

VALORE DELLE ABITAZIONI: DISPARITÀ REGIONALI

Erika GHIRALDO¹, Gennaro CHIROLA², Caterina ANDREUSSI³

SOMMARIO

La spesa per l'abitazione costituisce una voce rilevante del reddito delle famiglie; è quindi fondamentale verificare l'esistenza e valutare l'entità delle differenze geografiche tra i prezzi.

Test classici che fondano la loro costruzione su assunti distributivi di normalità e omoschedasticità appaiono strumenti inadeguati con i dati a disposizione. Per verificare la significatività statistica delle differenze riscontrate, l'analisi è condotta implementando un test di natura non parametrica che si basa sulla permutazione delle osservazioni all'interno del campione.

L'analisi viene svolta utilizzando dati a livello comunale, forniti dall'Osservatorio del Mercato Immobiliare dell' Agenzia del Territorio. L'esame è approfondito considerando sia l'intera base dati, sia sottoinsiemi di comuni che possono risultare omogenei rispetto a variabili di interesse.

¹ Agenzia del Territorio, Largo Leopardi 5, 00185, Roma, e-mail: erika.ghiraldo@agenziaterritorio.it.

² Agenzia del Territorio, Largo Leopardi 5, 00185, Roma, e-mail: gennaro.chirola@agenziaterritorio.it.

³ Agenzia del Territorio, Largo Leopardi 5, 00185, Roma, e-mail: caterina.andreussi@agenziaterritorio.it.

1 Introduzione

Quando si parla di disparità residenziale, generalmente si fa riferimento alle differenze che esistono nella qualità delle abitazioni (Sen, 2004). Seguendo la definizione di povertà, espressa dallo stesso Sen (Sen, 2000), come "la privazione delle capacità di base" la disuguaglianza residenziale è particolarmente rilevante in tutte le discussioni sulla povertà.

La qualità degli alloggi è un fattore che determina se condizioni di vita desiderabili sono immediatamente disponibili. Se si ipotizza che il prezzo⁴ medio delle abitazioni sia correlato con il tenore dei servizi o in altri termini sia rappresentativo di quelle condizioni di vita desiderabili, a cui si è fatto cenno, indagare le differenze di prezzo vuol dire esprimere una misura della disuguaglianza.

A tal fine risulta indispensabile, come in molti altri settori dell'economia, tenere elevato il livello di trasparenza del mercato immobiliare. Negli ultimi anni l'Osservatorio del Mercato Immobiliare⁵ (OMI) dell'Agenzia del Territorio è divenuto un importante punto di riferimento per gli operatori del mercato immobiliare e per quanti hanno interesse a queste informazioni. I dati resi disponibili dall'OMI sono divenuti una fonte ufficiale sempre più completa, puntuale e più facilmente comparabile. In questo senso si prestano a molteplici analisi e studi che permettono di migliorare l'interpretazione dei fenomeni territoriali. In particolare l'OMI pubblica dati sulle quotazioni immobiliari a livello di aggregazione territoriale sub-comunale. Con l'obiettivo di evidenziare le differenze che si riscontrano tra le regioni nei prezzi delle abitazioni questo lavoro propone uno studio prendendo a riferimento le quotazioni immobiliari rese disponibili dall'OMI. Il dettaglio minimo di riferimento è rappresentato dalla zona OMI⁶ che rappresenta il punto di partenza per la formazione delle quotazioni e per ognuna è calcolato un intervallo di quotazioni, a partire da indagini puntuali sul territorio.

E' bene premettere che le regioni sono entità amministrative che non necessariamente corrispondono ad aree omogenee del mercato immobiliare residenziale. Tuttavia, questa distinzione artificiale è di notevole interesse in quanto importante dal punto di vista politico. E', infatti, alla ripartizione regionale che il legislatore rivolge sempre più spesso la sua

⁴ Nel seguito dello studio i termini prezzo, valore e quotazione saranno utilizzati indistintamente come sinonimi.

⁵ L'Agenzia del Territorio è stata istituita a seguito della riforma del Ministero dell'Economia e delle Finanze, ed è operativa dal 1 gennaio 2001. E' un ente pubblico dotato di personalità giuridica e ampia autonomia regolamentare, amministrativa, patrimoniale, organizzativa, contabile e finanziaria. Tra i suoi compiti istituzionali vi è la gestione dell'Osservatorio del Mercato Immobiliare che cura la rilevazione e l'elaborazione delle informazioni di carattere tecnico-economico relative ai valori immobiliari, al mercato degli affitti ed ai tassi di rendita, nonché la pubblicazione e la diffusione di studi ed elaborazioni e la valorizzazione statistica degli archivi dell'Agenzia del Territorio.

⁶ Può essere definita come "una porzione continua del territorio comunale che riflette un comparto omogeneo del mercato immobiliare locale, nel quale si registra uniformità di apprezzamento per condizioni economiche e socio-ambientali" (Agenzia del Territorio, 2008).

attenzione per attuare leve di politica economica. Inoltre, la distinzione regionale permette studi di confronto con le altre regioni dell'Unione Europea.

L'articolo presenta spunti di originalità per il pregio della base dati che viene considerata e per la tecnica con cui è indagata la variabilità tra le ripartizioni territoriali. I dati di partenza, riferendosi a porzioni minime del territorio sono molto dettagliati e puntuali, mentre le differenze tra le medie sono testate attraverso una procedura non parametrica che supera i limiti posti dal rispetto degli assunti richiesti nei modelli classici.

Nel secondo paragrafo sarà presentata la base dati, nel terzo è spiegata la procedura statistica non parametrica che è utilizzata ai fini dell'analisi. I risultati dei test statistici applicati per testare la significatività delle differenze tra le quotazioni medie delle regioni sono discussi nel quarto paragrafo. Nel quinto si riportano gli esiti di un'indagine mirata a valutare possibili similitudini nei prezzi medi delle regioni. Infine l'ultimo paragrafo dell'elaborato è dedicato alle conclusioni evidenziando criticità e possibili ulteriori sviluppi futuri.

2 Presentazione della base dati

Nel 2010, nel mercato delle abitazioni, sono state rilevate circa 80.000 schede relative a prezzi di compravendita, offerte o valori⁷ riferiti alle zone OMI di circa 1.300 comuni scelti tra quelli con maggiore dinamica di mercato. Per ogni comune si è calcolata la quotazione media sulla base delle quotazioni relative a tutte le tipologie immobiliari. Le medie così ottenute sono state utilizzate per calcolare la quotazione media comunale che rappresenta la base dati di partenza di questa elaborazione. Si tratta, quindi, di una cross-section di 8.096 quotazioni medie del settore residenziale riferite a ciascun comune per l'anno 2010.

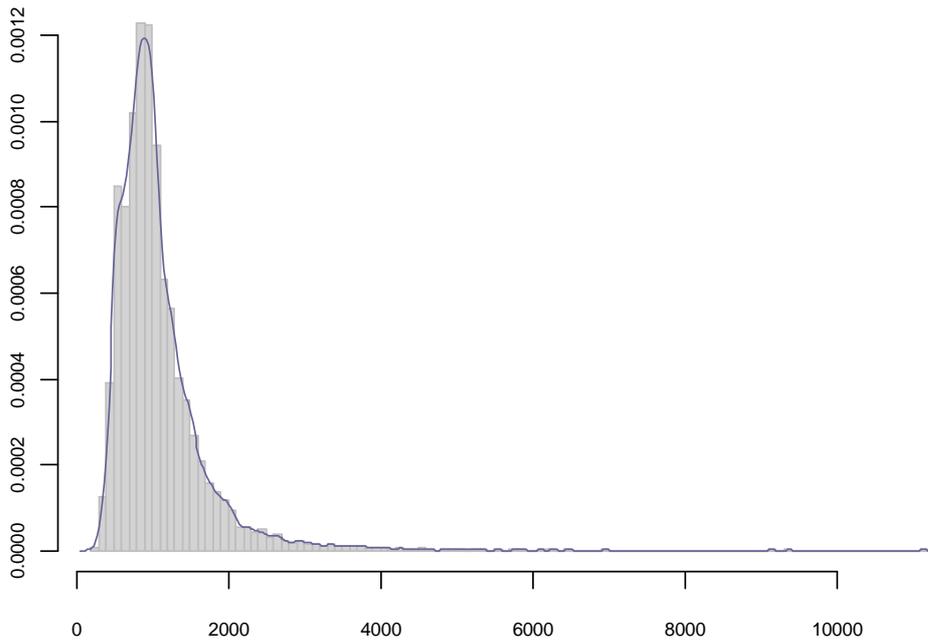
La distribuzione complessiva delle quotazioni, rappresentata attraverso l'istogramma e la curva della densità non parametrica stimata (Figura I), mostra una coda destra accentuata e valori molto concentrati intorno ai 1500 €/mq. Le quotazioni assumono valore minimo a 224 €/mq e massimo pari a 11.130 €/mq. La mediana risulta uguale a 952 €/mq mentre la media a 1.079 €/mq.

⁷ Per la formazione delle quotazioni immobiliari l'OMI svolge semestralmente un'indagine del mercato immobiliare avvalendosi di diverse fonti tra cui le principali sono rappresentate da:

- offerte provenienti da agenti immobiliari. A supporto del processo di rilevazione sono stati sottoscritti rapporti di collaborazioni con le maggiori associazioni di agenti immobiliari FIAIP, FIMAA, AICI;
- prezzi dedotti dagli atti di compravendita (il decreto-legge 4 luglio 2006, n. 223 rafforzando il comma 497 della finanziaria 2006 obbliga le parti contraenti ad indicare nell'atto il corrispettivo pattuito);
- valori ottenuti da stime immobiliari.

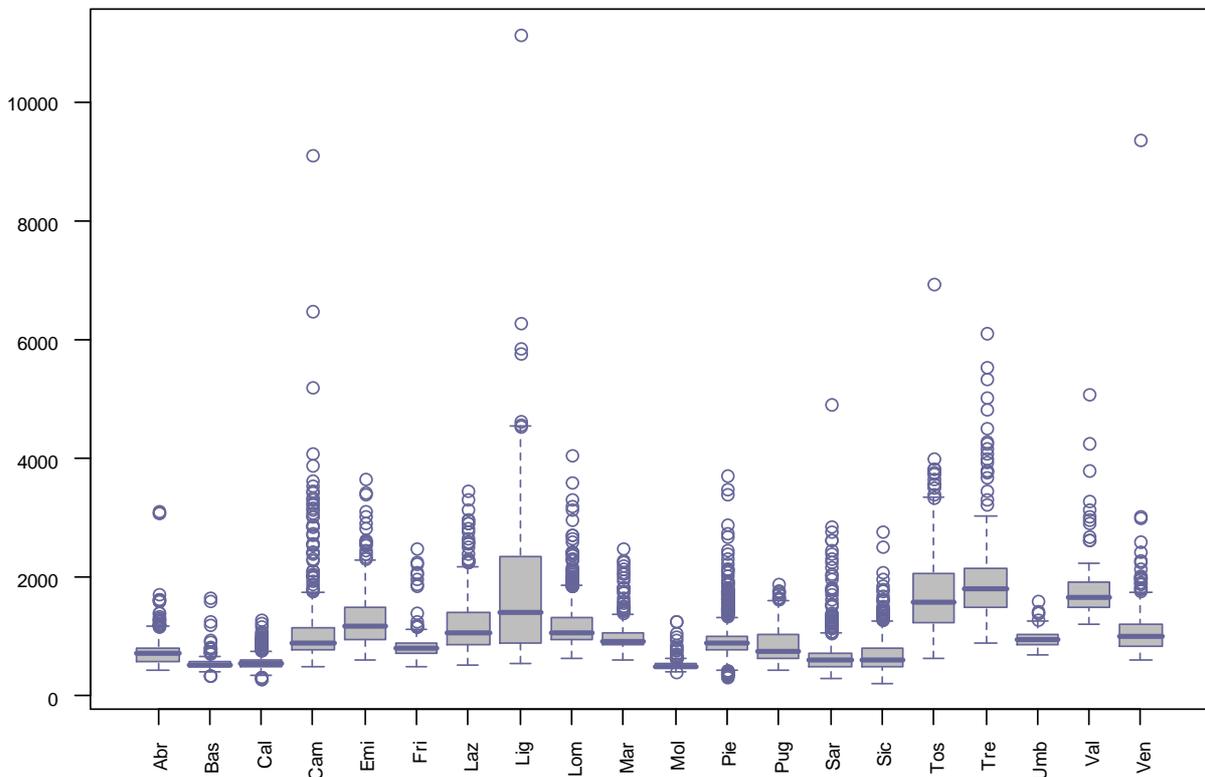
Per un approfondimento sulla formazione delle quotazioni immobiliari si può consultare il Manuale della banca dati dell'osservatorio del mercato immobiliare (Agenzia del Territorio, 2008).

Figura I: Distribuzione delle quotazioni medie comunali



Entrando nel dettaglio regionale la Figura II mostra la distribuzione delle quotazioni di ciascuna regione. La rappresentazione grafica evidenzia forti differenze in termini di mediane, di campi di variazione e di presenza di valori estremi. Pur rappresentando un primo approccio, questo grafico rappresenta il punto di partenza dello studio in quanto fornisce una misura, seppure solo visiva, dell'entità delle differenze che si riscontrano tra i raggruppamenti territoriali considerati.

Figura II: Boxplot della distribuzione dei valori comunali di ciascuna regione

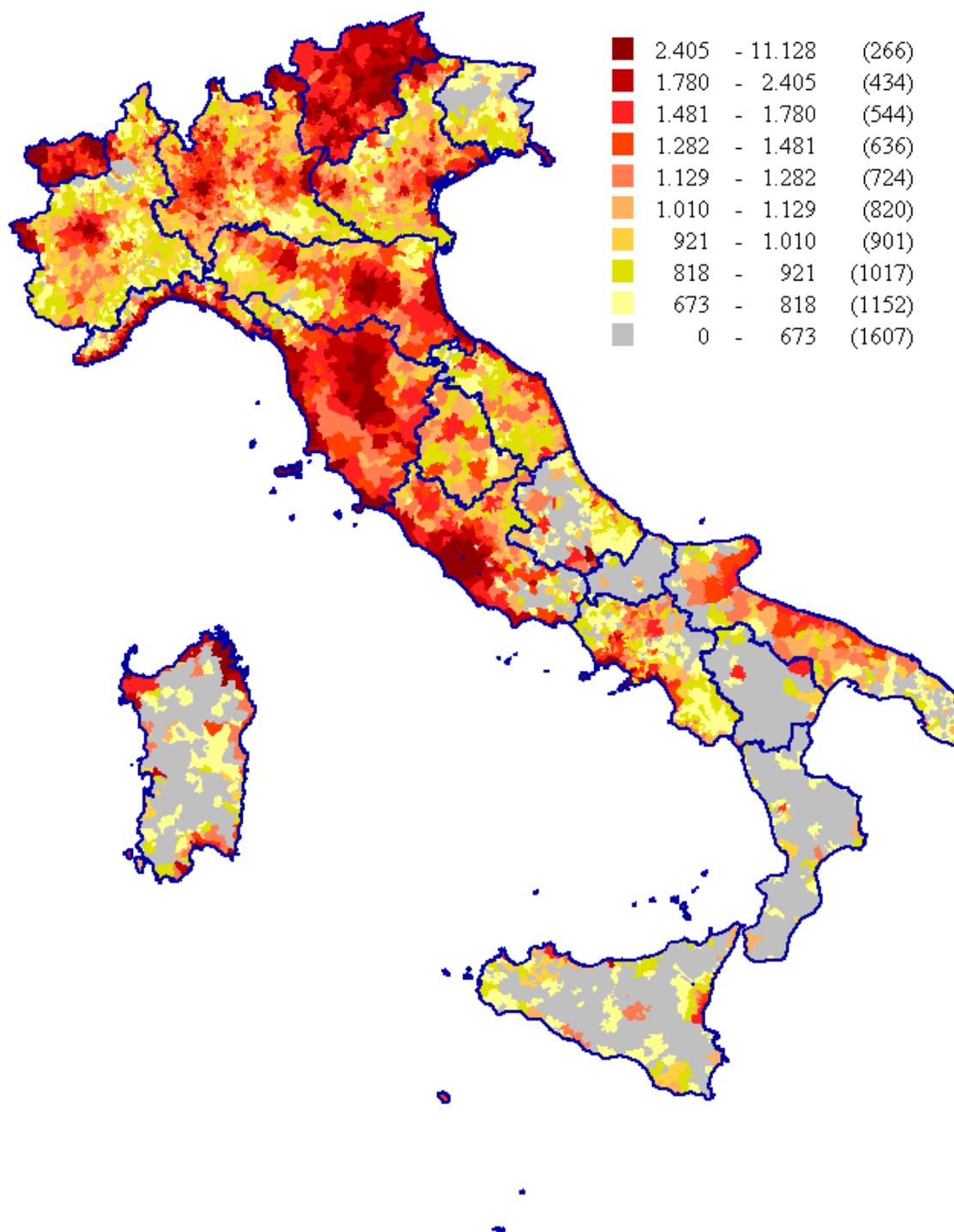


Le quotazioni medie regionali e alcune statistiche di sintesi sono mostrate in Tabella I. E' evidente già in questa fase che i prezzi medi sono molto diversi tra le regioni si passa, infatti, dal minimo del Molise, 554 € al metro quadro per arrivare a 1.954 € al metro quadro del Trentino Alto Adige. Anche la deviazione standard delle quotazioni per ciascuna regione risulta piuttosto elevata con un picco per la regione Liguria dove è pari a 1.310. Ciò è imputabile al fatto che per tale regione il range dei valori assunti risulta piuttosto ampio essendo pari a 553 €/m² la quotazione minima e a 11.128 €/m² la quotazione massima.

All'interno di ciascuna regione si trovano comuni molto diversi tra loro con la conseguenza che anche le quotazioni medie che in essi si rilevano risultano molto eterogenee. Un approfondimento delle distribuzioni delle quotazioni interne a ciascuna regione è fornito dalla mappa dei prezzi medi comunali (Figura III). Appare evidente una diversità tra regioni del Sud e del Nord, tuttavia, una lettura più attenta rileva differenze tra le stesse regioni del Sud e le stesse regioni del Nord. Al Sud la Puglia palesa una distribuzione delle quotazioni diversa dalle confinanti Basilicata e Molise, così come la Campania che esibisce sfumature più simili a qualche regione del Centro piuttosto che del Sud. Mentre al Nord il Friuli Venezia Giulia mostra maggiore similarità con qualche regione del Sud piuttosto che del Nord.

Questa prima analisi esplorativa evidenzia come i comuni presentino caratteristiche eterogenee che rendono difficile la definizione di un modello unico in grado di ricomprendere e spiegare tutte le relazioni di causa ed effetto che concorrono nella formazione dei prezzi. I comuni con prezzi medi molto superiori rispetto al resto della regione sono probabilmente turistici. Tuttavia, all'interno della distribuzione si trovano altri comuni turistici che non hanno quotazioni così elevate. Inoltre, a differenza di quanto ci si può attendere, non tutti i capoluoghi di provincia hanno valori che spiccano rispetto ad altri comuni minori.

Figura III: Mappa delle quotazioni medie comunali 2010



3 Procedura di test

Dal punto di vista statistico è importante verificare se le differenze osservate a livello empirico siano significative. Il confronto tra due o più gruppi viene solitamente effettuato attraverso la tecnica inferenziale nota come analisi della varianza (ANOVA).

L'ANOVA, attraverso la scomposizione della varianza nelle componenti within e between, permette infatti se il gruppo di appartenenza delle osservazioni influisce sul valore medio delle stesse.

Si ipotizzi il seguente modello per i prezzi rilevati

$$p_{ij} = \mu_j + \varepsilon_{ij}$$

dove p_{ij} indica la quotazione media dell' i -esimo comune, che fa parte di una generica j -esima regione, μ_j è la quotazione media regionale e ε_{ij} sono gli errori. Il sistema di ipotesi alla base del test ha come ipotesi nulla l'uguaglianza delle medie μ_j , mentre l'alternativa prevede che almeno una di esse sia diversa. Le ipotesi alla base della procedura ANOVA sono:

- indipendenza delle osservazioni: la quotazione media di un comune non influenza le quotazioni degli altri;
- gli errori sono normalmente distribuiti con media pari a zero, e quindi le osservazioni sono normalmente distribuite intorno alla media regionale;
- le varianze all'interno dei diversi gruppi sono omogenee.

Nel nostro caso come evidente dal grafico dei boxplot (Figura II) confermano l'implausibilità di questi assunti: sarebbe complicato giustificare l'assunto di indipendenza; inoltre, le osservazioni sembrano molto lontane dall'essere normalmente distribuite con la stessa varianza.

Tuttavia a maggior conferma le ipotesi di normalità e omoschedasticità sono state testate utilizzando test formali. L'ipotesi di normalità è stata testata in ciascun gruppo attraverso il test di Jarque e Bera e i p -value conducono sempre al rifiuto dell'ipotesi di normalità. L'ipotesi di omogeneità delle varianze è testata applicando il test Fligner-Killeen (William & all., 1981) che è un test non parametrico molto robusto anche quando i gruppi non seguono una distribuzione normale. Il test ha restituito un p -value pressoché nullo e quindi si rifiuta l'ipotesi di omogeneità delle varianze.

Al fine di testare lo stesso sistema di ipotesi dell'ANOVA, si applica un metodo non parametrico che non richiede particolari assunti circa la distribuzione dei dati.

Si è scelto di applicare un test di permutazione che nel seguito si indica sinteticamente come *permanova* del tipo suggerito da Anderson (2001).

Questo test condivide molte similarità con il suo omologo parametrico, ad esempio è possibile mantenere la stessa specificazione della statistica test, anche se potenzialmente è possibile usare statistiche basate su diverse definizioni di distanza. La differenza principale risiede nel fatto che la distribuzione di tale statistica non è standard, e viene approssimata tramite un metodo di simulazione.

Questa simulazione è disegnata per verificare se il valore della statistica test dipende dalla particolare assegnazione dei comuni alle diverse regioni o se piuttosto, immaginando una serie di assegnazioni casuale, tale statistica assumerebbe valori significativamente diversi.

La procedura è iterativa e prevede, ad ogni iterazione, l'estrazione casuale di variabili di assegnazione dei singoli comuni alle diverse regioni. Sulla base della nuova assegnazione si calcola il nuovo valore della statistica test.

Alla fine delle iterazioni, avremo un campione di valori della statistica test che possono essere messi a paragone con il valore originale, cioè quello ottenuto con l'assegnazione reale. Ad esempio, il p -value associato al test viene approssimato come la frequenza dei valori simulati che eccedono il valore osservato.

L'idea è quella di creare un set di dati basati sul "rimescolamento" di quelli realmente osservati, calcolare la statistica di riferimento e confrontarla con una statistica calcolata sul campione originale. Più in particolare si seguono i seguenti passi:

- si definisce un'ipotesi nulla H_0 ;
- si definisce una statistica test;
- si crea un set di dati basati sul "rimescolamento" di quelli realmente osservati (la modalità di "rimescolamento" viene definita in funzione dell'ipotesi nulla);
- si ricalcola la statistica di riferimento e si compara il valore con quello osservato. Si ripetono gli ultimi due punti molte volte (es. 100.000 volte);
- se la statistica osservata è maggiore del limite ottenuto nel 95% dei casi basati sul "rimescolamento", si rigetta H_0 .

Nel caso in esame si calcola la statistica F sul campione originale, si spostano casualmente le osservazioni tra i diversi gruppi, si ricalcola la statistica sulla nuova base e si ripete la procedura per un numero elevato di volte.

La distribuzione del test ottenuta al termine del ciclo non è più una distribuzione nota i cui quantili sono tabulati ma è una distribuzione che possiamo confrontare con il valore della statistica ottenuto sulla base originale.

4 Analisi dei risultati

Le differenze regionali nelle quotazioni medie sono molto evidenti, e qualsiasi tecnica statistica applicata a questo tipo di dati confermerà l'esistenza di questo fenomeno. Considerando i dati relativi a tutte le regioni, qualsiasi tipo di verifica di ipotesi rifiuterà l'ipotesi nulla di uguaglianza delle medie. Tale conclusione appare corretta, anche alla luce della sintesi fornita dal grafico di Figura II e dalla mappa di Figura III.

L'importanza della correttezza della tecnica risulta più evidente per sottoinsiemi di dati in cui le differenze tra le quotazioni medie sono meno evidenti. In tali sottoinsiemi, i differenti insiemi di assunti possono condurre a conclusioni diverse.

L'esistenza di differenze significative nelle quotazioni medie delle regioni è indagata, come detto, attraverso la tecnica del *permanova* applicata oltre che all'intero campione anche a diversi sottoinsiemi della base dati complessiva di partenza. Si vuole, infatti, indagare se le

diversità riscontrate a livello empirico e confermate dal test sull'intero campione siano imputabili ad una o più variabili che influenzano le quotazioni. Più in particolare si può ritenere che le differenze siano dovute alle eterogeneità nei comuni di ciascuna regione. Lo studio è condotto cercando di confrontare le regioni considerando gruppi di comuni simili secondo una qualche variabile di interesse.

Nel seguito sono discussi in dettaglio i risultati analitici per il caso del campione complessivo mentre i dettagli per tutte le altre ipotesi sono riportati in appendice.

Utilizzando la base dati complessiva si ottiene un valore della statistica F pari a 210,67. Un test ANOVA condurrebbe ad un netto rifiuto, come detto però tale analisi non può considerarsi valida in questo contesto in quanto le ipotesi del modello non sono rispettate.

Si procede quindi con il *permanova*, impostando 50.000 iterazioni si ottiene il campione della distribuzione della statistica test F rappresentato in Figura IV. Si può notare che il valore originario della statistica test è molto distante dall'intervallo dei valori simulati conducendo al rifiuto dell'ipotesi nulla.

Le differenze di quotazione visualizzate nella Tabella I sono quindi estremamente significative.

4.1 Analisi per classi di prezzo

Dopo aver analizzato l'intera distribuzione dei prezzi e aver ottenuto differenze significative al fine di approfondire l'indagine si procede con la creazione di una variabile che dipende solo dal prezzo. Un metodo automatico per definire una stratificazione è quello di effettuare una cluster analysis in modo tale da ottenere gruppi di quotazioni medie comunali omogenei sui quali effettuare separatamente il test.

La partizione della variabile quotazioni è stata ottenuta attraverso un algoritmo *k-medie*. Il numero dei cluster fornito all'algoritmo è pari a sette ed è stato scelto in base al criterio di informazione BIC (Schwarz, 1978) e confermato dall'*elbow method* (Thorndike, 1953).

Le distribuzioni per regione dei sette gruppi ottenuti sono visualizzate in Figura V. E' evidente che non si è fatto altro che dividere l'intervallo delle quotazioni in sottointervalli, nell'ipotesi che il confronto tra prezzi medi appartenenti ad un range omogeneo evidenzia qualche similitudine tra le regioni. Applicando separatamente il test a ciascun gruppo si ottengono sempre *p-value* che portano al rifiuto dell'ipotesi nulla e quindi alla conclusione che le differenze regionali sono significative. Se ne deduce che considerare solo comuni con quotazioni che cadono in uno stesso intervallo non rende simili le distribuzioni regionali.

4.2 *Analisi per classi di popolazione*

Nel paragrafo precedente si è cercata una variabile che definisse un gruppo di comuni simili sulla base di uno stesso livello di prezzi medi giungendo a conclusione che permangono significative differenze tra le regioni italiane. La variabile prezzo sembra non essere sufficientemente informativa pertanto potrebbe essere utile condurre l'analisi considerando ulteriori variabili potenzialmente influenti nella formazione delle quotazioni. Marginalizzando rispetto a tali fattori i divari riscontrati tra le regioni potrebbero attenuarsi.

Un primo naturale raggruppamento che a priori si può ritenere discriminante nella determinazione dei prezzi è la grandezza del comune misurata in termini di popolazione residente. In comuni di dimensioni simili si dovrebbe assistere ad un andamento dei prezzi comparabile. Considerando quindi all'interno di ciascuna regione solo comuni di classi dimensionali analoghe le quotazioni dovrebbero essere indipendenti dalla regione di appartenenza con la conseguenza che le medie dovrebbero risultare non significativamente diverse. Sulla base della popolazione residente nel 2008 (dati ISTAT) si sono formati tre gruppi di comuni: quelli con popolazione inferiore ai 5.000 abitanti, quelli con popolazione compresa tra i 5.000 e i 25.000 e quelli con popolazione superiore ai 25.000. In Tabella II si riportano in dettaglio il numero di comuni per ciascuna regione, le quotazioni medie e i differenziali di quotazione rispetto al livello medio nazionale per le tre classi di comuni.

Considerando separatamente i tre campioni il test ha evidenziato in tutti i casi differenze significative tra le regioni. Comuni appartenenti ad una stessa classe demografica forniscono distribuzioni regionali dei prezzi medi diverse. Se si torna alla mappa di Figura III si può considerare che pur prendendo per ciascuna regione sottoinsiemi di comuni, ad esempio, solo i comuni con meno di 5.000 abitanti le differenze cromatiche continuano a sussistere. Per chiarire, nessun sottogruppo di comuni dalla regione Trentino Alto Adige è analogo ad un qualsiasi gruppo di comuni della regione Calabria. Conferma si ottiene dalle mappe riportate in Figura VII che evidenziano le distribuzioni delle quotazioni per ogni classe di popolazione. In questo contesto, è di interesse riportare quanto emerso da un'analisi effettuata sulla posizione relativa di ciascuna regione rispetto alle quotazioni medie. Calcolando per ciascuna regione le quotazioni medie e assegnandone il rango⁸ discendente, si ottiene una graduatoria ordinata delle regioni. Tale graduatoria è stata calcolata sia sulla base dati originale sia nei tre gruppi di comuni. Sebbene le quotazioni medie di ciascuna regione siano diverse nei quattro casi il rango di ogni regione nel caso generale e nei casi di comuni con meno di 5.000 abitanti, con abitanti tra 5.000 e 25.000 e con più di 25.000 abitanti non subisce variazioni tali da rendere meno significative le differenze tra le ripartizioni regionali. Quanto detto è evidenziato nei grafici di Figura VI che mostrano la posizione di ciascuna regione

⁸ Il rango della variabile quotazione media regionale è stato calcolato come l'intero corrispondente alla posizione che la variabile occupa quando si passa dal campione (X_1, X_2, \dots, X_n) al campione ordinato in senso decrescente $(X_{(n)}, X_{(n-1)}, \dots, X_{(1)})$.

considerando il campione complessivo, posizione iniziale, e la posizione con le diverse ripartizioni dei comuni, posizione finale. In particolare, in ascissa si riporta il rango nella graduatoria iniziale della regione mentre in ordinata si riporta il rango che la regione occupa calcolando le quotazioni medie sul sottoinsieme di comuni. Quanto più i punti sono vicini alla bisettrice tanto più la posizione relativa iniziale e la posizione nel raggruppamento esaminato sono simili. L'analisi mostra che soprattutto nel caso di comuni piccoli, con meno di 5.000 abitanti, le differenze riscontrate sono minime. Sono, infatti, probabilmente i comuni minori a contribuire maggiormente alla formazione dei prezzi medi essendo anche il gruppo più numeroso. Piccoli scostamenti di posizione sono evidenziati tra i comuni di grandi dimensioni. Spicca la situazione della Basilicata che vede dimezzata la sua posizione in graduatoria in considerazione del fatto che i comuni della regione con più di 25.000 abitanti coincidono con i capoluoghi di provincia. Questa analisi dimostra, rafforzando quanto già emerso, un'indipendenza delle quotazioni medie regionali dal fattore classe di popolazione del comune.

Al termine del paragrafo si può concludere che anche considerando la variabile popolazione le differenze riscontrate non sono di entità tale da modificare l'assetto di fondo che si evidenzia nel caso più generale dell'intera base dati.

4.3 Analisi per soli capoluoghi di provincia

Nella realtà coesistono dimensioni urbane diverse a cui corrispondono anche funzioni economiche differenti. In genere si rileva che i centri maggiori accolgono funzioni più rare e specializzate in cui compaiono più rilevanti economie di scala (Camagni, 1993).

I centri più grandi coincidono per la maggior parte delle regioni con i capoluoghi di provincia che rappresentano un polo di attrazione rispetto alle città minori. E', quindi, di interesse verificare se considerando solo i capoluoghi di provincia si riscontrano o meno differenze tra le quotazioni medie delle regioni.

Applicando il test al campione composto dai soli capoluoghi di provincia le differenze tra regioni continuano ad essere significative.

Il numero di capoluoghi, le quotazioni medie e il differenziale rispetto al valore medio nazionale sono riportati in Tabella III. Anche per il sottoinsieme di comuni composto dai soli capoluoghi si è effettuata l'analisi dei ranghi rappresentata nell'ultimo grafico di Figura VI. E' evidente che anche in tal caso il ranking delle regioni rimane sostanzialmente invariato rispetto al campione originale. Piccole variazioni si notano per la regione Basilicata che guadagna posizioni anche se in misura inferiore rispetto al caso dei comuni con più di 25.000 abitanti. Tale risultato va imputato al fatto che i comuni grandi delle altre regioni non comprendono i soli capoluoghi. A guadagnare posizione è anche la regione Sardegna evidenziando la maggiore influenza dei suoi capoluoghi di provincia. Di contro la regione

Valle d'Aosta perde posizioni. Per quest'ultima va considerato che nella graduatoria originale beneficia del fatto di essere formata da una sola provincia.

I livelli medi delle quotazioni delle regioni sono significativamente diversi anche considerando la variabile capoluogo.

4.4 *Analisi del campione senza valori estremi*

Le distribuzioni delle quotazioni di ciascuna regione mostrano dei valori estremi, essenzialmente anche se non esclusivamente dovuti alla presenza di comuni turistici, nei quali si rilevano quotazioni molto distanti dalla media di regione. E' ragionevole ritenere che questi valori, influenzando la media, determinino differenze tra i raggruppamenti regionali che altrimenti non emergerebbero. In altri termini, si vuole verificare se si ottengono risultati diversi escludendo dall'analisi i valori che risultano estremi rispetto al resto dei valori della propria regione.

I valori estremi sono stati individuati usando la *boxplot rule* (Iglewicz & Hoaglin, 1993), ovvero considerando estreme, per ogni regione, le osservazioni che eccedono il valore dato da $Q_3 + 1,5 \cdot (Q_3 - Q_1)$, con Q_i che indica l' i -esimo quartile.

Si tratta di 521 comuni le cui quotazioni medie sono riportate in Tabella IV.

Il test, effettuato sul campione depurato di queste osservazioni, ha evidenziato comunque significative differenze tra le regioni. Anche eliminando dalle rispettive regioni quelle città che presentano valori molto elevati, le differenze regionali permangono, a significare che non sono solo i valori estremi ad influenzare l'analisi. Solo a titolo di esempio è possibile affermare che i risultati non cambiano se si toglie Portofino alla Liguria o Capri alla Campania.

4.5 *Analisi delle quotazioni medie pesate*

La media aritmetica semplice si mostra spesso inadeguata a rappresentare realtà ampie e complesse quali quelle immobiliari. E' utile attribuire ad ogni osservazione un peso specifico che ne aumenta o diminuisce l'incidenza nel calcolo al fine di fornire una statistica di sintesi più affidabile. Buone variabili di ponderazione per le quotazioni immobiliari sono rappresentate dallo stock⁹ di abitazioni e dalla superficie¹⁰ abitativa.

⁹ Lo stock immobiliare è il numero di unità immobiliari censite negli archivi catastali, distinto per tipologia catastale. Ai fini di questo elaborato sono considerate solo le abitazioni ovvero le unità censite nelle categorie da A1 ad A11 escluse le A10 (per un approfondimento si può consultare la pubblicazione "Statistiche Catastali", 2009).

¹⁰ Le superfici dello stock di abitazioni sono state stimate sulla base dei vani catastali e della superficie del vano medio comunale.

In questo paragrafo si commentano confronti tra le quotazioni medie regionali calcolate come medie semplici e come medie ponderate utilizzando come variabili lo stock e le superfici residenziali nel 2010.

Le quotazioni medie e i differenziali rispetto alla quotazione media nazionale¹¹ si riportano in Tabella V.

In questo caso non si è applicata la procedura di test *permanova* ma si è svolta solo l'analisi dei ranghi che comunque permette di concludere che la graduatoria tra le regioni non subisce variazioni significative al variare della procedura di calcolo e di ponderazione. In Figura VIII si riportano i confronti tra le posizioni delle regioni nel caso di medie calcolate come semplici medie aritmetiche delle quotazioni e calcolate come medie ponderate per lo stock e per la superficie.

5 Similitudini

Dopo avere testato l'esistenza di significative differenze tra le quotazioni medie delle regioni sia sul complesso dei comuni che per sottogruppi scelti in relazione ad alcune variabili ritenute potenzialmente discriminanti si vogliono analizzare le similarità tra regioni.

Dalla mappa di Figura III risulta evidente indagare le differenze tra tutte le regioni simultaneamente, in qualunque modo siano presi i comuni all'interno di ciascuna regione, conduce sempre alla conclusione che le diversità osservate sono significative. Se, tuttavia, si vuole effettuare un'indagine mirata a valutare anche le similarità si possono considerare sottogruppi di regioni.

Partendo da un'ispezione grafica della mappa (Figura III) e dei boxplot (Figura I) si sono effettuati diversi confronti tra le distribuzioni delle quotazioni medie comunali delle regioni. Effettuando test a coppie tra le regioni con la tecnica del *permanova* si sono ottenuti risultati in cui non si è potuto rifiutare l'ipotesi nulla e quindi si è concluso per un'uguaglianza tra le quotazioni medie delle regioni in esame. In Tabella VI si riportano i risultati dei test effettuati con i rispettivi livelli di significatività.

L'esame ha confermato per alcune regioni quanto ci si aspettava: è il caso della Basilicata, della Calabria e del Molise per i quali il test ha mostrato che non risulta una differenza significativa tra le quotazioni medie. E' probabile che questo gruppo di regioni presenti molte caratteristiche comuni, quali condizioni geografiche o economiche che ripercuotendosi sul mercato immobiliare determinano dinamiche simili. In questo caso si è testata anche l'ipotesi sull'uguaglianza congiunta delle medie di tutte e tre le regioni che è risultata non significativamente diversa. Attesa è anche la non significativa diversità tra Liguria e Toscana.

¹¹ In questo caso non sono stati considerati i comuni dove il catasto è gestito localmente. Sebbene, infatti, per tali comuni siano disponibili le quotazioni immobiliari non si dispone del dato relativo allo stock di unità abitative. Si tratta in particolare di tutti i comuni della regione Trentino Alto Adige, dei comuni delle province di Gorizia e Trieste e di due comuni della provincia di Belluno.

Le due regioni vantano entrambe la presenza di rinomate località turistiche. Trentino Alto Adige e Valle d'Aosta risultano simili forse perché accomunate dalla vocazione montana e dal noto stato di regione autonoma. Di facile previsione è anche la similarità emersa tra le due isole Sardegna e Sicilia. Sebbene non molto vicine geograficamente, comportamenti analoghi delle quotazioni si riscontrano per Lazio e Lombardia. Ad avvicinarle è probabilmente la presenza in entrambe di due importanti città quali Roma e Milano che rappresentano due rilevanti poli attrattivi. Si può affermare che l'influenza di queste metropoli va ben oltre i comuni confinanti interessando sicuramente la gran parte dei comuni della regione.

Meno facili da interpretare appaiono le similitudini tra alcune regioni che l'analisi ha evidenziato. Le quotazioni medie del Friuli Venezia Giulia appaiono simili a quelle della Puglia, mentre non c'è evidenza di differenza tra Piemonte e Umbria. Più articolata la situazione della Campania che mostra analogie con le Marche ma non con l'Umbria pur essendo questa simile alle Marche. Infine, La regione Emilia Romagna non mostra similarità con nessun'altra regione.

6 Conclusioni

Dai test effettuati le differenze nei livelli dei prezzi medi delle abitazioni tra regioni sono risultate statisticamente significative. Se consideriamo che "la ricchezza abitativa incorporata nella casa in proprietà e la sua trasmissione rappresentano un elemento importante di strutturazione delle disuguaglianze sociali" come evidenziato in Poggio (2009) la ripartizione territoriale appare una variabile discriminante nell'influenzare la formazione dei prezzi medi. Le diversità indagate, a parità di altri fattori, possono essere intese sia in termini di maggiore o minore valore dell'abitazione di proprietà sia in termini di maggiore o minore possibilità di accesso. Se si ipotizza che le condizioni abitative generali siano incorporate nei livelli delle quotazioni medie allora la qualità della vita degli individui è influenzata dalla collocazione regionale.

L'indagine condotta rappresenta un aspetto importante delle analisi sulle disparità, tema sul quale non si è raggiunto un livello esaustivo. Occorre tenere conto, comunque, del fatto che lo studio si pone come un primo approccio alla disamina del più ampio fenomeno delle divergenze. Si tratta, infatti, di uno dei molteplici aspetti che andrebbero esaminati all'interno del filone degli studi inerenti la disuguaglianza residenziale. Oltre a prendere in considerazione i diversi livelli delle quotazioni si dovrebbero esaminare anche variabili che colgono aspetti sociali, demografici ed economici che possono interagire nella formazione delle quotazioni. Tuttavia, l'esame congiunto del valore delle abitazioni e del PIL mostra netti divari tra le regioni come evidenziato nel lavoro sull'analisi del patrimonio (Agenzia del Territorio e Dipartimento delle Finanze, 2010).

Vale comunque la pena di ricordare che secondo recenti dati ISTAT nel 2008 il 68,5% del totale delle famiglie vive in case di proprietà cui si aggiunge un ulteriore 11,5% di famiglie che vivono in abitazioni in usufrutto o in uso gratuito. Il modello dell'aspirazione alla proprietà è stato, infatti, in Italia storicamente incentivato e considerato come un indicatore del benessere (Poggio, 2009).

Inoltre, nonostante benefici o svantaggi che si possono avere in una regione rispetto ad un'altra come evidenziato in uno studio dell'Agenzia del Territorio (Rapporto Immobiliare, 2005) per circa il 57% degli acquisti di abitazioni effettuati in un anno vi è coincidenza tra comune di ubicazione dell'immobile e comune di residenza del soggetto acquirente e solo per il 18% l'acquisto riguarda un immobile di una regione diversa rispetto a quella di residenza dell'acquirente.

In conclusione, grazie alla disponibilità di informazioni dettagliate, il lavoro può rappresentare un utile strumento per migliorare la conoscenza del fenomeno della disegualianza residenziale affidando a successivi sviluppi ulteriori approfondimenti.

7 Appendice

7.1 Tabelle dei risultati

Tabella I: Dati di sintesi per regione

<i>Regione</i>	<i>N. comuni</i>	<i>Media</i>	<i>Deviazione Standard</i>	<i>Differenziale rispetto alla media Italia</i>	<i>Minimo</i>	<i>Massimo</i>
<i>Abruzzo</i>	305	784	328	0,73	452	3.115
<i>Basilicata</i>	131	578	185	0,54	349	1.671
<i>Calabria</i>	409	576	145	0,53	295	1.283
<i>Campania</i>	550	1.105	709	1,02	500	9.127
<i>Emilia Romagna</i>	341	1.289	489	1,19	609	3.671
<i>Friuli Venezia Giulia</i>	219	850	297	0,79	500	2.500
<i>Lazio</i>	376	1.199	549	1,11	525	3.463
<i>Liguria</i>	235	1.827	1.310	1,69	553	11.128
<i>Lombardia</i>	1.546	1.164	328	1,08	636	4.050
<i>Marche</i>	246	1.029	317	0,95	605	2.491
<i>Molise</i>	136	554	140	0,51	398	1.280
<i>Piemonte</i>	1.209	949	317	0,88	338	3.730
<i>Puglia</i>	257	861	318	0,80	431	1.890
<i>Sardegna</i>	373	733	466	0,68	303	4.919
<i>Sicilia</i>	390	713	335	0,66	224	2.781
<i>Toscana</i>	287	1.753	752	1,62	640	6.957
<i>Trentino Alto Adige</i>	339	1.954	730	1,81	888	6.110
<i>Umbria</i>	92	980	163	0,91	711	1.604
<i>Valle d'Aosta</i>	74	1.884	707	1,75	1.200	5.080
<i>Veneto</i>	581	1.085	460	1,00	604	9.380
<i>Italia</i>	8.096	1.079	593	1,00	224	11.128

Figura IV: Istogramma del campione della distribuzione della statistica test F ottenuta dal permanova

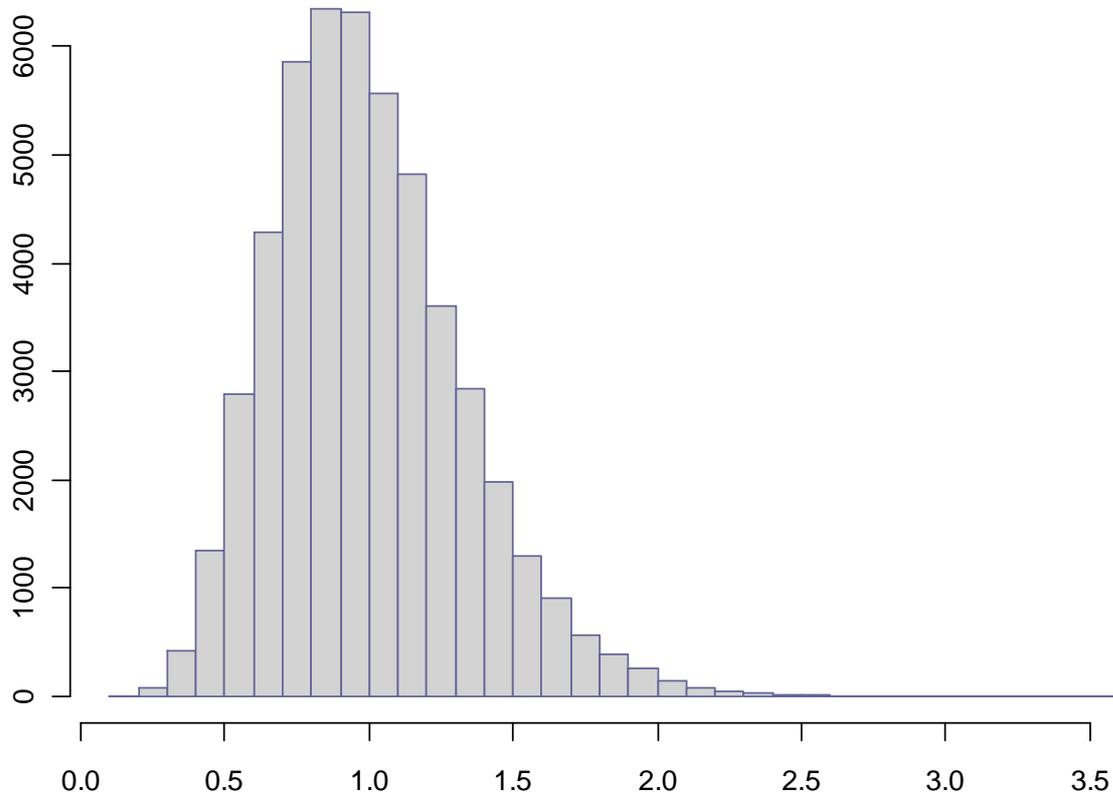


Tabella II: Dati di sintesi per classi di popolazione dei comuni

Classe di popolazione	Campione complessivo				< 5.000 ab.				5.000 25.000 ab.				> 25.000 ab.			
	Differenziale rispetto alla media Italia		Differenziale rispetto alla media Italia		Differenziale rispetto alla media Italia		Differenziale rispetto alla media Italia		Differenziale rispetto alla media Italia		Differenziale rispetto alla media Italia		Differenziale rispetto alla media Italia			
Regione	N. comuni	Media	N. comuni	Media	N. comuni	Media	N. comuni	Media	N. comuni	Media	N. comuni	Media	N. comuni	Media		
Abruzzo	305	784	0,73	251	741	0,76	45	932	0,74	9	1.224	0,75				
Basilicata	131	578	0,54	99	520	0,53	30	697	0,55	2	1.639	1,01				
Calabria	409	576	0,53	326	542	0,55	74	677	0,53	9	962	0,59				
Campania	550	1.105	1,02	334	921	0,94	166	1.319	1,04	50	1.624	1,00				
Emilia Romagna	341	1.289	1,19	153	1.049	1,07	164	1.408	1,11	24	2.005	1,24				
Friuli Venezia Giulia	219	850	0,79	158	766	0,78	56	1.038	0,82	5	1.411	0,87				
Lazio	376	1.199	1,11	253	984	1,01	95	1.525	1,20	28	2.039	1,26				
Liguria	235	1.827	1,69	183	1.524	1,56	44	2.937	2,32	8	2.656	1,64				
Lombardia	1.546	1.164	1,08	1093	1.088	1,11	406	1.309	1,03	47	1.682	1,04				
Marche	246	1.029	0,95	178	919	0,94	54	1.206	0,95	14	1.746	1,08				
Molise	136	554	0,51	124	529	0,54	10	725	0,57	2	1.277	0,79				
Piemonte	1.209	949	0,88	1075	904	0,93	110	1.250	0,99	24	1.558	0,96				
Puglia	257	861	0,80	84	694	0,71	132	852	0,67	41	1.234	0,76				
Sardegna	373	733	0,68	309	634	0,65	53	1.129	0,89	11	1.614	0,99				
Sicilia	390	713	0,66	198	600	0,61	149	793	0,63	43	956	0,59				
Toscana	287	1.753	1,62	135	1.465	1,50	124	1.927	1,52	28	2.371	1,46				
Trentino Alto Adige	339	1.954	1,81	307	1.921	1,97	28	2.180	1,72	4	2.884	1,78				
Umbria	92	980	0,91	60	898	0,92	25	1.084	0,86	7	1.305	0,80				
Valle d'Aosta	74	1.884	1,75	73	1.883	1,93	==	==	==	1	1.923	1,19				
Veneto	581	1.085	1,00	314	953	0,98	243	1.203	0,95	24	1.601	0,99				
Italia	8.096	1.079	1,00	5707	976	1,00	2.008	1.268	1,00	381	1.622	1,00				

Figura V: Boxplot delle quotazioni per ciascun cluster

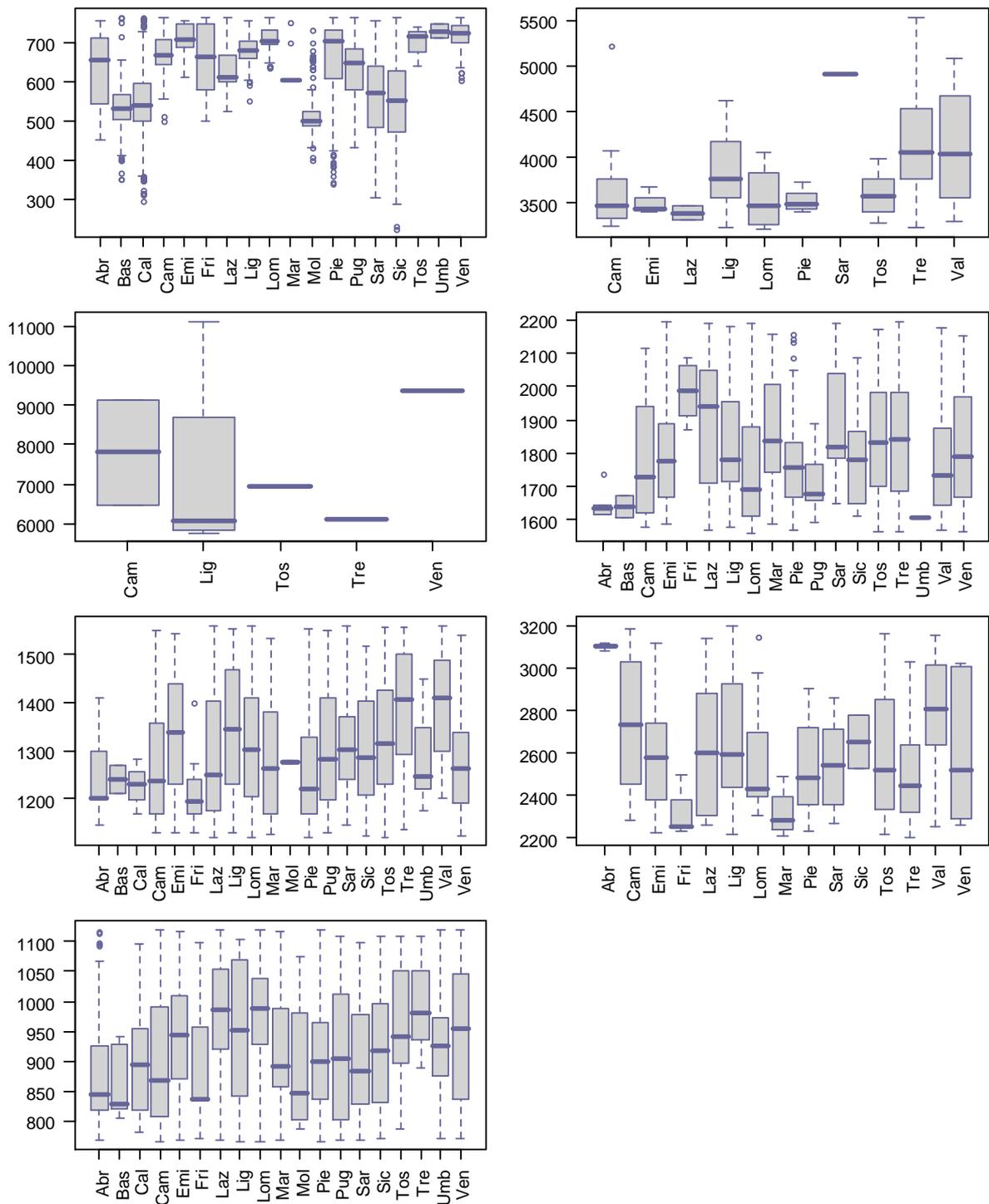


Figura VI: Ranghi delle quotazioni delle regioni per classi di popolazione e capoluoghi rispetto al campione complessivo

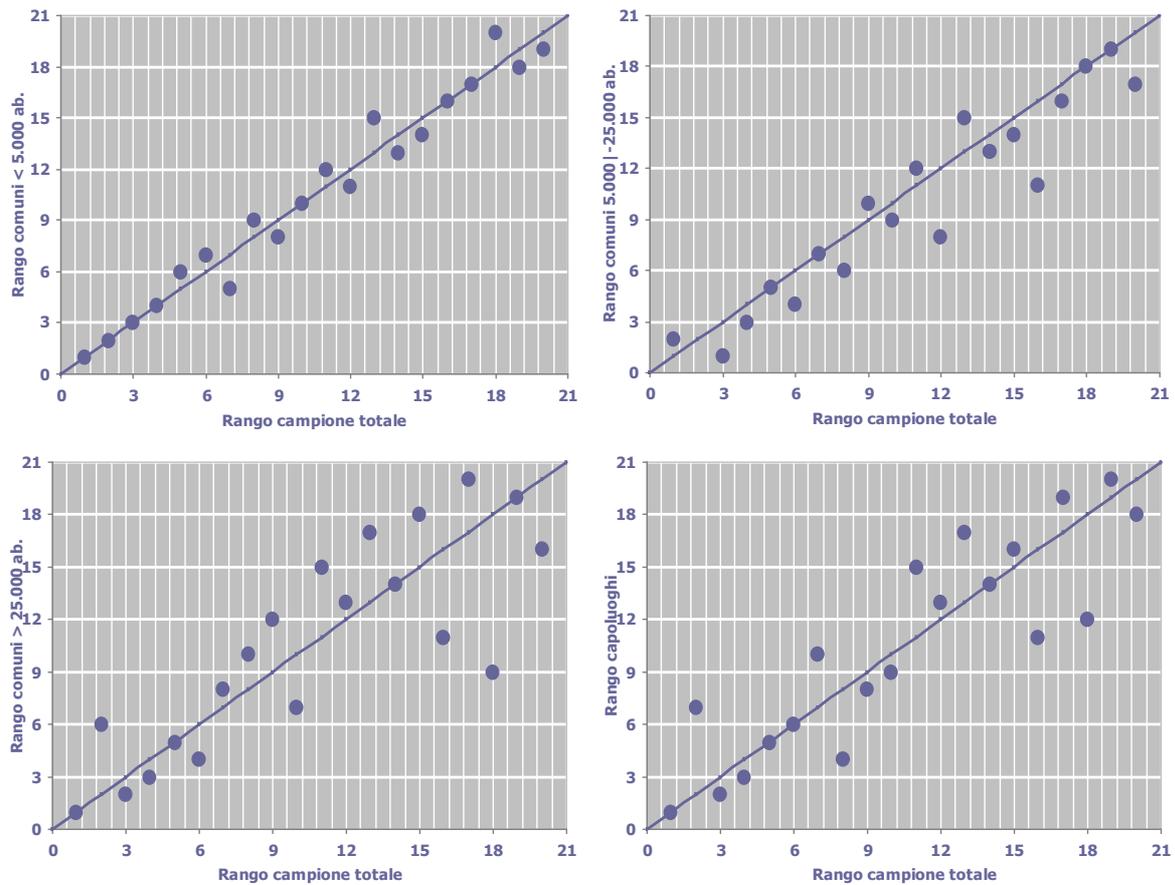
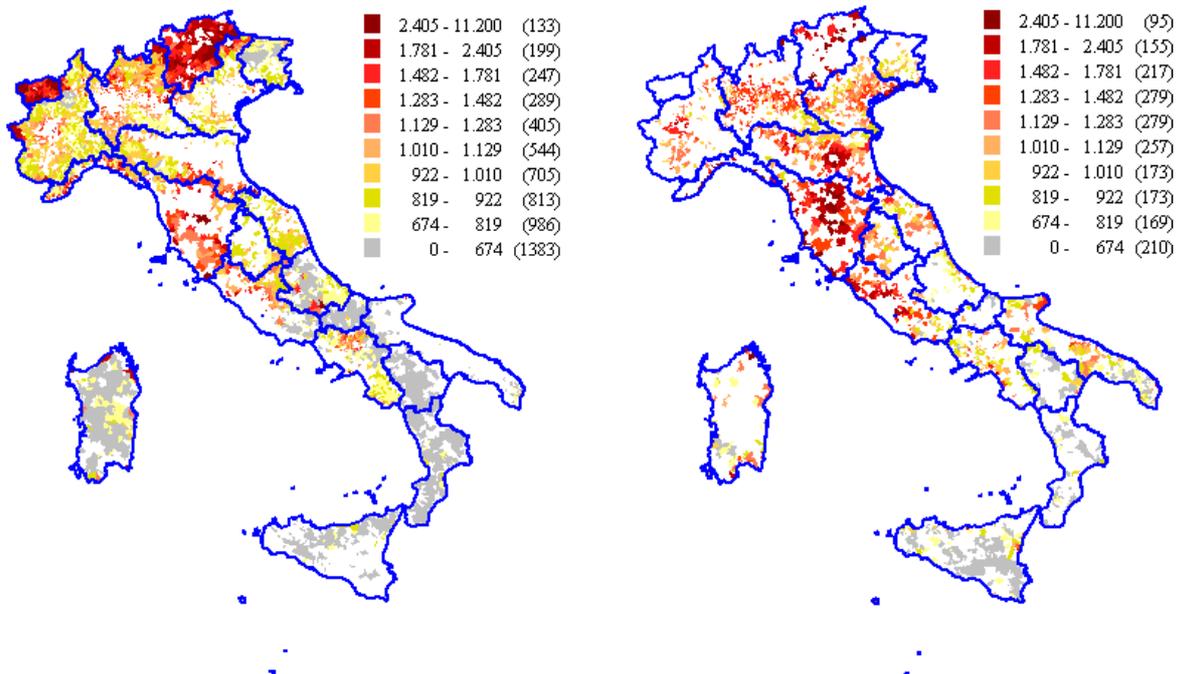


Figura VII: Mappe della distribuzione delle quotazioni per i comuni con popolazione <5.000, tra 5.000 e 25.000 e > di 25.000 rispettivamente



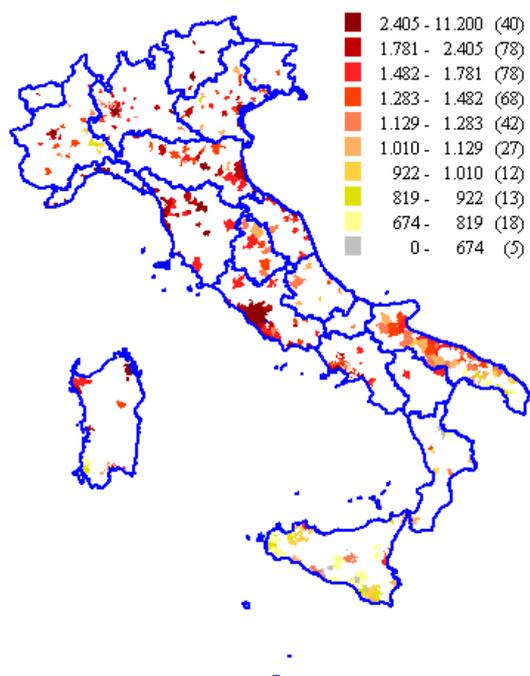


Tabella III: Dati di sintesi per regione e solo capoluoghi

Classe di popolazione	Campione complessivo			Capoluoghi		
	N. comuni	Media	Differenziale rispetto alla media Italia	N. comuni	Media	Differenziale rispetto alla media Italia
<i>Abruzzo</i>	305	784	0,73	4	1.285	0,73
<i>Basilicata</i>	131	578	0,54	2	1.639	0,93
<i>Calabria</i>	409	576	0,53	5	1.004	0,57
<i>Campania</i>	550	1.105	1,02	5	2.112	1,20
<i>Emilia Romagna</i>	341	1.289	1,19	9	2.036	1,16
<i>Friuli Venezia Giulia</i>	219	850	0,79	4	1.465	0,83
<i>Lazio</i>	376	1.199	1,11	5	2.019	1,15
<i>Liguria</i>	235	1.827	1,69	4	2.395	1,36
<i>Lombardia</i>	1546	1.164	1,08	11	1.816	1,03
<i>Marche</i>	246	1.029	0,95	4	1.816	1,03
<i>Molise</i>	136	554	0,51	2	1.174	0,67
<i>Piemonte</i>	1209	949	0,88	8	1.471	0,84
<i>Puglia</i>	257	861	0,80	5	1.270	0,72
<i>Sardegna</i>	373	733	0,68	4	1.703	0,97
<i>Sicilia</i>	390	713	0,66	9	1.145	0,65
<i>Toscana</i>	287	1.753	1,62	10	2.383	1,36
<i>Trentino Alto Adige</i>	339	1.954	1,81	2	3.273	1,86
<i>Umbria</i>	92	980	0,91	2	1.398	0,80
<i>Valle d'Aosta</i>	74	1.884	1,75	1	1.923	1,10
<i>Veneto</i>	581	1.085	1,00	7	1.861	1,06
Italia	8096	1.079	1,00	103	1.755	1,00

Tabella IV: dati di sintesi per regione dell'intero campione, dei soli valori estremi e senza valori estremi

Regione	Campione complessivo			Solo valori estremi			Senza valori estremi		
	N. comuni	Media	Differenziale rispetto alla media Italia	N. comuni	Media	Differenziale rispetto alla media Italia	N. comuni	Media	Differenziale rispetto alla media Italia
Abruzzo	305	784	0,73	26	1.563	0,74	279	711	0,70
Basilicata	131	578	0,54	15	978	0,47	116	526	0,52
Calabria	409	576	0,53	37	924	0,44	372	541	0,54
Campania	550	1.105	1,02	44	2.972	1,42	506	943	0,93
Emilia Romagna	341	1.289	1,19	16	2.810	1,34	325	1.214	1,20
Friuli Venezia Giulia	219	850	0,79	15	1.701	0,81	204	787	0,78
Lazio	376	1.199	1,11	21	2.686	1,28	355	1.111	1,10
Liguria	235	1.827	1,69	7	6.117	2,91	228	1.695	1,68
Lombardia	1.546	1.164	1,08	50	2.273	1,08	1.496	1.127	1,12
Marche	246	1.029	0,95	24	1.825	0,87	222	943	0,93
Molise	136	554	0,51	23	798	0,38	113	505	0,50
Piemonte	1.209	949	0,88	92	1.792	0,85	1.117	879	0,87
Puglia	257	861	0,80	9	1.732	0,82	248	830	0,82
Sardegna	373	733	0,68	48	1.697	0,81	325	591	0,59
Sicilia	390	713	0,66	26	1.625	0,77	364	648	0,64
Toscana	287	1.753	1,62	12	3.935	1,87	275	1.658	1,64
Trentino Alto Adige	339	1.954	1,81	19	4.280	2,04	320	1.816	1,80
Umbria	92	980	0,91	5	1.435	0,68	87	954	0,95
Valle d'Aosta	74	1.884	1,75	11	3.320	1,58	63	1.633	1,62
Veneto	581	1.085	1,00	21	2.483	1,18	560	1.032	1,02
Italia	8.096	1.079	1,00	521	2.100	1,00	7.575	1.009	1,00

Tabella V: Quotazioni medie ponderate per stock e superfici

Regione	Campione complessivo medie semplici			Campione complessivo medie pesate Stock			Campione complessivo medie pesate Superficie		
	N. comuni	Media	Differenziale rispetto alla media Italia	N. comuni	Media	Differenziale rispetto alla media Italia	N. comuni	Media	Differenziale rispetto alla media Italia
Abruzzo	305	784	0,73	305	1.060	0,67	305	1.040	0,68
Basilicata	131	578	0,54	131	809	0,51	131	834	0,55
Calabria	409	576	0,53	409	726	0,46	409	735	0,48
Campania	550	1.105	1,02	550	1.619	1,03	550	1.573	1,03
Emilia Romagna	341	1.289	1,19	341	1.783	1,13	341	1.716	1,13
Friuli Venezia Giulia	219	850	0,79	219	1.272	0,81	219	1.163	0,76
Lazio	376	1.199	1,11	376	2.430	1,54	376	2.358	1,55
Liguria	235	1.827	1,69	235	2.711	1,72	235	2.651	1,74
Lombardia	1546	1.164	1,08	1546	1.632	1,03	1546	1.557	1,02
Marche	246	1.029	0,95	246	1.443	0,91	246	1.408	0,93
Molise	136	554	0,51	136	748	0,47	136	764	0,50
Piemonte	1209	949	0,88	1209	1.484	0,94	1209	1.387	0,91
Puglia	257	861	0,80	257	1.108	0,70	257	1.089	0,72
Sardegna	373	733	0,68	373	1.361	0,86	373	1.277	0,84
Sicilia	390	713	0,66	390	962	0,61	390	972	0,64
Toscana	287	1.753	1,62	287	2.227	1,41	287	2.194	1,44
Trentino Alto Adige	339	1.954	1,81	==	==	==	==	==	==
Umbria	92	980	0,91	92	1.209	0,77	92	1.205	0,79
Valle d'Aosta	74	1.884	1,75	74	2.259	1,43	74	2.197	1,44
Veneto	581	1.085	1,00	581	1.464	0,93	581	1.406	0,92
Italia	8096	1.079	1,00	7757	1.578	1,00	7757	1.521	1,00

Figura VIII: Confronti tra rango delle regioni nel caso di medie semplici e ponderate per lo stock e per la superficie

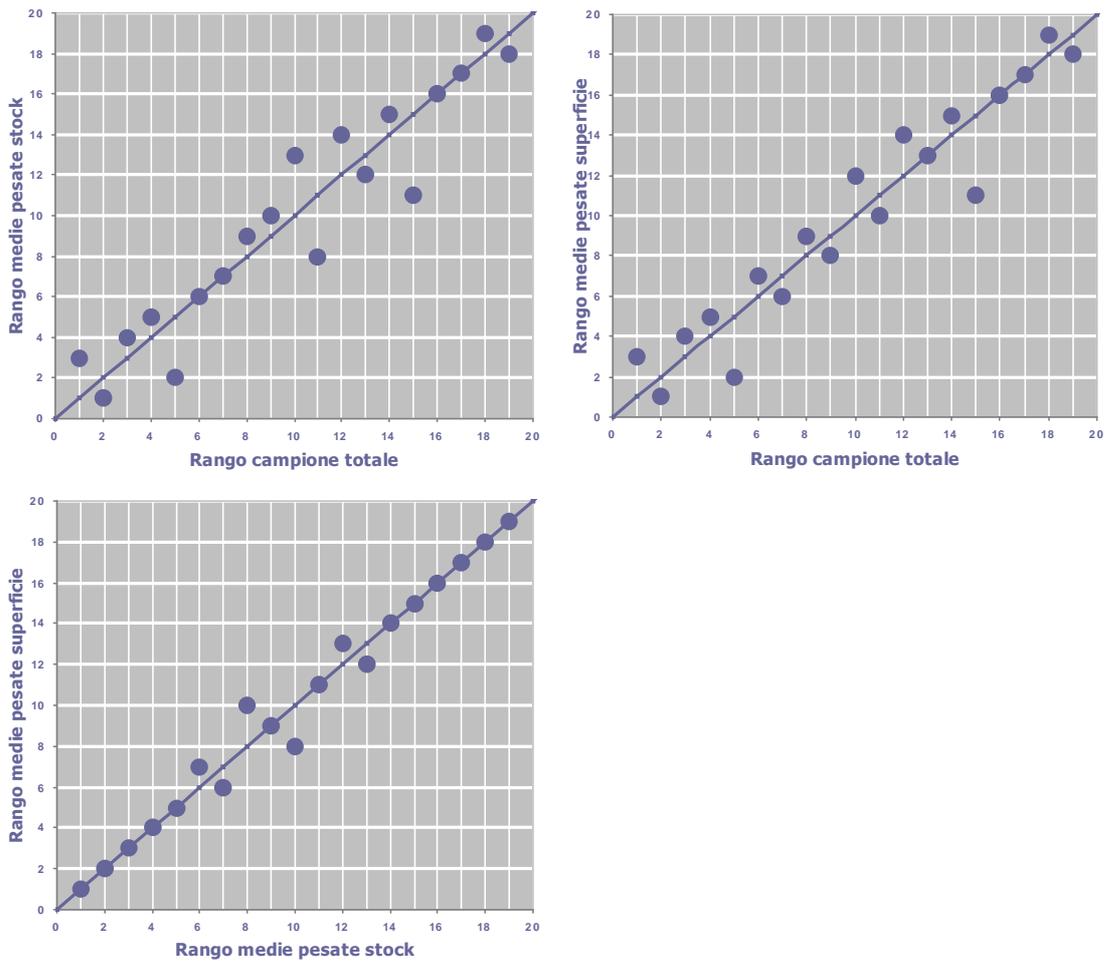


Tabella VI: Significatività del test di confronto tra coppie di regioni

	A b r	B a s	C a l	C a m	E m i	F r i	L a z	L i g	L o m	L a z	M a r	M o l	P i e	P u g	S a r	S i c	T o s	T r e	U m b	V a l	V e n	
Abr																						
Bas																						
Cal																						
Cam																						
Emi																						
Fri																						
Laz																						
Lig																						
Lom																						
Mar																						
Mol																						
Pie																						
Pug																						
Sar																						
Sic																						
Tos																						
Tre																						
Umb																						
Val																						
Ven																						

* Le celle bianche indicano un'elevata significatività (p-value < 0,01) mentre quelle in grigio una moderata significatività (0,01 < p-value < 0,1).

8 Bibliografia

- Agenzia del Territorio (2008). *Manuale della banca dati dell'osservatorio del mercato immobiliare*. (pag. 59). Disponibile sul sito www.agenziaterritorio.it
- Agenzia del Territorio (2005). *Rapporto immobiliare 2005*. Disponibile sul sito www.agenziaterritorio.it
- Agenzia del Territorio (2010). *Statistiche Catastali. Catasto Edilizio Urbano*. Disponibile sul sito www.agenziaterritorio.it
- Anderson, M.J. (2001 a). *A new method for non-parametric multivariate analysis of variance*. *Austral Ecology* 26: 32 - 46.
- Camagni R. (1993). *Principi di economia urbana e territoriale*, Carocci editore
- Conover, William J., Johnson, Mark E., and Johnson, Myrle M. (1981). *A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data*. *Technometrics* 23, 351–361.
- Dipartimento delle Finanze e Agenzia del Territorio (2010). *Gli Immobili in Italia. Distribuzione del patrimonio e dei redditi dei proprietari*. Disponibile sul sito www.agenziaterritorio.it
- Iglewicz, B. and Hoaglin, D. C. (1993). *How to Detect and Handle Outliers*, American Society for Quality Control, Milwaukee, WI.
- ISTAT (2010). *L'abitazione delle famiglie residenti in Italia*.
- Jarque, C. M., Bera, A. K. (1980). *Efficient tests for normality, homoscedasticity and serial independence of regression residuals*. *Economics Letters* 6, 255–259. Available at <http://ideas.repec.org/a/eee/eolet/v6y1980i3p255-259.html>.
- Poggio T. (2009), "Le principali dimensioni della disuguaglianza abitativa", in Andrea Brandolini, Chiara Saraceno e Antonio Schizzerotto (eds.), *Dimensioni della disuguaglianza in Italia: povertà, abitazione, salute*, Bologna, il Mulino.
- Schwarz G (1978). "Estimating the Dimension of a Model." *The Annals of Statistics*, 6, 461–464.
- Sen, Amartya K. 2004. "From Income Inequality to Economic Inequality," in C. Michael Henry (ed.) *Race, Poverty, and Domestic Policy*, pp. 59–82, New Haven and London: Yale University Press.
- Sen, Amartya. *Development as Freedom*. New York: Anchor, 2000. 1–53. Print.
- Thorndike, Robert L. (1953). "Who Belong in the Family?". *Psychometrika* 18 (4).

ABSTRACT

Spending on housing represents a large share of families' income and is therefore crucial to verify the existence and assess the degree of geographical differences in house prices.

Classical tests based on normality and homoskedasticity assumptions seem to be inappropriate tools for the data at hand. In the present work, these differences are examined using a particular method which allows to relax the strong assumptions we usually need in the parametric framework.

The analysis will be carried out using data at a city level, provided by the Osservatorio del Mercato Immobiliare of Agenzia del Territorio. Besides, the examination will be repeated using subsets of cities that appear to be homogeneous with respect to some variables of interest.