Economics 326
Methods of Empirical Research in Economics

Lecture 16: Large sample results: Consistency

Vadim Marmer
University of British Columbia

March 29, 2011

Why we need the large sample theory

- We have shown that the OLS estimator $\hat{\beta}$ has some desirable properties:

Why we need the large sample theory

- We have shown that the OLS estimator $\hat{\beta}$ has some desirable properties:
  - $\hat{\beta}$ is unbiased if the errors are strongly exogenous: $E(U|X) = 0$.

Why we need the large sample theory

- We have shown that the OLS estimator $\hat{\beta}$ has some desirable properties:
  - $\hat{\beta}$ is unbiased if the errors are strongly exogenous: $E(U|X) = 0$.
  - If in addition the errors are homoskedastic then $\widehat{Var}\left(\hat{\beta}\right) = s^2 / \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2$ is an unbiased estimator of the conditional variance of the OLS estimator $\hat{\beta}$.

Why we need the large sample theory

- We have shown that the OLS estimator $\hat{\beta}$ has some desirable properties:
  - $\hat{\beta}$ is unbiased if the errors are strongly exogenous: $E\left(U|X\right) = 0$.
  - If in addition the errors are homoskedastic then $\widehat{Var}\left(\hat{\beta}\right) = s^2 / \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2$ is an unbiased estimator of the conditional variance of the OLS estimator $\hat{\beta}$.
  - If in addition the errors are normally distributed (given X) then $T = \left(\hat{\beta} - \beta\right) / \sqrt{\widehat{Var}\left(\hat{\beta}\right)}$ has a t distribution which can be used for hypotheses testing.

Why we need the large sample theory

- If the errors are only weakly exogenous:

$$E\left(X_i U_i\right) = 0,$$

the OLS estimator is in general biased.

Why we need the large sample theory

- If the errors are only weakly exogenous:

$$E\left(X_i U_i\right) = 0,$$

  the OLS estimator is in general biased.
- If the errors are heteroskedastic:

$$E\left(U_i^2 | X_i\right) = h\left(X_i\right),$$

  the "usual" variance formula is invalid; we also do not have an unbiased estimator for the variance in this case.

Why we need the large sample theory

- If the errors are only weakly exogenous:

$$E\left(X_i U_i\right) = 0,$$

  the OLS estimator is in general biased.
- If the errors are heteroskedastic:

$$E\left(U_i^2 | X_i\right) = h\left(X_i\right),$$

  the "usual" variance formula is invalid; we also do not have an unbiased estimator for the variance in this case.
- If the errors are not normally distributed conditional on X then T- and F-statistics do not have t and F distributions under the null hypothesis.

Why we need the large sample theory

- If the errors are only weakly exogenous:

$$E\left(X_i U_i\right) = 0,$$

  the OLS estimator is in general biased.
- If the errors are heteroskedastic:

$$E\left(U_i^2 | X_i\right) = h\left(X_i\right),$$

  the "usual" variance formula is invalid; we also do not have an unbiased estimator for the variance in this case.
- If the errors are not normally distributed conditional on X then T- and F-statistics do not have t and F distributions under the null hypothesis.
- The asymptotic or large sample theory allows us to derive approximate properties and distributions of estimators and test statistics by assuming that the sample size n is very large.

Convergence in probability and LLN

- Let θ_n be a sequence of random variables indexed by the sample size n. We say that θ_n converges in probability if

$$\lim_{n \to \infty} P\left(|\theta_n - \theta| \geq \varepsilon\right) = 0 \text{ for all } \varepsilon > 0.$$

Convergence in probability and LLN

- Let θ_n be a sequence of random variables indexed by the sample size n. We say that θ_n converges in probability if

$$\lim_{n \to \infty} P\left(|\theta_n - \theta| \geq \varepsilon\right) = 0 \text{ for all } \varepsilon > 0.$$

- We denote this as $\theta_n \to_p \theta$ or $p\lim \theta_n = \theta$.

Convergence in probability and LLN

- Let θ_n be a sequence of random variables indexed by the sample size n. We say that θ_n converges in probability if

$$\lim_{n \to \infty} P\left(|\theta_n - \theta| \geq \varepsilon\right) = 0 \text{ for all } \varepsilon > 0.$$

- We denote this as $\theta_n \to_p \theta$ or $p\lim \theta_n = \theta$.
- An example of convergence in probability is a Law of Large Numbers (LLN):

Convergence in probability and LLN

- Let θ_n be a sequence of random variables indexed by the sample size n. We say that θ_n converges in probability if

$$\lim_{n\to\infty} P\left(|\theta_n - \theta| \geq \varepsilon\right) = 0 \text{ for all } \varepsilon > 0.$$

- We denote this as $\theta_n \to_p \theta$ or $p\lim \theta_n = \theta$.
- An example of convergence in probability is a Law of Large Numbers (LLN):
  Let $X_1, X_2, \ldots, X_n$ be a random sample such that $E\left(X_i\right) = \mu$ for all $i = 1, \ldots, n$, and define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Convergence in probability and LLN

- Let θ_n be a sequence of random variables indexed by the sample size n. We say that θ_n converges in probability if

$$\lim_{n\to\infty} P\left(|\theta_n - \theta| \geq \varepsilon\right) = 0 \text{ for all } \varepsilon > 0.$$

- We denote this as $\theta_n \to_p \theta$ or $p\lim \theta_n = \theta$.
- An example of convergence in probability is a Law of Large Numbers (LLN):
  Let $X_1, X_2, \ldots, X_n$ be a random sample such that $E\left(X_i\right) = \mu$ for all $i = 1, \ldots, n$, and define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Then, under certain conditions,

$$\bar{X}_n \to_p \mu.$$

LLN

- Let $X_1, \ldots, X_n$ be a sample of independent identically distributed (iid) random variables. Let $EX_i = \mu$. If $Var\left(X_i\right) = \sigma^2 < \infty$ then

$$\bar{X}_n \rightarrow_p \mu.$$

LLN

- Let $X_1, \ldots, X_n$ be a sample of independent identically distributed (iid) random variables. Let $EX_i = \mu$. If $Var(X_i) = \sigma^2 < \infty$ then

$$\bar{X}_n \rightarrow_p \mu.$$

- In fact when the data are iid, the LLN holds if

$$E|X_i| < \infty,$$

but we prove the result under a stronger assumption that $Var(X_i) < \infty$.

LLN

- Let $X_1, \ldots, X_n$ be a sample of independent identically distributed (iid) random variables. Let $EX_i = \mu$. If $Var\left(X_i\right) = \sigma^2 < \infty$ then

$$\bar{X}_n \rightarrow_p \mu.$$

- In fact when the data are iid, the LLN holds if

$$E\left|X_i\right| < \infty,$$

  but we prove the result under a stronger assumption that $Var\left(X_i\right) < \infty$.

Markov's inequality

- **Markov's inequality.** Let W be a random variable. For $\varepsilon > 0$ and $r > 0$,
$$P\left(|W| \geq \varepsilon\right) \leq \frac{E\left|W\right|^{r}}{\varepsilon^{r}}.$$

Markov's inequality

- **Markov's inequality.** Let W be a random variable. For $\varepsilon > 0$ and $r > 0$,

$$P\left(|W| \geq \varepsilon\right) \leq \frac{E\left|W\right|^r}{\varepsilon^r}.$$

- With $r = 2$, we have **Chebyshev's inequality**.

Markov's inequality

- **Markov's inequality.** Let W be a random variable. For $\varepsilon > 0$ and $r > 0$,
$$P\left(|W| \geq \varepsilon\right) \leq \frac{E\left|W\right|^r}{\varepsilon^r}.$$

- With $r = 2$, we have **Chebyshev's inequality**. Suppose that $EX = \mu$.

Markov's inequality

- **Markov's inequality.** Let W be a random variable. For $\varepsilon > 0$ and $r > 0$,

$$P\left(|W| \geq \varepsilon\right) \leq \frac{E\left|W\right|^r}{\varepsilon^r}.$$

- With $r = 2$, we have **Chebyshev's inequality**. Suppose that $EX = \mu$. Take $W \equiv X - \mu$ and apply Markov's inequality with $r = 2$.

Markov's inequality

- **Markov's inequality.** Let W be a random variable. For $\varepsilon > 0$ and $r > 0$,
$$P(|W| \geq \varepsilon) \leq \frac{E|W|^r}{\varepsilon^r}.$$

- With $r = 2$, we have **Chebyshev's inequality**. Suppose that $EX = \mu$. Take $W \equiv X - \mu$ and apply Markov's inequality with $r = 2$. For $\varepsilon > 0$,
$$P(|X - \mu| \geq \varepsilon) \quad \leq \quad \frac{E|X - \mu|^2}{\varepsilon^2}$$

Markov's inequality

- **Markov's inequality.** Let W be a random variable. For $\varepsilon > 0$ and $r > 0$,

$$P\left(|W| \geq \varepsilon\right) \leq \frac{E\left|W\right|^r}{\varepsilon^r}.$$

- With $r = 2$, we have **Chebyshev's inequality**. Suppose that $EX = \mu$. Take $W \equiv X - \mu$ and apply Markov's inequality with $r = 2$. For $\varepsilon > 0$,

$$\begin{aligned} P\left(|X - \mu| \geq \varepsilon\right) &\leq \frac{E\left|X - \mu\right|^2}{\varepsilon^2} \\ &= \frac{Var\left(X\right)}{\varepsilon^2}. \end{aligned}$$

Markov's inequality

- **Markov's inequality.** Let W be a random variable. For $\varepsilon > 0$ and $r > 0$,

$$P(|W| \geq \varepsilon) \leq \frac{E|W|^r}{\varepsilon^r}.$$

- With $r = 2$, we have **Chebyshev's inequality**. Suppose that $EX = \mu$. Take $W \equiv X - \mu$ and apply Markov's inequality with $r = 2$. For $\varepsilon > 0$,

$$\begin{aligned} P(|X-\mu| \geq \varepsilon) &\leq \frac{E|X-\mu|^2}{\varepsilon^2} \\ &= \frac{Var(X)}{\varepsilon^2}. \end{aligned}$$

- Probability of observing an outlier (a large deviation of X from its mean μ) can be bounded by the variance.

Proof of the LLN

Fix $\varepsilon > 0$ and apply Markov's inequality with $r = 2$:

Proof of the LLN

Fix $\varepsilon > 0$ and apply Markov's inequality with $r = 2$:

$$P\left(|\bar{X}_n - \mu| \geq \varepsilon\right)$$

Proof of the LLN

Fix $\varepsilon > 0$ and apply Markov's inequality with $r = 2$:

$$P\left(|\bar{X}_n - \mu| \geq \varepsilon\right) = P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right)$$

Proof of the LLN

Fix $\varepsilon > 0$ and apply Markov's inequality with $r = 2$:

$$P\left(|\bar{X}_n - \mu| \geq \varepsilon\right) = P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right)$$

$$= \quad P\left(\left|\frac{1}{n}\sum_{i=1}^{n} (X_i - \mu)\right| \geq \varepsilon\right)$$

Proof of the LLN

Fix $\varepsilon > 0$ and apply Markov's inequality with $r = 2$:

$$P\left(|\bar{X}_n - \mu| \geq \varepsilon\right) = P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right)$$

$$= P\left(\left|\frac{1}{n}\sum_{i=1}^{n} (X_i - \mu)\right| \geq \varepsilon\right)$$

$$\leq \frac{E\left(\frac{1}{n}\sum_{i=1}^{n} (X_i - \mu)\right)^2}{\varepsilon^2}$$

Proof of the LLN

Fix $\varepsilon > 0$ and apply Markov's inequality with $r = 2$:

$$P\left(|\bar{X}_n - \mu| \geq \varepsilon\right) = P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right)$$

$$= P\left(\left|\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)\right| \geq \varepsilon\right)$$

$$\leq \frac{E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)\right)^2}{\varepsilon^2}$$

$$= \frac{1}{n^2\varepsilon^2}\left(\sum_{i=1}^{n} E(X_i - \mu)^2 + \sum_{i=1}^{n}\sum_{j \neq i} E(X_i - \mu)(X_j - \mu)\right)$$

Proof of the LLN

Fix $\varepsilon > 0$ and apply Markov's inequality with $r = 2$:

$$P\left(|\bar{X}_n - \mu| \geq \varepsilon\right) = P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right)$$

$$= P\left(\left|\frac{1}{n}\sum_{i=1}^{n} (X_i - \mu)\right| \geq \varepsilon\right)$$

$$\leq \frac{E\left(\frac{1}{n}\sum_{i=1}^{n} (X_i - \mu)\right)^2}{\varepsilon^2}$$

$$= \frac{1}{n^2\varepsilon^2}\left(\sum_{i=1}^{n} E(X_i - \mu)^2 + \sum_{i=1}^{n}\sum_{j\neq i} E(X_i - \mu)(X_j - \mu)\right)$$

$$= \frac{1}{n^2\varepsilon^2}\left(\sum_{i=1}^{n} Var(X_i) + \sum_{i=1}^{n}\sum_{j\neq i} Cov(X_i, X_j)\right)$$

Proof of the LLN

Fix $\varepsilon > 0$ and apply Markov's inequality with $r = 2$:

$$
\begin{aligned}
P\left(|\bar{X}_n - \mu| \geq \varepsilon\right) &= P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right) \\
&= P\left(\left|\frac{1}{n}\sum_{i=1}^{n} (X_i - \mu)\right| \geq \varepsilon\right) \\
&\leq \frac{E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)\right)^2}{\varepsilon^2} \\
&= \frac{1}{n^2\varepsilon^2}\left(\sum_{i=1}^{n} E(X_i - \mu)^2 + \sum_{i=1}^{n}\sum_{j \neq i} E(X_i - \mu)(X_j - \mu)\right) \\
&= \frac{1}{n^2\varepsilon^2}\left(\sum_{i=1}^{n} Var(X_i) + \sum_{i=1}^{n}\sum_{j \neq i} Cov(X_i, X_j)\right) \\
&= \frac{n\sigma^2}{n^2\varepsilon^2}
\end{aligned}
$$

Proof of the LLN

Fix $\varepsilon > 0$ and apply Markov's inequality with $r = 2$:

$$P\left(|\bar{X}_n - \mu| \geq \varepsilon\right) = P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right)$$

$$= P\left(\left|\frac{1}{n}\sum_{i=1}^{n} (X_i - \mu)\right| \geq \varepsilon\right)$$

$$\leq \frac{E\left(\frac{1}{n}\sum_{i=1}^{n} (X_i - \mu)\right)^2}{\varepsilon^2}$$

$$= \frac{1}{n^2\varepsilon^2}\left(\sum_{i=1}^{n} E(X_i - \mu)^2 + \sum_{i=1}^{n}\sum_{j \neq i} E(X_i - \mu)(X_j - \mu)\right)$$

$$= \frac{1}{n^2\varepsilon^2}\left(\sum_{i=1}^{n} Var(X_i) + \sum_{i=1}^{n}\sum_{j \neq i} Cov(X_i, X_j)\right)$$

$$= \frac{n\sigma^2}{n^2\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

Proof of the LLN

Fix $\varepsilon > 0$ and apply Markov's inequality with $r = 2$:

$$P\left(|\bar{X}_n - \mu| \geq \varepsilon\right) = P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right)$$

$$= P\left(\left|\frac{1}{n}\sum_{i=1}^{n} (X_i - \mu)\right| \geq \varepsilon\right)$$

$$\leq \frac{E\left(\frac{1}{n}\sum_{i=1}^{n} (X_i - \mu)\right)^2}{\varepsilon^2}$$

$$= \frac{1}{n^2\varepsilon^2}\left(\sum_{i=1}^{n} E(X_i - \mu)^2 + \sum_{i=1}^{n}\sum_{j\neq i} E(X_i - \mu)(X_j - \mu)\right)$$

$$= \frac{1}{n^2\varepsilon^2}\left(\sum_{i=1}^{n} Var(X_i) + \sum_{i=1}^{n}\sum_{j\neq i} Cov(X_i, X_j)\right)$$

$$= \frac{n\sigma^2}{n^2\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \to 0 \text{ as } n \to \infty \text{ for all } \varepsilon > 0.$$

Averaging and variance reduction

- Let $X_1, \ldots, X_n$ be a sample and suppose that

$$E(X_i) = \mu \text{ for all } i = 1, \ldots, n,$$

Averaging and variance reduction

- Let $X_1, \ldots, X_n$ be a sample and suppose that

$$\begin{aligned} E\left(X_i\right) &= \mu \text{ for all } i = 1, \ldots, n, \\ Var\left(X_i\right) &= \sigma^2 \text{ for all } i = 1, \ldots, n, \end{aligned}$$

Averaging and variance reduction

- Let $X_1, \ldots, X_n$ be a sample and suppose that

$$\begin{aligned} E(X_i) &= \mu \text{ for all } i = 1, \ldots, n, \\ Var(X_i) &= \sigma^2 \text{ for all } i = 1, \ldots, n, \\ Cov(X_i, X_j) &= 0 \text{ for all } j \neq i. \end{aligned}$$

Averaging and variance reduction

- Let $X_1, \ldots, X_n$ be a sample and suppose that

$$\begin{aligned} E(X_i) &= \mu \text{ for all } i = 1, \ldots, n, \\ Var(X_i) &= \sigma^2 \text{ for all } i = 1, \ldots, n, \\ Cov(X_i, X_j) &= 0 \text{ for all } j \neq i. \end{aligned}$$

- Consider the mean of the average:

$$E(\bar{X}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

Averaging and variance reduction

- Let $X_1, \ldots, X_n$ be a sample and suppose that

$$\begin{aligned} E\left(X_i\right) &= \mu \text{ for all } i = 1, \ldots, n, \\ Var\left(X_i\right) &= \sigma^2 \text{ for all } i = 1, \ldots, n, \\ Cov\left(X_i, X_j\right) &= 0 \text{ for all } j \neq i. \end{aligned}$$

- Consider the mean of the average:

$$\begin{aligned} E\left(\bar{X}_n\right) &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} E\left(X_i\right) \end{aligned}$$

Averaging and variance reduction

- Let $X_1, \ldots, X_n$ be a sample and suppose that

$$
\begin{aligned}
E(X_i) &= \mu \text{ for all } i = 1, \ldots, n, \\
Var(X_i) &= \sigma^2 \text{ for all } i = 1, \ldots, n, \\
Cov(X_i, X_j) &= 0 \text{ for all } j \neq i.
\end{aligned}
$$

- Consider the mean of the average:

$$
\begin{aligned}
E(\bar{X}_n) &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n}\sum_{i=1}^{n} E(X_i) \\
&= \frac{1}{n}\sum_{i=1}^{n} \mu
\end{aligned}
$$

Averaging and variance reduction

- Let $X_1, \ldots, X_n$ be a sample and suppose that

$$\begin{aligned} E(X_i) &= \mu \text{ for all } i = 1, \ldots, n, \\ Var(X_i) &= \sigma^2 \text{ for all } i = 1, \ldots, n, \\ Cov(X_i, X_j) &= 0 \text{ for all } j \neq i. \end{aligned}$$

- Consider the mean of the average:

$$\begin{aligned} E(\bar{X}_n) &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} E(X_i) \\ &= \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{1}{n} n\mu = \mu. \end{aligned}$$

Averaging and variance reduction

- Consider the variance of the average:

$$Var\left(\bar{X}_n\right) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

Averaging and variance reduction

- Consider the variance of the average:

$$
\begin{aligned}
Var\left(\bar{X}_n\right) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right)
\end{aligned}
$$

Averaging and variance reduction

- Consider the variance of the average:

$$
\begin{aligned}
Var\left(\bar{X}_n\right) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n} Var\left(X_i\right) + \sum_{i=1}^{n}\sum_{j \neq i} Cov\left(X_i, X_j\right)\right)
\end{aligned}
$$

Averaging and variance reduction

- Consider the variance of the average:

$$
\begin{aligned}
Var\left(\bar{X}_n\right) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n} Var\left(X_i\right) + \sum_{i=1}^{n}\sum_{j\neq i} Cov\left(X_i, X_j\right)\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n} \sigma^2 + \sum_{i=1}^{n}\sum_{j\neq i} 0\right)
\end{aligned}
$$

Averaging and variance reduction

- Consider the variance of the average:

$$
\begin{aligned}
Var\left(\bar{X}_n\right) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n} Var\left(X_i\right) + \sum_{i=1}^{n}\sum_{j\neq i} Cov\left(X_i, X_j\right)\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n} \sigma^2 + \sum_{i=1}^{n}\sum_{j\neq i} 0\right) \\
&= \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}.
\end{aligned}
$$

Averaging and variance reduction

- Consider the variance of the average:

$$
\begin{aligned}
Var\left(\bar{X}_n\right) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n} Var\left(X_i\right) + \sum_{i=1}^{n}\sum_{j \neq i} Cov\left(X_i, X_j\right)\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n} \sigma^2 + \sum_{i=1}^{n}\sum_{j \neq i} 0\right) \\
&= \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}.
\end{aligned}
$$

- The variance of the average approaches zero as $n \to \infty$ if the

Convergence in probability: properties

- **Slutsky's Lemma.** Suppose that $\theta_n \rightarrow_p \theta$, and let g be a function continuous at θ.

Convergence in probability: properties

- **Slutsky's Lemma.** Suppose that $\theta_n \rightarrow_p \theta$, and let g be a function continuous at θ. Then,

$$g(\theta_n) \rightarrow_p g(\theta).$$

Convergence in probability: properties

- **Slutsky's Lemma.** Suppose that $\theta_n \to_p \theta$, and let g be a function continuous at θ. Then,

$$g(\theta_n) \to_p g(\theta).$$

  - If $\theta_n \to_p \theta$, then $\theta_n^2 \to_p \theta^2$.

Convergence in probability: properties

- **Slutsky's Lemma.** Suppose that $\theta_n \rightarrow_p \theta$, and let g be a function continuous at θ. Then,

$$g(\theta_n) \rightarrow_p g(\theta).$$

  - If $\theta_n \rightarrow_p \theta$, then $\theta_n^2 \rightarrow_p \theta^2$.
  - If $\theta_n \rightarrow_p \theta$ and $\theta \neq 0$, then $1/\theta_n \rightarrow_p 1/\theta$.

Convergence in probability: properties

- **Slutsky's Lemma.** Suppose that $\theta_n \rightarrow_p \theta$, and let g be a function continuous at θ. Then,

$$g\left(\theta_n\right) \rightarrow_p g\left(\theta\right).$$

  - If $\theta_n \rightarrow_p \theta$, then $\theta_n^2 \rightarrow_p \theta^2$.
  - If $\theta_n \rightarrow_p \theta$ and $\theta \neq 0$, then $1/\theta_n \rightarrow_p 1/\theta$.
- Suppose that $\theta_n \rightarrow_p \theta$ and $\lambda_n \rightarrow_p \lambda$.

Convergence in probability: properties

- **Slutsky's Lemma.** Suppose that $\theta_n \rightarrow_p \theta$, and let g be a function continuous at θ. Then,

$$g\left(\theta_n\right) \rightarrow_p g\left(\theta\right).$$

  - If $\theta_n \rightarrow_p \theta$, then $\theta_n^2 \rightarrow_p \theta^2$.
  - If $\theta_n \rightarrow_p \theta$ and $\theta \neq 0$, then $1/\theta_n \rightarrow_p 1/\theta$.
- Suppose that $\theta_n \rightarrow_p \theta$ and $\lambda_n \rightarrow_p \lambda$.Then,
  - $\theta_n + \lambda_n \rightarrow_p \theta + \lambda$.

Convergence in probability: properties

- **Slutsky's Lemma.** Suppose that $\theta_n \to_p \theta$, and let g be a function continuous at θ. Then,

$$g(\theta_n) \to_p g(\theta).$$

  - If $\theta_n \to_p \theta$, then $\theta_n^2 \to_p \theta^2$.
  - If $\theta_n \to_p \theta$ and $\theta \neq 0$, then $1/\theta_n \to_p 1/\theta$.

- Suppose that $\theta_n \to_p \theta$ and $\lambda_n \to_p \lambda$.Then,
  - $\theta_n + \lambda_n \to_p \theta + \lambda$.
  - $\theta_n \lambda_n \to_p \theta\lambda$.

Convergence in probability: properties

- **Slutsky's Lemma.** Suppose that $\theta_n \rightarrow_p \theta$, and let g be a function continuous at θ. Then,

$$g(\theta_n) \rightarrow_p g(\theta).$$

  - If $\theta_n \rightarrow_p \theta$, then $\theta_n^2 \rightarrow_p \theta^2$.
  - If $\theta_n \rightarrow_p \theta$ and $\theta \neq 0$, then $1/\theta_n \rightarrow_p 1/\theta$.
- Suppose that $\theta_n \rightarrow_p \theta$ and $\lambda_n \rightarrow_p \lambda$.Then,
  - $\theta_n + \lambda_n \rightarrow_p \theta + \lambda$.
  - $\theta_n \lambda_n \rightarrow_p \theta \lambda$.
  - $\theta_n / \lambda_n \rightarrow_p \theta / \lambda$ provided that $\lambda \neq 0$.

Consistency

- Let $\hat{\beta}_n$ be an estimator of β based on a sample of size n.

Consistency

- Let $\hat{\beta}_n$ be an estimator of β based on a sample of size n.
- We say that $\hat{\beta}_n$ is a consistent estimator of β if as $n \longrightarrow \infty$,

$$\hat{\beta}_n \longrightarrow_p \beta.$$

Consistency

- Let $\hat{\beta}_n$ be an estimator of β based on a sample of size n.
- We say that $\hat{\beta}_n$ is a consistent estimator of β if as $n \longrightarrow \infty$,

$$\hat{\beta}_n \longrightarrow_p \beta.$$

- Consistency means that the probability of the event that the distance between $\hat{\beta}_n$ and β exceeds $\varepsilon > 0$ can be made arbitrary small by increasing the sample size.

Consistency of OLS

- Suppose that:
  1. The data $\{(Y_i, X_i) : i = 1, \ldots, n\}$ are iid.

Consistency of OLS

- Suppose that:
  1. The data $\{(Y_i, X_i) : i = 1, \ldots, n\}$ are iid.
  2. $Y_i = \beta_0 + \beta_1 X_i + U_i$, where $E(U_i) = 0$.

Consistency of OLS

- Suppose that:
  1. The data $\{(Y_i, X_i) : i = 1, \ldots, n\}$ are iid.
  2. $Y_i = \beta_0 + \beta_1 X_i + U_i$, where $E(U_i) = 0$.
  3. $E(X_i U_i) = 0$.

Consistency of OLS

- Suppose that:
  1. The data $\{(Y_i, X_i) : i = 1, \ldots, n\}$ are iid.
  2. $Y_i = \beta_0 + \beta_1 X_i + U_i$, where $E(U_i) = 0$.
  3. $E(X_i U_i) = 0$.
  4. $0 < Var(X_i) < \infty$.

Consistency of OLS

- Suppose that:
  1. The data $\{(Y_i, X_i) : i = 1, \ldots, n\}$ are iid.
  2. $Y_i = \beta_0 + \beta_1 X_i + U_i$, where $E(U_i) = 0$.
  3. $E(X_i U_i) = 0$.
  4. $0 < Var(X_i) < \infty$.
- Let $\hat{\beta}_{0,n}$ and $\hat{\beta}_{1,n}$ be the OLS estimators of β_0 and β_1 respectively based on a sample of size n. Under Assumptions 1-4,

$$\hat{\beta}_{0,n} \to_p \beta_0,$$
$$\hat{\beta}_{1,n} \to_p \beta_1.$$

Consistency of OLS

- Suppose that:
  1. The data $\{(Y_i, X_i) : i = 1, \ldots, n\}$ are iid.
  2. $Y_i = \beta_0 + \beta_1 X_i + U_i$, where $E(U_i) = 0$.
  3. $E(X_i U_i) = 0$.
  4. $0 < Var(X_i) < \infty$.
- Let $\hat{\beta}_{0,n}$ and $\hat{\beta}_{1,n}$ be the OLS estimators of β_0 and β_1 respectively based on a sample of size n. Under Assumptions 1-4,

$$\hat{\beta}_{0,n} \to_p \beta_0,$$
$$\hat{\beta}_{1,n} \to_p \beta_1.$$

- The key identifying assumption is Assumption 3: $Cov(X_i, U_i) = 0$.

Proof of consistency

- Write

$$\hat{\beta}_{1,n} = \frac{\sum_{i=1}^{n} (X_i - \bar{X}_n)\, Y_i}{\sum_{i=1}^{n} (X_i - \bar{X}_n)^2}$$

Proof of consistency

- Write

$$\hat{\beta}_{1,n} = \frac{\sum_{i=1}^{n} (X_i - \bar{X}_n)\, Y_i}{\sum_{i=1}^{n} (X_i - \bar{X}_n)^2} \quad = \quad \beta_1 + \frac{\sum_{i=1}^{n} (X_i - \bar{X}_n)\, U_i}{\sum_{i=1}^{n} (X_i - \bar{X}_n)^2}$$

Proof of consistency

- Write

$$\hat{\beta}_{1,n} = \frac{\sum_{i=1}^{n} \left(X_i - \bar{X}_n\right) Y_i}{\sum_{i=1}^{n} \left(X_i - \bar{X}_n\right)^2} = \beta_1 + \frac{\sum_{i=1}^{n} \left(X_i - \bar{X}_n\right) U_i}{\sum_{i=1}^{n} \left(X_i - \bar{X}_n\right)^2}$$

$$= \beta_1 + \frac{\frac{1}{n} \sum_{i=1}^{n} \left(X_i - \bar{X}_n\right) U_i}{\frac{1}{n} \sum_{i=1}^{n} \left(X_i - \bar{X}_n\right)^2}.$$

Proof of consistency

- Write

$$\hat{\beta}_{1,n} = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)Y_i}{\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2} = \beta_1 + \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)U_i}{\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2} = \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)U_i}{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2}.$$

- We will show that

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)U_i \to_p 0,$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 \to_p Var\left(X_i\right),$$

Proof of consistency

- Write

$$\hat{\beta}_{1,n} = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)Y_i}{\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2} = \beta_1 + \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)U_i}{\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2} = \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)U_i}{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2}.$$

- We will show that

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)U_i \rightarrow_p 0,$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 \rightarrow_p Var\left(X_i\right),$$

- Since $Var\left(X_i\right) \neq 0$,

$$\hat{\beta}_{1,n} = \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)U_i}{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2} \rightarrow_p \beta_1 + \frac{0}{Var\left(X_i\right)} = \beta_1.$$

$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)U_i \rightarrow_p 0$

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)U_i$$

$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right) U_i \rightarrow_p 0$

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right) U_i = \frac{1}{n}\sum_{i=1}^{n} X_i U_i - \bar{X}_n \left(\frac{1}{n}\sum_{i=1}^{n} U_i\right).$$

$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)U_i \to_p 0$

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)U_i = \frac{1}{n}\sum_{i=1}^{n}X_iU_i - \bar{X}_n\left(\frac{1}{n}\sum_{i=1}^{n}U_i\right).$$

By the LLN,

$$\frac{1}{n}\sum_{i=1}^{n}X_iU_i \to_p E\left(X_iU_i\right) = 0,$$

$$\bar{X}_n \to_p E\left(X_i\right),$$

$$\frac{1}{n}\sum_{i=1}^{n}U_i \to_p E\left(U_i\right) = 0.$$

$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)U_i \to_p 0$

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)U_i = \frac{1}{n}\sum_{i=1}^{n}X_iU_i - \bar{X}_n\left(\frac{1}{n}\sum_{i=1}^{n}U_i\right).$$

By the LLN,

$$\frac{1}{n}\sum_{i=1}^{n}X_iU_i \to_p E\left(X_iU_i\right) = 0,$$

$$\bar{X}_n \to_p E\left(X_i\right),$$

$$\frac{1}{n}\sum_{i=1}^{n}U_i \to_p E\left(U_i\right) = 0.$$

Hence,

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)U_i = \frac{1}{n}\sum_{i=1}^{n}X_iU_i - \bar{X}_n\left(\frac{1}{n}\sum_{i=1}^{n}U_i\right) &\to_p 0 - E\left(X_i\right)\cdot 0 \\ &= 0.\end{aligned}$$

$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 \rightarrow_p Var\left(X_i\right)$

- First,

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2$$

$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 \rightarrow_p Var\left(X_i\right)$

- First,

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - 2\bar{X}_n X_i + \bar{X}_n^2\right)$$

$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 \rightarrow_p Var\left(X_i\right)$

- First,

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - 2\bar{X}_n X_i + \bar{X}_n^2\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\bar{X}_n \frac{1}{n}\sum_{i=1}^{n} X_i + \bar{X}_n^2 \end{aligned}$$

$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 \rightarrow_p Var\left(X_i\right)$

- First,

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - 2\bar{X}_n X_i + \bar{X}_n^2\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - 2\bar{X}_n\frac{1}{n}\sum_{i=1}^{n}X_i + \bar{X}_n^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - 2\bar{X}_n\bar{X}_n + \bar{X}_n^2 =
\end{aligned}$$

$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 \rightarrow_p Var\left(X_i\right)$

- First,

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - 2\bar{X}_n X_i + \bar{X}_n^2\right) \\
&= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\bar{X}_n \frac{1}{n}\sum_{i=1}^{n} X_i + \bar{X}_n^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\bar{X}_n\bar{X}_n + \bar{X}_n^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}_n^2.
\end{aligned}$$

$\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 \to_p Var\left(X_i\right)$

- First,

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - 2\bar{X}_n X_i + \bar{X}_n^2\right) \\
&= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\bar{X}_n \frac{1}{n}\sum_{i=1}^{n} X_i + \bar{X}_n^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\bar{X}_n\bar{X}_n + \bar{X}_n^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}_n^2.
\end{aligned}$$

- By the LLN, $\frac{1}{n}\sum_{i=1}^{n} X_i^2 \to_p E\left(X_i^2\right)$ and $\bar{X}_n \to_p EX_i$.

$\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}_n\right)^2 \to_p Var\left(X_i\right)$

- First,

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}_n\right)^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2-2\bar{X}_nX_i+\bar{X}_n^2\right)\\ &= \frac{1}{n}\sum_{i=1}^{n}X_i^2-2\bar{X}_n\frac{1}{n}\sum_{i=1}^{n}X_i+\bar{X}_n^2\\ &= \frac{1}{n}\sum_{i=1}^{n}X_i^2-2\bar{X}_n\bar{X}_n+\bar{X}_n^2=\frac{1}{n}\sum_{i=1}^{n}X_i^2-\bar{X}_n^2.\end{aligned}$$

- By the LLN, $\frac{1}{n}\sum_{i=1}^{n}X_i^2 \to_p E\left(X_i^2\right)$ and $\bar{X}_n \to_p EX_i$.
- By Slutsky's Lemma, $\bar{X}_n^2 \to_p \left(EX_i\right)^2$.

$\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}_n\right)^2 \to_p Var\left(X_i\right)$

- First,

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2-2\bar{X}_nX_i+\bar{X}_n^2\right)\\ &= \frac{1}{n}\sum_{i=1}^{n}X_i^2-2\bar{X}_n\frac{1}{n}\sum_{i=1}^{n}X_i+\bar{X}_n^2\\ &= \frac{1}{n}\sum_{i=1}^{n}X_i^2-2\bar{X}_n\bar{X}_n+\bar{X}_n^2=\frac{1}{n}\sum_{i=1}^{n}X_i^2-\bar{X}_n^2.\end{aligned}$$

- By the LLN, $\frac{1}{n}\sum_{i=1}^{n}X_i^2 \to_p E\left(X_i^2\right)$ and $\bar{X}_n \to_p EX_i$.
- By Slutsky's Lemma, $\bar{X}_n^2 \to_p (EX_i)^2$.
- Thus,

$$\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2=\frac{1}{n}\sum_{i=1}^{n}X_i^2-\bar{X}_n^2 \to_p E\left(X_i^2\right)-(EX_i)^2=Var\left(X_i\right).$$

Multiple regression

- Under similar conditions to 1-4, one can establish consistency of OLS for the multiple linear regression model:

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \ldots + \beta_k X_{k,i} + U_i,$$

where $EU_i = 0$.

Multiple regression

- Under similar conditions to 1-4, one can establish consistency of OLS for the multiple linear regression model:

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \ldots + \beta_k X_{k,i} + U_i,$$

  where $EU_i = 0$.
- The key assumption is that the errors and regressors are uncorrelated:

$$E(X_{1,i}U_i) = \ldots = E(X_{k,i}U_i) = 0.$$

Omitted variables and the inconsistency of OLS

- Suppose that the true model has two regressors:

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i,$$
$$E\left(X_{1,i} U_i\right) = E\left(X_{2,i} U_i\right) = 0.$$

Omitted variables and the inconsistency of OLS

- Suppose that the true model has two regressors:

$$
\begin{aligned}
Y_i &= \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i, \\
E\left(X_{1,i} U_i\right) &= E\left(X_{2,i} U_i\right) = 0.
\end{aligned}
$$

- Suppose that the econometrician includes only X_1 in the regression when estimating β_1:

Omitted variables and the inconsistency of OLS

- Suppose that the true model has two regressors:

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i, \\ E\left(X_{1,i} U_i\right) &= E\left(X_{2,i} U_i\right) = 0. \end{aligned}$$

- Suppose that the econometrician includes only X_1 in the regression when estimating β_1:

$$\tilde{\beta}_{1,n} = \frac{\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right) Y_i}{\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2}$$

Omitted variables and the inconsistency of OLS

- Suppose that the true model has two regressors:

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i,$$
$$E(X_{1,i} U_i) = E(X_{2,i} U_i) = 0.$$

- Suppose that the econometrician includes only X_1 in the regression when estimating β_1:

$$\begin{aligned}
\tilde{\beta}_{1,n} &= \frac{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_{1,n}\right) Y_i}{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_{1,n}\right)^2} \\
&= \frac{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_{1,n}\right) \left(\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i\right)}{\sum_{i=1}^{n} \left(X_{1,i} - \bar{X}_{1,n}\right)^2}
\end{aligned}$$

Omitted variables and the inconsistency of OLS

- Suppose that the true model has two regressors:

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i,$$
$$E\left(X_{1,i} U_i\right) = E\left(X_{2,i} U_i\right) = 0.$$

- Suppose that the econometrician includes only X_1 in the regression when estimating β_1:

$$\begin{aligned}
\tilde{\beta}_{1,n} &= \frac{\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right) Y_i}{\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2} \\
&= \frac{\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)\left(\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i\right)}{\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2} \\
&= \beta_1 + \beta_2 \frac{\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right) X_{2,i}}{\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2} + \frac{\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right) U_i}{\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2}.
\end{aligned}$$

Omitted variables and the inconsistency of OLS

$$\tilde{\beta}_{1,n} = \beta_1 + \beta_2 \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right) X_{2,i}}{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2} + \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right) U_i}{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2}.$$

Omitted variables and the inconsistency of OLS

$$\tilde{\beta}_{1,n} = \beta_1 + \beta_2 \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)X_{2,i}}{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2} + \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)U_i}{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2}.$$

- As before,

$$\begin{aligned}\frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)U_i}{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2} &= \frac{\frac{1}{n}\sum_{i=1}^{n}X_{1,i}U_i - \bar{X}_{1,n}\bar{U}_n}{\frac{1}{n}\sum_{i=1}^{n}X_{1,i}^2 - \bar{X}_{1,n}^2} \\ &\to_p \frac{0}{EX_{1,i}^2 - \left(EX_{1,i}\right)^2} \\ &= \frac{0}{Var\left(X_{1,i}\right)} = 0.\end{aligned}$$

Omitted variables and the inconsistency of OLS

$$\tilde{\beta}_{1,n} = \beta_1 + \beta_2 \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right) X_{2,i}}{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2} + \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right) U_i}{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2}.$$

Omitted variables and the inconsistency of OLS

$$\tilde{\beta}_{1,n} = \beta_1 + \beta_2 \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right) X_{2,i}}{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2} + \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right) U_i}{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2}.$$

- However,

$$\begin{aligned}\frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right) X_{2,i}}{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2} &= \frac{\frac{1}{n}\sum_{i=1}^{n} X_{1,i}X_{2,i} - \bar{X}_{1,n}\bar{X}_{2,n}}{\frac{1}{n}\sum_{i=1}^{n} X_{1,i}^2 - \bar{X}_{1,n}^2} \\ &\to_p \frac{E\left(X_{1,i}X_{2,i}\right) - \left(EX_{1,i}\right)\left(EX_{2,i}\right)}{EX_{1,i}^2 - \left(EX_{1,i}\right)^2} \\ &= \frac{Cov\left(X_{1,i}, X_{2,i}\right)}{Var\left(X_{1,i}\right)}.\end{aligned}$$

Omitted variables and the inconsistency of OLS

- We have,

$$\begin{aligned}\tilde{\beta}_{1,n} &= \beta_1 + \beta_2 \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right) X_{2,i}}{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2} + \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right) U_i}{\frac{1}{n}\sum_{i=1}^{n}\left(X_{1,i} - \bar{X}_{1,n}\right)^2} \\ &\to_p \beta_1 + \beta_2 \frac{Cov\left(X_{1,i}, X_{2,i}\right)}{Var\left(X_{1,i}\right)} + \frac{0}{Var\left(X_{1,i}\right)} \\ &= \beta_1 + \beta_2 \frac{Cov\left(X_{1,i}, X_{2,i}\right)}{Var\left(X_{1,i}\right)}.\end{aligned}$$

- Thus, $\tilde{\beta}_{1,n}$ is inconsistent unless:
  1. $\beta_2 = 0$ (the model is correctly specified).
  2. $Cov\left(X_{1,i}, X_{2,i}\right) = 0$ (the omitted variable is uncorrelated with the included regressor).

Omitted variables and the inconsistency of OLS

- In this example, the model contains two regressors:

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i,$$
$$E(X_{1,i} U_i) = E(X_{2,i} U_i) = 0.$$

- However, since X_2 is not controlled for, it goes into the error term:

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_{1,i} + V_i, \text{ where} \\ V_i &= \beta_2 X_{2,i} + U_i. \end{aligned}$$

Omitted variables and the inconsistency of OLS

- In this example, the model contains two regressors:

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i,$$
$$E(X_{1,i} U_i) = E(X_{2,i} U_i) = 0.$$

- However, since X_2 is not controlled for, it goes into the error term:

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_{1,i} + V_i, \text{ where} \\ V_i &= \beta_2 X_{2,i} + U_i. \end{aligned}$$

- For consistency of $\tilde{\beta}_{1,n}$ we need $Cov(X_{1,i}, V_i)$ to be equal to zero,

Omitted variables and the inconsistency of OLS

- In this example, the model contains two regressors:

$$
\begin{aligned}
Y_i &= \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i, \\
E\left(X_{1,i} U_i\right) &= E\left(X_{2,i} U_i\right) = 0.
\end{aligned}
$$

- However, since X_2 is not controlled for, it goes into the error term:

$$
\begin{aligned}
Y_i &= \beta_0 + \beta_1 X_{1,i} + V_i, \text{ where} \\
V_i &= \beta_2 X_{2,i} + U_i.
\end{aligned}
$$

- For consistency of $\tilde{\beta}_{1,n}$ we need $Cov\left(X_{1,i}, V_i\right)$ to be equal to zero, however,

$$
Cov\left(X_{1,i}, V_i\right) = Cov\left(X_{1,i}, \beta_2 X_{2,i} + U_i\right)
$$

Omitted variables and the inconsistency of OLS

- In this example, the model contains two regressors:

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i, \\ E(X_{1,i} U_i) &= E(X_{2,i} U_i) = 0. \end{aligned}$$

- However, since X_2 is not controlled for, it goes into the error term:

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_{1,i} + V_i, \text{ where} \\ V_i &= \beta_2 X_{2,i} + U_i. \end{aligned}$$

- For consistency of $\tilde{\beta}_{1,n}$ we need $Cov(X_{1,i}, V_i)$ to be equal to zero, however,

$$\begin{aligned} Cov(X_{1,i}, V_i) &= Cov(X_{1,i}, \beta_2 X_{2,i} + U_i) \\ &= Cov(X_{1,i}, \beta_2 X_{2,i}) + Cov(X_{1,i}, U_i) \end{aligned}$$

Omitted variables and the inconsistency of OLS

- In this example, the model contains two regressors:

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i, \\ E\left(X_{1,i} U_i\right) &= E\left(X_{2,i} U_i\right) = 0. \end{aligned}$$

- However, since X_2 is not controlled for, it goes into the error term:

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_{1,i} + V_i, \text{ where} \\ V_i &= \beta_2 X_{2,i} + U_i. \end{aligned}$$

- For consistency of $\tilde{\beta}_{1,n}$ we need $Cov\left(X_{1,i}, V_i\right)$ to be equal to zero, however,

$$\begin{aligned} Cov\left(X_{1,i}, V_i\right) &= Cov\left(X_{1,i}, \beta_2 X_{2,i} + U_i\right) \\ &= Cov\left(X_{1,i}, \beta_2 X_{2,i}\right) + Cov\left(X_{1,i}, U_i\right) \\ &= \beta_2 Cov\left(X_{1,i}, X_{2,i}\right) + 0 \end{aligned}$$

Omitted variables and the inconsistency of OLS

- In this example, the model contains two regressors:

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + U_i,$$
$$E\left(X_{1,i} U_i\right) = E\left(X_{2,i} U_i\right) = 0.$$

- However, since X_2 is not controlled for, it goes into the error term:

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_{1,i} + V_i, \text{ where} \\ V_i &= \beta_2 X_{2,i} + U_i. \end{aligned}$$

- For consistency of $\tilde{\beta}_{1,n}$ we need $Cov\left(X_{1,i}, V_i\right)$ to be equal to zero, however,

$$\begin{aligned} Cov\left(X_{1,i}, V_i\right) &= Cov\left(X_{1,i}, \beta_2 X_{2,i} + U_i\right) \\ &= Cov\left(X_{1,i}, \beta_2 X_{2,i}\right) + Cov\left(X_{1,i}, U_i\right) \\ &= \beta_2 Cov\left(X_{1,i}, X_{2,i}\right) + 0 \\ &\neq 0, \text{ unless } \beta_2 = 0 \text{ or } Cov\left(X_{1,i}, X_{2,i}\right) = 0. \end{aligned}$$