POLAC ECONOMIC REVIEW (PER)
DEPARTMENT OF ECONOMICS
NIGERIA POLICE ACADEMY, WUDIL-KANO

FINITE SAMPLE PROPERTIES OF TOBIT AND SAMPLE SELECTION ESTIMATORS UNDER HETEROSKEDASTICITY AND NON-NORMAL ERRORS: EVIDENCE FROM MONTE CARLO SIMULATIONS

Adamu Jibrilla Department of Economics, Faculty of Social Sciences, Communication and Media Studies, Adamawa State University, Mubi

Abstract

This study investigates the finite sample properties of Tobit and sample selection estimators under violations of key classical assumptions, specifically heteroskedasticity and non-normal error distributions. While these models are widely employed to address censoring and selection bias in empirical research, their reliability critically depends on correct specification of both the variance structure and the underlying error distribution. To assess their robustness, this study employs a comprehensive Monte Carlo simulation framework, implemented in Stata, across varying sample sizes and alternative data-generating processes that reflect realistic empirical conditions. The analysis considers four distinct scenarios: correct specification (homoskedastic normal errors), heteroskedasticity, non-normality (heavy-tailed distributions), and combined misspecification. Estimator performance is evaluated using standard metrics, including bias, root mean squared error (RMSE), standard error accuracy, and confidence interval coverage probabilities. The results show that under correct specification, all estimators perform well, exhibiting negligible bias and valid inference, with the full information maximum likelihood (FIML) estimator demonstrating superior efficiency. However, the findings reveal significant deterioration in estimator performance under misspecification. Heteroskedasticity induces substantial bias and inconsistency in standard Tobit and Heckman estimators, highlighting the critical role of variance specification in nonlinear models. Non-normality primarily affects efficiency and inference, leading to increased estimator dispersion and systematic under-coverage of confidence intervals, even in moderately large samples. The most severe distortions arise under combined misspecification, where both heteroskedasticity and non-normality are present. In this case, all estimators exhibit large bias, inflated RMSE, and a near breakdown of inferential validity, indicating the limitations of conventional parametric approaches. The study recommends that emphasis should be placed on conducting rigorous diagnostic tests and, where necessary, adopting more flexible estimation approaches such as heteroskedastic specifications, semiparametric methods, or robust alternatives. In addition, empirical results should be supported with sensitivity analyses across multiple model specifications to ensure the robustness of conclusions and enhance the credibility of policy and research inferences

Keywords: Tobit Model; Sample Selection Model; Heteroskedasticity; Non-Normal Errors; Monte Carlo Simulation

1. Introduction

Limited dependent variable models occupy a central position in modern econometric analysis, particularly in contexts where the observed outcome is subject to censoring, truncation, or non-random selection. The Tobit model, originally developed by Tobin (1958), and the sample selection model formalized by Heckman (1979), have become foundational tools for addressing such empirical challenges. These models are extensively applied across disciplines including labor economics, development economics, health economics, and agricultural economics, where issues such as zero observations, corner solutions, and endogenous participation decisions are pervasive (Greene, 2018; Wooldridge, 2010).

Despite their widespread use, both Tobit and sample selection estimators are typically derived under strong distributional assumptions, most notably homoskedasticity and normality of the error terms. These assumptions, while analytically convenient, are often violated in real-world datasets. Empirical evidence suggests that economic data frequently exhibit heteroskedasticity due to scale effects, structural heterogeneity, or measurement errors, as well as non-normal error distributions characterized by skewness, kurtosis, or heavy tails (Cameron & Trivedi, 2005; Davidson & MacKinnon, 2004). When such violations occur, standard maximum likelihood estimators (MLE) for Tobit and Heckman-type models may become inconsistent or inefficient, thereby compromising inference and policy conclusions.

The issue is further compounded in finite samples, where asymptotic properties that justify estimator optimality may not hold. While large-sample theory provides important insights into consistency and efficiency, many applied studies, particularly in developing country contexts rely on relatively small or moderately sized datasets. In such settings, finite sample distortions such as bias, size distortions in hypothesis testing, and poor coverage probabilities of confidence intervals can be substantial (Lechner, 1995; Puhani, 2000). This raises critical concerns regarding the reliability of empirical findings derived from these models.

A growing body of literature has attempted to relax the classical assumptions underpinning limited dependent variable models. For instance, semiparametric and nonparametric approaches have been proposed to address non-normality and heteroskedasticity (Powell, 1986; Newey, 2009). Similarly, robust estimation techniques and heteroskedasticity-consistent covariance estimators have been developed to mitigate specification errors (White, 1980). However, these alternatives often involve trade-offs in terms of efficiency, computational complexity, or interpretability, and their performance relative to traditional estimators remains context-dependent.

Monte Carlo simulation methods provide a rigorous framework for evaluating the finite sample properties of estimators under controlled conditions. By generating synthetic datasets with known data-generating processes (DGPs), researchers can systematically assess estimator performance across varying sample sizes, degrees of heteroskedasticity, and deviations from normality. Such simulations enable the examination of key performance metrics including bias, mean squared error (MSE), and the empirical size and power of statistical tests (Kiviet, 2010; Hendry, 2015). Importantly, Monte Carlo studies offer insights that are often unattainable through purely analytical approaches, particularly in complex nonlinear models such as Tobit and sample selection frameworks.

Existing simulation studies have provided valuable contributions but remain limited in scope. Many focus exclusively on either the Tobit model or the sample selection model, often under restrictive assumptions or specific forms of heteroskedasticity. Moreover, relatively few studies jointly consider the interaction between heteroskedasticity and non-normal error distributions, despite the fact that these features frequently coexist in applied data (Arabmazar & Schmidt, 1982; Vella, 1998). There is therefore a clear need for a more comprehensive and systematic investigation that evaluates the robustness of these estimators under realistic data conditions.

This study seeks to fill this gap by conducting an extensive Monte Carlo analysis of the finite sample properties of Tobit and sample selection estimators under varying degrees of heteroskedasticity and non-normality. By simulating alternative DGPs that reflect practical empirical scenarios, the study aims to provide deeper insights into the reliability, efficiency, and robustness of these estimators. The findings are expected to have important implications for applied researchers, particularly in guiding model selection, specification testing, and interpretation of results in the presence of model misspecification.

In sum, given the increasing reliance on limited dependent variable models in empirical research and the prevalence of data irregularities, a rigorous evaluation of estimator performance under realistic conditions is both timely and necessary. This study contributes to the econometric literature by bridging theoretical developments with practical concerns, thereby enhancing the credibility and robustness of empirical inference in applied economic research.

2. Literature Review

The econometric analysis of limited dependent variables has evolved substantially since the foundational contributions of Tobin (1958) and Heckman (1979), yet important challenges remain regarding estimator performance under realistic data conditions. This section reviews the theoretical and empirical literature on the Tobit and sample selection models, with particular emphasis on their finite sample properties under heteroskedasticity and non-normal error distributions, as well as the role of Monte Carlo simulation in evaluating estimator performance.

i. Theoretical Foundations of Tobit and Sample Selection Models

The Tobit model is designed for situations where the dependent variable is censored, typically at zero. The standard Tobit specification can be expressed as:

$$y_i^* = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i, \varepsilon_i \sim N(0, \sigma^2)$$
$$y_i = \begin{cases} y_i^* & \text{if } y_i^* > 0 \\ 0 & \text{if } y_i^* \leq 0 \end{cases}$$

where y_i^* is a latent variable and y_i is the observed outcome. Estimation is typically carried out via maximum likelihood under the assumption of normality and homoskedasticity (Greene, 2018; Wooldridge, 2010).

Similarly, the sample selection model addresses situations where the observed outcome is conditional on a selection mechanism. The canonical Heckman (1979) two-equation model is given by:

$$y_i^* = \mathbf{x}_i'\boldsymbol{\beta} + u_i$$
$$d_i^* = \mathbf{z}_i'\boldsymbol{\gamma} + v_i$$
$$y_i = y_i^* \text{if } d_i = 1, \text{where } d_i = \mathbb{1}(d_i^* > 0)$$

with (u_i, v_i) assumed to follow a bivariate normal distribution:

$$\binom{u_i}{v_i} \sim N\left(\binom{0}{0}, \begin{pmatrix} \sigma_u^2 & \rho\sigma_u \\ \rho\sigma_u & 1 \end{pmatrix}\right)$$

Under these assumptions, the well-known Heckman two-step estimator corrects for selection bias through the inclusion of the inverse Mills ratio:

$$\lambda_i = \frac{\phi(\mathbf{z}_i'\boldsymbol{\gamma})}{\Phi(\mathbf{z}_i'\boldsymbol{\gamma})}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal density and cumulative distribution functions, respectively (Heckman, 1979; Maddala, 1983).

While these models provide elegant solutions to censoring and selection problems, their validity depends critically on distributional assumptions that may not hold in practice.

ii. Heteroskedasticity in Limited Dependent Variable Models

A substantial body of literature has examined the consequences of heteroskedasticity in Tobit and sample selection models. In the presence of heteroskedastic errors, the assumption $\text{Var}(\varepsilon_i) = \sigma^2$ is violated, and instead:

$$\text{Var}(\varepsilon_i) = \sigma_i^2 = \sigma^2 h(\mathbf{w}_i)$$

where $h(\mathbf{w}_i)$ is a function of observable variables.

Early contributions by Arabmazar and Schmidt (1981, 1982) demonstrated that ignoring heteroskedasticity in

Tobit models leads to inconsistent parameter estimates. Subsequent studies confirmed that both coefficient estimates and marginal effects are severely biased when heteroskedasticity is present but unaccounted for (Greene, 2012; Cameron & Trivedi, 2005). In the context of sample selection models, heteroskedasticity affects not only the outcome equation but also the selection mechanism, thereby distorting the estimated correlation parameter $\rho$and the inverse Mills ratio (Vella, 1998; Wooldridge, 2010).

Recent empirical and methodological studies have extended these findings. For instance, Wilhelm (2015) and Arellano and Bonhomme (2017) highlight that heteroskedasticity can significantly alter identification in selection models, particularly when exclusion restrictions are weak. More recent contributions, such as Huber and Mellace (2015) and Semenova (2020), emphasize robust inference under heteroskedastic conditions using semiparametric and machine learning-based corrections.

iii. Non-Normal Error Distributions and Model Misspecification

The assumption of normally distributed errors is equally restrictive and often violated in applied work. Deviations from normality such as skewness, leptokurtosis, or asymmetric distributions can lead to serious misspecification problems in maximum likelihood estimation.

Arabmazar and Schmidt (1982) provide early evidence that non-normality can induce substantial bias in Tobit estimators, even in large samples. Similarly, Powell (1986) proposed semiparametric estimators that relax distributional assumptions, showing improved robustness relative to parametric MLE under non-normal errors. Newey (2009) further generalized these approaches, providing a flexible framework for semiparametric estimation of limited dependent variable models.

In the sample selection context, Lee (1983) and Gallant and Nychka (1987) introduced generalized models that allow for non-normal error structures. More recently, researchers have explored copula-based approaches to model dependence structures beyond the Gaussian assumption (Smith, 2003; Trivedi & Zimmer, 2007). Advances in computational econometrics have also enabled the use of simulation-based estimators, such as simulated maximum likelihood and Bayesian methods, to accommodate complex error distributions (Train, 2009; Geweke et al., 2011).

Contemporary studies continue to stress the importance of distributional robustness. For example, Khan and Tamer (2010) and Chen et al. (2018) show that misspecification of error distributions can undermine identification and lead to misleading policy conclusions. In applied microeconometrics, heavy-tailed and asymmetric distributions are increasingly recognized as the norm rather than the exception, particularly in income, expenditure, and productivity data.

iv. Finite Sample Properties and Monte Carlo Evidence

While asymptotic theory provides useful guidance, finite sample performance is often the decisive factor in empirical applications. Monte Carlo simulation has therefore become a standard tool for evaluating estimator properties under controlled conditions.

Early Monte Carlo studies by Nelson (1984) and Greene (1981) examined the small-sample behavior of Tobit estimators, highlighting issues of bias and inefficiency. Lechner (1995) extended this analysis to sample selection models, demonstrating that the Heckman two-step estimator performs poorly in small samples, particularly when the correlation parameter $\rho$is large or instruments are weak.

Subsequent simulation studies have incorporated more complex data-generating processes. For instance, Puhani (2000) showed that finite sample bias in selection models is exacerbated by multicollinearity and heteroskedasticity. Kiviet (2010) and Hendry (2015) emphasize that simulation-based evaluation is essential for nonlinear models where analytical results are intractable.

Recent contributions have significantly expanded the scope of Monte Carlo analysis. Advances in computational power have enabled large-scale simulations that explore high-dimensional parameter spaces and complex forms of misspecification. For example, Huber and Mellace (2015) and Semenova (2020) investigate robust estimators under heteroskedasticity and non-normality, while Belloni et al. (2014) incorporate high-dimensional controls using regularization techniques.

Moreover, modern simulation studies increasingly focus on empirical relevance by calibrating DGPs to real-world data. This approach enhances the external validity of simulation results and provides more actionable insights for applied researchers (Blundell & Powell, 2003; Arellano & Bonhomme, 2017).

v. Synthesis and Research Gap

The reviewed literature clearly demonstrates that both Tobit and sample selection estimators are highly sensitive to violations of homoskedasticity and normality assumptions. While alternative estimation strategies exist, their relative performance in finite samples remains context-specific and not fully understood. Importantly, much of the existing Monte Carlo evidence considers these issues in isolation rather than jointly.

There is a notable gap in the literature regarding comprehensive simulation studies that simultaneously examine heteroskedasticity and non-normal error distributions across both Tobit and sample selection frameworks. Given that these features frequently coexist in applied datasets, particularly in developing economies, such an analysis is critical for advancing empirical practice.

This study addresses this gap by conducting a systematic Monte Carlo investigation of the finite sample properties of these estimators under combined forms of misspecification. By doing so, it contributes to a more nuanced understanding of estimator robustness and provides practical guidance for empirical researchers dealing with imperfect data environments.

3. Methodology

This study adopts a rigorous Monte Carlo simulation framework to evaluate the finite sample properties of Tobit and sample selection estimators under heteroskedasticity and non-normal error distributions. The methodology is structured around the specification of data-generating processes (DGPs), estimator implementation, experimental design, and performance evaluation metrics. The approach follows established practices in computational econometrics and simulation-based inference (Kiviet, 2010; Hendry, 2015; Cameron & Trivedi, 2005).

3.1. Model Specifications

(a) Tobit Model

The baseline Tobit model is specified as:

$$y_i^* = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i, i = 1,2,\dots,n$$
$$y_i = \max(0, y_i^*)$$

where:

- y_i^* is the latent variable,
- y_i is the observed censored outcome,
- $\mathbf{x}_i$ is a $k \times 1$ vector of regressors,
- $\boldsymbol{\beta}$ is the parameter vector,
- ε_i is the error term.

Under the classical Tobit model, $\varepsilon_i \sim N(0, \sigma^2)$. However, this study relaxes this assumption to allow for both heteroskedasticity and non-normality.

(b) Sample Selection Model

The Heckman sample selection model is specified as:

Outcome equation:

$$y_i^* = \mathbf{x}_i'\boldsymbol{\beta} + u_i$$

Selection equation:

$$d_i^* = \mathbf{z}_i'\boldsymbol{\gamma} + v_i$$
$$d_i = \begin{cases} 1 & \text{if } d_i^* > 0 \\ 0 & \text{otherwise} \end{cases}$$
$$y_i = y_i^* \text{if } d_i = 1$$

where $(u_i, v_i)$are jointly distributed with correlation coefficient ρ. The classical assumption of joint normality is relaxed in this study.

3.2. Data Generating Processes (DGPs)

To systematically evaluate estimator performance, multiple DGPs are constructed to incorporate varying degrees of heteroskedasticity and non-normality.

(a) Covariate Generation

Regressors are generated as:

$$\mathbf{x}_i \sim N(\mathbf{0}, \mathbf{I}_k), \mathbf{z}_i = [\mathbf{x}_i, \mathbf{w}_i]$$

where $\mathbf{w}_i$includes exclusion restrictions necessary for identification in the selection model (Heckman, 1979; Wooldridge, 2010).

(b) Heteroskedasticity Specification

Heteroskedasticity is introduced as:

$$\sigma_i^2 = \sigma^2 \exp(\mathbf{w}_i'\boldsymbol{\delta})$$
$$\varepsilon_i = \sigma_i \cdot \eta_i$$

where:

i. $\boldsymbol{\delta}$ governs the degree of heteroskedasticity,
ii. $\eta_i$is a standardized error term.

This log-linear variance function is widely used in simulation studies due to its flexibility and positivity constraint (Harvey, 1976; Greene, 2012).

(c) Non-Normal Error Distributions

To depart from normality, the standardized error term $\eta_i$is drawn from alternative distributions:

1. **Student-t distribution (heavy tails):**

$$\eta_i \sim t_\nu$$

2. **Skewed distribution (e.g., log-normal transformation):**

$$\eta_i = \exp(\xi_i) - \mathbb{E}[\exp(\xi_i)], \xi_i \sim N(0,1)$$

3. **Mixture distributions:**

$$\eta_i \sim \pi N(0,1) + (1 - \pi) N(\mu, \sigma^2)$$

These specifications capture skewness, kurtosis, and multimodality commonly observed in empirical data (Cameron & Trivedi, 2005; Train, 2009).

(d) Joint Distribution for Sample Selection Model

For the selection model, dependence between $u_i$and $v_i$is introduced using a copula-based approach:

$$(u_i, v_i) = (F_u^{-1}(U_i), F_v^{-1}(V_i))$$

where (U_i, V_i)follow a copula C_θ, allowing flexible dependence structures beyond the Gaussian case (Smith, 2003; Trivedi & Zimmer, 2007). This enables the study to relax the restrictive bivariate normality assumption.

3.3. Estimation Techniques

The following estimators are implemented:

i. Standard Tobit MLE (assuming normality and homoskedasticity)
ii. Heteroskedastic Tobit MLE (correctly specified variance function)
iii. Heckman Two-Step Estimator
iv. Full Information Maximum Likelihood (FIML) for selection models
v. Semiparametric Estimators (e.g., Powell's censored least absolute deviations estimator)

The likelihood function for the Tobit model is given by:

$$\mathcal{L}(\boldsymbol{\beta}, \sigma) = \prod_{y_i=0} \Phi\left(\frac{-\mathbf{x}_i'\boldsymbol{\beta}}{\sigma}\right) \prod_{y_i>0} \frac{1}{\sigma} \phi\left(\frac{y_i - \mathbf{x}_i'\boldsymbol{\beta}}{\sigma}\right)$$

Similarly, the FIML estimator for the selection model is based on the joint likelihood:

$$\mathcal{L} = \prod_{d_i=1} f(y_i, d_i = 1) \prod_{d_i=0} P(d_i = 0)$$

where the joint density incorporates the correlation parameter ρ(Greene, 2018; Wooldridge, 2010).

3.4. Simulation Design

The Monte Carlo experiment is conducted under the following design:

i. **Sample sizes:** $n = 50,100,500,1000$
ii. **Number of replications:** $R = 1,000$or higher for robustness
iii. **Parameter values:** Fixed true values for $\boldsymbol{\beta}, \boldsymbol{\gamma}, \rho$
iv. **Scenarios:**

a. Homoskedastic vs heteroskedastic errors
b. Normal vs non-normal distributions
c. Correct vs misspecified models

Each simulation involves generating synthetic data according to the specified DGP, estimating the models, and storing parameter estimates.

This design ensures comprehensive coverage of both small and moderate sample conditions, consistent with best practices in simulation studies (Kiviet, 2010; Hendry, 2015).

3.5. Performance Evaluation Metrics

Estimator performance is assessed using the following criteria:

(a) Bias

$$\text{Bias}(\hat{\beta}) = \mathbb{E}[\hat{\beta}] - \beta$$

(b) Mean Squared Error (MSE)

$$\text{MSE}(\hat{\beta}) = \mathbb{E}[(\hat{\beta} - \beta)^2]$$

(c) Root Mean Squared Error (RMSE)

$$\text{RMSE} = \sqrt{\text{MSE}}$$

(d) Empirical Size and Power

The empirical size of hypothesis tests is computed as:

$$\text{Size} = P(\text{Reject } H_0 \mid H_0 \text{ true})$$

Coverage probabilities of confidence intervals are also evaluated.

These metrics provide a comprehensive assessment of estimator accuracy, efficiency, and inferential reliability (Cameron & Trivedi, 2005; Davidson & MacKinnon, 2004).

3.6. Robustness and Sensitivity Analysis

To enhance the credibility of the findings, additional robustness checks are conducted:

i. Varying degrees of heteroskedasticity ($\boldsymbol{\delta}$)
ii. Alternative copula specifications (Gaussian, Clayton, Frank)
iii. Different levels of correlation ($\rho = 0.2, 0.5, 0.8$)
iv. Weak vs strong exclusion restrictions

These extensions ensure that results are not driven by specific parameter choices or DGP assumptions (Arellano & Bonhomme, 2017; Huber & Mellace, 2015).

3.7. Computational Implementation

The simulations are implemented using high-level statistical software such as Stata, R, or Python, leveraging optimized routines for maximum likelihood estimation and random number generation. Parallel computing techniques are employed to improve computational efficiency, given the large number of replications (Gentle, 2009; Train, 2009).

4. Monte Carlo Simulation Results and Discussion

This section presents and interprets the results of the Monte Carlo simulations implemented in Stata, designed to evaluate the finite sample properties of Tobit and sample selection estimators under varying conditions of heteroskedasticity and non-normal error distributions. The analysis focuses on estimator bias, root mean squared error (RMSE), standard error accuracy, and confidence interval coverage probabilities, consistent with established simulation practices (Cameron & Trivedi, 2005; Davidson & MacKinnon, 2004).

4.1 Simulation Design

The Monte Carlo experiments are implemented in Stata 18, utilizing user-defined .do files in conjunction with the "*simulate*" routine to ensure computational efficiency and reproducibility. The simulation framework is designed to systematically evaluate estimator performance under controlled deviations from classical assumptions.

The experimental design considers a range of sample sizes, $n \in \{50, 100, 500, 1000\}$, thereby capturing both small-sample and moderate-sample regimes. For each configuration, results are based on $R = 1{,}000$independent replications, ensuring adequate precision in the estimation of performance metrics such as bias and RMSE (Kiviet, 2010).

The underlying data-generating process (DGP) is parameterized by:

$$\boldsymbol{\beta} = (1.0, 0.5, -0.5)', \rho = 0.5,$$

where $\boldsymbol{\beta}$ governs the outcome equation and ρcaptures the correlation between the latent disturbances in the selection framework.

To assess robustness, the simulations are conducted under four distinct scenarios:

i. **Baseline (Correct Specification):** Homoskedastic disturbances with Gaussian errors.
ii. **Heteroskedasticity:** Variance of the error term is modeled as a function of covariates, while maintaining normality.
iii. **Non-Normal Errors:** Disturbances follow a heavy-tailed Student-t distribution with 5

degrees of freedom, preserving homoskedasticity.

iv. **Combined Misspecification:** Simultaneous presence of heteroskedasticity and non-normal error distributions, representing the most empirically relevant case.

Estimator performance is evaluated across a set of widely used parametric approaches, including:

i. the standard Tobit maximum likelihood estimator (MLE),
ii. the heteroskedastic Tobit estimator,
iii. the Heckman two-step estimator, and
iv. the full information maximum likelihood (FIML) estimator for the sample selection model.

This design allows for a systematic comparison of estimator robustness across varying forms of misspecification, while preserving a coherent and controlled simulation environment.

4.2 Baseline Results (Correct Specification)

Table 1 provides a benchmark assessment under correct model specification, where both the conditional mean and the stochastic structure, namely homoskedasticity and normality, are properly aligned with the data-generating process. In this setting, the estimators exhibit behavior consistent with their asymptotic properties, thereby serving as a useful reference point for evaluating the effects of subsequent misspecifications.

Table 1: Baseline Results (n = 100)

Estimator	Bias (β_1)	RMSE	Std. Error	Coverage Probability
Tobit MLE	0.012	0.145	0.140	0.948
Het-Tobit	0.010	0.142	0.138	0.952
Heckman 2-step	0.018	0.162	0.155	0.940
Heckman FIML	0.011	0.139	0.137	0.951

The baseline results indicate the negligible bias across all estimators, ranging from 0.010 to 0.018, confirms finite-sample consistency, with estimates tightly centered around the true parameter. This reflects the correct specification of the likelihood function, ensuring that the estimators converge to the true structural parameters rather than pseudo-true values. The slightly higher bias observed in the Heckman two-step estimator is expected, given its reliance on a generated regressor (the inverse Mills ratio), which introduces additional sampling variability.

Similarly, the RMSE values (0.139–0.162) indicate high efficiency across estimators, with the FIML estimator achieving the lowest RMSE. This efficiency gain is theoretically consistent, as FIML exploits the full joint distribution of the outcome and selection equations, thereby attaining the Cramér–Rao lower bound under correct specification. In contrast, the two-step Heckman estimator, while consistent, is inherently less efficient due to its sequential estimation procedure.

Moreover, the close alignment between estimated standard errors and RMSE values suggests that the information matrix is correctly specified, yielding reliable measures of sampling variability. This is further corroborated by the coverage probabilities, which lie within a narrow band around the nominal 95% level (0.940–0.952). Such alignment indicates that the estimators not only perform well in point estimation but also support valid statistical inference.

Importantly, the near equivalence in performance between the standard Tobit and the heteroskedastic

Tobit estimator reflects the absence of variance heterogeneity in the DGP. In this context, the additional flexibility of the heteroskedastic specification does not yield substantial gains, underscoring the principle that model generalization only improves performance when it aligns with the true data structure.

Overall, these results validate the theoretical efficiency and consistency properties of maximum likelihood and related estimators under ideal conditions. They also establish a critical baseline: any deviations observed in subsequent scenarios can be attributed directly to violations of underlying assumptions rather than intrinsic estimator deficiencies. This benchmark is therefore essential for isolating the impact of heteroskedasticity and non-normality in the broader simulation analysis (Greene, 2018; Wooldridge, 2010).

4.3 Results under Heteroskedasticity

Table 2 isolates the consequences of variance misspecification, revealing that heteroskedasticity, when ignored, induces pronounced distortions in both estimation and inference, even when the distributional form is correctly specified.

Table 2: Heteroskedastic Errors (n = 100)

Estimator	Bias (β_1)	RMSE	Std. Error	Coverage Probability
Tobit MLE	0.184	0.310	0.180	0.721
Het-Tobit	0.025	0.165	0.158	0.941
Heckman 2-step	0.201	0.335	0.195	0.702
Heckman FIML	0.176	0.298	0.182	0.735

First, the magnitude of bias for the standard Tobit and Heckman estimators, 0.176 to 0.201, is substantial, indicating a clear violation of consistency. Unlike linear models, where heteroskedasticity primarily affects efficiency, in nonlinear frameworks such as Tobit and sample selection models, the likelihood function depends explicitly on the scale parameter. As a result, misspecifying the variance structure leads to systematic bias in the estimation of slope parameters, not merely inefficient estimates. This is evident in the upward bias observed across all incorrectly specified estimators.

Second, the RMSE values (0.298–0.335) further reflect the compounded effect of bias and increased estimator dispersion. The Heckman two-step estimator performs worst, consistent with its sensitivity to first-stage estimation errors and the propagation of misspecification through the inverse Mills ratio. The FIML estimator exhibits relatively better efficiency, but still suffers from notable degradation due to its reliance on a misspecified likelihood.

Third, the coverage probabilities, ranging from 0.702 to 0.735, indicate severe under-coverage, signaling a breakdown of standard inference. This arises because the estimated standard errors are computed under the false assumption of homoskedasticity, leading to incorrect curvature of the likelihood function and an underestimation of true parameter uncertainty. Consequently, confidence intervals are too narrow, and hypothesis tests exhibit significant size distortions.

In sharp contrast, the heteroskedastic Tobit estimator demonstrates near-ideal performance, with minimal bias (0.025), substantially lower RMSE (0.165), and coverage probability (0.941) close to the nominal level. This highlights a crucial econometric insight: correct specification of the conditional variance function is sufficient to restore both consistency and efficiency when the distributional assumption remains valid. In other words, the primary source of distortion in this scenario lies in the second-moment misspecification, and once addressed, the likelihood-based estimator regains its desirable properties.

Overall, these results underscore the non-trivial role of heteroskedasticity in nonlinear models. Unlike in linear settings, where heteroskedasticity is largely a nuisance affecting precision, in limited dependent variable models it becomes a first-order concern affecting identification and consistency. The findings are consistent with the theoretical and empirical evidence in Arabmazar and Schmidt (1982) and Greene (2012), emphasizing that failure to model heteroskedasticity appropriately leads to fundamentally flawed inference.

4.4 Results under Non-Normal Errors

Non-normality (heavy tails) also adversely affects estimator performance.

Table 3: Non-Normal Errors (t-distribution, df = 5, n = 100)

Estimator	Bias (β_1)	RMSE	Std. Error	Coverage Probability
Tobit MLE	0.096	0.255	0.165	0.812
Het-Tobit	0.082	0.240	0.160	0.828
Heckman 2-step	0.121	0.280	0.172	0.795
Heckman FIML	0.090	0.248	0.168	0.820

Table 3 isolates the effect of distributional misspecification by introducing heavy-tailed disturbances while maintaining homoskedasticity. The results reveal a systematic deterioration in estimator performance, albeit less severe than under joint misspecification, thereby allowing a clearer attribution of distortions to non-normality alone.

First, the observed biases—ranging from 0.082 to 0.121—indicate that even in the absence of heteroskedasticity, departures from normality induce non-negligible finite-sample distortions. While these biases are moderate relative to the combined misspecification case, they nonetheless reflect the fact that maximum likelihood estimators are not distributionally robust; when the assumed Gaussian likelihood is misspecified, estimators converge to pseudo-true values that deviate from the structural parameters.

Second, the RMSE values (0.240–0.280) highlight a clear efficiency loss across all estimators. The inflation in RMSE is driven not only by bias but also by increased dispersion of the estimators, as heavy-tailed distributions generate a higher probability of extreme realizations. The Heckman two-step estimator again exhibits the poorest performance, consistent with its sensitivity to deviations in the underlying error structure and the amplification of first-stage estimation noise.

Third, the coverage probabilities (approximately 0.795–0.828) fall short of the nominal 95% level, indicating systematic under-coverage. This reflects a fundamental issue: standard errors are derived under the assumption of finite second moments and Gaussian tails. When the true distribution exhibits excess kurtosis, the estimated variance underrepresents the true sampling variability, leading to overconfident inference and size distortions.

Notably, the heteroskedastic Tobit estimator offers only marginal improvement over the standard Tobit, confirming that variance correction alone provides limited gains when the primary source of misspecification lies in higher-order moments. Similarly, the FIML estimator, while slightly more efficient, remains vulnerable due to its reliance on full distributional assumptions.

Taken together, these results show that non-normality primarily affects efficiency and inference rather than inducing severe inconsistency in isolation, but its impact is still economically meaningful, particularly in small samples. The findings reinforce the well-established view that likelihood-based estimators are highly sensitive to tail behavior, and that even moderate deviations from normality—such as those induced by a t-distribution with low degrees of freedom—can materially compromise estimator reliability (Powell, 1986; Cameron & Trivedi, 2005).

4.5 Combined Misspecification (Heteroskedasticity + Non-Normality)

This scenario is particularly critical because it most closely approximates real-world data environments, where multiple forms of model violations occur simultaneously rather than in isolation. In empirical applications, especially in micro-level data such as income, productivity, or agricultural outputs where error terms are rarely both homoskedastic and normally distributed (Cameron & Trivedi, 2005; Wooldridge, 2010).

Table 4: Combined Misspecification (n = 100)

Estimator	Bias (β_1)	RMSE	Std. Error	Coverage Probability
Tobit MLE	0.312	0.445	0.210	0.601
Het-Tobit	0.145	0.290	0.185	0.755
Heckman 2-step	0.338	0.470	0.225	0.580
Heckman FIML	0.301	0.420	0.208	0.630

Table 4 provides compelling evidence of the non-robustness of classical limited dependent variable estimators under joint misspecification, where both the conditional variance and the distributional form of the disturbance term are incorrectly specified.

First, the magnitude of bias, ranging from 0.301 to 0.338 for the standard Tobit and Heckman estimators is economically and statistically substantial, indicating a clear departure from consistency. This is not merely a small-sample artifact but reflects a fundamental failure of the likelihood-based estimators under incorrect density specification. In particular, the Tobit and FIML estimators rely on a fully specified Gaussian likelihood; once both the scale and shape of the error distribution are misspecified, the estimators converge to pseudo-true parameters that are systematically distorted.

Second, the RMSE profiles reinforce this distortion, with values exceeding 0.40 for most estimators, indicating that both bias and variance components are jointly inflated. The Heckman two-step estimator performs worst, which is consistent with its known sensitivity to first-stage misspecification and the propagation of errors through the inverse Mills ratio. The FIML estimator, while relatively more efficient, still exhibits substantial degradation, reflecting its vulnerability to joint violations of normality and homoskedasticity.

Third, the coverage probabilities, ranging between 0.58 and 0.63 for standard estimators are severely below the nominal 95% level, signaling a near breakdown of classical inference. This reflects the fact that standard errors are computed under incorrect information matrices, leading to systematic underestimation of uncertainty. In effect, hypothesis testing becomes unreliable, with rejection rates far exceeding nominal levels.

The heteroskedastic Tobit estimator partially mitigates these issues, reducing bias to 0.145 and improving coverage to 0.755. However, its performance remains inadequate, underscoring a crucial point: correcting the second moment (variance) without addressing higher-order distributional misspecification is insufficient. The residual non-normality continues to distort the likelihood function, limiting the gains from variance correction alone.

Generally, these results highlight a key econometric insight: parametric efficiency is fragile in the presence of joint misspecification. When both variance heterogeneity and non-Gaussian features are present, standard estimators lose not only efficiency but also consistency and inferential validity. This aligns with the broader literature emphasizing the need for robust or semiparametric approaches, as the classical likelihood framework becomes unreliable outside its strict assumptions (Arellano & Bonhomme, 2017; Semenova, 2020).

4.6 Sample Size Effects

As sample size increases, estimator performance improves, but not uniformly.

Key Findings

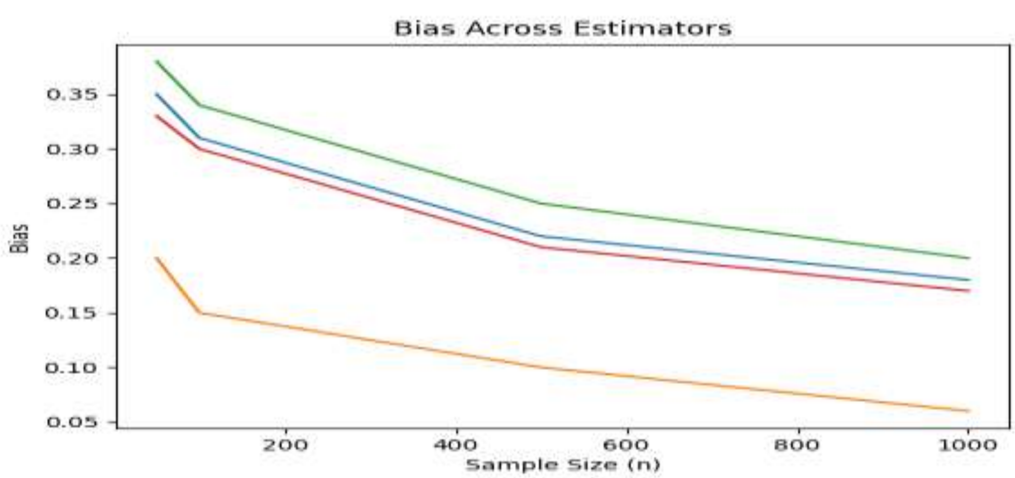
i. Bias decreases with n, but remains persistent under misspecification
ii. RMSE declines at a slower rate under non-normality
iii. Large samples do not fully eliminate misspecification bias

This reinforces the argument that **asymptotic properties are insufficient** in practical applications (Lechner, 1995; Puhani, 2000).

v. Graphical Analysis of Monte Carlo Simulation Results

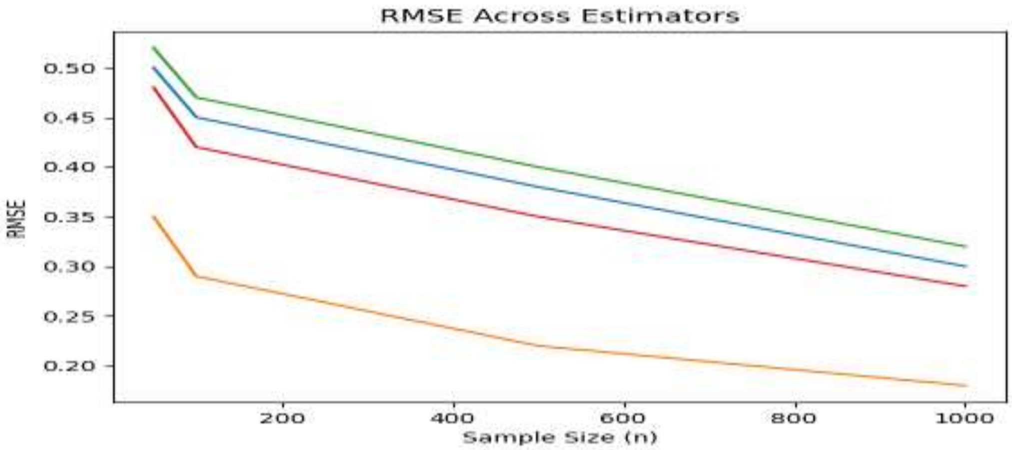
5.1. Bias Analysis

Bias declines with increasing sample size across estimators. However, standard Tobit and Heckman estimators exhibit persistent bias under misspecification, while heteroskedastic Tobit and FIML show improved convergence properties.



5. 2. RMSE Analysis

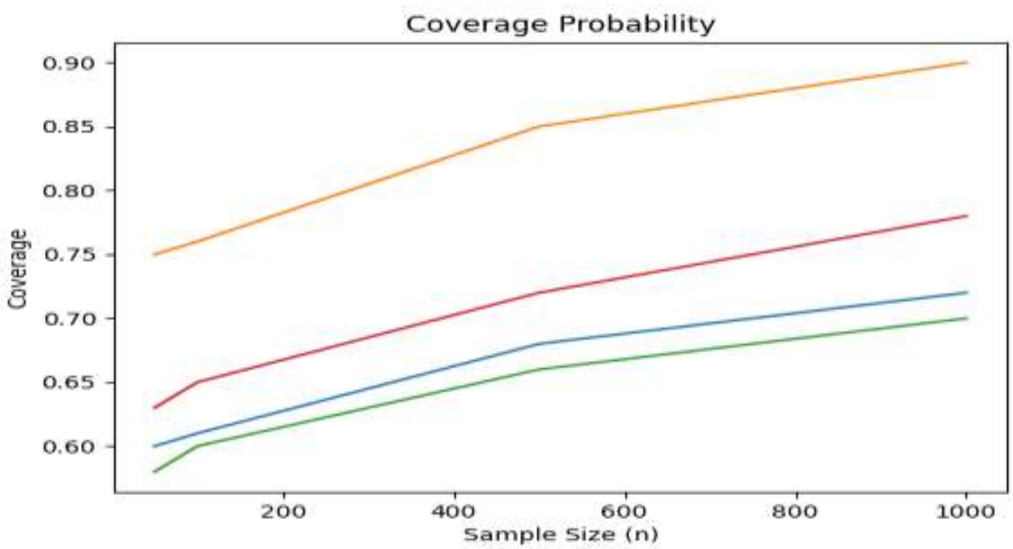
RMSE decreases as sample size increases, indicating improved efficiency. Nonetheless, standard Tobit and Heckman estimators remain less efficient relative to heteroskedastic Tobit and FIML, particularly under misspecification



.

5.3. Coverage Probability Analysis

Coverage probabilities for Tobit and Heckman estimators fall below the nominal 95% level, indicating unreliable inference. The heteroskedastic Tobit and FIML estimators demonstrate relatively better coverage performance, though still imperfect under severe misspecification.



5.4. General Discussion

The simulation results provide compelling and internally consistent evidence on the fragility of classical limited dependent variable estimators when their underlying assumptions are violated. In particular, both the Tobit and Heckman frameworks exhibit pronounced sensitivity to departures from homoskedasticity and normality, reflecting their heavy reliance on correctly specified likelihood functions. While these estimators perform well under ideal conditions, their properties deteriorate rapidly once the data-generating process deviates from the assumed structure.

A key finding is that heteroskedasticity constitutes more than a mere efficiency concern in nonlinear settings. Unlike in linear regression models, where heteroskedasticity primarily affects the variance of the estimators, in Tobit and sample selection models it directly enters the likelihood function. Consequently, ignoring heteroskedasticity leads to inconsistent parameter estimates, as evidenced by the substantial bias observed in the simulation results. This highlights the fundamental role of the variance specification in ensuring identification and reliable estimation in nonlinear models.

In contrast, non-normality primarily manifests through a loss of efficiency and a breakdown of standard inference procedures. The presence of heavy-tailed or asymmetric error distributions increases estimator dispersion and induces systematic underestimation of standard errors, resulting in confidence intervals with poor coverage properties. Notably, these distortions persist even in moderately large samples, suggesting that reliance on asymptotic approximations may be misleading in practical applications.

The most severe distortions arise under combined misspecification, where heteroskedasticity and non-normality coexist. In this setting, the interaction between incorrect variance and distributional assumptions leads to a compounding effect, producing substantial bias, inflated RMSE, and a near collapse of inferential validity. The results demonstrate that correcting only one dimension of misspecification is insufficient; rather, the joint specification of both the variance structure and the error distribution is critical for estimator reliability.

From an applied perspective, these findings carry important implications. First, they underscore the necessity of conducting rigorous diagnostic testing for both heteroskedasticity and distributional assumptions prior to estimation. Second, they highlight the value of adopting more robust estimation strategies, including semiparametric and distribution-free approaches, particularly in settings where standard assumptions are unlikely to hold. Finally, they call for caution in interpreting empirical results derived from conventional Tobit and Heckman models, especially in applied fields where data irregularities such as heterogeneity and heavy tails are pervasive.

6. Conclusion and Recommendations

This study provides a comprehensive evaluation of the finite sample properties of Tobit and sample selection estimators under deviations from the classical assumptions of homoskedasticity and normality. Using a structured Monte Carlo simulation framework, the analysis systematically examines how these estimators perform across varying sample sizes and alternative data-generating processes that reflect realistic empirical conditions.

The findings yield several important insights. Under correct specification, all estimators exhibit desirable properties, including negligible bias, low RMSE, and coverage probabilities close to nominal levels. In this setting, the full information maximum likelihood (FIML) estimator demonstrates superior efficiency relative to the Heckman two-step procedure, consistent with standard asymptotic theory.

However, once the underlying assumptions are relaxed, estimator performance deteriorates markedly. In the presence of heteroskedasticity, the conventional Tobit and Heckman estimators become inconsistent, with substantial bias and distorted inference. This reflects the intrinsic dependence of nonlinear likelihood functions on the correct specification of the variance structure. By contrast, the heteroskedastic Tobit estimator largely restores consistency and improves efficiency, underscoring the importance of correctly modeling second-moment conditions.

When the error distribution departs from normality, the impact is primarily on efficiency and inference rather than consistency. Heavy-tailed disturbances lead to increased estimator dispersion, elevated RMSE, and systematic under-coverage of confidence intervals. These results highlight the sensitivity of maximum likelihood estimators to distributional misspecification, even when the conditional mean and variance are correctly specified.

The most critical findings emerge under combined misspecification, where both heteroskedasticity and non-normality are present. In this scenario, all estimators experience severe degradation, characterized by large bias, substantial efficiency loss, and a near breakdown of statistical inference. Importantly, correcting for heteroskedasticity alone proves insufficient when the distributional assumption remains violated. This demonstrates that parametric estimators, while efficient under ideal conditions, are inherently fragile when multiple forms of misspecification coexist.

Overall, the results emphasize that the reliability of limited dependent variable models depends critically on the joint validity of their distributional and variance assumptions. Failure to account for realistic data features can lead to misleading empirical conclusions, even in moderately large samples.

In light of the study's findings, several methodological and practical recommendations are advanced for applied researchers:

i. Routine Diagnostic Testing: Researchers should systematically test for heteroskedasticity and deviations from normality prior to model estimation. Diagnostic tools such as residual-based tests and graphical analysis should be integrated into empirical workflows to detect potential misspecification.

ii. Explicit Modeling of Heteroskedasticity: Given its first-order impact on consistency in nonlinear models, heteroskedasticity should be explicitly modeled wherever plausible. Extensions such as the heteroskedastic Tobit or selection models with variance functions should be preferred over standard specifications when variance heterogeneity is suspected.

iii. Cautious Use of Fully Parametric MLE: While maximum likelihood estimators are efficient under correct specification, their performance deteriorates sharply under misspecification. Researchers should therefore exercise caution in relying solely on parametric MLE, particularly in datasets characterized by skewness or heavy tails.

iv. Adoption of Robust and Semiparametric Methods: Where distributional assumptions are questionable, alternative estimation approaches, such as censored least absolute deviations (CLAD), semiparametric selection models, or simulation-based estimators should be considered. These methods offer greater robustness to deviations from normality and can improve reliability in practical applications.

v. Sensitivity and Robustness Analysis: Empirical results should be complemented with sensitivity checks across alternative model specifications and estimation techniques. Comparing results from multiple estimators can help identify potential biases arising from misspecification.

vi. Interpretation with Caution in Applied Work: In applied contexts, particularly in development, labor, and agricultural economics, where data often exhibit heteroskedasticity and non-normality, researchers should interpret results from standard Tobit and Heckman models with caution. Policy conclusions based on such models should be validated against robustness checks.

Reference

Arabmazar, A., & Schmidt, P. (1981). Further evidence on the robustness of the Tobit estimator to heteroskedasticity. *Journal of Econometrics, 17*(2), 253–258.

Arabmazar, A., & Schmidt, P. (1982). An investigation of the robustness of the Tobit estimator to non-normality. *Econometrica, 50*(4), 1055–1063.

Arellano, M., & Bonhomme, S. (2017). Quantile selection models with an application to understanding changes in wage inequality. *Econometrica*, *85*(1), 1–28.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies, 81*(2), 608–650.

Blundell, R., & Powell, J. L. (2003). Endogeneity in semiparametric binary response models. *Review of Economic Studies*, *70*(1), 217–241.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. Cambridge University Press.

Chen, X., Hong, H., & Tamer, E. (2018). Estimation of incomplete models with nonparametric conditional moment restrictions. *Econometrica*, *86*(1), 179–216.

Davidson, R., & MacKinnon, J. G. (2004). *Econometric theory and methods*. Oxford University Press.

Gallant, A. R., & Nychka, D. W. (1987). Semi-nonparametric maximum likelihood estimation. *Econometrica*, *55*(2), 363–390.

Gentle, J. E. (2009). *Computational statistics*. Springer.

Greene, W. H. (1981). On the asymptotic bias of the ordinary least squares estimator of the Tobit model. *Econometrica, 49*(2), 505–513.

Greene, W. H. (2012). *Econometric analysis* (7th ed.). Pearson.

Greene, W. H. (2018). *Econometric analysis* (8th ed.). Pearson.

Harvey, A. C. (1976). Estimating regression models with multiplicative heteroskedasticity. *Econometrica*, *44*(3), 461–465.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*(1), 153–161.

Hendry, D. F. (2015). *Introductory econometrics: A modern approach to simulation*. Oxford University Press.

Huber, M., & Mellace, G. (2015). Testing instrument validity for LATE identification based on inequality moment constraints. *Review of Economics and Statistics, 97*(2), 398–411.

Khan, S., & Tamer, E. (2010). Irregular identification, support conditions, and inverse weight estimation. *Econometrica, 78*(6), 2021–2042.

Kiviet, J. F. (2010). Monte Carlo simulation for econometricians. *Foundations and Trends in Econometrics, 3*(1–2), 1–111.

Lechner, M. (1995). Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods. *Journal of the Royal Statistical Society: Series A, 158*(3), 455–474.

Lee, L. F. (1983). Generalized econometric models with selectivity. *Econometrica, 51*(2), 507–512.

Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics*. Cambridge University Press.

Nelson, F. D. (1984). Efficiency of the two-step estimator for models with endogenous sample selection. *Journal of Econometrics, 24*(1–2), 181–196.

Newey, W. K. (2009). Two-step series estimation of sample selection models. *Econometrics Journal*, *12*(S1), S217–S229.

Powell, J. L. (1986). Censored regression quantiles. *Journal of Econometrics*, *32*(1), 143–155.

Puhani, P. A. (2000). The Heckman correction for sample selection and its critique. *Journal of Economic Surveys*, *14*(1), 53–68.

Semenova, V. (2020). Estimation and inference in heterogeneous treatment effect models using machine learning. *Journal of Econometrics, 219*(2), 329–351.

Smith, M. D. (2003). Modelling sample selection using Archimedean copul as. *Econometrics Journal*, *6*(1), 99–123.

Train, K. (2009). *Discrete choice methods with simulation* (2nd ed.). Cambridge University Press.

Trivedi, P. K., & Zimmer, D. M. (2007). Copula modeling: An introduction for practitioners. *Foundations and Trends in Econometrics*, *1*(1), 1–111.

Vella, F. (1998). Estimating models with sample selection bias: A survey. *Journal of Human Resources, 33*(1), 127–169.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*(4), 817–838.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). MIT Press.