Instrumental Variables

Bas Machielsen

2025-06-11

The Goal of Causal Inference

Our fundamental goal is often to estimate the causal effect of a variable X on an outcome Y.

Consider the simple linear model:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

We want our estimate of β_1, which we call $\hat{\beta}_1$, to be an unbiased and consistent estimate of the true causal effect.

For Ordinary Least Squares (OLS) to be unbiased, we need the regressor X to be uncorrelated with the error term ϵ.

$$\mathbb{E}[X_i \epsilon_i] = 0 \quad \text{or} \quad \mathrm{Cov}(X_i, \epsilon_i) = 0$$

This is the assumption of **exogeneity**.

The Problem: Endogeneity

Endogeneity occurs when the exogeneity assumption is violated:

$$\text{Cov}(X_i, \epsilon_i) \neq 0$$

When X is endogenous, OLS is biased and inconsistent. Our $\hat{\beta}_1$ does not converge to the true β_1, even with infinite data.

Why does this happen? The error term ϵ contains all other factors that determine Y but are not included in the model. If any of these omitted factors are also correlated with X, we have a problem.

Sources of Endogeneity

- **Omitted Variable Bias (OVB)**
  - A variable W affects Y and is also correlated with X, but W is not included in the model.
  - *Example*: Effect of education (X) on wages (Y). Omitted "ability" (W) affects wages and is correlated with education.
- **Simultaneity / Reverse Causality**
  - X causes Y, but Y also causes X.
  - *Example*:Effect of police presence (X) on crime rates (Y). More police may reduce crime, but higher crime rates lead to more police being hired.
- **Measurement Error**
  - The variable X is measured with error.
  - *Example*: Effect of household income (X) on consumption (Y). Income reported in surveys (X) is often an imperfect measure of true income.

Illustrating Omitted Variable Bias

Let the true model be:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_i + u_i$$

where W_i is an unobserved variable (e.g., ability).

But we estimate the simple model:

$$Y_i = \alpha_0 + \alpha_1 X_i + \epsilon_i \quad (\text{where } \epsilon_i = \beta_2 W_i + u_i)$$

The OLS estimate for α_1 will be:

$$\text{plim}\,\hat{\alpha}_1 = \beta_1 + \beta_2 \cdot \frac{\text{Cov}(X_i, W_i)}{\text{Var}(X_i)}$$

The bias is $\beta_2 \cdot \frac{\text{Cov}(X_i, W_i)}{\text{Var}(X_i)}$. It is non-zero if W affects Y ($\beta_2 \neq 0$) and is correlated with X ($\text{Cov}(X, W) \neq 0$).

Intuition for Instrumental Variables

If X is endogenous, we can't use its correlation with Y to identify β_1.

- Part of the correlation between X and Y is the causal effect we want (β_1).
- Another part is the "bad" correlation due to endogeneity (e.g., OVB).

The IV idea: Find a third variable Z, the **instrument**, that can isolate the "good" part of the variation in X.

An instrument is a variable that is correlated with the endogenous regressor X, but is *not* correlated with the error term ϵ.

- It creates "as-if random" variation in X.
- It only affects Y *through* its effect on X.

Two Core IV Assumptions

For a variable Z to be a valid instrument for X in the model $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$, it must satisfy two conditions:

1. Relevance Condition (First Stage)

- The instrument Z must be correlated with the endogenous variable X: $\text{Cov}(Z_i, X_i) \neq 0$
- This is testable from the data. We can regress X on Z.

2. Exclusion Restriction (Exogeneity)

- The instrument Z must be uncorrelated with the error term ϵ: $\text{Cov}(Z_i, \epsilon_i) = 0$
- This means Z affects Y **only** through its effect on X.
- This is a theoretical assumption and is **not testable**. It requires deep institutional knowledge and a strong argument.

Visualize IV with a DAG

The Endogeneity Problem (OVB): U is an unobserved confounder.

The path $X \leftarrow U \rightarrow Y$ creates a spurious correlation between X and Y. OLS is biased.

IV Solution:

- **Relevance**: The arrow $Z \rightarrow X$.
- **Exclusion Restriction**: The *absence* of a direct arrow from Z to Y or from Z to U.

Potential Outcomes (Reminder)

Let's consider a binary treatment $X_i \in \{0, 1\}$.

- $Y_i(1)$: The potential outcome for individual i if they receive the treatment ($X_i = 1$).
- $Y_i(0)$: The potential outcome for individual i if they do not receive the treatment ($X_i = 0$).

The individual causal effect is $\tau_i = Y_i(1) - Y_i(0)$.

For each individual i, we can only observe one of the two potential outcomes:

$$Y_i^{\text{obs}} = X_i Y_i(1) + (1 - X_i) Y_i(0)$$

We never observe $Y_i(1)$ and $Y_i(0)$ for the same person.

IV in a Potential Outcomes Framework

Now let's introduce a binary instrument $Z_i \in \{0, 1\}$.

- $Z_i = 1$: Individual i is "encouraged" to get treatment.
- $Z_i = 0$: Individual i is "not encouraged".

We need to define potential outcomes for the *treatment* itself, based on the instrument:

- $X_i(1)$: The treatment status of individual i if they are encouraged ($Z_i = 1$).
- $X_i(0)$: The treatment status of individual i if they are not encouraged ($Z_i = 0$).

And we also have potential outcomes for Y: $Y_i(x)$.

Compliers, Always-Takers, Never-Takers

Based on how an individual's treatment status X_i responds to the instrument Z_i, we can classify the population into four groups.

Type	**If** $Z_i = 0$ $X_i(0)$	**If** $Z_i = 1$ $X_i(1)$
Complier *(Does what they are encouraged to do)*	0	1
Always-Taker *(Takes treatment regardless of encouragement)*	1	1
Never-Taker *(Never takes treatment regardless)*	0	0
Defier *(Does the opposite of encouragement)*	1	0

IV Assumptions in Potential Outcomes

Like potential outcomes, we never know for sure which group an individual belongs to. The two core assumptions have precise meanings in this framework, plus we need two more.

- **Independence of the Instrument**: $Z_i \perp\!\!\!\perp \{Y_i(1), Y_i(0), X_i(1), X_i(0)\}$
  - The instrument is "as-if" randomly assigned. It's independent of all potential outcomes and potential treatment decisions.
- **Exclusion Restriction**: $Y_i(x, z) = Y_i(x) \quad \text{for all } x, z$
  - The instrument Z does not have a direct effect on the outcome Y. It only works through X.
- **Relevance (First Stage)**: $\mathbb{E}[X_i|Z_i = 1] - \mathbb{E}[X_i|Z_i = 0] \neq 0$
  - The instrument must affect the treatment status for some individuals. This means there must be some Compliers.
- **Monotonicity**: $X_i(1) \geq X_i(0) \quad \text{for all } i$
  - The instrument encourages (or discourages) everyone weakly in the same direction. This assumption rules out the existence of **Defiers**.

IV Estimator

With these four assumptions, IV does **not** estimate the Average Treatment Effect ($ATE = \mathbb{E}[Y_i(1) - Y_i(0)]$) for the whole population.

Instead, IV estimates the **Local Average Treatment Effect (LATE)** (Imbens & Angrist, 1994).

$$\beta_{IV} = \mathbb{E}[Y_i(1) - Y_i(0) \mid X_i(1) > X_i(0)]$$

This is the average treatment effect **only for the subpopulation of Compliers**.

This is a crucial insight: The causal effect we get from IV is specific to the group of people who are induced into treatment by the instrument. It may not generalize to Always-Takers or Never-Takers.

The Wald Estimator

The simplest IV estimator is the Wald estimator, used when both the instrument Z and the treatment X are binary.

$$\hat{\beta}_{\text{Wald}} = \frac{\mathbb{E}[Y|Z=1] - \mathbb{E}[Y|Z=0]}{\mathbb{E}[X|Z=1] - \mathbb{E}[X|Z=0]}$$

Interpretation:

- **Numerator:** The "Reduced Form" effect. How much does the outcome change when the instrument is switched on?
  - Called the Intent-to-Treat (ITT) effect
- **Denominator:** The "First Stage" effect. How much does the treatment take-up change when the instrument is switched on?
  - Called the Share of Compliers

The Wald estimator is simply the sample-analogue of this formula.

LATE Theorem

Why does the Wald formula gives us the LATE?

The numerator is $\mathbb{E}[Y|Z=1] - \mathbb{E}[Y|Z=0]$

This difference in outcomes is only driven by the Compliers switching from $X = 0$ to $X = 1$. For everyone else (Always/Never-Takers), their treatment status doesn't change. So the numerator can be rewritten as
Numerator $= \Pr(\text{Complier}) \cdot \mathbb{E}[Y(1) - Y(0) \mid \text{Complier}] = \Pr(\text{Complier}) \cdot LATE$.

The denominator is $\mathbb{E}[X|Z=1] - \mathbb{E}[X|Z=0] = \Pr(\text{Complier})$

Putting it together:

$$\hat{\beta}_{\text{Wald}} = \frac{\Pr(\text{Complier}) \cdot \text{LATE}}{\Pr(\text{Complier})} = \text{LATE}$$

Proof of LATE Theorem

Let's prove that the Wald estimator identifies the LATE under the four key IV assumptions (Independence, Exclusion, Relevance, Monotonicity).

Wald $= \frac{\mathbb{E}[Y|Z=1]-\mathbb{E}[Y|Z=0]}{\mathbb{E}[X|Z=1]-\mathbb{E}[X|Z=0]}$

We will show that $\mathbb{E}[X|Z=1] - \mathbb{E}[X|Z=0] = \Pr(\text{Complier})$.

$$
\begin{aligned}
\mathbb{E}[X|Z=z] &= \mathbb{E}[X_i(z)|Z=z] && \text{(Def. of observed X)} \\
&= \mathbb{E}[X_i(z)] && \text{(by Independence assumption)}
\end{aligned}
$$

Proof of LATE Theorem (Cont.)

So the denominator is $\mathbb{E}[X_i(1)] - \mathbb{E}[X_i(0)]$. Let's expand this using the Law of Total Expectation over the population types (C=Complier, A=Always-Taker, N=Never-Taker). We assume Monotonicity, so there are no Defiers.

$$\begin{aligned}
\mathbb{E}[X_i(1)] &= \mathbb{E}[X_i(1)|C]\Pr(C) + \mathbb{E}[X_i(1)|A]\Pr(A) + \mathbb{E}[X_i(1)|N]\Pr(N) \\
&= (1)\Pr(C) + (1)\Pr(A) + (0)\Pr(N) = \Pr(C) + \Pr(A) \\
\mathbb{E}[X_i(0)] &= \mathbb{E}[X_i(0)|C]\Pr(C) + \mathbb{E}[X_i(0)|A]\Pr(A) + \mathbb{E}[X_i(0)|N]\Pr(N) \\
&= (0)\Pr(C) + (1)\Pr(A) + (0)\Pr(N) = \Pr(A)
\end{aligned}$$

Therefore, the denominator is:

$$(\Pr(C) + \Pr(A)) - \Pr(A) = \Pr(C)$$

Proof of LATE Theorem (Cont.)

The Numerator: Intent-to-Treat for Compliers

The numerator is $\mathbb{E}[Y|Z=1] - \mathbb{E}[Y|Z=0]$. Again, we use Independence to state that:

$$\mathbb{E}[Y|Z=z] = \mathbb{E}[Y_i(X_i(z))|Z=z] = \mathbb{E}[Y_i(X_i(z))]$$

Dissecting these two terms:

$$\begin{aligned}
\mathbb{E}[Y_i(X_i(1))] &= \mathbb{E}[Y_i(X_i(1))|C]\Pr(C) + \mathbb{E}[Y_i(X_i(1))|A]\Pr(A) + \mathbb{E}[Y_i(X_i(1))|N]\Pr(N) \\
&= \mathbb{E}[Y_i(1)|C]\Pr(C) + \mathbb{E}[Y_i(1)|A]\Pr(A) + \mathbb{E}[Y_i(0)|N]\Pr(N) \\
\mathbb{E}[Y_i(X_i(0))] &= \mathbb{E}[Y_i(X_i(0))|C]\Pr(C) + \mathbb{E}[Y_i(X_i(0))|A]\Pr(A) + \mathbb{E}[Y_i(X_i(0))|N]\Pr(N) \\
&= \mathbb{E}[Y_i(0)|C]\Pr(C) + \mathbb{E}[Y_i(1)|A]\Pr(A) + \mathbb{E}[Y_i(0)|N]\Pr(N)
\end{aligned}$$

Proof of LATE Theorem (Cont.)

Now, we take the difference. Notice the terms for Always-Takers and Never-Takers are identical and will cancel out:

$$\begin{aligned}
\text{Numerator} &= (\mathbb{E}[Y_i(1)|C]\Pr(C) + \dots) - (\mathbb{E}[Y_i(0)|C]\Pr(C) + \dots) \\
&= \mathbb{E}[Y_i(1)|C]\Pr(C) - \mathbb{E}[Y_i(0)|C]\Pr(C) \\
&= \Pr(C) \cdot (\mathbb{E}[Y_i(1)|C] - \mathbb{E}[Y_i(0)|C]) \\
&= \Pr(\text{Complier}) \cdot \mathbb{E}[Y_i(1) - Y_i(0)|\text{Complier}] \\
&= \Pr(\text{Complier}) \cdot \text{LATE}
\end{aligned}$$

(The Exclusion Restriction is implicitly used because Y only depends on X, not Z).

Putting It All Together

We have shown:

- **Numerator** $=$ Pr(Complier) $\times$ LATE
- **Denominator** $=$ Pr(Complier)

Therefore, the Wald estimator is:

$$\hat{\beta}_{\text{Wald}} = \frac{\mathbb{E}[Y|Z=1] - \mathbb{E}[Y|Z=0]}{\mathbb{E}[X|Z=1] - \mathbb{E}[X|Z=0]} = \frac{\text{Pr(Complier)} \times \text{LATE}}{\text{Pr(Complier)}} = \text{LATE}$$

This is a powerful result. It shows that the simple ratio of the reduced form effect to the first-stage effect precisely isolates the average causal effect for the specific subpopulation whose treatment status is actually manipulated by the instrument.

General IV Estimator

Let's return to our linear model: $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ where $\text{Cov}(X_i, \epsilon_i) \neq 0$. We have an instrument Z such that $\text{Cov}(Z_i, X_i) \neq 0$ and $\text{Cov}(Z_i, \epsilon_i) = 0$.

The IV estimator for β_1 is given by:

$\hat{\beta}_1^{\text{IV}} = \frac{\text{Cov}(Z,Y)}{\text{Cov}(Z,X)}$

Notice the similarity to the Wald estimator. The numerator is the covariance of the instrument and outcome (reduced form), and the denominator is the covariance of the instrument and the endogenous variable (first stage).

Analogy to OLS Estimator

This formula is very instructive. Let's compare it to the OLS estimator in a simple regression.

OLS Estimator:$\hat{\beta}_1^{\text{OLS}} = \frac{\text{Cov}(X,Y)}{\text{Cov}(X,X)} = \frac{\text{Cov}(X,Y)}{\text{Var}(X)}$ OLS uses the full covariance of X and Y. If X is endogenous, this covariance is "contaminated."

IV Estimator: $\hat{\beta}_1^{\text{IV}} = \frac{\text{Cov}(Z,Y)}{\text{Cov}(Z,X)}$ IV replaces the "bad" variation in X with the "good" variation in Z.

- It uses only the part of the variation in X that is induced by the exogenous instrument Z.
- It then scales the relationship between Z and Y by the relationship between Z and X.

General Case: Two-Stage Least Squares (2SLS)

What if we have multiple instruments or other exogenous control variables (W)? We use a procedure called Two-Stage Least Squares (2SLS).

Let the structural model be:

$$Y_i = \beta_0 + \beta_1 X_i + \gamma' W_i + \epsilon_i$$

(X_i is endogenous, W_i are exogenous controls).

Let the instruments be $Z_1, Z_2, ..., Z_k$.

The 2SLS procedure works in two steps. . .

2SLS: The First Stage

First Stage Regression: Regress the endogenous variable X on all the instruments Z and all other exogenous controls W.

$$X_i = \pi_0 + \pi_1 Z_{1i} + ... + \pi_k Z_{ki} + \delta' W_i + \nu_i$$

- From this regression, we obtain the **predicted values** for X, which we call $\hat{X}_i$.

$$\hat{X}_i = \hat{\pi}_0 + \hat{\pi}_1 Z_{1i} + ... + \hat{\pi}_k Z_{ki} + \hat{\delta}' W_i$$

- This $\hat{X}_i$ is the part of the variation in X that is explained by our exogenous variables.
- By construction, $\hat{X}_i$ is uncorrelated with the structural error ϵ_i (as long as our instruments are valid!).

2SLS: The Second Stage

Second Stage Regression: Regress the outcome variable Y on the **predicted** endogenous variable $\hat{X}$ and the other exogenous controls W.

$$Y_i = \beta_0 + \beta_1 \hat{X}_i + \gamma' W_i + \text{error}$$

- The coefficient $\hat{\beta}_1$ from this second stage regression is our 2SLS estimate.
- Since we used $\hat{X}_i$ instead of X_i, we have purged the endogeneity, and our estimate for β_1 is now consistent.

Note on standard errors: Do **not** perform these two steps manually in your statistical software. The standard errors from the second stage will be incorrect.

Finding Good Instruments

The credibility of any IV analysis rests entirely on the quality of the instrument.

- A good instrument must be both **relevant** and **valid** (exogenous).
- The exclusion restriction is a very strong assumption. You must provide a convincing story for why your instrument Z could not possibly affect Y except through its effect on X.

Often, instruments come from:

- Natural or quasi-random experiments (e.g., policy changes, lotteries).
- Institutional details or rules.
- Geographic or historical quirks.

The Problem of Weak Instruments

What happens if the **Relevance** condition is only barely met? i.e., $\text{Cov}(Z, X)$ is very close to zero.

If an instrument is weak, even a tiny violation of the exclusion restriction (a tiny $\text{Cov}(Z, \epsilon)$) can lead to a very large bias in the IV estimate.

$$\text{Bias}(\hat{\beta}_{IV}) \approx \frac{\text{Cov}(Z, \epsilon)}{\text{Cov}(Z, X)}$$

A small denominator leads to a large bias! Furthermore, the variance of the IV estimator will be very large.

Testing for Weak Instruments

We can and should test for instrument relevance.

- This is a test on the **first stage** regression:

$$X_i = \pi_0 + \pi_1 Z_{1i} + ... + \pi_k Z_{ki} + (\text{controls}) + \nu_i$$

- We perform an F-test on the joint significance of all instruments:

$$H_0 : \pi_1 = \pi_2 = ... = \pi_k = 0$$

A first-stage **F-statistic greater than 10** is often used as a benchmark to indicate that instruments are not weak (Stock, Wright, & Yogo, 2002). $F < 10$ is a serious red flag. You should always report the first-stage F-statistic in any IV analysis.

Testing the Exclusion Restriction

The exclusion restriction, $\text{Cov}(Z, \epsilon) = 0$, is the bedrock of IV and is **fundamentally untestable**.

- Its validity is based on economic theory, institutional knowledge, and careful reasoning.
- However, if you have **more instruments than endogenous variables** (the "overidentified" case), you can perform a partial test (Sargan-Hansen test).
  - *Intuition*: If all instruments are valid, they should all point to the same estimate of β_1. The test checks if the different instruments produce statistically different estimates.
  - A rejection of the null hypothesis suggests that at least one of your instruments is not valid (i.e., it is correlated with the error term).
  - This test requires you to believe that at least one instrument is valid to test the validity of the "extra" ones.

Classic IV Example: Angrist & Krueger (1991)

Research Question: What is the causal effect of an additional year of schooling on wages?

- Outcome (Y): Log weekly wages.
- Endogenous Variable (X): Years of schooling.
- Endogeneity Problem: Ability is an omitted variable. More able individuals tend to get more schooling and earn higher wages, biasing the OLS estimate upwards.
- The Instrument (Z): Quarter of Birth.

Angrist & Krueger (1991): The Instrument

- Why is Quarter of Birth a valid instrument?
- Institutional Detail: In the US, compulsory schooling laws required students to attend school until they turned 16 or 17.
  - Students born early in the year (e.g., Jan, Feb) start school at an older age. They turn 16 earlier in their school career and can legally drop out with slightly less education.
  - Students born late in the year (e.g., Oct, Nov) are younger when they start school. They are forced by the law to stay in school longer to reach their 16th birthday, resulting in slightly more education on average.
- IV Assumptions:
  - Relevance: Is quarter of birth correlated with years of schooling? Yes, the data showed a small but statistically significant relationship.
  - Exclusion restriction: Is it plausible that quarter of birth has no direct effect on wages, other than through its effect on schooling? It's hard to think of a reason why birth month would directly influence adult earnings, other than this institutional reason.

Angrist & Krueger (1991): Results

OLS Result: Found that an extra year of school was associated with about a 7.5% increase in wages.

- Likely an overestimate due to ability bias.
- Why is this an overestimate? Think about the OVB formula and the correlation between X and Z

2SLS Result: Using quarter of birth as an instrument, they found a very similar return to schooling, also around 7.5%.

- Interpretation (LATE):
  - This IV estimate is the *Local* Average Treatment Effect.
  - It represents the return to schooling for the *compliers*: those individuals who were induced to stay in school for an extra bit of time because of compulsory schooling laws.
  - These are likely people on the margin of dropping out of high school. The result may not apply to the returns to getting a college degree.

Other Example: Demand Estimation

Let's say we want to estimate the price elasticity of demand for avocados.

The Structural Model:

1. **Demand Curve:** We want to estimate this equation. α_1 is our parameter of interest (the elasticity).

$$Q_i^d = \alpha_0 + \alpha_1 P_i + u_i$$

   - u_i represents unobserved demand shocks (e.g., a new health trend, a guacamole festival).

2. **Supply Curve:** The market also has a supply curve.

$$Q_i^s = \beta_0 + \beta_1 P_i + v_i$$

   - v_i represents unobserved supply shocks (e.g., a localized pest).

The Endogeneity Issue

In market equilibrium, $Q_i^d = Q_i^s = Q_i$, and the price P_i adjusts to make this happen. This means price P_i is determined by *both* supply and demand factors.

- If there is a positive demand shock ($u_i > 0$), the demand curve shifts right. This leads to a higher equilibrium price **and** a higher quantity.
- Therefore, P_i is positively correlated with the demand error term u_i.
- This violates the core OLS assumption that regressors are uncorrelated with the error term: $Cov(P_i, u_i) \neq 0$.
- Running OLS of `Q` on `P` will give a biased estimate of α_1.

The Solution: IV

We need a variable, let's call it Z, that provides an *exogenous* source of variation in price. This variable is our **Instrument**.

A Good Instrument: An Exogenous Supply Shifter

Let's use a variable that only shifts the supply curve. A perfect example is **weather**. - Let Z_i be a measure of growing conditions (e.g., rainfall in avocado-growing regions). Good weather increases supply.

The Updated Structural Model

- **Demand Curve (Unchanged):** Weather doesn't directly affect how many avocados a person wants to buy at a given price.

$$Q_i = \alpha_0 + \alpha_1 P_i + u_i$$

- **Supply Curve (with Instrument):** Weather directly affects the quantity supplied.

$$Q_i = \beta_0 + \beta_1 P_i + \delta Z_i + v_i$$

The Two IV Conditions (in this context)

1. **Relevance:** The instrument must affect the endogenous variable. Good weather must actually change the price of avocados.
   - *Algebraically:* $\delta \neq 0$, which implies $Cov(Z_i, P_i) \neq 0$.
2. **Exclusion Restriction:** The instrument must be uncorrelated with the error term in the primary equation (demand). Good weather doesn't directly cause a craving for avocados.
   - *Algebraically:* $Cov(Z_i, u_i) = 0$. This is the crucial identifying assumption.

Identification: The Algebra of IV

How do we use our instrument Z and the exclusion restriction to find α_1?

Our goal is to estimate α_1 from the demand equation:

$$Q_i = \alpha_0 + \alpha_1 P_i + u_i$$

The problem is that $Cov(P_i, u_i) \neq 0$. But we have assumed $Cov(Z_i, u_i) = 0$. Let's use that!

Step 1: Take the covariance of the demand equation with our instrument Z_i.

$$Cov(Z_i, Q_i) = Cov(Z_i, \alpha_0 + \alpha_1 P_i + u_i)$$

Step 2: Use the properties of covariance to expand the right side.

$$Cov(Z_i, Q_i) = Cov(Z_i, \alpha_0) + Cov(Z_i, \alpha_1 P_i) + Cov(Z_i, u_i)$$

Identification: The Algebra of IV (Cont.)

Step 3: Simplify.

- $Cov(Z_i, \alpha_0) = 0$ (covariance with a constant is zero).
- $Cov(Z_i, \alpha_1 P_i) = \alpha_1 Cov(Z_i, P_i)$ (constants can be pulled out).
- $Cov(Z_i, u_i) = 0$ (this is our crucial **Exclusion Restriction** assumption!).

Step 4: See what's left.

$$Cov(Z_i, Q_i) = \alpha_1 Cov(Z_i, P_i)$$

Step 5: Solve for our parameter of interest, α_1.

$$\alpha_1 = \frac{Cov(Z_i, Q_i)}{Cov(Z_i, P_i)}$$

This is the Instrumental Variables estimator for α_1.

The Link to Reduced Form & First Stage

The expression $\frac{Cov(Z_i, Q_i)}{Cov(Z_i, P_i)}$ has a beautiful and intuitive interpretation. Let's look at two simple regressions.

1. **The First Stage:** How does our instrument (Z) affect our endogenous variable (P)?

$$P_i = \pi_{p0} + \pi_{p1} Z_i + \text{error}_p$$

   - The OLS estimate of π_{p1} is: $\hat{\pi}_{p1} = \frac{Cov(Z_i, P_i)}{Var(Z_i)}$.
   - This tells us how much price changes for a one-unit change in weather.

2. **The Reduced Form:** How does our instrument (Z) affect our final outcome (Q)?

$$Q_i = \pi_{q0} + \pi_{q1} Z_i + \text{error}_q$$

   - The OLS estimate of π_{q1} is: $\hat{\pi}_{q1} = \frac{Cov(Z_i, Q_i)}{Var(Z_i)}$.
   - This tells us how much quantity changes for a one-unit change in weather.

Ratio of RF and FS

Now, let's look at the ratio of these two coefficients:

$$\frac{\hat{\pi}_{q1}}{\hat{\pi}_{p1}} = \frac{\frac{Cov(Z_i, Q_i)}{Var(Z_i)}}{\frac{Cov(Z_i, P_i)}{Var(Z_i)}} = \frac{Cov(Z_i, Q_i)}{Cov(Z_i, P_i)}$$

This is exactly our IV estimator!

$$\hat{\alpha}_{1,IV} = \frac{\text{Reduced Form Effect}}{\text{First Stage Effect}} = \frac{\text{Effect of Z on Q}}{\text{Effect of Z on P}}$$

The structural parameter α_1 is **identified** because it can be expressed as the ratio of two estimable parameters from simple regressions. We have used the exogenous variation from the instrument to isolate the causal effect of price on quantity demanded.

Summary and Key Takeaways

- Problem: Endogeneity ($\text{Cov}(X, \epsilon) \neq 0$) makes OLS biased and inconsistent for estimating causal effects.
  - Solution: A valid Instrumental Variable (Z) is:
    - **Relevant**: Correlated with X.
    - **Exogenous**: Uncorrelated with the error term ϵ (the Exclusion Restriction).
- **Interpretation**: IV estimates the Local Average Treatment Effect (LATE) - the causal effect for the subpopulation of "compliers" whose behavior is changed by the instrument.
- **Estimation**: Use the Wald estimator for simple cases, and Two-Stage Least Squares (2SLS) for the general case.
- **In Practice**: Finding a valid instrument is the hardest part. Always check for weak instruments (First-stage F-statistic) and be prepared to defend your exclusion restriction.