

Metoda Zmiennych Instrumentalnych

- W dotychczas przedstawionych modelach przyjmowaliśmy założenie o braku korelacji między zmiennymi uwzględnionymi w specyfikacji modelu a składnikiem losowym

- W dotychczas przedstawionych modelach przyjmowaliśmy założenie o braku korelacji między zmiennymi uwzględnionymi w specyfikacji modelu a składnikiem losowym
- Jednak w wielu ważnych z punktu widzenia teorii ekonomicznej zastosowaniach takie założenie nie jest spełnione

- W dotychczas przedstawionych modelach przyjmowaliśmy założenie o braku korelacji między zmiennymi uwzględnionymi w specyfikacji modelu a składnikiem losowym
- Jednak w wielu ważnych z punktu widzenia teorii ekonomicznej zastosowaniach takie założenie nie jest spełnione
- W takim przypadku nie można udowodnić zgodności estymatora MNK

Endogenicność i egzogenicność

- W modelu

$$y = X\beta + \varepsilon$$

Endogenicność i egzogenicność

- W modelu

$$y = X\beta + \varepsilon$$

- zmiennymi **egzogenicznymi** nazywamy zmienne, które nie są skorelowane ze składnikiem losowym

Endogenicność i egzogenicność

- W modelu

$$y = X\beta + \varepsilon$$

- zmiennymi **egzogenicznymi** nazywamy zmienne, które nie są skorelowane ze składnikiem losowym
- zmiennymi **endogenicznymi** nazywamy zmienne, które są skorelowane ze składnikiem losowym

Równoczesność

Równoczesność

O problemie równoczesności mówimy, gdy występuje niezerowa korelacja pomiędzy zmienną objaśniającą x_i a równoczesnym błędem losowym ε_i

Równoczesność

Równoczesność

O problemie równoczesności mówimy, gdy występuje niezerowa korelacja pomiędzy zmienną objaśniającą x_i a równoczesnym błędem losowym ε_i

- Gdy $E(\varepsilon|X) \neq 0$ to

$$E(b) = E((X'X)^{-1}X'y|X) = (X'X)^{-1}X'E(\varepsilon|X)$$

Równoczesność

Równoczesność

O problemie równoczesności mówimy, gdy występuje niezerowa korelacja pomiędzy zmienną objaśniającą x_i a równoczesnym błędem losowym ε_i

- Gdy $E(\varepsilon|X) \neq 0$ to

$$E(b) = E((X'X)^{-1}X'y|X) = (X'X)^{-1}X'E(\varepsilon|X)$$

$$E(b) = \beta + (X'X)^{-1}X'E(\varepsilon|X) \neq \beta$$

Równoczesność

Równoczesność

O problemie równoczesności mówimy, gdy występuje niezerowa korelacja pomiędzy zmienną objaśniającą x_i a równoczesnym błędem losowym ε_i

- Gdy $E(\varepsilon|X) \neq 0$ to

$$E(b) = E((X'X)^{-1}X'y|X) = (X'X)^{-1}X'E(\varepsilon|X)$$

$$E(b) = \beta + (X'X)^{-1}X'E(\varepsilon|X) \neq \beta$$

- Więc estymator wektora parametrów jest obciążony

Przykład (1)

- Model Keynesowski gospodarki zakłada, że

$$\text{PKB} = \text{konsumpcja} + \text{inwestycje} + \text{export netto}$$

Przykład (1)

- Model Keynesowski gospodarki zakłada, że

$$\text{PKB} = \text{konsumpcja} + \text{inwestycje} + \text{export netto}$$

- Z drugiej strony Keynesowska funkcja konsumpcji zakłada, że
 $C_t = f(Y_t)$

Przykład (1)

- Model Keynesowski gospodarki zakłada, że

PKB = konsumpcja + inwestycje + export netto

- Z drugiej strony Keynesowska funkcja konsumpcji zakłada, że
 $C_t = f(Y_t)$
- Szacując jej parametry metodą KMRL

$$C_t = \beta_0 + \beta_1 Y_t + \varepsilon_t$$

Przykład (1)

- Model Keynesowski gospodarki zakłada, że

PKB = konsumpcja + inwestycje + export netto

- Z drugiej strony Keynesowska funkcja konsumpcji zakłada, że
 $C_t = f(Y_t)$
- Szacując jej parametry metodą KMRL

$$C_t = \beta_0 + \beta_1 Y_t + \varepsilon_t$$

- nie są spełnione założenia modelu, gdyż $\text{cov}(X_t, \varepsilon_t) \neq 0$

Przykład (2)

- Szacujemy model autoregresyjny postaci

$$y_t = f(y_{t-1}, y_{t-2}, \dots) + \varepsilon_t \quad (1)$$

Przykład (2)

- Szacujemy model autoregresyjny postaci

$$y_t = f(y_{t-1}, y_{t-2}, \dots) + \varepsilon_t \quad (1)$$

- ale

$$y_{t-1} = f(y_{t-2}, y_{t-3}, \dots) + \varepsilon_{t-1}$$

Przykład (2)

- Szacujemy model autoregresyjny postaci

$$y_t = f(y_{t-1}, y_{t-2}, \dots) + \varepsilon_t \quad (1)$$

- ale

$$y_{t-1} = f(y_{t-2}, y_{t-3}, \dots) + \varepsilon_{t-1}$$

- zatem $cov(y_{t-1}, \varepsilon_{t-1}) \neq 0$. Wobec tego w modelu (1) zmienne objaśniające są skorelowane z błędem losowym

Przykład (3)

- Zazwyczaj w zbiorach danych mikroekonomicznych brakuje informacji o zdolnościach respondentów

Przykład (3)

- Zazwyczaj w zbiorach danych mikroekonomicznych brakuje informacji o zdolnościach respondentów
- Mimo wszystko szacuje się równanie płacy typu Mincera

$$\ln(placa) = \beta_0 + \beta_1 plec + \beta_2 wiek + \beta_3 wiek^2 + \beta_4 wyksz + u$$

Przykład (3)

- Zazwyczaj w zbiorach danych mikroekonomicznych brakuje informacji o zdolnościach respondentów
- Mimo wszystko szacuje się równanie płacy typu Mincera

$$\ln(placa) = \beta_0 + \beta_1 plec + \beta_2 wiek + \beta_3 wiek^2 + \beta_4 wyksz + u$$

- Ponieważ w modelu pominięto zmienną niezależną zdolności to składnik losowy ma postać

$$u = \gamma_0 + \gamma_1 zdolnosc + \phi$$

Przykład (3)

- Z drugiej strony poziom wykształcenia jest determinowany przez zdolności respondenta. Zatem

$$\text{cov}(u, \text{wyksz}) = \text{cov}(\text{zdolności} + \phi, \text{wyksz}) =$$

Przykład (3)

- Z drugiej strony poziom wykształcenia jest determinowany przez zdolności respondenta. Zatem

$$\begin{aligned} \text{cov}(u, \text{wyksz}) &= \text{cov}(\text{zdolności} + \phi, \text{wyksz}) = \\ &= \text{cov}(\text{zdolności}, \text{wyksz}) + \underbrace{\text{cov}(\phi, \text{wyksz})}_{0} \neq 0 \end{aligned}$$

Przykład (3)

- Z drugiej strony poziom wykształcenia jest determinowany przez zdolności respondenta. Zatem

$$\begin{aligned} \text{cov}(u, \text{wyksz}) &= \text{cov}(\text{zdolności} + \phi, \text{wyksz}) = \\ &= \text{cov}(\text{zdolności}, \text{wyksz}) + \underbrace{\text{cov}(\phi, \text{wyksz})}_0 \neq 0 \end{aligned}$$

- Ponieważ zmienna pominięta jest dodatnio skorelowana z uzyskanym wykształceniem i wpływa dodatnio na zmienną zależną to parametr przy zmiennej będzie dodatnio obciążony

- Metoda Zmiennych Instrumentalnych pozwala na uzyskanie zgodnych estymatorów w przypadku występowania korelacji między zmiennymi objaśniającymi a składnikiem losowym

- Metoda Zmiennych Instrumentalnych pozwala na uzyskanie zgodnych estymatorów w przypadku występowania korelacji między zmiennymi objaśniającymi a składnikiem losowym
- Pozwala również uzyskać zgodne oszacowania parametrów w przypadku występowania problemu równoczesności

- Metoda Zmiennych Instrumentalnych pozwala na uzyskanie zgodnych estymatorów w przypadku występowania korelacji między zmiennymi objaśniającymi a składnikiem losowym
- Pozwala również uzyskać zgodne oszacowania parametrów w przypadku występowania problemu równoczesności
- Polega ona na zastąpieniu oryginalnych zmiennych instrumentami

- Metoda Zmiennych Instrumentalnych pozwala na uzyskanie zgodnych estymatorów w przypadku występowania korelacji między zmiennymi objaśniającymi a składnikiem losowym
- Pozwala również uzyskać zgodne oszacowania parametrów w przypadku występowania problemu równoczesności
- Polega ona na zastąpieniu oryginalnych zmiennych instrumentami
- Instrumenty powinny być skorelowane ze zmiennymi objaśniającymi ale nie powinny być skorelowane z błędem losowym

- Metoda Zmiennych Instrumentalnych pozwala na uzyskanie zgodnych estymatorów w przypadku występowania korelacji między zmiennymi objaśniającymi a składnikiem losowym
- Pozwala również uzyskać zgodne oszacowania parametrów w przypadku występowania problemu równoczesności
- Polega ona na zastąpieniu oryginalnych zmiennych instrumentami
- Instrumenty powinny być skorelowane ze zmiennymi objaśniającymi ale nie powinny być skorelowane z błędem losowym
- Znalezienie właściwych instrumentów jest najciekawszym, ale również najtrudniejszym etapem badania

- Oznaczmy przez Z macierz zmiennych instrumentalnych (instrumentów).

- Oznaczmy przez \mathbf{Z} macierz zmiennych instrumentalnych (instrumentów).
- Estymator MZI jest zgodny, gdy spełnione są następujące warunki

$$plim\left(\frac{1}{n}Z'\varepsilon\right) = plim\left(\frac{1}{n}\sum_i z'_i\varepsilon_i\right) = E(z'_i\varepsilon_i) = 0$$

- Oznaczmy przez \mathbf{Z} macierz zmiennych instrumentalnych (instrumentów).
- Estymator MZI jest zgodny, gdy spełnione są następujące warunki

$$plim\left(\frac{1}{n}Z'\varepsilon\right) = plim\left(\frac{1}{n}\sum_i z'_i\varepsilon_i\right) = E(z'_i\varepsilon_i) = 0$$

$$plim\left(\frac{1}{n}Z'X\right) = plim\left(\frac{1}{n}\sum_i z'_ix_i\right) = E(z'_ix_i) \neq 0$$

- Oznaczmy przez \mathbf{Z} macierz zmiennych instrumentalnych (instrumentów).
- Estymator MZI jest zgodny, gdy spełnione są następujące warunki

$$plim\left(\frac{1}{n}Z'\varepsilon\right) = plim\left(\frac{1}{n}\sum_i z'_i\varepsilon_i\right) = E(z'_i\varepsilon_i) = 0$$

$$plim\left(\frac{1}{n}Z'X\right) = plim\left(\frac{1}{n}\sum_i z'_ix_i\right) = E(z'_ix_i) \neq 0$$

- Dodatkowo $r(E(z'_ix_i)) = k$

- Oznaczmy przez \mathbf{Z} macierz zmiennych instrumentalnych (instrumentów).
- Estymator MZI jest zgodny, gdy spełnione są następujące warunki

$$plim\left(\frac{1}{n}Z'\varepsilon\right) = plim\left(\frac{1}{n}\sum_i z'_i\varepsilon_i\right) = E(z'_i\varepsilon_i) = 0$$

$$plim\left(\frac{1}{n}Z'X\right) = plim\left(\frac{1}{n}\sum_i z'_ix_i\right) = E(z'_ix_i) \neq 0$$

- Dodatkowo $r(E(z'_ix_i)) = k$

$$plim\left(\frac{1}{n}Z'Z\right) = plim\left(\frac{1}{n}\sum_i z'_iz_i\right) = E(z'_iz_i) \neq 0$$

- Metoda polega na zastąpieniu oryginalnych wartości zmiennych objaśniających wartościami dopasowanymi uzyskanymi z regresji pomocniczej wykorzystującej zmienne instrumentalne

- Metoda polega na zastąpieniu oryginalnych wartości zmiennych objaśniających wartościami dopasowanymi uzyskanymi z regresji pomocniczej wykorzystującej zmienne instrumentalne
- Macierz instrumentów Z musi zawierać conajmniej tyle zmiennych ile oryginalna macierz X

- Metoda polega na zastąpieniu oryginalnych wartości zmiennych objaśniających wartościami dopasowanymi uzyskanymi z regresji pomocniczej wykorzystującej zmienne instrumentalne
- Macierz instrumentów Z musi zawierać conajmniej tyle zmiennych ile oryginalna macierz X
- Ale nie w każdym przypadku konieczne jest posiadanie k nowych zmiennych

- Metoda polega na zastąpieniu oryginalnych wartości zmiennych objaśniających wartościami dopasowanymi uzyskanymi z regresji pomocniczej wykorzystującej zmienne instrumentalne
- Macierz instrumentów Z musi zawierać conajmniej tyle zmiennych ile oryginalna macierz X
- Ale nie w każdym przypadku konieczne jest posiadanie k nowych zmiennych
- Zmienne z macierzy X , które są nieskorelowane ze składnikiem losowym mogą same dla siebie stanowić instrumenty

- Metoda polega na zastąpieniu oryginalnych wartości zmiennych objaśniających wartościami dopasowanymi uzyskanymi z regresji pomocniczej wykorzystującej zmienne instrumentalne
- Macierz instrumentów Z musi zawierać conajmniej tyle zmiennych ile oryginalna macierz X
- Ale nie w każdym przypadku konieczne jest posiadanie k nowych zmiennych
- Zmienne z macierzy X , które są nieskorelowane ze składnikiem losowym mogą same dla siebie stanowić instrumenty
- W rezultacie potrzeba conajmniej tylu dodatkowych zmiennych instrumentalnych ile jest zmiennych skorelowanych ze składnikiem losowym

- Macierz instrumentów uzyskujemy poprzez rzutowanie wektora X na przestrzeń rozpiętą przez kolumny macierzy instrumentów Z

$$\hat{X} = Z \underbrace{(Z'Z)^{-1}Z'X}_{\beta} = Z\hat{\beta} = P_Z X$$

- Macierz instrumentów uzyskujemy poprzez rzutowanie wektora X na przestrzeń rozpiętą przez kolumny macierzy instrumentów Z

$$\hat{X} = Z \underbrace{(Z'Z)^{-1}Z'X}_{\beta} = Z\hat{\beta} = P_Z X$$

- P_Z jest macierzą rzutu, więc jest symetryczna i idempotentna

- Macierz instrumentów uzyskujemy poprzez rzutowanie wektora X na przestrzeń rozpiętą przez kolumny macierzy instrumentów Z

$$\hat{X} = Z \underbrace{(Z'Z)^{-1}Z'X}_{\beta} = Z\hat{\beta} = P_Z X$$

- P_Z jest macierzą rzutu, więc jest symetryczna i idempotentna
- Dysponując macierzą instrumentów \hat{X} budujemy estymator

$$\beta_{MZI} = (\hat{X}'\hat{X})^{-1}\hat{X}'y = (X'P_Z X)^{-1}X'P_Z y =$$

- Macierz instrumentów uzyskujemy poprzez rzutowanie wektora X na przestrzeń rozpiętą przez kolumny macierzy instrumentów Z

$$\hat{X} = Z \underbrace{(Z'Z)^{-1}Z'X}_{\beta} = Z\hat{\beta} = P_Z X$$

- P_Z jest macierzą rzutu, więc jest symetryczna i idempotentna
- Dysponując macierzą instrumentów \hat{X} budujemy estymator

$$\beta_{MZI} = (\hat{X}'\hat{X})^{-1}\hat{X}'y = (X'P_Z X)^{-1}X'P_Z y =$$

$$\beta_{MZI} = (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y$$

- Macierz instrumentów uzyskujemy poprzez rzutowanie wektora X na przestrzeń rozpiętą przez kolumny macierzy instrumentów Z

$$\hat{X} = Z \underbrace{(Z'Z)^{-1}Z'X}_{\beta} = Z\hat{\beta} = P_Z X$$

- P_Z jest macierzą rzutu, więc jest symetryczna i idempotentna
- Dysponując macierzą instrumentów \hat{X} budujemy estymator

$$\beta_{MZI} = (\hat{X}'\hat{X})^{-1}\hat{X}'y = (X'P_Z X)^{-1}X'P_Z y =$$

$$\beta_{MZI} = (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y$$

- jeżeli $r(Z) = r(X)$, czyli jest tyle nowych zmiennych, ile zmiennych w macierzy X skorelowanych ze składnikiem losowym, to

- Estymator MZI jest również nazywany estymatorem dwustopniowej metody najmniejszych kwadratów

- Estymator MZI jest również nazywany estymatorem dwustopniowej metody najmniejszych kwadratów
- W pierwszym kroku przeprowadzana jest regresja zmiennych endogenicznych na zmienne instrumentalne

- Estymator MZI jest również nazywany estymatorem dwustopniowej metody najmniejszych kwadratów
- W pierwszym kroku przeprowadzana jest regresja zmiennych endogenicznych na zmienne instrumentalne
- W drugim, oryginalne wartości zmiennych są zastępowane przez wartości dopasowane z pierwszego kroku i obliczane są oszacowania poszczególnych parametrów

- Estymator MZI jest również nazywany estymatorem dwustopniowej metody najmniejszych kwadratów
- W pierwszym kroku przeprowadzana jest regresja zmiennych endogenicznych na zmienne instrumentalne
- W drugim, oryginalne wartości zmiennych są zastępowane przez wartości dopasowane z pierwszego kroku i obliczane są oszacowania poszczególnych parametrów
- W kolejnym kroku pokażemy, że przy warunkach zdefiniowanych wcześniej estymator MZI jest zgodny

- Estymator MZI jest dany wzorem

$$\beta_{MZI} = (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y$$

- Estymator MZI jest dany wzorem

$$\beta_{MZI} = (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y$$

- więc

$$plim\beta_{MZI} = plim(X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y =$$

- Estymator MZI jest dany wzorem

$$\beta_{MZI} = (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y$$

- więc

$$plim\beta_{MZI} = plim(X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y =$$

- podstawiając $y = X\beta + \varepsilon$ mamy

$$plim\beta_{MZI} = \beta + plim(X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'\varepsilon =$$

- Na mocy tw. Słuckiego $plim(AB) = plim(A) \cdot plim(B)$, zatem

$$plim\beta_{MZI} = \beta + [plim(X'Z)plim(Z'Z)^{-1}plim(Z'X)]^{-1}$$

- Na mocy tw. Słuckiego $plim(AB) = plim(A) \cdot plim(B)$, zatem

$$plim\beta_{MZI} = \beta + [plim(X'Z)plim(Z'Z)^{-1}plim(Z'X)]^{-1}$$

$$plim(X'Z)plim(Z'Z)^{-1}plim(Z'\varepsilon)$$

- Na mocy tw. Słuckiego $plim(AB) = plim(A) \cdot plim(B)$, zatem

$$plim\beta_{MZI} = \beta + [plim(X'Z)plim(Z'Z)^{-1}plim(Z'X)]^{-1}$$

$$plim(X'Z)plim(Z'Z)^{-1}plim(Z'\varepsilon)$$

- Następnie całe wyrażenie mnożymy i dzielimy przez N

$$plim\beta_{MZI} = \beta + [plim(\frac{X'Z}{N})plim(\frac{Z'Z}{N})^{-1}plim(\frac{Z'X}{N})]^{-1}$$

$$plim(\frac{X'Z}{N})plim(\frac{Z'Z}{N})^{-1}plim(\frac{Z'\varepsilon}{N}) =$$

- korzystając z założeń wiemy, że

$$\text{plim}\beta_{MZI} = \beta + [E(x_i' z_i) E(z_i' z_i)^{-1} E(Z_i' X_i)]^{-1} E(x_i' z_i) E(z_i' z_i)^{-1} \cdot 0 = \beta$$

- korzystając z założeń wiemy, że

$$plim\beta_{MZI} = \beta + [E(x_i'z_i)E(z_i'z_i)^{-1}E(Z_i'X_i)]^{-1}E(x_i'z_i)E(z_i'z_i)^{-1} \cdot 0 = \beta$$

- Zatem estymator MZI dla wektora parametrów jest zgodny

- korzystając z założeń wiemy, że

$$plim\beta_{MZI} = \beta + [E(x_i'z_i)E(z_i'z_i)^{-1}E(Z_i'X_i)]^{-1}E(x_i'z_i)E(z_i'z_i)^{-1} \cdot 0 = \beta$$

- Zatem estymator MZI dla wektora parametrów jest zgodny
- Jeżeli dodatkowo

$$\frac{Z'\varepsilon}{N} \xrightarrow{D} \mathcal{N}(0, \sigma^2 E(z_i'x_i))$$

- korzystając z założeń wiemy, że

$$\text{plim}\beta_{MZI} = \beta + [E(x_i'z_i)E(z_i'z_i)^{-1}E(Z_i'X_i)]^{-1}E(x_i'z_i)E(z_i'z_i)^{-1} \cdot 0 = \beta$$

- Zatem estymator MZI dla wektora parametrów jest zgodny
- Jeżeli dodatkowo

$$\frac{Z'\varepsilon}{N} \xrightarrow{D} \mathcal{N}(0, \sigma^2 E(z_i'x_i))$$

- to rozkład estymatora MZI w granicy jest rozkładem normalnym

$$\sqrt{n}(\beta_{MZI} - \beta) \xrightarrow{D} \mathcal{N}(0, \sigma^2 \{E(z_i'x_i)[E(z_i'z_i)^{-1}]E(z_i'x_i)\}^{-1})$$

- korzystając z założeń wiemy, że

$$\text{plim}\beta_{MZI} = \beta + [E(x'_i z_i)E(z'_i z_i)^{-1}E(Z'_i X_i)]^{-1}E(x'_i z_i)E(z'_i z_i)^{-1} \cdot 0 = \beta$$

- Zatem estymator MZI dla wektora parametrów jest zgodny
- Jeżeli dodatkowo

$$\frac{Z'\varepsilon}{N} \xrightarrow{D} \mathcal{N}(0, \sigma^2 E(z'_i x_i))$$

- to rozkład estymatora MZI w granicy jest rozkładem normalnym

$$\sqrt{n}(\beta_{MZI} - \beta) \xrightarrow{D} \mathcal{N}(0, \sigma^2 \{E(z'_i x_i)[E(z'_i z_i)^{-1}]E(z'_i x_i)\}^{-1})$$

- przy czym zgodnym estymatorem wariancji-kowariancji b_{MZI} jest

$$\Sigma_b = \sigma^2(\hat{X}'\hat{X})^{-1} = \frac{e'e}{n}(\hat{X}'\hat{X})^{-1}$$

- Procedura testowa jest analogiczna to procedury dla estymatora MNK

- Procedura testowa jest analogiczna to procedury dla estymatora MNK
- Weryfikujemy $H_0 : H\beta = h$

- Procedura testowa jest analogiczna to procedury dla estymatora MNK
- Weryfikujemy $H_0 : H\beta = h$
- Statystyka testowa ma postać

$$F = \frac{(Hb_{MZI} - h)'[H(\hat{X}'\hat{X})^{-1}H'](Hb_{MZI} - h)}{\sigma^2} \sim \chi^2_J$$

- Procedura testowa jest analogiczna to procedury dla estymatora MNK
- Weryfikujemy $H_0 : H\beta = h$
- Statystyka testowa ma postać

$$F = \frac{(Hb_{MZI} - h)'[H(\hat{X}'\hat{X})^{-1}H'](Hb_{MZI} - h)}{\sigma^2} \sim \chi^2_J$$

- Można też wykorzystać statystykę w formie LR

$$F = \frac{(e'_R P_Z e_R - e' P_Z e) / J}{\sigma^2} \sim \chi^2_J$$

- Estymator MZI jest zgodny nawet gdy w modelu występują zmienne endogeniczne

- Estymator MZI jest zgodny nawet gdy w modelu występują zmienne endogeniczne
- Ale gdy nie ma takich zmiennych, a spełnione są założenia MNK to estymator MZI nie jest efektywny

Source	SS	df	MS
Model	1.5953e+09	8	199416342
Residual	1.4756e+10	4402	3352112.81
Total	1.6351e+10	4410	3707784.88

Number of obs	=	4411
F(8, 4402)	=	59.49
Prob > F	=	0.0000
R-squared	=	0.0976
Adj R-squared	=	0.0959
Root MSE	=	1830.9

placa	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
plec	159.7862	56.22739	2.84	0.005	49.55225	270.0202
wiek	-26.36377	67.93703	-0.39	0.698	-159.5545	106.827
wiek2	.3387365	.8445655	0.40	0.688	-1.317037	1.99451
_Iwykszt_2	-1006.02	84.69471	-11.88	0.000	-1172.064	-839.9757
_Iwykszt_3	-1155.616	127.8124	-9.04	0.000	-1406.193	-905.0398
_Iwykszt_4	-1415.621	79.40485	-17.83	0.000	-1571.295	-1259.948
_Iwykszt_5	-1809.905	108.9463	-16.61	0.000	-2023.494	-1596.315
wies	-121.6751	58.52882	-2.08	0.038	-236.4211	-6.929203
_cons	4327.832	1336.369	3.24	0.001	1707.876	6947.788

Instrumental variables (2SLS) regression

Number of obs = 3951
 Wald chi2(8) = 117.19
 Prob > chi2 = 0.0000
 R-squared = .
 Root MSE = 2321.5

placa	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
_Iwykszt_2	1103.416	1213.748	0.91	0.363	-1275.485 3482.318
_Iwykszt_3	-22.80231	2161.186	-0.01	0.992	-4258.65 4213.045
_Iwykszt_4	-2815.861	561.0983	-5.02	0.000	-3915.593 -1716.128
_Iwykszt_5	-777.9333	1288.779	-0.60	0.546	-3303.893 1748.026
plec	459.6071	157.8633	2.91	0.004	150.2007 769.0134
wiek	12.92527	98.07904	0.13	0.895	-179.3061 205.1567
wiek2	-.1063004	1.209343	-0.09	0.930	-2.47657 2.263969
wies	189.0222	113.2531	1.67	0.095	-32.9499 410.9942
_cons	3094.369	2012.577	1.54	0.124	-850.2097 7038.947

Instrumented: _Iwykszt_2 _Iwykszt_3 _Iwykszt_4 _Iwykszt_5
 Instruments: plec wiek wiek2 wies _Ipaeduc_2 _Ipaeduc_3 _Ipaeduc_4
 _Ipaeduc_5 _Ipaeduc_6 _Ipaeduc_7 _Ipaeduc_8

Variable	mzi	reg
_Iwykszt_2	1103.4163	-1006.0199***
_Iwykszt_3	-22.802307	-1155.6164***
_Iwykszt_4	-2815.8608***	-1415.6213***
_Iwykszt_5	-777.93328	-1809.9049***
plec	459.60705**	159.78622**
wiek	12.925274	-26.363765
wiek2	-.10630039	.33873649
wies	189.02217	-121.67513*
_cons	3094.3686	4327.8323**

legend: * p<0.05; ** p<0.01; *** p<0.001

- MZI można traktować jako uogólnienie MNK

- MZI można traktować jako uogólnienie MNK
- Wobec tego wszystkie testy stosowane przy MNK mogą być stosowane przy MZI

- MZI można traktować jako uogólnienie MNK
- Wobec tego wszystkie testy stosowane przy MNK mogą być stosowane przy MZI
- Testy specyficzne dla MZI weryfikują założenie o egzogeniczności zmiennych, oraz poprawności wykorzystanych instrumentów

- Przy estymacji MZI test Hausmana jest testem na egzogeniczność zmiennych

- Przy estymacji MZI test Hausmana jest testem na egzogeniczność zmiennych
- Estymator MZI jest estymatorem zawsze zgodnym

- Przy estymacji MZI test Hausmana jest testem na egzogeniczność zmiennych
- Estymator MZI jest estymatorem zawsze zgodnym
- Przy prawdziwej H_0 o egzogeniczności zmiennych X oba estymatory są zgodne, ale estymator MZI ma większą wariancję niż estymator MNK

- Przy estymacji MZI test Hausmana jest testem na egzogeniczność zmiennych
- Estymator MZI jest estymatorem zawsze zgodnym
- Przy prawdziwej H_0 o egzogeniczności zmiennych X oba estymatory są zgodne, ale estymator MZI ma większą wariancję niż estymator MNK
- Przy fałszywej H_0 estymator MZI jest zgodny i efektywny, a estymator MNK nie jest zgodny

- Przy estymacji MZI test Hausmana jest testem na egzogeniczność zmiennych
- Estymator MZI jest estymatorem zawsze zgodnym
- Przy prawdziwej H_0 o egzogeniczności zmiennych X oba estymatory są zgodne, ale estymator MZI ma większą wariancję niż estymator MNK
- Przy fałszywej H_0 estymator MZI jest zgodny i efektywny, a estymator MNK nie jest zgodny
- Statystyka testowa jest dana przez formę kwadratową

$$(\beta_{MZI} - \beta_{MNK})' \Sigma_{\beta_{MZI} - \beta_{MNK}}^{-1} (\beta_{MZI} - \beta_{MNK}) \xrightarrow{D} \chi^2(r(\Sigma))$$

- Przy estymacji MZI test Hausmana jest testem na egzogeniczność zmiennych
- Estymator MZI jest estymatorem zawsze zgodnym
- Przy prawdziwej H_0 o egzogeniczności zmiennych X oba estymatory są zgodne, ale estymator MZI ma większą wariancję niż estymator MNK
- Przy fałszywej H_0 estymator MZI jest zgodny i efektywny, a estymator MNK nie jest zgodny
- Statystyka testowa jest dana przez formę kwadratową

$$(\beta_{MZI} - \beta_{MNK})' \Sigma_{\beta_{MZI} - \beta_{MNK}}^{-1} (\beta_{MZI} - \beta_{MNK}) \xrightarrow{D} \chi^2(r(\Sigma))$$

- Gdy różnica jest duża sugeruje to wykorzystanie MNK

	---- Coefficients ----			
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	mzi	reg	Difference	S.E.
_Iwykszt_2	1103.416	-1006.02	2109.436	1210.789
_Iwykszt_3	-22.80231	-1155.616	1132.814	2157.404
_Iwykszt_4	-2815.861	-1415.621	-1400.24	555.4513
_Iwykszt_5	-777.9333	-1809.905	1031.972	1284.165
plec	459.6071	159.7862	299.8208	147.5103
wiek	12.92527	-26.36377	39.28904	70.73937
wiek2	-.1063004	.3387365	-.4450369	.8655751
wies	189.0222	-121.6751	310.6973	96.95695

b = consistent under H_0 and H_a ; obtained from ivregress
 B = inconsistent under H_a , efficient under H_0 ; obtained from regress

Test: H_0 : difference in coefficients not systematic

$$\begin{aligned} \text{chi2}(7) &= (b-B)' [(V_b-V_B)^{-1}] (b-B) \\ &= 15.54 \\ \text{Prob}>\text{chi2} &= 0.0297 \end{aligned}$$

- Test Sargana weryfikuje poprawność instrumentów

- Test Sargana weryfikuje poprawność instrumentów
- Zauważmy, że reszty modelu są równe

$$e = (I - X(X'P_zX)^{-1}X'P_z)\varepsilon$$

- Test Sargana weryfikuje poprawność instrumentów
- Zauważmy, że reszty modelu są równe

$$e = (I - X(X'P_zX)^{-1}X'P_z)\varepsilon$$

- Mnożąc obie strony przez P_z z lewej uzyskujemy

$$P_z e = (P_z - P_z X(X'P_zX)^{-1}X'P_z)\varepsilon = M_z \varepsilon$$

- Test Sargana weryfikuje poprawność instrumentów
- Zauważmy, że reszty modelu są równe

$$e = (I - X(X'P_zX)^{-1}X'P_z)\varepsilon$$

- Mnożąc obie strony przez P_z z lewej uzyskujemy

$$P_z e = (P_z - P_z X(X'P_zX)^{-1}X'P_z)\varepsilon = M_z \varepsilon$$

- Oczywiście M_z jest macierzą idempotentną, rzędu $p - k$

- Test Sargana weryfikuje poprawność instrumentów
- Zauważmy, że reszty modelu są równe

$$e = (I - X(X'P_zX)^{-1}X'P_z)\varepsilon$$

- Mnożąc obie strony przez P_z z lewej uzyskujemy

$$P_z e = (P_z - P_z X(X'P_zX)^{-1}X'P_z)\varepsilon = M_z \varepsilon$$

- Oczywiście M_z jest macierzą idempotentną, rzędu $p - k$
- Przy prawdziwej H_0 jest brak korelacji instrumentów z błędami losowymi

$$\frac{e'P_z e}{\sigma^2} \sim \chi_{p-k}^2$$

- Test Sargana weryfikuje poprawność instrumentów
- Zauważmy, że reszty modelu są równe

$$e = (I - X(X'P_ZX)^{-1}X'P_Z)\varepsilon$$

- Mnożąc obie strony przez P_Z z lewej uzyskujemy

$$P_Ze = (P_Z - P_ZX(X'P_ZX)^{-1}X'P_Z)\varepsilon = M_Z\varepsilon$$

- Oczywiście M_Z jest macierzą idempotentną, rzędu $p - k$
- Przy prawdziwej H_0 jest brak korelacji instrumentów z błędami losowymi

$$\frac{eP_Ze}{\sigma^2} \sim \chi_{p-k}^2$$

- Niestety przeprowadzenie testu jest wyłącznie możliwe, gdy liczba instrumentów przekracza liczbę zmiennych objaśniających

```
. estat overid
```

Tests of overidentifying restrictions:

Sargan (score) $\chi^2(3) = .933392$ (p = 0.8174)

Basman $\chi^2(3) = .930777$ (p = 0.8180)

- Brak jest podstaw do odrzucenia hipotezy zerowej o poprawności instrumentów