

A simple approximation to the average effect of the treatment on the treated in panel settings with selective enrollment

John Gardner*

May 2017

Abstract

This paper shows that in panel settings when the probability of receiving a treatment is low, simple comparisons of outcomes between those who are treated in a given period and those who are untreated in that period, but receive the treatment in some other period, approximately identify the average effect of the treatment on the treated. Monte Carlo simulations reveal that this approximate selection-correction substantially reduces selection bias compared to naive comparisons of outcomes between treated and untreated units.

Keywords: Causal effect, treatment effect, selection bias.

JEL Codes: C1, C2, C13, C14.

1 Introduction

One of the central concerns of empirical economics is the possibility that self-selection on unobserved characteristics introduces bias into comparisons between treated and untreated individuals, invalidating the causal interpretation of those comparisons. This paper proposes a simple selection-correction that can be used to approximate the Average Effect of the Treatment on the Treated (ATT) in panel settings where the probability of receiving the treatment is low. The intuition for this approach is that, when the probability of receiving the treatment is low, individuals who are treated in any period are likely to differ systematically from those that never are. Consequently, comparing outcomes for those who are treated in a given period with those who are untreated in that period, but receive the treatment in some other period, approximates the average effect of the treatment among the treated population.

To be clear at the outset, this approach can only approximate the ATT, and can only do so well in panel settings where the probability of receiving the treatment is small (though it may be a useful indicator of the direction and severity of selection bias even at moderate treatment rates). However, in applications it is often the case that relatively few individuals receive the treatment in any given period. Moreover, the approximate selection-correction approach is simple and intuitive,

*366 Holman Hall, P.O. Box 1848, Department of Economics, University, MS 38677. Email: jrgardne@olemiss.edu.

and can be applied absent an instrument or other source of exogenous variation in treatment status, as well as in settings where the effect of unobserved characteristics on outcomes varies over time (in which traditional fixed-effects methods are invalid).

In the remainder of this paper, I provide theoretical motivation for the approximate selection-correction, illustrate its properties using a series of Monte Carlo simulations, demonstrate its application by using it to approximate the effect of Northward migration on Southern-born blacks' educational attainment, and conclude.

2 Approximating the ATT

Let Y_{1it} represent the outcome that individual $i \in \{1, \dots, N\}$ would experience at time $t \in \{1, \dots, T\}$ if they received a binary treatment ($D_{it} = 1$), and Y_{0it} the outcome they would experience absent the treatment (Rubin, 1974). Using this notation, i 's observed outcome can be expressed $Y_{it} = D_{it}Y_{1it} + (1 - D_{it})Y_{0it}$, and the time- t Average Effect of the Treatment on the Treated (ATT) can be expressed $ATT_t = E(Y_{1it} - Y_{0it} | D_{it} = 1)$. It is well known that when treatment status D_{it} is not randomly assigned, comparisons of observed outcomes between treated and untreated individuals identify a combination of the ATT and a bias term reflecting the possibility that treated individuals might have experienced systematically different outcomes even absent the treatment.

Suppose that potential outcomes (Y_{0it}, Y_{1it}) are independent of treatment status D_{it} conditional on an unobserved time-invariant characteristic U_i .¹ Suppose also that successive realizations of D_{it} are independent conditional on U_i , so that the U_i explain any serial correlation in treatment status. In this case, conditional on U_i , the bias term vanishes, and treated-untreated outcome comparisons identify U -specific causal effects:

$$\begin{aligned} E(Y_{it} | D_{it} = 1, U_i) - E(Y_{it} | D_{it} = 0, U_i) &= \underbrace{E(Y_{1t} - Y_{0t} | D_{it} = 1, U_i)}_{\text{ATT}} \\ &+ \underbrace{[E(Y_{0t} | D_{it} = 1, U_i) - E(Y_{0t} | D_{it} = 0, U_i)]}_{\text{Bias}} = E(Y_{1t} - Y_{0t} | D_{it} = 1, U_i). \end{aligned}$$

These strata-specific causal effects, in turn, can be aggregated according to the distribution of U_i among the treated population to identify the overall ATT. The argument that follows can be applied within covariate strata if potential outcomes are independent of treatment status conditional on U_i and observed covariates X_{it} .

This approach to identification is infeasible when U_i is unobserved. However, consider approximating the U_i by a binary variable $Q_i \in \{l, h\}$ with mass function $P(Q_i = h) = \pi$ and satisfying the restriction that $p_{lt} = P(D_{it} = 1 | Q_i = l) = 0$ for all t . In other words, approximate the unobserved heterogeneity by assuming that there are two types of individuals: those that will never receive the

¹The time-invariance assumption is for simplicity and can be relaxed.

treatment and those that might.²

Under this approximation, some individuals can be identified as members of group h with probability one. To see this, let $D_i = 1 \left(\sum_{t=1}^T D_{it} > 0 \right)$ be an indicator for whether individual i ever receives the treatment, and let $p_{ht} = P(D_{it} = 1 | Q_i = h)$ denote the probability that an individual belonging to group h receives the treatment at time t . The probability that an individual who ever receives the treatment belongs to group h is therefore simply

$$P(Q_i = h | D_i = 1) = \frac{\pi \left[1 - \prod_{t=1}^T (1 - p_{ht}) \right]}{\pi \left[1 - \prod_{t=1}^T (1 - p_{ht}) \right] + (1 - \pi) \left[1 - \prod_{t=1}^T (1 - p_{lt}) \right]} = 1.$$

Furthermore, the average effect of the treatment on the treated can be approximated as

$$\begin{aligned} ATT_t &= \int E(Y_{1it} - Y_{0it} | D_{it} = 1, U_i = u) dP(U_i = u | D_{it} = 1) \\ &= \int [E(Y_{it} | D_{it} = 1, U_i = u) - E(Y_{it} | D_{it} = 0, U_i = u)] dP(U_i = u | D_{it} = 1) \\ &\approx \sum_{q \in \{l, h\}} [E(Y_{it} | D_{it} = 1, Q_i = q) - E(Y_{it} | D_{it} = 0, Q_i = q)] P(Q_i = q | D_{it} = 1) \\ &= E(Y_{it} | D_{it} = 1, D_i = 1) - E(Y_{it} | D_{it} = 0, D_i = 1). \end{aligned}$$

In summary, under the two-type approximation described above, the time- t ATT is the same as the average effect of the treatment for group h , and can therefore be identified by comparing treated and untreated outcomes at time t , among those who receive the treatment in *any* period. Of course, the usefulness of this approach to identification depends on the quality of the approximation that it provides to the underlying ATT, which may be poor if the effects of, and likelihood of receiving, the treatment are very heterogeneous.³ However, the lower the probabilities $P(D_{it} = 1)$, $t \in \{1, \dots, T\}$, of receiving the treatment are, the better this approximation will be. This is because when there is meaningful unobserved heterogeneity in treatment probabilities but the overall population probability is low, it must be the case that a large proportion of the population receives

²In the terminology of Angrist, Imbens and Rubin (1996), this approximation separates the never-takers from the compliers, defiers, and always-takers.

³This approximate selection-correction approach can also be viewed as an approximation to a formal latent-class analysis, in which individuals are assumed to belong to one of a finite number of latent classes, the latent-class distribution and class-specific treatment probabilities are chosen to maximize the sample log-likelihood, the number of classes is chosen to maximize either the Bayesian or Akaike Information Criterion, and individuals are assigned to latent classes according to their probabilities of membership in each class, conditional on their observed characteristics (see Leisch, 2004, for an introduction). At lower treatment rates, the approximate and formal latent-class methods produce similar estimates (as unreported simulation results show).

Though latent-class analysis can be applied regardless of the treatment probabilities, when those probabilities are high, latent-class membership is harder to identify, introducing classification error into treatment effect estimates obtained by comparing treated and untreated units assigned to the same classes (this problem can be solved by modeling observed outcomes along with the treatment decisions, or by correcting for the resulting classification errors; see Bolck, Croon and Hagenaars 2004). Similarly, the approximate method cannot reliably identify low types (and hence the average effect $ATE = E(Y_{1it} - Y_{0it})$ of the treatment), since those who are never treated may belong to either group.

the treatment with near-zero probability.⁴ Although this approximation argument is necessarily heuristic, I illustrate it with evidence from Monte Carlo simulations in the next section.

3 Monte Carlo simulations

I perform a number of Monte Carlo simulations to illustrate the properties of the approximation developed above. In each simulation, the data are generated by the following processes:

$$\begin{aligned} Y_{0it} &= bX_{it} + U_i + e_{0it} & Y_{1it} &= Y_{0it} + bX_{it} + U_i + e_{1it} \\ D_{it} &= 1(Y_{1it} - Y_{0it} > c) & Y_{it} &= D_{it}Y_{1it} + (1 - D_{it})Y_{0it}. \end{aligned}$$

I chose these processes to introduce selection bias into naive comparisons between treated and untreated individuals. In each simulation, X_{it} , e_{0it} , and e_{1it} are drawn from (independent) standard normal distributions. I vary the distribution from which the U_i are drawn, the coefficient b on the observed covariate X_{it} , and the enrollment threshold c , as described below.

For each configuration of the parameters, I simulate 1000 datasets consisting of observations on $\{Y_{0it}, Y_{1it}, Y_{it}, D_{it}, X_{it}\}$ for $i \in \{1, \dots, 1000\}$ and $t \in \{1, \dots, 5\}$. For each simulate, I compute the sample average ATT as the sample analog of $E(Y_{1it} - Y_{0it} | D_{it} = 1)$, a naive estimate of the ATT as the coefficient on D_{it} from a regression of Y_{it} on D_{it} and X_{it} , and an approximately selection-corrected estimate of the ATT as the coefficient on D_{it} from the same regression estimated using the sample of individuals who ever receive the treatment. Note that the first of these objects can only be computed in a simulation setting where both treated and untreated outcomes are observed. Because the environment is stationary, I only report results for the final period. I summarize the simulation results graphically in Figure 1 and numerically in Table 1.⁵

For the first four simulations, I set the enrollment threshold c equal to 2. As the row labelled “Rate” in Table 1 shows, this generates population-average treatment probabilities of about 10% (these refer to the probability of being treated in the fifth period, not the probability of being treated in any period). In simulation (1), I draw the U_i from a standard normal distribution and set the coefficient b on the observed covariate X_{it} to .1. The results are summarized in the first panel of Figure 1 and the first column of Table 1. A naive regression of observed outcomes on treatment status and the observed covariate overstates the average effect of the treatment on the treated considerably. In contrast, the approximately selection-corrected estimates are much closer to the true ATT, with the correction eliminating an average of about 73% $[=(4.07-3.03)/(4.07-2.64)]$ of the bias present in the naive estimate.

⁴E.g., consider a two-type model where a fraction π of the population enroll with probability p_h and $(1 - \pi)$ enroll with probability $p_l < p_h$. In broad terms, a small population probability $p = P(D_{it} = 1) = \pi p_h + (1 - \pi)p_l$ implies that either $p_l \approx 0$ and $p_h \approx p/\pi > p$ or $p_l \approx p_h \approx p$. In the latter case, unobserved heterogeneity would not introduce much selection bias, however, so only the former case is of interest here.

⁵In the plots in Figure 1, the bold lines represent the medians, the boxes represent the interquartile range, and the “whiskers” represent the minimum and maximum estimates (with outliers excluded according to R’s default algorithm).

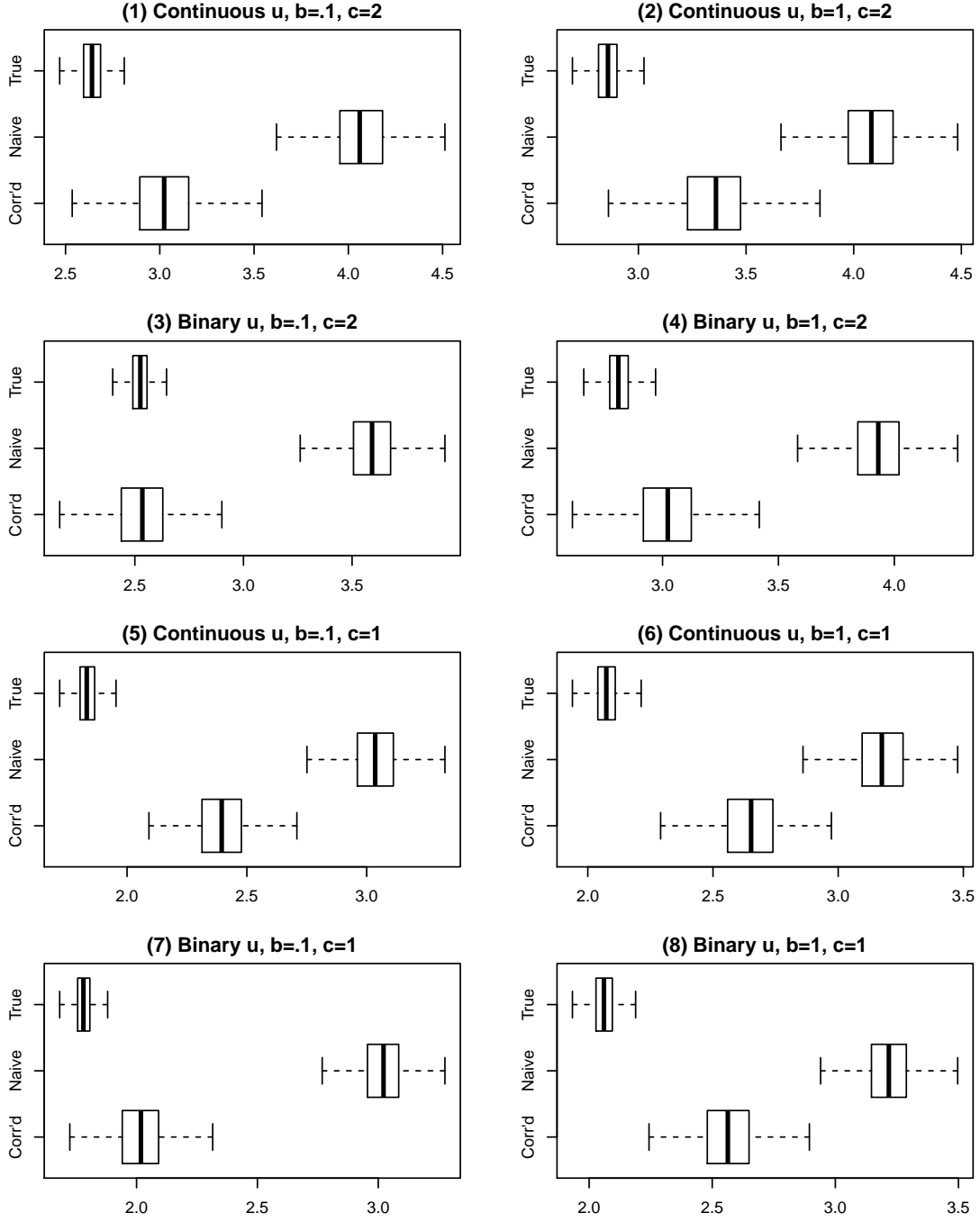


Figure 1: Distributions of ATTs from 1000 simulations. “True” refers the the actual ATT, “Naive” the coefficient from a regression of observed outcomes on treatment status and the observed covariate X , and “Cor’d” the same regression among the sample of those who are ever treated. Estimates are for the last of five time periods. “ c ” denotes the assumed threshold c at which individuals enroll. “ b ” denotes the coefficient on the observed covariate X .

Table 1: Simulation summaries

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
True	2.64 (0.07)	2.86 (0.07)	2.53 (0.05)	2.81 (0.06)	1.83 (0.04)	2.08 (0.05)	1.78 (0.04)	2.06 (0.05)
Naive	4.07 (0.17)	4.08 (0.16)	3.59 (0.13)	3.93 (0.13)	3.04 (0.11)	3.18 (0.12)	3.02 (0.09)	3.22 (0.1)
Corrected	3.03 (0.19)	3.35 (0.18)	2.53 (0.14)	3.02 (0.15)	2.39 (0.12)	2.65 (0.13)	2.02 (0.11)	2.57 (0.12)
R sq.	0 (0)	0.13 (0.02)	0 (0)	0.14 (0.02)	0 (0)	0.19 (0.02)	0 (0)	0.18 (0.02)
Rate	0.08 (0.01)	0.12 (0.01)	0.08 (0.01)	0.13 (0.01)	0.24 (0.01)	0.28 (0.02)	0.26 (0.01)	0.29 (0.01)
Threshold	2	2	2	2	1	1	1	1
Het. type	Cts	Cts	Binary	Binary	Cts	Cts	Binary	Binary
b	0.1	1	0.1	1	0.1	1	0.1	1

Notes—Means and standard deviations from 1000 simulations. All table entries refer to the last of five time periods. “True” denotes the average actual ATT. “Naive” denotes the average coefficient from a regression of observed outcomes on treatment status and the observable covariate X . “Corrected” denotes the average coefficient from a regression of observed outcomes and treatment status and the observed covariate X among the subsample of units that ever receive the treatment. “Rate” denotes the average probability of receiving the treatment. “R sq.” denotes the average R-squared from a regression of treatment status on X . “Threshold” denotes the assumed threshold c at which individuals enroll. “b” denotes the coefficient on the observed covariate X . “Het type” is “Cts” when U is normally distributed and “Binary” when U is drawn from a binary distribution.

To illustrate the impact of *observed* heterogeneity on the quality of the approximation, in simulation (2) I set $b = 1$. As the row labelled “R sq.” in Table 1 shows, this increases the average R^2 from a regression of D_{it} on X_{it} from nearly zero to about .13. This additional source of variation in the probability of enrollment decreases the quality of the approximate selection-correction, which now eliminates an average of about 60% of the bias in the naive regression estimates. This pattern suggests that, when observed covariates matter, the approximate selection-correction should be applied within observed strata.

For simulations (3) and (4), I repeat the analysis of simulations (1) and (2), drawing the $U_i \in \{-1, 1\}$ from an equiprobable binary distribution. The results show that the two-type selection correction provides a much better approximation to the true ATT when there is less variation in the underlying unobserved heterogeneity. Although the empiricist is unlikely to have prior information about such heterogeneity, these results illustrate the theoretical underpinnings of the approximation approach developed here.

To show how the quality of the approximate selection-correction varies with the probability of receiving the treatment, in simulations (5)-(8) I recreate simulations (1)-(4) with a lower enrollment threshold of $c = 1$, generating average treatment probabilities of about 25%. Consistent with the reasoning above, the performance of the approximate selection-correction declines as the probability of enrollment increases. In simulation (5), for example, the approximate correction only eliminates about half of the bias from the naive regression. Nevertheless, simulations (5)-(8) suggest that the approximate correction may be useful even at moderate treatment probabilities as an indicator of the direction and potential severity of bias due to selection on unobservables.

4 Application: The effect of the Great Migration on educational attainment

The African American Great Migration, during which millions of blacks left their birthplaces in the Southern US in favor of cities in the North, is thought to have had wide-reaching consequences for black-white relative economic progress in the 20th-century (see Tolnay, 2003, for a review). However, the possibility that the Great Migrants were a select group that might have also experienced better socioeconomic outcomes in the South calls the casual interpretation of comparisons between migrants and non-migrants into question. To illustrate its application, I use the approximate selection-correction to estimate the effect of Northward migration on educational attainment. My analysis draws on the Three-Generation Survey of Black American Families (Jackson and Tucker, 1997), an intergenerational dataset that reports socioeconomic outcomes, state of birth, and state of residence when surveyed for three generations of a number of black families.⁶

The method developed in this paper is appropriate for this question for several reasons. First,

⁶I exclude observations for which the first-generation member of the family was born in a Northern state. Throughout my analysis, I use the Census Bureau’s definition of the North and South. To account for the possibility that Southern-born members of the third generation (who are the youngest in the sample) might have migrated North after they were interviewed, I exclude third-generation respondents who were 18 or younger at the time of the survey.

to the extent that family-specific unobserved characteristics affect the educational attainment of parents and their children, they likely do so differently over time, violating the assumptions of fixed-effects methods. Second, the fractions of Southern-born individuals who migrate North are relatively small, at about 11% for the first generation sample, 18% for the second generation, and 12% for the third, suggesting that the correction will approximate the ATT well. Third, there are few clearly predetermined explanatory variables and no obviously exogenous sources of variation in migration with which to implement traditional regression or instrumental variables methods.

Table 2 presents the results of the analysis. The first column reports results from a naive regression of educational attainment in years on indicators for living in the North and being male for Southern-born members of the first generation. The coefficient on North implies an estimate of the ATT of about 2.14 years. The second column reports approximately selection-corrected results from the same regression, estimated among first-generation members of families who ever leave the South. Consistent with the notion that migrants are positively selected on education, the implied ATT is lower, at about 1.74. Columns (3) and (4) repeat this exercise for the second generation. The naive estimate is 1.44, while the corrected estimate is a statistically insignificant 1.02, again suggesting positive selection.⁷ For comparison, I also report in column (5) fixed-effects estimates from a pooled sample of all three generations. Contrary to intuition, the implied ATT of 3.76 exceeds those from the naive regressions, suggesting that a fixed-effects specification is inappropriate in this intergenerational setting.

Table 2: Northward migration and education

	Gen 1	Gen 1 (Corr'd)	Gen 2	Gen 2 (Corr'd)	FE
Intercept	6.73*** (0.21)	7.22*** (0.36)	10.34*** (0.23)	10.60*** (0.50)	9.41*** (2.14)
North	2.14*** (0.55)	1.74** (0.57)	1.44*** (0.39)	1.02 (0.54)	3.76*** (0.44)
Male	-0.78* (0.38)	-1.04 (0.54)	-0.76* (0.38)	-0.18 (0.53)	0.41 (0.30)
R ²	0.05	0.09	0.05	0.03	0.37
Num. obs.	348	135	348	135	1044

Notes—Regressions of educational attainment in years on sex and an indicator for Northward migration for two generations of Southern-born black families. “Corr’d” denotes estimates obtained using the sample of families that ever migrate to the North. “FE” denotes a regression with family fixed effects. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

⁷Because North-South migration is rare in the data, there are almost no untreated members of the third generation among families who ever leave the South, making it impossible to repeat this analysis for the third generation.

5 Conclusion

This paper shows that in panel settings when enrollment is selective, the ATT can be approximated using outcomes comparisons between treated and untreated units among the population that receives the treatment in any period. Monte Carlo simulations reveal that this approximate selection-correction can substantially reduce the bias present in naive comparisons of outcomes between treated and untreated individuals, particularly when the treatment probability is low, and may be useful as an indicator of the direction and severity of selection bias even when the treatment probability is modest. An application of this method suggests that, while Northward migration increased Southern-born blacks' educational attainment, naive comparisons of education between migrants and non-migrants overstate this effect.

References

- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 2007. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444-455.
- Bolck, Annabel, Marcel Croon, and Jacques Hagenaars. 2004. "Estimating Latent Structure Models with Categorical Variables: One-Step versus Three-Step Estimators." *Political Analysis* 12 (1): 3-27.
- Jackson, James S., and M. Belinda Tucker. 1997. *Three-Generation National Survey of Black American Families, 1979-1981*. ICPSR09288-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].
- Rubin, Donald. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66 (5): 688-701.
- Tolnay, Stewart E. 2003. "The African American 'Great Migration' and Beyond." *Annual Review of Sociology* 29 (1): 209-32.
- Leisch, Friederich. 2004. "FlexMix: A General Framework for Finite Mixture Models and Latent Class Regression in R." *Journal of Statistical Software* 11(1). <https://www.jstatsoft.org/article/view/v011i08>.