Causal Effects of Smoking on Birth Outcomes

Cindy Lu

December 21, 2020

Abstract

Estimating the causal effect of smoking on birth outcomes is difficult since omitted / unobservable variables affecting birth outcomes are very likely correlated with a mother's smoking decision [1]. The data is sourced from Abrevaya (2006), consisting of the third panel of the Natality Data Sets and a matching algorithm proxy variable. Three models are considered in this paper: the Fixed Effects (FE), Random Effects (RE), and Pooled OLS. Following Abrevaya's true regression specification of birth weight and gestation, the fixed effects regressions, which control for individual heterogeneity, yield significantly different results from pooled and random effects regressions in the full panel [1]. Accounting for the bias introduced by 'false matches', all three models are fit on sub-samples based on the proxy variable. All three models are comparable on the mismatched sample, and the RE model is more appropriate based on the Hausman test for both birth outcome dependent variables. However, the FE estimator of the smoking effect on birth weight is again significantly different from the OLS and RE estimators using the correctly matched data.

1 Summary Statistics

Following Table V in Abrevaya (2006), the summary statistics are computed for the variables to be included in the regression model specified previously; these include mother's age, marital status, whether they smoke, the child's gender (1 if male), race, education metrics, the Kessner indices, and prenatal visits.

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
dmage	296,218	28.48	5.447	13	50
smoke	296,218	0.130	0.336	0	1
male	296,218	0.513	0.500	0	1
married	296,218	0.869	0.337	0	1
hsgrad	296,218	0.294	0.456	0	1
somecoll	296,218	0.234	0.423	0	1
collgrad	296,218	0.378	0.485	0	1
agesq	296,218	840.8	312.8	169	2,500
black	296,218	0.0742	0.262	0	1
adeqcode2	296,218	0.169	0.375	0	1
adeqcode3	296,218	0.0398	0.195	0	1
novisit	296,218	0.00745	0.0860	0	1
pretri2	296,218	0.113	0.317	0	1
pretri3	296,218	0.0199	0.140	0	1

Figure 1: Summary statistics of the panel dataset

The regressors with greater variability are age, the child's gender, age squared, and indicators for being a smoker, gender, marital status, education level, the Kessner index equalling 2, and the first prenatal visit occurring in the 2nd trimester. The variables with less variation are the indicators for race, the Kessner index equalling 3, no prenatal visit, and the first prenatal visit occurring during the 3rd trimester, all of which had standard deviations less than 0.3.

2 Pooled OLS Estimates

The first baseline models are to fit pooled OLS regressions on the data, the panel structure notwithstanding. Adding the year, state, and birth order dummy variables will control for temporal, location, and family size effects in regressions. So 50 state (there were 51 states and US territories), 8 year, and 15 birth order dummy variables were added. The reason to leave one category out for these three types of indicators is to avoid the "trap of dummy variables", which will result in perfect collinearity. Consequences include unreliable parameter estimates, for the model will be unidentifiable. The Pooled model results for both dependent variables, birth weight, and gestation time, are given in Figure 2.

OLS Results

VARIABLES	(1) birthweight	(2) gestation
smoke	-243.2657***	-0.1362***
	(3.064)	(0.013)
male	126.7022***	-0.1250***
	(1.885)	(0.008)
dmage	7.0643***	0.0431***
	(1.686)	(0.007)
agesq	-0.1184***	-0.0010***
	(0.029)	(0.000)
hsgrad	60.5220***	0.0318*
	(3.907)	(0.016)
somecoll	91.3413***	0.0349*
	(4.334)	(0.018)
collgrad	100.8931***	0.0718***
	(4.568)	(0.019)
married	64.4328***	0.1587***
	(3.412)	(0.014)
black	-252.0416***	-0.6038***
	(3.981)	(0.017)
adeqcode2	-100.9336***	-0.4039***
	(3.858)	(0.016)
adeqcode3	-176.4819***	-0.6856***
	(9.043)	(0.038)
novisit	-26.4902*	-0.2900***
	(14.116)	(0.059)
pretri2	89.1162***	0.5328***
	(4.620)	(0.019)
pretri3	154.6603***	0.8337***
	(11.140)	(0.047)
Constant	3,125.6808***	38.7487***
	(23.918)	(0.100)
Observations	296,218	296,218
R-squared	0.097	0.019
MSE	2.143	2.143

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 2: Pooled OLS estimates of birth outcome measures

Some concerns arise when fitting this pooled model. The coefficient of interest of both gestation time and birth weight is the coefficient on *smoke*. Based on Figure 2, the ceteris paribus effect of being a smoker has a strong negative effect on birth weight; children with smoking mothers are born approximately 243.27 grams lighter than children with non-smoking mothers. The ceteris paribus effect of being a smoker also decreases gestation length by 0.136 weeks or approximately 1 day. Although smoking is highly statistically significant, the magnitudes are not reliable.

We check the R^2 values, which in both cases are extremely low. The pooled regression model explains just under 10% and not even 2% of the variation in the birth weight and gestation time regressands. Additionally, there are concerns about model misspecification, due to a "lumping together different individuals at different times", which "camouflage(s) the heterogeneity" that exists among the individuals. Moreover, if the "uniqueness" of each individual mother is "subsumed in the composite error term", then this leads to a non-zero correlation between explanatory variables and the error term u_{it}. Breaking this exogeneity condition of the Gauss-Markov assumptions will lead to both biased and inconsistent OLS estimates [2].

3 Random Effects Estimates

Instead of pooling, which assumes constant coefficients across time and cross-section, we allow for individual intercepts, drawn randomly from larger populations [2]. While they have a common mean value, the differences among individual intercepts, α_i are incorporated into the error term [2]. Based on Figure 3, the coefficient on *smoke* implies that children born to smokers weigh 219.22 grams less than those with nonsmokers, and the gestation period is again almost 24 hours shorter for children born to smokers.

REM Results

VARIABLES	(1) birthweight	(2) gestation
smoke	-219.2264***	-0.1302***
	(3.1557)	(0.0132)
male	129.2549***	-0.1220***
	(1.7468)	(0.0078)
dmage	2.5905	0.0431***
	(1.8029)	(0.0074)
agesq	-0.0420	-0.0010***
	(0.0305)	(0.0001)
hsgrad	66.5698***	0.0329*
	(4.5994)	(0.0178)
somecoll	100.0716***	0.0361*
	(5.0749)	(0.0197)
collgrad	110.6024***	0.0719***
	(5.3356)	(0.0207)
married	69.5116***	0.1572***
	(3.9999)	(0.0155)
black	-248.7042***	-0.5995***
	(4.7003)	(0.0181)
adeqcode2	-95.0867***	-0.3928***
	(3.6373)	(0.0160)
adeqcode3	-164.0689***	-0.6521***
	(8.4805)	(0.0374)
novisit	-34.0691**	-0.3256***
	(13.3968)	(0.0587)
pretri2	80.3360***	0.5100***
	(4.3368)	(0.0191)
pretri3	138.2234***	0.7896***
	(10.3987)	(0.0460)
Constant	3,174.6902***	38.7515***
	(25.7319)	(0.1050)
Observations	296,218	296,218
Number of momid3	141,929	141,929

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 3: Modeling birth outcome measures using Random Effects

4 Fixed Effects Estimates

Like the previous RE model, we allow each cross-sectional unit to have different intercepts for the Fixed Effects model, capturing the heterogeneity. However, unlike the Random Effects model, the individual specific coefficients remain time-invariant. We also note that based on 4, coefficients on the race, marital status, and different levels of education indicators have not been estimated. This is because these time-invariant variables or fixed effect terms α_i get washed out and vanish via the fixed effect transformation, which entails subtracting from each individual equation the time-demeaned data [2]. One key assumption is that the fixed effects estimator allows for arbitrary correlation between α_i and the explanatory variables in any time period, just as with first differencing [3]. Thus, any explanatory variable that is constant over time for all individuals i gets swept away by the fixed effects [2]. An individual's race is fixed, and so if mothers in this sample did not obtain any new education or change their marital status, these would be dropped from the estimation procedure.

FEM Results

	(1)	(2)
VARIABLES	birthweight	gestation
smoke	-144.0363***	-0.0800***
	(4.7515)	(0.0227)
male	133.5766***	-0.1064***
	(2.0784)	(0.0099)
dmage	-15.9802***	0.0900***
	(3.9624)	(0.0189)
agesq	0.3169***	-0.0009***
	(0.0526)	(0.0003)
adeqcode2	-84.4280***	-0.3414***
	(4.4540)	(0.0213)
adeqcode3	-143.9103***	-0.5036***
	(10.2823)	(0.0492)
novisit	-42.3514**	-0.4762***
	(16.5677)	(0.0792)
pretri2	66.5551***	0.4013***
	(5.2653)	(0.0252)
pretri3	111.9026***	0.5874***
	(12.4919)	(0.0597)
Constant	3,534.5035***	37.7997***
	(75.6293)	(0.3615)
Observations	296,218	296,218
R-squared	0.0496	0.0098
Number of momid3	141,929	141,929

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 4: Fixed effects model of birth outcome functions for birth weight and gestation

Figure 4 presents the fixed effect estimators. The coefficient estimate on *smoke* suggests that children of smokers are 144.036 grams lighter than nonsmokers' children, ceteris paribus. The gestation period is shorter by 0.08 weeks for smokers than nonsmokers, with all other variables held constant.

5 Comparing the FE and RE estimates

Comparing the FE and RE estimates involves examining the assumption about correlations between the cross-section specific error α_i and the regressors X. If the α_i and regressors are uncorrelated, the RE model is appropriate. Otherwise, the FE model is preferred. The null hypothesis stated by the Hausman test is that the FE and RE models are not significantly different.

We get highly significant test statistics from the two Hausman tests. The tests suggest that these coefficients look different since the χ^2 values were 615.11 and 221.21 for the birth weight and gestation functions, respectively. Because the p-value is extremely small, we reject the null at both the 5% and 10% levels. This suggests that the correlation between regressors and the cross-sectional error α_i is not zero. The consistency properties state that if the underlying model was Fixed Effects, then the Random Effects estimators are inconsistent [2]. Based on this Hausman test result, it strongly suggests the fixed effects was the true model.

6 Regressions using a Proxy for Correctly Matched Subsamples

However, knowing that there were incorrect matches, the FE estimators can be heavily biased due to omitted variables. Thus a proxy for matches was constructed. We consider the two subsamples separately. In this subsample, there is a high chance of mismatch, so $proxy = 0$. The results are shown in Figure 5 for the three specifications for birth weight, while Figure 6 presents the results for gestation.

	(1)	(2)	(3)
VARIABLES	OLS	Random Effects	Fixed Effects
smoke	-234.1988***	-233.2780***	-213.6243***
	(10.9746)	(10.9840)	(15.7865)
male	127.8833***	127.6973***	124.0041***
	(7.8826)	(7.8711)	(10.8181)
dmage	6.2363	6.0953	-16.7035
	(7.7279)	(7.8686)	(23.3144)
agesq	-0.1467	-0.1443	0.0642
	(0.1361)	(0.1386)	(0.3736)
hsgrad	50.9802***	51.1621***	
	(15.8247)	(16.2125)	
somecoll	85.2566***	85.5788***	
	(18.2675)	(18.7086)	
collgrad	98.9980***	99.3775***	
	(20.0331)	(20.5148)	
married	60.4981***	60.7496***	
	(11.9197)	(12.2058)	
black	-272.6568***	-272.6049***	
	(15.8780)	(16.2710)	
adeqcode2	-112.6482***	-112.0960***	-101.4259***
	(15.8773)	(15.8605)	(21.9786)
adeqcode3	-200.1963***	-197.7964***	-151.0463***
	(33.1438)	(33.1057)	(45.8500)
novisit	-190.0863***	-190.1065***	-188.1245***
	(50.6807)	(50.6214)	(70.0832)
pretri2	86.5422***	85.9864***	73.9577***
	(18.1658)	(18.1419)	(25.0173)
pretri3	190.5878***	186.9589***	117.6595**
	(40.0074)	(39.9362)	(54.6107)
Constant	3,218.0678***	3,220.3770***	3,714.0051***
	(106.6484)	(108.6438)	(393.4113)
Observations	18,610	18,610	18,610
R-squared	0.1102		0.0465
Number of momid3		9,305	9,305

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 5: OLS, Fixed Effect and Random Effects models of birth weight for the sub-sample $proxy = 0$

VARIABLES	(1) OLS	(2) Random Effects	(3) Fixed Effects
smoke	-0.1178**	-0.1168**	-0.1258*
	(0.0484)	(0.0484)	(0.0707)
male	-0.1463***	-0.1474***	-0.1756***
	(0.0348)	(0.0347)	(0.0485)
dmage	0.0960***	0.0998***	0.1040
	(0.0341)	(0.0343)	(0.1044)
agesq	-0.0021***	-0.0021***	-0.0013
	(0.0006)	(0.0006)	(0.0017)
hsgrad	-0.0258	-0.0305	
	(0.0698)	(0.0705)	
somecoll	-0.0187	-0.0238	
	(0.0806)	(0.0813)	
collgrad	-0.0080	-0.0126	
	(0.0884)	(0.0892)	
married	0.1793***	0.1863***	
	(0.0526)	(0.0531)	
black	-0.7751***	-0.7556***	
	(0.0699)	(0.0707)	
adeqcode2	-0.3939***	-0.4633***	-0.4364***
	(0.0685)	(0.0700)	(0.0984)
adeqcode3	-0.1943**	-0.7541***	-0.5942***
	(0.0889)	(0.1461)	(0.2053)
novisit	-1.7241***	-1.1670***	-1.3304***
	(0.1915)	(0.2234)	(0.3139)
pretri2	0.4668***	0.5698***	0.4956***
	(0.0772)	(0.0801)	(0.1120)
pretri3		0.8479***	0.4611*
		(0.1763)	(0.2446)
Constant	38.0998***	38.0385***	37.5277***
	(0.4703)	(0.4739)	(1.7619)
Observations	18,610	18,610	18,610
R-squared	0.0311		0.0148
Number of momid3		9,305	9,305

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 6: Three models of gestation period for the sub-sample $proxy = 0$

According to Abrevaya, the amount of omitted variable bias in this mismatched sub-sample can be quantified. Since the OLS estimate of the smoking effect is around 250 grams and the fixed effects estimates for his first two panels (not available) are roughly 180 grams, the 'true' or unbiased fixed effects estimate is $\approx$ 75 grams [1]. The same calculation applied to this paper's panel data also yields a 'true' fixed effects estimate of around 75 grams, using an inconsistency factor of 0.4 [1].

As expected, for the incorrectly matched $proxy$ =0 sub-sample, the FE estimates of the smoking effect on birth weight and gestation are quite similar to the estimates yielded by OLS and RE. The signs and magnitudes are comparable for this sub-sample. However, there are some differences in statistical significance between the RE and FE models. For example, the effect of smoking and having the first prenatal visit in the last trimester is statistically significant in the RE model at the 0.05 level, while in the FE model, both coefficients are significant at the 0.1 level. The mother's age and age squared are highly statistically significant in the RE model, while they are insignificant in the FE model.

For the correct matches where $proxy = 1$, Figures 7 and 8 present the coefficient estimates by the three models for birth weight and gestation, respectively.

VARIABLES	(1) OLS	(2) Random Effects	(3) Fixed Effects
smoke	-247.0125***	-207.3953***	-66.9852***
	(5.6147)	(5.9684)	(9.9740)
male	125.3874***	127.9296***	131.3339***
	(3.5848)	(3.2045)	(3.7558)
dmage	5.2857	-2.5556	-34.8302***
	(3.2953)	(3.6629)	(8.9441)
agesq	-0.0799	0.0565	0.7490***
	(0.0563)	(0.0624)	(0.1269)
hsgrad	68.0834***	78.4546***	
	(7.0730)	(8.4552)	
somecoll	95.6833***	110.7186***	
	(7.8992)	(9.4028)	
collgrad	104.2993***	121.1936***	
	(8.3477)	(9.9132)	
married	66.7949***	76.4131***	
	(6.9408)	(8.3034)	
black	-253.8969***	-249.1199***	
	(7.4633)	(8.9744)	
adeqcode2	-114.1360***	-103.0468***	-85.9894***
	(7.2997)	(6.6851)	(8.0926)
adeqcode3	-212.9357***	-187.7314***	-151.2456***
	(16.4142)	(14.9634)	(18.0396)
novisit	21.8525	-0.2521	-26.4243
	(25.0935)	(23.2436)	(28.7338)
pretri2	99.7560***	84.1555***	63.1881***
	(8.5688)	(7.8152)	(9.4009)
pretri3	185.1812***	158.9985***	123.9769***
	(19.9432)	(18.0582)	(21.5212)
Constant	3,133.0308***	3,212.0915***	3,722.7037***
	(46.7394)	(52.3004)	(166.5603)
Observations	80,370	80,370	80,370
R-squared	0.1078		0.0526
Number of momid3		40,185	40,185

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 7: OLS, Fixed Effect and Random Effects models of birth weight for the subsample $proxy = 1$

According to Abrevaya, this sub-sample is the most reliable of the samples considered since the birth pairs are likely to be correctly matched, whereas the birth pairs are almost certainly mismatched if the proxy variable is zero [1]. However, we note that the fixed effects estimates are smaller than the other two model estimates and smaller than the fixed effects based on the larger sample displayed in Figure 4. For gestation, however, based on the following table (Table 8), the smoking effect estimated by FE is the opposite sign and is no longer statistically significant, unlike the results obtained in Table 4 's second column. Abrevaya also notes that the results for $proxy = 1$ cast doubt on the fixed effects results from the overall sample that had indicated a significant negative effect of smoking [1].

VARIABLES	(1) OLS	(2) Random Effects	(3) Fixed Effects
smoke	-0.1105***	-0.0980***	0.0466
	(0.0239)	(0.0251)	(0.0509)
male	-0.1322***	-0.1291***	-0.1164***
	(0.0152)	(0.0149)	(0.0192)
dmage	0.0630***	0.0606***	0.1128**
	(0.0140)	(0.0150)	(0.0457)
agesq	-0.0014***	-0.0013***	-0.0006
	(0.0002)	(0.0003)	(0.0006)
hsgrad	0.0016	0.0045	
	(0.0301)	(0.0330)	
somecoll	0.0079	0.0120	
	(0.0336)	(0.0368)	
collgrad	0.0479	0.0513	
	(0.0355)	(0.0388)	
married	0.2223***	0.2243***	
	(0.0295)	(0.0324)	
black	-0.6520***	-0.6499***	
	(0.0317)	(0.0349)	
adeqcode2	-0.4184***	-0.4005***	-0.3226***
	(0.0310)	(0.0307)	(0.0413)
adeqcode3	-0.7056***	-0.6290***	-0.3234***
	(0.0698)	(0.0689)	(0.0921)
novisit	-0.1473	-0.2624**	-0.7564***
	(0.1067)	(0.1060)	(0.1467)
pretri2	0.5282***	0.4987***	0.3775***
	(0.0364)	(0.0360)	(0.0480)
pretri3	0.8170***	0.7450***	0.4566***
	(0.0848)	(0.0834)	(0.1099)
Constant	38.5694***	38.5930***	36.9477***
	(0.1987)	(0.2129)	(0.8502)
Observations	80,370	80,370	80,370
R-squared	0.0216		0.0091
Number of momid3		40,185	40,185

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 8: Three models of gestation period for the subsample $proxy = 1$

7 Hausman Test on the Mismatched Sub-sample

We first compare the FE and RE models of birth weight. Using the Hausman test statistic, we test the null hypothesis, which states that FE and RE estimators do not have systematic differences. The χ^2 test statistic yields a value of 28.66, with a $p-$value of approximately 0.1917. We cannot reject the null hypothesis here. Therefore, this supports the original assumption that the unobserved effects term α_i is uncorrelated with the included regressors. In conclusion, we accept the null hypothesis, and RE is a more preferred model since our original uncorrelated unobserved effects assumption is supported.

Moving on to the Hausman test of the FE and GE gestation models, the χ^2 statistic was 32.42, with a $p-$value of 0.0917. It is statistically significant at the 0.1 significance level, but not at the 0.05 significance level. In this situation, we conclude that FE could be a more appropriate model than the RE model.

8 Hausman Test on the Correctly Matched Subsample

When the matches are correct, the results of the two Hausman tests on the $proxy = 1$ sub-sample agree with the results reported in section 5. The assumption that there exists no correlation between the unobserved effects term and the regressors is tenuous. The estimated coefficients using FE and RE on this sub-sample are systematically different. The χ^2 test statistic of the FE and RE models for birth weight is 424.13, with a very low $p-$value. When the dependent variable is gestation, the χ^2 value was 128.22 with a very low $p-$value. Both test statistics were statistically highly significant, giving evidence against the null hypothesis. To conclude, we reject the null hypothesis, and this implies that the fixed effects model is appropriate for estimating the smoking effect on both dependent variables.

9 Conclusions & Preferred Specification

This paper examined three panel data methods to estimate the causal effects of smoking on birth outcomes despite the omitted variable bias.

In conclusion, this paper's findings suggest that the fixed effect regression is the most preferred model overall., especially on the correctly matched sub-sample. Running the Hausman test on the three data sets (mismatched, correctly matched, and Abrevaya's full third panel), the test statistic gives evidence against the null on the correctly matched and third panel data. Rejecting the null implies that the RE model is not appropriate in determining the causal effect of smoking on birth weight and gestation on the correctly matched sample; it is highly likely that the cross-sectional specific error α_i was correlated with included regressors.

If the fixed model were the preferred model to use on the correctly matched samples, then the FE estimators imply that the deleterious effect of smoking on birth weight is highly statistically significant, even though this estimator's magnitude is less than the corresponding RE and pooled OLS estimators. However, on the $proxy = 1$ sub-sample, the FE estimate of the effect of smoking on gestation is positive and not statistically significant. In contrast, the OLS and RE estimators are negative and statistically highly significant. Yet, when comparing the FE and RE models of gestation on the correctly matched sample via the Hausman test, we find that the χ^2 test statistic is highly significant. So as the FE estimator suggests, the effect of smoking on gestation is not significantly different from zero.

In terms of statistical properties, even if the underlying models were random effects or pooled OLS, the FE estimates are always consistent [2]. On the other hand, if the underlying population model was either RE or FE, the pooled OLS will yield inconsistent estimators [2]. Abrevaya arrives at the same conclusion [1]. Because pooled OLS suffers from its inability to account for individual heterogeneity and underestimates standard errors, the pooled OLS model is not the preferred specification. Only when the slope coefficients are constant across subjects and exogeneity is satisfied are the pooled OLS estimators consistent [2]. Additionally, if the true model is fixed effects, then the random effects estimators are inconsistent [2].

One implication of this finding is highlighting the importance of matching. If this source of omitted variable bias can be minimized, then the FE model would emerge more clearly as the preferred model. This implication agrees with findings in Abrevaya [1]. While Abrevaya does not consider the RE model as an alternative, this paper suggests that on incorrectly matched samples, RE could be an appropriate alternative to the FE and OLS models, for the Hausman test failed to reject the null hypothesis, which assumes that the unobserved error term is uncorrelated with the regressors.

References

[1] Jason Abrevaya. "Estimating the effect of smoking on birth outcomes using a matched panel data approach". In: *Journal of Applied Econometrics* 21.4 (2006), pp. 489–519. DOI: `https://doi.org/10.1002/jae.851`. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/jae.851`. URL: `https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.851`.

[2] Damodar N. Gujarati. *Econometrics by example*. Palgrave, 2015.

[3] Adrian Woolridge. *Introductory Econometrics*. CENGAGE LEARNING EMEA, 2014.