Limited Dependent Variables

Martin Conyon

December 3, 2024

Introduction to Binary Response Models

Introduction

- Binary dependent variables are prevalent in the social sciences.
- Also known as discrete choice models.
- A dummy variable can take only two values: 0 and 1.

Expected Value and Interpretation

- The expected value of the dependent variable, conditional on covariates, is equal to the probability of the outcome being 1.
- Mathematically:

$$\mathbb{E}[Y|X] = \Pr(Y = 1|X)$$

- The predicted value is interpreted as a predicted probability.

Linear Probability Model (LPM)

Linear Probability Model

- When a dummy variable is the dependent variable, Ordinary Least Squares (OLS) can be used.
- The Linear Probability Model (LPM) is specified as:

$$\Pr(Y = 1|X) = X\beta$$

Assumptions and Interpretation

- The LPM relies on Gauss-Markov assumptions, particularly homoskedasticity.
- **Interpretation of Parameters:**
- **Each coefficient β_j represents the average change in the probability of success when X_j increases by one unit, holding other variables constant.**

Shortcomings of LPM

- The model is linear and unconstrained.
- Predicted probabilities can fall outside the [0, 1] interval.
- Induces heteroskedasticity, requiring robust standard errors for inference.

Logit Model

Overview of Logit Model

- Designed for binary response variables.
- Ensures predicted probabilities are restricted between 0 and 1.
- Utilizes a nonlinear link function to map predictors to probability scale.
- Instance of generalized linear models (GLMs).

Linking Function

- The Logit model uses the logistic function as the linking function.
- The probability of success is modeled as:

$$\Pr(Y = 1|X) = \frac{1}{1 + e^{-X\beta}}$$

- The logistic function ensures that predicted probabilities lie within the [0, 1] interval.

Log-Odds and Odds Ratio

- The Logit model can be expressed in terms of log-odds:

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta X$$

  where $p = \Pr(Y = 1|X)$.
- **Interpretation:**
- α **is the intercept term.**
- β **represents the change in log-odds of success for a one-unit increase in X.**

Mathematical Relationship

- Starting from the log-odds equation:

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta X$$

- To express odds in terms of α and β:

$$\frac{p}{1-p} = e^{\alpha+\beta X}$$

- Solving for p:

$$p = \frac{e^{\alpha+\beta X}}{1+e^{\alpha+\beta X}} = \frac{1}{1+e^{-(\alpha+\beta X)}}$$

Odds Ratio

- The odds ratio is the exponentiated coefficient:

$$e^{\beta}$$

- Represents the multiplicative change in odds for a one-unit increase in X.
- **Example:**
- **If $\beta = 0.5$, then $e^{0.5} \approx 1.65$.**
- **This means the odds of success increase by 65% for a one-unit increase in X.**

Interpreting Coefficients

- Each coefficient β_j in the Logit model represents the change in the log-odds of success for a one-unit increase in X_j, holding other variables constant.
- Mathematically:

$$\frac{\partial \ln\left(\frac{p}{1-p}\right)}{\partial X_j} = \beta_j$$

- To interpret in terms of odds ratio:

$$\text{Odds Ratio} = e^{\beta_j}$$

Log-Odds vs. Probability

- Log-odds provide a linear relationship between predictors and the logit of the probability.
- Direct interpretation of coefficients is straightforward in log-odds space.
- However, interpreting changes in probability requires understanding the nonlinear link function.

Reporting Log-Odds vs. Marginal Effects

- **Log-Odds:**
- **Directly report coefficients and interpret changes in log-odds or odds ratios.**
- **Useful for understanding the relationship between predictors and the outcome in log-odds space.**
- **Marginal Effects:**
- **Measure the impact of a small change in a covariate on the probability of success.**
- **Provide a more intuitive interpretation in probability terms.**
- **Often preferred for presenting results to non-technical audiences.**

Advantages of Reporting Marginal Effects

- Easier to understand and interpret in practical terms.
- Directly relates to changes in probability, which is often of primary interest.
- Facilitates comparison across different models and studies.

Disadvantages of Reporting Log-Odds

- Less intuitive for those unfamiliar with log-odds.
- Can be misleading if interpreted directly as probability changes.
- Requires transformation for probability-based interpretation.

Probit Model

Overview of Probit Model

- Designed for binary response variables.
- Ensures predicted probabilities are restricted between 0 and 1.
- Utilizes a nonlinear link function based on the cumulative distribution function (CDF) of the standard normal distribution.
- Instance of generalized linear models (GLMs).

Linking Function

- The Probit model uses the standard normal CDF as the linking function.
- The probability of success is modeled as:

$$\Pr(Y = 1|X) = \Phi(X\beta)$$

  where Φ denotes the standard normal CDF.
- The Probit link ensures that predicted probabilities lie within the [0, 1] interval.

Latent Variable Interpretation

- The Probit model assumes an underlying latent variable Y^*:

$$Y^* = X\beta + \epsilon$$

  where $\epsilon \sim N(0, 1)$.

- The observed variable Y is defined as:

$$Y = \begin{cases} 1 & \text{if } Y^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

Interpretation of Coefficients

- Each coefficient β_j represents the change in the z-score (standard deviations) of the latent variable Y^* for a one-unit increase in X_j, holding other variables constant.
- Unlike Logit, the Probit model does not have a natural odds ratio interpretation.
- Interpretation in terms of probabilities requires calculating marginal effects.

Model Comparison: Logit vs. Probit

- Both models ensure predicted probabilities lie within [0, 1].
- Logit uses the logistic function; Probit uses the standard normal CDF.
- Logit coefficients can be directly interpreted in terms of log-odds ratios; Probit coefficients require marginal effects for interpretation.
- In practice, Logit and Probit models often yield similar predictions.

Logit Model Assumptions and Diagnostics

Assumptions of Logit Models

- **Independence of Irrelevant Alternatives (IIA)**
- **No Perfect Multicollinearity**
- **Linearity in the Latent Variable Space**

Independence of Irrelevant Alternatives (IIA)

- **Definition**: The relative odds of choosing between any two alternatives remain unaffected by the presence or absence of other alternatives.
- **Intuition**: If you are deciding between coffee and tea, introducing juice as a third option should not change the odds of choosing coffee over tea.
- **Relevance**: Primarily important in multinomial Logit models where multiple choices exist.

No Perfect Multicollinearity

- **Definition**: Predictors (independent variables) should not be perfectly correlated with each other.
- **Intuition**: If two variables provide the exact same information, the model cannot distinguish their individual effects.
- **Example**: Including both ”total income” and ”income from salary” where one is a subset of the other.
- **Implication**: Perfect multicollinearity leads to infinite estimates and makes the model estimation impossible.

Linearity in the Latent Variable Space

- **Definition**: The relationship between the independent variables and the latent variable (log-odds) is linear.
- **Intuition**: While the probability itself changes nonlinearly, the underlying log-odds respond linearly to predictors.
- **Mathematical Representation**:

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

where $p = \Pr(Y = 1|X)$.

Formal Tobit Model

- The Tobit model assumes an underlying latent variable Y^*:

$$Y^* = X\beta + \epsilon$$

  where $\epsilon \sim N(0, \sigma^2)$
- The observed variable Y is defined as:

$$Y = \begin{cases} Y^* & \text{if } Y^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

Expected Value of Tobit Model

- The expected value of Y given X is:

$$\mathbb{E}[Y|X] = X\beta\Phi\left(\frac{X\beta}{\sigma}\right) + \sigma\phi\left(\frac{X\beta}{\sigma}\right)$$

  where Φ is the standard normal CDF and ϕ is the standard normal PDF.

Estimation of Tobit Models

- Tobit models are estimated using Maximum Likelihood Estimation (MLE).
- The likelihood function accounts for both the probability of Y being censored and the density of Y being observed.
- The log-likelihood function for the Tobit model is:

$$\ln \mathcal{L} = \sum_{i=1}^{n} \left[Y_i \ln \phi \left(\frac{Y_i - X_i\beta}{\sigma} \right) + (1 - Y_i) \ln \Phi \left(\frac{Y_i - X_i\beta}{\sigma} \right) \right]$$

Interpretation of Tobit Model Parameters

- The coefficients β represent the change in the latent variable Y^* for a one-unit change in the predictor.
- Effects on the observed variable Y are more complex due to censoring.
- Interpretation requires consideration of both the probability of censoring and the expected value when not censored.

Partial Effects in Tobit Models

- Partial effects measure the impact of a small change in a covariate on the expected value of Y.
- For continuous variables:

$$\frac{\partial \mathbb{E}[Y|X]}{\partial X_j} = \beta_j \Phi\left(\frac{X\beta}{\sigma}\right) + \frac{X\beta\beta_j}{\sigma}\phi\left(\frac{X\beta}{\sigma}\right) + \beta_j \phi\left(\frac{X\beta}{\sigma}\right)$$

- Partial effects depend on the level of the covariates and are nonlinear.

Estimation Challenges

- Convergence issues can arise during the numerical optimization process in MLE.
- Choice of starting values can affect the likelihood of achieving convergence.
- Model specification must account for the nature and extent of censoring in the data.

Interpretation and Partial Effects

Partial Effects in Logit and Probit Models

- Partial effects measure the impact of a small change in a covariate on the probability of success.
- For continuous variables:

$$\frac{\partial \Pr(Y = 1|X)}{\partial X_j} = G'(X\beta) \cdot \beta_j$$

  where G is the link function (logistic for Logit, standard normal CDF for Probit).
- Partial effects are nonlinear and depend on the values of X.

Interpreting Partial Effects

- Represents the change in the probability of success when X_j increases by one unit, holding other variables constant.
- For discrete variables:
- Calculated as the change in probability when the variable increases by one discrete unit.

Aggregating Partial Effects

Why Aggregate Partial Effects?

- Due to nonlinearity, partial effects vary across observations.
- Aggregating provides summary statistics that are easier to interpret and report.

Partial Effects at the Average

- Compute partial effects using the average values of the covariates.
$$\left.\frac{\partial \Pr(Y = 1|X)}{\partial X_j}\right|_{X=\bar{X}}$$
- Provides a single estimate of the partial effect evaluated at the mean of the covariates.

Average Partial Effects

- Take the average of the partial effects across all observations.

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \Pr(Y=1|X_i)}{\partial X_j}$$

- Reflects the average impact of a covariate on the probability of success across the entire sample.

Marginal Effects vs. Average Partial Effects

Odds

- **Odds**: The ratio of the probability of an event occurring to the probability of it not occurring.
- Mathematically:

$$\text{Odds} = \frac{p}{1 - p}$$

Log-Odds (Logit)

- **Log-Odds (Logit)**: The natural logarithm of the odds.
- Expressed as:

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta X$$

  where $p = \Pr(Y = 1|X)$.

Exponential Function

- **Exponential Function**: Used to convert log-odds back to odds.
- Given log-odds $\alpha + \beta X$:

$$\frac{p}{1-p} = e^{\alpha+\beta X}$$

From Log-Odds to Probability

- Solving for p:

$$p = \frac{e^{\alpha+\beta X}}{1 + e^{\alpha+\beta X}} = \frac{1}{1 + e^{-(\alpha+\beta X)}}$$

Marginal Effects: Intuitive Explanation

- **Marginal Effect**: The change in the probability of the outcome variable resulting from a one-unit change in a predictor variable, holding other variables constant.
- Reflects how sensitive the probability is to changes in a predictor at a specific point.

Average Partial Effects: Intuitive Explanation

- **Average Partial Effect (APE)**: The average of the marginal effects across all observations in the sample.
- Provides an overall measure of the impact of a predictor on the probability of the outcome.

Marginal Effects: Mathematical Definition

- For a Logit model:

$$\frac{\partial p}{\partial X_j} = p(1-p)\beta_j$$

- For a Probit model:

$$\frac{\partial p}{\partial X_j} = \phi(X\beta)\beta_j$$

  where ϕ is the standard normal density function.

Average Partial Effects: Mathematical Definition

- Calculated by taking the average of the individual partial effects across all observations.
- For Logit:

$$\text{APE}_j = \frac{1}{n}\sum_{i=1}^{n} p_i(1-p_i)\beta_j$$

- For Probit:

$$\text{APE}_j = \frac{1}{n}\sum_{i=1}^{n} \phi(X_i\beta)\beta_j$$

Comparison Between Marginal Effects and Average Partial Effects

- **Marginal Effects**:
  - Impact of a predictor at a specific point.
  - Can vary across observations due to nonlinearity.
- **Average Partial Effects**:
  - A single summary measure across the entire sample.
  - Useful for generalizing the effect of predictors.

Reporting Considerations

- **Log-Odds**:
  - Useful for understanding the relationship in logit space.
  - Not directly interpretable in terms of probability changes.
- **Marginal Effects**:
  - Provide a more intuitive understanding in probability terms.
  - Easier to communicate to non-technical audiences.

Advantages of Reporting Marginal Effects

- Easier to understand and interpret in practical terms.
- Directly relates to changes in probability, which is often of primary interest.
- Facilitates comparison across different models and studies.

Disadvantages of Reporting Log-Odds

- Less intuitive for those unfamiliar with log-odds.
- Can be misleading if interpreted directly as probability changes.
- Requires transformation for probability-based interpretation.

Heckman Selection Model

Introduction to Selection Bias

- **Selection Bias**: Occurs when the sample is not randomly selected from the population, leading to biased estimates.
- Common in labor economics where wage is only observed for employed individuals.
- Ignoring selection bias can result in inconsistent and biased parameter estimates.

Motivation for Heckman Selection Model

- Traditional OLS estimates are biased when sample selection is non-random.
- Example: Estimating the effect of education on wages using only employed individuals ignores those not employed.
- **James Heckman** introduced a two-step procedure to correct for selection bias.

Heckman Selection Model Overview

- Combines a selection equation and an outcome equation.
- Accounts for the non-random selection into the sample.
- Requires at least one variable that affects selection but not the outcome (exclusion restriction).

Formal Model Specification

- **Selection Equation**:

$$S^* = Z\gamma + \epsilon_s$$

  where $S = 1$ if selected (e.g., employed), and $S = 0$ otherwise.

- **Outcome Equation**:

$$Y^* = X\beta + \epsilon_y$$

  where Y is observed only if $S = 1$.

- Assumes:
  - Errors (ϵ_s, ϵ_y) follow a bivariate normal distribution.
  - Correlation between selection and outcome errors: $\rho \neq 0$.

Expected Value with Selection Bias

- Without correcting for selection:

$$\mathbb{E}[Y|S=1,X] = X\beta$$

- With selection bias:

$$\mathbb{E}[Y|S=1,X] = X\beta + \rho\sigma_y \frac{\phi(Z\gamma)}{\Phi(Z\gamma)}$$

  where ϕ and Φ are the standard normal PDF and CDF respectively.

- The term $\rho\sigma_y\lambda(Z\gamma)$ is known as the **Inverse Mills Ratio**.

Selection Bias Illustrated

- Sample consists of employed and not-employed individuals.
- Education levels differ between employed and not-employed.
- OLS on employed only:

$$\hat{Y} = X\beta + \text{Bias}$$

- Bias arises from omitted variables correlated with both selection and outcome.

Heckman's Two-Step Procedure

- **Stage 1**: Selection Equation
  - Estimate the probability of selection using a Probit model.
  - Generate the Inverse Mills Ratio:

$$\lambda(Z\gamma) = \frac{\phi(Z\gamma)}{\Phi(Z\gamma)}$$

- **Stage 2**: Outcome Equation
  - Include the Inverse Mills Ratio as an additional regressor.
  - Corrects for selection bias:

$$Y = X\beta + \rho\sigma_y\lambda(Z\gamma) + u$$

Identification Issues

- Requires at least one **exclusion restriction**: A variable that affects selection but not the outcome.
- Ensures the Inverse Mills Ratio is identifiable.
- Without exclusion restrictions, the model suffers from identification problems.

Intuition Behind Heckman's Correction

- The Inverse Mills Ratio captures the non-random selection into the sample.
- By including it in the outcome equation, we control for the factors influencing both selection and the outcome.
- This adjustment removes the bias caused by the correlation between selection and outcome errors.

Limitations of Heckman Selection Model

- Relies on the normality assumption of errors.
- Sensitive to the choice of exclusion restrictions.
- Requires correct model specification in both selection and outcome equations.
- Potential for multicollinearity if Inverse Mills Ratio is highly correlated with other regressors.

Heckman Model in Limited Dependent Variable Framework

- Part of **Limited Dependent Variable** models because the outcome is only partially observed.
- Deals with scenarios where the dependent variable is censored or truncated.
- Extends beyond binary outcomes to continuous outcomes with sample selection.

Heckman's Contribution to Econometrics

- Provided a method to correct for sample selection bias.
- Enhanced the reliability of parameter estimates in non-randomly selected samples.
- Expanded the toolkit for handling limited dependent variables in econometric analysis.

Resolving Selection Bias

- OLS ignores the selection process, leading to biased estimates.
- Heckman's two-step procedure accounts for the selection mechanism.
- Ensures consistent and unbiased parameter estimates by incorporating the Inverse Mills Ratio.

Mathematical Resolution of Selection Bias

- Original Outcome Equation:

$$Y^* = X\beta + u$$

- Selection Equation:

$$S^* = Z\gamma + \epsilon_s$$

- Observed Outcome:

$$Y = \begin{cases} Y^* & \text{if } S = 1 \\ \text{Not Observed} & \text{otherwise} \end{cases}$$

- After Two-Step Correction:

$$Y = X\beta + \rho\sigma_y\lambda(Z\gamma) + u$$

Practical Implementation in Stata

- Step 1: Estimate the Selection Equation using Probit.
- Step 2: Generate the Inverse Mills Ratio.
- Step 3: Include the Inverse Mills Ratio in the Outcome Equation.
- Alternatively, use the built-in 'heckman' command for a streamlined two-step estimation.

Example Interpretation

- Significant Inverse Mills Ratio indicates the presence of selection bias.
- Coefficients on regressors in the outcome equation represent the adjusted effects accounting for selection.
- Non-significant Inverse Mills Ratio suggests that selection bias may not be a concern.

Summary

- Heckman Selection Model addresses sample selection bias in econometric analysis.
- Utilizes a two-step procedure to correct for non-random selection.
- Essential for obtaining unbiased and consistent parameter estimates in the presence of selection.
- Part of the broader class of Limited Dependent Variable models.

Propensity Score Matching

Introduction to Propensity Score Matching

- **Propensity Score Matching (PSM)**: A statistical technique used to control for selection bias in observational studies.
- Addresses the challenge of estimating treatment effects when treatment assignment is not random.
- Essential for ensuring comparability between treatment and control groups in the presence of confounding variables.

Why Introduce Propensity Score Matching?

- In many studies, treatment assignment (1 for treatment, 0 for control) is not randomized.
- Non-random assignment can lead to biased estimates of treatment effects.
- PSM aims to create a balanced comparison by matching treated units with similar control units based on covariates.

Intuitive Explanation of PSM

- Imagine two groups: treated (T=1) and control (T=0).
- Without matching, these groups may differ systematically in important characteristics.
- PSM involves pairing each treated unit with one or more control units that have similar propensity scores.
- This matching process balances the distribution of covariates, mimicking randomization.

Visualizing PSM Intuition

- Consider the distribution of propensity scores for both treated and control groups.
- PSM ensures that for each treated unit, there exists a control unit with a similar propensity score.
- This overlap reduces selection bias and allows for a more accurate estimation of treatment effects.

Examples of PSM in Various Fields

- **Finance**:
    - Assessing the impact of a financial regulation on firm performance.
    - Matching regulated firms with non-regulated firms based on size, industry, and other characteristics.

Examples of PSM in Various Fields (Cont.)

- **Economics**:
    - Evaluating the effect of job training programs on employment outcomes.
    - Matching participants with non-participants based on education, experience, and demographics.

Examples of PSM in Various Fields (Cont.)

- **Accounting**:
  - Determining the impact of adopting a new accounting standard on financial reporting quality.
  - Matching firms that adopted the standard with similar firms that did not.

Mathematics of Propensity Score Matching

- **Propensity Score Definition**:

$$e(X) = \Pr(T = 1|X)$$

where T is the treatment indicator and X represents covariates.

Mathematics of Propensity Score Matching (Cont.)

- **Average Treatment Effect (ATE)**:

$$\text{ATE} = \mathbb{E}[Y(1) - Y(0)]$$

- **Average Treatment Effect on the Treated (ATT)**:

$$\text{ATT} = \mathbb{E}[Y(1) - Y(0) | T = 1]$$

Balancing Covariates with PSM

- After matching, the distribution of covariates should be similar between treated and control groups.
- **Covariate Balance**:

$$\mathbb{E}[X|T=1] = \mathbb{E}[X|T=0]$$

- Ensures that differences in outcomes can be attributed to the treatment rather than underlying covariates.

Limited Dependent Variable Analysis

- PSM deals with binary treatment variables, a form of limited dependent variable.
- Ensures that the selection into treatment is appropriately addressed.
- Connects to models like Logit and Probit where the treatment assignment is modeled.

Estimating the Propensity Score

- **Logit Model**:

$$e(X) = \Pr(T = 1|X) = \frac{1}{1 + e^{-(\alpha+\beta X)}}$$

- **Probit Model**:

$$e(X) = \Pr(T = 1|X) = \Phi(\alpha + \beta X)$$

where Φ is the standard normal CDF.

Estimating the Propensity Score (Cont.)

- **Other Methods**:
  - **Machine Learning Techniques**: Random forests, gradient boosting.
  - **Nearest Neighbor Matching**: Pair treated units with control units based on nearest propensity scores.
  - **Kernel Matching**: Weight control units based on their distance from treated units.

Why Choose Logit or Probit for Estimating Propensity Scores?

- **Logit Model**:
  - Provides odds ratios, which are easy to interpret.
  - Suitable when the relationship between covariates and treatment is logistic.

Why Choose Logit or Probit for Estimating Propensity Scores? (Cont.)

- **Probit Model**:
  - Based on the standard normal distribution.
  - Often used when the latent variable interpretation is preferred.
- **Considerations**:
  - Choice depends on the nature of the data and underlying assumptions.
  - Both models aim to accurately estimate the propensity score for effective matching.

Advantages of Propensity Score Matching

- Reduces selection bias by balancing covariates between treated and control groups.
- Mimics randomization, making causal inferences more credible.
- Flexible and can be combined with various matching algorithms.
- Does not require modeling the outcome variable directly.

Limitations of Propensity Score Matching

- Only accounts for observed covariates; unobserved confounders can still bias estimates.
- Requires overlap in propensity scores between treated and control groups (common support).
- Matching quality depends on the specification of the propensity score model.
- Can be computationally intensive with large datasets.

Steps in Propensity Score Matching

1. Estimate the Propensity Score using Logit, Probit, or other suitable models.
2. Perform Matching using methods like Nearest Neighbor, Caliper, or Kernel Matching.
3. Assess Covariate Balance to ensure matched samples are comparable.
4. Estimate the Treatment Effect by comparing outcomes between matched treated and control groups.

Implementing PSM in Stata

- **Step 1**: Estimate the Propensity Score using Logit or Probit.
  - Example using Logit:

```
logit treatment X1 X2 X3
```

Implementing PSM in Stata (Cont.)

- **Step 2**: Perform Matching.
  - Nearest Neighbor Matching:

```
teffects psmatch (outcome) (treatment X1 X2 X3)
```

Implementing PSM in Stata (Cont.)

- **Step 3**: Assess Balance.
  - Use graphical and statistical methods to ensure covariate balance.
  - Example:

```
pstest X1 X2 X3, graph
```

Implementing PSM in Stata (Cont.)

- **Step 4**: Estimate Treatment Effect.
    - After matching, compare outcomes to estimate the treatment effect.

Why Choose Propensity Score Matching Over Other Methods?

- **Compared to OLS**:
  - PSM reduces selection bias by balancing covariates.
  - OLS may still suffer from omitted variable bias.

Why Choose Propensity Score Matching Over Other Methods? (Cont.)

- **Compared to Instrumental Variables (IV)**:
    - PSM does not require valid instruments.
    - IV requires exclusion restrictions which can be hard to justify.

Why Choose Propensity Score Matching Over Other Methods? (Cont.)

- **Compared to Heckman Selection Model**:
  - PSM is non-parametric and does not rely on distributional assumptions.
  - Heckman requires an exclusion restriction and assumes normality.

Why Choose Propensity Score Matching Over Other Methods? (Cont.)

- **Overall**:
  - PSM is intuitive, relatively easy to implement, and provides a clear framework for reducing selection bias.

Summary of Propensity Score Matching

- PSM is a robust method for addressing selection bias in observational studies.
- By matching treated and control units with similar propensity scores, PSM aims to mimic randomization.
- Essential for making causal inferences when random assignment is not possible.
- Requires careful estimation of the propensity score and assessment of covariate balance.
- Complements other econometric techniques like Logit, Probit, and Heckman models.