ORIGINAL PAPER

Non-random Study Attrition: Assessing Correction Techniques and the Magnitude of Bias in a Longitudinal Study of Reentry from Prison

Meghan M. Mitchell[1] · Chantal Fahmy[2] · Kendra J. Clark[3] · David C. Pyrooz[3]

Abstract

Objectives Longitudinal data offer many advantages to criminological research yet suffer from attrition, namely in the form of sample selection bias. Attrition may undermine reaching valid inferences by introducing systematic differences between the retained and attrited samples. We explored (1) if attrition biases correlates of recidivism, (2) the magnitude of bias, and (3) how well methods of correction account for such bias.

Methods Using data from the LoneStar Project, a representative longitudinal sample of reentering men in Texas, we examined correlates of recidivism using official measures of recidivism under four sample conditions: full sample, listwise deleted sample, multiply imputed sample, and two-stage corrected sample. We compare and contrast the results regressing rearrest on a range of covariates derived from a pre-release baseline interview across the four sample conditions.

Results Attrition bias was present in 44% of variables and null hypothesis significance tests differed for the correlates of recidivism in the full and retained samples. The bias was substantial, altering effect sizes for recidivism by a factor as large as 1.6. Neither the Heckman correction nor multiple imputation adequately corrected for bias. Instead, results from listwise deletion most closely mirrored the results of the full sample with 89% concordance.

Conclusions It is vital that researchers examine attrition-based selection bias and recognize the implications it has on their data when generating evidence of theoretical, policy, or practical significance. We outline best practices for examining the magnitude of attrition and analyzing longitudinal data affected by sample selection.

Keywords Attrition · Heckman correction · Multiple imputation · Listwise deletion · Sample selection bias

---

✉ Meghan M. Mitchell
  mmitchell@ucf.edu

1 Department of Criminal Justice, University of Central Florida, 12805 Pegasus Dr., Orlando, FL 32816, USA

2 Department of Criminology and Criminal Justice, The University of Texas at San Antonio, San Antonio, USA

3 Department of Sociology, University of Colorado Boulder, Boulder, USA

Introduction

Panel studies rely on interviewing and re-interviewing the same respondents at multiple points in time. Such research designs offer incredible value to the social sciences, including the field of criminology (Liberman 2008). First, panel studies bolster causal inferences because they assure temporal ordering. Second, they yield more robust conclusions than cross-sectional analyses by isolating the relative and cumulative effects of multiple factors over time. Finally, since panel studies permit the study of dynamic change, they set the stage for the emergence of developmental and life-course criminology. In short, longitudinal research in criminology has not only been used to answered difficult questions of the past, but it is critically important to advancing key aims of the field, including theory testing, evaluating policies and practices, and improving research methodology (for critiques of longitudinal research with respect to these aims, see Cartwright 2007; Cullen et al. 2019; Hirschi and Gottfredson 1983; Sampson 2010).

The focus of this study is on a core methodological limitation of longitudinal methods, which has implications for etiological and evaluation research alike: *attrition*. Attrition occurs when respondents drop out permanently from a study or temporarily skip an interview, ultimately introducing missing data. No longitudinal studies are immune to attrition. A loss of respondents could lead to lower statistical power and Type II errors in analyses. More importantly, even a moderate amount of attrition could alter the conclusions of a study by changing the composition or characteristics of the sample (Gustavson et al. 2012; Hansen et al. 1990). This is known as a form of sample selection bias (Campbell and Stanley 1963), which becomes problematic when the source of attrition is non-random, particularly due to variables relevant to outcomes of interest, such as recidivism (Clark et al. 2020). Missing data could have profound implications for research, policy, and practice, making it necessary for researchers to observe, assess, and correct any bias.

In this study, we examine approaches to address sample selection bias, including listwise deletion, multiple imputation, and the Heckman two-step correction.[1] Although the Heckman method has been used extensively in many disciplines, it maintains a rocky history in criminology (Bushway et al. 2007; Heckman 1979). This may be partially due to the complicated and sometimes problematic application of the Heckman two-step correction (Bushway et al. 2007). In other fields, such as econometrics, the Heckman selection model is widely used "often without questioning its validity" (Yu and Gastwirth 2010, p. 8). We apply this method—along with other popular correction methods—to an area of study that has long commanded the interest of researchers, policymakers, and practitioners: recidivism.

Using data from a longitudinal sample of imprisoned people with a baseline pre-release interview and two post-release interviews, our purpose is threefold. First, we explore if attrition biases the correlates of official records of arrest as a form of recidivism. Second, we examine the magnitude of bias for each correlate. Third, we test how well methods of correction account for sample selection bias. To do so, we combine survey and operational data (which we term "metadata") from the LoneStar Project to create four conditions of

---

[1] Other approaches include Manski bounds, Cosslett's selection model, Newey's series estimator, Powell's two-step semiparametric estimator, Robinson's estimator, along with other statistical methods such as xtARGLS, support vector regression, cluster-based estimating, and kernel mean matching (Winship and Mare 1992). Many of these models with limited dependent variables are estimated using LIMDEP's statistical software package.

sample selection: (1) a *full sample condition*, where we have the entire sample of respondents to illustrate what would happen in a perfect world without attrition ($N=791$); (2) a *retained condition*, where we examine the correlates of recidivism with incomplete data, that is, only among the respondents retained in the wave 3 sample, the most common condition since attrition is unavoidable ($n=506$); (3) a *multiply imputed condition* using multiple imputation chained equations ($N=791$); and (4) a *corrected condition* ($N=791$), where we use a Heckman correction to examine statistical corrections for sample selection bias introduced by attrition. By using a measure of official arrest as the outcome variable we have complete data for all respondents which allows us to identify the consequences of sample selection bias on the correlates of recidivism.

Attrition and Sample Selection Bias in Longitudinal Research

Not all longitudinal studies are created equally. Confidence in the findings derived from longitudinal studies is undermined by respondents participating in earlier waves of data collection but failing to complete later, or final waves (Asendorpf et al. 2014). Attrition tends to compound throughout waves. Once a respondent misses a wave, they are at a higher risk of missing subsequent waves (Davidov et al. 2006; Lugtig 2014). This is particularly the case for panel studies with repeated measures gleaned from the same set of individuals over two or more specific time points (Cordray and Polk 1983; Ribisl et al. 1996; Young et al. 2007). As the importance of such data grows in criminology, so too do concerns about attrition.

Attrition is "one of the most important sources of nonsampling error in panel surveys," a problem all longitudinal data collection efforts must reckon with (Gustavson et al. 2012; Lugtig 2014, p. 700). If the presence of attrition is non-random, in that some people have a higher likelihood of dropping out than others, then the potential for bias heightens (Fitzgerald et al. 1998). The sources of attrition may be unavoidable in situations of respondent mortality or refusals, but other sources of bias may be affected by the research design and can be manipulated by researchers (Ribisl et al. 1996; Thornberry et al. 1993). Respondent tracking and retention strategies are commonly used to limit attrition in panel studies, including gathering a contact card, sending postcards, and offering incentives (Bolanos et al. 2012; McLaughlin et al. 2014; Mutti et al. 2014).

The consequences of attrition go beyond the loss of respondents and a corresponding reduction in statistical power. Sample selection bias reflects the fact that those who attrite from the study may have unique characteristics that other respondents do not share as well as reasoning for non-participation (Barry 2005; Claus et al. 2002; David et al. 2013; Fumagalli et al. 2013; Odierna and Schmidt 2009; Ribisl et al. 1996). Consequently, the remaining sample is no longer representative of the original sample, and thus not representative of the population to which inferences could be made (Carkin and Tracy 2015; Crisanti et al. 2014). Random sampling of respondents begins to lose its utility once non-random attrition occurs (Goodman and Blum 1996).

There is much evidence pointing to differences between retained and attrited samples. These include factors such as racial/ethnic minority status (Butler et al. 2013; Magruder et al. 2009; Thornberry et al. 1993), gender (Badawi et al. 1999; Snow et al. 2007; Young et al. 2006), employment status (Claus et al. 2002; Crisanti et al. 2014), educational attainment (Badawi et al. 1999; Lugtig 2014; Thornberry et al. 1993), disease severity and health status (Chatfield et al. 2005; Lugtig 2014; Schmidt et al. 2000; Young et al. 2006), substance use (Crisanti et al. 2014; Snow et al. 1992; Thornberry et al. 1993), and

psychological distress (Chang et al. 2009; Claus et al. 2002). If these factors are not taken into account and corrected, the interpretability and credibility of the results may be undermined (Barry 2005; Chatfield et al. 2005), as both internal and external validity are compromised (Berk 1983; Boys et al. 2003; Crisanti et al. 2014). Overall, data that are missing non-randomly—especially as it relates to the outcome of interest—present a skewed picture of the patterns in criminal justice research, an obstacle to yielding a more complete understanding of crime (Wadsworth and Roberts 2008).

Reentering Populations and Study Attrition

Some study populations of concern to criminologists are more difficult to retain than others. Formerly incarcerated people constitute a classic example of a high-risk, hard-to-reach population because of their transient lifestyles and erratic behavior (Crisanti et al. 2014; Curtis 2010; Western 2018). Many formerly incarcerated individuals recidivate because they are ill-equipped to handle the challenges of reentry (Shinkfield and Graffam 2007). Upon leaving prison, parole violations, arrests, and reincarceration make them much more difficult to track (Coen et al. 1996; Kinner 2006).[2] Moreover, they tend to distrust researchers whom they believe could be working with law enforcement (Eidson et al. 2016; Western et al. 2017). Panel studies of formerly incarcerated people often show high rates of study attrition and nonresponse (Western et al. 2017), even after an in-prison baseline assessment (e.g., Fahmy et al. 2019). As attrition rates in longitudinal panel studies of reentering individuals vary (e.g., *Serious and Violent Offender Reentry Initiative (SVORI)* or *Returning Home Study*), so too do their sample sizes, project infrastructure, and resources. Formerly incarcerated people may also be weakly attached to mainstream social institutions which only adds to inaccessibility in standard data collection procedures (Brame and Piquero 2003; Western et al. 2017). Moreover, study participation is likely a low priority for previously incarcerated persons owing to the various post-release stressors. If researchers are more likely to retain certain groups it could bias the results, and more importantly, the implications of the research. Given the difficulty of retaining a high-risk sample over time, it is necessary for social scientists to use methods for exploring and correcting systematic sample selection bias.

Estimation Methods for Correcting Attrition

Researchers have long understood that sample selection bias presents issues in sociological and criminological research (Berk 1983; Stolzenberg and Relles 1990). When the observations that remain are dependent upon the outcome of interest, sample selection may lead to biased inferences to the population and thus, materialize into a critical data issue (Berk 1983; Bushway et al. 2007). Though this issue is ubiquitous in panel studies, researchers often fail to evaluate differences between the retained and attrited samples; although,

---

[2] There are circumstances, however, in which reincarceration may actually make someone easier to locate and study as they are in custody (Fahmy et al. 2019). But, overall, it is more difficult to make inferences about this group than those in the general population.

researchers could correct for biases in their data using a variety of methods, depending on their study aims. We detail three commonly used methods, all of which we test in our data.[3]

Listwise Deletion

Listwise deletion is the simplest and least computationally intensive way to address missing data; however, there is debate on whether it adequately addresses selection issues. In listwise deletion, cases which contain missing data on variables that are included in the model are dropped from analyses entirely. The benefits of this technique are that it is parsimonious and does not require the manipulation or creation of data the way more complicated methods require (Tabachnick and Fidell 2001). One concern about this technique is that, depending on the magnitude of missing data, sample sizes can decrease significantly and vary across models. Another concern is that if the data are not missing randomly, it culminates in biased results (Tabachnick and Fidell 2001). In this case, the results and implications of the analyses would be limited to the sample included in the models and not generalizable to the broader population.

In our study, listwise deletion would limit our analysis to include the 506 respondents in the LoneStar Project who completed the wave 3 interview. The 285 other respondents who have missing data on wave 3 would be dropped from the model. Our results would be applicable only to individuals who completed wave 3 of the study, or at best, the *type* of respondent who completes follow-up waves of longitudinal studies. However, this is not important for the purposes of this study because we are not attempting to draw conclusions about predictors of our outcome variable, arrest, but instead trying to determine which technique best approximates the full sample analysis. In theory, listwise deletion does not attempt to deal with issues of selection bias. However, we include it in our comparison to help determine whether more complicated techniques have any added value over the most parsimonious option.

Multiple Imputation

Imputation strategies vary in utility (Wadsworth and Roberts 2008), but criminologists commonly employ multiple-chained imputation. Multiple imputation by chained equations (MICE) uses one imputation model for each variable with missing values to generate a series of imputed datasets that are void of missing data and can be used for analysis. MICE is the preferred method for imputation because its algorithm allows each individual variable with missing data its own imputation model. This means that it can impute for different

---

[3] In some cases, researchers rely on propensity score matching (PSM) to correct for bias in their samples (Dehejia and Wahba 2002), though this method is not relevant here and is beyond the scope of the paper. PSM matches subjects based on relevant variables and should reduce the potential bias due to confounding variables (Rosenbaum and Rubin 1983). Although some researchers suggest that PSM holds promise over a listwise deleted sample (Dehejia and Wahba 1998; Lennox et al. 2012), the technique requires a strict set of assumptions that are difficult to meet in many cases (Dehejia 2005), it may not be appropriately estimated, and does not provide a universal correction for selection (Dehejia 2005; Smith and Todd 2005). For instance, PSM's conditional independence assumption is predicated on the idea that assignment in the treatment group (e.g., retained versus attrited) is based on relevant *observed* characteristics (Campbell et al. 2020; Tucker 2010). Because of this assumption, however, PSM cannot account for the hidden bias created by *unobserved* characteristics (Tucker 2010; Wolfolds and Siegel 2019). When the goal is to control for endogeneity that arises from unobservable characteristics (such as in our case), the Heckman selection model is superior (Lennox et al. 2012).

types of variables (i.e., continuous, binary, ordered/unordered categorical, etc.) while remaining without the logical bounds of potential outcomes for each variable (Royston and White 2011). It does this by regressing each individual variable with missing data on all other variables in the model while restricting the estimation to cases with observed values of the missing variable. MICE allows for the specification of variable type to ensure the estimation at this stage is conducted using the most appropriate method according to the variable type. The missing values are then replaced with the imputed values, which is determined using simulations from the predictive distribution that was estimated from the initial regression analysis. This process is repeated for each of the variables with missing data to create a full imputed dataset (Royston and White 2011). The final step includes repeating the entire process numerous times to create multiple imputed datasets—recommendations vary between 5 and 20 cycles (Van Buuren et al. 1999). In turn, estimates for the complete sample are generated by pooling across the imputed datasets.

In our models, we impose missingness on official arrest for our 285 participants who did not complete the wave 3 follow-up survey. In order to avoid losing 285 cases in our analysis because of attrition, as we did in the listwise deleted sample, we can utilize MICE to replace the missing data on our rearrest variable with imputed data. We replaced the missing values of rearrest (x_1) in the 285 participants without wave 3 interviews by using the coefficients and regression equation to impute the missing value based on the variables which were observed for those with missing values for rearrest (x_1). This process was repeated ten times to ensure stability in the imputed estimates.

When used appropriately (i.e., data are missing completely at random; Allison 2000), multiple imputation can help address selection bias by creating a complete data set, ultimately allowing every participant to be included in analyses. In the case of the LoneStar Project, although the data are missing respondents at wave 2 and/or wave 3, we have full data on all respondents from the in-prison, baseline interviews. Thus, we attempt to contrast multiple imputation techniques against the Heckman two-step correction with the full sample condition.

Heckman Two-Step Correction

Across many disciplines in the social sciences, the Heckman correction pioneered how selection issues were addressed in applied research (Bascle 2008; Certo et al. 2016; LaLonde 1986; Peel 2014; Tucker 2010). In fact, most estimation methods today are, in some form, an extension of Heckman's original (1976, 1979) two-step approach (Yu and Gastwirth 2010). In criminological research, the use of Heckman's two-step correction and related models[4] have not experienced as large of a growth as in economics, accounting, and finance. Bushway et al. (2007) reviewed publications appearing in *Criminology* from 1986 to 2005 and found 25 studies that used the Heckman correction. We updated and extended their review by looking at the Heckman correction in leading generalist and specialized criminology outlets from 2006 to 2019.[5] Of the 21 articles (see Appendix 1) that used the

---

[4] Although the Heckman two-step correction can be completed in two stages, the preferred method involves a Full Information Maximum Likelihood (FIML) model which estimates the equations simultaneously to reduce model error.

[5] We used the following search terms in a Boolean fashion: Heckman AND ("two-step" OR "two-stage" OR "Two-step correction" OR "Two-stage correction" OR "two step" OR " two stage" OR "Two step correction" OR "Two stage correction" OR "selection" OR "correction" OR "Mills ratio" OR "Mill's ratio" OR "rho") NOT "two-stage least squares" and searched the following journals: *Criminology*, *Criminal*

Heckman correction, 17 were related to criminal justice case processing (there was no clear pattern for the remaining four articles). And, no papers used Heckman to account for selection bias introduced via attrition—the focus of our paper.

The original conception of the Heckman two-step correction (Heckman 1976) involves estimating a probit model for selection (i.e., Stage 1 selection model), creating an inverse Mill's ratio, and using the inverse Mill's ratio as a covariate in the model estimating a continuous outcome (i.e., Stage 2 substantive model) to "correct" for sample selection bias. Specifically adapted for full-information maximum likelihood (FIML), this model simultaneously estimates a selection probit model and a substantive logit model for the dependent variable.

As with any post-data collection correction method, there are pitfalls if not utilized correctly. The Heckman estimator is only useful insofar as its assumptions and conditions are met and the data are appropriate (Bascle 2008; Bushway et al. 2007; Wolfolds and Siegel 2019). In short, the error terms must follow a bivariate normal distribution; the sample must be large; collinearity should be absent; and exclusion restrictions must be powerful and relevant—an issue that has not been suitably considered in past research (Bascle 2008). Otherwise, Heckman's method is inappropriate. When the above criteria are not met, a simple ordinary least squares (OLS) regression performs more reliably and with less bias (Wolfolds and Siegel 2019). Indeed, pitfalls of the Heckman estimation method are essentially the counterfactuals of the conditions necessary for its "magic," as Bushway and colleagues put it.

These Heckman-related pitfalls can be categorized in terms of criticisms brought about since the original formation of the method (Heckman 1976). Criticisms include (1) sensitivity to the deviations from the distributional assumptions of the error terms (Little and Rubin 1987; Puhani 2000; Stolzenberg and Relles 1990; Vella 1998); (2) a reliance on untestable assumptions (Duan et al. 1983; Puhani 2000); (3) predictive power and/or efficiency of a maximum likelihood (ML) or OLS estimation compared to the Heckman (Little and Rubin 1987; Puhani 1997, 2000; Stolzenberg and Relles 1990; Vella 1998); (4) failure in the presence of outliers (Yu and Gastwirth 2010); and lastly (5) collinearity in the presence of weak or nonexistent exclusion restrictions (Anderson 2017; Leung and Yu 1996; Little and Rubin 1987; Puhani 2000; Vella 1998). In terms of weak or absent exclusion restrictions, such variables are extremely difficult to find in practice (Leung and Yu 1996; Yu and Gastwirth 2010), but "the most important difference for the performance of the alternative estimators arises from the existence of exclusion restrictions" (Puhani 2000, p. 64), a point that is critical to the current study.

Current Study

We examine the effects of non-random sample selection as a function of attrition on recidivism in longitudinal research in the context of three key research questions:

Footnote 5 (continued)

Justice and Behavior, *Journal of Developmental and Life-Course Criminology*, *Journal of Experimental Criminology*, *Justice Quarterly*, *Journal of Quantitative Criminology*, and *Journal of Research in Crime and Delinquenc*y. The initial search resulted in 67 articles. After careful review of the articles for relevance and the use of a Heckman correction, 46 articles were removed, resulting in a final sample of 21 articles.

1. How does attrition bias the correlates of recidivism?
2. What is the magnitude of bias for each correlate?
3. Given the bias within these correlates, how well do methods of correction account for the sampling selection bias introduced by attrition?

We tackle these questions using four sample selection conditions. The first is an "ideal world" where no attrition exists, which we term the *full sample condition*. Because we have complete data for this condition (i.e., data for all 802 respondents), we have no attrition and thus no sample selection bias. This condition, and the associated parameters, serve as our benchmark for variables related to recidivism.

The second condition illustrates the "real world" of research where attrition exists, which we term the *retained condition*. This condition is the most commonly relied upon data by researchers. For this condition, we limit the sample to only those who were retained in the study after release from prison and reexamine the covariates of recidivism—thus, respondents are listwise deleted. Should the estimates vary in direction, significance, or effect size, it would lend credence to the assertion that a correction is needed.

The third condition we termed the *multiply imputed condition.* This condition is used to examine if bias in selection can be corrected using multiply imputed chained equations. This condition serves as a complete dataset although that completion is only achieved via imputation.

The final condition, which we term the *corrected condition*, is used to explore correction for sample selection bias through the use of a Heckman correction. The corrected condition combines the complete and partial sample by accounting for non-random attrition through the use of robust exclusion restriction variables that we outline below.

It is rare for publicly available longitudinal datasets to have the capacity to examine outcomes based on the four sample conditions we presented above. In an earlier example, LaLonde (1986) set out to accomplish similar goals as our own. His aim was to determine how well econometric techniques replicated the estimates of those produced in a field experiment. Similar to LaLonde (1986), the unique circumstances of our data allow us to determine how well statistical methods account for selection bias and how it operates in comparison to analysis conducted on the full, unmanipulated sample. As such, this gives us a strong basis to explore the implications of sample selection bias on recidivism and test our ability to correct for such bias.

Methods

Data

We use data from the LoneStar Project, or the Texas Study of Trajectories, Associations, and Reentry. The study consisted of three waves of data collection: baseline (within one week prior to release), wave 2 (approximately one month post-release), and wave 3 (approximately ten months post-release). The sampling frame for the baseline interview included a disproportionate stratified random sample of all men scheduled for release from the largest Texas Department of Criminal Justice (TDCJ) regional release center. Because the primary focus of the LoneStar Project was on prison gangs and reentry, TDCJ-identified gang members were oversampled by a factor of five to support subgroup analyses by race, gang type, and gang status. The LoneStar Project's final sample of 802 males includes

368 gang and 434 non-gang respondents, so sampling weights are used for all analyses. Our analytic sample consisted of 791 respondents.[6]

The current study relies on baseline (in-prison) variables as predictors of retention and recidivism. The project's design involved in-person computer-assisted personal interviews (CAPI). Wave 1 data collection began April 2016 and was completed by December 2016. Interviews with respondents in general population housing (95% of the sample) occurred daily in an enclosed, public area of the unit. Interviews with respondents in high custody housing (the remaining 5% of the sample) occurred in a secure visiting room in the administrative segregation wing of a prison (all of whom left prison from the aforementioned release center). Interviews were conducted by trained interviewers using laptop computers and lasted approximately two hours; few people refused to participate ($< 5\%$). Respondents were not compensated for their involvement in the study for baseline interviews. Interview data were linked with administrative records from state agencies, including TDCJ and the Texas Department of Public Safety (DPS). Together, these data paint a detailed picture of self-reported beliefs and behaviors, coupled with official data on criminal justice system contact.

The interview and administrative-linked data are ideal to answer our research questions. First, the LoneStar Project data include unique study metadata, variables, and perceptions of participation that we use as exclusion restrictions (see below). Second, because we have a complete sample and recidivism data for all respondents, we can compare the retained, imputed, and corrected samples to determine the effect of bias on recidivism. Third, the interview data contain a wide range of measures, allowing us to take stock of potential bias across correlates that are not based on official data. Finally, we have two years of post-release measures of arrest, which allows for the exploration of criminal justice system contact over an extensive period of time.

Sample Selection Condition

Retention at wave 3 was used as our sample selection dependent variable. This variable was coded as 1 = "yes" for those who completed a wave 3 interview ($n = 506$). Otherwise, respondents were coded as 0 = "no" if they did not participate in wave 3.

Exclusion Restrictions

Regardless of the type of Heckman correction used, it is crucial to identify proper exclusion restrictions (Bushway et al. 2007; Certo et al. 2016). Exclusion restrictions are variables associated with the process of selection but *not* the substantive model of interest. For example, well-populated collateral contact cards are important for study retention (Clark et al. 2020), but there is no theoretical reason to expect that such information is associated

---

[6] Although we understand that Little's (1988) Missing Completely at Random (MCAR) test does not consider unobservable data and cannot determine whether missing data is truly MCAR, we ran the test to help justify our decision to listwise delete missing data. The test statistic was non-significant ($p = 1.00$), which provides support for our decision. Additionally, only 1.37% of cells were missing, which gives us confidence that that the missing data were sparse enough to utilize listwise deletion. For variables that were missing more than one response, regressions were estimated with a dummy variable adjustment. Coefficients were compared between the mean/mode replaced models and the listwise deleted models and no statistically significant differences existed.

with recidivism. Properly specified exclusion restrictions are critical for reducing the potential for collinearity (Bushway et al. 2007). Indeed, failure to include exclusion restrictions or the use of weak exclusion restrictions will inevitably lead to biased and inconsistent estimates (Anderson 2017). Researchers must consider theoretical and substantive justifications when considering which variables to use as exclusion restrictions in order to ensure the selection process has been accurately modeled (Berk 1983; Bushway et al. 2007; Stolzenberg and Relles 1990). When the same predictors are used to model both the selection process and the outcome, it is problematic because of the correlation between the correction term and the included variables and may inflate perceived explanatory power.

Lennox et al. (2012) demonstrated that most studies fail to adequately report exclusion restrictions. In a review of accounting research, they found a large number of studies failed to include exclusion restrictions (14 out of 75) or do not report the first stage model (7 out of 75). And, 95% of studies failed to provide proper justification—theoretically or empirically—for their chosen restrictions. In a meta-analysis of research on strategy management using the Heckman correction, Wolfolds and Siegel (2019) revealed that fewer than one-third of the 165 papers reviewed claimed the use of a valid exclusion restrictions, which may actually worsen rather than correct estimates. Data limitations typically result in researchers including only a single variable as an exclusion criterion. When estimating a Heckman correction, it is vital that at least one predictor in the selection equation is not included in the final estimation equation (Certo et al. 2016; Stevens et al. 2011).

The metadata from the LoneStar Project allow us to incorporate *nine* variables that are unique to the selection equation and not typically available in other datasets. We draw upon the work of Clark and colleagues (2020), who used data from the LoneStar Project to examine the correlates of retention. They identified a grouping of such correlates— what they termed "preventive measures"—related to retention/attrition yet unrelated to recidivism. We use the following variables as exclusion restrictions: (1) if the respondent received any reminders (sent via text, email, and phone calls) before their interview; (2) the total number of calls to that respondent and their collaterals; (3) whether a concerted retention effort was deployed in an attempt to retain them; (4) how many contacts they provided for themselves and collaterals; (5) whether they completed the wave 2 interview; 6) respondent carelessness when completing the interview; (7) whether or not the interviewer perceived the respondent was being honest in their responses; (8) respondent commitment to study participation; and (9) whether respondents participated in the interview for passive, internal, external, or unclassifiable reasons.[7]

We relied on these exclusion restrictions because previous research has found that they are directly related to study retention, our selection condition (e.g., Clark et al. 2020; Barber et al. 2016; Deeg et al. 2002; Odierna and Bero 2014), but only indirectly related to recidivism, our prediction outcome, through the selection process. For example, prior research has found that many participants join studies for altruistic intentions and that positive inclinations as well as positive experiences with the study team are some of the strongest predictors of retention (Price et al. 2016). The LoneStar Project placed a premium on rapport-building, interview experiences, and study branding (Mitchell et al. 2018). In other

[7] In response to a reviewer's comment, we have run additional analyses and created a table explicating the use of our exclusion restrictions from the LoneStar Project's metadata. We closely examined the findings of linear probability models comparing our exclusion restrictions' ability to predict wave 3 retention versus rearrest from the Clark et al. (2020) paper and are confident that our exclusion restrictions are appropriately justified (see Appendix 2).

words, respondents were invested in study participation. Although there may be reason to believe that there is overlap among those involved with the criminal justice system who also lack interest in study participation post-prison, to the best of our knowledge we have not observed theory or research in the extensive literature on prisoner reentry make a claim for study investment as a predictor of recidivism, as measured by rearrest (Gendreau et al. 1996; Mears and Cochran 2015; Visher and Travis 2003). These nine variables are valid and meet all the criteria necessary for their use, thus lending themselves as viable exclusion restrictions.

Dependent Variable

Recidivism in the form rearrest after release from prison as derived from official records is our outcome. *Arrest* was dichotomously coded (1 = "yes"; 0 = "no").[8] We right-censor the data at December 18, 2018, which represents two years following the conclusion of the baseline data collection. All models control for the number of days post-release through right-censoring.

Independent Variables

We organized the correlates of recidivism into four broad categories: personal attributes, criminal justice involvement, prison and reentry experiences, and attitudes. See Table 1 for variable descriptions and response options.

Personal Attributes

Prior research has found that age (Alper et al. 2018), race/ethnicity (Alper et al. 2018; McGovern et al. 2009), mental health issues (Barrett et al. 2014; Skeem et al. 2011), ties to social institutions (Berg and Huebner 2011; Uggen 2000), and days within the community post-release (Alper et al. 2018) are associated with recidivism. As such, we control for any association they may have with arrest rates.

Criminal Justice Involvement

Prior criminogenic behaviors are strongly correlated with future criminal justice involvement (Katsiyannis et al. 2018; Padfield and Maruna 2006). Hence, it was vital to control for prior criminal justice contact (i.e., total incarcerations, prior arrests) and characteristics of the current incarceration (i.e., incarceration length and incarcerating offense).

Prison and Reentry Experiences

Prison experiences and reintegration barriers can influence the likelihood of success upon reentry. Indeed, researchers have found that gang members (Huebner et al. 2007; Pyrooz et al. 2021) and those housed in restrictive housing (Butler et al. 2017; Clark and Duwe

---

[8] We also created a continuous measure, *arrest count,* representing the number of times a respondent was arrested after release from prison. This estimate was necessary for one modeling strategy—the TPM.

Table 1 Description of variables

	Description	Response options
Dependent variables		
Arrest count[a]	Number of arrests from prison release and 12/18/18	
Arrested	Arrested from prison release and 12/18/18	1 = Yes; 0 = No
Personal attributes		
Age (years)[a]	At the time of the baseline interview	Continuous
Race/ethnicity	Self-reported identity	1 = Yes; 0 = No
White (reference)		
Black		
Latino		
Other	Includes American Indian, Alaskan Native, Asian, East Indian and any respondents with mixed race	
Mental health	Self-rated mental health on average	3 = Excellent; 0 = Poor
Pre-prison education	Number of school years completed prior to incarceration	Continuous
Pre-prison employment	Employed prior to incarceration	1 = Yes; 0 = No
Days within community	Number of days from release to 12/18/18	Continuous
Criminal justice involvement		
Total incarcerations[b]	The number of incarcerations in TDCJ	Continuous
Prior arrests[b]	Number of arrests in Texas prior to current incarceration	Continuous
Incarceration length (years)	Incarceration duration for current incarceration	Continuous
Incarcerating offense	Offense respondent was incarcerated for	1 = Yes; 0 = No
Violent offense (reference)		
Property offense		
Drug offense		
Other offense		
Prison and reentry experiences		
Prison gang member	Self-reported gang status	1 = Yes; 0 = No
Restrictive housing	In restrictive housing immediately prior to their release	1 = Yes; 0 = No
Prison visits	See Appendix 4	See Appendix 4
Misconduct	Variety score self-reported within six months prior to release	0 through 1
Victimization	Variety score self-reported within six months prior to release	0 through 1
Exposure to violence[c]	Number of times a respondent was exposed to four violent incidents within the six months prior to release	Continuous
Delinquent peers	See Appendix 4	See Appendix 4
Parole	Released on parole	1 = yes; 0 = no
Release plan	Released with a reentry plan	1 = yes; 0 = no
Halfway house	Released to a halfway house	1 = yes; 0 = no
Attitudes	See Appendix 4	See Appendix 4
Procedural justice		
Low self-control		

Table 1 (continued)

Description	Response options
Stress	
Social support	
Social capital	

The *days out* variable is the number of days from release from prison until 12/18/18. This variable is included as a control but not presented in the tables

[a]DPS Data; [b]TDCJ data; [c]Due to the skewed nature of the variable, estimates were divided by 100 to decrease the scale of the variable

2018; Mears and Bales 2009) may have greater challenges upon reentry. Prison visitation has been linked with decreases in recidivism post-release (Mitchell et al. 2016). Involvement in misconduct or experiences of victimization and violence may negatively affect a person's ability to successfully reintegrate (Daquin et al. 2016; Hsieh et al. 2018; Lugo et al. 2017; Taylor 2015). Therefore, continuous indicators measuring involvement in violent misconduct or victimization and exposure to violence in the six months prior to release are used. Given the influence that peers can have on the reentry process (Taxman 2017), a mean scale of delinquent peers was created.

Having a plan for release or support during this transition can help to foster reintegration between prison and the free world (Clark 2016; Costanza et al. 2015; Wong et al. 2018). As such, three dichotomous controls were used to represent various supports and control forces designed to increase successful reintegration, including parole status, a reentry plan, and residing in a halfway house.

Attitudes

Respondent attitudes and beliefs can influence recidivism. For example, incarcerated people who were treated in a procedurally just way may be more likely to comply with the law based on their prior interactions (Beijersbergen et al. 2016) and low self-control has also been linked to criminal behavior post-release (Malouf et al. 2014). Hence, we used measures of procedural justice and low self-control. Stress upon reentry may influence recidivism rates (Spohr et al. 2016; Thoits 1995; Wimberly and Engstrom 2018) and social support has been found to aid in the reentry process and potentially decrease involvement in antisocial behavior (Berg and Huebner 2011). Consequently, a stress scale and family social support scale were created. Also, an indicator of social capital was created and included four items measuring readiness for change, three items measuring locus of control, and two items measuring self-efficacy.

Analytic Strategy

Multiple analyses were used to examine the extent of selection bias and determine the degree to which the models can account for such bias. To do this we specify four sampling conditions: (1) a complete sample, (2) a retained sample, (3) an imputed sample, and (4) a Heckman corrected sample.

We used two methods to evaluate our first research question, how attrition biases correlates of recidivism. First, we determined if there was a statistical difference between the

characteristics of the full and retained samples using independent samples *t*-tests (LaLonde 1986). If no bias existed, we would expect the predictors of recidivism in the retained sample to be similar to the predictors in the full sample.

Second, we estimated a Two-Part Model (TPM) to examine bias. A TPM uses a probit model to estimate selection (i.e., retention) and an ordinary least squares regression to estimate the outcome of interest (i.e., rearrest count, see footnote 9). This type of model provides estimates of recidivism conditional on selection into the retained subsample of the population. In the absence of selection or adequate exclusion restrictions, an uncorrected TPM is sufficient for modeling an outcome using a specific subsample of a population (Bushway et al. 2007; Duan et al. 1983). However, if the TPM is biased and the point estimate is wrong, efficiency is no longer relevant because the TPM may produce misleadingly small standard errors. Since it is not possible to test the independence of error terms between the selection and substantive models, Bushway et al. (2007) suggested an alternative approach. This involves comparing the results of the TPM to the results of other models which attempt to correct for selection bias to determine how much of an influence selection may have on model estimates. If the estimates do not vary from those of the TPM, selection is not an issue. Thus, we compared the results from a TPM to the results of models using Heckman's two-step model and variations existed (see Appendix 3).

Since selection bias was determined to be an issue, we proceeded with our second research question: how much is each correlate affected by attrition bias? Stolzenberg and Relles' (1997) modeling technique which, once combined with the coefficients from the TPM, demonstrate the amount of bias that is present within each covariate. In short, Stolzenberg and Relles' (1997) method offers an intuitive approach to understanding the severity of sample selection bias. Their technique allows for the investigation of the level of selection bias in coefficients and then to bound the estimates. This is an important first step to understanding the size and underlying reasons for the selection problem in order to better grapple with the various approaches to address it (Bushway et al. 2007; Stolzenberg and Relles 1997). Bushway et al. (2007) note that this methodology is often underused, despite its ability to determine if a Heckman correction is necessary, and their hope was to reacquaint criminologists to this method for assessing selection bias prior to applying correction methods. To do so, the following equation (Stolzenberg and Relles 1997) is used to estimate the predicted bias for each independent variable, represented by $\aleph_{B_1}$, which is the unknown correlation between the selection and substantive equation error terms. The independent variable of interest is represented by x_1 and all other variables in the equation are identified as $x_2 \ldots x_p$. Z represents the covariates of the selection equation and $\hat{Z}$ is the predicted values from the selection equation.

$$\aleph_{B_1} \approx -\sigma \rho_{\varepsilon\delta} \rho_{x_1 \hat{Z}_{x_2 \cdots x_p}} \lambda' \left(T - \hat{\alpha}_1 \overline{Z}_1 - \cdots - \hat{\alpha}_q \overline{Z}_q \right) \left(S\hat{Z}/S_{x_1} \right) \frac{\sqrt{1 - (R_{\hat{Z} \cdot x_2 \cdots x_p})^2}}{1 - (R_{x_1 \cdot x_2 \cdots x_P})^2}$$

To interpret bias estimates derived from Stolzenberg and Relles' (1997) technique, the predicted bias is divided by the sampling error for each covariate, represented by $\aleph_{\beta_1}/S_{\beta_1}$. Then $\aleph_{\beta_1}/S_{\beta_1}$ is multiplied by the standard error for each covariate from a two-part regression model. The final estimate demonstrates the predicted bias in the dependent variable (i.e., arrest count) for each covariate if bias is not accounted for within the models.

The third research question examined how well each method of correction could account for sample selection bias as a result of attrition. To do so, Linear Probability Models (LPMs) were estimated for the full, retained, imputed, and corrected samples. Results

from these models were compared for consistency of coefficients and equality of coefficient tests were conducted. LPMs allow for simple interpretation and coefficient comparison by modeling the probability of a dichotomous outcome as a linear function of covariates (Breen et al. 2018; Greene 2011).[9]

For the imputed sample, a dichotomous variable for rearrest by wave 3 was generated, which was nearly identical to that used in the full sample model. The only difference was that this variable was coded as missing for those who did not complete wave 3. We used multiple imputation with chained equations (MICE) to impute data for the cases missing on wave 3 rearrest and ran an identical LPM to the full sample using the imputed datasets. In a sense, this simulated a scenario where we had missing data due to attrition that would allow us to compare how analyses with multiple imputation compared to those using the full sample data.

An LPM for a Heckman simultaneous full-information maximum likelihood (FIML) model was estimated for the dichotomous dependent variable: arrest.[10] The selection model included a probit regression for retention with controls and metadata correlates predictive of retention but not arrest (Van de Ven and Van Praag 1981). Then, the substantive model involved estimating a logit regression of the dependent variable while accounting for the independent variables. Two statistics are of particular importance—the Wald's test of independent equations and *rho*. The Wald's test of independent equations determines if the equations are conditional on retention. If it is significant, the null hypothesis is rejected, and we would conclude that the equations are not independent and selection bias from attrition is a problem. The *rho* estimate is a correlation that quantifies the dependence between the outcome and selection models. If it is equal to zero, the selection and prediction models are independent and can be analyzed separately, suggesting selection is not an issue. Conversely, if the *rho* estimate is not equal to zero then the use of a Heckman model is justified.

Multiple concerns exist when estimating a Heckman correction. Strong exclusion restrictions are needed, but it is also important to examine the distribution of the error terms, the presence of multicollinearity, and concerns about endogeneity. The variance inflation factors (VIFs) were calculated by estimating a probit model for selection, creating an inverse Mill's ratio, and estimating a regression to examine the VIFs. No multicollinearity issues existed; the average VIF was 1.37 (range: 1.05–1.99), which falls well below levels of concern (Field 2009; Menard 1995). Moreover, the error terms were normally distributed and endogeneity was examined using endogenous treatment effects (Cattaneo 2010; Cerulli 2014; Wooldridge 2010) and the Hausman Test (Davidson and MacKinnon 1993). The results from the treatment effects indicated that when no one is retained, the average probability of arrest is 0.25, which is 0.04 higher when someone is retained than when someone is not retained. This effect is not statistically different from zero ($p = 0.366$),

---

[9] Although probit models are statistically available for the analysis of a binary outcome, in order for us to compare coefficients across models, as suggested by a reviewer, LPMs were the most appropriate.

[10] Consistent with traditional Heckman models, we attempted to use a continuous measure of arrest. Due to the overdispersion of zeros indicating no arrests in our data, this measure was not normally distributed. A heckpoisson command exists in Stata, but our data did not meet poisson distribution assumptions. Given the challenges with normality, a heckprobit was also assessed in our analyses; however, we ultimately decided to run a FIML LPM Heckman model with a binary outcome in order to compare coefficients across models. Therefore, we estimate a FIML LPM Heckman in order to compare equality of coefficients as well as a heckprobit to maintain the integrity of the Heckman correction using a binary outcome. As demonstrated by Tables 4 and 5, statistically significant coefficients did not vary between Heckman models.

suggesting that retention is independent of arrest outcomes. Similar results are found for the Hausman test ($p = 0.160$), indicating that retention is exogenous to arrest, and issues of endogeneity are limited.

Results

Descriptive and Predictive Differences Based on Sample Selection Conditions

Descriptive statistics and independent samples *t*- tests partitioned by the full, retained, and attrition samples are presented in Table 2. It is evident that there are differences between the respective samples. Arrest differs between those who were retained and those who were not by 13 percentage points (retained = 35 percent; attrited = 48 percent). The arrest count also differed. Whereas the retained group maintained 0.64 arrests, there were 0.82 arrests recorded among the attrition group (Cohen's $d = 0.139$). Differences between the groups extended from the dependent to the independent variables, including age, incarceration length, prison gang membership, delinquent peers, and having a reentry plan. All of these differences were statistically significant at *p*-values at or below 0.05.

The results from the *t*-tests demonstrate that correlates of arrest vary between the full and retained sample, suggesting that attrition may bias results when not properly accounted for.[11] However, the substantive implications of those biases on recidivism is unclear. In other words, how much do these biases affect the magnitude of the correlates of recidivism?

The Magnitude of Sample Selection via Attrition on Correlates of Recidivism

Table 3 provides the results from the Stolzenberg and Relles' (1997) technique to assess bias for all correlates of arrest. Column $\aleph_\beta$ represents the estimate of selection bias; whereas, the column labeled $\aleph_{\beta_1}/S_{\beta_1}$ presents the selection bias divided by the sampling error of the covariate. Although the aforementioned columns are needed for computation, the third column is most important for interpretation because it illustrates the predicted bias in arrests for covariates. For illustrative purposes, an arrest statistic of 0.45 would indicate a one unit increase in a given coefficient biasing the effect by 0.45 arrests. When modeling bias, the correlation between the error terms in the selection and substantive equations is unknown. Consistent with Bushway et al. (2007), we allow $\rho_{\varepsilon\delta}$—the correlation of error terms—to vary between 0.25 and 1; when the correlation increases, so too does the amount of selection bias.

We use a threshold of 0.35 for bias because it represents $\pm 50\%$ change in arrests ($M = 0.70$). The effects of selection bias for 19 of the 28 correlates was modest—that is less than a 0.35 change in arrest—even when the error terms were perfectly correlated. The largest bias was observed for restrictive housing, which ranged from -0.06 with a correlation of 0.25 to -0.22 when the error terms are perfectly correlated. Once combined with the standard errors from the uncorrected TPM (see Appendix 3), the effect size for restrictive housing would be biased by -0.29 to -1.14 arrests. Other large effects were

---

[11] It is possible that this variation is due to the loss of analytical power between the full and retained sample. It becomes more difficult to detect statistically significant differences when a sample changes from 791 to 506 people.

Table 2 Descriptive statistics and t-tests for full, retained, and attrition samples

	Full sample (N=791)		Retained (n=506)		Attrition (n=285)		t-test
	M/%	SD	M/%	SD	M/%	SD	
Dependent variable							
Arrest count	0.70	1.34	0.64	1.42	0.82	1.17	
Arrested	39%		35%		48%		*
Personal attributes							
Age (in years)	40.26	12.12	41.09	12.30	38.71	11.64	*
Race/ethnicity							
White	33%		32%		34%		
Black	28%		29%		24%		
Other	1%		1%		1%		
Latino	31%		28%		35%		
Mental health	2.19	0.81	2.18	0.80	2.20	0.83	
Pre-prison education	11.05	2.24	11.18	2.25	10.81	2.22	
Pre-prison employment	63%		66%		58%		
Criminal justice involvement							
Total incarcerations	1.84	1.23	1.82	1.19	1.89	1.29	
Prior arrests	8.78	8.03	8.42	5.82	9.46	11.02	
Incarceration length	4.37	5.32	4.94	5.85	3.31	3.95	*
Incarcerating offense							
Violent offense	39%		41%		37%		
Property offense	19%		17%		23%		
Drug offense	15%		16%		13%		
Other offense	26%		26%		27%		
Prison and reentry experiences							
Prison gang member	21%		17%		26%		*
Restrictive housing	3%		4%		2%		
Prison visits	57%		55%		62%		
Misconduct	0.14	0.29	0.13	0.28	0.15	0.31	
Victimization	0.20	0.38	0.21	0.39	0.18	0.36	
Exposure to violence	6.53	23.96	5.72	20.83	8.07	28.91	
Delinquent peers	0.77	0.68	0.72	0.64	0.88	0.73	*
Parole	86%		88%		84%		
Release plan	60%		64%		51%		*
Halfway house	18%		18%		18%		
Attitudes							
Procedural justice	1.50	0.59	1.53	0.59	1.46	0.60	
Low self-control	1.37	0.75	1.35	0.73	1.41	0.79	
Stress	0.85	0.56	0.85	0.56	0.85	0.56	
Social support	2.33	0.74	2.35	0.73	2.28	0.75	
Social capital	2.38	0.37	2.37	0.36	2.39	0.37	

Notes. Weighted statistics. *t*-test between attrition and retention groups

Abbreviations. M=Mean; SD=Standard deviation; N=Number of respondents; n=Subsample of respondents

* $p \leq 0.05$

Table 3 Stolzenberg and Relles estimates of selection bias in arrest count for all correlates ($N=791$)

	$\rho_{\varepsilon\delta}=0.25$			$\rho_{\varepsilon\delta}=0.50$			$\rho_{\varepsilon\delta}=0.75$			$\rho_{\varepsilon\delta}=1.00$		
	$\aleph_\beta$	$\aleph_{\beta_1}/S_{\beta_1}$	Arrests	$\aleph_\beta$	$\aleph_{\beta_1}/S_{\beta_1}$	Arrests	$\aleph_\beta$	$\aleph_{\beta_1}/S_{\beta_1}$	Arrests	$\aleph_\beta$	$\aleph_{\beta_1}/S_{\beta_1}$	Arrests
Personal attributes												
Age (in years)	−0.001	−0.198	−0.001	−0.002	−0.396	−0.001	−0.002	−0.593	−0.002	−0.003	−0.791	−0.003
Race/ethnicity												
Black	−0.029	−0.288	−0.097	−0.058	−0.576	−0.195	−0.086	−0.864	−0.292	−0.115	−1.152	−0.389
Other	0.057	0.128	0.107		0.257	0.214	0.170	0.385	0.322	0.227	0.513	0.429
Latino	0.013	0.139	0.016	0.026	0.278	0.032	0.038	0.417	0.048	0.051	0.557	0.064
Mental health	0.002	0.031	0.029	0.003	0.062	0.058	0.005	0.092	0.087	0.007	0.123	0.117
Pre-prison education	−0.004	−0.220	−0.021	−0.009	−0.440	−0.042	−0.013	−0.660	−0.063	−0.018	−0.880	−0.084
Pre-prison employment	−0.040	−0.441	−0.144	−0.080	−0.881	−0.287	−0.120	−1.322	−0.431	−0.160	−1.762	−0.574
Criminal justice involvement												
Total incarcerations	0.002	0.044	0.002	0.003	0.088	0.005	0.005	0.132	0.007	0.006	0.176	0.010
Incarceration length	0.002	0.412	−0.001	0.005	0.824	−0.002	0.007	1.236	−0.003	0.010	1.648	−0.005
Prior arrests	−0.005	−0.686	−0.016	−0.011	−1.372	−0.033	−0.016	−2.058	−0.049	−0.022	−2.744	−0.066
Incarcerating offense												
Property	0.030	0.267	0.110	0.059	0.535	0.221	0.089	0.802	0.331	0.118	1.069	0.441
Drug	−0.045	−0.383	0.008	−0.091	−0.766	0.017	−0.136	−1.149	0.025	−0.181	−1.532	0.034
Other	−0.005	−0.044	0.133	−0.009	−0.088	0.267	−0.014	−0.131	0.400	−0.018	−0.175	0.533
Prison and reentry experiences												
Prison gang member	0.041	0.457	0.213	0.082	0.915	0.426	0.123	1.372	0.639	0.164	1.830	0.852
Restrictive housing	−0.055	−0.301	−0.285	−0.110	−0.602	−0.570	−0.164	−0.903	−0.856	−0.219	−1.204	−1.141
Prison visits	0.012	0.193	0.101	0.024	0.387	0.202	0.036	0.580	0.303	0.048	0.774	0.404
Misconduct	0.034	0.224	0.181	0.067	0.449	0.363	0.101	0.673	0.544	0.135	0.898	0.725
Victimization	0.018	0.147	0.064	0.035	0.294	0.129	0.053	0.441	0.193	0.071	0.588	0.257
Exposure to violence	0.000	0.099	0.001	0.000	0.198	0.002	0.000	0.297	0.002	0.000	0.396	0.003
Delinquent peers	0.031	0.487	0.007	0.062	0.973	0.014	0.093	1.460	0.021	0.123	1.946	0.028

Table 3 (continued)

	$\rho_{\epsilon\delta} = 0.25$			$\rho_{\epsilon\delta} = 0.50$			$\rho_{\epsilon\delta} = 0.75$			$\rho_{\epsilon\delta} = 1.00$		
	$\aleph_\beta$	$\aleph_{\beta_1}/S_{\beta_1}$	Arrests	$\aleph_\beta$	$\aleph_{\beta_1}/S_{\beta_1}$	Arrests	$\aleph_\beta$	$\aleph_{\beta_1}/S_{\beta_1}$	Arrests	$\aleph_\beta$	$\aleph_{\beta_1}/S_{\beta_1}$	Arrests
Parole	−0.017	−0.135	−0.041	−0.033	−0.270	−0.082	−0.050	−0.404	−0.123	−0.067	−0.539	−0.164
Release plan	−0.062	−0.678	−0.086	−0.123	−1.355	−0.171	−0.185	−2.033	−0.257	−0.246	−2.711	−0.343
Halfway house	0.021	0.172	0.028	0.043	0.345	0.056	0.064	0.517	0.084	0.085	0.689	0.112
Attitudes												
Procedural justice	−0.011	−0.152	−0.047	−0.022	−0.304	−0.093	−0.033	−0.456	−0.140	−0.044	−0.608	−0.186
Low self-control	−0.007	−0.126	−0.030	−0.015	−0.251	−0.059	−0.022	−0.377	−0.089	−0.029	−0.503	−0.119
Stress	−0.004	−0.045	−0.044	−0.007	−0.089	−0.087	−0.011	−0.134	−0.131	−0.014	−0.179	−0.174
Social support	−0.012	−0.192	−0.042	−0.024	−0.384	−0.084	−0.036	−0.577	−0.126	−0.048	−0.769	−0.168
Social capital	0.053	0.440	−0.037	0.107	0.880	−0.074	0.160	1.320	−0.110	0.214	1.759	−0.147

Weighted estimates. Days out was a control in each model. Mean arrest count is 0.70

$\rho_{\epsilon\delta}$ = Estimated correlation between error terms; $\aleph_\beta$ = Estimate of selection bias; $\aleph_{\beta_1}/S_{\beta_1}$ = Ratio of selection bias to the sampling error of the covariate; Arrests = Predicted bias in arrest rate for covariate calculated as $((\aleph_{\beta_1}/S)_{\beta_1} \times SE_{TPM})_{\beta_1}$; TPM = Two-part model in Appendix 3

seen for prison gang membership (0.85 arrests), misconduct (0.73), pre-prison employment (0.57), and incarcerated for an "other" offense (0.53), when error terms were perfectly correlated. The implications of bias were smaller, albeit important, for "other" race (0.43) or Black (-0.39), incarcerated for a property offense (0.44), receiving prison visits (0.40), and having a release plan (-0.34). The results from these analyses demonstrate that there is a non-trivial amount of selection bias within the correlates of arrest. Not accounting for this could result in biased regression estimates and inaccurate implications, which is the focus of our second research question.

Comparisons of Corrections for Sample Selection Bias

If a technique was able to account for *all* bias within a model, the corrected model results would mirror those of the full sample. This is not the case for any of the correction techniques, although some models performed better than others.

We start by examining the consistency of statistically significant coefficients between sample conditions and the complete sample condition (see Table 4). The listwise deleted model of retained respondents and the multiply imputed model had four effects that were consistently significant (i.e., age, incarcerating property offense, incarcerating drug offense, and delinquent peers) with the full sample condition. Both the listwise deleted and the imputed models had two effects that were new compared to the full model (i.e., total incarcerations and parole) and prior arrest was no longer significant in either model. In addition, the imputed model predicted one additional significant effect for misconduct.

Before comparing the Heckman model, it is important to note that according to the Wald's test of independence, Heckman equations are conditional on retention; therefore, it is appropriate to control for selection bias within the model (Wald$=7.37$, $p=<0.05$). The *rho* statistic (0.42) also confirms the use of a Heckman correction, suggesting that the selection and substantive equations are not independent of one another. Although the Heckman correction appears to statistically account for bias and is necessary due to the dependence between the selection and substantive equations, the substantive effects of this correction tell a different story. Three effects (i.e., age, property incarcerating offense, and drug incarcerating offense) were consistently significant between the Heckman and the full models; whereas, other race, total incarcerations, and parole were newly significant in the corrected model. Also, incarceration length and delinquent peers were no longer significant.

To further highlight the agreement between the full, retained, and corrected samples, we present concordance statistics on the bottom of Table 4.[12] Between the full and corrected samples, 82% of the findings were consistent across models—this was the lowest of all estimates. In fact, the agreement between the full and retained sample was 7 percentage points higher at 89%, whereas the agreement between the full and the imputed sample was

---

[12] These estimates were calculated by first coding each variable to determine how many of the two coefficients across each model were significant and noting the agreement between those coefficients. For each variable across models, if both coefficients were significant or if both coefficients were non-significant, it was coded as 2 (i.e., "agreement"). If one coefficient was significant and the other was non-significant, it was coded as 0 (i.e., "disagreement"). This step is conducted at the variable level, so each variable within the model had an agreement estimate, which were later summed. Step two requires calculating the percent agreement between the models which involved summing agreement/disagreement estimates from each variable and dividing them by the total number of coefficients across models. That equation was $ModelAgreement = \frac{Sum\ of\ Agreement}{2 \times Number\ of\ Variables} \times 100$.

Table 4 Linear probability models predicting arrest within the full, retained, imputed, and Heckman corrected sample

	Full (N=791)		Retained (n=506)		Imputed (N=791)		Corrected (N=791)	
	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Personal attributes								
Age (in years)	−0.01*	0.00	−0.01*	0.00	−0.01*	0.00	−0.01*	0.00
Race/ethnicity [a]								
Black	−0.03	0.05	−0.01	0.06	−0.03	0.06	0.00	0.06
Other	−0.13	0.10	−0.18	0.09	−0.26	0.15	−0.20*	0.09
Latino	−0.04	0.05	0.00	0.06	−0.02	0.07	−0.01	0.06
Mental health	−0.01	0.03	0.00	0.03	−0.01	0.03	0.00	0.03
Pre-prison education	0.00	0.01	−0.01	0.01	−0.01	0.01	0.00	0.01
Pre-prison employment	−0.05	0.04	−0.04	0.05	−0.05	0.05	−0.02	0.05
Criminal justice involvement								
Total incarcerations	0.04	0.02	0.06*	0.03	0.07*	0.02	0.06*	0.03
Prior arrests	0.01*	0.00	0.01	0.00	0.00	0.00	0.01	0.00
Incarceration length	0.00	0.00	0.00	0.00	−0.01	0.00	0.00	0.00
Incarcerating offense [b]								
Property	0.15*	0.06	0.20*	0.08	0.17*	0.06	0.19*	0.08
Drug	0.25*	0.07	0.24*	0.09	0.21*	0.08	0.24*	0.08
Other	0.09	0.05	0.02	0.06	0.01	0.06	0.01	0.06
Prison and reentry experiences								
Prison gang member	0.07	0.05	0.04	0.06	0.05	0.06	0.02	0.06
Restrictive housing	0.01	0.11	0.06	0.12	0.06	0.11	0.09	0.12
Prison visits	−0.05	0.03	−0.05	0.04	−0.04	0.04	−0.06	0.04
Misconduct	0.08	0.08	0.14	0.11	0.21*	0.09	0.12	0.11
Victimization	0.00	0.06	0.00	0.08	−0.01	0.07	0.04	0.07
Exposure to violence	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Delinquent peers	0.09*	0.03	0.08*	0.04	0.08*	0.04	0.07	0.04

Table 4 (continued)

	Full ($N=791$)		Retained ($n=506$)		Imputed ($N=791$)		Corrected ($N=791$)	
	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Parole	−0.10	0.06	−0.24*	0.08	−0.20*	0.08	−0.24*	0.08
Release plan	−0.04	0.04	0.07	0.05	0.02	0.05	0.08	0.05
Halfway house	−0.08	0.05	−0.01	0.06	0.01	0.07	−0.01	0.06
Attitudes								
Procedural justice	0.07	0.04	0.06	0.04	0.04	0.04	0.06	0.04
Low self-control	−0.03	0.03	−0.02	0.04	−0.01	0.04	−0.02	0.04
Stress	0.01	0.04	−0.03	0.05	−0.05	0.05	−0.04	0.05
Social support	−0.01	0.03	−0.03	0.04	−0.01	0.04	−0.02	0.04
Social capital	−0.01	0.06	0.00	0.08	0.06	0.07	−0.01	0.08
Constant	0.19	0.36	0.08	0.42	0.21	0.45	−0.04	0.41
rho	–	–	–	–	–	–	0.42	–
Wald's test of indep. eq	–		–		–		7.37*	
Concordance estimates [c]			89%		85%		82%	
	$F=(29,761)=8.42$*		$F=(29,476)=6.22$*		$F=(29,572.9)=6.33$*		$\chi^2:(29,791)=145.06$*	
	Adjusted r2=0.18		Adjusted r2=0.20				Log pseudo.: −624.22	

Weighted statistics. Days out was a control in each model and with non-significant p-values

Coef. Coefficients, *SE* Standard error, *N* Number of respondents, *n* Subsample of respondents, *Psuedo* Pseudolikelihood, *Indep.* Independence

[a] Reference white; [b] Reference violent offense; [c] Concordance estimates represent concordance with the full sample

* $p \leq 0.05$

Table 5 Summary of findings of the equality of coefficient comparisons and Heckman probit

	t-tests	S&R	Full vs. Listwise	Full vs. MI	Full vs. Heckman	Heckman Probit[a]	
			Z	Z	Z	Coef.	SE
Dependent variable							
Arrest count	—						
Arrested	*						
Personal attributes							
Age (in years)	*	—	0.00	0.00	0.00	-0.02*	0.01
Race/ethnicity							
White (reference)	—						
Black	—	*	-0.26	0.00	-0.38	0.03	0.19
Other	—	*	0.37	0.72	0.52	-1.45*	0.48
Latino	—	—	-0.51	-0.23	-0.38	-0.01	0.18
Mental health	—	—	-0.24	0.00	-0.24	-0.02	0.11
Pre-prison education	—	—	0.71	0.71	0.00	-0.02	0.04
Pre-prison employment	—	*	-0.16	0.00	-0.47	-0.09	0.16
Criminal justice involvement							
Total incarcerations	—	—	-0.55	-1.06	-0.55	0.20*	0.08
Prior arrests	*	—	0.00	0.00	0.00	-0.02	0.02
Incarceration length	—	—	0.00	0.00	0.00	0.02	0.01
Incarcerating offense							
Violent offense (reference)	—						
Property offense	—	*	-0.50	-0.24	-0.40	0.52*	0.22
Drug offense	—	—	0.09	0.38	0.09	0.71*	0.24
Other offense	—	*	0.90	1.02	1.02	0.04	0.21
Prison and reentry experiences							
Prison gang member	*	*	0.38	0.26	0.64	0.07	0.18
Restrictive housing	—	*	-0.31	-0.32	-0.49	0.35	0.37
Prison visits	—	*	0.00	-0.20	0.20	-0.20	0.13
Misconduct	—	*	-0.44	-1.08	-0.29	0.34	0.31
Victimization	—	—	0.00	0.11	-0.40	0.12	0.23
Exposure to violence	—	—	0.00	0.00	0.00	0.00	0.00
Delinquent peers	*	—	0.20	0.20	0.40	0.22	0.12
Parole	—	—	1.40	1.00	1.40	-0.72*	0.25
Release plan	*	—	-1.72	-0.94	-1.87	0.26	0.15
Halfway house	—	—	-0.90	-1.05	-0.90	-0.08	0.22
Attitudes							
Procedural justice	—	—	0.18	0.53	0.18	0.24	0.13
Low self-control	—	—	-0.20	-0.40	-0.20	-0.05	0.13
Stress	—	—	0.62	0.94	0.78	-0.11	0.16
Social support	—	—	0.40	0.00	0.20	-0.09	0.12
Social capital	—	—	-0.10	-0.76	0.00	0.02	0.24

Number of respondents; for all models was 791 except the number of respondents for the retained sample was 506. An * for the Stolzenberg & Relles (S&R) column represents a bias of ± 0.35 or greater when the error terms were perfectly correlated. [a] Model constant coefficient $= -1.58$; SE $= 1.35$. Wald's test of independent equations $= 7.75*$; *rho* $= 0.57$; concordance estimates $= 82\%$; χ^2: (29,791) $= 91.62*$; log pseudolikelihood: -608.50

Abbreviations: $-$ = Non-significant effect or S&R bias less than ± 0.35; ▨ = Not included in model. S&R = Stolzenberg & Relles; Full = Full sample; Listwise = Listwise deleted sample; MI = Multiple imputation sample; Coef. = Coefficients; SE = Standard error; Z = z-score

* *p* or Z ≤ 0.05

slightly lower at 85%. These estimates suggest that out of all corrections, listwise deletion is producing results that are most similar to the full sample.

Although the consistency of significant correlates varies across models, it is important to examine if those variations are substantively significant. In doing so, we present the equality of coefficient comparisons across all models in Table 5. When comparing the coefficients between the full, listwise, imputed, and Heckman models, none of the coefficients were significantly different from one another. These findings suggest that all three correction samples approximate the full sample closely.[13]

---

[13] We are aware that coefficient comparisons, similar to Paternoster et al. (1998), are designed for and assume independent samples. That is not the case for our data. However, this was the most viable way of comparing coefficients across models since we are not able to estimate seemingly unrelated estimation (SUR). SUR is implausible because equations have to be balanced in terms of number of observations,

Discussion

There are many data sources available to criminologists to examine the etiology of criminal behavior and evaluate responses to crime. Longitudinal research design is one that has proven to be beneficial to the field when seeking to understand leading issues in crime and justice. Yet, attrition could undermine the utility of such data, making it critical to determine the extent to which it undermines conclusions and whether there are viable solutions. In this study we used multiple techniques to explore and correct for potential selection bias, which leads to three main takeaway points.

First, attrition-induced selection bias exists within longitudinal data and produces bias in the correlates of recidivism. Our study design allowed us to examine the bias between the full and retained sample, which was noteworthy given we had complete data on arrests for all respondents. Our findings demonstrated that the retained sample was systematically different than the full sample. Selection bias was observed in bivariate associations and, sometimes, the bias would change the outcome variable by a factor as large as 1.6. The means between the retained and attrited sample were statistically different for six correlates (see Table 5). The error produced by not accounting for bias can be substantial, as demonstrated by the Stolzenberg and Relles' (1997) technique. In some instances, selection bias did not affect the correlate, but for 36% of the correlates, bias could increase or decrease the rearrest count by 43% or more. Bias as a result of attrition is problematic and would go unrecognized if researchers fail to explore the implications of bias within their data. Though our findings rely on LoneStar Project data, it is likely that selection bias alters the conclusions reached about correlates covering a wide range of outcomes across longitudinal studies that experience attrition across waves.

Second, correcting for selection bias may be simpler than previously thought. Listwise deletion proved to be the most accurate and useful means of correcting for attrition. Correspondingly, Brame and Paternoster (2003) note that the most typical response for dealing with missing data in criminological literature is listwise deletion. Indeed, we observed 89% concordance in statistical significance for the coefficients across the full and listwise sample models and perfect concordance in the equality of coefficients. Although the imputed analysis fared better than the Heckman analysis in terms of concordance, imputed data are problematic because the chained equations are based on the cases which have partial data for that collection. Also, imputation does not account for those cases who may have missed an entire wave of data collection (Allison 2002). Therefore, imputation to address attrition would not be ideal, even had it produced results equivalent to the full sample.

The Heckman correction produced less consistent results with the full sample than did the retained sample. Although no coefficients were significantly different from one another, we observed the lowest levels of agreement across models in the Heckman correction model. Disparities in levels of agreement are problematic given most researchers typically interpret the findings based on statistically significant coefficients, of which only 82% matched between the two models. In other words, the Heckman correction performed poorer than simply using the retained model, which is in line with Winship and Mare's

Footnote 13 (continued)

the *sureg* command in Stata does not allow for conditional statements (such as weights) or use of the same dependent variable, and the SUR method is used to model parameters of all equations simultaneously; thus, we have no way of fitting different model types. Due to these limitations, please use caution when interpreting the findings.

([1992](#)) contention that the Heckman correction is not a panacea. Instead, this modeling technique should be used as a tool to examine if selection bias may be present, but it does not necessarily replicate the results that would be found if bias did not exist.

Our final takeaway highlights the role researchers play in advancing the utility of longitudinal research. Our results demonstrated that it is vital for researchers to more carefully consider attrition-induced selection bias when working with longitudinal data. Even basic sensitivity analyses can go a long way in critically examining the assumptions made by incomplete data (Brame and Paternoster 2003). Our review of studies in recent criminological literature using the Heckman correction illustrated that rarely, if ever, is the Heckman correction used to study recidivism or reentry experiences (see Appendix 1). Instead, most of the studies relying on the correction use court processing outcomes. Selection bias introduced through attrition should be explored and corrected (Wolfolds and Siegel 2019). As illustrated through this study, statistical methods exist that adequately determine the presence and magnitude of bias (i.e., Stolzenberg and Relles estimate) and these methods can lead to a better understanding of how bias via attrition may influence their covariates. Because researchers and practitioners rely heavily on longitudinal research for policy and practical implications, our hope is that researchers use these findings to establish a new standard of practice in longitudinal research that examines and accounts for attrition bias.

The steps researchers take moving forward will depend on the adequacy of exclusion restrictions. A Heckman correction model can only be used to confirm the presence of selection bias if the data allow for exclusion restrictions. Because of this, we highlight the importance of collecting metadata during longitudinal data collections so that these variables are available for analysis, should a correction be necessary (Brame and Paternoster 2003). In situations where exclusion restrictions are available the steps that should be taken are: (1) compare *t*-tests between correlates of the full and retained sample; (2) run a two-part model followed by Stolzenberg and Relles (1997) to determine the covariate-specific magnitude of bias[14]; (3) report results from models estimated on the retained sample with intentional transparency about which variables may or may not be accurately estimated based on the results of the Stolzenberg and Relles models; (4) compare the equality of coefficients; and (5) estimate a Heckman correction model, noting the bias estimates (e.g., *rho*, Wald's statistic, or Mill's ratio) and any variation from the retained model. We recommend proceeding in this manner even though the Heckman correction was inadequate. Transparency is needed to explore potential bias, the magnitude of such bias, and how it affects correlates using multiple methods (Wolfolds and Siegel 2019). Furthermore, exploring the utility of the Heckman correction using different data sets will increase the validity of our findings and increase our understanding of when and where the method may or may not be necessary.

In situations where exclusion restrictions are not available, researchers should still (1) compare *t*-tests between correlates of the full and retained sample; (2) report results from models estimated on the retained sample with transparency about which variables may or may not be accurately estimated based on the results of the *t*-tests; and (3) compare coefficients across models. Although *t*-tests and coefficient comparisons will demonstrate bias in each covariate independently rather than the set of covariates as a whole, this will allow

---

[14] Stolzenberg and Relles (1997) require a continuous outcome variable. Although this may prohibit its use for some research questions, we encourage researchers to move away from binary outcomes which limit the variation and restrain the social world to a binary.

researchers to identify potential variation enabling them to determine how seriously their results may be biased and make recommendations for policy and practice accordingly.

In closing, this research highlighted the importance of examining selection bias as a result of attrition in longitudinal studies. The results confirm that attrition biases covariates of recidivism outcomes and in some instances those biases are large in magnitude. The Heckman correction is no universal remedy. Paradoxically, listwise deletion produced the most consistent findings with the full sample. Indeed, in many cases, "more complicated methods [to examine missing data issues] would not be warranted" (Brame and Paternoster 2003, p. 75). Selection bias will continue to plague longitudinal research if investigators do not capitalize on the tools to examine the implications of attrition. Correcting for such bias is not only applicable to research on sentencing and court outcomes. Exploring and reporting the potential effects selection bias can have on results will increase transparency and certainty of research conclusions. We believe these conclusions will empower researchers to examine how their longitudinal data may be biased when retention rates are less than ideal—which is far too often the case.

Appendix 1

See Table 6.

Table 6 Systematic review of Heckman correction from 2006 to 2019

Authors	Year	Journal	Key dependent variable(s)
Griffin & Wooldredge	2006	*Crim*	Sentence length
Johnson	2006	*Crim*	Sentence length
Ulmer & Bradley	2006	*Crim*	Sentence length
Felson & Pare	2007	*JQ*	(1) Arrest; (2) conviction; (3) incarceration
Kingsnorth & MacIntosh	2007	*JQ*	(1) Case filed; (2) filed as felony; (3) charge was dismissed; (4) charge was reduced
Wooldredge	2007	*JRCD*	Sentence length
Johnson, Ulmer, & Kramer	2008	*Crim*	(1) Federal guidelines departures; (2) magnitude of departures
Johnson & Betsinger	2009	*Crim*	(1) Federal guidelines departures; (2) incarceration; (3) sentence length
Giordano, Lonardo, Manning, & Longmore	2010	*Crim*	Relationship qualities
Ulmer, Eisenstein, & Johnson	2010	*JQ*	Sentence length
Xie & McDowall	2010	*Crim*	Racial turnover
Feldmeyer & Ulmer	2011	*JRCD*	Sentence length
Ulmer, Light, & Kramer	2011	*JQ*	(1) Likelihood of an imprisonment sentence; (2) sentence length
Steiner, Makarios, Travis III, & Meade	2012	*JQ*	(1) Felony arrest; (2) time to felony arrest
Kutateladze, Andiloro, Johnson, & Spohn	2014	*Crim*	(1) Case acceptance; (2) pretrial detention; (3) case dismissal; (4) custodial plea offer; (5) incarceration
Turanovic & Pratt	2014	*JQC*	Repeat victimization
Feldmeyer, Warren, Siennick, & Neptune	2015	*JRCD*	(1) Sentence disposition; (2) sentence length
Hauser & Peck	2017	*JQ*	(1) Incarceration; (2) sentence length
Daigle & Teasdale	2018	*CJB*	Physical assault victimization
Peck, Leiber, & Beaudry-Cyr	2019	*CJB*	(1) Intake; (2) adjudication; (3) judicial disposition
Petersen	2019	*JQ*	(1) Time-to-plea; (2) punishment outcome of guilty pleas

Crim Criminology, *JQ* Justice Quarterly, *JRCD* Journal of Research in Crime and Delinquency, *CJB* Criminal Justice and Behavior

Appendix 2

See Table 7.

Table 7 LPMs comparing our exclusion restrictions' ability to predict wave 3 retention versus rearrest (N=791)

	Wave 3 retention			Rearrest		
	Coef	SE	p	Coef	SE	p
Exclusion restrictions						
Reminder	0.03	(0.04)	0.44	0.00	(0.04)	0.96
Total calls	0.00	(0.00)	0.15	0.00	(0.00)	0.75
W2 completion	0.31*	(0.08)	0.00	− 0.06	(0.06)	0.34
Concerted retention	− 0.22*	(0.05)	0.00	0.03	(0.07)	0.66
Contact count W1/W2	0.03*	(0.01)	0.01	0.00	(0.01)	0.96
Respondent carelessness scale	0.01	(0.04)	0.89	− 0.03	(0.05)	0.55
Interviewer lie scale	− 0.01	(0.03)	0.68	− 0.04	(0.04)	0.27
Respondent commitment	0.02	(0.05)	0.66	0.03	(0.06)	0.58
Respondent study interest						
Internal	− 0.02	(0.04)	0.65	0.00	(0.05)	0.98
External	− 0.07	(0.04)	0.10	0.08	(0.06)	0.13
Passive	− 0.03	(0.07)	0.65	0.03	(0.09)	0.74
Constant	0.40			0.53		
R-squared	0.31			0.02		

Weighted statistics

Coef. Coefficients, *SE* robust standard error; *p* P-value, *W1* baseline, *W2* wave 2

*$p \leq 0.05$

Appendix 3

See Table 8.

Table 8 Uncorrected two-part regression model predicting arrest count

	Two-part model probit (N = 791)			Two-part model regression (N = 364)		
	Coef	SE	p	Coef	SE	p
Personal attributes						
Age (in years)	− 0.03	0.01	0.00	− 0.01	0.01	0.47
Race/ethnicity [a]	− 0.10	0.16	0.52			
Black	− 0.49	0.43	0.26	0.04	0.22	0.86
Other	− 0.12	0.15	0.42	− 0.21	0.50	0.67
Latino	− 0.03	0.08	0.76	− 0.04	0.26	0.89
Mental health	0.00	0.03	0.87	− 0.18	0.14	0.21
Pre-prison education	− 0.17	0.13	0.20	0.07	0.06	0.21
Pre-prison employment	− 0.03	0.01	0.00	− 0.29	0.21	0.17
Criminal justice involvement						
Total incarcerations	0.13	0.06	0.04	− 0.04	0.08	0.61
Incarceration length	− 0.01	0.01	0.45	0.04	0.01	0.00
Prior arrests	0.03	0.01	0.00	− 0.01	0.02	0.61
Incarcerating offense [b]						
Property	0.41	0.18	0.02	0.23	0.27	0.40
Drug	0.69	0.20	0.00	0.08	0.26	0.75
Other	0.27	0.17	0.11	0.10	0.28	0.73
Prison and reentry experiences						
Prison gang member	0.20	0.14	0.15	0.32	0.41	0.44
Restrictive housing	0.06	0.31	0.84	0.18	0.63	0.77
Prison visits	− 0.17	0.09	0.06	0.13	0.19	0.48
Misconduct	0.21	0.24	0.37	− 0.42	0.34	0.22
Victimization	− 0.01	0.18	0.94	− 0.07	0.25	0.77
Exposure to violence	0.00	0.00	0.11	0.00	0.00	0.76
Delinquent peers	0.27	0.10	0.01	0.26	0.14	0.06
Parole	− 0.31	0.19	0.10	− 0.47	0.49	0.34
Release plan	− 0.16	0.13	0.20	− 0.35	0.24	0.15
Halfway house	− 0.31	0.19	0.10	− 0.39	0.22	0.08
Attitudes						
Procedural justice	0.20	0.11	0.08	0.22	0.24	0.36
Low self-control	− 0.12	0.10	0.25	− 0.02	0.13	0.87
Stress	0.04	0.13	0.78	− 0.18	0.26	0.48
Social support	− 0.06	0.10	0.59	0.03	0.13	0.82
Social capital	0.02	0.20	0.92	0.64	0.42	0.13
Constant	− 0.90	1.09	0.41	0.96	1.54	0.53
χ^2/F Statistic	(29, 791) = 130.35*			(29, 334) = 2.24*		
Pseudo r^2/Adjusted r^2	0.18			0.11		

Weighted statistics. Days out was a control in this model. [a] Reference white; [b] Reference violent offense

Coef. coefficients, *SE* standard error

*$p \leq 0.05$

Appendix 4: Description of Constructs and Coding

Prison visits (mean score of family members and friends visiting); response options: *(Never, A few times, Monthly, Bi-weekly, Weekly, Daily)*.

- *Lead in question:* In the last six months, did you have visitation privileges? *(No, Yes)*
- [IF YES] Please tell me if you have been visited by the following people:

  - Any family members?
  - Any friends who are not gang members?

Delinquent peers (mean score); response options: *(No, Yes)*.
   … Think about your closest friends before being incarcerated. A close friend is some-one you could call in an emergency, someone you can trust.
   Have any of those friends ever…

- been arrested?
- been convicted of a crime?
- been in a correctional facility, such as a jail, prison, or juvenile correctional facility?
- had problems with drugs or alcohol?
- Are any of those friends currently in a correctional facility?

Procedural justice (mean score); response options: *(Always, Most of the time, Some-times, Never)*; alpha = 0.89.
   How often do police officers…

- give people a chance to tell their side of the story before they make decisions?
- treat people fairly?
- respect people's rights?
- make decisions that are good for everyone in the community?
- clearly explain the reasons for their actions and decisions?
- treat people with dignity and respect?
- try to do what is best for the people they are dealing with?

Low self-control (mean score); response options: (1 = *not at all like you*, 2 = *A lit-tle bit like you*, 3 = *Somewhat like you*, 4 = *More so like you*, 5 = *very much like you*); alpha = 0.80.

- You are good at resisting temptation. (reverse coded)
- You have a hard time breaking bad habits.
- You say inappropriate things.
- You do certain things that are bad for you if they are fun.
- You refuse things that are bad for you. (reverse coded)
- Pleasure and fun sometimes keeps you from getting work done.
- You have trouble concentrating.
- You are able to work effectively toward long-term goals. (reverse coded)
- Sometimes you can't stop yourself from doing something, even if you know it is wrong.
- You often act without thinking through all the alternatives.
- You have iron self-discipline. (reverse coded)

Stress (mean score); response options: *(All of the time, Most of the time, Sometimes, None of the time)*; alpha=0.64.

In the past month, how often have you felt…

- that you were unable to control the important things in your life?
- confident about your ability to handle your personal problems? (reverse coded)
- that things were going your way? (reverse coded)
- difficulties were piling up so high that you could not overcome them?
- worried or stressed about your upcoming reentry to the community?

Social support (mean score); response options: *(Strongly agree, Agree, Disagree, Strongly disagree)*; alpha=0.95.

You have someone in your family who…

- is willing to help you make decisions.
- really tries to help you.
- can give you the emotional help and support you need.
- provide help or advice on finding a place to live.
- provide help or advice on finding a job.
- provide support for dealing with a substance abuse problem.
- provide transportation to work or other appointments if needed.
- provide financial support.

Social capital (mean score); response options: *(Strongly agree, Agree, Disagree, Strongly disagree)*; alpha=0.74.

- You can do just about anything you really set your mind to.
- You often feel helpless dealing with the problems of life. (reverse coded)
- You have little control over the things that happen to you. (reverse coded)
- There is really no way you can solve some of the problems you have. (reverse coded)
- What happens to you in the future mostly depends on you.
- You are tired of the problems caused by the crimes you committed.
- You want to get your life straightened out.
- You think you will be able to stop committing crimes when released.
- You will give up friends and hangouts that get you into trouble.

References

Allison PD (2000) Multiple imputation for missing data: a cautionary tale. Soc Methods Res 28(3):301–309

Allison PD (2002) Missing data. Sage, Thousand Oaks

Alper M, Durose MR, Markman J (2018) 2018 Update on prisoner recidivism: a 9-year follow-up period (2005–2014). U.S. Department of Justice, Bureau of Justice Statistics, Washington, DC

Anderson B (2017) Selection models and weak instruments. Musings. Retrieved from https://info.umkc.edu/drbanderson/selection-models-and-weak-instruments/

Asendorpf JB, Van De Schoot R, Denissen JJA, Hutteman R (2014) Reducing bias due to systematic attrition in longitudinal studies: the benefits of multiple imputation. Int J Behav Dev 38(5):453–460. https://doi.org/10.1177/0165025414542713

Badawi MA, Eaton WW, Myllyluoma J, Weimer LG, Gallo J (1999) Psychopathology and attrition in the Baltimore ECA 15-year follow-up 1981–1996. Soc Psychiatry Psychiatr Epidemiol 34(2):91–98. https://doi.org/10.1007/s001270050117

Barber J, Kusunoki Y, Gatny H, Schulz P (2016) Participation in an intensive longitudinal study with weekly web surveys over 2.5 years. Jl Med Internet Res 18(6):1–12. https://doi.org/10.2196/jmir.5422

Barrett DE, Katsiyannis A, Zhang D, Zhang D (2014) Delinquency and recidivism: a multicohort, matched-control study of the role of early adverse experiences, mental health problems, and disabilities. J Emot Behav Disord 22(1):3–15. https://doi.org/10.1177/1063426612470514

Barry AE (2005) How attrition impacts the internal and external validity of longitudinal research. J Sch Health 75(7):267–270. https://doi.org/10.1111/j.1746-1561.2005.00035.x

Bascle G (2008) Controlling for endogeneity with instrumental variables in strategic management research. Strateg Organ 6(3):285–327. https://doi.org/10.1177/1476127008094339

Beijersbergen KA, Dirkzwager AJE, Nieuwbeerta P (2016) Reoffending after release: Does procedural justice during imprisonment matter? Crim Justice Behav 43(1):63–82. https://doi.org/10.1177/0093854815609643

Berg MT, Huebner BM (2011) Reentry and the ties that bind: an examination of social ties, employment, and recidivism. Justice Q 28(2):382–410. https://doi.org/10.1080/07418825.2010.498383

Berk RA (1983) An introduction to sample selection bias in sociological data. Am Sociol Rev 48(3):386–398

Bolanos F, Herbeck D, Christou D, Lovinger K, Pham A, Raihan A et al (2012) Using Facebook to maximize follow-up response rates in a longitudinal study of adults who use methamphetamine. Subst Abuse Res Treat 6:1–11

Boys A, Marsden J, Stillwell G, Hatchings K, Griffiths P, Farrell M (2003) Minimizing respondent attrition in longitudinal research: practical implications from a cohort study of adolescent drinking. J Adolesc 26(3):363–373

Brame R, Paternoster R (2003) Missing data problems in criminological research: two case studies. J Quant Criminol 19(1):55–78. https://doi.org/10.4324/9781315089256-5

Brame R, Piquero AR (2003) Selective attrition and the age-crime relationship. J Quant Criminol 19(2):107–127

Breen R, Karlson KB, Holm A (2018) Interpreting and understanding logits, probits, and other nonlinear probability models. Ann Rev Sociol 44(1):39–54. https://doi.org/10.1146/annurev-soc-073117-041429

Bushway S, Johnson BD, Slocum LA (2007) Is the magic still there? The use of the Heckman two-step correction for selection bias in criminology. J Quant Criminol 23(2):151–178

Butler J, Quinn SC, Fryer CS, Garza MA, Kim KH, Thomas SB (2013) Characterizing researchers by strategies used for retaining minority participants: Results of a national survey. Contemp Clin Trials 36(1):61–67

Butler HD, Steiner B, Makarios MD, Travis LF (2017) Assessing the effects of exposure to supermax confinement on offender postrelease behaviors. Prison J 97(3):275–295. https://doi.org/10.1177/0032885517703925

Campbell DT, Stanley JC (1963) Experimental and quasi-experimental designs for research. Rand McNally, Chicago

Campbell CM, Labrecque RM, Weinerman M, Sanchagrin K (2020) Gauging detention dosage: assessing the impact of pretrial detention on sentencing outcomes using propensity score modeling. J Crim Just 70:1–14. https://doi.org/10.1016/j.jcrimjus.2020.101719

Carkin DM, Tracy PE (2015) Adjusting for unit non-response in surveys through weighting. Crime Delinq 61(1):143–158. https://doi.org/10.1177/0011128714556739

Cartwright N (2007) Are RCTs the gold standard? BioSocieties 2(1):11–20. https://doi.org/10.1017/s1745855207005029

Cattaneo MD (2010) Efficient semiparametric estimation of multi-valued treatment effects under ignorability. J Econom 155(2):138–154. https://doi.org/10.1016/j.jeconom.2009.09.023

Certo ST, Busenbark JR, Woo H-S, Semadeni M (2016) Sample selection bias and Heckman models in strategic management research. Strateg Manag J 37:2639–2657. https://doi.org/10.1002/smj

Cerulli G (2014) Ivtreatreg: a command for fitting binary treatment models with heterogeneous response to treatment and unobservable selection. Stata J 14(3):453–480. https://doi.org/10.1177/1536867x1401400301

Chang MW, Brown R, Nitzke S (2009) Participant recruitment and retention in a pilot program to prevent weight gain in low-income overweight and obese mothers. BMC Public Health 9:1–11. https://doi.org/10.1186/1471-2458-9-424

Chatfield MD, Brayne CE, Matthews FE (2005) A systematic literature review of attrition between waves in longitudinal studies in the elderly shows a consistent pattern of dropout between differing studies. J Clin Epidemiol 58(1):13–19. https://doi.org/10.1016/j.jclinepi.2004.05.006

Clark VA (2016) Predicting two types of recidivism among newly released prisoners: first addresses as "launch pads" for recidivism or reentry success. Crime Delinq 62(10):1364–1400. https://doi.org/10.1177/0011128714555760

Clark VA, Duwe G (2018) From solitary to the streets: the effect of restrictive housing on recidivism. Corrections. https://doi.org/10.1080/23774657.2017.1416318

Clark KJ, Mitchell MM, Fahmy C, Pyrooz DC, Decker SH (2020) What if they are all high-risk for attrition? Correlates of retention in a longitudinal study of reentry from prison. Int J Offender Ther Comp Criminol. https://doi.org/10.1177/0306624X20967934

Claus RE, Kindleberger LR, Dugan MC (2002) Predictors of attrition in a longitudinal study of substance abusers. J Psychoact Drugs 34(1):69–74

Coen AS, Patrick DC, Shern DL (1996) Minimizing attrition in longitudinal studies of special populations: an integrated management approach. Eval Program Plan 19(4):309–319

Cordray S, Polk K (1983) The implications of respondent loss in panel studies of deviant behavior. J Res Crime Delinq 20(2):214–242. https://doi.org/10.1177/002242788302000205

Costanza SE, Cox SM, Kilburn JC (2015) The impact of halfway houses on parole success and recidivism. J Sociol Res 6(2):39–55. https://doi.org/10.5296/jsr.v6i2.8038

Crisanti AS, Case BF, Isakson BL, Steadman HJ (2014) Understanding study attrition in the evaluation of jail diversion programs for persons with serious mental illness or co-occurring substance use disorders. Crim Justice Behav 41(6):772–790

Cullen FT, Pratt TC, Graham A (2019) Why longitudinal research is hurting criminology. The Criminologist 5:63

Curtis R (2010) Getting good data from people that do bad things: Effective methods and techniques for conduting research with hard-to-reach and hidden populations. In: Bernasco W (ed) Offenders on offending: learning about crime from criminals. Willan Publishing, London, pp 141–158

Daquin JC, Daigle LE, Listwan SJ (2016) Vicarious victimization in prison: examining the effects of witnessing victimization while incarcerated on offender reentry. Crim Justice Behav 43(8):1018–1033

David MC, Alati R, Ware RS, Kinner SA (2013) Attrition in a longitudinal study with hard-to-reach participants was reduced by ongoing contact. J Clin Epidemiol 66(5):575–581

Davidov E, Yang-Hansen K, Gustafsson JE, Schmidt P, Bamberg S (2006) Does money matter? A theory-driven growth mixture model to explain travel-mode choice with experimental data. Methodology 2(3):124–134. https://doi.org/10.1027/1614-2241.2.3.124

Davidson R, MacKinnon JG (1993) Estimation and inference in econometrics. Oxford University Press, New York

Deeg DJH, Van Tilburg T, Smit JH, De Leeuw ED (2002) Attrition in the longitudinal aging study Amsterdam: the effect of differential inclusion in side studies. J Clin Epidemiol 55(4):319–328. https://doi.org/10.1016/S0895-4356(01)00475-9

Dehejia RH (2005) Does matching overcome LaLonde's critique of non-experimental estimators? A postscript. In: Nber working paper series

Dehejia RH, Wahba S (1998) Causal effects in non-experimental studies: re-evaluating the evaluation of training programs (No. 6586)

Dehejia RH, Wahba S (2002) Propensity score-matching methods for nonexperimental causal studies. Rev Econ Stat 84(1):151–161

Duan N, Manning WG, Morris CN, Newhouse JP (1983) A comparison of alternative models for the demand for medical care. J Bus Econ Stat 1(2):115–126. https://doi.org/10.1080/07350015.1983.10509330

Eidson JL, Roman CG, Cahill M (2016) Successes and challenges in recruiting and retaining gang members in longitudinal research: lessons learned from a multisite social network study. Youth Violence Juvenile Just. https://doi.org/10.1177/1541204016657395

Fahmy C, Clark K, Mitchell MM, Decker SH, Pyrooz DC (2019) Method to the madness: tracking and interviewing respondents in a longitudinal study of prisoner reentry. Sociol Methods Res. https://doi.org/10.1177/0049124119875962

Field A (2009) Discovering statistics using SPSS. Sage, Thousand Oaks

Fitzgerald J, Gottschalk P, Moffitt R (1998) An analysis of sample attrition in panel data: the Michigan Panel Study of income dynamics. Population 33(2):251–299

Fumagalli L, Laurie H, Lynn P (2013) Experiments with methods to reduce attrition in longitudinal surveys. J R Stat Soc A Stat Soc 176(2):499–519

Gendreau P, Little T, Goggin C (1996) A meta-analysis of the predictors of adult offender recidivism: What works! Criminology 34(4):575–607

Goodman JS, Blum TC (1996) Assessing the non-random sampling effects of subject attrition in longitudinal research. J Manag 22(4):627–652. https://doi.org/10.1016/s0149-2063(96)90027-6

Greene WH (2011) Econometric analysis, 7th edn. Prentice Hall, Upper Saddle River

Gustavson K, Von Soest T, Karevold E, Roysamb E (2012) Attrition and generalizability in longitudinal studies: findings from a 15-year population-based study and a Monte Carlo simulation study. BMC Public Health 12:918–929. https://doi.org/10.1186/1471-2458-12-918

Hansen WB, Tobler NS, Graham JW (1990) Attrition in substance abuse prevention research: a meta-analysis of 85 longitudinally followed cohorts. Eval Rev 14(6):677–685

Heckman JJ (1976) The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. Ann Econ Soc Meas 5(4):475–492

Heckman JJ (1979) Sample selection bias as a specification error. Econometrica 47(1):153–161. https://doi.org/10.1007/S10021-01

Hirschi T, Gottfredson M (1983) Age and the explanation of crime. Am J Sociol 89(3):552–584

Hsieh M-L, Hamilton Z, Zgoba KM (2018) Prison experience and reoffending: exploring the relationship between prison terms, institutional treatment, infractions, and recidivism for sex offenders. Sex Abuse 30(5):556–575. https://doi.org/10.1177/1079063216681562

Huebner BM, Varano SP, Bynum TS (2007) Gangs, guns, and drugs: recidivism among serious, young offenders. Criminol Public Policy 6(2):187–222

Katsiyannis A, Whitford DK, Zhang D, Gage NA (2018) Adult recidivism in United States: a meta-analysis 1994–2015. J Child Fam Stud 27(3):686–696. https://doi.org/10.1007/s10826-017-0945-8

Kinner SA (2006) Continuity of health impairment and substance misuse among adult prisoners in Queensland. Australia Int J Prison Health 2(2):101–113. https://doi.org/10.1080/17449200600935711

LaLonde RJ (1986) Evaluating the econometric evaluations of training programs with experimental data. Am Econ Rev 76(4):604–620

Lennox CS, Francis JR, Wang Z (2012) Selection models in accounting research. Account Rev 87(2):589–616. https://doi.org/10.2308/accr-10195

Leung SF, Yu S (1996) On the choice between sample selection and two-part models. J Econom 72:197–229

Liberman AM (2008) The long view of crime: a synthesis of longitudinal research. Springer, Washington, DC

Little RJA (1988) A test of missing completely at random for multivariate data with missing values. J Am Stat Assoc 83(404):1198–1202

Little RJ, Rubin DB (1987) Statistical analysis with missing data. Wiley, New York

Lugo M, Wooldredge J, Pompoco A, Sullivan C, Latessa EJ (2017) Assessing the impact of unit management programs on institutional misconduct and prison "returns." Just Q. https://doi.org/10.1080/07418825.2017.1357741

Lugtig P (2014) Panel attrition: separating stayers, fast attriters, gradual attriters, and lurkers. Sociol Methods Res 43(4):699–723. https://doi.org/10.1177/0049124113520305

Magruder KM, Ouyang B, Miller S, Tilley BC (2009) Retention of under-represented minorities in drug abuse treatment studies. Clin Trials 6(3):252–260

Malouf ET, Schaefer KE, Witt EA, Moore KE, Stuewig J, Tangney JP (2014) The brief self-control scale predicts jail inmates' recidivism, substance dependence, and post-release adjustment. Pers Soc Psychol Bull 40(3):334–347. https://doi.org/10.1177/0146167213511666

McGovern V, Demuth S, Jacoby JE (2009) Racial and ethnic recidivism risks: a comparison of postincarceration rearrest, reconviction, and reincarceration among White, Black, and Hispanic releasees. Prison J 89(3):309–327. https://doi.org/10.1177/0032885509339507

McLaughlin P, White N, King K, Hann RG, Williams RJ, Schopflocher D et al (2014) QLS front-line retention manual: methods for achieving a 94% cohort retention rate in longitudinal research

Mears DP, Bales WD (2009) Supermax incarceration and recidivism. Criminology 47(4):1131–1166

Mears DP, Cochran JC (2015) Prisoner reentry in the era of mass incarceration. Sage, Thousand Oaks

Menard S (1995) Applied logistic regression analysis: Sage university series on quantitative applications in the social sciences. Sage, Thousand Oaks

Mitchell MM, McCullough K, Wu J et al. (2018) Survey research with gang and non-gang members in prison: operational lessons from the LoneStar Project. Trends Organ Crim. https://doi.org/10.1007/s12117-018-9331-1

Mitchell MM, Spooner K, Jia D, Zhang Y (2016) The effect of prison visitation on reentry success: a meta-analysis. J Crim Just 47:74–83. https://doi.org/10.1016/j.jcrimjus.2016.07.006

Mutti S, Kennedy RD, Thompson ME, Fong GT (2014) Prepaid monetary incentives-predictors of taking the money and completing the survey: results from the International Tobacco Control (ITC) four-country survey. Sociol Methods Res 43(2):338–355

Odierna DH, Bero LA (2014) Retaining participants in outpatient and community-based health studies: researchers and participants in their own words. SAGE Open 4(4):1–11. https://doi.org/10.1177/2158244014554391

Odierna DH, Schmidt LA (2009) The effects of failing to include hard-to-reach respondents in longitudinal surveys. Am J Public Health 99(8):1515–1521. https://doi.org/10.2105/AJPH.2007.111138

Padfield N, Maruna S (2006) The revolving door at the prison gate: exploring the dramatic increase in recalls to prison. Criminol Crim Just 6(3):329–352. https://doi.org/10.1177/1748895806065534

Paternoster R, Brame R, Mazerolle P, Piquero A (1998) Using the correct statistical test for the equality of regression coefficients. Criminology 36(4):859–866

Peel MJ (2014) Addressing unobserved endogeneity bias in accounting studies: control and sensitivity methods by variable type. Acc Bus Res 44(5):545–571. https://doi.org/10.1080/00014788.2014.926249

Price SM, Park CH, Brenner RA, Lo A, Adams S, Baetz RA, Li T (2016) Participant retention in a longitudinal study: Do motivations and experiences matter? Survey Practice 9(4):1–10. https://doi.org/10.29115/sp-2016-0022

Puhani PA (1997) Foul or fair? The Heckman correction for sample selection and its critique: a short survey. In: ZEW Discussion Papers (No. 97–07). Mannheim, Germany

Puhani PA (2000) The Heckman correction for sample selection and its critique. J Econ Surv 14(1):53–68. https://doi.org/10.1111/1467-6419.00104

Pyrooz DC, Clark KJ, Tostlebe JJ, Decker SH, Orrick E (2021) Gang affiliation and prisoner reentry: discrete-time variation in recidivism by current, former, and non-gang status. J Res Crime Delinq 58(2):192–234. https://doi.org/10.1177/0022427820949895

Ribisl KM, Walton MA, Mowbray CT, Luke DA, Davidson WS, Bootsmiller BJ (1996) Minimizing participant attrition in panel studies through the use of effective retention and tracking strategies: review and recommendations. Eval Program Plann 19(1):1–25

Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. Biometrika 70(1):41–55. https://doi.org/10.1017/CBO9780511810725.016

Royston P, White IR (2011) Multiple imputation by chained equations (MICE): Implementation in Stata. J Stat Softw 45(4):1–20

Sampson RJ (2010) Gold standard myths: observations on the experimental turn in quantitative criminology. J Quant Criminol 26(4):489–500. https://doi.org/10.1007/s10940-010-9117-3

Schmidt JA, Gruman C, King MB, Wolfson LI (2000) Attrition in an exercise intervention: a comparison of early and later dropouts. J Am Geriatr Soc 48(8):952–960. https://doi.org/10.1111/j.1532-5415.2000.tb06894.x

Shinkfield AJ, Graffam J (2007) Community reintegration of ex-prisoners: type and degree of change in variables influencing successful reintegration. Int J Offender Ther Comp Criminol 53(1):29–42. https://doi.org/10.1177/0306624x07309757

Skeem JL, Manchak S, Peterson JK (2011) Correctional policy for offenders with mental illness: creating a new paradigm for recidivism reduction. Law Hum Behav 35(2):110–126. https://doi.org/10.1007/s10979-010-9223-7

Smith JA, Todd PE (2005) Does matching overcome LaLonde's critique of nonexperimental estimators? J Econom 125(1–2):305–353. https://doi.org/10.1016/j.jeconom.2004.04.011

Snow DL, Tebes JK, Arthur MW (1992) Panel attrition and external validity in adolescent substance use research. J Consult Clin Psychol 60(5):804–807

Snow WM, Connett JE, Sharma S, Murray RP (2007) Predictors of attendance and dropout at the Lung Health Study 11-year follow-up. Contemp Clin Trials 28(1):25–32. https://doi.org/10.1016/j.cct.2006.08.010

Spohr SA, Suzuki S, Marshall B, Taxman FS, Walters ST (2016) Social support quality and availability affects risk behaviors in offenders. Health Just 4(2):1–10. https://doi.org/10.1186/s40352-016-0033-y

Stevens T, Morash M, Chesney-Lind M (2011) Are girls getting tougher, or are we tougher on girls? Probability of arrest and juvenile court oversight in 1980 and 2000. Just Q 28(5):719–744. https://doi.org/10.1080/07418825.2010.532146

Stolzenberg RM, Relles DA (1990) Theory testing in a world of constrained research design: the significance of Heckman's censored sampling bias correction for nonexperimental research. Sociol Methods Res 18(4):395–415

Stolzenberg RM, Relles DA (1997) Tools for intuition about sample selection bias and its correction. Am Sociol Rev 62(3):494–507

Tabachnick BG, Fidell LS (2001) Using multivariate statistics, 4th edn. Allyn & Bacon, Needham Heights

Taxman FS (2017) Are you asking me to change my friends? Criminol Public Policy 16(3):775–782. https://doi.org/10.1111/1745-9133.12328

Taylor CJ (2015) Recent victimization and recidivism: the potential moderating effects of family support. Violence Vict 30(2):342–360

Thoits PA (1995) Stress, coping, and social support processes: Where are we? What next? J Health Soc Behav 35:53–79. https://doi.org/10.2307/2626957

Thornberry TP, Bjerregaard B, Miles W (1993) The consequences of respondent attrition in panel studies: a simulation based on the Rochester Youth Development Study. J Quant Criminol 9(2):127–158. https://doi.org/10.1007/BF01071165

Tucker JW (2010) Selection bias and econometric remedies in accounting and finance research. J Account Lit 29:31–57

Uggen C (2000) Work as a turning point in the life course of criminals: a duration model of age, employment, and recidivism. Am Sociol Rev 65(4):529–546

Van Buuren S, Boshuizen HC, Knook DL (1999) Multiple imputation of missing blood pressure covariates in survival analysis. Stat Med 18(6):681–694

Van de Ven W, Van Praag B (1981) The demand for deductibles in private health insurance: a probit model with sample selection. J Econom 17(2):229–252. https://doi.org/10.1016/0304-4076(81)90028-2

Vella F (1998) Estimating models with sample selection bias: a survey. J Hum Resour 33(1):127–169

Visher CA, Travis J (2003) Transitions from prison to community: understanding individual pathways. Ann Rev Sociol 29:89–113. https://doi.org/10.1146/annurev.soc.29.010202.095931

Wadsworth T, Roberts JM (2008) When missing data are not missing: a new approach to evaluating supplemental homicide report imputation strategies. Criminology 46(4):841–870. https://doi.org/10.1111/j.1745-9125.2008.00129.x

Western B (2018) Homeward: life in the year after prison. Russell Sage Foundation, New York

Western B, Braga AA, Kohl R (2017) A longitudinal survey of newly-released prisoners: methods and design of the Boston Reentry Study. Fed Probat 81(1):32–40. https://doi.org/10.3868/s050-004-015-0003-8

Wimberly AS, Engstrom M (2018) Stress, substance use, and yoga in the context of community reentry following incarceration. J Correct Health Care 24(1):96–103. https://doi.org/10.1177/1078345817726536

Winship C, Mare RD (1992) Models for sample selection bias. Ann Rev Sociol 18:327–350

Wolfolds SE, Siegel J (2019) Misaccounting for endogeneity: the peril of relying on the Heckman two-step method without a valid instrument. Strateg Manag J 40(3):432–462. https://doi.org/10.1002/smj.2995

Wong JS, Bouchard J, Gushue K, Lee C (2018) Halfway out: an examination of the effects of halfway houses on criminal recidivism. Int J Offender Ther Comp Criminol. https://doi.org/10.1177/0306624X18811964

Wooldridge JM (2010) Econometric analysis of cross section and panel data, 2nd edn. MIT Press, Cambridge

Young AF, Powers JR, Bell SL (2006) Attrition in longitudinal studies: Who do you lose? Aust N Z J Public Health 30(4):353–361. https://doi.org/10.1111/j.1467-842X.2006.tb00849.x

Young AF, Powers J, Wheway V (2007) Working with longitudinal data: Attrition and retention, data quality, measures of change and other analytical issues. Int J Multiple Res Approach 1(2):175–187

Yu B, Gastwirth JL (2010) How well do selection models perform? Assessing the accuracy of art auction pre-sale estimates. Stat Sin 20(2):837–852