Robustness Null Findings

Robust Null Findings on Offspring Sex and Political Orientation

Byungkyu Lee, *Columbia University*
Dalton Conley, *Princeton University*

In an earlier paper, we deployed the European Social Survey (ESS) and the General Social Survey (GSS) to conduct the largest analysis to date examining the question of whether child sex affects parent political orientation. We found null effects in contrast to earlier, smaller studies. In the current issue, Hopcroft (2016) argues that our null findings may have been obtained due to sample restrictions and measurement error arising from the fact that we used the sex of the first child "residing" in home rather than the sex of the first "biological" child.

We believe that in their comment, Hopcroft largely restates the limitations we have already discussed in our original manuscript, adding only details of the GSS and ESS codebook for the readers. More importantly, she has confused identification issues (e.g., measurement error) with inference issues (e.g., sample restrictions) that we discussed in detail in three pages (Lee and Conley 2016, 1112–14). The goal of our systematic sample selection was to reduce the potential attenuation bias arising from measurement error. Further, by ignoring the period/country variations we showed, this critic missed one of the main points of our paper—we asked why we might observe contradictory findings in the UK and the United States in the first place. We concluded that such results are more likely due to publication bias (or possibly period heterogeneity) rather than to treatment effect heterogeneities or country differences. Nevertheless, we are open to the possibility that we made mistakes in our original paper. In the present response, we have decided to play devil's advocate by taking the opportunity to revisit our case.

There seems to be a straightforward way to measure the sex of the first child; one can ask respondents about the sex/age and biological status of all children they have ever had and infer the sex of the first child from the resulting roster.

We thank the editor of *Social Forces* for providing us an important opportunity to revisit our case and check the likely bias of null findings we initially reported. Replication materials will be made available at *http://dataverse.harvard.edu/dataverse/bk*. *Please direct correspondence to Byungkyu Lee, Columbia University, Department of Sociology, 501 Knox Hall, New York, NY 10027, USA. E-mail: byungkyu.lee@columbia.edu or Dalton Conley, Princeton University, Department of Sociology, 153 Wallace Hall, Princeton, NJ 08544. E-mail: dconley@princeton.edu*

Social Forces 95(2) 899–908, December 2016
doi: 10.1093/sf/sow068
Advance Access publication on 12 October 2016

This is how the 1994 GSS collected the relevant information (Conley and Rauscher 2013). Despite the possibility of coding errors, interviewer bias, respondent fatigue, and recall bias, one can assume that this approach is by far the gold standard by which one can measure the sex of the first child. By contrast, in our paper, we used the information about the sex of all children who are still living in the household, since a complete fertility roster was not available. We tried to minimize measurement error bias as best we could, by including in the sample only those parents whose oldest child is under seventeen in the ESS and the GSS data, following the strategy of prior studies (e.g., Oswald and Powdthavee 2010). Additionally, in the GSS data, we attempted to capture all children each respondent has ever had only by including those who have the same number of kids they have in their life and in their household (though this strategy was not available in the entire ESS data) – thus lowering the chances that we were picking up stepchildren.

In this response, we focus on the 1994 GSS data to show the actual amount of measurement error by comparing the bio sample (that was only available in the 1994 GSS data) and the cohab sample (which we called the "analytical sample" in the original report), and its likely consequences for the inferences we made. Moreover, we revisit the second round of ESS data (years 2005–2006) that asked respondents if there is any child who does not live together with the respondents, and we check how the additional exclusion of those respondents might affect our original conclusions. Additionally, we show the effects of having biological versus adopted/stepchildren on parental political orientation. Finally, we simulate measurement error processes by extrapolating the predicted equation for measurement bias in the 1994 GSS data to other periods and show whether the model-based correction of measurement error can alter the temporally-mixed patterns we reported in figure 2 of the original report.

Issues of identification and inference in the 1994 GSS data

Table 1 shows how much measurement error is reduced by our sample exclusion criteria. We consider the sex of the first child residing in the household as having an error if it does not match the sex of the first biological child as identified by children roster in the 1994 GSS data. Before applying filters, the proportion mismatch between the sex of the first biological and cohab child is 32.6 percent. Note that the amount of error ends up being 7.6 percent in the final analytic sample. The relative rarity of those cases shows that any bias arising from measurement error should be minimal in the 1994 GSS data. Further, we calculate the treatment effect by the difference of the proportion of Republican party identifiers for parents who have the first child as a daughter versus a son. It shows that the treatment effects in the analytic sample either from using the biological child (= 4.8 percent) or the cohab child (= 3.7 percent) are close to the treatment effect in the biological sample (= 5.0 percent). Therefore, at least in the 1994 GSS data, our sampling strategy can effectively minimize bias arising from measurement error.

A second, related issue is whether inferences based on the analytic sample have adequate external validity. This is an issue not only for our analytic sample but

Table 1. Reduction in Measurement Error for the Sex of the First Child by Systematic Sample Exclusion in the 1994 GSS

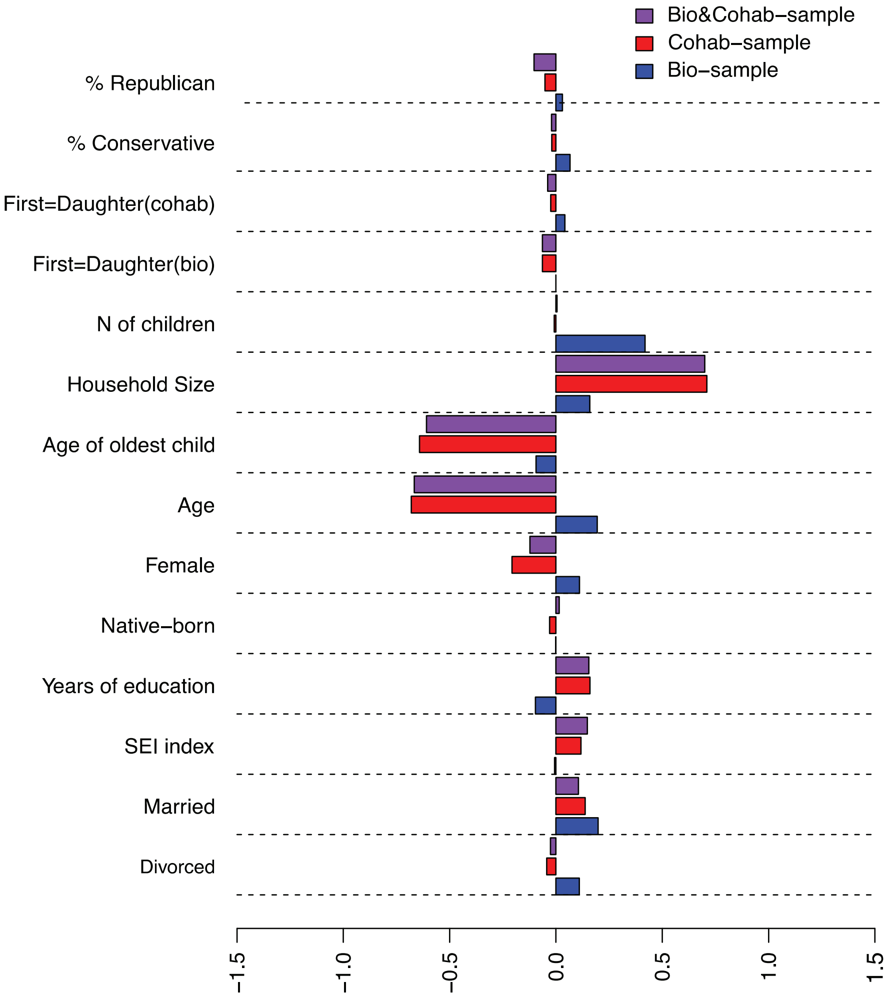
	Biological children sample	Before applying the filter[a]	Final analytic sample
Sample size	1092	584	183
% First biological daughter	46.9%	44.7%	43.7%
% First cohab daughter	48.3%	48.3%	44.3%
% Mismatch of the sex of the first child[b]		32.6%	7.6%
% Republican ($=\mu$)	29.9%	28.2%	23.9%
Treatment effect ($\tau = \mu_{daughter} - \mu_{son}$) of biological child	5.0%	3.1%	4.8%
Treatment effect ($\tau = \mu_{daughter} - \mu_{son}$) of cohab child	1.2%	1.2%	3.7%

Note:
[a]Applying the filter, we excluded the following categories as we described in the online appendix in the original report (Lee and Conley 2016): (1) Generational age gap < 10 ($N = 4$), (2) R has two or more children who are the same age ($N = 1$), (3) Missing data for any child's age or gender ($N = 4$), (d) R's oldest child is older than 16 ($N = 111$), the number of children ≠ the number of cohab children ($N = 60$).
[b]Percent mismatch is twice the proportion of parents whose first biological child's sex is not equal to the sex of one's first child residing in the household, since mismatches that are opposite sex represent only half of all mismatches by chance.

also for the biological sample used in the earlier study (Conley and Rauscher 2013). Figure 1 shows the standardized mean differences of variables in three different samples (bio sample and cohab sample, and the intersection of two) with the full sample (also see table S1). As we noted in our original report, the analytic sample (i.e., cohab sample) consists of younger families and relatively larger households. Interestingly, the bio sample also exhibits some significant differences. Namely, the bio sample includes older people who have are more children and are more female than the full sample. Could these seemingly different sample characteristics make in-sample causal estimates differ?

Table 2 reports the results for OLS models for the party identification across samples. We replicate the effects of having a firstborn daughter on party identification in model 1, which is statistically significant at $p < 0.05$. Model 2 shows results for the analytic sample. The size of coefficient (= 0.277) is almost identical to the original estimate (= 0.265), though it has a wider confidence interval due to reduced sample size. Also, within the same sample (bio and cohab), models 3 and 4 show that the sex of the first biological and first cohabiting child yield both very small and similarly sized estimates (0.159 versus 0.106). The differences between the two point estimates, 0.012 (= 0.277 – 0.265) or 0.053 (= 0.159 – 0.106), are quite small, given that the standard deviation of the

Figure 1. Inference issues due to sample exclusion: mean differences of covariates (full sample – each sample) in standard deviation units



Note: The differences of the means between the full sample ($N = 2{,}992$) and three different analytic samples are divided by the standard deviations of the full sample. The bio sample includes respondents who provide responses for the child roster, the cohab sample includes respondents who were part of the final analytic sample in Lee and Conley (2016), and the bio and cohab sample includes the intersection of the two. The figure is created based on table S1 in the online appendix.

Republican Index in 1994 is 2. Given larger standard errors due to reduced sample size, the lack of statistical significance could simply be due to small sample size in our analytic sample for the 1994 data. However, as we have shown in the original paper, increasing the sample size to 5,179 by combining different GSS surveys across time still does not yield a significant effect. In sum, these almost identical (and small) effect sizes confirm that it is highly unlikely that different characteristics of samples or variable specifications would make a difference.

Table 2. OLS Results for the Effect of the Sex of the First Child on Party Identification in the 1994 GSS

Dependent variable	Republican Scale Index			
	Bio	Cohab	Bio&Cohab	Bio&Cohab
Sample	Model 1	Model 2	Model 3	Model 4
First child = Daughter (bio)	0.265* (0.124)		0.159 (0.290)	
First child = Daughter (cohab)		0.277 (0.205)		0.106 (0.289)
N	1051	347	173	173

Note: The dependent variable is the party identification index that ranges from −3 (= strong Democrat) to 3 (= strong Republican). Controls include parents' sex, age, nativity status, years of education, and age of oldest child. Standard errors are in parentheses. *$p < 0.05$.

Revisiting the second round of the ESS data

Hopcroft argued that "the variable sex of first child from the European Social Survey … has similar measurement issues. They may even be worse, because in the ESS there is no other question asking about the respondent's number of children apart from the household roster" (p. 000). Fortunately, the second round of ESS data asked respondents about non-cohabiting children. The exact wording of the question is "Do you have any children, of any age, who do not currently live in your household? Please include any step-, adopted, foster, or partner's children." Among the final analytic sample in the second round of the ESS data, 11.4 percent of respondents said they have children who do not currently live in the household. We check whether exclusion of those respondents changes the effect of the first (cohab) child's sex on parental conservative index in table 3. A comparison of models 1 and 2 shows that the effect in the original analytic sample (= 0.331) becomes smaller after excluding those "biased" sample (= 0.246), which is also consistent in models 3 and 4, controlling for pre-treatment covariates. It is against Hopcroft's expectations that the effect would get bigger for a more unbiased sample, though coefficients in both samples are not statistically significant. Moreover, given the large standard deviation for the conservative indicator in the second round of the ESS (= 43), both effects and their difference are negligible. Again, the almost identical effect sizes tell us that those measurement errors have a minimal impact on our conclusions, at least for the second round of the ESS data.

Biological, adopted/stepchildren, and parental political orientation

There is one critical-but-untested limitation to our results. As we stated, "one of limitations is that we could not isolate the effects of a biological child versus

Table 3. OLS Results for the Effect of the Sex of the First Child on Political Ideology in the Second Round of the ESS

Dependent variable:	Conservative Ideology Index (from −100 to 100)			
Controls?	No	No	Yes	Yes
Exclude those who have any non-cohab child	No	Yes	No	Yes
	Model 1	Model 2	Model 3	Model 4
First child = Daughter (cohab)	0.331 (0.885)	0.246 (0.933)	0.342 (0.883)	0.251 (0.931)
N	8325	7376	8325	7376

Note: Controls include parents' sex, age, nativity status, years of education, and age of oldest child in addition to survey year dummies and country indicators. Standard errors are in parentheses. No estimates are statistically significant at $p < 0.1$.

adopted/stepchild with the samples we used. It is possible that there exist indeed treatment effects, but they are cancelled out by the association between political ideology and willingness to adopt or marry into a family with daughters (or sons)" (1121). Thus, we additionally test the idea of whether adoption of child or marrying into a family may bias the treatment effect.

Among 1,076 respondents, 8.5 percent (= 91) have an adopted/stepchild in their children roster in the 1994 GSS biological sample. One could come up with a theory that those who adopt female children are more liberal because they seek to redress gender inequality through their family decisions, or that Republicans would be more likely to adopt a child of any gender because they espouse traditional family values (and, by extension, more female children since they are more available for adoption internationally). However, the 1994 GSS data belie both these theories. Table 4 shows that having any adopted/stepchildren does not alter the coefficient of the first child's sex nor show a significant effect on either party identification nor political ideology. Although the biological status of children is necessary for "identification" of causal effects because it ensures that the sex of the firstborn is randomly assigned (under some very weak assumptions), this does not necessarily mean that the effect of non-biological children on parental ideology is biased. It is an empirical question whether those selection biases will systematically play a role, depending on context. We acknowledge that it is still possible that having an adopted/stepchild may have a significant effect in other periods or other countries, but it did not in the United States in 1994.

Measurement error and period heterogeneity

Finally, we conduct one last critical test for the robustness of the null findings we reported. One could argue that correcting the small amount of measurement error for the key right-hand-side variable (and thereby reducing attenuation bias) could make reportedly insignificant effects of the first child's sex become significant. Against this claim, we simulate the measurement error process in

Table 4. OLS Results for the Effect of Having Any Adopted/Stepchild on Parental Political Orientation in the 1994 GSS Bio Sample

	Republican Scale Index				Conservative Scale Index			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
(A) First child = Daughter (bio)	0.265*		0.268*	0.303*	0.176*		0.176*	0.186*
	(0.124)		(0.124)	(0.130)	(0.0864)		(0.0864)	(0.0905)
(B) Have any adopted/Stepchild		0.333	0.340	0.527^{+}		−0.0801	−0.0773	−0.0245
		(0.224)	(0.224)	(0.302)		(0.156)	(0.155)	(0.211)
(A)*(B)				−0.409				−0.114
				(0.444)				(0.308)
N	1051	1051	1051	1051	1027	1027	1027	1027

Note: Controls include parents' sex, age, nativity status, years of education, and age of oldest child. OLS models without controls yield the same results. Standard errors are in parentheses. $^{**}p < 0.01$ $^{*}p < 0.05$ $^{+}p < 0.1$.

other survey periods based on the predicted equation for measurement error in 1994. In doing so, we first estimate logistic regression models for the mismatch of the sex of the first biological and cohab child (model 1) and the sex of the first biological child (model 2) on a set of covariates including female, age, native-born, years of education, the household size, and the total number of children (see table S2 for the full regression tables in the appendix). And, then, we simulate coefficients for each covariate from the multivariate normal distribution with mean $\hat{\beta}$ and variance matrix $\sigma^2 V_\beta$ ($\hat{\beta}$ is the vector of estimated parameters, V_β is the unscaled estimation covariance matrix, $\sigma = \hat{\sigma}\sqrt{(n-k)/X}$ where $\hat{\sigma}^2$ is the residual variance, and X is a random draw from the χ^2 distribution with $n-k$ degrees of freedom) using the *arm* package in **R** (Gelman and Hill 2007), based on which we predict the probability of a mismatch for the first biological child's sex per individual in each year with the same set of variables. We then correct the sex of the first child either based on the predicted probability for the mismatch or the biological child's actual sex. And, then we re-estimate the effects of corrected sex of the first child on party identification. We repeat this process 1,000 times.

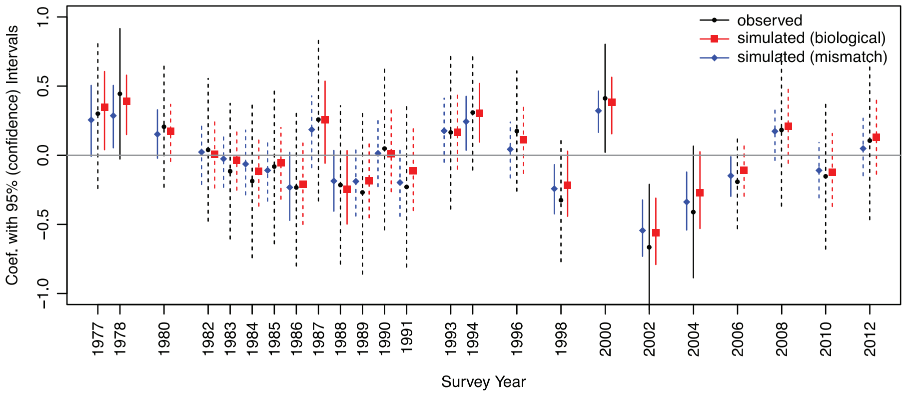
Figure 2 shows that the simulated coefficients from two different measurement models largely overlap with the observed coefficient and its 95 percent confidence interval. The figure suggests a minimal role for measurement error. That said, in some years, such as 1977, 1988, 1994, and 1998, correcting for measurement error leads insignificant effects to become significant. However, the directions of effects make interpretation complicated. For example, having a daughter makes parents lean toward the Republican Party in 1977 and 1994 but toward the Democratic Party in 1988 and 1998. These different directions of effects cancel each other out to produce essentially null effects.

What next?

We have shown that Hopcroft's claims about the possible role of measurement error and sample selection, while theoretically possible are empirically unwarranted based on reanalysis of the 1994 GSS data and the second round of ESS data. Also, we additionally discover that having an adopted/stepchild does not bias the effect of the first child's sex nor significantly influence parental political orientation in the United States in 1994. Finally, we show that correcting for measurement error can recover the significance of effects in some years toward opposite directions but not in most periods. The last point resonates with the following claim that we made in the original paper:

> As evinced, the previous findings (i.e., Conley and Rauscher 2013; Hopcroft 2005) that used the 1994 GSS sample, which contains the information about all biological children, are hard to generalize to other periods. This may be due to the peculiarity of the 1994 GSS sample. Or, alternatively, as we have shown in Figure 2 and Figure S3, it could be the case that parents react to their children by focusing on particularly salient issues at the moment in time, thereby adapting their political attitudes corresponding to the time-varying political context in

Figure 2. Simulated period variation: the effects of the first child's sex on parental party identification by extrapolation of the predicted equation in the 1994 GSS sample to other survey years



Note: Black lines indicate the observed bivariate regression estimates of effects of having an oldest daughter on party identification (Republican Index) with 95 percent confidence intervals. Blue lines show the 95th-percentile range of 1,000 estimated coefficients when the sex of the first child is predicted by logistic regression models for the first biological child's sex. Red lines show the 95th-percentile range of 1,000 estimated coefficients when the sex of the first child is predicted by logistic regression models for the mismatch of the first child's sex between biological and residing kids. The predicted equations are presented in table S2. Statistically significant effects (at $p < 0.1$) are displayed by solid lines.

> combination with the sex of their offspring (Highton and Kam 2011). (Lee and Conley 2016, 1124, note 22)

There is indeed a compelling analytical motivation behind the burgeoning literature on the role of child's gender in various parental outcomes. However, despite its methodological advantage for the identification of causal effects, offspring gender may simply not affect everything we theorize that it will. As for political orientation in particular (which is known to be durable and consists of multiple dimensions of issues and attitudes), it seems likely that publication bias explains the surfeit of reported effects (in both directions). Indeed, one of the present authors was on the original study that used the 1994 GSS and showed tantalizing effects of child gender on parental partisanship. We are thus not in any way predisposed to report a null effect, thereby refuting our own earlier work. But we must yield to the weight of evidence and conclude that the original study was a statistical aberration. We believe that scientific progress will be made if one builds a strong theory and proposes an empirical evidence to support it. Unfortunately, there is no evidence that offspring gender affects parental political orientation.

Supplementary Material

Supplementary material is available at *Social Forces* online, http://sf.oxfordjournals.org/.

References

Conley, Dalton and Emily Rauscher. 2013. "The Effect of Daughters on Partisanship and Social Attitudes Toward Women." *Sociological Forum* 28(4):700–718.

Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.

Highton, Benjamin and Cindy D. Kam. 2011. "The Long-Term Dynamics of Partisanship and Issue Orientations." *Journal of Politics* 73(1):202–15.

Hopcroft, Rosemary L. 2005. "Parental Status and Differential Investment in Sons and Daughters: Trivers–Willard Revisited." *Social Forces* 83(3):1111–36.

Hopcroft, Rosemary, "Comment on Lee and Conley (2016) 'Does the Gender of Offspring Affect Parental Political Orientation?' Lee and Conley's null findings are likely due to measurement error in their measure of sex of first child," *Social Forces* 95(2):893–897.

Oswald, Andrew J. and Nattavudh Powdthavee. 2010. "Daughters and Left-Wing Voting." *Review of Economics and Statistics* 92(2):213–27.