Binary Outcomes: LPM, Logit, and Probit

16 Nov 2026 | Chapter 11

ECO 351 – Introduction to Econometrics

Shi Qiu – Fall 2026

Recall: an estimate is not the whole conclusion

Last class, a ten-cent beer-tax increase had a fitted difference of -0.064 deaths per 10,000 residents. Its 95% interval was approximately $[-0.136, 0.008]$.

1. Does the interval establish that the effect is exactly zero?
2. Would removing a changing enforcement policy from the data make the tax estimate automatically causal?

We will keep the distinction between **a fitted comparison**, **uncertainty**, and **causal assumptions** when the outcome changes.

New outcome, the same discipline of interpretation

An interval containing zero does not prove no effect. Leaving a confounder unobserved does not remove it.

Now consider **2,380 Boston mortgage applications**. Define $Y = 1$ for denial and $Y = 0$ for approval.

```
. tabulate deny
       deny |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      2,095       88.03       88.03
          1 |        285       11.97      100.00
------------+-----------------------------------
      Total |      2,380      100.00
```

$X = PIR$: debt payments relative to income. $PIR = 0.30$ means payments equal 30% of income.

Why regression can model a zero-one outcome

For a zero-one outcome, the sample mean is the proportion of ones:

$$\overline{Y} = \frac{285}{2380} \approx 0.120.$$

At a particular $PIR = x$, the population conditional mean is

$$E(Y \mid X = x) = 1\,P(Y = 1 \mid X = x) + 0\,P(Y = 0 \mid X = x).$$

Thus $E(Y \mid X = x) = P(Y = 1 \mid X = x)$. Write this probability as $p(x)$.

The observed outcome is zero or one. The conditional mean is a probability, which can lie between them.

Start with a linear probability model

The **linear probability model (LPM)** specifies $p(x) = \beta_0 + \beta_1 x$ over the relevant range.

```
. regress deny pir, vce(robust)
Linear regression                               Number of obs     =      2,380
------------------------------------------------------------------------------
             |               Robust
        deny | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         pir |   .6035349   .0984826     6.13   0.000     .4104144    .7966555
       _cons |  -.0799096   .0319666    -2.50   0.012    -.1425949   -.0172243
------------------------------------------------------------------------------
```

The fitted line is $\widehat{p}(x) \approx -0.080 + 0.604x$.

$\widehat{\beta}_1$ is in probability units per one-unit increase in PIR.

Compare two applicants in probability units

Using the **rounded fitted LPM**,

$$\widehat{p}(x) = -0.080 + 0.604x,$$

PIR	Fitted probability	As a percent
0.30	0.1012	10.12%
0.40	0.1616	16.16%

$$\Delta\widehat{p} = 0.1616 - 0.1012 = 0.0604 = 6.04 \text{ percentage points.}$$

This is a fitted comparison, not a claim that changing one applicant's debt causes denial.

Would robust standard errors fix this prediction?

A lending analyst uses $\widehat{p}(x) = -0.080 + 0.604x$.

1. Calculate the prediction at $PIR = 0.05$. Is it a valid probability?
2. The analyst proposes adding `vce(robust)` to fix it. Would that change the prediction?
3. Under this line, would a 0.10 PIR increase have a different probability effect at a high starting PIR?

Distinguish a limitation of the conditional-mean model from a problem with its standard errors.

Two restrictions of the straight line

At $PIR = 0.05$,

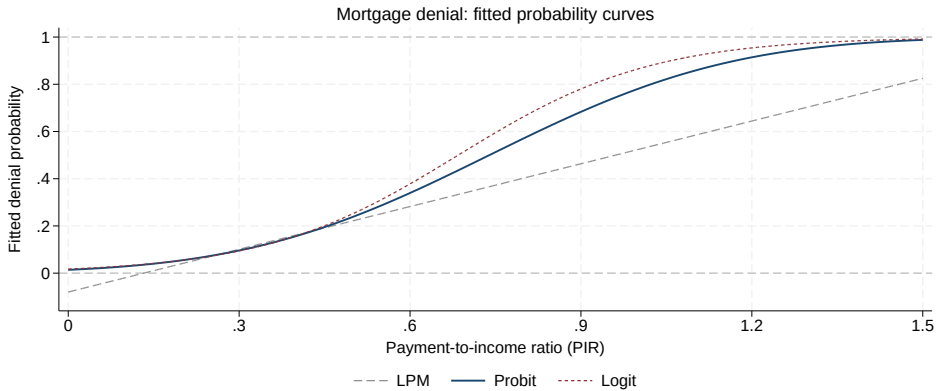
$$\widehat{p}(0.05) = -0.080 + 0.604(0.05) = -0.0498.$$

- A negative probability is impossible; the LPM does not enforce bounds.
- Every 0.10 PIR increase changes the fitted probability by 0.0604.

For binary Y, $\text{Var}(Y \mid X) = p(X)[1 - p(X)]$, which generally varies with X.

Robust standard errors address this variance issue. They repair neither the bounds nor the constant slope.

Keep probabilities inside their natural bounds



The wide PIR range illustrates model shape; extreme values are sparsely observed. Bounded curves flatten near zero and one.

Probit: turn an index into a probability

Let $\Phi(z) = P(Z \leq z)$ for a standard normal variable Z.

$$p(x) = \Phi(\eta), \qquad \eta = \beta_0 + \beta_1 x.$$

1. Form the **index** η using the regressor and coefficients.
2. Apply the normal CDF Φ to obtain a probability.
3. For a comparison, repeat at the other x and subtract.

β_1 changes the index per unit of x; it is not a constant probability effect.

Stata: estimate the probit index

```
. probit deny pir, vce(robust)
Probit regression                                       Number of obs =  2,380
------------------------------------------------------------------------------
             |               Robust
        deny | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         pir |   2.967907   .4653217     6.38   0.000     2.055893    3.879921
       _cons |  -2.194159   .1649761   -13.30   0.000    -2.517506   -1.870812
------------------------------------------------------------------------------
Note: 0 failures and 1 success completely determined.
```

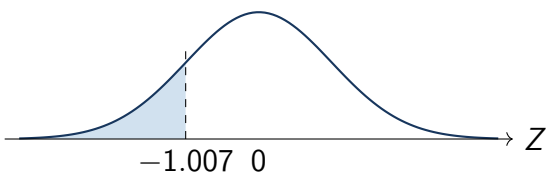
$$\widehat{\eta}(x) = -2.194159 + 2.967907x.$$

A 0.10 PIR increase changes the **index** by 0.296791. We still need Φ to find the probability change.

From an index to an area under the normal curve

At $PIR = 0.40$, the fitted index is

$$\widehat{\eta} = -2.194159 + 2.967907(0.40) \approx -1.007.$$



$$\widehat{p}(0.40) = \Phi(-1.007) \approx 0.157.$$

The shaded left-tail area is the fitted denial probability: about **15.7%**.

Stata: obtain probabilities at the comparison values

```
. margins, at(pir=(0.30 0.40 0.50))
Adjusted predictions                                    Number of obs = 2,380
------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         _at |
          1  |   .0961532   .0069017    13.93   0.000      .082626    .1096803
          2  |   .1569683   .0101166    15.52   0.000     .1371402    .1767965
          3  |   .2387884    .024485     9.75   0.000     .1907987    .2867781
------------------------------------------------------------------------------
```

This follows the fitted probit. Rows 1–3 set PIR to 0.30, 0.40, and 0.50.

The output contains probabilities; multiply a *difference* by 100 to report percentage points.

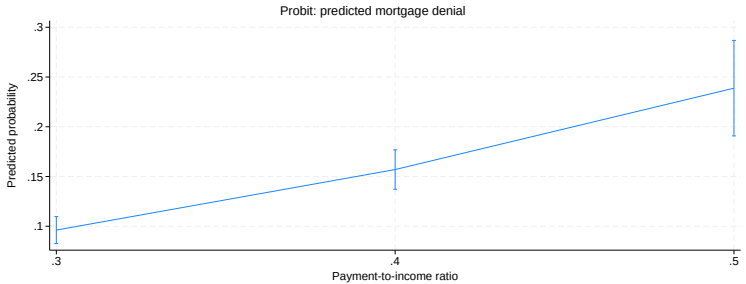
Student task: two equal increases, two different answers

Use the three fitted probit probabilities:

PIR	$\widehat{p}$	Denial probability
0.30	0.0961532	9.61532%
0.40	0.1569683	15.69683%
0.50	0.2387884	23.87884%

1. Compare $0.30 \to 0.40$ and $0.40 \to 0.50$ in percentage points.
2. Which increase crosses the steeper part of the fitted curve?
3. A colleague uses 2.967907(0.10) for each answer. Explain exactly what that calculation measures.

The starting probability changes the fitted response



$$100(0.1569683 - 0.0961532) = 6.08 \text{ percentage points},$$

$$100(0.2387884 - 0.1569683) = 8.18 \text{ percentage points}.$$

The same index change crosses a steeper part of the CDF in the second comparison.

Logit uses another bounded transformation

Logit uses the logistic CDF, denoted Λ:

$$p(x) = \Lambda(\eta) = \frac{1}{1 + \exp(-\eta)}, \qquad \eta = \beta_0 + \beta_1 x.$$

- Probit applies the normal CDF; logit applies the logistic CDF.
- Both turn an index into a probability between zero and one.
- Coefficients use different index scales. Compare probabilities for the **same change in PIR**.

Stata: the logit coefficient is on an index scale

```
. logit deny pir, vce(robust)
Logistic regression                                     Number of obs =  2,380
------------------------------------------------------------------------------
             |               Robust
        deny | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         pir |   5.884498   .9999333     5.88   0.000     3.924664    7.844331
       _cons |  -4.028432   .3589036   -11.22   0.000     -4.73187   -3.324994
------------------------------------------------------------------------------
```

The fitted logit index is $\widehat{\eta}(x) = -4.028432 + 5.884498x$.

The larger slope is *not* evidence of a probability effect nearly twice as large as probit's.

Stata: compare probabilities, not index coefficients

```
. margins, at(pir=(0.30 0.40 0.50))
Adjusted predictions                                            Number of obs = 2,380
------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         _at |
          1  |   .0942269   .0071278    13.22   0.000      .0802567    .1081972
          2  |   .1578074   .0105965    14.89   0.000      .1370387    .1785762
          3  |   .2523377   .0302665     8.34   0.000      .1930163     .311659
------------------------------------------------------------------------------
```

For PIR 0.30 $\rightarrow$ 0.40, logit gives about **6.36 percentage points**; probit gives **6.08**.

Similar fitted comparisons do not establish a causal effect or prove either functional form correct.

Which sentence would you put in the report?

A report states:

> *The probit coefficient is 2.968, so a 0.10 PIR increase raises denial probability by 29.68 percentage points. Bounded predictions make this effect causal.*

Rewrite it using the $0.30 \rightarrow 0.40$ comparison. Address both the **scale** and the **causal claim**.

Next class, 18 November: how coefficients are estimated, uncertainty for probability differences, and adjusted comparisons.