To be, or not to be:
Extending Difference-in-Differences to Binary Outcomes

Alonso M. Guerrero Castañeda

Department of Statistical Science, Duke University

April 5th, 2024

Abstract

The traditional difference-in-difference (DiD) set-up assumes continuous outcomes and usually employs linear regression to estimate the average treatment effect on the treated. We extend the DiD set-up to binary outcomes and propose an alternative estimator based on logistic regression. We use simulations to compare the relative performance of the traditional linear regression DiD estimator and this new proposed logistic-based DiD estimator. We show that both estimators behave similarly in terms of bias and empirical coverage when the outcome's probability of success is around 50%. However, the former becomes biased outside that 50% vicinity, while the latter remains unbiased and maintains good coverage.

1 Introduction

Difference-in-differences (DiD) is a widely used tool in econometrics and causal inference for addressing confounding in observational studies (Roth et al., 2023)[6]. While DiD typically assumes continuous outcomes and employs linear regression, applying this method to binary outcomes may introduce bias, especially with rare outcomes (Strumpf et al., 2017)[7]. This project aims to adapt DiD for settings with binary outcomes by proposing an alternative estimation method based on logistic regression. By simulating data and comparing both estimation methods in terms of bias and coverage, we seek to provide insights into their relative performance.

This research builds upon the work of Li & Li (2019)[5], who extended DiD estimation to count data and conducted simulations. We follow their approach, adapting it to our specific context of binary outcomes. Our proposed estimator, based on logistic regression, also resembles the one formulated in Lechner (2010) [4] for probit regression.

We find that when the proportion of the outcome is around 50%, bias and coverage are similar for both the linear regression approach and the proposed logistic regression approach. However, as one gets away from the 50%, the logistic regression approach remains unbiased while the linear regression approach increases in bias and decreases in empirical coverage. This is in line with well-known behavior of linear probability models as compared to logistic probability models (Strumpf et al., 2017)[7].

2 Literature Review

Difference-in-differences (DiD) has been enjoying a methodological revival in the last few years. Roth et al. (2023)[6] synthesized recent advances in the econometrics of DiD. They identify three strands of DiD literature: (1) settings where there are more than two time periods and units are treated at different points in times, (2) potential violations to the parallel trends assumptions, and (3) alternative sampling assumptions. For example, Callaway and Sant'Anna (2021)[1] extended DiD to settings with multiple time periods and, as of March 2024, has been cited almost 4,200 times.

In this paper we focus on a different issue: DiD in the context of binary outcomes. This is part of a larger discussion on non-linear DiD. This is an important subject matter given the prominence of binary outcomes in the field of economics, psychology, health, among others. Yet, empirical applications of DiD tend to rely heavily on linear probability model instead of logistic models when it comes to binary data (Strumpf et al., 2017)[7]. Gomila (2021)[3] goes as far as to generally advice against the use of logit or probit for binary outcomes in the field of psychology, given the direct causal interpretability of linear regression.

Lechner (2010)[4] acknowledges that nonlinear models pose a challenge for the traditional DiD set-up. Simply running a generalized two-way fixed effects model does not recover any meaningful causal estimate. He points out there are two approaches one could take: (1) using nonlinear models with standard common trend assumption, or (2) using nonlinear models with a modified common trend assumption. We will follow the former. He presents both a general form and application for the case of probit model (rather than logit as we do). He does admit that, as of the time of publishing, "there seems to be no applied literature using such a specification in this way" (pg.200)[4]. Li and Li (2019)[5] develop a generalized framework for double-robust DiD estimators in the context of count outcome data. We have partially adapted their approach and propose a estimator specifically designed for binary outcome data. Our proposed estimator matches the general form presented in both Lechner (2010) and Li and Li (2019).

More recently, Wooldridge (2023)[8] has proposed flexible strategies for DiD settings where the nature of the response variable may warrant a nonlinear model. His approach to the issue of binary outcomes specifically was to modify the parallel trends assumption and allow for a non-linear version. This is similar to Lechner's aforementioned second approach involving a modified trend assumption. Unlike Wooldridge (2023), here we do not relax the linear parallel trend assumption, but rather, propose a new estimator based on the findings of Lechner (2010) and Li and Li (2019).

3 Methodology

We adopt the framework outlined by Li and Li (2019)[5] with modifications for binary outcomes. Consider two periods (t and $t+1$), denoting before and after treatment. The treatment is only administered to one group ($G_i = 1$) in the after period. We denote $\theta_1 = \mathbb{E}[Y_{i,t+1}(1)|G_i = 1]$ (expected value of the outcome for group 1 under treatment at $t+1$) and $\theta_0 = \mathbb{E}[Y_{i,t+1}(0)|G_i = 1]$ (the expected counterfactual outcome of group 1 in the absence of treatment at $t+1$). We are interested in estimating the average treatment effect on the treated (ATT), which can be written as $\tau = \theta_1 - \theta_0$.

Given that we do observe Group 1 being treated, under the stable unit treatment value assumption (SUTVA), θ_1 is non-parametrically identified as $\theta_1 = \mathbb{E}[Y_{i,t+1}|G_i = 1]$, with a consistent estimator $\hat{\theta}_1 = \frac{\sum G_i Y_{i,t+1}}{\sum G_i}$.

On the other hand, θ_0 is more challenging since it is not directly observed. We need the parallel trend assumption to go on:

$$\mathbb{E}[Y_{i,t+1}(0) - Y_{i,t}(0)|X_i, G_i = 1] = \mathbb{E}[Y_{i,t+1}(0) - Y_{i,t}(0)|X_i, G_i = 0]$$

Applying the law of total of expectations, we can rewrite θ_0 as $\mathbb{E}[Y_{i,t}|G_i = 1] + \mathbb{E}_X[\mathbb{E}[Y_{i,t+1}(0) - Y_{i,t}(0)|X_i, G_i = 0]|G_i = 1]$. The first term (which is directly observed) can be consistently estimated with $\frac{\sum_i G_i Y_{i,t}}{\sum_i G_i}$. The second term requires a model specification. In our case, we are dealing with binary outcomes, so we will assume a Bernoulli model. This is a modification from the original Li and Li (2019)[5] methodology where the authors use negative binomial for count outcomes.

$$Y_{i,t}(0)|X_i, G_i = 0 \sim \text{Bernoulli}\left(\mu = \frac{1}{1 + \exp(-X'\beta)}\right) \tag{1}$$

$$Y_{i,t+1}(0)|X_i, G_i = 0 \sim \text{Bernoulli}\left(\nu = \frac{1}{1 + \exp(-X'\gamma)}\right) \tag{2}$$

The MLE estimates for $\hat{\beta}$ and $\hat{\gamma}$ do not have a closed form solution, so they are estimated numerically via logistic regression. Under SUTVA and the parallel trends assumption,

$$\hat{\theta}_0^{reg} = \frac{\sum_i G_i Y_{i,t}}{\sum_i G_i} + \frac{\sum_i G_i \{\nu(X_i, \hat{\gamma}) - \mu(X_i, \hat{\beta})\}}{\sum_i G_i}$$

If the model is correctly specified, $\hat{\theta}_0^{reg}$ is a consistent estimator for θ_0. And so, $\hat{\tau}^{reg} = \hat{\theta}_1 - \hat{\theta}_0^{reg}$ is also a consistent estimator.

4 Simulations and Results

To illustrate the performance of our proposed estimator against the performance of a regular linear regression DiD estimator, we conduct simulations under a two-period two-group design. Each simulation has $N = 1,000$ units. Following Li and Li (2019)[5], each unit has a binary covariate X_1 and a continuous covariate X_2, generated as follows:

$$X_1 \sim \text{Bernoulli}(0.25), \quad X_2|X_1 \sim \text{Normal}(2 + 6X_1, 2^2)$$

We simulate the treatment group label G_i from a Bernoulli distribution with success probability being the propensity score:

$$\text{logit}\{e(\mathbf{X})\} = -2 + X_1 - 0.2X_2 + 0.04X_2^2$$

We now need to simulate the binary outcomes from Bernoulli models. In order to better understand patterns in bias with regards to the "rarity" of the outcome, we will assume the following:

$$\begin{aligned} Y_t(0)\,|\mathbf{X}, G = 0 &\sim \text{Bernoulli}\left(\mu_{00}(\mathbf{X})\right), \quad Y_t(0)|\,\mathbf{X}, G = 1 \sim \text{Bernoulli}\left(\mu_{01}(\mathbf{X})\right), \\ Y_{t+1}(0)\,|\mathbf{X}, G = 0 &\sim \text{Bernoulli}\left(\nu_{00}(\mathbf{X})\right), \quad Y_{t+1}(1)|\,\mathbf{X}, G = 1 \sim \text{Bernoulli}\left(\nu_{11}(\mathbf{X})\right), \end{aligned}$$

with

$$\begin{aligned} \mu_{00}(\mathbf{X}) &= \text{expit}\left(k + 0.5X_1 + 0.05X_2\right) \\ \mu_{01}(\mathbf{X}) &= \text{expit}\left(k - 0.5 + 0.5X_1 + 0.05X_2\right) \\ \nu_{00}(\mathbf{X}) &= \text{expit}\left(k - 0.5 + 0.5X_1 + 0.05X_2\right) \\ \nu_{11}(\mathbf{X}) &= \text{expit}\left(k + 0.5X_1 + 0.05X_2\right) \end{aligned}$$

where we try values of k between -9.0 and 3.0. The smaller k is, the rarer the outcome is. Under the parallel trend, the mean function of the counterfactual binary outcome for the treated observations is $\nu_{01}(\mathbf{X}) = \nu_{00}(\mathbf{X}) + \mu_{01}(\mathbf{X}) - \mu_{00}(\mathbf{X})$.

We simulate 200 replicates based on the models specified above (a batch of 200 per each value of k). To compute the ATT, τ, we use the following two outcome approaches[1]:

Linear regression approach: we consider running the standard additive fixed effects regression model, where we include group assignment, period dummies, their interactions and covariates:

$$Y_{it} = \alpha + \gamma G_i + \delta_T + \tau G_i \mathbb{1}\{T = t + 1\} + \beta \mathbf{X_i} + \epsilon_{iT}$$

with $\epsilon_{iT}|G_i, T, X_i \sim N(0, \sigma^2)$. Thus, $\hat{\tau}^{lm}$ and its 95% confidence interval (CI) are directly estimated from running the linear regression based on the model above.

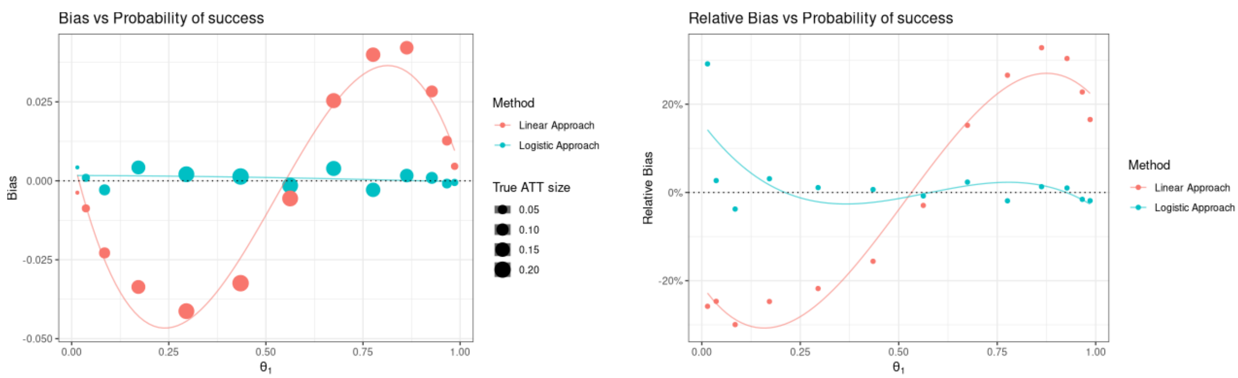
Logistic regression approach: we use our proposed estimator with correctly specified mean functions $\mu(\mathbf{X})$ and $\nu(\mathbf{X})$, where each is estimated using logistic regression:

$$\hat{\theta}_0^{reg} = \frac{\sum_i G_i Y_{i,t}}{\sum_i G_i} + \frac{\sum_i G_i\{\nu(X_i, \hat{\gamma}) - \mu(X_i, \hat{\beta})\}}{\sum_i G_i}$$

Once again, $\hat{\theta}_1$ is directly observed. Thus, our resulting ATT estimate is $\hat{\tau}^{reg} = \hat{\theta}_1 - \hat{\theta}_0^{reg}$. We will use the nonparametric bootstrap (Efron & Tibshirani, 1993)[2] to obtain the associated 95% CI.

Figure 1 and 2 summarize the results of our simulations. The linear regression approach performs relatively well in terms of bias and coverage of the 95% confidence interval when θ_1, the probability of

[1] Outcome approach refers to the fact that we are imposing parametrical assumptions on the outcome. Li and Li (2019) also consider the propensity score weighting approach (imposing parametrical assumption on the propensity score model) and the double-robust approach (the coming together of both outcome and propensity score weighting approaches). It is worth noting that there would not be a need for any modelling at all if the parallel assumption were to hold unconditional of pre-treatment covariates $\mathbf{X}$.



(a) Bias results

(b) Relative Bias results

Figure 1: Comparison of bias and relative bias results for both linear and logistic regression approaches.

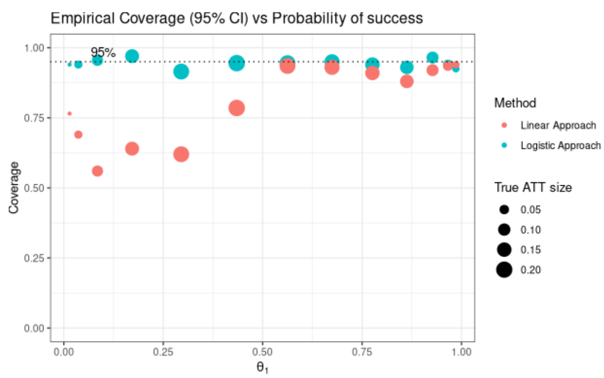


Figure 2: Empirical coverage of the 95% confidence intervals for both linear and logistic regression approaches.

success of the treatment group in period $t+1$, is close to 55%. It tends to under-estimate the true effect the most around 25% and it tends to over-estimate the true effect the most around 80%. Results indicate that the logistic regression approach maintains stability in bias and coverage across various scenarios, particularly when the probability of success in the treatment group is not extreme. However, limitations are observed at the extreme ends of the outcome probability spectrum. This is mostly because the probability of simulating either all failures or all successes is fairly high. We would need to a larger sample size for these edge cases to avoid this situation.

5 Conclusion

To conclude, in this study, we proposed an extension of the Difference-in-Differences (DiD) estimator to handle binary outcomes, addressing a limitation of the traditional approach which assumes continuous outcomes. Following Li and Li (2019)[5], we adapted the methodology by incorporating logistic regression to estimate treatment effects, comparing it with the standard linear regression approach. Through simulations, we demonstrated that the logistic regression approach maintains stability in bias and coverage across various scenarios, particularly when the probability of success in the treatment group is not extreme.

In summary, our findings underscore the importance of considering appropriate estimation techniques when applying DiD to settings with binary outcomes. The proposed logistic regression approach offers a robust alternative, since the standard linear regression approach suffers from bias particularly when the probability of success is not around 50%. Future research could explore further refinements, such as extending this solely outcome regression approach to a double robust estimation approach, as Li and Li (2019)[5] did for count outcomes, but for binary outcomes.

6 Reproducibility

Code and data are available in the following repository: https://github.com/alonsomgc/DiD-BinaryOutcome/.

References

[1] Brantly Callaway and Pedro H.C. Sant'Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230, 2021. Themed Issue: Treatment Effect 1.

[2] Bradley Efron and Robert J. Tibshirani. *An Introduction to the Bootstrap*. Chapman Hall, 1993.

[3] Robin Gomila. Logistic or linear? estimating causal effects of experimental treatments on binary outcomes using regression analysis. *Journal of Experimental Psychology: General*, 150(4):700–709, 2021.

[4] Micael Lechner. *The Estimation of Causal Effects by Difference-in-Difference Methods*, volume 4. Now Publishers Inc, 2011.

[5] Fan Li and Fan Li. Double-robust estimation in difference-in-differences with an application to traffic safety evaluation, 2019.

[6] Jonathan Roth, Pedro H.C. Sant'Anna, Alyssa Bilinski, and John Poe. What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2):2218–2244, 2023.

[7] Erin C Strumpf, Sam Harper, Jay S Kaufman, and Michael Oakes. Fixed effects and difference in differences. *Methods in social epidemiology*, 1, 2017.

[8] Jeffrey M Wooldridge. Simple approaches to nonlinear difference-in-differences with panel data. *The Econometrics Journal*, 26(3):C31–C66, 08 2023.