Functional Forms

When a Straight Line Won't Do

Jake Anderson

March 21, 2026

Outline

1. Motivation: The Return to Education
2. The Log-Linear Model
3. The Linear-Log Model
4. The Log-Log Model
5. The Quadratic Model
6. Choosing a Functional Form
7. Prediction in Log Models
8. Summary

Outline

1. Motivation: The Return to Education
2. The Log-Linear Model
3. The Linear-Log Model
4. The Log-Log Model
5. The Quadratic Model
6. Choosing a Functional Form
7. Prediction in Log Models
8. Summary

A Question About Wages

Does each additional year of schooling raise wages by…

- the same **dollar amount**? (e.g., $3,000/year)
- or the same **percentage**? (e.g., 8%/year)

A Question About Wages

Does each additional year of schooling raise wages by...

- the same **dollar amount**? (e.g., $3,000/year)
- or the same **percentage**? (e.g., 8%/year)

These are very different claims:

- Constant dollar amount $\Longrightarrow$ going from 12 to 13 years adds the same as going from 16 to 17
- Constant percentage $\Longrightarrow$ going from 16 to 17 adds more *dollars* than 12 to 13, because the base is higher

A Question About Wages

Does each additional year of schooling raise wages by...

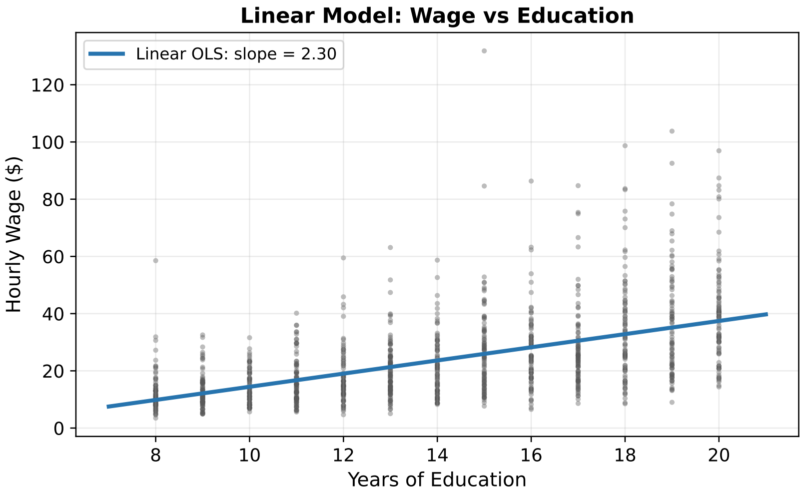
- the same **dollar amount**? (e.g., \$3,000/year)
- or the same **percentage**? (e.g., 8%/year)

These are very different claims:

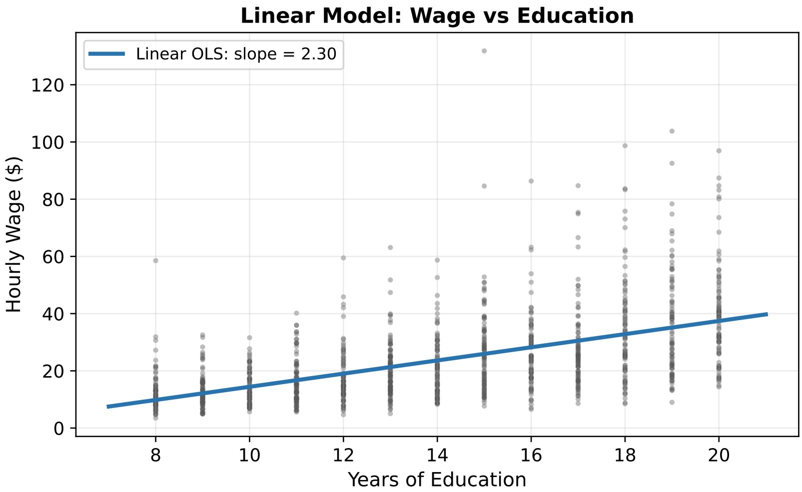
- Constant dollar amount $\implies$ going from 12 to 13 years adds the same as going from 16 to 17
- Constant percentage $\implies$ going from 16 to 17 adds more *dollars* than 12 to 13, because the base is higher

$\implies$ The choice of functional form determines the economic story your regression tells.

Wage vs. Education: The Linear Model

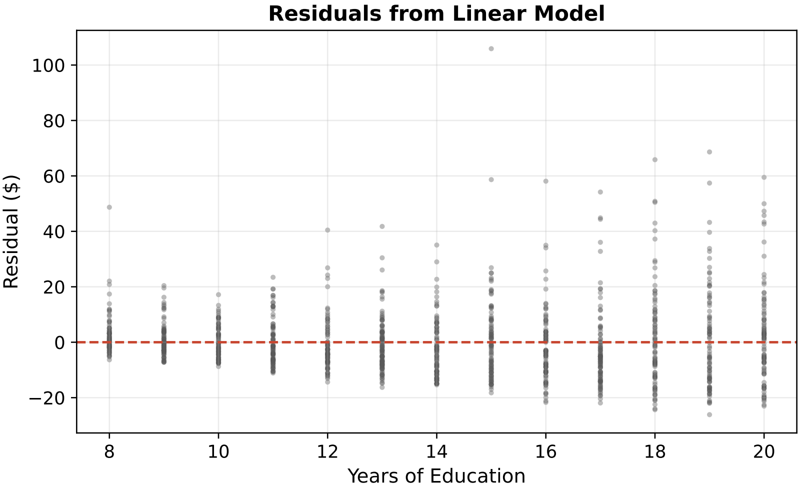


Wage vs. Education: The Linear Model

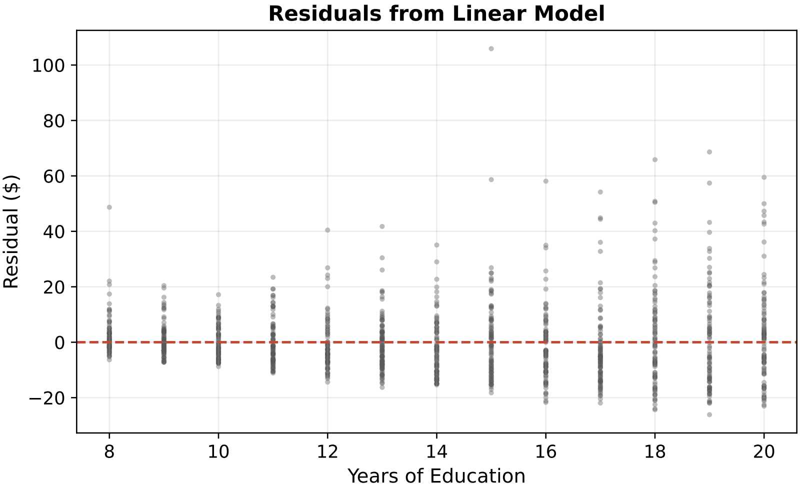


The linear model says: each year of education adds a fixed dollar amount to wages. Does that look right?

Check the Residuals



Check the Residuals

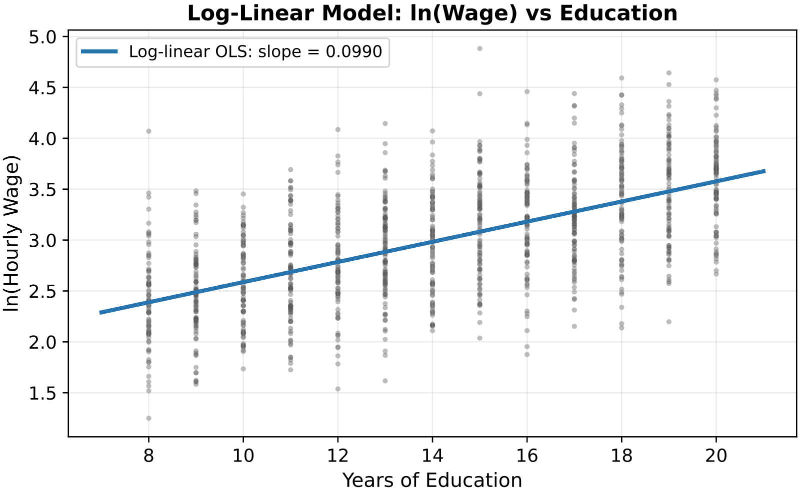


The residuals **fan out** as education increases. Two problems:

- The variance of residuals grows with x (heteroskedasticity)

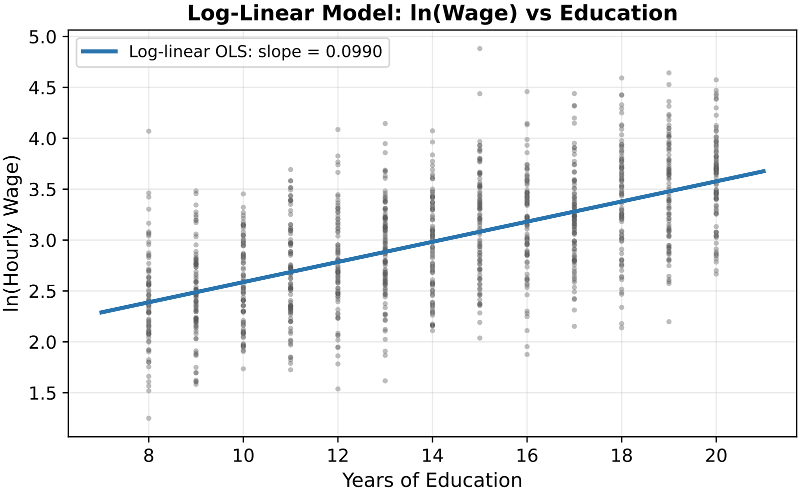
A Better Idea: Model ln(Wage)

What if we take the natural log of wages first, then regress on education?

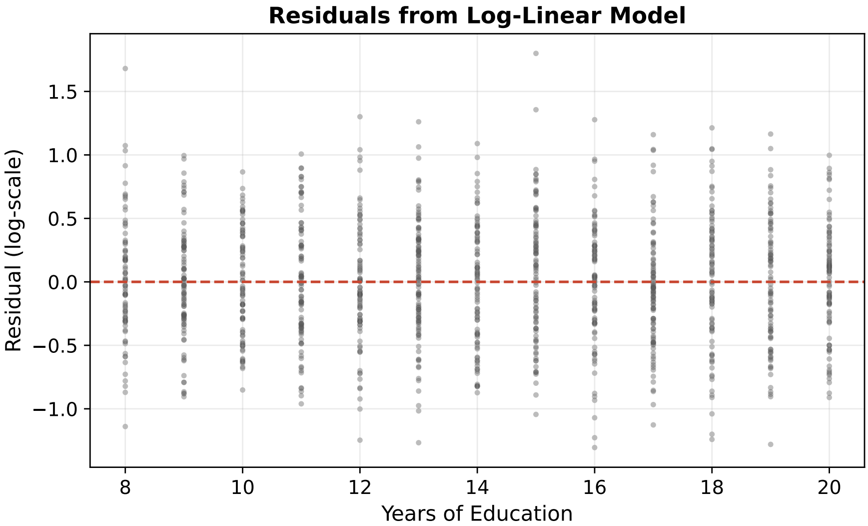


A Better Idea: Model ln(Wage)

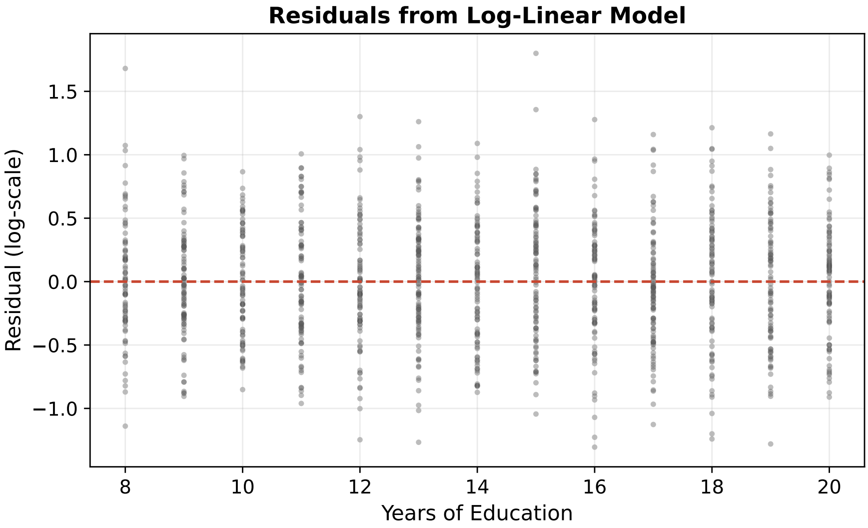
What if we take the natural log of wages first, then regress on education?



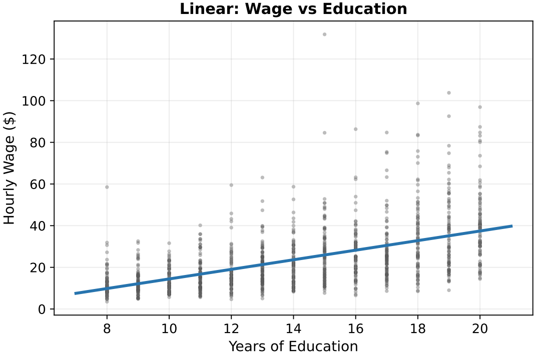
Residuals from the Log-Linear Model

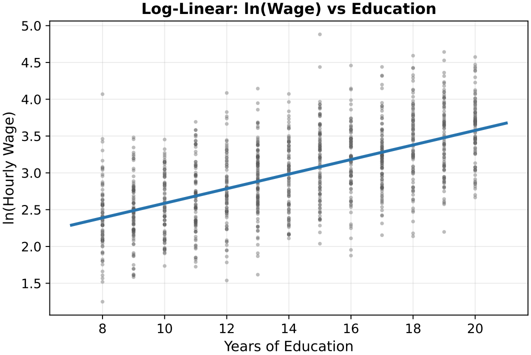


Residuals from the Log-Linear Model

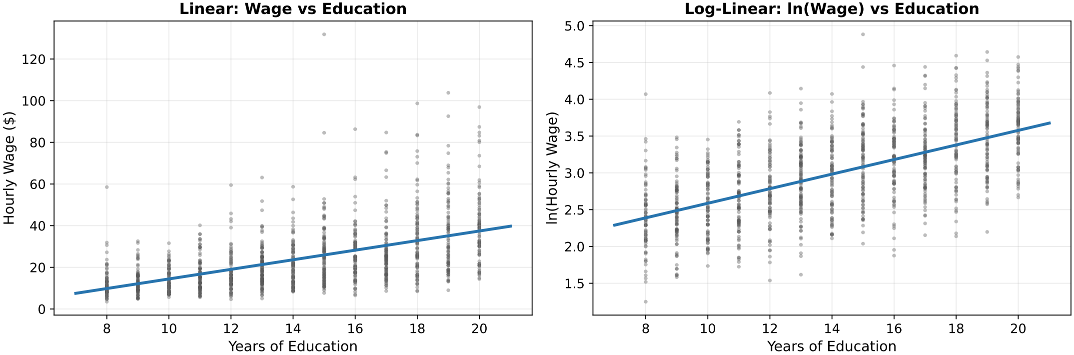


Side-by-Side Comparison





Side-by-Side Comparison



The right panel is a tighter fit. But more importantly, the log-linear model embeds a specific economic theory: the return to education is a constant **percentage**, not a constant dollar amount.

Outline

1 Motivation: The Return to Education

2 The Log-Linear Model

3 The Linear-Log Model

4 The Log-Log Model

5 The Quadratic Model

6 Choosing a Functional Form

7 Prediction in Log Models

8 Summary

From Pictures to Coefficients

We just saw that ln(wage) is approximately linear in education:

$$\ln(\widehat{\text{WAGE}}) = 1.597 + 0.099 \times \text{EDUC}$$

From Pictures to Coefficients

We just saw that $\ln(\text{wage})$ is approximately linear in education:

$$\ln(\widehat{\text{WAGE}}) = 1.597 + 0.099 \times \text{EDUC}$$

The slope is 0.099. But what does that *mean*?

- It is *not* a dollar amount (the left-hand side is a log, not dollars).
- We need to translate the log-scale coefficient into a percentage statement about wages.

From Pictures to Coefficients

We just saw that $\ln(\text{wage})$ is approximately linear in education:

$$\ln(\widehat{\text{WAGE}}) = 1.597 + 0.099 \times \text{EDUC}$$

The slope is 0.099. But what does that *mean*?

- It is *not* a dollar amount (the left-hand side is a log, not dollars).
- We need to translate the log-scale coefficient into a percentage statement about wages.

$\Longrightarrow$ The next few slides derive the interpretation rule.

Log-Linear Specification

The **log-linear** model:

$$\ln(y) = \beta_1 + \beta_2 x + e$$

Log-Linear Specification

The **log-linear** model:

$$\ln(y) = \beta_1 + \beta_2 x + e$$

Interpretation: A 1-unit increase in x changes y by approximately $100\beta_2\%$.

Log-Linear Specification

The **log-linear** model:

$$\ln(y) = \beta_1 + \beta_2 x + e$$

Interpretation: A 1-unit increase in x changes y by approximately $100\beta_2\%$.

Where does the "$100\beta_2\%$" come from?

$$\ln(y_1) - \ln(y_0) = \beta_2(x_1 - x_0) = \beta_2 \cdot \Delta x$$

$$\ln\left(\frac{y_1}{y_0}\right) = \beta_2 \cdot \Delta x$$

$$\underbrace{100 \times \ln\left(\frac{y_1}{y_0}\right)}_{\approx \%\Delta y} \approx 100\beta_2 \cdot \Delta x$$

The approximation $100 \times \ln(y_1/y_0) \approx \%\Delta y$ works well when the log change is small: roughly $|\beta_2 \cdot \Delta x| < 0.20$, i.e., percentage changes under about 20%.

$\implies$ Multiply the coefficient by 100 to get the approximate percentage effect.

Log-Linear Specification

The **log-linear** model:

$$\ln(y) = \beta_1 + \beta_2 x + e$$

Interpretation: A 1-unit increase in x changes y by approximately $100\beta_2\%$.

Where does the "$100\beta_2\%$" come from?

$$\ln(y_1) - \ln(y_0) = \beta_2(x_1 - x_0) = \beta_2 \cdot \Delta x$$

$$\underbrace{100 \times \ln\left(\frac{y_1}{y_0}\right)}_{\approx \%\Delta y} \approx 100\beta_2 \cdot \Delta x$$

The approximation $100 \times \ln(y_1/y_0) \approx \%\Delta y$ works well when the log change is small: roughly $|\beta_2 \cdot \Delta x| < 0.20$, i.e., percentage changes under about 20%.

$\implies$ Multiply the coefficient by 100 to get the approximate percentage effect.

Log-Linear Specification

The **log-linear** model:

$$\ln(y) = \beta_1 + \beta_2 x + e$$

Interpretation: A 1-unit increase in x changes y by approximately $100\beta_2\%$.

Where does the "$100\beta_2\%$" come from?

$$\ln(y_1) - \ln(y_0) = \beta_2(x_1 - x_0) = \beta_2 \cdot \Delta x$$

$$\ln\left(\frac{y_1}{y_0}\right) = \beta_2 \cdot \Delta x$$

$$\underbrace{100 \times \ln\left(\frac{y_1}{y_0}\right)}_{\approx \%\Delta y} \approx 100\beta_2 \cdot \Delta x$$

Log-Linear Specification

The **log-linear** model:

$$\ln(y) = \beta_1 + \beta_2 x + e$$

Interpretation: A 1-unit increase in x changes y by approximately $100\beta_2\%$.

Where does the "$100\beta_2\%$" come from?

$$\ln(y_1) - \ln(y_0) = \beta_2(x_1 - x_0) = \beta_2 \cdot \Delta x$$

$$\ln\left(\frac{y_1}{y_0}\right) = \beta_2 \cdot \Delta x$$

$$\underbrace{100 \times \ln\left(\frac{y_1}{y_0}\right)}_{\approx \%\Delta y} \approx 100\beta_2 \cdot \Delta x$$

The approximation $100 \times \ln(y_1/y_0) \approx \%\Delta y$ works well when the log change is small: roughly $|\beta_2 \cdot \Delta x| < 0.20$, i.e., percentage changes under about 20%.

$\Longrightarrow$ Multiply the coefficient by 100 to get the approximate percentage effect.

The Mincer Equation: Economic Foundation

Why would $\ln(\text{wage})$ be linear in education? Think of education as an investment.

If each year of schooling yields a constant rate of return r:

$$\text{WAGE} = \text{WAGE}_0 \times (1 + r)^{\text{EDUC}}$$

The Mincer Equation: Economic Foundation

Why would $\ln(\text{wage})$ be linear in education? Think of education as an investment.

If each year of schooling yields a constant rate of return r:

$$\text{WAGE} = \text{WAGE}_0 \times (1 + r)^{\text{EDUC}}$$

Taking logs of both sides:

$$\ln(\text{WAGE}) = \underbrace{\ln(\text{WAGE}_0)}_{\beta_1} + \underbrace{\ln(1 + r)}_{\beta_2} \times \text{EDUC}$$

The Mincer Equation: Economic Foundation

Why would $\ln(\text{wage})$ be linear in education? Think of education as an investment.

If each year of schooling yields a constant rate of return r:

$$\text{WAGE} = \text{WAGE}_0 \times (1+r)^{\text{EDUC}}$$

Taking logs of both sides:

$$\ln(\text{WAGE}) = \underbrace{\ln(\text{WAGE}_0)}_{\beta_1} + \underbrace{\ln(1+r)}_{\beta_2} \times \text{EDUC}$$

Since $\ln(1+r) \approx r$ for small r, the slope coefficient directly estimates the rate of return to education.

The Mincer Equation: Economic Foundation

Why would $\ln(\text{wage})$ be linear in education? Think of education as an investment.

If each year of schooling yields a constant rate of return r:

$$\text{WAGE} = \text{WAGE}_0 \times (1 + r)^{\text{EDUC}}$$

Taking logs of both sides:

$$\ln(\text{WAGE}) = \underbrace{\ln(\text{WAGE}_0)}_{\beta_1} + \underbrace{\ln(1 + r)}_{\beta_2} \times \text{EDUC}$$

Since $\ln(1 + r) \approx r$ for small r, the slope coefficient directly estimates the rate of return to education.

$\implies$ The log-linear wage equation is a compound-interest model. This is the **Mincer wage equation**, the workhorse of labor economics.

Example: The Return to Education

Using CPS wage data ($N = 1{,}200$):

$$\ln(\widehat{\text{WAGE}}) = 1.597 + 0.099 \times \text{EDUC}$$

$$\text{(se)} \quad (0.070) \quad (0.005)$$

Example: The Return to Education

Using CPS wage data ($N = 1{,}200$):

$$\ln(\widehat{\text{WAGE}}) = 1.597 + 0.099 \times \text{EDUC}$$

$$\text{(se)} \quad (0.070) \quad (0.005)$$

Interpretation: Each additional year of education increases wages by approximately $100 \times 0.099 = 9.9\%$.

Example: The Return to Education

Using CPS wage data ($N = 1{,}200$):

$$\ln(\widehat{\text{WAGE}}) = 1.597 + 0.099 \times \text{EDUC}$$

$$\text{(se)} \quad (0.070) \quad (0.005)$$

Interpretation: Each additional year of education increases wages by approximately $100 \times 0.099 = 9.9\%$.

Both the coefficient estimate and the standard errors come from applying OLS to the $\ln(\text{wage})$ regression. The Mincer model gives us the economic reason to expect a log-linear relationship; the data tell us $\hat{r} \approx 0.099$.

When to Use the Log-Linear Model

The log-linear form is natural when:

1. The dependent variable is **always positive** (wages, prices, expenditures)
2. Theory suggests **constant percentage effects** (returns to education, growth rates, depreciation)
3. Residuals from a linear model fan out (the log transform stabilizes variance)

When to Use the Log-Linear Model

The log-linear form is natural when:

1. The dependent variable is **always positive** (wages, prices, expenditures)
2. Theory suggests **constant percentage effects** (returns to education, growth rates, depreciation)
3. Residuals from a linear model fan out (the log transform stabilizes variance)

Growth models are a special case. If output grows at rate g per year:

$$Y_t = Y_0(1+g)^t \quad \Longrightarrow \quad \ln(Y_t) = \underbrace{\ln(Y_0)}_{\beta_1} + \underbrace{\ln(1+g)}_{\beta_2} \cdot t$$

When to Use the Log-Linear Model

The log-linear form is natural when:

1. The dependent variable is **always positive** (wages, prices, expenditures)
2. Theory suggests **constant percentage effects** (returns to education, growth rates, depreciation)
3. Residuals from a linear model fan out (the log transform stabilizes variance)

Growth models are a special case. If output grows at rate g per year:

$$Y_t = Y_0(1+g)^t \quad \Longrightarrow \quad \ln(Y_t) = \underbrace{\ln(Y_0)}_{\beta_1} + \underbrace{\ln(1+g)}_{\beta_2} \cdot t$$

$\Longrightarrow$ The slope in a log-linear model with time as x estimates the constant growth rate.

Outline

1. Motivation: The Return to Education
2. The Log-Linear Model
3. The Linear-Log Model
4. The Log-Log Model
5. The Quadratic Model
6. Choosing a Functional Form
7. Prediction in Log Models
8. Summary

Linear-Log Specification

The **linear-log** model:

$$y = \beta_1 + \beta_2 \ln(x) + e$$

Linear-Log Specification

The **linear-log** model:

$$y = \beta_1 + \beta_2 \ln(x) + e$$

Interpretation: A 1% increase in x changes y by approximately $\beta_2/100$ units.

Linear-Log Specification

The **linear-log** model:

$$y = \beta_1 + \beta_2 \ln(x) + e$$

Interpretation: A 1% increase in x changes y by approximately $\beta_2/100$ units.

Derivation:

$$\Delta y = \beta_2[\ln(x_1) - \ln(x_0)] = \beta_2 \cdot \ln\left(\frac{x_1}{x_0}\right)$$

$$\Delta y \approx \beta_2 \cdot \frac{\%\Delta x}{100}$$

$\implies$ Divide the coefficient by 100 to get the unit change in y from a 1% increase in x.

This model is natural when y has **diminishing returns** to x: doubling income from \$20k to \$40k has the same effect on food spending as doubling from \$40k to \$80k.

Linear-Log Specification

The **linear-log** model:

$$y = \beta_1 + \beta_2 \ln(x) + e$$

Interpretation: A 1% increase in x changes y by approximately $\beta_2/100$ units.

Derivation:

$$\Delta y = \beta_2[\ln(x_1) - \ln(x_0)] = \beta_2 \cdot \ln\left(\frac{x_1}{x_0}\right)$$

$$\Delta y \approx \beta_2 \cdot \frac{\%\Delta x}{100}$$

Linear-Log Specification

The **linear-log** model:

$$y = \beta_1 + \beta_2 \ln(x) + e$$

Interpretation: A 1% increase in x changes y by approximately $\beta_2/100$ units.

Derivation:

$$\Delta y = \beta_2[\ln(x_1) - \ln(x_0)] = \beta_2 \cdot \ln\left(\frac{x_1}{x_0}\right)$$

$$\Delta y \approx \beta_2 \cdot \frac{\%\Delta x}{100}$$

$\implies$ Divide the coefficient by 100 to get the unit change in y from a 1% increase in x.

Linear-Log Specification

The **linear-log** model:

$$y = \beta_1 + \beta_2 \ln(x) + e$$

Interpretation: A 1% increase in x changes y by approximately $\beta_2/100$ units.

Derivation:

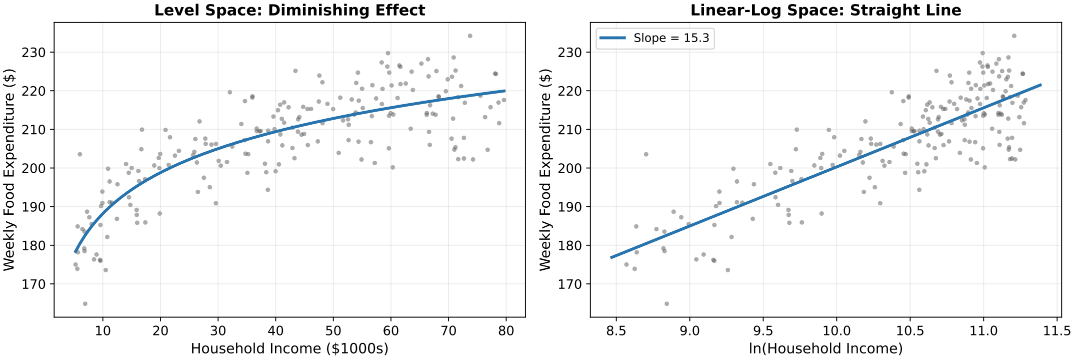
$$\Delta y = \beta_2[\ln(x_1) - \ln(x_0)] = \beta_2 \cdot \ln\left(\frac{x_1}{x_0}\right)$$

$$\Delta y \approx \beta_2 \cdot \frac{\%\Delta x}{100}$$

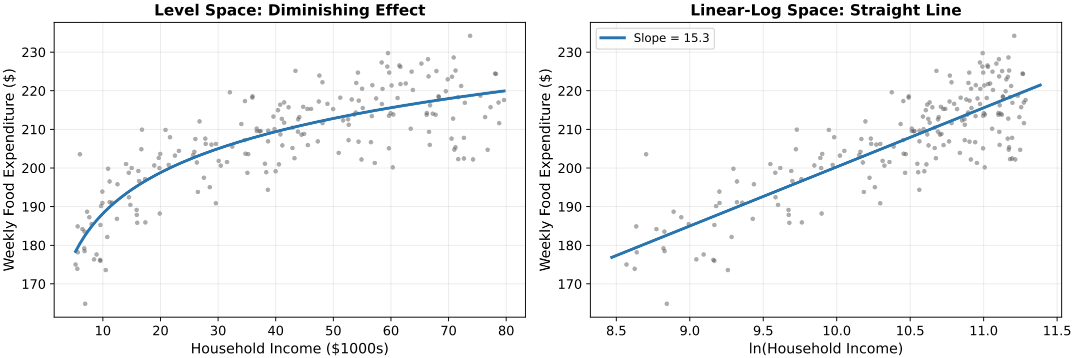
$\implies$ Divide the coefficient by 100 to get the unit change in y from a 1% increase in x.

This model is natural when y has **diminishing returns** to x: doubling income from $20k to $40k has the same effect on food spending as doubling from $40k to $80k.

Example: Food Expenditure and Income



Example: Food Expenditure and Income



Left panel: in levels, food expenditure rises steeply at low incomes then flattens. Right panel: plotting against ln(income) straightens the relationship. A 1% increase in income raises food spending by about $\hat{\beta}_2/100$ dollars.

Outline

1 Motivation: The Return to Education

2 The Log-Linear Model

3 The Linear-Log Model

4 The Log-Log Model

5 The Quadratic Model

6 Choosing a Functional Form

7 Prediction in Log Models

8 Summary

What Log-Linear Cannot Answer

Suppose we want to know: *if the price of chicken rises by 10%, how much does quantity demanded fall?*

What Log-Linear Cannot Answer

Suppose we want to know: *if the price of chicken rises by 10%, how much does quantity demanded fall?*

The log-linear model $\ln(Q) = \beta_1 + \beta_2 P$ answers a different question: what happens when price rises by \$1? The effect depends on the price level.

What Log-Linear Cannot Answer

Suppose we want to know: *if the price of chicken rises by 10%, how much does quantity demanded fall?*

The log-linear model $\ln(Q) = \beta_1 + \beta_2 P$ answers a different question: what happens when price rises by $1? The effect depends on the price level.

We need a model where both sides are in percentages, so that the answer is the same at every price level.

What Log-Linear Cannot Answer

Suppose we want to know: *if the price of chicken rises by 10%, how much does quantity demanded fall?*

The log-linear model $\ln(Q) = \beta_1 + \beta_2 P$ answers a different question: what happens when price rises by \$1? The effect depends on the price level.

We need a model where both sides are in percentages, so that the answer is the same at every price level.

$\Longrightarrow$ Log both sides: put $\ln(Q)$ on the left and $\ln(P)$ on the right.

Log-Log Specification

The **log-log** model:

$$\ln(y) = \beta_1 + \beta_2 \ln(x) + e$$

Log-Log Specification

The **log-log** model:

$$\ln(y) = \beta_1 + \beta_2 \ln(x) + e$$

Interpretation: A 1% increase in x changes y by approximately β_2%.

Log-Log Specification

The **log-log** model:

$$\ln(y) = \beta_1 + \beta_2 \ln(x) + e$$

Interpretation: A 1% increase in x changes y by approximately β_2%.

Derivation:

$$\ln(y_1) - \ln(y_0) = \beta_2[\ln(x_1) - \ln(x_0)]$$

$$\underbrace{100 \times [\ln(y_1) - \ln(y_0)]}_{\approx \%\Delta y} \approx \beta_2 \times \underbrace{100 \times [\ln(x_1) - \ln(x_0)]}_{\approx \%\Delta x}$$

Recall from microeconomics: elasticity $= (\partial y/\partial x) \cdot (x/y)$. In a linear model, elasticity $= \beta_2(x/y)$, which varies across observations. In the log-log model, elasticity is constant at β_2 everywhere.

Both $x > 0$ and $y > 0$ are required for the log transformation.

Log-Log Specification

The **log-log** model:

$$\ln(y) = \beta_1 + \beta_2 \ln(x) + e$$

Interpretation: A 1% increase in x changes y by approximately β_2%.

Derivation:

$$\ln(y_1) - \ln(y_0) = \beta_2[\ln(x_1) - \ln(x_0)]$$

$$\underbrace{100 \times [\ln(y_1) - \ln(y_0)]}_{\approx \%\Delta y} \approx \beta_2 \times \underbrace{100 \times [\ln(x_1) - \ln(x_0)]}_{\approx \%\Delta x}$$

Log-Log Specification

The **log-log** model:

$$\ln(y) = \beta_1 + \beta_2 \ln(x) + e$$

Interpretation: A 1% increase in x changes y by approximately β_2%.

Derivation:

$$\ln(y_1) - \ln(y_0) = \beta_2[\ln(x_1) - \ln(x_0)]$$

$$\underbrace{100 \times [\ln(y_1) - \ln(y_0)]}_{\approx \%\Delta y} \approx \beta_2 \times \underbrace{100 \times [\ln(x_1) - \ln(x_0)]}_{\approx \%\Delta x}$$

Recall from microeconomics: elasticity $= (\partial y/\partial x) \cdot (x/y)$. In a linear model, elasticity $= \beta_2(x/y)$, which varies across observations. In the log-log model, elasticity is constant at β_2 everywhere.

Log-Log Specification

The **log-log** model:

$$\ln(y) = \beta_1 + \beta_2 \ln(x) + e$$

Interpretation: A 1% increase in x changes y by approximately β_2%.

Derivation:

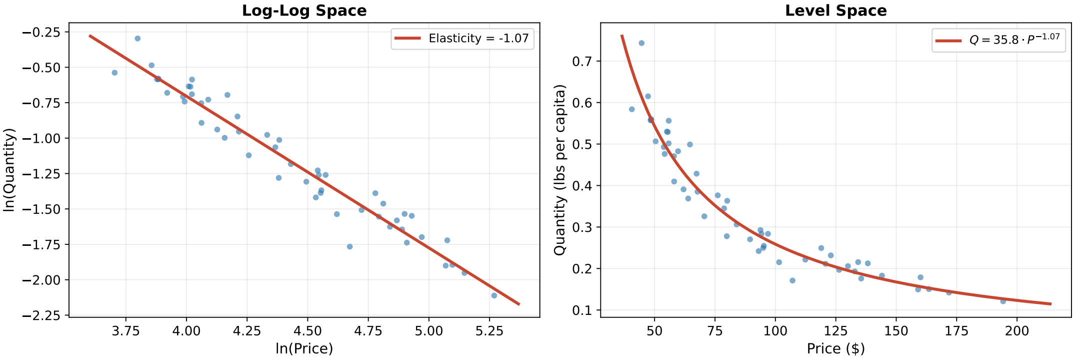
$$\ln(y_1) - \ln(y_0) = \beta_2[\ln(x_1) - \ln(x_0)]$$

$$\underbrace{100 \times [\ln(y_1) - \ln(y_0)]}_{\approx \%\Delta y} \approx \beta_2 \times \underbrace{100 \times [\ln(x_1) - \ln(x_0)]}_{\approx \%\Delta x}$$

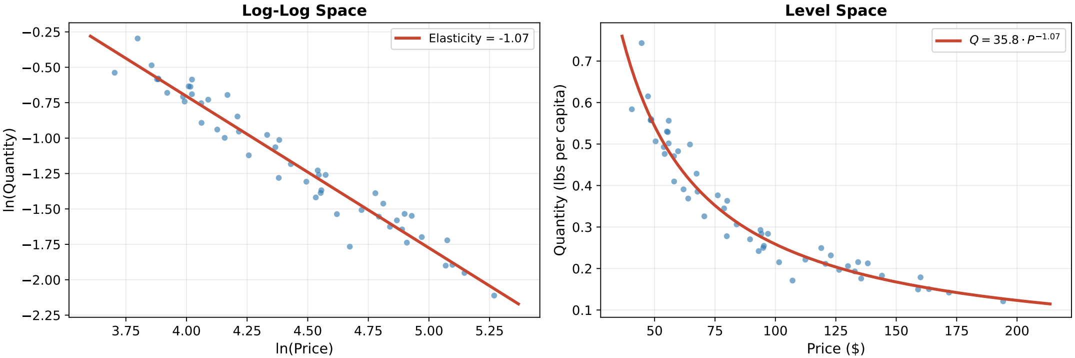
Recall from microeconomics: elasticity $= (\partial y / \partial x) \cdot (x/y)$. In a linear model, elasticity $= \beta_2(x/y)$, which varies across observations. In the log-log model, elasticity is constant at β_2 everywhere.

Both $x > 0$ and $y > 0$ are required for the log transformation.

Example: Demand Elasticity



Example: Demand Elasticity



A constant elasticity of -1.1 means: a 1% price increase reduces quantity demanded by about 1.1%. The demand is slightly elastic ($|\beta_2| > 1$).

Why Economists Love the Log-Log Model

1. **Elasticities are the natural language of economics.** Demand curves, production functions, Engel curves are all characterized by elasticities.

Why Economists Love the Log-Log Model

1. **Elasticities are the natural language of economics.** Demand curves, production functions, Engel curves are all characterized by elasticities.
2. **The elasticity is constant.** In the linear model, elasticity varies with (x, y): $\varepsilon = \beta_2(x/y)$. The log-log model fixes this at β_2 everywhere.

Why Economists Love the Log-Log Model

1. **Elasticities are the natural language of economics.** Demand curves, production functions, Engel curves are all characterized by elasticities.
2. **The elasticity is constant.** In the linear model, elasticity varies with (x, y): $\varepsilon = \beta_2(x/y)$. The log-log model fixes this at β_2 everywhere.
3. **Cobb-Douglas production functions** take this form. A is total factor productivity, K is capital, L is labor:

$$Y = AK^{\alpha}L^{\beta} \quad \Longrightarrow \quad \ln Y = \ln A + \alpha \ln K + \beta \ln L$$

Why Economists Love the Log-Log Model

1. **Elasticities are the natural language of economics.** Demand curves, production functions, Engel curves are all characterized by elasticities.
2. **The elasticity is constant.** In the linear model, elasticity varies with (x, y): $\varepsilon = \beta_2(x/y)$. The log-log model fixes this at β_2 everywhere.
3. **Cobb-Douglas production functions** take this form. A is total factor productivity, K is capital, L is labor:

$$Y = AK^{\alpha}L^{\beta} \quad \Longrightarrow \quad \ln Y = \ln A + \alpha \ln K + \beta \ln L$$

Shape of the curve in levels depends on β_2:

- $\beta_2 > 1$: increasing at an increasing rate
- $0 < \beta_2 < 1$: increasing at a decreasing rate (diminishing returns)
- $\beta_2 < 0$: inverse relationship (demand curves)

Outline

1 Motivation: The Return to Education
2 The Log-Linear Model
3 The Linear-Log Model
4 The Log-Log Model
5 The Quadratic Model
6 Choosing a Functional Form
7 Prediction in Log Models
8 Summary

Why Quadratics?

Some relationships are not just nonlinear but **non-monotonic**: the effect of x on y changes sign.

Classic example: **experience and wages**.

- Early career: each year of experience adds to wages (learning, skill accumulation)
- Late career: the gains taper off and may reverse (skill depreciation, obsolescence)

Why Quadratics?

Some relationships are not just nonlinear but **non-monotonic**: the effect of x on y changes sign.

Classic example: **experience and wages**.

- Early career: each year of experience adds to wages (learning, skill accumulation)
- Late career: the gains taper off and may reverse (skill depreciation, obsolescence)

Log models cannot capture this. They are always monotonic (always increasing or always decreasing). We need a term that allows the slope to change sign.

Why Quadratics?

Some relationships are not just nonlinear but **non-monotonic**: the effect of x on y changes sign.

Classic example: **experience and wages**.

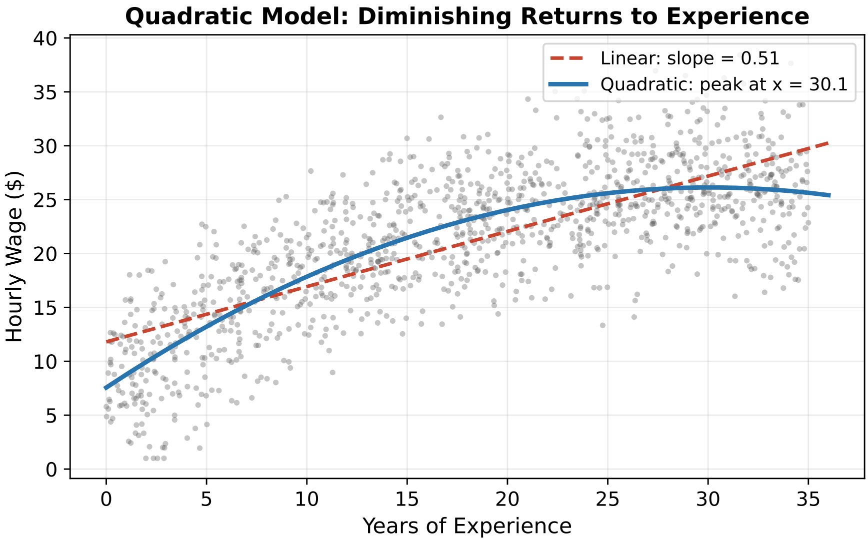
- Early career: each year of experience adds to wages (learning, skill accumulation)
- Late career: the gains taper off and may reverse (skill depreciation, obsolescence)

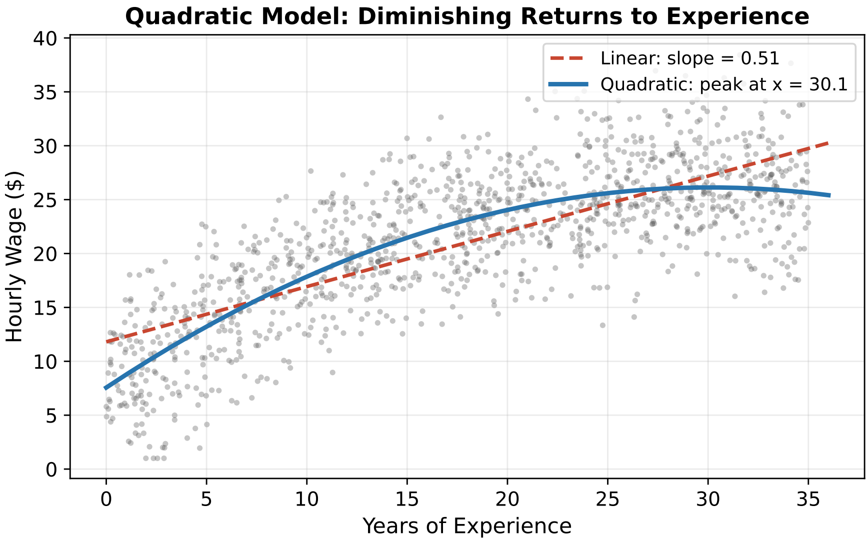
Log models cannot capture this. They are always monotonic (always increasing or always decreasing). We need a term that allows the slope to change sign.

The **quadratic model**:

$$y = \beta_1 + \beta_2 x + \beta_3 x^2 + e$$

Experience and Wages: Quadratic Fit




Quadratic Model: Diminishing Returns to Experience
Linear: slope = 0.51
Quadratic: peak at x = 30.1
Hourly Wage ($)
Years of Experience
0
5
10
15
20
25
30
35
40

The Marginal Effect Depends on x

In the quadratic model $y = \beta_1 + \beta_2 x + \beta_3 x^2$, the marginal effect is:

$$\frac{dy}{dx} = \beta_2 + 2\beta_3 x$$

The Marginal Effect Depends on x

In the quadratic model $y = \beta_1 + \beta_2 x + \beta_3 x^2$, the marginal effect is:

$$\frac{dy}{dx} = \beta_2 + 2\beta_3 x$$

This is a **linear function of** x. Unlike the linear model (where $dy/dx = \beta_2$ is constant), here the effect of one more unit of x depends on where you are.

The Marginal Effect Depends on x

In the quadratic model $y = \beta_1 + \beta_2 x + \beta_3 x^2$, the marginal effect is:

$$\frac{dy}{dx} = \beta_2 + 2\beta_3 x$$

This is a **linear function of** x. Unlike the linear model (where $dy/dx = \beta_2$ is constant), here the effect of one more unit of x depends on where you are.

Turning point: Set the marginal effect to zero:

$$\beta_2 + 2\beta_3 x^* = 0 \quad \Longrightarrow \quad x^* = -\frac{\beta_2}{2\beta_3}$$

The Marginal Effect Depends on x

In the quadratic model $y = \beta_1 + \beta_2 x + \beta_3 x^2$, the marginal effect is:

$$\frac{dy}{dx} = \beta_2 + 2\beta_3 x$$

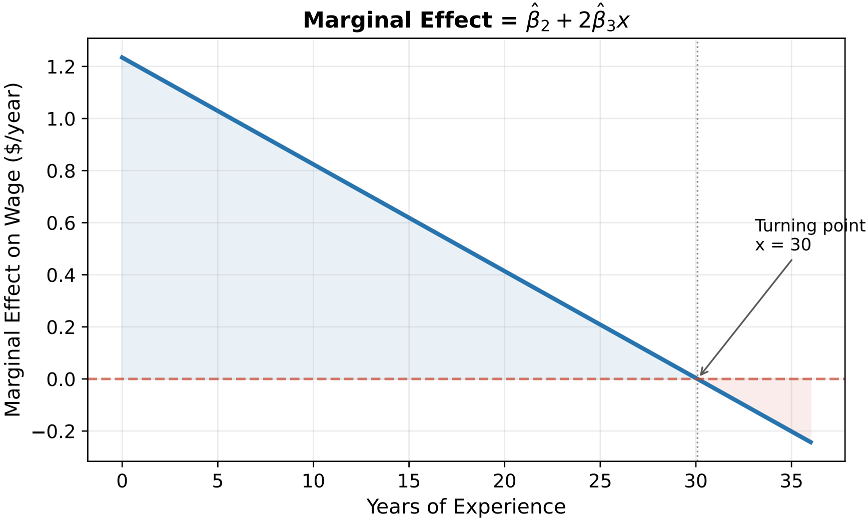
This is a **linear function of** x. Unlike the linear model (where $dy/dx = \beta_2$ is constant), here the effect of one more unit of x depends on where you are.

Turning point: Set the marginal effect to zero:

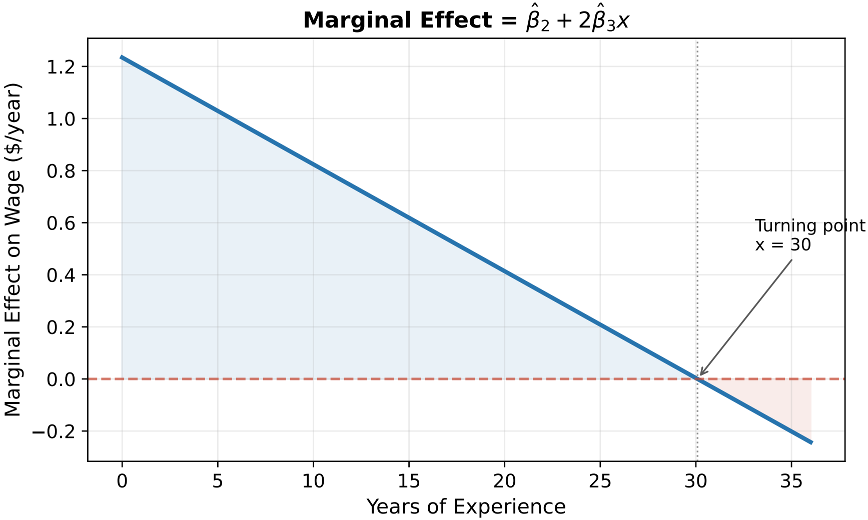
$$\beta_2 + 2\beta_3 x^* = 0 \quad \Longrightarrow \quad x^* = -\frac{\beta_2}{2\beta_3}$$

- If $\beta_3 < 0$: inverted-U (maximum at x^*). Wages peak at x^* years of experience.
- If $\beta_3 > 0$: U-shape (minimum at x^*).

Visualizing the Marginal Effect



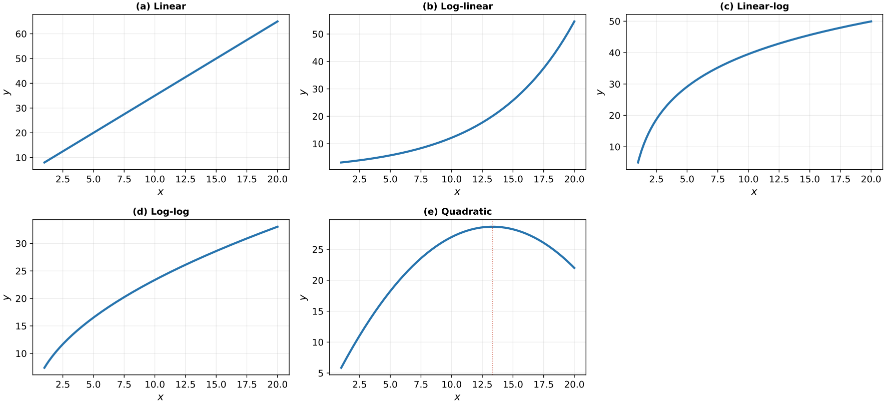
Visualizing the Marginal Effect



Outline

1. Motivation: The Return to Education
2. The Log-Linear Model
3. The Linear-Log Model
4. The Log-Log Model
5. The Quadratic Model
6. Choosing a Functional Form
7. Prediction in Log Models
8. Summary

Five Functional Forms at a Glance



Coefficient Interpretation: Summary Table

Model	**Equation**	**Slope** (dy/dx)	**Interpretation of** β_2
Linear	$y = \beta_1 + \beta_2 x$	β_2	1-unit $\Delta x \implies \beta_2$-unit Δy

Coefficient Interpretation: Summary Table

Model	**Equation**	**Slope** (dy/dx)	**Interpretation of** β_2
Linear	$y = \beta_1 + \beta_2 x$	β_2	1-unit $\Delta x \implies \beta_2$-unit Δy
Log-linear	$\ln(y) = \beta_1 + \beta_2 x$	$\beta_2 y$	1-unit $\Delta x \implies \approx 100\beta_2\%\ \Delta y$

Coefficient Interpretation: Summary Table

Model	Equation	Slope (dy/dx)	Interpretation of β_2
Linear	$y = \beta_1 + \beta_2 x$	β_2	1-unit $\Delta x \implies \beta_2$-unit Δy
Log-linear	$\ln(y) = \beta_1 + \beta_2 x$	$\beta_2 y$	1-unit $\Delta x \implies \approx 100\beta_2\%\ \Delta y$
Linear-log	$y = \beta_1 + \beta_2 \ln(x)$	β_2/x	$1\%\ \Delta x \implies \approx \beta_2/100\ \Delta y$

Coefficient Interpretation: Summary Table

Model	Equation	Slope (dy/dx)	Interpretation of β_2
Linear	$y = \beta_1 + \beta_2 x$	β_2	1-unit $\Delta x \implies \beta_2$-unit Δy
Log-linear	$\ln(y) = \beta_1 + \beta_2 x$	$\beta_2 y$	1-unit $\Delta x \implies \approx 100\beta_2\%\ \Delta y$
Linear-log	$y = \beta_1 + \beta_2 \ln(x)$	β_2 / x	$1\%\ \Delta x \implies \approx \beta_2/100\ \Delta y$
Log-log	$\ln(y) = \beta_1 + \beta_2 \ln(x)$	$\beta_2(y/x)$	$1\%\ \Delta x \implies \beta_2\%\ \Delta y$

Coefficient Interpretation: Summary Table

Model	Equation	Slope (dy/dx)	Interpretation of β_2
Linear	$y = \beta_1 + \beta_2 x$	β_2	1-unit $\Delta x \implies \beta_2$-unit Δy
Log-linear	$\ln(y) = \beta_1 + \beta_2 x$	$\beta_2 y$	1-unit $\Delta x \implies \approx 100\beta_2\%\ \Delta y$
Linear-log	$y = \beta_1 + \beta_2 \ln(x)$	β_2/x	1% $\Delta x \implies \approx \beta_2/100\ \Delta y$
Log-log	$\ln(y) = \beta_1 + \beta_2 \ln(x)$	$\beta_2(y/x)$	1% $\Delta x \implies \beta_2\%\ \Delta y$
Quadratic	$y = \beta_1 + \beta_2 x + \beta_3 x^2$	$\beta_2 + 2\beta_3 x$	Marginal effect varies with x

Coefficient Interpretation: Summary Table

Model	**Equation**	**Slope** (dy/dx)	**Interpretation of** β_2
Linear	$y = \beta_1 + \beta_2 x$	β_2	1-unit $\Delta x \implies \beta_2$-unit Δy
Log-linear	$\ln(y) = \beta_1 + \beta_2 x$	$\beta_2 y$	1-unit $\Delta x \implies \approx 100\beta_2\% \ \Delta y$
Linear-log	$y = \beta_1 + \beta_2 \ln(x)$	β_2/x	1% $\Delta x \implies \approx \beta_2/100 \ \Delta y$
Log-log	$\ln(y) = \beta_1 + \beta_2 \ln(x)$	$\beta_2(y/x)$	1% $\Delta x \implies \beta_2\% \ \Delta y$
Quadratic	$y = \beta_1 + \beta_2 x + \beta_3 x^2$	$\beta_2 + 2\beta_3 x$	Marginal effect varies with x

Note: the log approximations use $\ln(1 + r) \approx r$, which is accurate when $|\beta_2 \cdot \Delta x| < 0.20$, i.e., percentage changes under about 20%.

How to Choose: Theory + Diagnostics

Step 1: Economic theory

- Does the relationship have diminishing returns? $\implies$ log-log, linear-log, or quadratic
- Is the effect best expressed in percentages? $\implies$ log-linear
- Does the effect change sign? $\implies$ quadratic
- Is the parameter an elasticity? $\implies$ log-log
- Does doubling x have the same dollar effect regardless of level? $\implies$ linear-log

How to Choose: Theory + Diagnostics

Step 1: Economic theory

- Does the relationship have diminishing returns? $\Longrightarrow$ log-log, linear-log, or quadratic
- Is the effect best expressed in percentages? $\Longrightarrow$ log-linear
- Does the effect change sign? $\Longrightarrow$ quadratic
- Is the parameter an elasticity? $\Longrightarrow$ log-log
- Does doubling x have the same dollar effect regardless of level? $\Longrightarrow$ linear-log

Step 2: Residual diagnostics

- Plot residuals vs. x: look for patterns (U-shape $\Longrightarrow$ missed quadratic; fan $\Longrightarrow$ try log)
- Compare fits using R^2 only if the dependent variable is the same
- If dependent variables differ (y vs. $\ln y$), use the generalized $R_g^2 = [\text{corr}(y, \hat{y})]^2$

How to Choose: Theory + Diagnostics

Step 1: Economic theory

- Does the relationship have diminishing returns? $\implies$ log-log, linear-log, or quadratic
- Is the effect best expressed in percentages? $\implies$ log-linear
- Does the effect change sign? $\implies$ quadratic
- Is the parameter an elasticity? $\implies$ log-log
- Does doubling x have the same dollar effect regardless of level? $\implies$ linear-log

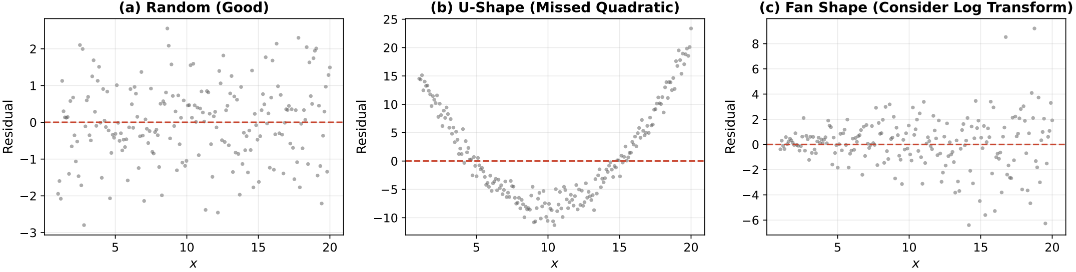
Step 2: Residual diagnostics

- Plot residuals vs. x: look for patterns (U-shape $\implies$ missed quadratic; fan $\implies$ try log)
- Compare fits using R^2 only if the dependent variable is the same
- If dependent variables differ (y vs. $\ln y$), use the generalized $R_g^2 = [\text{corr}(y, \hat{y})]^2$

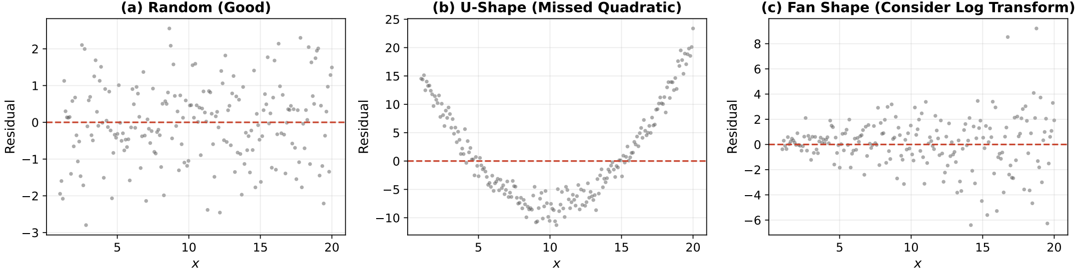
Step 3: Common sense

- Does the turning point of a quadratic fall inside or outside the data range?
- Does the model produce nonsensical predictions (e.g., negative wages)?

Residual Patterns and What They Tell You



Residual Patterns and What They Tell You



- **(a) Random:** No pattern. The functional form is appropriate.
- **(b) U-Shape:** Missed quadratic or higher-order term. Add x^2.
- **(c) Fan:** Variance grows with x. Try logging y.

Comparing R^2 Across Models

You can compare R^2 directly when the **dependent variable is the same**:

- Linear (y) vs. Quadratic (y): compare R^2 directly
- Log-linear ($\ln y$) vs. Log-log ($\ln y$): compare R^2 directly

Comparing R^2 Across Models

You can compare R^2 directly when the **dependent variable is the same**:

- Linear (y) vs. Quadratic (y): compare R^2 directly
- Log-linear ($\ln y$) vs. Log-log ($\ln y$): compare R^2 directly

You **cannot** compare R^2 when the dependent variable differs:

- Linear (y) vs. Log-linear ($\ln y$): R^2 measures different things

Comparing R^2 Across Models

You can compare R^2 directly when the **dependent variable is the same**:

- Linear (y) vs. Quadratic (y): compare R^2 directly
- Log-linear ($\ln y$) vs. Log-log ($\ln y$): compare R^2 directly

You **cannot** compare R^2 when the dependent variable differs:

- Linear (y) vs. Log-linear ($\ln y$): R^2 measures different things

$\Longrightarrow$ Use the **generalized** R^2 instead:

$$R_g^2 = [\text{corr}(y,\ \hat{y})]^2$$

This measures how well the model predicts y in its *original units*, regardless of what transformation was used internally.

Outline

1. Motivation: The Return to Education
2. The Log-Linear Model
3. The Linear-Log Model
4. The Log-Log Model
5. The Quadratic Model
6. Choosing a Functional Form
7. Prediction in Log Models
8. Summary

Predicting y When You Estimated $\ln(y)$

If you estimated $\ln(y) = \hat{\beta}_1 + \hat{\beta}_2 x$, your fitted values predict $\ln(y)$, not y itself.

Predicting y When You Estimated $\ln(y)$

If you estimated $\ln(y) = \hat{\beta}_1 + \hat{\beta}_2 x$, your fitted values predict $\ln(y)$, not y itself.

Natural predictor:

$$\hat{y}_n = \exp(\hat{\beta}_1 + \hat{\beta}_2 x)$$

Predicting y When You Estimated $\ln(y)$

If you estimated $\ln(y) = \hat{\beta}_1 + \hat{\beta}_2 x$, your fitted values predict $\ln(y)$, not y itself.

Natural predictor:

$$\hat{y}_n = \exp(\hat{\beta}_1 + \hat{\beta}_2 x)$$

Problem: $\hat{y}_n$ systematically **underpredicts** y.

Why? Because $E(e^Z) \neq e^{E(Z)}$ (Jensen's inequality). The exponential function is convex, so exponentiating the mean gives you less than the mean of the exponentiated values.

Predicting y When You Estimated $\ln(y)$

If you estimated $\ln(y) = \hat{\beta}_1 + \hat{\beta}_2 x$, your fitted values predict $\ln(y)$, not y itself.

Natural predictor:

$$\hat{y}_n = \exp(\hat{\beta}_1 + \hat{\beta}_2 x)$$

Problem: $\hat{y}_n$ systematically **underpredicts** y.

Why? Because $E(e^Z) \neq e^{E(Z)}$ (Jensen's inequality). The exponential function is convex, so exponentiating the mean gives you less than the mean of the exponentiated values.

Think of it this way: $\ln(y)$ fluctuates symmetrically around the fitted line. But when you exponentiate, large positive residuals blow up more than large negative residuals shrink. The average in levels is pulled above $e^{\text{average in logs}}$.

$\Longrightarrow$ The natural predictor misses this upward pull.

The Corrected Predictor

If $\ln(y) \sim N(\mu, \sigma^2)$, then $E(y) = e^{\mu+\sigma^2/2}$.

The Corrected Predictor

If $\ln(y) \sim N(\mu, \sigma^2)$, then $E(y) = e^{\mu+\sigma^2/2}$.

Corrected predictor:

$$\hat{y}_c = \exp\left(\hat{\beta}_1 + \hat{\beta}_2 x + \frac{\hat{\sigma}^2}{2}\right) = \hat{y}_n \cdot e^{\hat{\sigma}^2/2}$$

The Corrected Predictor

If $\ln(y) \sim N(\mu, \sigma^2)$, then $E(y) = e^{\mu + \sigma^2/2}$.

Corrected predictor:

$$\hat{y}_c = \exp\left(\hat{\beta}_1 + \hat{\beta}_2 x + \frac{\hat{\sigma}^2}{2}\right) = \hat{y}_n \cdot e^{\hat{\sigma}^2/2}$$

The correction factor $e^{\hat{\sigma}^2/2}$ is always ≥ 1. It shifts the prediction upward by an amount that depends on the residual variance $\hat{\sigma}^2$.

The Corrected Predictor

If $\ln(y) \sim N(\mu, \sigma^2)$, then $E(y) = e^{\mu + \sigma^2/2}$.

Corrected predictor:

$$\hat{y}_c = \exp\left(\hat{\beta}_1 + \hat{\beta}_2 x + \frac{\hat{\sigma}^2}{2}\right) = \hat{y}_n \cdot e^{\hat{\sigma}^2/2}$$

The correction factor $e^{\hat{\sigma}^2/2}$ is always ≥ 1. It shifts the prediction upward by an amount that depends on the residual variance $\hat{\sigma}^2$.

- $\hat{\sigma}^2$ is the variance of the OLS residuals from the $\ln(y)$ regression
- Larger residual variance $\implies$ bigger correction needed
- In large samples ($N > 30$), prefer the corrected predictor

Outline

1. Motivation: The Return to Education
2. The Log-Linear Model
3. The Linear-Log Model
4. The Log-Log Model
5. The Quadratic Model
6. Choosing a Functional Form
7. Prediction in Log Models
8. Summary

What We Covered

1. **Linear models can be wrong** even when OLS “works.” Residual plots reveal misspecification.

What We Covered

1. **Linear models can be wrong** even when OLS "works." Residual plots reveal misspecification.
2. **Log-linear** ($\ln y = \beta_1 + \beta_2 x$): each unit of x changes y by $\approx 100\beta_2\%$. Application: returns to education.

What We Covered

1. **Linear models can be wrong** even when OLS “works.” Residual plots reveal misspecification.
2. **Log-linear** ($\ln y = \beta_1 + \beta_2 x$): each unit of x changes y by $\approx 100\beta_2\%$. Application: returns to education.
3. **Linear-log** ($y = \beta_1 + \beta_2 \ln x$): a 1% increase in x changes y by $\beta_2/100$ units. Application: food expenditure vs. income.

What We Covered

1. **Linear models can be wrong** even when OLS "works." Residual plots reveal misspecification.
2. **Log-linear** ($\ln y = \beta_1 + \beta_2 x$): each unit of x changes y by $\approx 100\beta_2\%$. Application: returns to education.
3. **Linear-log** ($y = \beta_1 + \beta_2 \ln x$): a 1% increase in x changes y by $\beta_2/100$ units. Application: food expenditure vs. income.
4. **Log-log** ($\ln y = \beta_1 + \beta_2 \ln x$): β_2 is the constant elasticity. Application: demand curves, production functions.

What We Covered

1. **Linear models can be wrong** even when OLS "works." Residual plots reveal misspecification.
2. **Log-linear** ($\ln y = \beta_1 + \beta_2 x$): each unit of x changes y by $\approx 100\beta_2\%$. Application: returns to education.
3. **Linear-log** ($y = \beta_1 + \beta_2 \ln x$): a 1% increase in x changes y by $\beta_2/100$ units. Application: food expenditure vs. income.
4. **Log-log** ($\ln y = \beta_1 + \beta_2 \ln x$): β_2 is the constant elasticity. Application: demand curves, production functions.
5. **Quadratic** ($y = \beta_1 + \beta_2 x + \beta_3 x^2$): marginal effect $= \beta_2 + 2\beta_3 x$. Turning point at $x^* = -\beta_2/(2\beta_3)$. Application: experience-wage profiles.

What We Covered

1. **Linear models can be wrong** even when OLS "works." Residual plots reveal misspecification.
2. **Log-linear** ($\ln y = \beta_1 + \beta_2 x$): each unit of x changes y by $\approx 100\beta_2\%$. Application: returns to education.
3. **Linear-log** ($y = \beta_1 + \beta_2 \ln x$): a 1% increase in x changes y by $\beta_2/100$ units. Application: food expenditure vs. income.
4. **Log-log** ($\ln y = \beta_1 + \beta_2 \ln x$): β_2 is the constant elasticity. Application: demand curves, production functions.
5. **Quadratic** ($y = \beta_1 + \beta_2 x + \beta_3 x^2$): marginal effect $= \beta_2 + 2\beta_3 x$. Turning point at $x^* = -\beta_2/(2\beta_3)$. Application: experience-wage profiles.
6. **Choose functional form** based on economic theory, residual diagnostics, and interpretability.

What We Covered

1. **Linear models can be wrong** even when OLS "works." Residual plots reveal misspecification.
2. **Log-linear** ($\ln y = \beta_1 + \beta_2 x$): each unit of x changes y by $\approx 100\beta_2\%$. Application: returns to education.
3. **Linear-log** ($y = \beta_1 + \beta_2 \ln x$): a 1% increase in x changes y by $\beta_2/100$ units. Application: food expenditure vs. income.
4. **Log-log** ($\ln y = \beta_1 + \beta_2 \ln x$): β_2 is the constant elasticity. Application: demand curves, production functions.
5. **Quadratic** ($y = \beta_1 + \beta_2 x + \beta_3 x^2$): marginal effect $= \beta_2 + 2\beta_3 x$. Turning point at $x^* = -\beta_2/(2\beta_3)$. Application: experience-wage profiles.
6. **Choose functional form** based on economic theory, residual diagnostics, and interpretability.
7. **Prediction in log models:** use the corrected predictor $\hat{y}_c = \hat{y}_n \cdot e^{\hat{\sigma}^2/2}$ to avoid systematic underprediction.

Looking Ahead

- **Multiple regression** (next topic) will let us combine functional forms: $\ln(\text{WAGE}) = \beta_1 + \beta_2\text{EDUC} + \beta_3\text{EXPER} + \beta_4\text{EXPER}^2$

Looking Ahead

- **Multiple regression** (next topic) will let us combine functional forms:
  $\ln(\text{WAGE}) = \beta_1 + \beta_2\text{EDUC} + \beta_3\text{EXPER} + \beta_4\text{EXPER}^2$
- The Mincer equation is log-linear in education and quadratic in experience simultaneously

Looking Ahead

- **Multiple regression** (next topic) will let us combine functional forms: $\ln(\text{WAGE}) = \beta_1 + \beta_2\text{EDUC} + \beta_3\text{EXPER} + \beta_4\text{EXPER}^2$
- The Mincer equation is log-linear in education and quadratic in experience simultaneously
- Formal tests for choosing between functional forms (e.g., RESET test) come in later chapters

Thank you!

jakeanderson@g.ucla.edu