# The Tobit Model (Censored Regression)

## When 40% of Your Data Is Piled Up at Zero

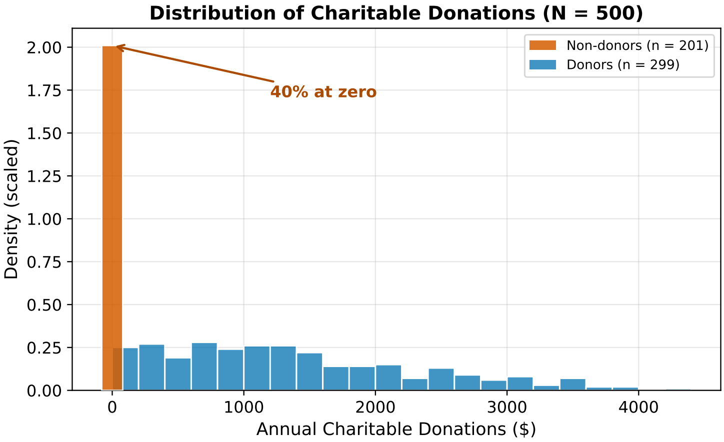
Jake Anderson

March 3, 2026

# Outline

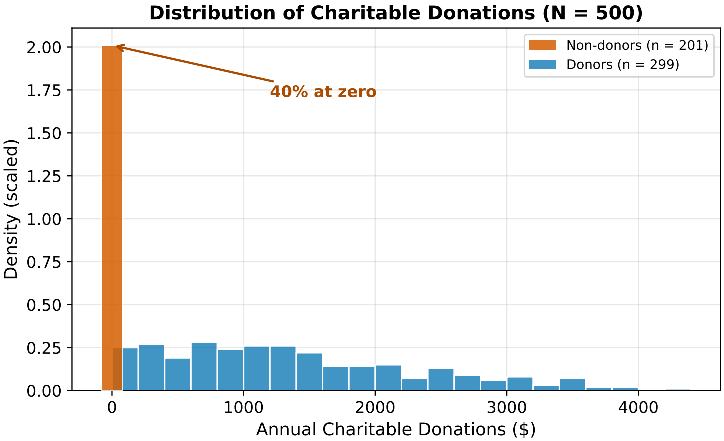
# The Data: Charitable Donations

A researcher surveys **500 households** and records **annual charitable donations** (\$). Covariates include household income (\$1000s), years of education, and number of children.



# The Data: Charitable Donations

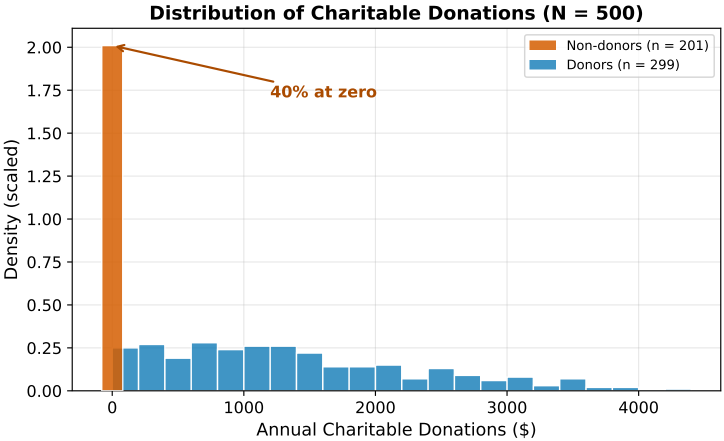
A researcher surveys **500 households** and records **annual charitable donations** ($). Covariates include household income ($1000s), years of education, and number of children.




What do you notice about this distribution?

# The Data: Charitable Donations

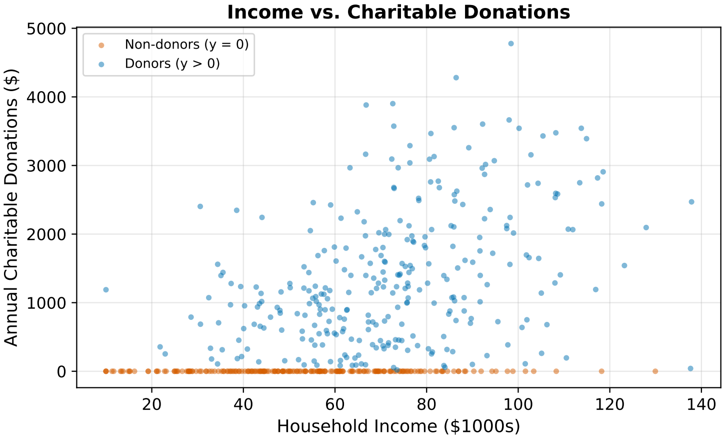
A researcher surveys **500 households** and records **annual charitable donations** ($). Covariates include household income ($1000s), years of education, and number of children.




What do you notice about this distribution?

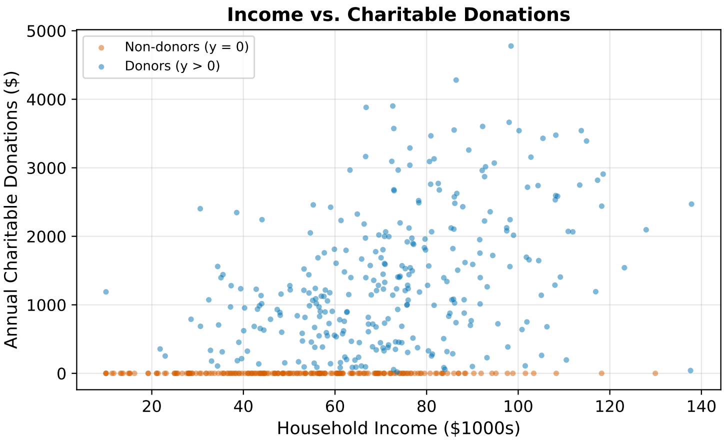
# Income and Donations

Let's plot household income against donations:



# Income and Donations

Let's plot household income against donations:

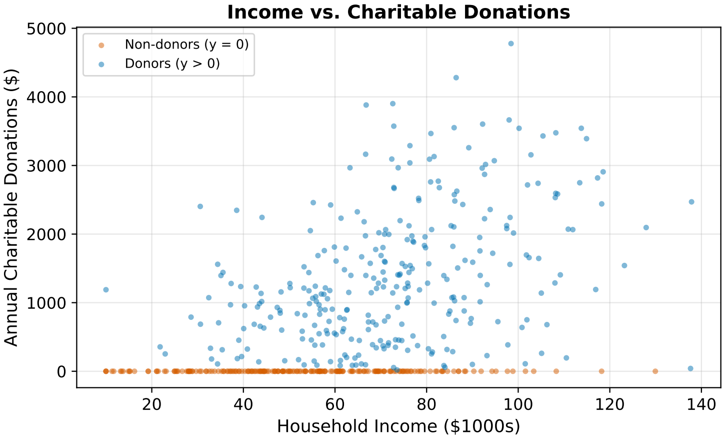



Two features stand out:

- A thick band of orange points **piled up along** $y = 0$, concentrated at lower incomes

# Income and Donations

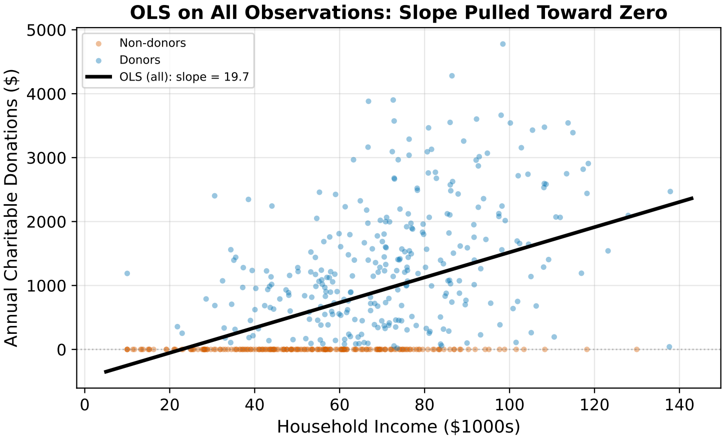
Let's plot household income against donations:



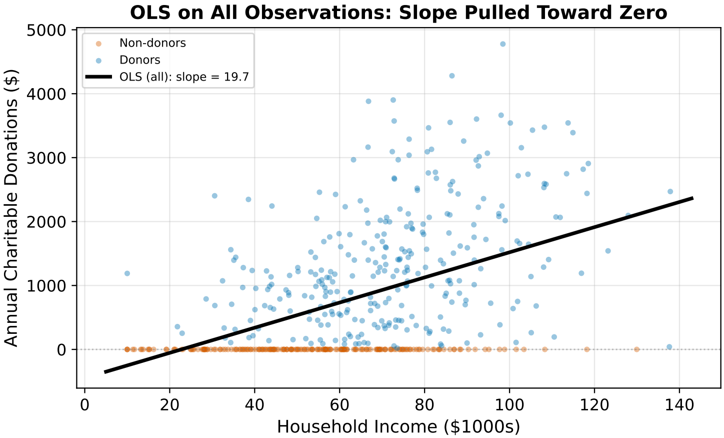

Two features stand out:

- A thick band of orange points **piled up along** $y = 0$, concentrated at lower incomes

# OLS on All Observations: Slope Pulled Down

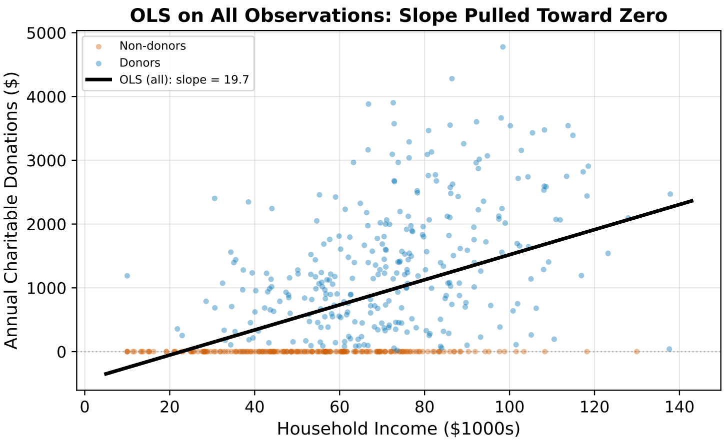


# OLS on All Observations: Slope Pulled Down




OLS slope = 19.7 dollars per $1000 income. But the true effect in the underlying model is **30 dollars per $1000 income**.

# OLS on All Observations: Slope Pulled Down

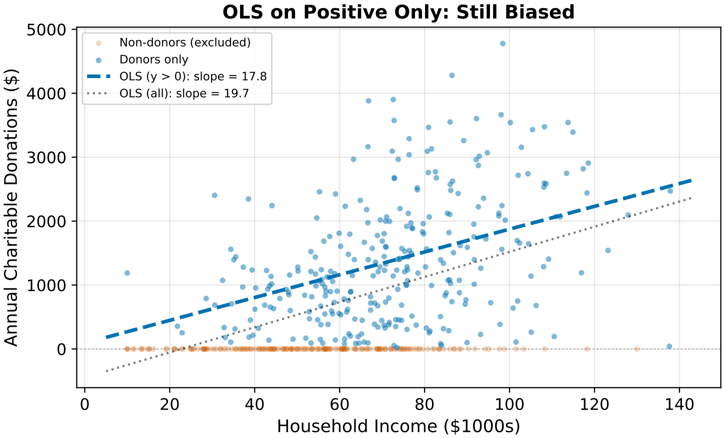



OLS slope = 19.7 dollars per $1000 income. But the true effect in the underlying model is **30 dollars per $1000 income**.

Why is OLS attenuated? The 200 non-donors all sit at $y = 0$ regardless of their income. OLS treats

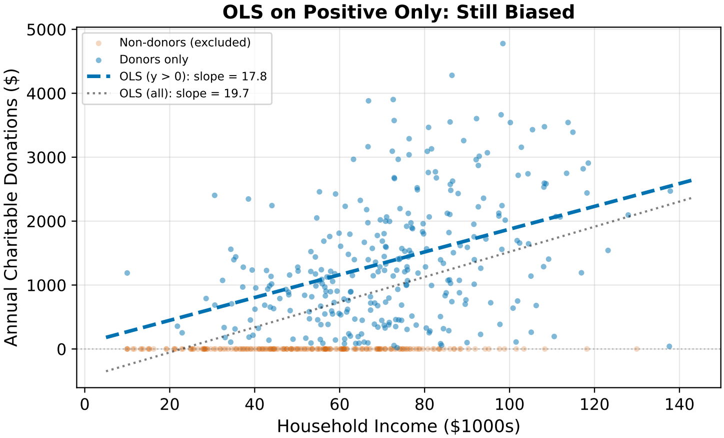
# OLS on Positive Observations Only: Still Biased

Maybe we should drop the zeros and run OLS on donors only?



# OLS on Positive Observations Only: Still Biased

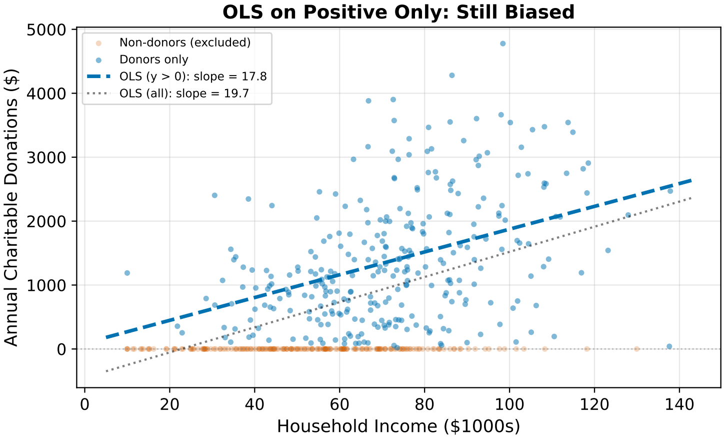
Maybe we should drop the zeros and run OLS on donors only?




OLS on donors only: slope = 17.8 (true = 30). Even worse!

# OLS on Positive Observations Only: Still Biased

Maybe we should drop the zeros and run OLS on donors only?




OLS on donors only: slope $= 17.8$ (true $= 30$). Even worse!

**The problem:** by conditioning on $y > 0$, we have selected a non-random subsample. Among

# Outline

# Neither OLS Works: Where Does That Leave Us?

| Approach | Slope estimate | True slope |
|---|---|---|
| OLS on all observations | 19.7 | 30 |
| OLS on positive only | 17.8 | 30 |

# Neither OLS Works: Where Does That Leave Us?

| Approach | Slope estimate | True slope |
|---|---|---|
| OLS on all observations | 19.7 | 30 |
| OLS on positive only | 17.8 | 30 |

**OLS on all:** treats the zeros as legitimate data points, pulling the slope toward zero.

# Neither OLS Works: Where Does That Leave Us?

| Approach | Slope estimate | True slope |
|---|---|---|
| OLS on all observations | 19.7 | 30 |
| OLS on positive only | 17.8 | 30 |

**OLS on all:** treats the zeros as legitimate data points, pulling the slope toward zero.

**OLS on positives:** throws away 40% of the data and introduces sample selection bias.

# Neither OLS Works: Where Does That Leave Us?

| Approach | Slope estimate | True slope |
|---|---|---|
| OLS on all observations | 19.7 | 30 |
| OLS on positive only | 17.8 | 30 |

**OLS on all:** treats the zeros as legitimate data points, pulling the slope toward zero.

**OLS on positives:** throws away 40% of the data and introduces sample selection bias.

$\Longrightarrow$ Both approaches ignore the **mechanism** that generates the zeros. We need a model that understands *why* some households donate zero.

# What Would a Better Model Need?

A model for corner solution outcomes should:

# What Would a Better Model Need?

A model for corner solution outcomes should:

1. **Explain the zeros:** some households *would* donate if they could, but their desired amount is negative (they are constrained to zero). This is what creates the pile-up that **attenuates the OLS-on-all slope** from 30 down to 19.7

# What Would a Better Model Need?

A model for corner solution outcomes should:

1. **Explain the zeros:** some households *would* donate if they could, but their desired amount is negative (they are constrained to zero). This is what creates the pile-up that **attenuates the OLS-on-all slope** from 30 down to 19.7
2. **Use all the data:** both zeros and positives carry information about the income effect. Dropping the zeros introduces the **sample selection bias** that made OLS-on-positives even worse (17.8)

# What Would a Better Model Need?

A model for corner solution outcomes should:

1. **Explain the zeros:** some households *would* donate if they could, but their desired amount is negative (they are constrained to zero). This is what creates the pile-up that **attenuates the OLS-on-all slope** from 30 down to 19.7
2. **Use all the data:** both zeros and positives carry information about the income effect. Dropping the zeros introduces the **sample selection bias** that made OLS-on-positives even worse (17.8)
3. **Recover the true slope:** the underlying relationship between income and desired donations, not the censored version OLS estimates

# What Would a Better Model Need?

A model for corner solution outcomes should:

1. **Explain the zeros:** some households *would* donate if they could, but their desired amount is negative (they are constrained to zero). This is what creates the pile-up that **attenuates the OLS-on-all slope** from 30 down to 19.7
2. **Use all the data:** both zeros and positives carry information about the income effect. Dropping the zeros introduces the **sample selection bias** that made OLS-on-positives even worse (17.8)
3. **Recover the true slope:** the underlying relationship between income and desired donations, not the censored version OLS estimates

$\Longrightarrow$ We need a model that distinguishes between the **desired** outcome and the **observed** outcome. This is the latent variable idea you already know from logit/probit.

# Outline

# The Idea: Desired vs. Observed Donations

We saw two problems: the pile-up at zero attenuates the OLS slope, and dropping zeros introduces selection. Both stem from the same source: **we observe donations, not desires**.

# The Idea: Desired vs. Observed Donations

We saw two problems: the pile-up at zero attenuates the OLS slope, and dropping zeros introduces selection. Both stem from the same source: **we observe donations, not desires**.

Imagine each household has a *desired* donation that could be positive or negative. A household with low income and no particular charitable inclination might "want" to donate $-\$500$ (they would take donations back if they could). But donations cannot be negative, so these households are constrained to zero.

# The Idea: Desired vs. Observed Donations

We saw two problems: the pile-up at zero attenuates the OLS slope, and dropping zeros introduces selection. Both stem from the same source: **we observe donations, not desires**.

Imagine each household has a *desired* donation that could be positive or negative. A household with low income and no particular charitable inclination might "want" to donate $-\$500$ (they would take donations back if they could). But donations cannot be negative, so these households are constrained to zero.

If we could observe the desires directly, OLS would work perfectly: there would be no pile-up, no selection, just a clean linear relationship.

# The Idea: Desired vs. Observed Donations

We saw two problems: the pile-up at zero attenuates the OLS slope, and dropping zeros introduces selection. Both stem from the same source: **we observe donations, not desires**.

Imagine each household has a *desired* donation that could be positive or negative. A household with low income and no particular charitable inclination might "want" to donate $-\$500$ (they would take donations back if they could). But donations cannot be negative, so these households are constrained to zero.

If we could observe the desires directly, OLS would work perfectly: there would be no pile-up, no selection, just a clean linear relationship.

$\Longrightarrow$ The Tobit model reconstructs those unobserved desires. It posits a **latent variable** for what each household *wants* to give, and an **observation rule** that censors negative desires to zero.

# Latent Variable Framework

Define the **latent (unobserved)** variable $y_i^*$ as each household's *desired* donations:

$$y_i^* = \beta_0 + \beta_1 \,\text{Income}_i + \beta_2 \,\text{Educ}_i + \beta_3 \,\text{Children}_i + \varepsilon_i, \qquad \varepsilon_i \sim N(0, \sigma^2)$$

# Latent Variable Framework

Define the **latent (unobserved)** variable $y_i^*$ as each household's *desired* donations:

$$y_i^* = \beta_0 + \beta_1 \,\text{Income}_i + \beta_2 \,\text{Educ}_i + \beta_3 \,\text{Children}_i + \varepsilon_i, \qquad \varepsilon_i \sim N(0, \sigma^2)$$

$y_i^*$ can be **positive or negative**. A household with low income might have $y_i^* = -\$500$: they would "take back" donations if they could.

# Latent Variable Framework

Define the **latent (unobserved)** variable $y_i^*$ as each household's *desired* donations:

$$y_i^* = \beta_0 + \beta_1 \,\text{Income}_i + \beta_2 \,\text{Educ}_i + \beta_3 \,\text{Children}_i + \varepsilon_i, \qquad \varepsilon_i \sim N(0, \sigma^2)$$

$y_i^*$ can be **positive or negative**. A household with low income might have $y_i^* = -\$500$: they would "take back" donations if they could.

But donations cannot be negative. The **observation rule** censors the latent variable:

$$y_i = \max(0,\ y_i^*) = \begin{cases} y_i^* & \text{if } y_i^* > 0 \quad \text{(donor)} \\ 0 & \text{if } y_i^* \leq 0 \quad \text{(non-donor, censored)} \end{cases}$$

# Latent Variable Framework

Define the **latent (unobserved)** variable $y_i^*$ as each household's *desired* donations:

$$y_i^* = \beta_0 + \beta_1 \,\text{Income}_i + \beta_2 \,\text{Educ}_i + \beta_3 \,\text{Children}_i + \varepsilon_i, \qquad \varepsilon_i \sim N(0, \sigma^2)$$

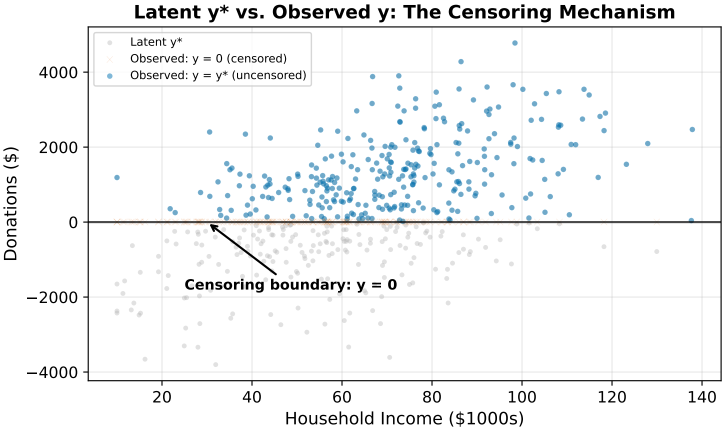
$y_i^*$ can be **positive or negative**. A household with low income might have $y_i^* = -\$500$: they would "take back" donations if they could.

But donations cannot be negative. The **observation rule** censors the latent variable:

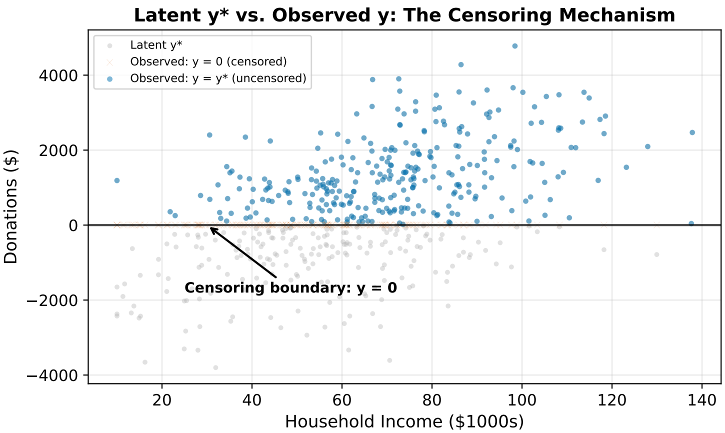
$$y_i = \max(0,\ y_i^*) = \begin{cases} y_i^* & \text{if } y_i^* > 0 \quad \text{(donor)} \\ 0 & \text{if } y_i^* \leq 0 \quad \text{(non-donor, censored)} \end{cases}$$

$\Longrightarrow$ The zeros are not real zeros. They are **censored observations** where the true desired donation is negative.

# Seeing the Censoring Mechanism

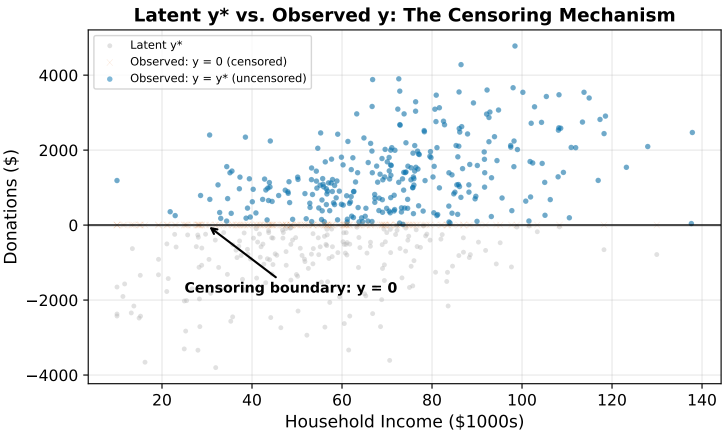


# Seeing the Censoring Mechanism




Gray points: latent $y_i^*$ (including negative values). Blue points: observed $y_i = y_i^*$ for donors. Orange crosses: observed $y_i = 0$ for non-donors.

# Seeing the Censoring Mechanism




Gray points: latent $y_i^*$ (including negative values). Blue points: observed $y_i = y_i^*$ for donors. Orange crosses: observed $y_i = 0$ for non-donors.

The censoring "folds" all negative latent values onto zero. This is what creates the pile-up, and this is

# The Tobit Likelihood: Two Pieces

Since $y_i$ comes from two different processes, the likelihood has two pieces.

# The Tobit Likelihood: Two Pieces

Since $y_i$ comes from two different processes, the likelihood has two pieces.

Recall: $\phi(\cdot)$ is the standard normal PDF and $\Phi(\cdot)$ is the standard normal CDF.

# The Tobit Likelihood: Two Pieces

Since $y_i$ comes from two different processes, the likelihood has two pieces.

Recall: $\phi(\cdot)$ is the standard normal PDF and $\Phi(\cdot)$ is the standard normal CDF.

**Uncensored observations** ($y_i > 0$): we observe the actual value, so the contribution is the normal density:

$$f(y_i) = \frac{1}{\sigma}\,\phi\left(\frac{y_i - \beta_0 - \beta_1 \text{Income}_i - \beta_2 \text{Educ}_i - \beta_3 \text{Children}_i}{\sigma}\right)$$

# The Tobit Likelihood: Two Pieces

Since $y_i$ comes from two different processes, the likelihood has two pieces.

Recall: $\phi(\cdot)$ is the standard normal PDF and $\Phi(\cdot)$ is the standard normal CDF.

**Uncensored observations** $(y_i > 0)$: we observe the actual value, so the contribution is the normal density:

$$f(y_i) = \frac{1}{\sigma}\,\phi\left(\frac{y_i - \beta_0 - \beta_1 \text{Income}_i - \beta_2 \text{Educ}_i - \beta_3 \text{Children}_i}{\sigma}\right)$$

**Censored observations** $(y_i = 0)$: we only know $y_i^* \leq 0$, so the contribution is the probability of censoring:

$$P(y_i = 0) = P(y_i^* \leq 0) = \Phi\left(\frac{-(\beta_0 + \beta_1 \text{Income}_i + \beta_2 \text{Educ}_i + \beta_3 \text{Children}_i)}{\sigma}\right)$$

# The Tobit Likelihood: Two Pieces

Since $y_i$ comes from two different processes, the likelihood has two pieces.

Recall: $\phi(\cdot)$ is the standard normal PDF and $\Phi(\cdot)$ is the standard normal CDF.

**Uncensored observations** ($y_i > 0$): we observe the actual value, so the contribution is the normal density:

$$f(y_i) = \frac{1}{\sigma}\,\phi\left(\frac{y_i - \beta_0 - \beta_1 \text{Income}_i - \beta_2 \text{Educ}_i - \beta_3 \text{Children}_i}{\sigma}\right)$$

**Censored observations** ($y_i = 0$): we only know $y_i^* \leq 0$, so the contribution is the probability of censoring:

$$P(y_i = 0) = P(y_i^* \leq 0) = \Phi\left(\frac{-(\beta_0 + \beta_1 \text{Income}_i + \beta_2 \text{Educ}_i + \beta_3 \text{Children}_i)}{\sigma}\right)$$

Why the negative sign inside $\Phi$? We need $P(\varepsilon_i \leq -\text{XB}_i)$. Flipping the sign converts $y_i^* \leq 0$ into a standard CDF evaluation.

# The Log-Likelihood

Define the shorthand $\text{XB}_i \equiv \beta_0 + \beta_1 \, \text{Income}_i + \beta_2 \, \text{Educ}_i + \beta_3 \, \text{Children}_i$ for the linear index.

# The Log-Likelihood

Define the shorthand $\mathrm{XB}_i \equiv \beta_0 + \beta_1\,\mathrm{Income}_i + \beta_2\,\mathrm{Educ}_i + \beta_3\,\mathrm{Children}_i$ for the linear index.

The log-likelihood sums over two groups: first the uncensored observations (donors), then the censored observations (non-donors):

$$\ell = \sum_{i:\; y_i > 0} \left[ \ln \phi\left( \frac{y_i - \mathrm{XB}_i}{\sigma} \right) - \ln \sigma \right] + \sum_{i:\; y_i = 0} \ln \Phi\left( \frac{-\mathrm{XB}_i}{\sigma} \right)$$

# The Log-Likelihood

Define the shorthand $\text{XB}_i \equiv \beta_0 + \beta_1 \,\text{Income}_i + \beta_2 \,\text{Educ}_i + \beta_3 \,\text{Children}_i$ for the linear index.

The log-likelihood sums over two groups: first the uncensored observations (donors), then the censored observations (non-donors):

$$\ell = \sum_{i:\; y_i > 0} \left[ \ln \phi\left( \frac{y_i - \text{XB}_i}{\sigma} \right) - \ln \sigma \right] + \sum_{i:\; y_i = 0} \ln \Phi\left( \frac{-\text{XB}_i}{\sigma} \right)$$

This combines **two types of contributions**: the first sum handles the continuous part (like OLS with normal errors), and the second sum handles the discrete part (like probit). Unlike a true mixture model, we observe which group each household belongs to.

# The Log-Likelihood

Define the shorthand $\text{XB}_i \equiv \beta_0 + \beta_1 \, \text{Income}_i + \beta_2 \, \text{Educ}_i + \beta_3 \, \text{Children}_i$ for the linear index.

The log-likelihood sums over two groups: first the uncensored observations (donors), then the censored observations (non-donors):

$$\ell = \sum_{i:\ y_i > 0} \left[ \ln \phi\left( \frac{y_i - \text{XB}_i}{\sigma} \right) - \ln \sigma \right] + \sum_{i:\ y_i = 0} \ln \Phi\left( \frac{-\text{XB}_i}{\sigma} \right)$$

This combines **two types of contributions**: the first sum handles the continuous part (like OLS with normal errors), and the second sum handles the discrete part (like probit). Unlike a true mixture model, we observe which group each household belongs to.

Software maximizes $\ell$ over $(\beta_0, \beta_1, \beta_2, \beta_3, \sigma)$ numerically, just as in logit/probit.

# The Log-Likelihood

Define the shorthand $\text{XB}_i \equiv \beta_0 + \beta_1\,\text{Income}_i + \beta_2\,\text{Educ}_i + \beta_3\,\text{Children}_i$ for the linear index.

The log-likelihood sums over two groups: first the uncensored observations (donors), then the censored observations (non-donors):

$$\ell = \sum_{i:\ y_i>0} \left[ \ln \phi\left( \frac{y_i - \text{XB}_i}{\sigma} \right) - \ln \sigma \right] + \sum_{i:\ y_i=0} \ln \Phi\left( \frac{-\text{XB}_i}{\sigma} \right)$$

This combines **two types of contributions**: the first sum handles the continuous part (like OLS with normal errors), and the second sum handles the discrete part (like probit). Unlike a true mixture model, we observe which group each household belongs to.

Software maximizes $\ell$ over $(\beta_0, \beta_1, \beta_2, \beta_3, \sigma)$ numerically, just as in logit/probit.

$\Longrightarrow$ The Tobit likelihood combines a regression component and a binary component into a single model.

# Numeric Example: Our Donation Data

Tobit MLE on the 500-household sample (true parameters in parentheses):

# Numeric Example: Our Donation Data

Tobit MLE on the 500-household sample (true parameters in parentheses):

| **Parameter** | **Tobit estimate** | **True value** |
|---|---|---|
| $\hat{\beta}_0$ (intercept) | $-3324$ | $-3500$ |
| $\hat{\beta}_1$ (Income, per \$1000) | 32.3 | 30 |
| $\hat{\beta}_2$ (Education, per year) | 119 | 150 |
| $\hat{\beta}_3$ (Children, per child) | $-16$ | $-100$ |
| $\hat{\sigma}$ | 1249 | 1200 |

# Numeric Example: Our Donation Data

Tobit MLE on the 500-household sample (true parameters in parentheses):

| Parameter | Tobit estimate | True value |
|---|---|---|
| $\hat{\beta}_0$ (intercept) | $-3324$ | $-3500$ |
| $\hat{\beta}_1$ (Income, per \$1000) | 32.3 | 30 |
| $\hat{\beta}_2$ (Education, per year) | 119 | 150 |
| $\hat{\beta}_3$ (Children, per child) | $-16$ | $-100$ |
| $\hat{\sigma}$ | 1249 | 1200 |

The income coefficient recovers **32.3** (true 30), much closer than OLS on all (19.7) or OLS on positives (17.8).

# Numeric Example: Our Donation Data

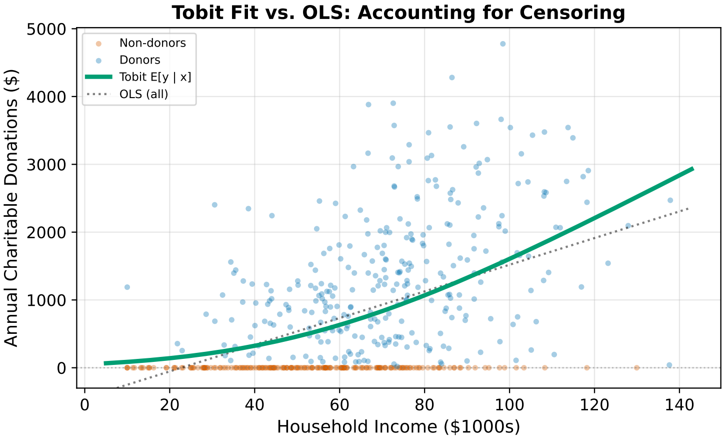
Tobit MLE on the 500-household sample (true parameters in parentheses):

| **Parameter** | **Tobit estimate** | **True value** |
|---|---|---|
| $\hat{\beta}_0$ (intercept) | $-3324$ | $-3500$ |
| $\hat{\beta}_1$ (Income, per \$1000) | 32.3 | 30 |
| $\hat{\beta}_2$ (Education, per year) | 119 | 150 |
| $\hat{\beta}_3$ (Children, per child) | $-16$ | $-100$ |
| $\hat{\sigma}$ | 1249 | 1200 |

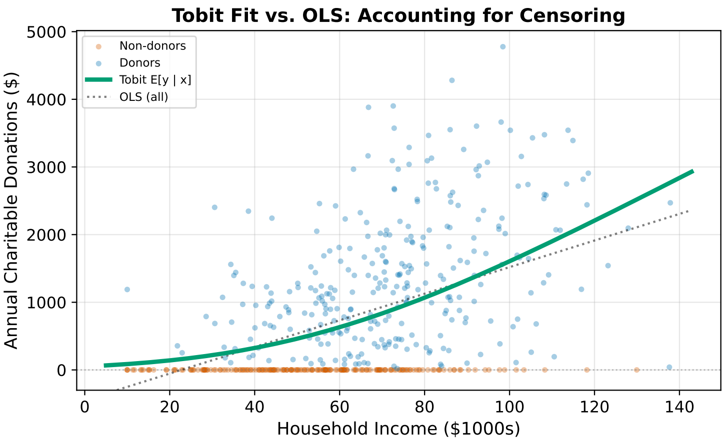
The income coefficient recovers **32.3** (true 30), much closer than OLS on all (19.7) or OLS on positives (17.8).

$\Longrightarrow$ By modeling the censoring mechanism explicitly, Tobit recovers the **latent** slope that OLS cannot.

# Tobit Fit on the Data

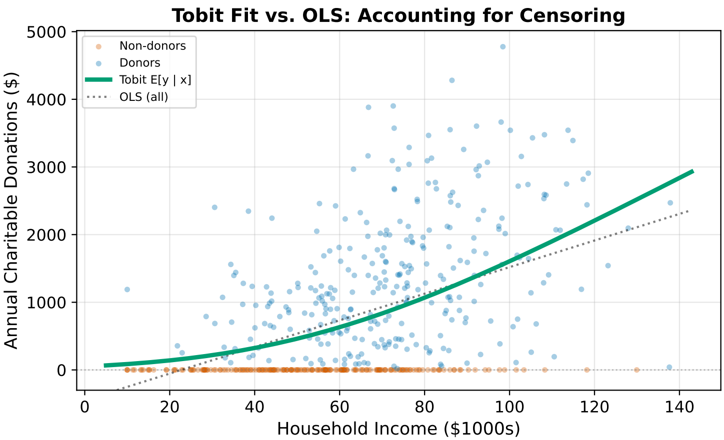


# Tobit Fit on the Data




The green Tobit curve shows **unconditional** E[$y_i$ | Income$_i$] (including zeros), holding education and children at their means. OLS (dotted) misses the nonlinear shape entirely.

# Tobit Fit on the Data




The green Tobit curve shows **unconditional** $\mathrm{E}[y_i \mid \mathrm{Income}_i]$ (including zeros), holding education and children at their means. OLS (dotted) misses the nonlinear shape entirely.

Notice the curve is flat near zero for low incomes (most households are censored, so additional income

# Three Questions, Three Marginal Effects

In OLS, $\hat{\beta}_1$ is the marginal effect, full stop. In Tobit, $\hat{\beta}_1 = 32.3$ is the effect on the **latent** variable $y^*$. But we observe $y = \max(0, y^*)$, so different research questions call for different marginal effects:

# Three Questions, Three Marginal Effects

In OLS, $\hat{\beta}_1$ is the marginal effect, full stop. In Tobit, $\hat{\beta}_1 = 32.3$ is the effect on the **latent** variable $y^*$. But we observe $y = \max(0, y^*)$, so different research questions call for different marginal effects:

1. **"What does the household want to give?"** The effect on latent desired donations $y^*$, including negative desires. Use this when you want the structural parameter of the underlying model

# Three Questions, Three Marginal Effects

In OLS, $\hat{\beta}_1$ is the marginal effect, full stop. In Tobit, $\hat{\beta}_1 = 32.3$ is the effect on the **latent** variable $y^*$. But we observe $y = \max(0, y^*)$, so different research questions call for different marginal effects:

1. **"What does the household want to give?"** The effect on latent desired donations $y^*$, including negative desires. Use this when you want the structural parameter of the underlying model
2. **"Does an extra dollar of income bring new households into the donor pool?"** The effect on $P(y > 0)$, the extensive margin. Use this when the policy goal is to increase *participation*

# Three Questions, Three Marginal Effects

In OLS, $\hat{\beta}_1$ is the marginal effect, full stop. In Tobit, $\hat{\beta}_1 = 32.3$ is the effect on the **latent** variable $y^*$. But we observe $y = \max(0, y^*)$, so different research questions call for different marginal effects:

1. **"What does the household want to give?"** The effect on latent desired donations $y^*$, including negative desires. Use this when you want the structural parameter of the underlying model
2. **"Does an extra dollar of income bring new households into the donor pool?"** The effect on $P(y > 0)$, the extensive margin. Use this when the policy goal is to increase *participation*
3. **"How much more does the average household actually give, including those stuck at zero?"** The effect on unconditional observed $E[y]$. Use this when you want the overall impact on total giving

# Three Questions, Three Marginal Effects

In OLS, $\hat{\beta}_1$ is the marginal effect, full stop. In Tobit, $\hat{\beta}_1 = 32.3$ is the effect on the **latent** variable $y^*$. But we observe $y = \max(0, y^*)$, so different research questions call for different marginal effects:

1. **"What does the household want to give?"** The effect on latent desired donations $y^*$, including negative desires. Use this when you want the structural parameter of the underlying model
2. **"Does an extra dollar of income bring new households into the donor pool?"** The effect on $P(y > 0)$, the extensive margin. Use this when the policy goal is to increase *participation*
3. **"How much more does the average household actually give, including those stuck at zero?"** The effect on unconditional observed $E[y]$. Use this when you want the overall impact on total giving

$\Longrightarrow$ Each question has its own formula. Let's see them.

# Marginal Effect 1: Latent Desired Donations

**Question:** What does the household *want* to give?

# Marginal Effect 1: Latent Desired Donations

**Question:** What does the household *want* to give?

$$\frac{\partial\, \mathrm{E}[y_i^*]}{\partial\, \mathrm{Income}_i} = \beta_1 = 32.3$$

# Marginal Effect 1: Latent Desired Donations

**Question:** What does the household *want* to give?

$$\frac{\partial \, \mathrm{E}[y_i^*]}{\partial \, \mathrm{Income}_i} = \beta_1 = 32.3$$

This is the simplest: the coefficient itself. It tells us the effect on the latent outcome, as if nobody were constrained.

# Marginal Effect 1: Latent Desired Donations

**Question:** What does the household *want* to give?

$$\frac{\partial\, \mathrm{E}[y_i^*]}{\partial\, \mathrm{Income}_i} = \beta_1 = 32.3$$

This is the simplest: the coefficient itself. It tells us the effect on the latent outcome, as if nobody were constrained.

**Use this when** you are interested in the structural relationship between income and the desire to donate, ignoring the censoring constraint. This is the parameter OLS was trying (and failing) to estimate.

# Marginal Effects 2 and 3: Probability and Observed Amount

**Question 2:** Does an extra dollar of income bring new donors?

# Marginal Effects 2 and 3: Probability and Observed Amount

**Question 2:** Does an extra dollar of income bring new donors?

$$\frac{\partial P(y_i > 0)}{\partial \, \text{Income}_i} = \phi\left(\frac{\text{XB}_i}{\sigma}\right) \cdot \frac{\beta_1}{\sigma}$$

**Use this when** the policy goal is to expand the donor pool (extensive margin).

# Marginal Effects 2 and 3: Probability and Observed Amount

**Question 2:** Does an extra dollar of income bring new donors?

$$\frac{\partial P(y_i > 0)}{\partial \text{Income}_i} = \phi\left(\frac{\text{XB}_i}{\sigma}\right) \cdot \frac{\beta_1}{\sigma}$$

**Use this when** the policy goal is to expand the donor pool (extensive margin).

**Question 3:** How much more does the average household actually give?

# Marginal Effects 2 and 3: Probability and Observed Amount

**Question 2:** Does an extra dollar of income bring new donors?

$$\frac{\partial\, P(y_i > 0)}{\partial\, \text{Income}_i} = \phi\left(\frac{\text{XB}_i}{\sigma}\right) \cdot \frac{\beta_1}{\sigma}$$

**Use this when** the policy goal is to expand the donor pool (extensive margin).

**Question 3:** How much more does the average household actually give?

$$\frac{\partial\, \text{E}[y_i]}{\partial\, \text{Income}_i} = \Phi\left(\frac{\text{XB}_i}{\sigma}\right) \cdot \beta_1$$

**Use this when** you want the total impact on giving, combining both channels.

# Marginal Effects 2 and 3: Probability and Observed Amount

**Question 2:** Does an extra dollar of income bring new donors?

$$\frac{\partial\, P(y_i > 0)}{\partial\, \text{Income}_i} = \phi\left(\frac{\text{XB}_i}{\sigma}\right) \cdot \frac{\beta_1}{\sigma}$$

**Use this when** the policy goal is to expand the donor pool (extensive margin).

**Question 3:** How much more does the average household actually give?

$$\frac{\partial\ \text{E}[y_i]}{\partial\, \text{Income}_i} = \Phi\left(\frac{\text{XB}_i}{\sigma}\right) \cdot \beta_1$$

**Use this when** you want the total impact on giving, combining both channels.

$\Longrightarrow$ Effects 2 and 3 depend on *where* you evaluate them (which household). A household far from the censoring threshold has different marginal effects than one near it.

# Numeric Marginal Effects: Low-Income Household

Using $\hat{\beta}_1 = 32.3$ and $\hat{\sigma} = 1249$:

# Numeric Marginal Effects: Low-Income Household

Using $\hat{\beta}_1 = 32.3$ and $\hat{\sigma} = 1249$:

**Low-income household** (Income $= 30$, Educ $= 14$, Children $= 2$):

- $\text{XB} = -3324 + 32.3 \times 30 + 119 \times 14 - 16 \times 2 = -721$
- $\text{XB}/\sigma = -721/1249 = -0.58$

# Numeric Marginal Effects: Low-Income Household

Using $\hat{\beta}_1 = 32.3$ and $\hat{\sigma} = 1249$:

**Low-income household** (Income = 30, Educ = 14, Children = 2):

- $\text{XB} = -3324 + 32.3 \times 30 + 119 \times 14 - 16 \times 2 = -721$
- $\text{XB}/\sigma = -721/1249 = -0.58$
- ME on latent $y^*$: $\beta_1 = \mathbf{32.3}$ dollars per \$1000 income

# Numeric Marginal Effects: Low-Income Household

Using $\hat{\beta}_1 = 32.3$ and $\hat{\sigma} = 1249$:

**Low-income household** (Income $= 30$, Educ $= 14$, Children $= 2$):

- $\text{XB} = -3324 + 32.3 \times 30 + 119 \times 14 - 16 \times 2 = -721$
- $\text{XB}/\sigma = -721/1249 = -0.58$
- ME on latent $y^*$: $\beta_1 = \mathbf{32.3}$ dollars per \$1000 income
- ME on $P(y > 0)$: $\phi(-0.58) \times 32.3/1249 = \mathbf{0.009}$ (0.9 percentage points per \$1000)

# Numeric Marginal Effects: Low-Income Household

Using $\hat{\beta}_1 = 32.3$ and $\hat{\sigma} = 1249$:

**Low-income household** (Income $= 30$, Educ $= 14$, Children $= 2$):

- $\text{XB} = -3324 + 32.3 \times 30 + 119 \times 14 - 16 \times 2 = -721$
- $\text{XB}/\sigma = -721/1249 = -0.58$
- ME on latent $y^*$: $\beta_1 = \mathbf{32.3}$ dollars per \$1000 income
- ME on $P(y > 0)$: $\phi(-0.58) \times 32.3/1249 = \mathbf{0.009}$ (0.9 percentage points per \$1000)
- ME on $\text{E}[y]$: $\Phi(-0.58) \times 32.3 = 0.28 \times 32.3 = \$\mathbf{9.1}$ per \$1000 income

# Numeric Marginal Effects: Low-Income Household

Using $\hat{\beta}_1 = 32.3$ and $\hat{\sigma} = 1249$:

**Low-income household** (Income $= 30$, Educ $= 14$, Children $= 2$):

- $\text{XB} = -3324 + 32.3 \times 30 + 119 \times 14 - 16 \times 2 = -721$
- $\text{XB}/\sigma = -721/1249 = -0.58$
- ME on latent $y^*$: $\beta_1 = \mathbf{32.3}$ dollars per \$1000 income
- ME on $P(y > 0)$: $\phi(-0.58) \times 32.3/1249 = \mathbf{0.009}$ (0.9 percentage points per \$1000)
- ME on $\text{E}[y]$: $\Phi(-0.58) \times 32.3 = 0.28 \times 32.3 = \$\mathbf{9.1}$ per \$1000 income

$\Longrightarrow$ For this household (only 28% chance of donating), most of the income effect is "absorbed" by the probability margin. The unconditional effect (\$9.1) is far below the latent effect (\$32.3).

# Numeric Marginal Effects: High-Income Household

**High-income household** (Income $= 100$, Educ $= 16$, Children $= 1$):

- $\text{XB} = -3324 + 32.3 \times 100 + 119 \times 16 - 16 \times 1 = 1794$
- $\text{XB}/\sigma = 1794/1249 = 1.44$

# Numeric Marginal Effects: High-Income Household

**High-income household** (Income $= 100$, Educ $= 16$, Children $= 1$):

- $\text{XB} = -3324 + 32.3 \times 100 + 119 \times 16 - 16 \times 1 = 1794$
- $\text{XB}/\sigma = 1794/1249 = 1.44$
- ME on latent $y^*$: $\beta_1 =$ **32.3** dollars per \$1000 income

# Numeric Marginal Effects: High-Income Household

**High-income household** (Income $= 100$, Educ $= 16$, Children $= 1$):

- $\text{XB} = -3324 + 32.3 \times 100 + 119 \times 16 - 16 \times 1 = 1794$
- $\text{XB}/\sigma = 1794/1249 = 1.44$
- ME on latent $y^*$: $\beta_1 = \mathbf{32.3}$ dollars per \$1000 income
- ME on $P(y > 0)$: $\phi(1.44) \times 32.3/1249 = \mathbf{0.004}$ (0.4 pp per \$1000)

# Numeric Marginal Effects: High-Income Household

**High-income household** (Income $= 100$, Educ $= 16$, Children $= 1$):

- $\text{XB} = -3324 + 32.3 \times 100 + 119 \times 16 - 16 \times 1 = 1794$
- $\text{XB}/\sigma = 1794/1249 = 1.44$
- ME on latent $y^*$: $\beta_1 =$ **32.3** dollars per \$1000 income
- ME on $P(y > 0)$: $\phi(1.44) \times 32.3/1249 =$ **0.004** (0.4 pp per \$1000)
- ME on E[$y$]: $\Phi(1.44) \times 32.3 = 0.925 \times 32.3 =$ \$**29.9** per \$1000 income

# Numeric Marginal Effects: High-Income Household

**High-income household** (Income = 100, Educ = 16, Children = 1):

- $\text{XB} = -3324 + 32.3 \times 100 + 119 \times 16 - 16 \times 1 = 1794$
- $\text{XB}/\sigma = 1794/1249 = 1.44$
- ME on latent $y^*$: $\beta_1 = \mathbf{32.3}$ dollars per \$1000 income
- ME on $P(y > 0)$: $\phi(1.44) \times 32.3/1249 = \mathbf{0.004}$ (0.4 pp per \$1000)
- ME on $\text{E}[y]$: $\Phi(1.44) \times 32.3 = 0.925 \times 32.3 = \$\mathbf{29.9}$ per \$1000 income

$\Longrightarrow$ This household is nearly certain to donate (92%), so almost all of the latent effect passes through to the observed outcome. The unconditional ME (\$29.9) is close to the raw coefficient (\$32.3).

# Comparing Marginal Effects Across Households

| | **Latent** $y^*$ | $P(y > 0)$ | $\mathrm{E}[y]$ |
|---|---|---|---|
| Low-income ($\mathrm{XB}/\sigma = -0.58$) | \$32.3 | 0.9 pp | \$9.1 |
| High-income ($\mathrm{XB}/\sigma = 1.44$) | \$32.3 | 0.4 pp | \$29.9 |

# Comparing Marginal Effects Across Households

| | **Latent** $y^*$ | $P(y > 0)$ | E[$y$] |
|---|---|---|---|
| Low-income ($XB/\sigma = -0.58$) | \$32.3 | 0.9 pp | \$9.1 |
| High-income ($XB/\sigma = 1.44$) | \$32.3 | 0.4 pp | \$29.9 |

The latent effect is always the same (\$32.3). But the observed effects differ dramatically:

# Comparing Marginal Effects Across Households

| | **Latent** $y^*$ | $P(y > 0)$ | $\mathrm{E}[y]$ |
|---|---|---|---|
| Low-income ($XB/\sigma = -0.58$) | \$32.3 | 0.9 pp | \$9.1 |
| High-income ($XB/\sigma = 1.44$) | \$32.3 | 0.4 pp | \$29.9 |

The latent effect is always the same (\$32.3). But the observed effects differ dramatically:

- For the low-income household, income mainly affects *whether* they donate (extensive margin)
- For the high-income household, income mainly affects *how much* they donate (intensive margin)

# Comparing Marginal Effects Across Households

| | **Latent** $y^*$ | $P(y > 0)$ | E[$y$] |
|---|---|---|---|
| Low-income ($XB/\sigma = -0.58$) | \$32.3 | 0.9 pp | \$9.1 |
| High-income ($XB/\sigma = 1.44$) | \$32.3 | 0.4 pp | \$29.9 |

The latent effect is always the same (\$32.3). But the observed effects differ dramatically:

- For the low-income household, income mainly affects *whether* they donate (extensive margin)
- For the high-income household, income mainly affects *how much* they donate (intensive margin)

$\Longrightarrow$ Where a household sits relative to the censoring threshold determines which margin dominates.

# Boundary Condition: When Tobit Reduces to OLS

Look at the unconditional marginal effect formula again:

$$\frac{\partial\, \mathrm{E}[y_i]}{\partial\, x_k} = \Phi\left(\frac{\mathrm{XB}_i}{\sigma}\right) \cdot \beta_k$$

# Boundary Condition: When Tobit Reduces to OLS

Look at the unconditional marginal effect formula again:

$$\frac{\partial \, \mathrm{E}[y_i]}{\partial \, x_k} = \Phi\left(\frac{\mathrm{XB}_i}{\sigma}\right) \cdot \beta_k$$

**When $\mathbf{XB}_i/\sigma$ is very large** (almost everyone donates):

- $\Phi(\mathrm{XB}_i/\sigma) \to 1$
- The marginal effect $\to \beta_k$
- Tobit behaves like OLS

# Boundary Condition: When Tobit Reduces to OLS

Look at the unconditional marginal effect formula again:

$$\frac{\partial \, E[y_i]}{\partial \, x_k} = \Phi\left(\frac{XB_i}{\sigma}\right) \cdot \beta_k$$

**When $\mathbf{XB}_i/\sigma$ is very large** (almost everyone donates):

- $\Phi(XB_i/\sigma) \to 1$
- The marginal effect $\to \beta_k$
- Tobit behaves like OLS

**When $\mathbf{XB}_i/\sigma$ is very negative** (almost no one donates):

- $\Phi(XB_i/\sigma) \to 0$
- The marginal effect $\to 0$
- Income increases mostly change the *probability* of donating, not the amount

# Boundary Condition: When Tobit Reduces to OLS

Look at the unconditional marginal effect formula again:

$$\frac{\partial\, \mathrm{E}[y_i]}{\partial\, x_k} = \Phi\left(\frac{\mathrm{XB}_i}{\sigma}\right) \cdot \beta_k$$

**When $\mathbf{XB}_i/\sigma$ is very large** (almost everyone donates):

- $\Phi(\mathrm{XB}_i/\sigma) \to 1$
- The marginal effect $\to \beta_k$
- Tobit behaves like OLS

**When $\mathbf{XB}_i/\sigma$ is very negative** (almost no one donates):

- $\Phi(\mathrm{XB}_i/\sigma) \to 0$
- The marginal effect $\to 0$
- Income increases mostly change the *probability* of donating, not the amount

$\Longrightarrow$ If censoring is rare (few zeros), Tobit and OLS give similar answers. The more censoring, the more Tobit differs from OLS.

# The McDonald-Moffitt Decomposition

A policymaker wants to increase total charitable giving. Should they target existing donors to give more (intensive margin), or convert non-donors into donors (extensive margin)? The McDonald-Moffitt decomposition answers exactly this.

# The McDonald-Moffitt Decomposition

A policymaker wants to increase total charitable giving. Should they target existing donors to give more (intensive margin), or convert non-donors into donors (extensive margin)? The McDonald-Moffitt decomposition answers exactly this.

McDonald and Moffitt (1980) showed that the unconditional marginal effect (Effect #3 from two slides ago) can be split into two channels via the product rule:

# The McDonald-Moffitt Decomposition

A policymaker wants to increase total charitable giving. Should they target existing donors to give more (intensive margin), or convert non-donors into donors (extensive margin)? The McDonald-Moffitt decomposition answers exactly this.

McDonald and Moffitt (1980) showed that the unconditional marginal effect (Effect #3 from two slides ago) can be split into two channels via the product rule:

$$\underbrace{\frac{\partial \, \mathrm{E}[y]}{\partial x_k}}_{\text{total effect}} = \underbrace{P(y > 0) \cdot \frac{\partial \, \mathrm{E}[y \mid y > 0]}{\partial x_k}}_{\text{how much more do donors give?}} + \underbrace{\mathrm{E}[y \mid y > 0] \cdot \frac{\partial \, P(y > 0)}{\partial x_k}}_{\text{how many new donors?}}$$

# The McDonald-Moffitt Decomposition

A policymaker wants to increase total charitable giving. Should they target existing donors to give more (intensive margin), or convert non-donors into donors (extensive margin)? The McDonald-Moffitt decomposition answers exactly this.

McDonald and Moffitt (1980) showed that the unconditional marginal effect (Effect #3 from two slides ago) can be split into two channels via the product rule:

$$\underbrace{\frac{\partial\, \mathrm{E}[y]}{\partial\, x_k}}_{\text{total effect}} = \underbrace{P(y > 0) \cdot \frac{\partial\, \mathrm{E}[y \mid y > 0]}{\partial\, x_k}}_{\text{how much more do donors give?}} + \underbrace{\mathrm{E}[y \mid y > 0] \cdot \frac{\partial\, P(y > 0)}{\partial\, x_k}}_{\text{how many new donors?}}$$

**Channel 1 (intensive):** Among current donors, income raises donations by $X.
**Channel 2 (extensive):** Higher income brings new households into the donor pool, each contributing $\mathrm{E}[y \mid y > 0]$.

# The McDonald-Moffitt Decomposition

A policymaker wants to increase total charitable giving. Should they target existing donors to give more (intensive margin), or convert non-donors into donors (extensive margin)? The McDonald-Moffitt decomposition answers exactly this.

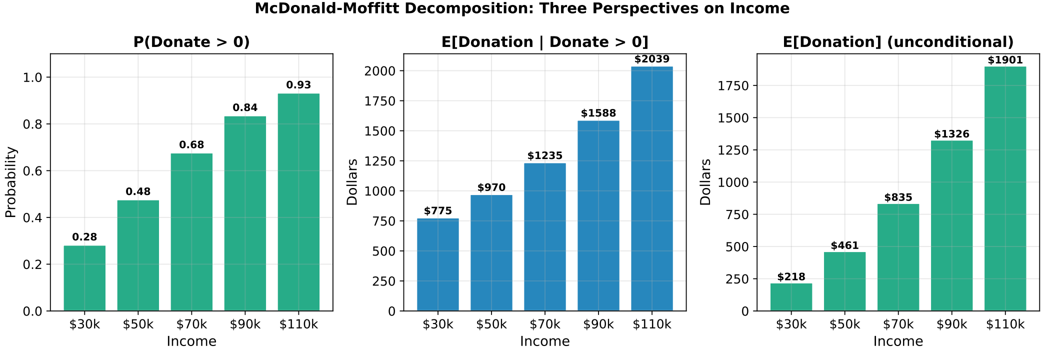
McDonald and Moffitt (1980) showed that the unconditional marginal effect (Effect #3 from two slides ago) can be split into two channels via the product rule:

$$\underbrace{\frac{\partial\, \mathrm{E}[y]}{\partial\, x_k}}_{\text{total effect}} = \underbrace{P(y>0)\cdot\frac{\partial\, \mathrm{E}[y\mid y>0]}{\partial\, x_k}}_{\text{how much more do donors give?}} + \underbrace{\mathrm{E}[y\mid y>0]\cdot\frac{\partial\, P(y>0)}{\partial\, x_k}}_{\text{how many new donors?}}$$

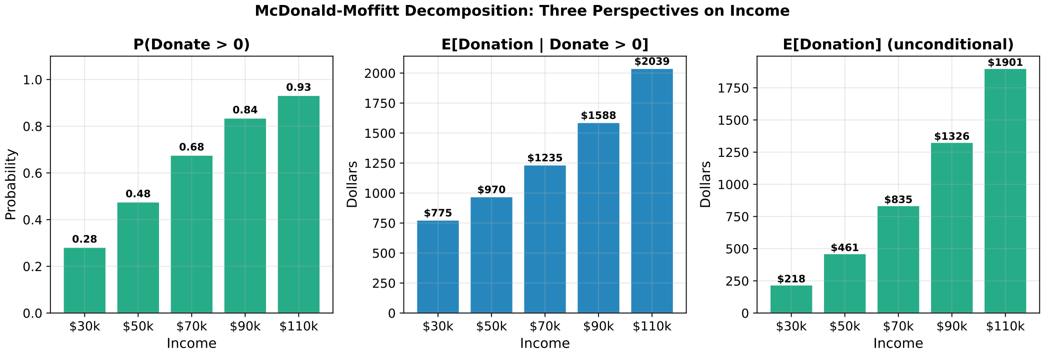
**Channel 1 (intensive):** Among current donors, income raises donations by $X.
**Channel 2 (extensive):** Higher income brings new households into the donor pool, each contributing $\mathrm{E}[y \mid y > 0]$.

$\Longrightarrow$ A single coefficient $\beta_1$ drives *both* the probability of participation and the level of giving. This is the defining feature (and restriction) of the Tobit model.

# Visualizing the Decomposition



# Visualizing the Decomposition




At Income = \$30k, only 28% donate; at \$110k, 93% donate. Among donors, average giving rises from \$775 to \$2,039. Both channels contribute to unconditional E[$y$] rising from \$218 to \$1,901.

# Outline

# Censoring vs. Truncation: They Sound Similar but Differ

These two terms describe different data situations. Consider the same donation study:

# Censoring vs. Truncation: They Sound Similar but Differ

These two terms describe different data situations. Consider the same donation study:

**Censoring** (Tobit):

- The researcher surveys *all* 500 households
- For non-donors, she records $y_i = 0$ along with their income, education, and children
- She knows the non-donors exist and can count them

# Censoring vs. Truncation: They Sound Similar but Differ

These two terms describe different data situations. Consider the same donation study:

**Censoring** (Tobit):

- The researcher surveys *all* 500 households
- For non-donors, she records $y_i = 0$ along with their income, education, and children
- She knows the non-donors exist and can count them

**Truncation**:

- The researcher only has records from a charity's donor database
- She observes the 299 donors $(y_i > 0)$ and their characteristics
- Non-donors are *completely absent*: she does not even know how many there are

# Censoring vs. Truncation: They Sound Similar but Differ

These two terms describe different data situations. Consider the same donation study:

**Censoring** (Tobit):

- The researcher surveys *all* 500 households
- For non-donors, she records $y_i = 0$ along with their income, education, and children
- She knows the non-donors exist and can count them

**Truncation**:

- The researcher only has records from a charity's donor database
- She observes the 299 donors ($y_i > 0$) and their characteristics
- Non-donors are *completely absent*: she does not even know how many there are

$\Longrightarrow$ With censoring, the zeros are **in the data**. With truncation, the zeros are **missing entirely**.

# When Do You Face Each Situation?

**Censored examples** (use Tobit):

- Charitable donations: non-donors report \$0, all households surveyed
- Hours worked: non-workers report 0 hours, all individuals in the sample
- Expenditure on durable goods: many households spend \$0 on new cars

# When Do You Face Each Situation?

**Censored examples** (use Tobit):

- Charitable donations: non-donors report \$0, all households surveyed
- Hours worked: non-workers report 0 hours, all individuals in the sample
- Expenditure on durable goods: many households spend \$0 on new cars

**Truncated examples** (use truncated regression):

- Firm profits: only firms that survived to be surveyed appear; failed firms are gone
- Scholarship amounts: only recipients are in the data; rejected applicants are absent
- Wages: only observed for employed workers; non-workers absent from the data. (In practice, wages are usually better handled by the **Heckman selection model**, since the decision to work is a separate process from wage determination.)

# When Do You Face Each Situation?

**Censored examples** (use Tobit):

- Charitable donations: non-donors report \$0, all households surveyed
- Hours worked: non-workers report 0 hours, all individuals in the sample
- Expenditure on durable goods: many households spend \$0 on new cars

**Truncated examples** (use truncated regression):

- Firm profits: only firms that survived to be surveyed appear; failed firms are gone
- Scholarship amounts: only recipients are in the data; rejected applicants are absent
- Wages: only observed for employed workers; non-workers absent from the data. (In practice, wages are usually better handled by the **Heckman selection model**, since the decision to work is a separate process from wage determination.)

$\Longrightarrow$ If the zeros are in your data, it is censoring. If you only see the positive values and do not know how many zeros were excluded, it is truncation.

# Outline

# The Tobit Model Rests on Two Assumptions

**Assumption 1: Normal, homoskedastic errors.**

$$\varepsilon_i \sim N(0, \sigma^2), \quad \text{independent of covariates}$$

# The Tobit Model Rests on Two Assumptions

**Assumption 1: Normal, homoskedastic errors.**

$$\varepsilon_i \sim N(0, \sigma^2), \quad \text{independent of covariates}$$

Two ways this can fail:

- **Non-normality:** skewed or heavy-tailed errors
- **Heteroskedasticity:** $\text{Var}(\varepsilon_i)$ depends on covariates. Unlike OLS, where heteroskedasticity only affects standard errors, in Tobit it makes the **coefficient estimates themselves inconsistent**

# The Tobit Model Rests on Two Assumptions

**Assumption 1: Normal, homoskedastic errors.**

$$\varepsilon_i \sim N(0, \sigma^2), \quad \text{independent of covariates}$$

Two ways this can fail:

- **Non-normality:** skewed or heavy-tailed errors
- **Heteroskedasticity:** $\text{Var}(\varepsilon_i)$ depends on covariates. Unlike OLS, where heteroskedasticity only affects standard errors, in Tobit it makes the **coefficient estimates themselves inconsistent**

**Assumption 2: Same mechanism governs participation and amount.**
The same $\beta_1$ determines both *whether* a household donates and *how much*. This is a strong restriction.

# The Tobit Model Rests on Two Assumptions

**Assumption 1: Normal, homoskedastic errors.**

$$\varepsilon_i \sim N(0, \sigma^2), \quad \text{independent of covariates}$$

Two ways this can fail:

- **Non-normality:** skewed or heavy-tailed errors
- **Heteroskedasticity:** $\text{Var}(\varepsilon_i)$ depends on covariates. Unlike OLS, where heteroskedasticity only affects standard errors, in Tobit it makes the **coefficient estimates themselves inconsistent**

**Assumption 2: Same mechanism governs participation and amount.**
The same $\beta_1$ determines both *whether* a household donates and *how much*. This is a strong restriction.

**Example where this fails:** a very wealthy household may decide whether to donate based on social pressure ($\beta_1^{\text{participate}}$), but the amount depends on tax incentives ($\beta_1^{\text{amount}} \neq \beta_1^{\text{participate}}$).

# The Tobit Model Rests on Two Assumptions

**Assumption 1: Normal, homoskedastic errors.**

$$\varepsilon_i \sim N(0, \sigma^2), \quad \text{independent of covariates}$$

Two ways this can fail:

- **Non-normality:** skewed or heavy-tailed errors
- **Heteroskedasticity:** $\text{Var}(\varepsilon_i)$ depends on covariates. Unlike OLS, where heteroskedasticity only affects standard errors, in Tobit it makes the **coefficient estimates themselves inconsistent**

**Assumption 2: Same mechanism governs participation and amount.**
The same $\beta_1$ determines both *whether* a household donates and *how much*. This is a strong restriction.

**Example where this fails:** a very wealthy household may decide whether to donate based on social pressure ($\beta_1^{\text{participate}}$), but the amount depends on tax incentives ($\beta_1^{\text{amount}} \neq \beta_1^{\text{participate}}$).

$\Longrightarrow$ When Assumption 2 fails, the Tobit model forces a single coefficient to represent two distinct processes. What alternatives exist?

# When Assumption 2 Fails: Alternatives

**Two-Part Model** (also called the “hurdle” model):

- Part 1: Probit or logit for $P(y > 0)$ with coefficients $\gamma$
- Part 2: OLS or truncated regression for $\mathrm{E}[y \mid y > 0]$ with coefficients $\delta$
- $\gamma$ and $\delta$ are estimated **separately**, so the participation and amount decisions can differ

# When Assumption 2 Fails: Alternatives

**Two-Part Model** (also called the "hurdle" model):

- Part 1: Probit or logit for $P(y > 0)$ with coefficients $\gamma$
- Part 2: OLS or truncated regression for $\mathrm{E}[y \mid y > 0]$ with coefficients $\delta$
- $\gamma$ and $\delta$ are estimated **separately**, so the participation and amount decisions can differ

**Heckman Selection Model**:

- For situations where the zeros are not corner solutions but **sample selection**
- Example: wages are only observed for people who choose to work; the decision to work is a separate equation
- Adds a selection correction (inverse Mills ratio) to the outcome equation

# When Assumption 2 Fails: Alternatives

**Two-Part Model** (also called the "hurdle" model):

- Part 1: Probit or logit for $P(y > 0)$ with coefficients $\gamma$
- Part 2: OLS or truncated regression for $\mathrm{E}[y \mid y > 0]$ with coefficients $\delta$
- $\gamma$ and $\delta$ are estimated **separately**, so the participation and amount decisions can differ

**Heckman Selection Model**:

- For situations where the zeros are not corner solutions but **sample selection**
- Example: wages are only observed for people who choose to work; the decision to work is a separate equation
- Adds a selection correction (inverse Mills ratio) to the outcome equation

$\Longrightarrow$ Tobit assumes one mechanism. The two-part model relaxes that. The Heckman model is for a fundamentally different problem: selection, not censoring.

# Decision Flowchart: Choosing the Right Model

1. Is your outcome non-negative with a pile-up at zero?

# Decision Flowchart: Choosing the Right Model

1. Is your outcome non-negative with a pile-up at zero?
2. **Are the zeros corner solutions?** (The person *wants* a negative value but is constrained to zero.)
   - Yes, and you believe the same $\beta$ governs both participation and amount $\implies$ **Tobit**

# Decision Flowchart: Choosing the Right Model

1. Is your outcome non-negative with a pile-up at zero?
2. **Are the zeros corner solutions?** (The person *wants* a negative value but is constrained to zero.)
   - Yes, and you believe the same $\beta$ governs both participation and amount $\implies$ **Tobit**
   - Yes, but participation and amount may have different drivers $\implies$ **Two-Part Model**

# Decision Flowchart: Choosing the Right Model

1. Is your outcome non-negative with a pile-up at zero?
2. **Are the zeros corner solutions?** (The person *wants* a negative value but is constrained to zero.)
   - Yes, and you believe the same $\beta$ governs both participation and amount $\Longrightarrow$ **Tobit**
   - Yes, but participation and amount may have different drivers $\Longrightarrow$ **Two-Part Model**
3. **Are the zeros from sample selection?** (The outcome *exists* but you do not observe it.)
   - A worker has a wage, but you only observe it if they participate in the labor market $\Longrightarrow$ **Heckman Selection Model**

# Decision Flowchart: Choosing the Right Model

1. Is your outcome non-negative with a pile-up at zero?
2. **Are the zeros corner solutions?** (The person *wants* a negative value but is constrained to zero.)
   - Yes, and you believe the same $\beta$ governs both participation and amount $\Longrightarrow$ **Tobit**
   - Yes, but participation and amount may have different drivers $\Longrightarrow$ **Two-Part Model**
3. **Are the zeros from sample selection?** (The outcome *exists* but you do not observe it.)
   - A worker has a wage, but you only observe it if they participate in the labor market $\Longrightarrow$ **Heckman Selection Model**
4. **Is your outcome a count?** (Non-negative integers: 0, 1, 2, . . . )
   - $\Longrightarrow$ **Poisson / Negative Binomial**, not Tobit

# Outline

# Summary: Back to Charitable Donations

1. **The data problem:** 40% of households donate \$0, creating a spike at zero with a continuous positive tail. This is a **corner solution**

# Summary: Back to Charitable Donations

1. **The data problem:** 40% of households donate \$0, creating a spike at zero with a continuous positive tail. This is a **corner solution**
2. **OLS fails twice:** OLS on all data gives slope $= 19.7$ (attenuated by the pile-up). OLS on positives gives slope $= 17.8$ (distorted by sample selection). True latent slope $= 30$

# Summary: Back to Charitable Donations

1. **The data problem:** 40% of households donate \$0, creating a spike at zero with a continuous positive tail. This is a **corner solution**
2. **OLS fails twice:** OLS on all data gives slope $= 19.7$ (attenuated by the pile-up). OLS on positives gives slope $= 17.8$ (distorted by sample selection). True latent slope $= 30$
3. **The Tobit model** uses a latent variable $y^* = \beta_0 + \beta_1 \text{Income} + \cdots + \varepsilon$ with $y = \max(0, y^*)$. Its likelihood combines a probit component (zeros) and a regression component (positives)

# Summary: Back to Charitable Donations

1. **The data problem:** 40% of households donate $0, creating a spike at zero with a continuous positive tail. This is a **corner solution**
2. **OLS fails twice:** OLS on all data gives slope $= 19.7$ (attenuated by the pile-up). OLS on positives gives slope $= 17.8$ (distorted by sample selection). True latent slope $= 30$
3. **The Tobit model** uses a latent variable $y^* = \beta_0 + \beta_1 \text{Income} + \cdots + \varepsilon$ with $y = \max(0, y^*)$. Its likelihood combines a probit component (zeros) and a regression component (positives)
4. **Tobit recovers the latent slope:** $\hat{\beta}_1 = 32.3 \approx 30$. Three types of marginal effects tell you the effect on latent $y^*$, on $P(y > 0)$, and on $\text{E}[y]$

# Summary: Back to Charitable Donations

1. **The data problem:** 40% of households donate \$0, creating a spike at zero with a continuous positive tail. This is a **corner solution**
2. **OLS fails twice:** OLS on all data gives slope $= 19.7$ (attenuated by the pile-up). OLS on positives gives slope $= 17.8$ (distorted by sample selection). True latent slope $= 30$
3. **The Tobit model** uses a latent variable $y^* = \beta_0 + \beta_1 \text{Income} + \cdots + \varepsilon$ with $y = \max(0, y^*)$. Its likelihood combines a probit component (zeros) and a regression component (positives)
4. **Tobit recovers the latent slope:** $\hat{\beta}_1 = 32.3 \approx 30$. Three types of marginal effects tell you the effect on latent $y^*$, on $P(y > 0)$, and on $\text{E}[y]$
5. **The McDonald-Moffitt decomposition** splits the total effect into a participation channel (new donors) and an amount channel (existing donors give more)

# Summary: Back to Charitable Donations

1. **The data problem:** 40% of households donate \$0, creating a spike at zero with a continuous positive tail. This is a **corner solution**
2. **OLS fails twice:** OLS on all data gives slope $= 19.7$ (attenuated by the pile-up). OLS on positives gives slope $= 17.8$ (distorted by sample selection). True latent slope $= 30$
3. **The Tobit model** uses a latent variable $y^* = \beta_0 + \beta_1 \text{Income} + \cdots + \varepsilon$ with $y = \max(0, y^*)$. Its likelihood combines a probit component (zeros) and a regression component (positives)
4. **Tobit recovers the latent slope:** $\hat{\beta}_1 = 32.3 \approx 30$. Three types of marginal effects tell you the effect on latent $y^*$, on $P(y > 0)$, and on $\text{E}[y]$
5. **The McDonald-Moffitt decomposition** splits the total effect into a participation channel (new donors) and an amount channel (existing donors give more)
6. **Tobit assumes one mechanism** for participation and amount. When that fails, use a two-part model. When zeros come from selection rather than censoring, use Heckman

# Comparison Table: Tobit vs. Alternatives

| | **Tobit** | **Two-Part** | **Heckman** |
|---|---|---|---|
| Zero mechanism | corner solution | corner solution | selection |
| Same $\beta$ for participation & amount? | yes | no (separate $\gamma$, $\delta$) | no (separate equations) |
| Normality required? | yes | only if Part 2 uses truncated regression | yes (or semi-parametric) |
| Typical example | donations, hours worked | health expenditures | wages (if employed) |

# Comparison Table: Tobit vs. Alternatives

| | **Tobit** | **Two-Part** | **Heckman** |
|---|---|---|---|
| Zero mechanism | corner solution | corner solution | selection |
| Same $\beta$ for participation & amount? | yes | no (separate $\gamma$, $\delta$) | no (separate equations) |
| Normality required? | yes | only if Part 2 uses truncated regression | yes (or semi-parametric) |
| Typical example | donations, hours worked | health expenditures | wages (if employed) |

$\Longrightarrow$ The correct choice depends on the economic mechanism generating the zeros: censoring from a corner solution, or selection from a missing-data process.

Thank you!

jakeanderson@g.ucla.edu