

UNIVERSITÀ DELLA VALLE D'AOSTA
UNIVERSITÉ DE LA VALLÉE D'AOSTE

DIPARTIMENTO DI SCIENZE ECONOMICHE E POLITICHE

Corso di Laurea Magistrale in
Economia e Politiche del territorio e dell'impresa
Curriculum mercato e impresa

TESI DI LAUREA MAGISTRALE

IL PREZZO DELLE CASE
IN UNA GRANDE CITTÀ

RELATORE: Gianluigi GORLA

CANDIDATO: Fanny LALE MURIX

N° MATRICOLA: 18 G01 182

Anno Accademico 2019/2020

INDICE

INTRODUZIONE.....	3
CAPITOLO 1: LA TEORIA ECONOMICA.....	4
1.1 L'ECONOMIA URBANA.....	4
1.2 ANALISI DELLA LETTERATURA SCIENTIFICA	4
1.2.1 La teoria dello stato isolato di Von Thünen	5
1.2.2 La teoria del <i>trade-off</i>	8
1.2.3 La relazione tra il mercato fondiario e il mercato immobiliare: l'uso residenziale della terra in una città monocentrica.....	10
1.3 DALLA TEORIA ALLA REALTÀ – LA PRESENZA DI ETEROGENEITÀ.....	16
CAPITOLO 2: METODOLOGIA	17
2.1 IL MODELLO DEI PREZZI EDONICI.....	17
2.1.1 Storia del metodo.....	18
2.1.2 Concetti alla base del metodo.....	19
2.1.3 La funzione di prezzo edonico	20
2.1.4 Vantaggi e svantaggi del modello.....	21
2.2 APPLICAZIONE DEL METODO DEI PREZZI EDONICI AL MERCATO IMMOBILIARE.....	22
2.2.1 Alcuni esempi di applicazioni.....	22
CAPITOLO 3: PRESENTAZIONE DEI DATI.....	25
3.1 LA SCELTA DELLA CITTÀ	25
3.2 LA SCELTA DELLE VARIABILI	26
3.3 LA RACCOLTA DEI DATI.....	28
3.4 L'ELABORAZIONE DEI DATI.....	29
CAPITOLO 4: ANALISI DEI DATI E RISULTATI	31
4.1 ANALISI DESCRITTIVA DEI DATI.....	31
4.1.1 La completezza della base dati	31
4.1.2 Tabelle delle frequenze	34

4.1.3 Statistiche descrittive per le variabili continue	40
4.2 REGRESSIONI E PRINCIPALI RISULTATI.....	42
4.2.1 Relazione tra prezzo al metro quadrato e localizzazione.....	42
4.2.2 Relazione tra prezzo al metro quadrato e distanza.....	45
4.2.3 Modello di regressione con più variabili.....	47
CONCLUSIONI	51
RIFERIMENTI BIBLIOGRAFICI.....	53
SITOGRAFIA	54
RINGRAZIAMENTI.....	55

INTRODUZIONE

Il presente lavoro di tesi si propone di analizzare i prezzi di vendita delle case in una grande città, attraverso l'individuazione delle principali variabili da cui essi dipendono.

Dopo una breve descrizione della disciplina in esame, l'economia urbana, si analizzeranno le principali teorie di localizzazione utili alla comprensione del caso oggetto di studio: la teoria della rendita fondiaria di Von Thünen (1826) e la teoria del trade-off. In base a queste teorie, in una città monocentrica, i prezzi degli immobili dipendono dalla loro distanza dal centro e dalle spese legate al pendolarismo: più un'abitazione è distante dal centro, più sono elevati i costi di trasporto, dunque minore è la disponibilità a pagare delle famiglie per una residenza. Tale teoria individua una curva di prezzo delle residenze decrescente (capitolo 1).

Nella realtà, tuttavia, le ipotesi assunte dal modello del prezzo delle abitazioni fondato sull'accessibilità risultano abbastanza irrealistiche. I territori non sono isomorfi, ma presentano attributi che rendono alcune zone più accessibili delle altre. Anche le abitazioni e gli edifici in cui esse si trovano sono eterogenei perché presentano caratteristiche differenti. Ciò determina delle oscillazioni attorno alla curva di prezzo degli immobili ricavata dalla teoria basata sull'accessibilità (capitolo 1, paragrafo 1.3).

Il prezzo delle abitazioni dipende dunque da molteplici caratteristiche. Per tener conto anche delle altre variabili, oltre alla distanza dal centro, si utilizza il metodo dei prezzi edonici. Nel secondo capitolo si vedrà infatti come, grazie a questo modello, è possibile attribuire a ciascuna caratteristica un prezzo ombra.

Nella seconda parte del presente lavoro (capitoli 3 e 4), si implementa il modello dei prezzi edonici al mercato residenziale di Milano, città che ha ancora forti caratteristiche monocentriche. Attraverso i siti di alcune agenzie immobiliari, sono state raccolte informazioni sulle abitazioni in vendita nella città e sono state inserite in una base dati.

Successivamente, i dati raccolti sono stati elaborati e analizzati con regressioni grazie al programma statistico SPSS.

Dalle analisi emerge un forte legame con la teoria dell'accessibilità: la distanza dal Duomo di Milano, considerato come centro della città, è una variabile chiave per spiegare i prezzi delle abitazioni.

CAPITOLO 1: LA TEORIA ECONOMICA

Prima di iniziare con l'esposizione delle principali teorie sulla localizzazione, si propone (nel paragrafo 1.1) una definizione dell'ambito di riferimento del presente lavoro di tesi: l'economia urbana. Successivamente, nel paragrafo 1.2, si analizzeranno, in particolare, la teoria dello stato isolato di Von Thünen e la teoria del trade-off. Si vedranno, infine, nel paragrafo 1.3, le differenze tra il modello teorico e la realtà.

1.1 L'ECONOMIA URBANA

L'economia urbana è la branca dell'economia che si occupa di studiare la formazione, la struttura e la dinamica delle città. È un termine abbastanza recente, utilizzato solamente a partire dalla metà del secolo scorso. Alla base di questa disciplina vi sono gli studi sulla creazione delle città. In generale, una città si forma quando una zona particolarmente accessibile e facilmente raggiungibile diventa un luogo importante per i traffici, quindi sempre più individui decidono di localizzarsi in quell'area perché ne traggono benefici.

La crescita dei centri urbani, con il tempo, porta alla necessità di occuparsi anche della pianificazione urbanistica della localizzazione residenziale e della gestione del traffico.

L'economia urbana si occupa quindi oggi soprattutto delle aree urbanizzate e della gestione dello spazio residenziale.

Vedremo nel prossimo paragrafo alcune delle prime teorie di localizzazione, per poi procedere con la presentazione della teoria del trade-off, teoria su cui si basano anche le teorie moderne sulla localizzazione residenziale.

1.2 ANALISI DELLA LETTERATURA SCIENTIFICA

Il presente paragrafo si propone di illustrare alcuni modelli della letteratura scientifica utili per spiegare il prezzo delle abitazioni. Partendo dalla teoria della rendita fondiaria di Von Thünen, si dimostrerà che, in una città monocentrica, la variabile chiave per la determinazione dei prezzi delle case è la distanza dal centro. Risulta dunque fondamentale analizzare il modello del prezzo delle abitazioni fondato sull'accessibilità, approssimata dalla distanza dal centro.

1.2.1 La teoria dello stato isolato di Von Thünen

Il primo economista ad aver parlato di localizzazione è stato il tedesco Johann Von Thünen (1780-1850) che nel 1826 ha scritto “*Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie*”, lo “Stato Isolato”. La sua è considerata la prima teoria economica di organizzazione dello spazio. È una teoria che cerca di spiegare in che modo si localizzano le attività agricole attorno ai centri urbani.

Per semplificare, Thünen ipotizza un unico centro urbano isolato, in cui si trova il mercato centrale, situato esattamente al centro di una vasta area omogenea, uniformemente fertile e senza caratteristiche che possano rendere una zona più vantaggiosa delle altre.

La conclusione del suo studio è che le diverse attività di produzione agricola si distribuiscono attorno al centro urbano in modo concentrico. Le coltivazioni più redditizie stanno verso il centro della città, perché possono pagare una rendita più alta. Man mano che ci si allontana dal centro si trovano le coltivazioni sempre meno redditizie, perché la rendita è decrescente. La rendita è qui definita come la differenza di valore tra due terreni uguali, localizzati a una diversa distanza dalla città: un terreno situato vicino al centro ha una rendita maggiore perché ha costi di trasporto minori. I costi di trasporto sono il punto principale delle teorie di localizzazione. La rendita di localizzazione è il fattore che determina la suddivisione delle attività nelle diverse aree.

Secondo la teoria di Von Thünen, i beni deperibili e pesanti vengono prodotti vicino alla città, quelli durevoli e più leggeri nelle aree periferiche.

Si formano così 6 zone concentriche attorno al centro città:

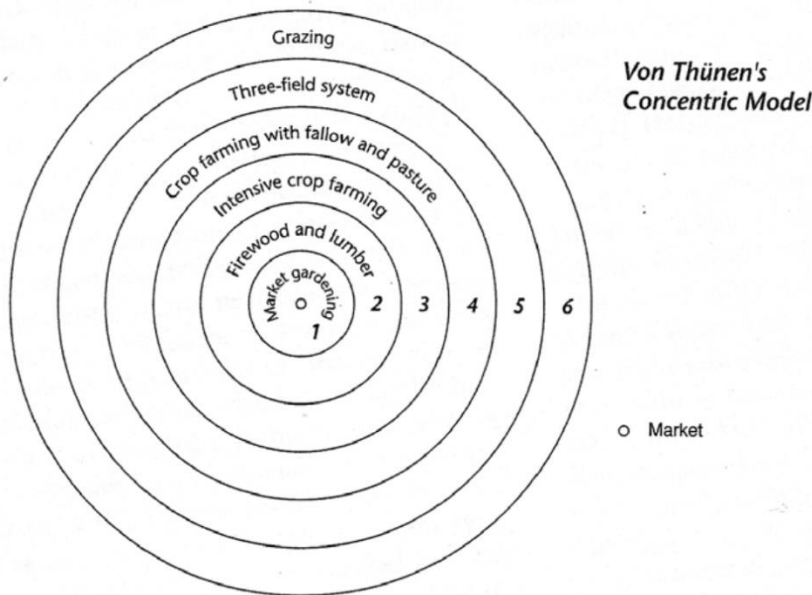
1. la zona più vicina al centro del mercato, detta di “colture libere”, in cui si coltivano i prodotti meno adatti al trasporto, perché più delicati. Nel primo cerchio il prezzo del terreno è più elevato, ma ci sono meno costi di trasporto e meno rischi di trasportare prodotti delicati o deperibili;
2. la zona dove si concentrano i boschi, necessari a quel tempo per far fronte alla grande domanda di legname;
3. la zona dedicata alle colture intensive;
4. la zona dedicata alla rotazione poliennale;
5. la zona dedicata alla rotazione triennale, in cui la terra è tenuta per un anno incolta (o a maggese);
6. la zona dedicata all'allevamento estensivo.

Oltre questi cerchi la terra resta incolta perché, a causa dei costi di trasporto troppo elevati, non conviene coltivarla.

La figura 1.1 sotto riportata rappresenta le zone presentate da Von Thünen.

In realtà, la rappresentazione è successiva, ma può essere utile per comprendere il suo ragionamento, dato che si possono facilmente notare le 6 aree concentriche che si creano attorno al centro.

Figura 1.1 Rappresentazione semplificata dello stato isolato



Fonte: web

A differenza di Ricardo, secondo cui la rendita fondiaria dipendeva quasi esclusivamente dalla fertilità del terreno e la localizzazione del terreno non era importante, per Von Thünen la rendita di ogni terreno dipende dalla sua posizione rispetto al centro, perché la distanza fa variare notevolmente i costi di trasporto per raggiungere il mercato centrale.

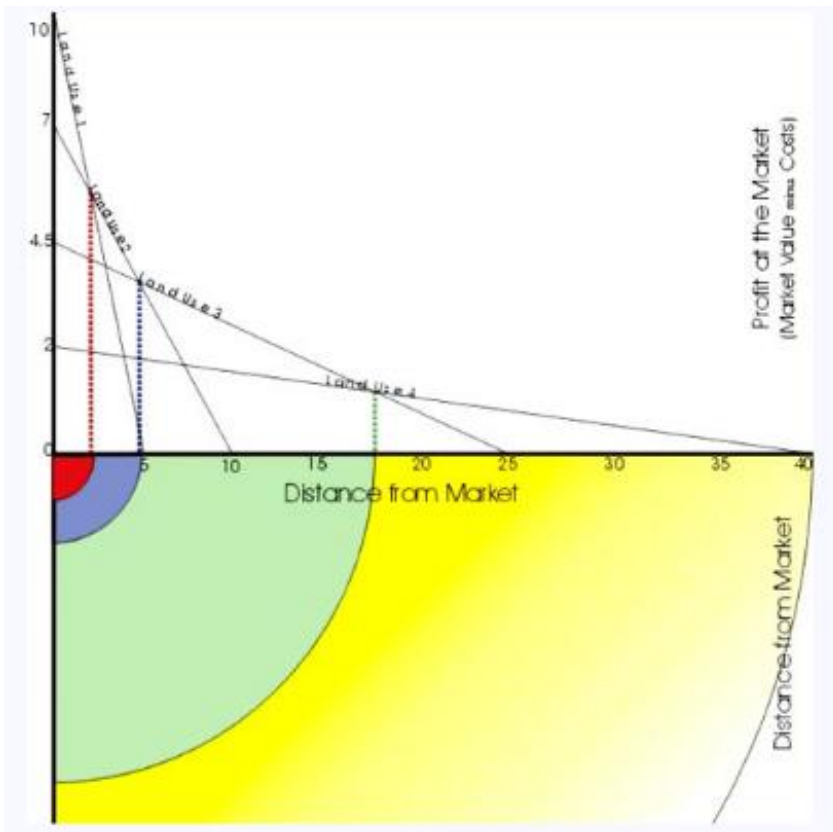
Von Thünen definisce la rendita fondiaria come differenza tra ricavi di produzione e costi di produzione e di trasporto. La formula della rendita fondiaria è quindi:

$$R = Q(p - c) - Qtd$$

con R=rendita fondiaria, Q=quantità, p=prezzo unitario, c=costo unitario, t=costo di trasporto per unità di prodotto, d=distanza del terreno dal mercato centrale.

La Figura 1.2 di seguito riportata rappresenta la rendita in funzione della distanza dal mercato.

Figura 1.2 La rendita fondiaria secondo il modello dello stato isolato



Fonte: Crosier (2001)

La rendita dipende solamente dalla distanza dal mercato, quindi può essere rappresentata come una funzione lineare (una retta) decrescente. Più la distanza aumenta, più la rendita diminuisce, fino a diventare nulla quando $Q_t = Q(p - c)$.

Le varie attività agricole si localizzano dove minimizzano i costi di trasporto, quindi dove massimizzano la rendita fondiaria. Il punto in cui cambia l'attività agricola è quello in cui la rendita marginale uguaglia il costo marginale.

Anche se il modello di Von Thünen presenta delle semplificazioni abbastanza irrealistiche, è comunque considerato un buon punto di partenza per altri studi sulle scelte localizzative. Attualmente, con la globalizzazione, le innovazioni e la maggiore efficienza dei trasporti questo modello non è più utilizzabile, ma è stato un'ottima base per le successive analisi sulla localizzazione e il ragionamento sottostante è ancora valido. Egli ha infatti dimostrato che la rendita fondiaria non dipende dalla fertilità del terreno ma dalla distanza dal centro, ovvero dalla sua accessibilità (la rendita del terreno aumenta all'aumentare dell'accessibilità).

1.2.2 La teoria del *trade-off*

La teoria del trade-off è ancora oggi alla base degli studi di economia urbana sulla localizzazione residenziale. In base a questa teoria, in una città monocentrica, la scelta di dove localizzarsi dipende dalla relazione tra il costo per l'abitazione, che diminuiscono all'aumentare della distanza dal centro, e i costi di trasporto, che invece aumentano con la distanza.

Alla base della teoria del trade-off ci sono alcune importanti assunzioni:

- la città è monocentrica: esiste, cioè, un unico centro direzionale in cui tutti si recano per lavoro;
- la città si trova in un'area omogenea, senza caratteristiche che rendano una zona più accessibile, facendola quindi preferire alle altre;
- tutti i trasporti sono ugualmente efficienti; in questo modo tutte le zone periferiche possono raggiungere il centro con gli stessi tempi e modi;
- non vi sono esternalità;
- gli immobili possono essere adeguati alle esigenze delle famiglie senza nessun costo aggiuntivo
- l'unica variabile rilevante per le famiglie nella scelta di localizzazione è il costo di trasporto dall'abitazione al centro, le altre variabili sono assunte irrilevanti.

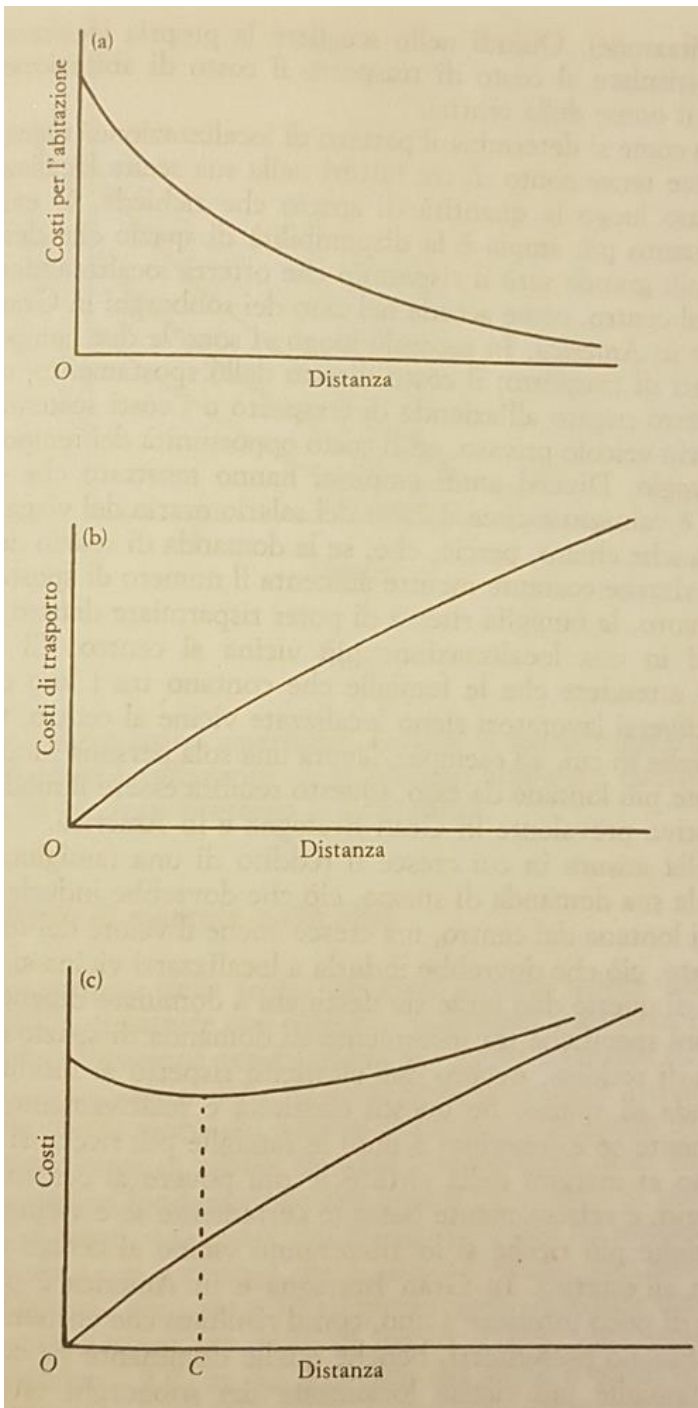
Dunque, se l'obiettivo delle famiglie è quello di minimizzare il costo di trasporto, si crea allora una domanda elevata di localizzazione al centro della città. Tale situazione provoca un aumento dei prezzi del suolo e delle abitazioni nelle zone vicine al centro, mentre restano bassi nelle aree più periferiche.

Esiste perciò un trade-off tra la distanza e localizzazione e, pertanto, tra distanza e prezzo delle abitazioni.

Dai grafici della Figura 1.3 di seguito riportata, si può infatti vedere la curva dei costi dell'abitazione che decresce all'aumentare della distanza. L'inclinazione negativa della curva è data dall'aumentare dei costi di trasporto con la distanza: chi è distante dal centro ha costi di trasporto elevati, quindi è disposto a pagare meno per la casa.

Le famiglie, quando devono scegliere la loro localizzazione, confrontano i costi per l'abitazione (grafico a) e i costi di trasporto (grafico b) per varie localizzazioni. La scelta ricade sul luogo in cui vengono minimizzati tali costi (punto C nel grafico c).

Figura 1.3 Relazioni tra costi e distanza dal centro



Fonte: Evans (1988)

1.2.3 La relazione tra il mercato fondiario e il mercato immobiliare: l'uso residenziale della terra in una città monocentrica

Nel paragrafo 1.2.2 si è visto che l'obiettivo delle famiglie è quello di minimizzare il costo di trasporto, quindi si crea, in questo modo, una domanda elevata di localizzazione al centro della città. Tale situazione provoca un aumento dei prezzi del suolo e delle abitazioni nelle zone vicine al centro, mentre restano bassi nelle aree più periferiche.

Si può dunque affermare che il prezzo della terra è alto perché c'è un'elevata domanda di case e, quindi, un'elevata domanda di terra su cui costruire. Il prezzo alto della terra è perciò il risultato (non la causa) dei prezzi alti delle case.

In questo paragrafo si vuole approfondire l'uso residenziale della terra in una città monocentrica, attraverso l'impiego di un semplice modello del settore residenziale.

Come si è visto analizzando la teoria del trade-off, la funzione di prezzo delle case (o *Housing-Price Function*) mostra la relazione tra i prezzi delle case e la loro distanza dal centro.

Definendo il prezzo delle abitazioni come il prezzo al metro quadrato al mese che una famiglia deve pagare per l'alloggio, la funzione di prezzo degli immobili mostra quanto una famiglia è disposta a pagare per una casa nelle diverse zone della città.

Quindi, ad esempio, se una famiglia affitta una casa di 100 metri quadrati a 500 euro al mese, il prezzo della casa è 5 euro al metro quadrato (500€/100mq). In relazione alla distanza dal centro di questa abitazione, e quindi ai costi che la famiglia deve sostenere per il pendolarismo, essa sarà disposta o meno a sceglierla.

Esistono due tipi di funzione di prezzo delle case: lineare (se si ipotizza che le case siano tutte uguali e che le famiglie abbiano lo stesso budget da spendere) e convessa (con ipotesi più realistiche).

Per semplificare, si spiegherà in primis il modello lineare, per poi abbandonare le ipotesi e procedere all'analisi della funzione convessa.

La funzione di prezzo delle abitazioni è lineare se, per semplificare, si assumono le seguenti ipotesi:

- 1- Tutte le abitazioni sono identiche. Ad esempio, ogni casa della città ha 100 metri quadrati.
- 2- Ogni famiglia ha un budget fisso da spendere tra casa e pendolarismo. Ad esempio, 500 euro al mese.

- 3- Il costo mensile del pendolarismo per ogni kilometro è uguale per tutte le famiglie. Ad esempio, 20 euro al mese per ogni kilometro, quindi una famiglia distante 1 kilometro dal centro paga 20 euro al mese di viaggio per lavorare, paga 40 euro al mese se è a 2 kilometri, ecc.

Si ricordi, inoltre, che in questo capitolo si fa riferimento ad una città monocentrica.

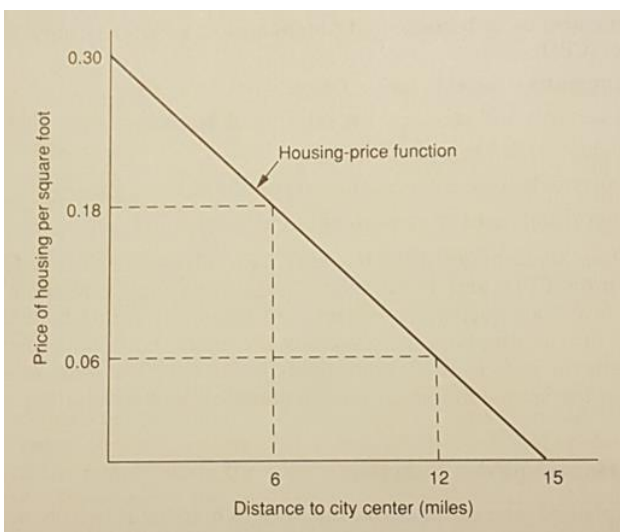
La funzione di prezzo indica quanto è disposta a pagare una famiglia per un'abitazione nelle diverse zone della città.

Sapendo che al centro della città i costi di pendolarismo sono nulli, la famiglia può spendere l'intero budget (500 euro) per la casa. Per una casa di 100 metri quadrati, il prezzo è 5 euro per metro quadrato.

A una distanza di 6 kilometri dal centro, i costi di pendolarismo sono 120 euro, quindi la famiglia ha 380 euro da spendere per la casa (3,8 euro per metro quadrato).

La seguente Figura 1.4 mostra un esempio di Housing-Price Function. L'autore ipotizza abitazioni identiche di 1,000 piedi quadrati, un budget di 300\$ al mese per l'abitazione e un costo di trasporto di 20\$ al mese per ogni miglio di distanza. Dunque, al centro una famiglia può spendere l'intero budget per la casa (300\$), 30 cents per piede quadrato. A 6 miglia di distanza, si spendono 18 cents per piede quadrato (300\$ - 120\$ di trasporto = 180\$ per la casa). Come si può facilmente vedere, il prezzo degli immobili decresce all'aumentare della distanza dal centro e, di conseguenza, dei costi di trasporto.

Figura 1.4 Esempio di Housing-Price Function



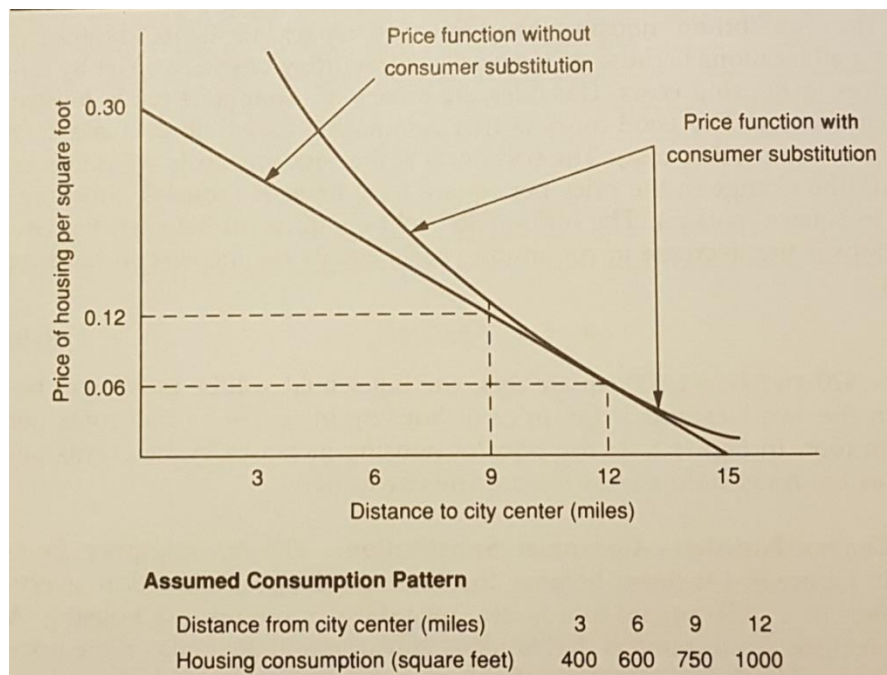
Fonte: Sullivan (1990)

La funzione di prezzo delle abitazioni è inclinata negativamente perché, per ridurre i costi di trasporto, le famiglie sono incentivate a spostarsi verso il centro ma, in questo modo, portano ad un aumento della domanda di case nella zona centrale e, di conseguenza, un aumento dei prezzi. Nelle zone periferiche, invece, la domanda di case diminuisce e, pertanto, diminuiscono anche i prezzi.

La funzione lineare di cui si è appena parlato assume che le case siano identiche. Un'ipotesi più realistica è invece quella secondo cui una famiglia, se si muove verso il centro città, dato che pagherà un prezzo più alto per la casa, cercherà un'abitazione più piccola, con meno metri quadrati.

La Figura 1.5 sotto riportata mostra che, in generale, spostandosi verso il centro città che è più caro, la famiglia tenderà ad occupare case più piccole, che richiedono aumenti progressivamente maggiori di prezzo per metro quadrato per compensare la riduzione dei costi di pendolarismo che rimane fissa per ogni kilometro di distanza. La funzione di prezzo delle abitazioni è, perciò, convessa.

Figura 1.5 Housing-Price Function lineare e convessa



Fonte: Sullivan (1990)

Siccome il prezzo delle case (P) e il consumo di case (H) variano in base alla distanza dal centro (u), il trade-off tra i costi della casa e del pendolarismo può essere scritto come:

$$\Delta u \cdot t = -\Delta P(u) \cdot H(u)$$

Per una piccola variazione della distanza dal centro (Δu), la variazione dei costi del pendolarismo (u volte il costo del pendolarismo per kilometro) uguaglia la variazione del prezzo delle case moltiplicata per il consumo di case.

La pendenza della funzione può quindi essere scritta come una semplice equazione:

$$\frac{\Delta P(u)}{\Delta u} = -\frac{t}{H(u)}$$

Man mano che la famiglia si sposta verso il centro città, il consumo di case diminuisce, facendo aumentare la pendenza della *Housing-Price Function*.

Per calcolare quanto rapidamente il prezzo delle case diminuisce man mano che la distanza dal centro aumenta, si calcola la variazione percentuale del prezzo delle case per ogni kilometro. L'*Housing-Price Gradient* si ottiene dividendo l'equazione precedente per il prezzo delle case, P :

$$\frac{\Delta P/P}{\Delta u} = -\frac{t}{H \cdot P}$$

Si può dunque riassumere il discorso dicendo che, nella scelta di localizzazione, si cerca di analizzare il trade-off che si crea tra un costo residenziale decrescente con la distanza e una spesa crescente per il trasporto.

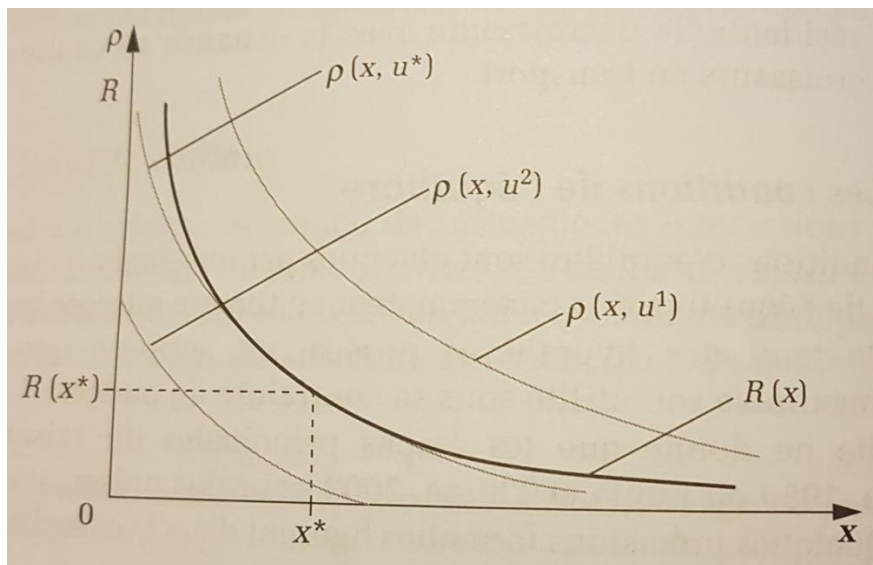
Tuttavia, come vedremo in seguito, nel paragrafo 1.3, nella realtà la curva di prezzo delle abitazioni non è così "perfetta", ma il suo andamento presenta delle oscillazioni, in aumento o in diminuzione, perché le zone e le abitazioni sono eterogenee e hanno delle caratteristiche che influenzano i prezzi.

Tenuto conto delle ipotesi assunte, un'altra modalità per individuare la scelta di localizzazione è attraverso l'analisi dell'utilità delle famiglie perché si suppone che ogni famiglia cerchi di massimizzare la propria utilità sotto un vincolo di budget.

Per tale analisi si utilizza una funzione di prezzo di offerta che esprime il prezzo massimo $\rho(x, u)$ che una famiglia è disposta a offrire per unità di suolo a ogni distanza x dal centro per assicurarsi un livello di utilità u . Questa funzione, decrescente, è rappresentata da un insieme di curve di offerta (simili alle curve di indifferenza del consumatore in microeconomia).

La Figura 1.6 sotto riportata mostra un esempio di tre curve di offerta $\rho(x, u)$ per tre diversi livelli di utilità.

Figura 1.6 Curve di offerta per livello di utilità



Fonte: Huriot, Bourdeau-Lepage (2009)

Esistono infinite curve di questo tipo, perché il livello di utilità varia. Su ogni curva il livello di utilità è costante, quindi su quella curva il residente è indifferente nella scelta di localizzazione.

Per un dato livello di utilità, più la distanza aumenta, più la spesa in trasporto è alta, quindi resta meno budget a disposizione per la terra. Il risultato è che il prezzo massimo offerto per unità di suolo diminuisce.

La natura particolare di queste curve di offerta è che una posizione più bassa corrisponde a una utilità più grande.

La localizzazione di equilibrio, che massimizza l'utilità, è quindi il punto in cui la curva dei prezzi (data) è tangente alla curva di offerta più bassa. Questa localizzazione x^* è tale per cui l'utilità è uguale a u^* e il prezzo di offerta di equilibrio $\rho(x^*, u^*)$ è uguale al prezzo del suolo $R(x^*)$.

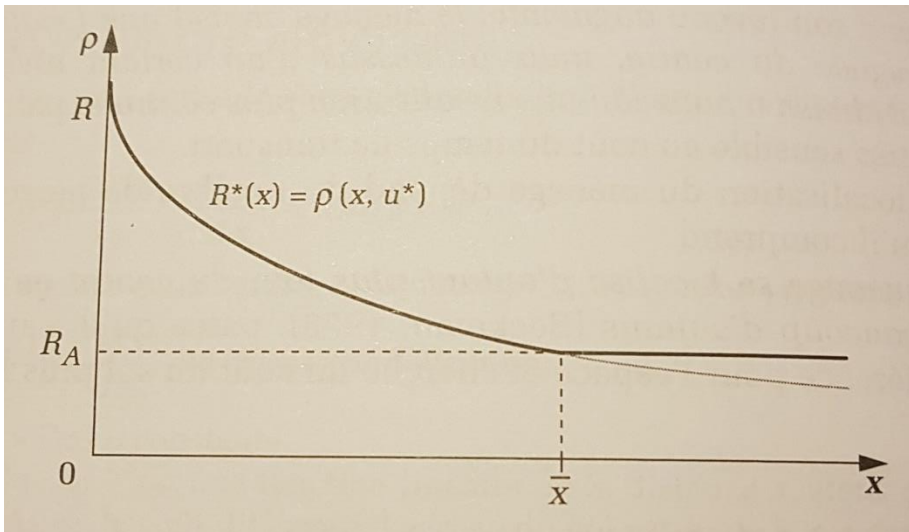
Le proprietà dell'equilibrio risultano dalla risoluzione di un problema di massimizzazione dell'utilità sotto un vincolo di bilancio. Si dimostra che l'equilibrio residenziale della famiglia è tale per cui, a partire dalla localizzazione di equilibrio, uno spostamento marginale implica una variazione di spesa in suolo che compensa esattamente la variazione di spesa in trasporto.

Alle condizioni enunciate, c'è quindi una stretta complementarità tra spesa per il suolo e spesa per il trasporto.

In conclusione, si può quindi affermare che la configurazione dello spazio residenziale monocentrico deriva dall'equilibrio urbano delle famiglie.

Per determinare l'equilibrio simultaneo di tutte le famiglie, bisogna considerare il prezzo della terra agricola, R_A , considerato come costo opportunità dell'utilizzo del suolo. Si può risiedere nella città solamente se si può pagare più di questo prezzo. Fissato anche il livello comune di utilità massima u^* che significa fissare il prezzo di offerta di equilibrio comune a tutte le famiglie $\rho(x, u^*)$. Da questo si deduce la funzione di prezzo di equilibrio $R^*(x)$ perché, a condizione che tutto il suolo della città sia utilizzato, il prezzo del suolo è dappertutto uguale al prezzo di offerta di equilibrio di ciascuno: $R^*(x) = \rho(x, u^*)$. L'intersezione di $R^*(x)$ con R_A indica il limite della città \bar{x} . Oltre questo limite, il prezzo del suolo è uguale al prezzo della terra agricola $R^*(x) = R_A$. La Figura 1.7 rappresenta quanto appena esposto.

Figura 1.7 L'equilibrio urbano



Fonte: Huriot, Bourdeau-Lepage (2009)

Siccome ogni funzione di prezzo di offerta è decrescente, le condizioni di equilibrio implicano immediatamente che, all'interno della città, il prezzo di equilibrio del suolo decresca andando dal centro verso la periferia.

1.3 DALLA TEORIA ALLA REALTÀ – LA PRESENZA DI ETEROGENEITÀ

I modelli analizzati presentano delle ipotesi abbastanza irrealistiche. Nella realtà, il territorio non è isomorfo e gli spazi urbani presentano degli attributi particolari che li rendono eterogenei. Indipendentemente dalla distanza dal centro, esistono delle caratteristiche (naturali, storiche, sociali) che rendono alcuni luoghi più accessibili e attrattivi degli altri. La distribuzione nell'area della città di queste peculiarità e la diversa sensibilità delle famiglie verso di esse possono modificare la struttura residenziale della città così come è presentata nel modello sopra esposto.

Oltre a differenze degli spazi urbani, anche gli immobili sono eterogenei e le famiglie hanno gusti diversi nella scelta della loro abitazione. Gli immobili differiscono non soltanto per la loro superficie, ma anche per caratteristiche proprie sia dell'unità abitativa sia dell'edificio in cui essa si colloca. Tali caratteristiche determinano delle oscillazioni attorno alla curva di prezzo (perfetta) presentata dal modello che tiene conto solamente dell'accessibilità, approssimata dalla distanza dal centro.

Riassumendo, si può dunque dire che per poter calcolare perché i prezzi al metro quadro degli immobili variano è necessario tener conto di diverse variabili:

- caratteristiche dell'abitazione (superficie, tipologia di alloggio, piano in cui si trova, classe energetica, presenza di balcone, cantina, posto auto, ...);
- caratteristiche dell'edificio in cui l'unità abitativa si colloca (anno di costruzione, stato, presenza dell'ascensore, ...);
- caratteristiche del contesto locale (qualità dell'aria, tasso di criminalità, servizi pubblici, ...);
- accessibilità (distanza dal centro e da altri punti di interesse, ...).

Come vedremo nel capitolo successivo, per tener conto di tutte queste caratteristiche e attribuire a ciascuna di esse un prezzo è necessario utilizzare il metodo dei prezzi edonici.

CAPITOLO 2: METODOLOGIA

In questo capitolo si presenta il modello che sarà utilizzato per studiare il caso proposto nei capitoli successivi: il metodo dei prezzi edonici, il quale consente di stimare il prezzo dei beni composti da diverse caratteristiche. Grazie a questo metodo è quindi possibile riuscire a stimare il contributo di ciascuna caratteristica al prezzo del bene. Quest'ultimo è infatti facilmente individuabile, perché è il prezzo che si vede quando si acquista un bene. Il valore di ciascuna caratteristica di cui si compone il bene è invece nascosto nel prezzo finale e per questo motivo viene anche definito "prezzo ombra".

Il modello sarà in primo luogo analizzato da un punto di vista teorico, senza far riferimento a un bene particolare (paragrafo 2.1). Successivamente si vedrà la sua applicazione al mercato immobiliare (paragrafo 2.2), settore in cui rientra il caso proposto nei capitoli seguenti.

2.1 IL MODELLO DEI PREZZI EDONICI

Sappiamo che il prezzo di un terreno o un immobile dipende dalle sue caratteristiche quali, ad esempio, la prossimità al centro e l'accessibilità ad altri luoghi di interesse, la sua estensione, la qualità dell'aria, la sicurezza, ... ma tutti questi fattori si combinano tra loro in maniera complessa e non è facile capire quanto ognuno di essi incida sul valore del bene. Un modo per calcolarlo è attraverso l'applicazione del metodo dei prezzi edonici.

I beni spesso sono composti da una serie di caratteristiche, quindi il loro prezzo dipende dall'insieme di esse. Il metodo dei prezzi edonici si basa sul presupposto che le persone riescano a individuare le varie caratteristiche del bene e a percepire che le medesime peculiarità, combinate in altri modi, non procurerebbero loro la stessa utilità e soddisfazione. Lo scopo di questo metodo è dunque quello di isolare i vari attributi del bene e di stimare, attraverso una regressione, quanto ognuno di essi incida sul prezzo finale del bene.

2.1.1 Storia del metodo

Come accade per molti metodi, anche quello dei prezzi edonici è apparso inizialmente nell'ambito dell'economia agricola quando, nel 1928, Frederick V. Waugh volle studiare le caratteristiche degli ortaggi per capire quali influenzavano maggiormente i consumatori durante l'acquisto. Egli studiò il prezzo degli asparagi a Boston individuandone principalmente tre caratteristiche qualitative: il colore, la misura del gambo e l'uniformità delle punte. Tuttavia, il termine "*hedonic pricing method*" è successivo.

Nel 1939, Court introduce questo concetto applicandolo al settore automobilistico, studiando come variano i prezzi delle autovetture in base alle loro diverse caratteristiche.

Il metodo dei prezzi edonici è stato successivamente applicato in molti altri settori. È stato, ad esempio, impiegato da Gregory Chow nel 1967 per studiare il prezzo dei computer in relazione ai cambiamenti tecnologici e da Chanel e Gerard-Varet nel 1995 per misurare i prezzi nel mercato dell'arte. È molto diffuso anche in studi sul prezzo del vino come, ad esempio, quelli di Combris, Lecocq e Visser (1997 e 2000), Bombrun e Sumner (2003), Schamel (2003), Schamel e Anderson (2003).

Un bene al quale il metodo dei prezzi edonici viene tipicamente applicato è quello del mercato immobiliare.

In questo ambito, il metodo viene collegato soprattutto a Lancaster (1966) e a Rosen (1974). Lancaster ritiene che l'utilità si ricavi non dalla quantità di beni consumati, ma dalla combinazione delle caratteristiche di questi beni. Ad esempio: non consumiamo un'auto, ma una potenza, un confort, un colore, ... non consumiamo un terreno, ma una posizione, una superficie e i vantaggi che derivano dall'averne un terreno... non consumiamo una casa, ma le caratteristiche dell'abitazione. Rosen nel 1974 ha applicato il metodo alle residenze urbane: ha stimato il prezzo delle residenze in relazione alla distanza dal centro e a numerose altre caratteristiche. Nel paragrafo 2.1.2 si riprende questo concetto e si approfondirà lo studio di Lancaster e Rosen.

Molti beni possono dunque essere considerati non solo beni in quanto tali ma come un insieme di caratteristiche. Come abbiamo detto, non si può vedere direttamente il prezzo di ognuna di esse, ma solamente il prezzo finale del bene. Per studiare il contributo di ciascuna sul valore finale si utilizza la funzione di prezzo edonico che vedremo nel paragrafo 2.1.3.

Prima di studiare la funzione di prezzo edonico è importante sottolineare alcuni concetti fondamentali alla base del metodo.

2.1.2 Concetti alla base del metodo

Abbiamo già ribadito che un bene si compone di una combinazione di caratteristiche diverse. Anche in ambito del marketing, infatti, si cercano strategie di differenziazione per consentire ai consumatori di distinguere i prodotti dei diversi concorrenti. La differenziazione di prodotto è l'ipotesi alla base del modello edonico. I consumatori, infatti, possono distinguere i prodotti grazie alle loro caratteristiche e acquistano il bene in quanto insieme di attributi. Questa è l'idea alla base della teoria di Lancaster che abbiamo accennato in precedenza. La scelta del consumatore su un determinato bene piuttosto che un altro dipende infatti dall'insieme delle sue caratteristiche. L'utilità che i consumatori riescono a trarre dal bene è data, quindi, dalle caratteristiche che esso incorpora e può essere descritta come:

$$U_k = U(Y_1, Y_2, \dots, Y_n)$$

con U_k che indica l'utilità data dal bene k e le Y che rappresentano le varie caratteristiche del bene. L'utilità di un bene dipende quindi dalle utilità delle varie caratteristiche di cui si compone. Il consumatore sceglie con l'obiettivo di massimizzare questa utilità, quindi sceglie il prodotto le cui peculiarità sono combinate in modo da procurargli la maggiore soddisfazione. La scelta di un'automobile, ad esempio, dipende dal comfort, dall'affidabilità, dalla sicurezza, dallo stile e da tutti gli altri attributi che la connotano, e dall'importanza che i consumatori attribuiscono loro.

Lancaster affronta il tema dell'utilità e delle diverse caratteristiche che compongono un bene, ma non si occupa di studiarne i prezzi. Il primo a parlare espressamente di metodo dei prezzi edonici è Rosen, nel 1974. Egli riprende la definizione di utilità di Lancaster e aggiunge che attraverso una regressione è possibile stimare quanto ogni caratteristica di un bene contribuisce al prezzo finale del bene in esame.

Il metodo dei prezzi edonici si fonda, quindi, su due assunzioni molto importanti:

1. il prezzo di una determinata tipologia di beni dipende prima di tutto dall'andamento del mercato, poi ogni bene è influenzato dalle sue varie caratteristiche;
2. tutte le caratteristiche devono poter essere valutate economicamente.

2.1.3 La funzione di prezzo edonico

Il prezzo di un bene può essere scritto come funzione delle sue caratteristiche quindi bisogna, in primo luogo, isolare i diversi attributi e attribuire a ciascuno di essi una misura. Per applicare il metodo dei prezzi edonici si utilizza un modello di regressione multipla in cui il prezzo è funzione delle sue caratteristiche. La funzione di prezzo edonico che indica la relazione tra il prezzo del bene e le sue caratteristiche è: $P = f(c)$ con P che indica il prezzo del bene e c che rappresenta il vettore delle diverse caratteristiche. Supponendo n caratteristiche, il vettore diventa $c = (c_1, c_2, \dots, c_n)$, quindi la funzione di prezzo edonico può essere riscritta come $P = f(c_1, c_2, \dots, c_n)$.

Questa funzione viene poi tradotta in una regressione di cui si andranno a stimare i parametri (i coefficienti delle variabili). Il prezzo ombra di uno specifico attributo si determina guardando la variazione del prezzo del bene che si ottiene facendo variare una caratteristica e mantenendo costanti tutte le altre. Quindi se abbiamo ad esempio una funzione di regressione $p = \alpha x + \beta y + \gamma z$ dove x, y, z sono tre caratteristiche del bene e α, β, γ sono i tre coefficienti corrispondenti, se solo la variabile x varia (y e z sono fisse) il coefficiente α misura la variazione di prezzo p quando x varia di una unità. I coefficienti indicano il contributo, positivo o negativo, di ciascuna caratteristica al prezzo finale del bene. Quindi, ad esempio, si interpreta α come il contributo al prezzo del bene della caratteristica x , cioè come prezzo implicito di questa caratteristica. Matematicamente, questo prezzo implicito corrisponde alla derivata prima parziale della funzione di prezzo rispetto alla variabile che misura la caratteristica.

La scelta della forma (lineare, log-lineare, logaritmica, ...) data alla funzione di prezzo è importante, anche perché la relazione tra il prezzo e le caratteristiche potrebbe non essere lineare.

Nel caso in cui si utilizzi la forma lineare, la funzione è del tipo:

$$P = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \gamma_1 Y_1 + \gamma_2 Y_2 + \dots + \gamma_m Y_m + \lambda_1 Z_1 + \lambda_2 Z_2 + \dots + \lambda_l Z_l + \dots$$

dove P indica il prezzo delle abitazioni; X, Y, Z, \dots sono le variabili delle caratteristiche considerate e $\alpha, \beta, \gamma, \lambda, \dots$ sono i prezzi ombra degli attributi da stimare. In questo caso, il prezzo edonico di una caratteristica, ottenuto dalla derivata del prezzo rispetto alla caratteristica stessa, è una costante.

Nella forma log-lineare, invece, tutte le variabili continue sono in forma logaritmica, mentre si lasciano lineari le variabili dummies. Pertanto, la funzione è del tipo:

$$\ln(P) = \alpha + \beta_1 \ln X_1 + \beta_2 \ln X_2 + \dots + \beta_n \ln X_n + \gamma_1 \ln Y_1 + \gamma_2 \ln Y_2 + \dots + \gamma_m \ln Y_m + \dots$$

In questo caso, i coefficienti stimati (α , β , γ , λ , ...) forniscono l'elasticità del prezzo del bene ai singoli attributi considerati.

2.1.4 Vantaggi e svantaggi del modello

Come ogni tecnica, anche il metodo dei prezzi edonici presenta sia vantaggi che svantaggi.

Per quanto concerne i vantaggi si può, in primo luogo, dire che esso concede abbastanza libertà nella scelta delle caratteristiche del bene da inserire nel modello. In secondo luogo, esso si fonda sulle preferenze dei consumatori e sulle loro scelte concrete e può essere modificato, aggiornato e adattato anche successivamente. Inoltre, come abbiamo visto, esso può essere impiegato in diversi settori, dall'ambito agricolo a quello automobilistico, a quello immobiliare. Infine, si tratta di un modello universale, ampiamente diffuso e utilizzato.

Per quanto riguarda, invece, gli svantaggi bisogna innanzitutto sottolineare il fatto che applicare il modello richiede un'ampia e spesso variegata base di dati, quindi è necessario molto tempo per implementarlo, per raccogliere tutti i dati e le informazioni e classificare le diverse variabili. Inoltre, è importante assicurarsi di aver inserito tutte le caratteristiche rilevanti, di trovare eventuali correlazioni tra alcune variabili e, soprattutto, come accennato in precedenza, è fondamentale scegliere la forma della funzione di prezzo edonico adeguata.

La seguente tabella riassume i vantaggi e gli svantaggi di cui abbiamo parlato.

Tabella 2.1 - Vantaggi e svantaggi del modello dei prezzi edonici

VANTAGGI	SVANTAGGI
Libertà	Necessari molti dati e tempo
Si fonda su scelte concrete	Rischio di non inserire caratteristiche rilevanti
Adattabilità	Correlazione tra diverse variabili
Versatilità	Scegliere la forma adatta
Universalità	

2.2 APPLICAZIONE DEL METODO DEI PREZZI EDONICI AL MERCATO IMMOBILIARE

Abbiamo visto che i territori non sono omogenei, come assunto nel modello di Von Thünen, ma presentano caratteristiche che rendono alcune zone più accessibili di altre, facendo così aumentare il valore del terreno. Lo stesso vale nel mercato immobiliare: il prezzo delle case varia in base all'accessibilità della zona in cui si trovano, alle caratteristiche dell'alloggio (superficie, presenza del balcone, classe energetica, ...), alle caratteristiche dell'edificio in cui esso si trova (anno di costruzione, presenza dell'ascensore, ...) e alle caratteristiche del quartiere dove è ubicato l'immobile. La teoria basata sull'accessibilità, di cui abbiamo parlato nel primo capitolo, esprime il prezzo come una curva determinata dal trade-off tra distanza e costi di trasporto. In realtà, questa curva non è così "perfetta" perché presenta delle oscillazioni determinate dalle diverse caratteristiche delle varie zone e degli immobili. La distanza dal centro, da sola, non spiega l'intero valore dell'abitazione. Per studiare il prezzo degli immobili è dunque necessario utilizzare il metodo dei prezzi edonici, il quale consente di inserire anche altre variabili oltre alla distanza e stimare il prezzo ombra di ciascuna.

2.2.1 Alcuni esempi di applicazioni

Alcuni autori hanno già applicato il metodo dei prezzi edonici al mercato immobiliare.

Abbiamo già visto nel paragrafo 2.1.1 Lancaster e Rosen.

Un altro studio è, ad esempio, quello di Cavailhès che, nel 2005, ha misurato l'impatto delle caratteristiche di un'abitazione sul canone di locazione annuale (sulla base dei dati dell'indagine sulle abitazioni dell'INSEE del 1996). Ha studiato diverse aree urbane della Francia e le ha suddivise per dimensione. Considera diverse caratteristiche riguardo a:

1. l'alloggio: superficie abitabile, impianto sanitario, buono stato della casa,...
2. l'edificio: anno di costruzione, qualità, degrado, numero di appartamenti nell'immobile, presenza di un posto auto coperto e di un giardino, ...
3. la posizione dell'abitazione: distanza dal centro, esposizione dell'alloggio, vista ...
4. l'ambiente circostante: criminalità, accessibilità ai servizi, qualità dell'aria, qualità dei servizi offerti, ...

Il risultato dello studio è che il prezzo al metro quadro è più alto nelle abitazioni più piccole e anche che il prezzo medio aumenta molto in base all'aumento della dimensione della città. Quindi, in una città grande, il prezzo medio di affitto è più alto che in una città piccola e, in città della stessa dimensione, le abitazioni più piccole hanno un canone

di affitto al metro quadro più alto rispetto alle case grandi. Dal suo studio emerge anche che le caratteristiche dell'immobile incidono molto sul canone di locazione, soprattutto il degrado dell'edificio. In questo studio, è stato studiato il canone di locazione e, pertanto, il costo della distanza dal centro sembra sottostimato, nel senso che la distanza dal centro è maggiormente rilevante sui prezzi delle case in vendita e, invece, incide meno sui canoni di locazione. Inoltre, si dimostra che in periferia le famiglie sono poco sensibili alle condizioni dell'ambiente, quindi la variabile del contesto risulta rilevante solamente per il centro città.

Un altro studio è quello di Gravel, Martinez e Trannoy (1997) che hanno studiato i prezzi delle case nei 33 comuni più popolati della Val d'Oise (secondo il censimento del 1990) in relazione alle caratteristiche delle abitazioni attraverso il modello dei prezzi edonici. La tabella di seguito riportata mostra il risultato del loro lavoro. Come si può facilmente notare, le caratteristiche che fanno aumentare il prezzo dell'abitazione sono, ad esempio, la casa singola (non in condominio), il numero di locali, il posto auto, la cucina arredata, il balcone, il giardino, la cantina, la presenza di un salone per eventi nei dintorni, la distanza dall'autostrada, ... La caratteristica più importante è la casa singola, la quale fa aumentare il prezzo di ben 161000F. Le caratteristiche che, invece, fanno ridurre il prezzo sono: le tasse sulla casa più alte, la distanza da Parigi, il tasso di alunni bocciati e il tasso di criminalità. La variabile che fa maggiormente diminuire il prezzo dell'abitazione è la distanza da Parigi, con 18000F.

Figura 2.1 Prezzi ombra (in Franchi)

Caractéristiques	Prix implicite
Maison individuelle	+ 161 000 F
1 pièce supplémentaire	+ 41 000 F
1 parking	+ 76 000 F
1 cuisine équipée	+ 70 000 F
1 balcon	+ 54 000 F
1 jardin	+ 53 000 F
1 cave	+ 31 000 F
1 point de taux de la taxe d'habitation	- 8 300 F
1 km de distance de Paris	- 18 000 F
1 km de distance de Roissy	+ 5 900 F
1 km de distance de l'autoroute	+ 8 200 F
1 cinéma pour 1 000 habitants	+ 18 000 F
1 centre culturel pour 1 000 habitants	+ 33 000 F
Autre salle de spectacles par millier d'habitants	+ 40 000 F
1 terrain de petits jeux à proximité	+ 3 500 F
1 point de taux de redoublement en 5e pour 100 élèves	- 10 800 F
1 point de taux de délinquance pour 1 000 habitants	- 1 600 F

Fonte: Huriot, Lise Bourdeau-Lepage (2009), da Gravel et al. (1997)

Nei capitoli successivi si utilizza il metodo dei prezzi edonici per studiare i prezzi di vendita delle case nella città di Milano.

CAPITOLO 3: PRESENTAZIONE DEI DATI

Dopo aver esposto la teoria di base dell'economia urbana e illustrato il metodo dei prezzi edonici, si è voluto applicare questo modello a una città italiana. L'obiettivo del presente lavoro è quello di valutare, in una grande città italiana, quanto le caratteristiche delle abitazioni incidano sul prezzo finale.

In questo capitolo, si descrivono le variabili inserite nel modello e si presenta la base dati.

3.1 LA SCELTA DELLA CITTÀ

La scelta è ricaduta su una grande città per poter avere un numero abbastanza elevato di case in vendita. La città scelta è Milano perché ha ancora forti caratteristiche monocentriche; vi è, infatti, il Duomo al centro, e un primo "cerchio" che si sviluppa attorno ad esso e rappresenta la zona del centro storico, poi un secondo "cerchio" fino alle mura spagnole (o Circonvallazione interna), un terzo fino alle mura medievali (detta anche circonvallazione dei Navigli) e infine fino alla Circonvallazione esterna.

La città di Milano è stata suddivisa in 32 zone:

ZONA	DENOMINAZIONE	ZONA	DENOMINAZIONE
01	Centro	17	Cimiano, Crescenzago, Adriano
02	Arco Della Pace, Arena, Pagano	18	Bicocca, Niguarda
03	Genova, Ticinese	19	Solari, Washington
04	Quadronno, Palestro, Guastalla	20	Affori, Bovisa
05	Garibaldi, Moscova, Porta Nuova	21	San Siro, Trenno
06	Fiera, Sempione, City Life, Portello	22	Bisceglie, Baggio, Olmi
07	Navigli	23	Ripamonti, Vigentino
08	Porta Romana, Cadore, Montenero	24	Forlanini
09	Porta Venezia, Indipendenza	25	Città Studi, Susa
10	Centrale, Repubblica	26	Maggiolina, Istria
11	Cenisio, Sarpi, Isola	27	Precotto, Turro
12	Viale Certosa, Cascina Merlata	28	Udine, Lambrate
13	Bande Nere, Inganni	29	Pasteur, Rovereto
14	Famagosta, Barona	30	Ponte Lambro, Santa Giulia
15	Abbiategrosso, Chiesa Rossa	31	Corvetto, Rogoredo
16	Porta Vittoria, Lodi	32	Napoli, Soderini

3.2 LA SCELTA DELLE VARIABILI

Come si è visto nei capitoli precedenti, i prezzi delle case variano, oltre che in base all'andamento del mercato immobiliare, in relazione a diversi aspetti:

- caratteristiche dell'abitazione;
- caratteristiche dell'edificio in cui la casa si colloca;
- caratteristiche del contesto locale;
- accessibilità.

Per ognuna di queste determinanti sono perciò state individuate delle variabili.

Le caratteristiche del contesto locale (criminalità, inquinamento, ...), siccome non sono disponibili per ogni zona in cui la città è stata suddivisa, sono sintetizzate all'interno delle variabili dummies create. Come si vedrà nel capitolo successivo, infatti, tali attributi distingueranno le varie zone e renderanno differente la localizzazione in una di esse piuttosto che nelle altre.

Per quanto riguarda le caratteristiche dell'abitazione, sono state individuate le seguenti variabili:

- prezzo di vendita richiesto;
- indirizzo: utile per individuare la zona e calcolare le distanze;
- spese condominiali (€/mese);
- superficie (mq): intesa come ammontare in metri quadrati commerciali dell'abitazione;
- tipologia (villa, mansarda, attico, appartamento, ...);
- numero di locali;
- numero di bagni;
- piano in cui si trova la casa;
- casa arredata (sì/no/parzialmente);
- presenza della fibra ottica (sì/no);
- riscaldamento autonomo o centralizzato;
- riscaldamento a radiatori o a pavimento;
- riscaldamento alimentato a metano, a gas, pompa di calore, ...;
- climatizzazione;
- classe energetica;
- accesso disabili (sì/no);
- presenza del balcone (sì/no);

- presenza della cantina privata (sì/no);
- posto auto privato (sì/no);
- giardino privato (sì/no).

Per quanto riguarda le caratteristiche dell'edificio in cui l'unità abitativa si colloca, sono state individuate le seguenti variabili:

- anno di costruzione dell'edificio, per valutarne il degrado strutturale;
- stato dell'edificio: nuovo/in costruzione, ottimo/ristrutturato, buono/abitabile, da ristrutturare;
- presenza dell'ascensore (sì/no);
- edificio storico o di pregio (sì/no);
- servizio portineria (sì/no/mezza giornata);
- presenza di un giardino comune (sì/no).

Per quanto riguarda l'accessibilità, essa è stata approssimata dalla distanza da alcuni luoghi principali, considerando il percorso più corto:

- Duomo di Milano (distanza misurata in km);
- fermata della metro più vicina (in km);
- stazione centrale (in km)
- aeroporto di Linate (in km e tempo in auto o taxi)

Successivamente, anche la variabile "zona" (di cui si è parlato nel paragrafo 3.1) è stata inserita nella base dati, per individuare in quale area della città si trova l'abitazione in esame e considerare intrinsecamente le caratteristiche del luogo.

Infine, è stata aggiunta anche la variabile "intermediario" per indicare da quale sito di agenzia immobiliare sono stati individuati i prezzi delle case e le relative informazioni. I due siti immobiliari considerati sono: immobiliare.it e casa.it.

Alcune variabili inserite nella base dati sono descrivibili con la presenza o assenza di una determinata caratteristica, quindi non sono variabili continue ma dicotomiche. In questo caso la variabile, d , assume valori $d = 0$ e $d = 1$. Se $d = 0$ allora $P(d = 0) = T$ mentre se $d = 1$ allora $P(d = 1) = T + \beta(d)$.

Si è quindi ottenuto un modello del tipo:

$$P = f(X_1, X_2, X_3, \dots, X_n, Y_1, Y_2, \dots, Y_m, Z_1, Z_2, \dots, Z_l, d_1, d_2, \dots, d_g, v)$$

dove X_i rappresenta le caratteristiche dell'abitazione, Y_i le caratteristiche dell'edificio, Z_i la zona, d_i l'accessibilità approssimata dalle distanze, e v l'intermediario.

3.3 LA RACCOLTA DEI DATI

Come evidenziato a livello teorico, uno dei principali problemi da risolvere per avere un buon modello riguarda la raccolta dei dati, la quale deve essere assai precisa e completa, con un elevato numero di unità statistiche.

Al fine di implementare il metodo dei prezzi edonici con il presente lavoro, sono state raccolte informazioni per 200 unità statistiche nel periodo dal 07 agosto 2020 al 20 febbraio 2021.

Come si è precedentemente detto, l'individuazione dei dati è avvenuta tramite la piattaforma online di alcune agenzie immobiliari: immobiliare.it e casa.it perché si è ritenuto che la maggior parte delle famiglie cerchi casa attraverso questi siti e che, quindi, le informazioni siano abbastanza attendibili.

La tecnica utilizzata per reperire i dati è stata quella di inserire nel portale delle agenzie immobiliari la zona in cui si cercava l'abitazione e vedere i risultati. Le abitazioni sono successivamente state selezionate sulla base della quantità di informazioni disponibili. Ad esempio, se per un'abitazione non c'erano la maggior parte delle variabili, essa è stata eliminata. Non sono state prese in considerazione anche le case il cui prezzo non era indicato o era "trattativa riservata".

Sono stati individuati dati per ciascuna delle variabili riferite alla zona, all'intermediario, all'abitazione e all'edificio sopra elencate.

Come si è detto, le caratteristiche del contesto locale (istruzione, verde pubblico, servizi, criminalità, ...) sono state considerate attraverso la variabile "zona".

Le variabili sull'accessibilità sono state elaborate in un secondo momento. Grazie all'indirizzo, infatti, si sono calcolate le distanze dai luoghi sopra indicati attraverso Google Maps che ne ha individuato il percorso più breve. La distanza che ha creato maggiori difficoltà è quella relativa alla fermata della metro più vicina perché è stata individuata visivamente sulla mappa della città di Milano.

3.4 L'ELABORAZIONE DEI DATI

Una volta raccolti i dati, essi sono stati elaborati prima di essere inseriti nel programma statistico per le analisi di regressione.

Dalle variabili “prezzo” di vendita richiesto per l’abitazione e “superficie” (mq) è stata realizzata la variabile “prezzo al metro quadrato” (p_mq). Si è scelto di utilizzare quest’ultima come variabile da spiegare, in modo da avere un parametro più omogeneo per tutte le abitazioni, data l’eterogeneità dei prezzi e delle superfici.

Dato che le case inserite nella base dati sono tutte attici o appartamenti, la variabile “tipologia” è stata trasformata in “attico” per renderla dicotomica: valore 1 se è attico, 0 se è appartamento.

La variabile “piano” in cui l’abitazione si colloca comprendeva anche la voce “piano rialzato” a cui è stato attribuito il valore 0,5 per renderla una variabile totalmente quantitativa.

La variabile “edificio arredato” prevede come risposte “sì”, “no” o “parzialmente” e, perciò, per renderla dicotomica, è stata suddivisa in due variabili differenti:

- parzialmente arredato (parz_arr): valore 1 se è parzialmente arredato, 0 negli altri due casi;
- totalmente arredato (tot_arr): valore 1 quando la risposta è “sì”, valore 0 negli altri due casi.

Alla variabile dicotomica “fibra ottica” è stato attribuito valore 1 in caso di presenza della fibra, valore 0 in caso di assenza di fibra.

La variabile “riscaldamento autonomo o centralizzato” è stata resa dicotomica trasformandola in “riscaldamento autonomo” (risc_aut) e attribuendo valore 1 se è autonomo, valore 0 se è centralizzato.

Lo stesso è stato fatto anche per la variabile “riscaldamento a radiatori o a pavimento”. La variabile è diventata “no_radiatori” quindi è stato attribuito valore 1 in caso di riscaldamento a pavimento, e valore 0 per il riscaldamento a radiatori.

La variabile relativa al tipo di alimentazione del riscaldamento è diventata anch’essa una dummy definita “energie alternative e pompa di calore” (energ_alt e pc), con valore 0 se è alimentato a metano, a gpl o a gas, e valore 1 negli altri casi (pompa di calore, teleriscaldamento, solare, fotovoltaico).

La variabile “classe energetica” è stata suddivisa in 3 classi: classe energetica A, classe energetica B-C-D, classe energetica E-F-G e sono state considerate come variabili dummy.

Altre variabili dicotomiche sono: accesso disabili (acc_dis), balcone (balc), cantina (cant), posto auto o box (pa_box), giardino privato (giard_priv). Esse assumono, quindi, valore 1 quando sono presenti e valore 0 altrimenti.

La variabile “anno di costruzione dell’edificio” è stata suddivisa in 6 classi, successivamente trasformate in dummies: ante 1900, 1900-1940, 1959-1960, 1970-1980, 1990-2000, 2010-2020.

Lo “stato dell’edificio” è stato suddiviso in 3 classi: nuovo/in costruzione e ottimo/ristrutturato (statoncr), buono/abitabile, e da ristrutturare (statodaris). Sono anch’esse variabili dummies.

Le variabili “ascensore”, “edificio storico o di pregio” e “giardino comune” sono dicotomiche chiamate rispettivamente “lift”, “preg” e “giar_com”. Hanno valore 1 quando è presente.

La variabile “servizio portineria”(port) prevede 3 possibili risposte: no, mezza giornata o intera giornata. Nella trasformazione in dummy, è stato attribuito valore 0 in caso non ci sia servizio, e valore 1 quando è presente, senza considerare la differenza tra mezza giornata e giornata intera.

Per quanto riguarda la zona, sono state create 32 colonne, una per ogni zona, in modo da rendere anche questa variabile una dummy.

Anche la variabile “intermediario” è una dummy che assume valore 0 per immobiliare.it e valore 1 per casa.it.

Le distanze sono invece rimaste variabili numeriche, in chilometri o in minuti.

Nel capitolo successivo si analizzeranno queste variabili e si procederà con le regressioni.

CAPITOLO 4: ANALISI DEI DATI E RISULTATI

Nel capitolo precedente si è visto quali sono i dati e le variabili utilizzati per applicare il metodo dei prezzi edonici al caso di Milano e si è spiegato il motivo della scelta di questa città. In seguito, è stata descritta la modalità di raccolta dei dati e della loro elaborazione.

In questo capitolo, attraverso l'impiego del programma statistico SPSS, si analizzano i dati da un punto di vista descrittivo e, successivamente, si procede con l'implementazione del metodo dei prezzi edonici e la presentazione dei risultati.

4.1 ANALISI DESCRITTIVA DEI DATI

Questo paragrafo si propone di analizzare i dati da un punto di vista descrittivo studiando, soprattutto, le frequenze, le medie e i valori massimi e minimi.

4.1.1 La completezza della base dati

La base dati è stata in primo luogo analizzata dal punto di vista della completezza, per valutare quanti casi mancanti ci sono per ognuna delle variabili considerate. Come si può facilmente notare dalla Tabella 4.1, la base dati è abbastanza completa; si riscontrano dati mancanti in maniera rilevante soprattutto per le variabili “servizio portineria” (port) e “giardino comune” (giar_com) con rispettivamente 70 e 98 valori mancanti. Altri dati mancanti, seppur in livello minore, sono presenti per le variabili “presenza dell'ascensore” (lift), “tipo di alimentazione del riscaldamento” (energ_alt e pc), “riscaldamento a radiatori o a pavimento” (no_radiatori), “classe energetica” (classenergA e classenergEFG) e tutte variabili relative all'anno di costruzione dell'edificio. Lo “stato dell'edificio” ha 1 solo caso mancante. Le altre variabili hanno tutti i dati.

Nel presente lavoro, la mancanza di un dato è stata considerata come “*missing value*” anche se, in alcuni casi, quando sul sito di vendita della casa non è specificata la presenza di una determinata caratteristica, si presume che essa non sia presente.

Tabella 4.1 Analisi completezza della base dati

N	Attico	Parzialm. arredato	Totalm. arredato	Fibra	Riscaldam. autonomo	No radiatori
Valido	200	200	200	200	200	165
Mancante	0	0	0	0	0	35

N	Energie alternative o pompa di calore	Classe energetica A	Classe energetica E F G	Accesso disabili	Balcone	Cantina
Valido	160	190	190	200	200	200
Mancante	40	10	10	0	0	0

N	Posto auto o box	Giardino privato	Ante 1900	1900 - 1940	1950 - 1960	1970 - 1980
Valido	200	200	189	189	189	189
Mancante	0	0	11	11	11	11

N	1990 - 2000	2010 - 2020	Stato nuovo o in costruzione	Stato da ristrutturare	Ascensore	Storico o di pregio
Valido	189	189	199	199	168	200
Mancante	11	11	1	1	32	0

N	Servizio portineria	Giardino comune	z01	z02	z03	z04
Valido	130	102	200	200	200	200
Mancante	70	98	0	0	0	0

N	z05	z06	z07	z08	z09	z10
Valido	200	200	200	200	200	200
Mancante	0	0	0	0	0	0

N	z11	z12	z13	z14	z15	z16
Valido	200	200	200	200	200	200
Mancante	0	0	0	0	0	0

N	z17	z18	z19	z20	z21	z22
Valido	200	200	200	200	200	200
Mancante	0	0	0	0	0	0

N	z23	z24	z25	z26	z27	z28
Valido	200	200	200	200	200	200
Mancante	0	0	0	0	0	0

N	z29	z30	z31	z32	Intermediario
Valido	200	200	200	200	200
Mancante	0	0	0	0	0

4.1.2 Tabelle delle frequenze

Successivamente, sono state create le tabelle delle frequenze per tutte le variabili dummies, per individuarne i risultati. Di seguito si riportano le principali tabelle e si illustrano brevemente.

Come mostra la Tabella 4.2 sotto riportata, la variabile “attico”, che indica se la tipologia di abitazione è attico (quando assume valore 1) o appartamento (quando assume valore 0), assume valore 1 solamente in 11 casi. Quindi il 94,5% delle abitazioni considerate è un appartamento.

Tabella 4.2 Tipo di immobile (1=attico, 0=appartamento)

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	0	189	94,5	94,5	94,5
	1	11	5,5	5,5	100,0
	Totale	200	100,0	100,0	

Di seguito è riportata la tabella che riassume le frequenze relative alle due variabili sull'arredamento dell'abitazione. Come si può notare, delle 200 case oggetto di analisi, 44 sono parzialmente arredate (il 22%) e 40 (il 20%) sono totalmente arredate, pertanto 116 non sono arredate.

Tabella 4.3 Stato dell'abitazione: presenza dell'arredamento

Stato dell'abitazione	Frequenza	Percentuale
Arredata	40	20
Parzialmente arredata	44	22
Non arredata	116	58
Totale	200	100

Fonte: elaborazione su dati SPSS

La tabella che segue mostra che la fibra ottica è presente in 118 abitazioni (il 59% dei casi).

Tabella 4.4 Fibra ottica (1=presente, 0=assente)

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	0	82	41,0	41,0	41,0
	1	118	59,0	59,0	100,0
	Totale	200	100,0	100,0	

Per quanto riguarda il riscaldamento, dalle tabelle 4.5, 4.6 e 4.7 di seguito riportate, risulta che 53 abitazioni su 200 hanno il riscaldamento autonomo, mentre il restante 73,5% ha il riscaldamento centralizzato. Inoltre, esclusi i 35 casi mancanti, 124 abitazioni (62%) hanno i radiatori e 41 (il 20,5%) hanno il riscaldamento a pavimento o ad aria. Per alimentare il riscaldamento, esclusi i 40 dati mancanti, 133 abitazioni (il 66,5%) utilizzano metano, gas o gpl, mentre solamente 27 (13,5%) utilizzano energie alternative o pompa di calore.

Tabella 4.5 Riscaldamento: 1=autonomo, 0=centralizzato

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	0	147	73,5	73,5	73,5
	1	53	26,5	26,5	100,0
	Totale	200	100,0	100,0	

Tabella 4.6 Riscaldamento: 1=a pavimento o ad aria, 0=a radiatori

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	0	124	62,0	75,2	75,2
	1	41	20,5	24,8	100,0
	Totale	165	82,5	100,0	
Mancante	Sistema	35	17,5		
	Totale	200	100,0		

Tabella 4.7 Riscaldamento: 1=energie alternative o pompa di calore, 0=metano, gas o gpl

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	0	133	66,5	83,1	83,1
	1	27	13,5	16,9	100,0
	Totale	160	80,0	100,0	
Mancante	Sistema	40	20,0		
	Totale	200	100,0		

Come è stato detto nel precedente capitolo, la classe energetica, definita dalle lettere dalla A alla G, è stata suddivisa in 3 classi. La seguente tabella (4.8) indica le frequenze per le varie classi. Escludendo i 10 casi mancanti, solamente 19 abitazioni (9,5%) sono classificate A, mentre 142 (71%) rientrano nella classe energetica E, F o G. Pertanto, 29 abitazioni sono classificate B, C o D.

Tabella 4.8 Classe energetica

	Frequenza	Percentuale
Classe energetica A	19	9,5
Classe energetica B C D	29	14,5
Classe energetica E F G	142	71,0
Casi mancanti	10	5,0
Totale	200	100,0

Fonte: elaborazione su risultati SPSS

Le tabelle di seguito riportate (Tabelle 4.9 – 4.13) mostrano in quante abitazioni sono presenti i servizi indicati. L'accesso disabili è previsto in 49 residenze su 200 casi analizzati (corrispondente al 24,5%), il balcone in 136 case (68%), la cantina in 126 (63%), il posto auto o box in 48 (24%) e il giardino privato in 11 casi (5,5%).

Tabella 4.9 Accesso disabili (1=presente; 0=assente)

	Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido 0	151	75,5	75,5	75,5
1	49	24,5	24,5	100,0
Totale	200	100,0	100,0	

Tabella 4.10 Balcone (1=presente, 0=assente)

	Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido 0	64	32,0	32,0	32,0
1	136	68,0	68,0	100,0
Totale	200	100,0	100,0	

Tabella 4.11 Cantina (1=presente, 0=assente)

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	0	74	37,0	37,0	37,0
	1	126	63,0	63,0	100,0
	Totale	200	100,0	100,0	

Tabella 4.12 Posto auto o box (1=presente, 0=assente)

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	0	152	76,0	76,0	76,0
	1	48	24,0	24,0	100,0
	Totale	200	100,0	100,0	

Tabella 4.13 Giardino privato (1=presente, 0=assente)

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	0	189	94,5	94,5	94,5
	1	11	5,5	5,5	100,0
	Totale	200	100,0	100,0	

Per quanto riguarda le caratteristiche dell'edificio, si è detto che la variabile "anno di costruzione" è stata suddivisa in sei classi. Pertanto, la suddivisione delle abitazioni nelle varie classi di età dell'edificio è rappresentata nella tabella che segue (Tabella 4.14).

Tabella 4.14 Anno di costruzione dell'immobile

	Frequenze	Percentuale
Ante 1900	6	3,0
1900 - 1940	49	24,5
1950 - 1960	69	34,5
1970 - 1980	26	13,0
1990 - 2000	10	5,0
2010 - 2020	29	14,5
Casi mancanti	11	5,5
Totale	200	100,0

Fonte: elaborazione su risultati SPSS

Per quanto riguarda lo stato dell'edificio, sono state realizzate tre variabili *dummies*. Prendendo come categoria di riferimento la variabile relativa allo stato dell'immobile buono, abitabile, ottimo e ristrutturato, la Tabella 4.15 di seguito riportata mostra che, tolto un caso mancante, 120 abitazioni (60%) sono nuove o in costruzione, 21 (10,5%) sono da ristrutturare e, pertanto, quelle abitabili sono 58 (29%).

Tabella 4.15 Stato dell'edificio

	Frequenze	Percentuale
Nuovo/in costruzione	120	60,0
Ottimo/ristrutturato/ Buono/abitabile	58	29,0
Da ristrutturare	21	10,5
Casi mancanti	1	0,5
Totale	200	100,0

Fonte: elaborazione su risultati SPSS

Altre caratteristiche dell'edificio sono rappresentate dalle variabili relative alla presenza dell'ascensore, del servizio portineria e del giardino comune. Nel campione selezionato, come mostrano le Tabelle 4.16, 4.18 e 4.19, 167 immobili (83,5%) hanno l'ascensore, 128 (64%) hanno il servizio portineria e 97 (48,5%) hanno il giardino comune. Vi sono, inoltre, 14 edifici (7%) storici o di pregio (Tabella 4.17).

Tabella 4.16 Ascensore (1=presente, 0=assente)

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	0	1	,5	,6	,6
	1	167	83,5	99,4	100,0
	Totale	168	84,0	100,0	
Mancante	Sistema	32	16,0		
Totale		200	100,0		

Tabella 4.17 Edificio storico o di pregio (1=sì, 0=no)

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	0	186	93,0	93,0	93,0
	1	14	7,0	7,0	100,0
	Totale	200	100,0	100,0	

Tabella 4.18 Servizio portineria (1=presente, 0=assente)

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	0	2	1,0	1,5	1,5
	1	128	64,0	98,5	100,0
	Totale	130	65,0	100,0	
Mancante	Sistema	70	35,0		
Totale		200	100,0		

Tabella 4.19 Giardino comune (1=presente; 0=assente)

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	0	5	2,5	4,9	4,9
	1	97	48,5	95,1	100,0
	Totale	102	51,0	100,0	
Mancante	Sistema	98	49,0		
Totale		200	100,0		

Per quanto riguarda le caratteristiche del contesto locale, si è già detto che esse sono state comprese all'interno della suddivisione della città di Milano in 32 zone. Il numero di abitazioni considerate in ogni zona è mostrato nella Tabella 4.20 di seguito riportata.

Tabella 4.20 Numero e percentuale di abitazioni in ogni zona

zona	n	%	zona	n	%	zona	n	%	zona	n	%
01	18	9,0	09	6	3,0	17	6	3,0	25	4	2,0
02	7	3,5	10	7	3,5	18	7	3,5	26	5	2,5
03	8	4,0	11	7	3,5	19	5	2,5	27	4	2,0
04	7	3,5	12	5	2,5	20	5	2,5	28	5	2,5
05	7	3,5	13	8	4,0	21	5	2,5	29	4	2,0
06	7	3,5	14	5	2,5	22	7	3,5	30	5	2,5
07	9	4,5	15	7	3,5	23	6	3,0	31	4	2,0
08	5	2,5	16	7	3,5	24	4	2,0	32	4	2,0

Fonte: elaborazione su risultati SPSS

Infine, per quanto riguarda l'intermediario, la Tabella 4.21 mostra che 175 abitazioni (87,5%) sono state selezionate dal sito immobiliare.it, mentre 25 (12,5%) da casa.it.

Tabella 4.21 Intermediario (1=casa.it, 0=immobiliare.it)

	Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido 0	175	87,5	87,5	87,5
1	25	12,5	12,5	100,0
Totale	200	100,0	100,0	

4.1.3 Statistiche descrittive per le variabili continue

In questo paragrafo si analizzano le variabili continue. Nella Tabella 4.24, di seguito riportata, sono indicati, per ogni variabile continua, il numero di casi validi (tolto quelli mancanti), il valore minimo, il valore massimo, la media e la deviazione standard.

Come si può facilmente notare, il prezzo delle abitazioni selezionate varia da 95.000 euro a 4.350.000 euro, con un prezzo medio di 575.952,25 euro. Anche la superficie è molto varia: da 25 metri quadrati a 304 metri quadrati, con una media di 101,81 mq. Pertanto, il prezzo al metro quadrato varia da 1.753 euro a 14.309 euro, con un prezzo medio di 5.293,90 euro per ogni metro quadrato. Inoltre, le residenze studiate hanno spese condominiali assai differenti tra loro; infatti, se per alcune non ci sono spese, per altre è necessario pagare fino a 717 euro al mese. Esse ammontano, in media, a 178,04 euro al mese. Anche il numero di locali e di bagni è molto differente: alcune abitazioni, infatti, hanno addirittura 9 locali e 5 bagni. In generale, comunque, la media è di 2,89 locali e 1,46 bagni. Le case sono situate a diversi piani degli edifici: dal piano terra, fino al piano 16, con una media di 2,5.

Per quanto riguarda le distanze, la variabile ritenuta più importante è la distanza dalla fermata della metro, perché è un indicatore dell'accessibilità, dato che consente di raggiungere in breve tempo tutti i luoghi principali. Da questo punto di vista, le abitazioni non hanno la stessa accessibilità. Alcune abitazioni, infatti, distano solamente 50 metri da una fermata della metro, mentre altre distano quasi 5 chilometri. Tuttavia, la media indica, in generale, una buona accessibilità, dato

che la distanza media è inferiore a 1 chilometro. Altra distanza fondamentale è quella dal centro, approssimato al Duomo di Milano. Le residenze oggetto di analisi distano, in media, circa 3,8 chilometri dal Duomo, ma vi sono notevoli differenze anche in questo caso: l'abitazione più vicina si trova a soli 450 metri dal Duomo, mentre quella più distante è a 9,3 chilometri. La stazione centrale dista, in media, circa 5 chilometri (11 minuti in auto); la casa più vicina è situata a 550 metri (2 minuti in auto), quella più distante a 13 chilometri (24 minuti). L'aeroporto di Linate è raggiungibile, in media, in 13,5 minuti in taxi (9,5 chilometri); 4 minuti (3,5 chilometri) dall'abitazione più vicina, 26 minuti (18,3 km) da quella più distante.

Tabella 4.22 Statistiche descrittive variabili continue

	N	Minimo	Massimo	Media	Deviazione std.
prezzo	200	95000	4350000	575952,25	557412,356
spesemese	191	0	717	178,04	136,903
mq	200	25	304	101,81	52,235
p_mq	200	1753	14309	5293,90	2350,802
nlocali	200	1	9	2,89	1,003
nbagni	199	1	5	1,46	,664
piano	199	,0	16,0	2,533	2,3688
Linate_km	200	3,5	18,3	9,520	3,0800
Linate-taxi	200	4,00	26,00	13,5850	4,87252
Centrale_km	200	,550	13,000	5,17175	2,575794
Centrale_auto	200	2,00	24,00	11,2850	4,35299
MM_km	200	,050	4,900	,87203	,753989
Duomo_km	200	,450	9,300	3,79925	1,976771
Numero di casi validi (listwise)	190				

4.2 REGRESSIONI E PRINCIPALI RISULTATI

Nel presente paragrafo si descrivono le principali regressioni oggetto di studio e si analizzano i principali risultati. Sono state effettuate sia regressioni lineari, sia log-lineari.

4.2.1 Relazione tra prezzo al metro quadrato e localizzazione

Sono stati, in primo luogo, analizzati i prezzi al metro quadrato delle abitazioni in relazione alla loro zona di localizzazione. Pertanto, è stato realizzato un modello di regressione inserendo come variabile dipendente il prezzo al metro quadrato (p_{mq}) delle case e come predittori le 32 zone in cui è stata suddivisa la città, assumendo come categoria di riferimento la zona 1 (Centro).

Come si può notare dalla Tabella 4.23 sotto riportata, la localizzazione, da sola, spiega il 72% del modello ($R^2_{\text{corretto}}=0,722$), quindi si tratta di un buon modello, anche per la significatività delle variabili dummies. Il risultato, mostrato nella Tabella 4.24, indica che nella zona Centro i prezzi al metro quadrato sono più alti. Infatti, i coefficienti non standardizzati β sono tutti negativi (tranne che per la categoria di riferimento). Ciò significa che, come visto nella teoria economica (capitolo 1), i prezzi delle abitazioni sono più elevati al centro e diminuiscono con la distanza.

Tabella 4.23 Zone. Forma lineare

Modello	R	R- quadrato	R- quadrato adattato	Errore std. della stima
1	,875 ^a	,765	,722	1240,297

Tabella 4.24 Coefficienti. Forma lineare

Modello	Coefficienti non standardizzati		Coefficienti standardizzati	t	Sign.
	B	Errore standard	Beta		
1 (Costante)	9494,916	292,341		32,479	,000
z02	-2716,921	552,472	-,213	-4,918	,000
z03	-2074,723	527,025	-,173	-3,937	,000
z04	-2489,170	552,472	-,195	-4,506	,000
z05	-735,742	552,472	-,058	-1,332	,185
z06	-3556,834	552,472	-,279	-6,438	,000
z07	-3573,696	506,349	-,316	-7,058	,000
z08	-3073,237	627,001	-,205	-4,901	,000
z09	-3207,798	584,682	-,233	-5,486	,000
z10	-4295,625	552,472	-,337	-7,775	,000
z11	-3224,654	552,472	-,253	-5,837	,000
z12	-4616,571	627,001	-,307	-7,363	,000
z13	-5705,904	527,025	-,477	-10,827	,000
z14	-5576,203	627,001	-,371	-8,893	,000
z15	-5773,386	552,472	-,452	-10,450	,000
z16	-5699,394	552,472	-,447	-10,316	,000
z17	-6537,685	584,682	-,476	-11,182	,000
z18	-6445,004	552,472	-,505	-11,666	,000
z19	-3674,360	627,001	-,245	-5,860	,000
z20	-6009,648	627,001	-,400	-9,585	,000
z21	-5178,459	627,001	-,345	-8,259	,000
z22	-6974,222	552,472	-,547	-12,624	,000
z23	-5081,625	584,682	-,370	-8,691	,000
z24	-7021,156	685,600	-,419	-10,241	,000
z25	-3951,801	685,600	-,236	-5,764	,000
z26	-5198,595	627,001	-,346	-8,291	,000
z27	-6261,337	685,600	-,374	-9,133	,000
z28	-5894,500	627,001	-,392	-9,401	,000
z29	-4996,956	685,600	-,298	-7,288	,000
z30	-6973,502	627,001	-,464	-11,122	,000
z31	-6552,930	685,600	-,391	-9,558	,000
z32	-4335,802	685,600	-,259	-6,324	,000

Lo stesso modello è stato, in seguito, realizzato con la forma log-lineare. Si è pertanto inserito, come variabile dipendente, il logaritmo del prezzo al metro quadrato delle abitazioni ($\ln p_{mq}$) e sono state mantenute, come variabili dummies, le zone. Come si nota dalla Tabella 4.25, con la forma log-lineare il modello migliora perché l' R^2 corretto aumenta a 0,767. Il risultato, mostrato nella Tabella 4.26, indica, anche in questo caso, che al Centro (categoria di riferimento) i prezzi sono maggiori, perché tutti i coefficienti β sono negativi.

Tabella 4.25 Zone. Forma log-lineare

Modello	R	R- quadrato	R- quadrato adattato	Errore std. della stima
1	,896 ^a	,804	,767	,20974

Tabella 4.26 Coefficienti. Forma log-lineare

Modello	Coefficienti non standardizzati		Coefficienti standardizzati	t	Sign.
	B	Errore standard	Beta		
1 (Costante)	9,127	,049		184,632	,000
z02	-,332	,093	-,141	-3,550	,001
z03	-,226	,089	-,102	-2,535	,012
z04	-,281	,093	-,119	-3,010	,003
z05	-,078	,093	-,033	-,835	,405
z06	-,446	,093	-,189	-4,773	,000
z07	-,455	,086	-,218	-5,319	,000
z08	-,374	,106	-,135	-3,531	,001
z09	-,406	,099	-,160	-4,102	,000
z10	-,580	,093	-,246	-6,211	,000
z11	-,395	,093	-,168	-4,233	,000
z12	-,670	,106	-,241	-6,321	,000
z13	-,903	,089	-,408	-10,130	,000
z14	-,882	,106	-,317	-8,319	,000
z15	-,924	,093	-,392	-9,895	,000
z16	-,902	,093	-,382	-9,650	,000
z17	-1,154	,099	-,454	-11,676	,000
z18	-1,122	,093	-,475	-12,010	,000
z19	-,467	,106	-,168	-4,407	,000
z20	-1,008	,106	-,363	-9,503	,000
z21	-,774	,106	-,279	-7,304	,000
z22	-1,323	,093	-,560	-14,157	,000
z23	-,747	,099	-,294	-7,559	,000
z24	-1,323	,116	-,427	-11,414	,000
z25	-,514	,116	-,166	-4,433	,000
z26	-,767	,106	-,276	-7,238	,000
z27	-1,056	,116	-,341	-9,109	,000
z28	-,943	,106	-,339	-8,890	,000
z29	-,723	,116	-,233	-6,233	,000
z30	-1,341	,106	-,483	-12,645	,000
z31	-1,144	,116	-,369	-9,867	,000
z32	-,607	,116	-,196	-5,239	,000

4.2.2 Relazione tra prezzo al metro quadrato e distanza

Successivamente, sono stati realizzati dei modelli per studiare anche la relazione tra i prezzi al metro quadrato e le distanze individuate.

In primo luogo, si è studiato l'impatto sul prezzo al metro quadrato della distanza dal Duomo. Avendo scelto la forma log-lineare, il modello è del tipo:

$$\ln(p_{mq}) = a + \beta \ln(\text{Duomo_km}).$$

Anche questa variabile è significativa e porta a un modello predittivo abbastanza buono, con un R^2 corretto = 0,685. Il risultato, mostrato nella Tabella 4.28, indica un coefficiente β negativo per la variabile in esame. Ciò significa che i prezzi al metro quadrato diminuiscono all'aumentare della distanza dal Duomo di Milano. Anche in questo caso, quindi, vengono confermate le attese e la teoria di riferimento.

Tabella 4.27 Distanza dal Duomo. Forma log-lineare

Modello	R	R- quadrato	R- quadrato adattato	Errore std. della stima
1	,829 ^a	,687	,685	,24392

Tabella 4.28 Coefficienti forma log-lineare

Modello		Coefficienti non standardizzati		Coefficienti standardizzati	t	Sign.
		B	Errore standard	Beta		
1	(Costante)	9,146	,036		252,254	,000
	lnDuomo_km	-,569	,027	-,829	-20,840	,000

In seguito, è stata aggiunta al modello anche la distanza dalla fermata della metropolitana. Il risultato mostra un leggero miglioramento dell' R^2 corretto che passa da 0,685 a 0,694 (Tabella 4.29). Anche in questo caso, i coefficienti β sono negativi (Tabella 4.30) e indicano la riduzione dei prezzi all'aumentare della distanza. Il modello è dunque in questo caso:

$$\ln(p_{mq}) = 9,098 - 0,55 \ln(\text{Duomo_km}) - 0,052 \ln(\text{MM_km}).$$

Tabella 4.29 Distanza dal Duomo e dalla fermata della metropolitana. Forma log-lineare

Modello	R	R- quadrato	R- quadrato adattato	Errore std. della stima
1	,835 ^a	,697	,694	,24070

Tabella 4.30 Coefficienti. Forma log-lineare

Modello		Coefficienti non standardizzati		Coefficienti standardizzati	t	Sign.
		B	Errore standard	Beta		
1	(Costante)	9,098	,040		225,020	,000
	InDuomo_km	-,550	,028	-,800	-19,602	,000
	InMM_km	-,052	,021	-,103	-2,518	,013

Si è poi, successivamente, introdotta anche la variabile relativa alla distanza dall'aeroporto di Linate, facendo aumentare ancora l' R^2 corretto a 0,713 (Tabella 4.31). In questo caso, il coefficiente β della distanza dall'aeroporto ha segno positivo, quindi i prezzi al metro quadrato aumentano all'aumentare della distanza da Linate. Il modello diventa quindi:

$$\ln(p_{mq}) = 8,703 - 0,579 \ln(\text{Duomo_km}) - 0,047 \ln(\text{MM_km}) + 0,196 \ln(\text{Linate_km}).$$

Tabella 4.31 Distanze. Forma log-lineare

Modello	R	R- quadrato	R- quadrato adattato	Errore std. della stima
1	,847 ^a	,717	,713	,23293

Tabella 4.32 Coefficienti. Forma log-lineare

Modello		Coefficienti non standardizzati		Coefficienti standardizzati	t	Sign.
		B	Errore standard	Beta		
1	(Costante)	8,703	,111		78,182	,000
	InDuomo_km	-,579	,028	-,844	-20,513	,000
	InMM_km	-,047	,020	-,093	-2,337	,020
	InLinate_km	,196	,052	,150	3,790	,000

4.2.3 Modello di regressione con più variabili

Dopo aver provato a inserire diverse variabili, si è giunti a un modello con $R^2=0,804$ (Tabella 4.34). Le variabili inserite, le cui statistiche descrittive sono indicate nella Tabella 4.33, riguardano sia le distanze, sia le zone, sia alcune caratteristiche dell'abitazione. Il numero di residenze considerate è 188 perché, con il metodo listwise, sono state eliminate le unità con più dati mancanti.

Come si nota dalla Tabella 4.35, i prezzi al metro quadrato diminuiscono all'aumentare della distanza dal Duomo di Milano e dalla fermata della metropolitana e, invece, aumentano all'aumentare della distanza dall'aeroporto di Linate.

Inoltre, è interessante sottolineare che il prezzo al metro quadrato diminuisce all'aumentare della superficie dell'abitazione. Ciò significa che, a parità di altre condizioni, una casa grande ha un prezzo al metro quadrato inferiore rispetto a una casa piccola.

L'attico, rispetto all'appartamento, fa aumentare i prezzi al metro quadrato, così come la presenza di più bagni, di una classe energetica A e di una cantina.

Come ci si aspettava, un'abitazione da ristrutturare ha invece prezzi al metro quadrato inferiori.

Per quanto riguarda le zone, risulta interessante notare che, in questo modello, la zona 1 (il Centro) ha un coefficiente β negativo, quindi i prezzi al metro quadrato diminuiscono. Essi aumentano, invece, nella zona 5 (Garibaldi, Moscova, Porta Nuova), poiché è un'area di tendenza, per poi diminuire nuovamente nelle altre quattro zone inserite nella regressione.

Il modello così ottenuto è il seguente:

$$\ln(p_{mq}) = 9,193 - 0,656\ln(\text{Duomo_km}) - 0,035\ln(\text{MM_km}) + 0,193\ln(\text{Linate_km}) - 0,121\ln(\text{mq}) + 0,185(\text{attico}) + 0,089(\text{nbagni}) + 0,159(\text{classenergA}) + 0,097(\text{cantina}) - 0,131(\text{daristrutt}) - 0,35(z01) + 0,166(z05) - 0,212(z15) - 0,289(z24) - 0,073(z30) - 0,147(z31)$$

Tabella 4.33 Statistica descrittiva delle variabili considerate

	Media	Deviazione std.	N
Inp_mq	8,4999	,42577	188
InDuomo_km	1,1499	,63184	188
InMM_km	-,5209	,85179	188
InLinate_km	2,2016	,32206	188
Inmq	4,5233	,46960	188
attico	,06	,235	188
nbagni	1,47	,674	188
classenergA	,10	,302	188
cantina	,63	,483	188
daristrutt	,11	,309	188
z01	,09	,288	188
z05	,04	,190	188
z15	,03	,161	188
z24	,02	,126	188
z30	,02	,145	188
z31	,02	,145	188

Tabella 4.34 Forma log-lineare. Modello con più variabili

Modello	R	R- quadrato	R- quadrato adattato	Errore std. della stima
1	,906 ^a	,820	,804	,18832

Tabella 4.35 Coefficienti. Forma log-lineare

Modello	Coefficienti non standardizzati		Coefficienti standardizzati	t	Sign.
	B	Errore standard	Beta		
1 (Costante)	9,193	,205		44,739	,000
InDuomo_km	-,656	,039	-,974	-16,813	,000
InMM_km	-,035	,018	-,070	-1,943	,054
InLinate_km	,193	,057	,146	3,372	,001
Inmq	-,121	,048	-,133	-2,513	,013
attico	,185	,065	,102	2,840	,005
nbagni	,089	,032	,141	2,799	,006
classenergA	,159	,049	,113	3,231	,001
cantina	,097	,032	,110	3,035	,003
daristrutt	-,131	,048	-,095	-2,759	,006
z01	-,350	,074	-,236	-4,714	,000
z05	,166	,077	,074	2,148	,033
z15	-,212	,087	-,080	-2,435	,016
z24	-,289	,126	-,085	-2,294	,023
z30	-,073	,108	-,025	-,676	,500
z31	-,147	,101	-,050	-1,448	,149

Si rende necessario evidenziare il coefficiente negativo della zona 1 (Centro). Teoricamente, avvicinandosi al centro i prezzi dovrebbero essere maggiori, quindi il coefficiente dovrebbe essere positivo. Si tratta, invece, di un tentativo del modello di correggere la sovrastima dei prezzi in questa zona. Infatti, se si considera solo l'accessibilità, il prezzo al metro quadrato nella zona 1 risulta sovrastimato.

Ad esempio, per semplificare, si consideri solamente una variabile: la distanza dal Duomo. Il modello sarebbe del tipo $\ln(p_{mq}) = a - \beta \ln(\text{Duomo_km})$ e, inserendo i coefficienti β trovati, diventa:

$$\ln(p_{mq}) = 9,193 - 0,656 \ln(\text{Duomo_km})$$

Ciò significa che, nella zona 1 il prezzo al metro quadrato è $e^{9,193} = 9828$ euro e che, allontanandosi dal centro, se la distanza raddoppia, il prezzo diminuisce del 65%, perché il prezzo al metro quadrato è:

$$p_{mq} = \frac{e^{9,193}}{(\text{Duomo_km})^{0,656}}$$

Pertanto, se la distanza dal centro è 2 km, il prezzo al metro quadrato è:

$$p_{mq} = \frac{e^{9,193}}{2^{0,656}} = 6237 \text{ euro.}$$

Se ci si avvicina a 1 km dal Duomo, il prezzo al metro quadrato è, come visto in precedenza, quello della zona 1:

$$p_{mq} = \frac{e^{9,193}}{1^{0,656}} = 9828 \text{ euro}$$

Avvicinandosi ancora al Duomo, però, si può notare che il prezzo al metro quadrato è sovrastimato, perché aumenta notevolmente al diminuire della distanza. Infatti,

a 500 metri dal Duomo il prezzo diventa: $p_{mq} = \frac{e^{9,193}}{0,5^{0,656}} = 15.487$ euro circa

e a 100 metri dal Duomo il prezzo è: $p_{mq} = \frac{e^{9,193}}{0,1^{0,656}} = 44.510$ euro circa.

Questi prezzi al metro quadrato non sono quelli realmente osservati. Pertanto, il modello tenta di correggere questi valori sovrastimati.

Infine, trattandosi di un modello log-lineare, è importante ricordare che i coefficienti stimati β indicano l'elasticità del prezzo al metro quadrato dell'abitazione ai singoli attributi considerati. Ciò significa che, ad esempio, il prezzo al metro quadrato è il 18,5% in più se si tratta di un attico invece che di un appartamento, aumenta dell'8,9% se l'abitazione ha un bagno in più e vi è un premio del 15,9% per la classe energetica A. Il prezzo al metro quadrato è, invece, il 13,1% in meno se la casa è da ristrutturare.

CONCLUSIONI

Si può quindi concludere il presente lavoro con alcune considerazioni finali sull'analisi effettuata.

Il metodo dei prezzi edonici è uno strumento rilevante ai fini di questo studio e le analisi di regressione hanno consentito di ricavare informazioni utili. L'analisi si è pertanto basata su questo modello per studiare l'influenza sul prezzo al metro quadrato delle abitazioni delle varie caratteristiche localizzative e strutturali individuate.

Dall'analisi sono emersi risultati abbastanza in linea con ciò che ci si aspettava dalla teoria e dalla letteratura scientifica: i prezzi delle case dipendono dagli attributi dell'abitazione stessa, dell'edificio, del contesto locale e dell'accessibilità. Inoltre, essendo Milano una città con ancora forti caratteristiche monocentriche, il modello delle abitazioni ispirato alla teoria della rendita fondiaria di Von Thünen assume un ruolo fondamentale, dato che i risultati hanno mostrato l'importanza della variabile relativa alla distanza dal Duomo. Pertanto, l'accessibilità al centro, approssimata dalla distanza dal centro, è fondamentale per la città di Milano.

Si ritiene importante evidenziare, inoltre, la distanza dall'aeroporto che, a differenza delle altre distanze, risulta di segno opposto. Ciò può essere dovuto al fatto che esso si trova nella periferia e che, quindi, all'aumentare della distanza da esso ci si avvicina, invece, al centro. Inoltre, il risultato può dipendere anche dal fatto che un aeroporto, pur essendo importante dal punto di vista dell'accessibilità, produce esternalità negative (es. rumori, traffico nelle vicinanze, ...) e, pertanto, come evidenziato dai risultati, i prezzi al metro quadrato delle abitazioni aumentano all'aumentare della distanza dall'aeroporto.

Risulta interessante notare anche la relazione tra prezzi e superficie. Infatti, i risultati hanno mostrato che i prezzi al metro quadrato diminuiscono all'aumentare della superficie. Ciò significa che le case piccole hanno un prezzo al metro quadrato maggiore delle abitazioni più grandi.

Inoltre, si ritiene rilevante sottolineare il valore del coefficiente della zona 5 che, a differenza di quanto ci si potrebbe aspettare, risulta positivo. In questo caso, infatti, i prezzi aumentano nonostante la zona non sia nel centro. Ciò è dovuto al fatto che la zona di corso Garibaldi è un'area di tendenza e, pertanto, i prezzi sono maggiori. In questo caso si può notare l'importanza delle caratteristiche del contesto locale

che sono state sintetizzate all'interno delle variabili rappresentate dalle zone. Risulta qui evidente che i prezzi, quindi, non dipendono solamente dalla distanza dal centro, ma anche dalle particolarità della zona.

Infine, ma non per importanza, si è vista la capacità del modello di correggere la sovrastima dei prezzi al metro quadrato della zona del Centro città inserendo, in questo caso, un coefficiente β negativo.

RIFERIMENTI BIBLIOGRAFICI

Nadia Donati, *Tesi di Laurea. Il mercato delle residenze a Trento: una valutazione basata sui prezzi edonici*, Università degli Studi di Trento, 1990/1991

Alan W. Evans, *Economia Urbana*, il Mulino, 1985

Jean-Marie Huriot, Lise Bourdeau-Lepage, *Economie des villes contemporaines*, Economica, 2009

Arthur M. Sullivan, *Urban Economics*, Irwin, 1990

SITOGRAFIA

Scott Crosier, *Johann-Heinrich von Thünen: Balancing Land-Use Allocation with Transport Costs*, CSISS, 2001

Edi Defrancesco, Samuele Trestini, *Il ruolo del nome del vitigno nella valorizzazione di un vino sul mercato al consumo: primi risultati di un'indagine sul Tocai*, *Agriregionieuropa* anno 4 n°12, Mar 2008

Sofia Massarotto, *Tesi di Laurea. Prezzi edonici: teoria e studi empirici applicati alla valutazione di opere d'arte e manoscritti musicali*, Università degli Studi di Padova, 2017/2018

Luca Stanca, *Le determinanti dei prezzi delle abitazioni: aspetti microeconomici*, 2008

RINGRAZIAMENTI

Vorrei, in primo luogo, ringraziare il Professor Gianluigi Gorla per la sua disponibilità e per il suo aiuto nella stesura del presente lavoro. Non esistono parole per esprimere la mia gratitudine per il tempo e l'attenzione che egli ha dedicato al mio lavoro e, soprattutto, per la passione che dimostra nel suo lavoro e per condiviso con me parte della sua conoscenza.

Un ringraziamento particolare va anche alla mia famiglia, senza la quale tutto ciò non sarebbe stato possibile. Grazie per avermi sostenuta anche in questo periodo di distanza e, per questo, rendermi fiera di voi. Spero di poter rendere voi fieri di me.

Ringrazio Jonathan, che ha sempre creduto in me e mi è rimasto accanto anche quando credevo di non farcela e lo sconforto mi rendeva insopportabile.

Infine, grazie anche a tutti i miei amici, quelli di sempre e quelli che ho incontrato durante questo percorso, che aspettano il momento di poterci riabbracciare e di festeggiare insieme a me questo traguardo.