

6089 ECONOMETRIA

A.A. 2009 – 2010

Lezione Info2

Obiettivo di questa lezione applicata è studiare come le caratteristiche di una comunità locale influiscono sul prezzo delle case vendute/acquistate. Per farlo, analizziamo i dati su 506 comunità locali nell'area metropolitana di Boston. Alcuni dati provengono dal mercato immobiliare (prezzo mediano delle abitazioni, numero medio delle stanze); altri descrivono le caratteristiche di ogni comunità (qualità dell'aria, vicinanza a servizi importanti, qualità dell'offerta scolastica, tasse sugli immobili). La nostra variabile dipendente è il prezzo mediano delle abitazioni. Vogliamo capire se (e come) le altre variabili influenzano i prezzi. I dati sono nel file *hprice.dta*.

1. Operazioni preliminari

- Create la directory "C:\econometria\" e copiatevi il file *hprice.dta*.
- Nella finestra di comando di Stata digitate "**cd c:\econometria**": da questo momento questa diventa la directory di riferimento per le operazioni che svolgeremo con Stata.
- Create un file in cui saranno contenuti tutti i passaggi svolti durante la lezione digitando "**log using info2.log, replace**".

2. Un primo sguardo ai dati e analisi di regressione

- Aprirete il file *vote.dta* utilizzando File → Open, oppure scrivendo nella riga di comando: "**use hprice,clear**".
- Il comando "**describe**" ("**des**") fornisce una descrizione del contenuto del dataset (numero di osservazioni, tipo e descrizione (*label*) delle variabili), mentre il comando "**summarize**" ("**sum**") fornisce alcune statistiche descrittive delle variabili.
- Come primo passo per l'analisi di regressione, trasformiamo alcune variabili continue in logaritmiche, per esempio il prezzo mediano (la nostra variabile dipendente): **gen lprice=log(price)**. Facciamo lo stesso in modo da creare le variabili "ldist" e lnox".
- Stimate $lprice_i = \beta_1 + \beta_2 \ln ox_i + \beta_3 ldist_i + \beta_4 rooms_i + \beta_5 stratio_i + \varepsilon_i$, usando il comando OLS "**reg**".
 - Commentate il segno e la dimensione dei coefficienti stimati.
 - Commentate i valori dell' R^2 e del test F.
 - In media, un aumento di una stanza che impatto ha (in percentuale) sul prezzo mediano?

3. Test RESET sulla forma funzionale

- Chiediamoci adesso se la specificazione del modello stimato al punto precedente sia la migliore che potevamo introdurre. In particolare, la forma funzionale è correttamente specificata? Per rispondere, applichiamo il test RESET con il comando: **estat ovtest**. Che cosa ci dice il risultato del test?

- b) Stimiamo, allora, un modello di regressione in cui aggiungiamo un termine quadratico del numero di stanze (“**gen rooms2=rooms^2**”) e il logaritmo delle tasse sugli immobili (“**gen lproptax=log(proptax)**”) come regressori.
- c) Eseguiamo il test RESET sul nuovo modello. Che cosa ci dice il risultato del test? Quale modello preferiremo?
- d) Sapendo che STATA implementa il test RESET inserendo i valori fittati di “lprice” alla seconda, alla terza e alla quarta nel modello completo, si replichino i risultati del test senza usare il comando “**ovtest**”.
- e) Nel modello stimato al punto b) si commentino congiuntamente i coefficienti di “rooms” e “rooms2”. In particolare:
 - Qual è l’effetto marginale di una stanza in più per un valore medio (tra comunità) di stanze?
 - L’impatto del numero medio di stanze sul prezzo mediano cambia di segno a un certo livello di stanze? Quale? È rilevante?

4. Eteroschedasticità e OLS pesati

- a) Stimiamo di nuovo: **reg lprice lnox ldist rooms rooms2 stratio lproptax**.
- b) Creiamo la serie dei residui della regressione (**predict resid, r**) e quella dei valori fittati (**predict lprice_hat**). Ottenete il grafico dei residui sui valori fittati: **twoway (scatter resid lprice_hat)**. Cosa suggerisce?
- c) Adesso, vogliamo testare formalmente la presenza del problema sopra evidenziato: eteroschedasticità nei termini di errore. Applichiamo quindi il test di Breusch-Pagan per forme di eteroschedasticità dipendenti dalla scala delle osservazioni; l’ipotesi nulla coincide con l’assunzione di omoschedasticità: $H_0 : Var(u | X) = \sigma_u^2$. Usiamo il comando “**estat hettest**”. In particolare, si possono usare i valori fittati della regressione precedente (“**estat hettest,iid**”) o l’insieme dei regressori (“**estat hettest,iid rhs**”) per implementare il test. Che cosa ci dicono i risultati di queste due versioni del test di Breusch-Pagan?
- d) Per affrontare il problema di eteroschedasticità, possiamo fare due cose:
 - Ristimare il modello con una versione aggiustata dello stimatore della matrice di varianza e covarianza: **reg lprice lnox ldist rooms rooms2 stratio lproptax,vce(robust)**. (Si commentino queste stime rispetto a quelle precedenti).
 - Applicare una stima FGLS con i minimi quadrati pesati, per esempio assumendo che la scala del problema di eteroschedasticità sia catturata dal numero medio di stanze (“rooms”): **reg lprice lnox ldist rooms rooms2 stratio lproptax [aw=1/rooms2]**. (Si commentino queste stime rispetto a quelle precedenti).

5. Operazioni di chiusura

Salvate il file dei dati con le nuove variabili costruite e chiudete il log file: **save hprice_new.dta,replace; log close**.