Causal Inference Replication 3

Matthew Borelli

4/12/2020

INTRODUCTION

Betsey Stevenson and Justin Wolfers project, "Bargaining in the Shadow of the Law: Divorce Laws and Family Distress", focuses on the expansion of divorce laws in the U.S. as they relate to changes in domestic abuse. They determine the causal effect of unilateral divorce laws on family violence metrics, particularly domestic violence rates and suicide rates. Stevenson and Wolfers utilize the natural experiment of differentiated timing in the rollout of unilateral divorce laws in U.S. states from the mid-1970's to the mid-1980's in order to exploit the following declines in divorce rates. Their data for this project comes from a few sources. Suicide data comes from the National Center for Health Statistics, a census of death certificates with codes for suicide deaths. Domestic violence data is from the Family Violence Surveys taken in 1976 and 1985, with data being collected from household interviews that asked couples how they resolve conflict. Lastly, their homicide data comes from the FBI Uniform Crime Reports (UCR). UCR data is created from a voluntary police reporting system, which collects incident-level data including information on the offender, victim, and their relationship. The authors note that these data sources are not fully reliable for two reasons: They rely on the voluntary participation of agencies, and there are coding breaks stemming from changes in the definition of of the victim-perpetrator relationship.

Stevenson and Wolfers found that unilateral divorce has statistically significant effects for women, while having no evidence of any benefit for men. They estimate that unilateral divorce laws are associated with a 10% decrease on average for the rate of intimate murders against women, 30% decreases in domestic violence for men and women, and 5-10% decreases on average of the suicide rate for women, with larger long-term effects. Figure 2 is Stevenson and Wolfers'key figure, as is shows the decline in intimate homicides holds under each of the potential definitions (spouse, non-stranger, and family member). I generally find these problems convincing, but there is one potential problem that would be good to look further at. In class, we discussed that estimating δ with differential timing and heterogeneous treatment effects over time produces biased estimates, even possibly changing the sign of the result entirely. While the intuition of the paper and results make logical sense, there is certainly justified concern that the estimates found in Stevenson and Wolfer' paper are biased.

Background

Unilateral divorce, also known as no fault divorce, is a dissolution of a marriage by only one side of the party that can be carried out for any reason. This contrasts with fault-based divorces, the standard that existed prior. In fault-based divorce proceedings, judges only granted divorces if there were some fault, such as adultery or domestic violence, that could be proved to the court. No fault divorces, as implied by their name, do not have to have any cause established and can be initiated by either party. Prior to unilateral divorce laws, people in bad or violent marriages did not have favorable choices for leaving. Leaving without becoming divorced was possible but would leave partners unable to claim assets from the marriage or even get remarried.

The authors consider the likelihood that more of these violent relationships ended with either suicide or intimate homicide and that unilateral divorce laws, which allow abused spouses to leave the marriage with

a more favorable position, would lower the amount of these relationships that end in negative ways. This contradicts the main predictions of the Coase theorem, in which unilateral divorce laws only transfer the property right or remarriage from the partner who wants to remain married to the one who doesn't, that marriages only end in divorce if the marriage is jointly suboptimal. Under Coase theorem, there are no"inefficient marriages",so changes to divorce laws that allow for unilateral divorce should not affect divorce rates. Thus, Stevenson and Wolfers hypothesis will only be proven correct if the Coase theorem is violated.

Estimation Method

The following equation is the estimation equation that Stevenson and Wolfers utilize

$$Suicide\,rate_{s,t} = \sum_k \beta_k Unilateral_{s,t}^k + \sum_s \eta_s State_s + \sum_t \lambda_t Year_t + Controls_{s,t} + \epsilon_{s,t}$$

Generally, this equation estimates the causal effect of unilateral divorce laws on suicide rates while accounting for state, year, and other controls. To break down this equation, there are thee subscripts (k, t, s), three summations, three coefficients, and the error term that should be explained.

- k denotes the number of years since the adoption of unilateral divorce.
- s denotes what state the observation is from.
- t denotes the year an observation is from.
- $\sum_k$ is summation of all numbers of years since a state has passed unilateral divorce laws.
- $\sum_s$ is summation of all states in the model.
- $\sum_t$ is summation of all years in the data set.
- β_k is the effect of a one unit change in the $Unilateral_{s,t}^k$ variable, a binary variable set to 1 is a state adopted unilateral divorce laws k years ago.
- η_s is the fixed effect of $State_s$.
- λ_t is the fixed effect of year = $Year_t$.
- $\epsilon_{s,t}$ is the error term for each $State_s$ and $Year_t$ combination.

The main parameter of interest is β_k, the effect of unilateral divorce laws k years after they were adopted, on suicide rates. At $k = 1$ for example, we estimate β_1 which is the causal effect of unilateral divorce laws on the suicide rate in states which adopted unilateral divorce laws one year before.

Pre-Analysis

In differential timing models, there are generally three timing categories that we can expect to find in the population, which in this case is the U.S. states.

1. Never-Treated: States that never adopted unilateral divorce laws, coded as NRS.
2. Always-Treated: States that adopted unilateral divorce laws before the sample period, coded as PRE.
3. Treated: States that adopted unilateral divorce laws after the sample period or never.

Table 1: Adoption Year of Unilateral Divorce by State

State	Year	State	Year	State	Year	State	Year
Alabama	1971	Arizona	1973	Arkansas	NRS	California	1970
Coloado	1971	Connecticut	1973	Delaware	NRS	District of Columbia	1977
Florida	1971	Georgia	1973	Idaho	1971	Illinois	1984
Indiana	1973	Iowa	1970	Kansas	1969	Kentucky	1972
Lousiana	PRE	Maine	1973	Maryland	PRE	Massachussetts	1975
Michigan	1972	Minnesota	1974	Mississippi	NRS	Missouri	1973
Montana	1975	Nebraska	1972	Nevada	1973	New Hampshire	1971
New Jersey	1971	New Mexico	1973	New York	NRS	North Carolina	PRE
North Dakota	1971	Ohio	1974	Oklahoma	PRE	Oregon	1973
Pennsylvania	1980	Rhode Island	1976	South Carolina	1969	South Dakota	1985
Tennesseee	NRS	Texas	1974	Utah	PRE	Vermont	PRE
Virginia	PRE	Washington	1973	West Virginia	PRE	Wisconsin	1977
Wyoming	1977						

Additional distinctions can be made depending on the timing of treatment during the sample period, such as creating a distinction between early and late adoption, but this is optional. Table 1 lists all of the states in the sample set along with the year that they adopted unilateral divorce laws.

For this replication, the "Never-Treated" group of states includes Arkansas, Delaware, Mississippi, New York, and Tennessee. We would expect that these states should not experience any significant change of trend in suicide or homicide rates over the time period we are looking at, since they never adopted unilateral divorce laws (at least in sample of time from 1964-1996). The "Always-Treated"group of states includes Louisiana, Maryland, North Carolina, Oklahoma, Utah, Vermont, Virginia, and West Virginia. With these states, we would expect that the suicide and domestic violence trends don't change during our sample period as well. Every other state is in the "Treatment" group, meaning that those states started the sample period untreated, adopted unilateral divorce laws between 1964 and 1996, and ended the time period as treated. Most of the states in this category adopted unilateral divorce laws in the 1971-1976 year range, with some outliers on either side. If unilateral divorce laws have a significant impact on suicide rates and homicide rates, we would expect that the trend in treatment states should change from the pre-treatment to post-treatment period for each state.

Results

We will estimate the above equation using OLS regression, as well as estimating a similar equation for intimate homicide rates amongst women. For the dependent variable, the first set of estimations will use the number of suicides per 100,000 people for a given state and year. The second set of estimations will use the number of intimate homicides per 100,000 people for a given state and year. For both sets of estimations, the independent variable is an indicator for whether the observation is taking place after that state has adopted unilateral divorce laws. We estimate each dependent variable using three estimation methods

1. OLS regression with state and year fixed effects
2. OLS regression with state and year fixed effects and clustered standard errors by state
3. OLS regression with state and year fixed effects, clustered standard errors by state, and a trend term interacted with each state fixed effect.

Table 2: Difference-in-Difference Estimation

Outcome	**Suicide Rate**			**Homocide Rate**		
	1	**2**	**3**	**1**	**2**	**3**
$Post - Adoption$	-3.08**	-3.08	-2.75	-0.15*	-0.15	-0.16
	(1.112)	(2.431)	(2.402)	(0.089)	(0.15)	(0.152)
Controls	x	x	x	x	x	x
Clustered SE		x	x		x	x
Trend			x			x

Standard errors in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

Table 2 shows the estimates for the post-adoption interaction varible, a simple estimation for the effect of unilateral divorce laws on suicide and homicide rates. Model 1 for each dependent variable gives significant estimates of a negative effect, at the 5% level for suicide rate and the 10% level for homicide rates. However, in models 2 and 3 we see no statistical significance for the estimations. Clustering standard errors for each state and adding in a trend/state interaction raised the standard errors of our estimates. This means that we do not have strong evidence that unilateral divorce laws lower suicide rates or intimate homicide rates, contradicting the findings of Stevenson and Wolfers.

Our next route to estimate δ the Variance Weighted Averate Treatment of Treated (VWATT), which is described by this equation:

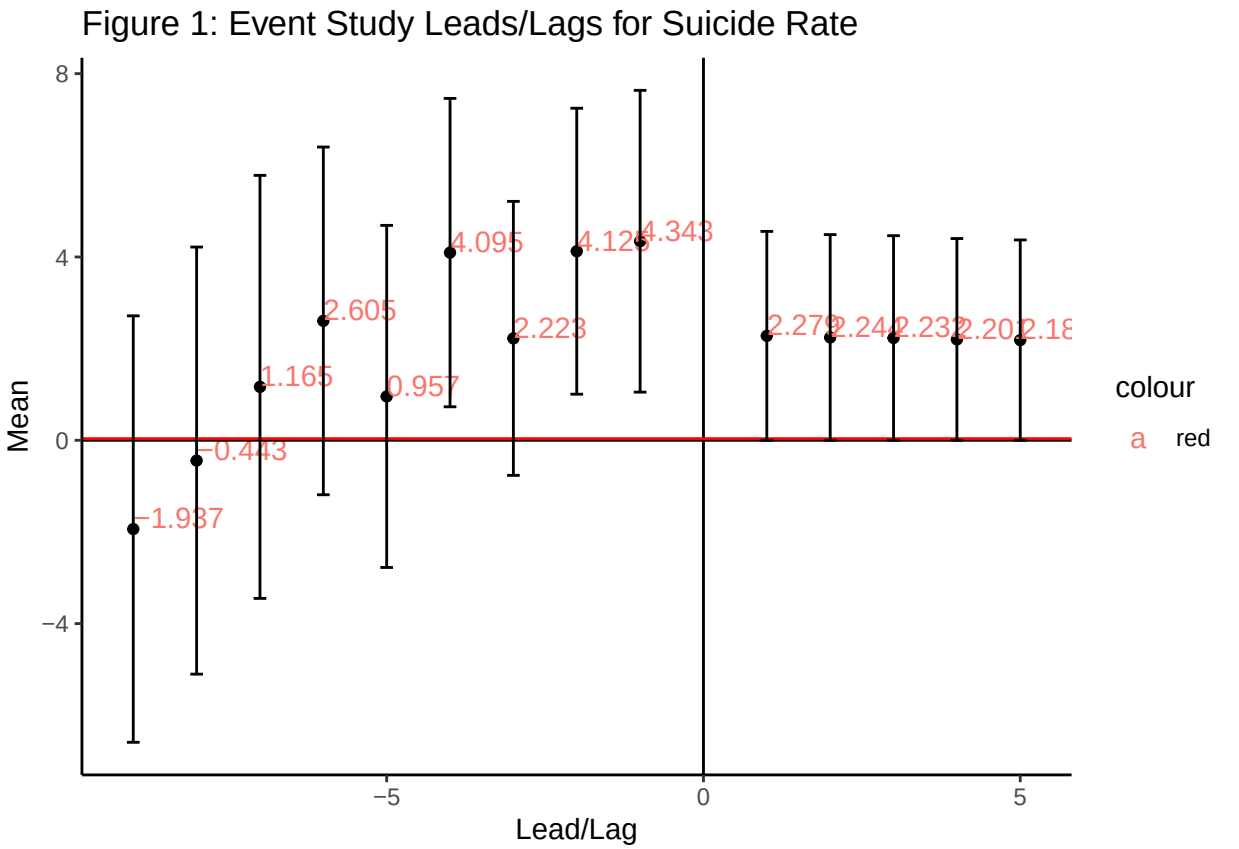
$$p\lim \ \hat{\delta}^{DD}_{n\to\infty} = VWATT + VWCT + \Delta ATT$$

In order to identify the VWATT, we have to make two identifying assumptions:

1. Variance Weighted Common Trends (VWCT) = 0
2. Change in Average Treatment of Treated (ΔATT) = 0

To handle the ΔATT assumption, you can test if treatment effects differ over time or state, which we did above and can say that we believe this to be 0. The best way to confirm that VWCT is equal to 0 is to confirm that states have common trends before they are treated. To test the plausibility of this assumption, we create an event study to see if there are significant differences in the pre-treatment period trends.

Figure 1 shows the results of the event study model for suicide, marking the effects of each lead and lag for 9 leads and 5 lags. The coefficients listed for each mean are the effects of each amount of time before/after treatment, compared to the control states. For this figure, the control states is the set of Never-Treated states. For our assumption of VWCT = 0 to hold true, we need to see that there was not a significant difference in the pre-treatment leads between treatment states and control states. Most of the coefficients for the leads are not statistically significant, with the exception of leads 1 and 2. Those leads are statistically significant since their error bars do not overlap with zero. This means we do have some evidence that VWCT $\neq$ 0, which means we might have biased estimates of VWATT. This is concerning, as we cannot reasonably state what the causal effect of unilateral divorce laws on suicide rates and intimate homicide rates are if we believe that our estimates are biased. Even so, if our results were not potentially biased, if you look at the lags you can see that we can't claim that those differences are statistically significant either. In fact, the treatment sates appear to get closer to the control states after adopting unilateral divorce laws, which is the opposite of what we hypothesized.



Next we perform a Bacon Decomposition in order to split the treatment effects into individual 2x2 difference-in-difference numbers. Figure 2 plots each of the treatment effects against their calculated weight, with more important estimates having a higher weight. There are four types of treatment effects that the Bacon decomposition categorizes each observation into:

1. Earlier vs. Later Treated
2. Later vs. Always Treated
3. Later vs. Earlier Treated
4. Treated vs. Untreated

Each point in Figure 2 is color-coded by its type which lets us see how the categories are distributed in both weight and estimates. Group 2 has mostly high weights that lean negative, which is corroborated in Table 3 as it has the highest combined weight of the four categories. On the other hand Group 1 generally has low weights and has estimates spread fairly evenly around zero, which means Group 1 has both a low weight and an estimate near 0 which has little impact on the average estimate of the DD coefficient.

The weighted average of the groups in the Bacon Decomposition is theoretically the same as the DD coefficient that we estimated using OLS earlier. To prove this, we can sum the weighted averages using this formula:

$$\hat{\delta}^{DD} = \sum_{i=1}^{4} Weight_i * AverageEstimate_i$$

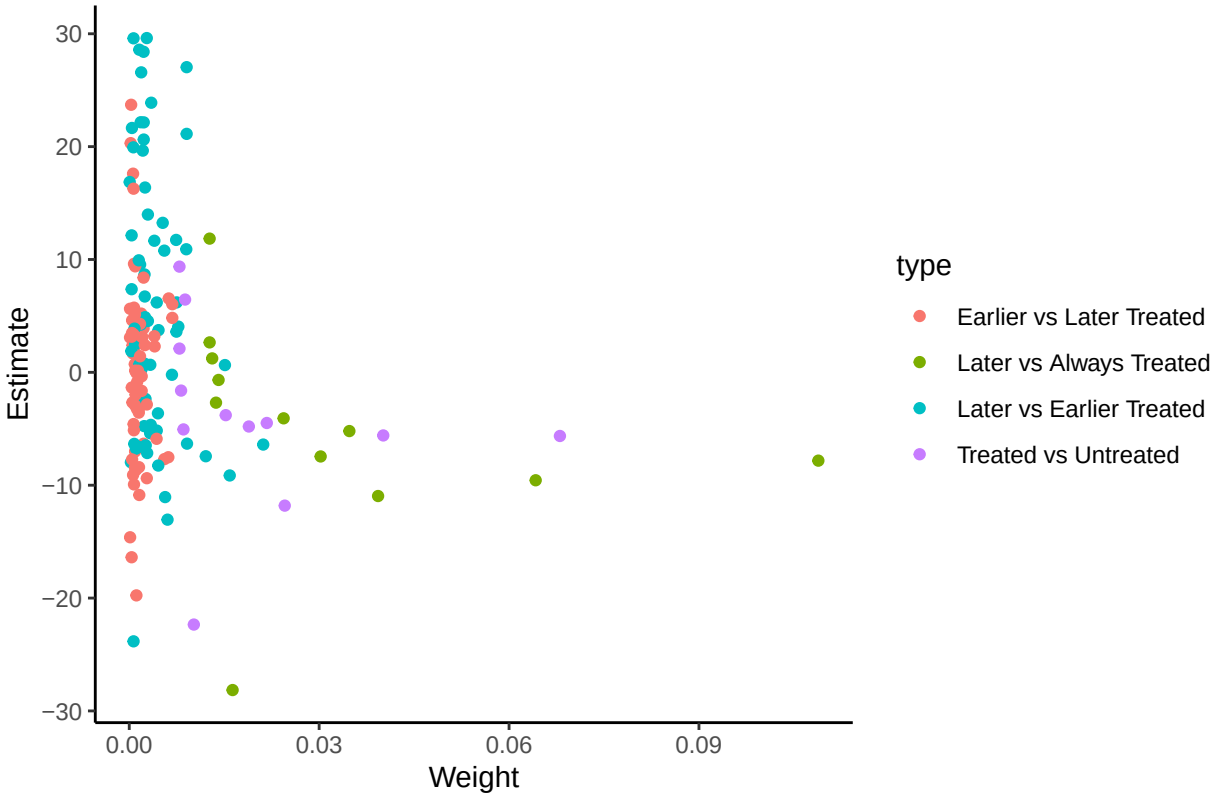
Where i represents the Bacon Decomposition groups listed previously. Summing the weighted averages together gives us an estimate of -1.666. This does not alight with our previous DD estimates for the suicide

Table 3: Bacon Decomposition Weights and Estimates

Type	Weight	Average Estimate
Earlier vs. Later Treated	0.11065	-0.18677
Later vs. Always Treated	0.38443	-7.04369
Later vs. Earlier Treated	0.26464	3.51197
Treated vs. Untreated	0.24027	0.533091

rate, which was -2.75 for the model with state-clustered standard errors and trend interacted with state fixed effects. For this to have happened, there must have been some modelling error on either side of this estimate.

Figure 2: Bacon Decomposition Weights and Estimates



Conclusion

There are two situations in which the static DD parameter is a biased estimate of the true ATT: If ΔATT $\neq 0$, meaning that the average treatment effect changes over either location or time; or if VWCT $\neq 0$, which means that we didn't have common trends that balanced out in the pre-treatment period. We can and did test for ΔATT $= 0$, but we can never test for VWCT $= 0$ because it involves counterfactuals to reality. The VWCT assumption is not testable because it is based on Y^0 for treatment groups in the post-treatment period, but we only have pre-treatment data for those. Normally in order to justify this assumption, you check to see if treatment groups pre-treatment are approximately the same as the control group and if they are we can reasonably believe this assumption. However, through the analysis in this replication, we saw that leads in the two years before adopting unilateral divorce laws were significantly different between treatment states and control states, meaning we cannot reasonably assume that VWCT $= 0$.

This is very concerning as it means we will likely experience bias in our estimates that we cannot account for, and therefore not be able to report true causal effects. In their project, Stevenson and Wolfers found significant evidence that unilateral divorce laws led to lower rates of suicide, intimate homicide, and domestic violence in women in the U.S. In this replication however, we found no significant evidence of this effect. The data used for this replication was not as expansive as the data used in the original paper, which included case-by-case demographics details that could not be controlled for using the aggregated data supplied for this replication. In any case, while future replications could still justifiably find significant evidence that unilateral divorce laws reduce the negative outcomes of marriages, we could not in this replication through utilization of the Bacon Decomposition.