson and Rossin-Slater (2016) do not report the results from any tests for endogeneity and do not report the OLS coefficients using actual date of birth to define the treatment. Therefore, the reader is not able to evaluate whether endogeneity is a significant problem in their context. This is particularly problematic given the focus on long term outcomes for which the endogeneity of actual birth date may not matter.

3 Measurement Error of Gestation Age

The second concern is related to the use of gestation age to generate the exogenous treatment variable based on the expected date of birth. The expected date of birth is defined as the date of conception plus 280 days. The problem is that the date of conception is not known, so the authors calculate it using the gestation age, which is reported in the data. Gestation age is itself merely an estimate, so the conception age as calculated is measured with error. This is not a problem for this identification strategy unless the nature of the measurement error differs for the treatment and control group.

The initial estimate of conception date is 14 days after the start of the last menstrual cycle. This estimate is later refined based on prenatal ultrasounds. If the estimated gestational age based on the ultrasound is significantly different from that based on the last menstrual cycle, then the ultrasound estimate is used as the preferred measure.[1] Significantly, the accuracy of the estimate of gestational age based on ultrasound decreases as the pregnancy progresses. The most accurate estimate of gestational age is obtained from an ultrasound at between 10-12 weeks. By the 24 week mark, the estimate based on the ultrasound is no more precise than that based on last menstrual cycle (Verburg et al., 2008). The problem for this identification strategy is that individuals in the treatment group may delay or skip prenatal appointments if they coincide with the treatment (e.g., death of a close relative, hurricane, etc.). Particularly if these missed appointments are early in the pregnancy, the treatment group may have poorly estimated gestation ages relative to the control group. Persson and Rossin-Slater (2016) find that mothers in the treatment group are less likely to have adequate prenatal care which corresponds to approximately one missed prenatal exam for mothers who lose a close relative. They claim that this difference in prenatal care is not sufficient to explain the differences in outcomes, but it could affect the estimate of gestational age.

Some of the results are consistent with the hypothesis that mothers in the treatment group have poorly measured gestational ages. They find that the treatment is associated with an increased probability of pre-term birth and an increased probability of c-sections

[1] See American College of Obstetricians and Gynecologists (2014) for details.

(Table 2, Table A5). Since pregnancies that end in a c-section right censor some of the birth outcomes, the higher rate of c-sections could explain the decrease in average birth weight and size measurements. In a study comparing the dating of pregnancies using first and second trimester ultrasounds from the same pregnancy, Thagaard et al. (2016) find that dating based on second trimester ultrasounds leads to a significantly lower rate of post-term pregnancies (2.9% versus 2.1%) and a significantly higher rate of pre-term pregnancies (4.6% versus 5.4%).[2] In Persson and Rossin-Slater (2016) the treatment group has a pre-term rate of 5.34% versus a 4.69% rate for the control group (Table 1), and the causal effect of treatment is estimated to be an increase of 0.6%.

This issue of measurement error in the estimation of gestation age may or may not have much of a practical impact on the results, but authors who wish to use this identification strategy should address whether their treatment and control groups are likely to have similarly measured gestation ages.[3] This is particularly important when the outcome is defined in terms of gestation age such as whether the child is born pre-term (defined as being born prior to 37 weeks of gestation).

4 Conclusion

Estimating the causal effect of *in utero* conditions on birth and later life outcomes presents several challenges relating to identification. This comment addresses one identification strategy that has been used several times in the literature, which is to use expected date of birth to define treatment groups because the actual date of birth may be endogenous. Depending on the outcome being studied, actual date of birth may not be endogenous. Also since the expected date of birth is such a strong predictor of actual date of birth, the endogeneity may make little difference in practice. Researchers can easily address this issue by conducting a regression based test for endogeneity. The second concern is that gestation age may be poorly measured in the treatment group in a way which biases the distribution of the expected birth date in the population. Ultimately, more work is needed on the part of researchers to show that the issue of measurement error in the estimation of gestation age does not affect the estimates using this identification strategy. I believe that this paper could be greatly improved if the authors were able to satisfactorily address these concerns.

[2]The study was limited to spontaneous deliveries, which makes it difficult to compare the post-term numbers to a sample that includes right censored pregnancies (induced labor and c-sections).

[3]The treatment need not lead only lead to a greater measurement error in the treatment group. If the treatment is defined as the presence of a health clinic, as in Rossin-Slater (2013), then the treatment group may have more accurately measured gestation age.

References

Black, S., P. Devereux, and K. Salvanes (2016). Does grief transfer across generations? bereavements during pregnancy and child outcomes. *American Economic Journal: Applied Economics 8*(1).

Currie, J. and M. Rossin-Slater (2013). Weathering the storm: Hurricanes and birth outcomes. *Journal of Health Economics 32*(3), 487–503.

Persson, P. and M. Rossin-Slater (2016). Family ruptures, stress, and the mental health of the next generation. *American Economic Review forthcoming*.

American College of Obstetricians and Gynecologists (2014). Committee opinion no. 611: Method for estimating due date. *Obstetricics and Gynecology 124*, 863–866.

Rossin-Slater, M. (2013). Wic in your neighborhood: New evidence on the impacts of geographic access to clinics. *Journal of Public Economics 102*(0), 51–69.

Thagaard, I., L. Krebs, U. Lausten-Thomsen, S. Larsen, J.-C. Holm, M. Christiansen, and T. Larsen (2016). Dating of pregnancy in first versus second trimester in relation to post-term birth rate: a cohort study. *PLOS ONE 11*(1).

Verburg, B. O., E. A. P. Steegers, M. D. Ridder, R. J. M. Snijders, E. Smith, A. Hofman, H. A. Moll, V. W. V. Jaddoe, and J. C. M. Witteman (2008). New charts for ultrasound dating of pregnancy and assessment of fetal growth: longitudinal data from a population-based cohort study. *Ultrasound in Obstetrics & Gynecology 31*(4), 388–396.