

Worksheet 1: Deriving the Wald Estimator

Gov 51 · Scott Cunningham

Spring 2026

This worksheet derives the Wald estimator step by step. We only need three properties of covariance to get there.

Three properties we will use

P1. Linearity: $\text{Cov}(aX + bW, Y) = a \text{Cov}(X, Y) + b \text{Cov}(W, Y)$

P2. Constants: $\text{Cov}(c, Y) = 0$ for any constant c

P3. Independence: If $Z \perp\!\!\!\perp u$, then $\text{Cov}(Z, u) = 0$

Setup. The true model is $Y_i = \beta_0 + \beta_1 D_i + u_i$. We want β_1 but cannot trust OLS because D correlates with u . Our instrument Z satisfies: **(A1)** $\text{Cov}(Z, D) \neq 0$ and **(A2)** $Z \perp\!\!\!\perp u$.

Step 1. Write down the structural equation.

$$Y_i = \beta_0 + \beta_1 D_i + u_i$$

This is what we believe about the world: β_1 is the causal effect of D on Y , and u is everything else. We cannot estimate β_1 by OLS because D and u are correlated.

Step 2. Take the covariance of both sides with Z .

$$\text{Cov}(Y, Z) = \text{Cov}(\beta_0 + \beta_1 D + u, Z)$$

This is the key move. We multiply both sides by the instrument. If Z is a valid instrument, this will let us isolate β_1 .

Step 3. Apply P1 (linearity) to split the right side into three terms.

$$\text{Cov}(Y, Z) = \text{Cov}(\beta_0, Z) + \beta_1 \text{Cov}(D, Z) + \text{Cov}(u, Z)$$

Covariance is linear, so we can separate the three terms. We also pull β_1 out front because it is a constant scalar.

Step 4. Apply P2: the intercept β_0 is a constant, so $\text{Cov}(\beta_0, Z) = 0$.

$$\text{Cov}(Y, Z) = \beta_1 \text{Cov}(D, Z) + \text{Cov}(u, Z)$$

The intercept drops out entirely. It doesn't vary across observations, so it has no covariance with anything.

Step 5. Apply A2 and P3: since $Z \perp\!\!\!\perp u$, we have $\text{Cov}(u, Z) = 0$.

$$\text{Cov}(Y, Z) = \beta_1 \text{Cov}(D, Z)$$

This is the exclusion restriction doing its work. Because Z has no direct path to Y (only through D), the covariance between Z and the error is zero. If this assumption fails, the proof breaks here.

Step 6. Divide both sides by $\text{Cov}(D, Z)$. This is valid because A1 guarantees $\text{Cov}(D, Z) \neq 0$.

$$\beta_1 = \frac{\text{Cov}(Y, Z)}{\text{Cov}(D, Z)}$$

This is the Wald estimator: the covariance of the outcome with the instrument, divided by the covariance of the treatment with the instrument.