

**La valutazione degli attributi di un
parcheggio pubblico:
un'analisi del caso di un Comune del
Canton Ticino con l'utilizzo di un
esperimento a scelta discreta**

Roberto Stoppa - www.tiresia.ch
Bellinzona, settembre 2007

Indice

1	Introduzione	5
2	La metodologia e il quadro teorico	7
2.1	Le analisi che si basano sui modelli di scelta	7
2.2	I modelli di scelta (Choice Models)	7
2.3	I dati delle preferenze dichiarate (SP data)	8
2.4	Il design	8
2.4.1	I design fattoriali pieni (full) o parziali (fractional)	9
2.5	L'esperimento di scelta con preferenze dichiarate (Stated Choice Experiment)	10
2.5.1	Il design utilizzato per l'esperimento SP	11
3	Il modello teorico e il modello econometrico	14
3.1	Il modello teorico: il Random Utility Model (RUM)	14
3.2	Il modello econometrico: il Logit Multinomial (MNL)	17
4	I dati	19
4.1	La raccolta dei dati	19
4.2	Il piano di campionamento e il campione	19
4.3	Le risposte alle domande generali sui parcheggi	20
4.4	Le statistiche descrittive	21
4.5	Le variabili	21
4.5.1	La variabile spiegata	21
4.5.2	Le variabili esplicative	22
5	I risultati	22
5.1	La stima del modello Logit Multinomial con tutti gli attributi e tutti i livelli dell'esperimento	22
5.2	Il modello delle previsioni dei comportamenti individuali con tutti gli attributi e i livelli dell'esperimento	24
5.3	Multinomial Logit vs. Mixed Logit: un'evidenza empirica	25
5.4	La disponibilità a pagare: un confronto fra il modello Logit e il Mixed Logit	28
6	Conclusioni	28

Non abitiamo il mondo, ma la sua descrizione. [Immanuel Kant (1724-1804)]

Abstract

The aim of the following analysis is to evaluate the willingness to accept a public parking with some specific characteristics. This came as a reaction to the proposal by the Municipio of a Ticino's village that suggested to build an economic evidence of the acceptance of the new parking policy among the population. In order to collect data, we applied a **Stated Choice Experiment**. Such method allows the citizens' preferences to be elicited by a choice process.

Data were collected in May 2007 through a **mail survey** (with the help of a questionnaire). Citizens were selected with a **simple random sample stratification procedure** and we sent a questionnaire to 710 of them (the percentage of answer was 55% that is, 387 individuals took part to the analysis and 270 of them did the choice experiment). Each questionnaire contained 25 questions, five of which concerned the parking problem. Two of these five questions were dedicated to the **Stated Choice Experiment**. In each of the two choice exercises people had to choose between two **unlabeled alternatives**, either Alternative A or Alternative B. Each alternative was described through three **attributes**: *cost* (in Swiss Francs per month) *distance* (in walking minutes) and *availability* (the probability to find a parking). Attribute *cost* allowed three **levels** while *distance* and *availability* only two. This means that a **full factorial design** involved twelve alternatives. Considering that an experiment with a mail survey might be a great challenge, we restricted the design. We therefore used a **fractional factorial design** including only two choice experiments in each questionnaire. In so doing individuals were asked to make a choice between two alternatives, which means that the **Choice Set** was composed by two alternatives.

The results of our analysis of the **willingness to accept a parking**, based on a stated preferences (SP) dataset, collected in a Ticino's village, was estimated using a **Multinomial Logit** and a **Mixed Logit** models with the latter capable of **accomodation random heterogeneity** in citizens' tastes. For the analysis of the accomodation of the heterogeneity only two attributes (cost and distance) were considered. Comparing the two models, this accomodation led to significant conclusions. The estimation of a Mixed Logit model requires a computational process that we solved using **Biogeme**, the Bierlaire's Optimization Toolbox for GEV Model Estimation.

Citizens are willing to pay up to Chf. 2.70 in order to save one minute walking time to reach the parking from their home residence. The **willingness to pay (WTP)** is different between people who live downtown (*nucleo*) and people who live outside.

1 Introduzione

Lo sviluppo di una politica dei parcheggi pubblici efficiente, oltre ad essere una componente importante per la pianificazione di una zona, sia essa una città, un paese (come nel nostro caso un Comune del Canton Ticino), un'area industriale, ecc., può essere considerato un elemento che riveste un certo peso quando si parla di qualità della vita, soprattutto in una zona di tipo residenziale.

Una petizione promossa da un gruppo di cittadini ha chiesto al Municipio locale di rivedere la futura politica dei parcheggi comunali la quale si basa su un modello di gestione dei posteggi ad anelli concentrici ponendo più o meno limitazioni alla sosta a dipendenza dell'ubicazione del posteggio rispetto alla zona centrale del nucleo. Generalmente per i nuclei vi è la preferenza ad usare un parcheggio in base a due fattori: la sua ubicazione rispetto al nucleo (diretto influsso sulla comodità) e la durata di sosta, rispettivamente il costo.

L'introduzione del nuovo regolamento comunale implicherebbe di fatto un'applicazione delle tariffe in funzione degli anelli, dove la zona maggiormente tariffata risulta quella centrale del paese, vale a dire il nucleo.

Il ragionamento effettuato dal Municipio si basa sul principio che un parcheggio pubblico non è un bene pubblico in quanto non presenta le particolarità dell'escludibilità e della rivalità nel consumo. In effetti la presenza o meno di queste due caratteristiche (escludibilità e rivalità nel consumo) dà luogo ad una classificazione dei beni nel seguente modo:

Tabella 1: Classificazione dei beni

		Rivalità	Rivalità
		si	no
Escludibilità	si	Beni privati	Beni di club
Escludibilità	no	Risorse collettive	Beni pubblici

I beni pubblici non sono né escludibili né rivali nel consumo. Il problema principale relativo ai beni pubblici è quello del *free rider*, cioè di un individuo che gode dei vantaggi del bene senza pagarne il prezzo. Quello del *free rider* è, d'altronde, un comportamento razionale in presenza di non escludibilità.

Come conseguenza vi è una esternalità nel senso che chi fornisce un servizio pubblico (o un bene pubblico) crea un'esternalità a favore di chi ne gode senza pagare.

Nel decidere se fornire un bene pubblico, l'ente pubblico deve stimare il valore monetario che i cittadini attribuiscono al bene o servizio pubblico. Una soluzione economica per risolvere tali problemi potrebbe essere quella della vendita all'asta.

In questa nostra analisi abbiamo voluto raccogliere delle informazioni (dati) fra i cittadini del Comune per capire qual è il loro "comportamento nel consumo del bene parcheggio".

Per raccogliere queste informazioni abbiamo applicato un modello disaggre-

gato di analisi con l'ausilio di un esperimento chiamato *Stated Preferences*¹.

Dai dati raccolti ci si potrebbe perciò aspettare che, come in molti altri campi d'analisi, ci siano delle differenze importanti fra i decisori (vale a dire i cittadini del Comune) nel loro processo di scelta in funzione a cambiamenti nei vari attributi di una data alternative. Queste differenze di scelta sono conosciute come *variazione nel gusto (taste variation): their existence signals a departure from a purely homogeneous population of decision-makers* (Hess, Polak, (2004)). Se una parte di questa variazione può essere spiegata in modo deterministico (per esempio suddividendo la popolazione in gruppi mutualmente esclusivi) in molte situazioni esiste anche una variazione puramente casuale nei gusti all'interno dei gruppi.

Le caratteristiche relative all'ubicazione del parcheggio comunale rispetto al luogo di domicilio, così come quelle del parco veicoli e della disponibilità di parcheggi privati, rappresentano alcuni dei fattori che possono determinare l'eterogeneità nelle preferenze espresse dagli individui. Considerare tale eterogeneità nella stima dei parametri di un modello per l'analisi delle preferenze consente di ottenere una rappresentazione più realistica del fenomeno in esame. In questa nostra analisi cercheremo di "catturare" questa eterogeneità nella scelta applicando un modello a parametri casuali (**Mixed Logit**). L'eterogeneità nei parametri viene trattata con l'assunzione che tali parametri non siano fissi ma seguano una determinata distribuzione² nella popolazione. In questa nostra analisi faremo dapprima un approfondimento analitico del modello **Multinomial Logit** il quale è il punto di partenza per capire il modo in cui i modelli a parametri casuali permettono di rappresentare l'eterogeneità che caratterizza le preferenze. A differenza del Mixed Logit, il Multinomial Logit assume che i parametri siano fissi e che vi sia dunque omogeneità nei gusti (Train, 2003).

Con il metodo di raccolta dati che si basa sulle preferenze dichiarate (Stated Preferences), che permette di ottenere un certo numero di scelte per ogni individuo, la disponibilità di un parcheggio pubblico viene messa in relazione con il concetto di utilità ricevuta. L'assunto è che il parcheggio pubblico determina un certo livello di utilità per il cittadino che dipende dalle caratteristiche dello stesso parcheggio. L'ipotesi è che: maggiore è l'utilità associata ad un parcheggio maggiore sarà la probabilità che un individuo gli assegni la propria preferenza. È per questo che si parla di *indagini di preferenze dichiarate (Stated Preferences)* le quali prevedono la somministrazione di interviste ad individui circa le loro preferenze tra diverse opzioni del bene o del servizio al fine di stimarne le funzioni di utilità.

¹In inglese viene chiamato SP dataset

²Sta al ricercatore scegliere la forma distributiva per ottenere tali parametri con una procedura di stima. Le distribuzioni che vengono solitamente prese in considerazione sono: la uniforme (U), la triangolare (T) la normale (N) e la lognormale (LN).

2 La metodologia e il quadro teorico

2.1 Le analisi che si basano sui modelli di scelta

Come evidenziato nell'introduzione, l'analisi effettuata mira a valutare le preferenze dei cittadini di un Comune del Canton Ticino dei diversi attributi che caratterizzano la scelta di un parcheggio pubblico. Per soddisfare questo obiettivo abbiamo usato un'indagine di tipo congiunta (*Conjoint Analysis*). Questa tecnica viene applicata in diversi campi di ricerca e, a partire dagli anni '70, l'impiego si è esteso da ambiti relativamente semplici da analizzare a tematiche sempre più complesse, per esempio: per l'analisi delle preferenze dei servizi di trasporto, della conservazione di beni pubblici, della somministrazione di servizi di assistenza sanitaria, delle diverse tipologie di attività turistiche, ecc.

Questa tecnica consente di stimare l'importanza relativa che le caratteristiche di un bene o di un servizio rappresentano per gli individui. Infatti l'analisi congiunta si basa sull'assunzione che un bene o un servizio possono essere scomposti in diversi componenti (chiamati attributi) affinché ne possa venir valutato il beneficio che ognuno genera agli individui (teoria della domanda lancasteriana)³.

Per valutare l'importanza di queste caratteristiche occorre però avere a disposizione dei dati. Da parte nostra abbiamo dovuto raccogliere i dati direttamente "sul campo" chiedendoli ai cittadini del Comune. Per ottenerli abbiamo utilizzato dei modelli di scelta (Choice Models (CM), che hanno permesso di raccogliere dei dati Stated Preferences (SP) i quali sono stati organizzati tramite un esperimento che faceva riferimento ad un design. Il tutto si basa su un concetto teorico ben definito chiamato Random Utility Model (RUM).

2.2 I modelli di scelta (Choice Models)

I Choice Models (CM) si collocano nell'ambito dei metodi di analisi della domanda del consumatore. Questi modelli vengono utilizzati in settori e campi di indagine differenti. In particolare, sono stati applicati nelle scienze politiche, con studi sulle votazioni in diverse elezioni, in economia dei trasporti con analisi su alternativi mezzi di trasporto ecc. Gli approcci Choice Models sono basati sull'idea che ogni bene può essere ben descritto e rappresentato nei termini dei suoi attributi, o caratteristiche, e i livelli che essi assumono. Ciò significa ad esempio, che un lago può essere rappresentato con riferimento alla qualità delle acque, alle caratteristiche ecologiche e a quelle di accessibilità. Queste caratteristiche vengono chiamate attributi. I CM sono in grado di fornire informazioni di rilievo con riferimento ai valori dei beni non di mercato, come nel nostro caso. In primo luogo delineano quali siano gli attributi significativi in termini di valore che viene attribuito dagli individui al bene in questione.

L'ambito di ricerca relativo alla valutazione economica di parcheggi pubblici non è molto diffuso anche se come evidenziano Hess e Polak (2004): *Parking policy is an important component of contemporary travel demand management*

³Per esempio nel caso di una valutazione nel campo dell'economia del turismo potrebbero essere utilizzati gli attributi: prossimità di un'attrazione primaria, presenza di un'area naturalistica protetta, costo dell'albergo, ecc.

policies. The effectiveness of many parking policy measures depends on influencing parking type choice, so that understanding the factors affecting these choices is of considerable practical importance. Yet, academic interest in this issue has been, at best, intermittent.

Gli autori Axhausen e Polak (1991) hanno effettuato uno studio SP per analizzare il tipo di parcheggio e anche Van der Goot (1982) e Hunt (1988) hanno effettuato studi per valutare la tipologia di parcheggi.

La nostra analisi non ha lo scopo di valutare i differenti tipi di parcheggi ma vuole analizzare quali sono gli attributi importanti nella decisione di utilizzo di un parcheggio pubblico così come la disponibilità dei cittadini a pagare per usufruire di un parcheggio pubblico.

2.3 I dati delle preferenze dichiarate (SP data)

Le indagini Stated Preferences fanno riferimento a quella famiglia di tecniche di raccolta dati che prevedono di intervistare gli individui per capire le loro preferenze riguardo un insieme di alternative diverse, con lo scopo di stimare una funzione di utilità. Le alternative non sono altro che descrizioni di beni o servizi che si differenziano fra loro per i livelli delle caratteristiche che li compongono. Le alternative sono perciò una sorta di situazioni costruite ad hoc dal ricercatore. Agli individui viene perciò chiesto di esprimere una propria scelta dichiarando le sue preferenze circa le alternative che gli vengono proposte. Esistono tre sistemi per esprimere le proprie preferenze: ordinare le alternative (ranking), assegnare un valore alle varie alternative (rating) oppure semplicemente scegliere l'alternativa preferita (choice). Considerando che uno dei nostri obiettivi è quello di valutare la disponibilità a pagare da parte dei cittadini per avere a disposizione un parcheggio pubblico (bene non di mercato) e sul quale non esistono dati da poter analizzare, con la tecnica delle Stated Preferences possiamo raccogliere i dati necessari sottoponendo delle ipotetiche alternative di scelta agli individui. Hensher, Rose e Greene (2005) evidenziano che: *If life were kind, then the alternative paradigm, that of SP data, would offer only advantages, thus making the choice of data paradigm to employ easy (in contrast to RP data). Unfortunately, life is not so kind and therefore SP data could have weaknesses.* Per esempio un elemento abbastanza delicato dei dati SP è quello: *... SP data represents choices "made" or stated given hypothetical situations ... this may lead to situations in which personal constraints are not considered as constraints at the time of choice (Sure, I'll take two Ferraris). The task of the analyst is therefore to make the hypothetical scenarios as realistic as possible.*

2.4 Il design

Il design rappresenta quell'insieme di regole che, in un'analisi di tipo congiunta e più precisamente di tipo choice (e non ranking o rating):

- selezionano i livelli e gli attributi da utilizzare nella descrizione delle alternative;
- selezionano le alternative da proporre per ogni esercizio di scelta.

Esistono diversi design, a dipendenza del numero di attributi, livelli e alternative impiegati nell'indagine. L'efficienza di un design è inversamente proporzionale alla varianza dell'utilità parziale degli attributi che esso consente di

stimare, ovvero: più il design è efficiente, più precisa sarà la stima degli attributi e/o dei livelli.

Il design presenta le proprietà dell'ortogonalità e del bilanciamento dei livelli.

Un design è ortogonale quando la probabilità che si verifichi qualunque combinazione di un attributo a 2 livelli e uno di 3 livelli è pari a un dodicesimo. L'ortogonalità del design consente di evitare che le preferenze degli intervistati dipenda dalla maggiore probabilità che un'alternativa ha di essere estratta rispetto ad un'altra. In altre parole la probabilità che sia scelta l'alternativa A piuttosto che la B dipende unicamente dal fatto che l'individuo preferisce A a B e non dal fatto che l'alternativa A ha una maggiore probabilità di essere estratta dall'insieme delle alternative possibili rispetto a B.

Il bilanciamento dei livelli prevede invece che ogni livello di ogni attributo sia presente negli esercizi di scelta lo stesso numero di volte. Per esempio se un attributo è definito da 4 livelli, ognuno dovrà comparire un quarto delle volte per ogni esercizio di scelta.

A volte risulta impossibile rispettare contemporaneamente ortogonalità e bilanciamento dei livelli: se ad esempio un attributo ha un numero di livelli pari a 3 e un altro attributo ha un numero di livelli pari a 2 e il numero di esercizi di scelta non è un multiplo di 6 ecco che il bilanciamento non è rispettato.

Come alcuni autori hanno evidenziato, non sempre l'ortogonalità del design garantisce la massima *efficienza* delle stime. Rose e Bliemen (2004) sottolineano come i casi di mancata risposta alterino l'ortogonalità del design generando problemi di *efficienza*.

Hensher e Truong, (1983), inoltre, sottolineano l'importanza di considerare il realismo degli esercizi di scelta nel definire il design, in particolare eliminando quelle alternative che risultano inverosimili o insensate.

2.4.1 I design fattoriali pieni (full) o parziali (fractional)

È possibile distinguere due tipi di design: il design *full factorial* e quello *parziale*. Nel design full, l'insieme delle alternative usate nell'esperimento è dato da tutte le combinazioni possibili dei livelli di tutti gli attributi indagati. Questo tipo di design consente di analizzare sia gli effetti principali (*main effect*) che tutti gli effetti di interazione 2 o 3 livelli di interazione (*two or three-way interaction effects*). Il design parziale al contrario utilizza solo una parte di tutte le possibili combinazioni dei livelli degli attributi, permettendo di analizzare solo gli effetti principali ed eventualmente alcuni effetti di interazione di grado pari o superiore al secondo.

L'effetto principale (*o main effect*) relativo a ciascun attributo misura l'influenza che quell'attributo esercita sul grado di preferenza percepita dall'intervistato per l'alternativa in esame, fermi restando i livelli di specificazione degli altri attributi utilizzati per descrivere il profilo. Gli effetti di interazione (*o*

interaction effect) rappresentano invece l'influenza che due o più attributi esercitano congiuntamente sulle preferenze degli individui.

A priori è molto difficile stabilire se gli effetti di interazione sono importanti nella formazione delle preferenze degli intervistati e quindi se sia più opportuno impiegare un design fattoriale completo o frazionato. Una cosa è certa ed è stata dimostrata: *mediamente i main effects spiegano dal 70% al 90% della varianza delle osservazioni raccolte mentre i 2-way interazioni spiegano invece una percentuale variabile dal 5% al 15%* (Dawes and Corrigan 1974).

2.5 L'esperimento di scelta con preferenze dichiarate (Stated Choice Experiment)

Una volta definito il problema da analizzare e soprattutto il tipo di analisi che abbiamo voluto effettuare (esperimenti di scelta tramite un questionario postale), dobbiamo valutare quali sono le variabili che più rappresentano la scelta di un posteggio pubblico. Per questo abbiamo effettuato una ricerca della letteratura più rilevante e va subito detto che analisi riguardanti i parcheggi non sono molto diffuse. Un riferimento molto importante per noi è stato il lavoro svolto da Hess e Polak (2004).

Al fine di strutturare l'indagine occorre perciò definire una serie di alternative di parcheggio, caratterizzate ognuna da attributi ritenuti fondamentali. Le alternative vengono costruite sulla base di diversi livelli degli attributi considerati. Nel nostro caso la tabella seguente evidenzia gli attributi con i relativi livelli che abbiamo ritenuto determinanti per la scelta di un parcheggio pubblico e che saranno utilizzati negli esperimenti di scelta.

Come abbiamo già evidenziato in precedenza l'esperimento veniva effettuato tramite un questionario postale e di conseguenza abbiamo cercato di semplificarlo in quanto l'introduzione di un numero di attributi e di scelte elevato avrebbe condotto ad una "autoselezione" dei rispondenti abbastanza marcata segnando così una riduzione di dati disponibili.

In base a questa nostra considerazione abbiamo perciò optato per l'inserimento di solo due esercizi di scelta per ogni questionario che corrispondevano alla domanda 20 rispettivamente 21 dell'intero questionario.

Se la scelta degli attributi è stata fatta osservando la letteratura, la scelta dei livelli è stata effettuata in base ai colloqui che abbiamo avuto con le Autorità Comunali. In particolare per gli attributi costo e distanza. Con la nuova proposta di regolamento il Municipio proponeva una serie di tariffe per ottenere l'autorizzazione a parcheggiare per un tempo indeterminato. In base alle proposte fatte dal Municipio abbiamo optato per un costo nullo (gratis) come l'attuale situazione, Fr. 30 che rispecchia la tassa che vorrebbe far applicare il Municipio per i parcheggi del nucleo e Fr 15 che risulta una via di mezzo fra le due situazioni. Per quanto riguarda l'attributo distanza abbiamo calcolato due distanze medie di circa 300 e 600 metri le quali sono percorribili a piedi in circa 3 rispettivamente 6 minuti. A questi due attributi ne abbiamo inserito

Tabella 2: Attributi e livelli

Attributi	Livelli degli attributi
Distanza	3 minuti a piedi dal luogo di domicilio 6 minuti a piedi dal luogo di domicilio
Costo	Gratuito 15 franchi al mese 30 franchi al mese
Disponibilit�	Media Alta

un terzo che indica la probabilit  di trovare un parcheggio (probabilit  media o bassa) considerando il fatto che il Municipio non garantisce la disponibilit  di un parcheggio anche con il nuovo regolamento. L'attributo costo   espresso in franchi svizzeri al mese, l'attributo distanza in minuti a piedi fra il luogo di domicilio e il parcheggio.

2.5.1 Il design utilizzato per l'esperimento SP

Dopo aver definito gli attributi con i rispettivi livelli diventa fondamentale la struttura del design dell'esperimento di scelta. In particolare bisogner  considerare quale tipo di design utilizzare, scegliere la specificit  del modello e non da ultimo considerare la nostra idea di "calibrare al questionario postale" la dimensione dell'esperimento.

Considerando che all'aumentare del numero dei livelli e/o degli attributi aumenta in modo esponenziale il numero di alternative generate dal design fattoriale completo   opportuno tenerne conto quando bisogner  stabilire il numero di esercizi di scelta che dovranno essere sottoposti agli intervistati. Questo fatto   molto importante per la nostra analisi. Considerando che per la nostra analisi abbiamo optato per il questionario postale che conteneva in totale 25 domande, abbiamo deciso di sottoporre **2 esercizi di scelta per individuo**. Questo motivo ci ha portati ad escludere gi  in partenza l'ipotesi di utilizzare un design pieno (full factorial). Per ridurre il numero di alternative da utilizzare non abbiamo incluso negli esperimenti di scelta le alternative dominanti (caratterizzate dai livelli migliori) o dominate (caratterizzate dai livelli peggiori). Nel nostro caso queste due alternative sono la numero 1 e la numero 6. In questo modo il totale delle alternative da utilizzare sono 9. Oltre a ci  abbiamo escluso l'alternativa numero 8 che   un'interazione di terzo livello la quale non ci fornisce informazioni rilevanti per questa analisi ma soprattutto spiega una minima parte della varianza.

La Tabella 3 evidenzia il Full Factorial Design ci  quello dove vengono esplicitate tutte le possibili combinazioni di attributi e livelli. Nel nostro caso il totale delle combinazioni (o alternative)   $(3^1 \cdot 2^2) = 12$. Come evidenziato in precedenza vi   la possibilit  di "ridurre" il design togliendo degli elementi non interessanti per l'analisi e soprattutto togliendo delle combinazioni "triviali" in

un processo di scelta, vale a dire le alternative dominate e quelle dominanti. Nel nostro caso abbiamo perciò eliminato le combinazioni 1 e 6 (per motivi triviali) e 8 (per motivi di informazione limitata): rimangono 9 *alternative* che possono essere utilizzate per l'esperimento.

Tabella 3: Full factorial Design

Alt.	ID Alt.	Distanza	Costo	Disponibilità	Effetti
1	...	3 min.	gratis	alta	dominante ⁴
2	D	10 min.	Fr. 15	alta	2-interac.
3	B	3 min.	Fr. 30	alta	princ.
4	H	10 min.	gratis	media	2-interac.
5	G	3 min.	Fr. 15	media	2-interac.
6	...	10 min.	Fr. 30	media	dominata ⁵
7	C	3 min.	gratis	media	princip.
8	L	10 min.	Fr. 15	media	3-interac.
9	F	3 min.	Fr. 30	media	2-interac.
10	A	10 min.	gratis	alta	princip.
11	E	3 min.	Fr. 15	alta	princip.
12	I	10 min.	Fr. 30	alta	2 interac.

A questo punto abbiamo bisogno di definire il Choice Set vale a dire determinare quante alternative sottoporre agli individui per la loro scelta. Il nostro esperimento è effettuato sottoponendo agli individui due alternative *non etichettate* (*unlabeled*) ad ogni esercizio di scelta. Di conseguenza questo ci porta ad avere 36 accoppiamenti possibili (pairwise alternatives) come evidenziato nella Tabella 4.

Tabella 4: Coppie di alternative (Pairwise alternatives)

	A	B	C	D	E	F	G	H	I
A		A-B	A-C	A-D	A-E	A-F	A-G	A-H	A-I
B			B-C	B-D	B-E	B-F	B-G	B-H	B-I
C				C-D	C-E	C-F	C-G	C-H	C-I
D					D-E	D-F	D-G	D-H	D-I
E						E-F	E-G	E-H	E-I
F							F-G	F-H	F-I
G								G-H	G-I
H									H-I

In grassetto sono evidenziate quelle coppie di alternative che non sono state prese in considerazione nell'esperimento di scelta. In totale abbiamo utilizzato 20 coppie di alternative. Considerando che in ogni questionario erano proposti 2 esercizi di scelta abbiamo generato 10 tipi di questionari differenti. Questi 10 tipi di questionari differenti sono stati inviati in maniera proporzionale alle 710 persone (71 questionari per ogni coppia di esercizi di scelta). Nella Tabella 5 sono evidenziate le coppie di esercizi di scelta.

Tabella 5: Coppie di alternative **spedite**

Domanda 20	No. quest. inv.	Domanda 21	No. quest. inv.
A-B	71	C-D	71
G-H	71	F-I	71
A-C	71	F-D	71
A-E	71	G-D	71
A-F	71	H-D	71
A-G	71	C-E	71
B-C	71	E-H	71
F-H	71	B-D	71
I-H	71	B-H	71
B-G	71	G-I	71

Tabella 6: Coppie di alternative **ricevute**

Domanda 20	No. quest. ric.	Domanda 21	No. quest. ric.
A-B	30	C-D	30
G-H	28	F-I	25
A-C	26	F-D	23
A-E	26	G-D	25
A-F	26	H-D	26
A-G	31	C-E	29
B-C	26	E-H	25
F-H	28	B-D	29
I-H	24	B-H	25
B-G	23	G-I	17
Totale	266		254

La Tabella 6 evidenzia invece il numero di coppie di alternative ritornate. A parte per l'alternativa FI osserviamo che vi è stato un buon equilibrio nelle risposte ottenute in funzione degli esercizi di scelta considerando il tipo di esperimento che abbiamo voluto effettuare (Stated Choice con questionario postale). In questa sede non ci addentriamo nella ponderazione delle alternative ma utilizziamo i dati ottenuti per effettuare le stime del modello.

La Tabella 8 evidenzia che ogni alternativa non è etichettata: *Experiments that use generic titles for the alternatives are called unlabeled experiments. The decision as to whether to use labeled or unlabeled experiments is an important one. One of the main benefits of using unlabeled experiments is that they do not require the identification and use of all alternatives within the universal set of alternatives. A further benefit in the use of unlabeled experiments is the IID. The IID assumption impose the restriction that the alternatives used in the modeling process be uncorrelated. This assumption is less likely to be met under labeled experiments than under unlabeled experiments. To explain, we note that a label attached to an alternative acts somewhat like an attribute*

for that alternative (albeit an attribute whose level is constant across treatment combinations) (Hensher, Rose, Greene, 2005).

Un ultimo elemento da prendere in considerazione è il numero di esercizi da sottoporre agli individui: *We have here the trade-off between collecting more data, that we obtain by delivering many combinations, and the quality of data, that we obtain by reducing the dimension of the questionnaire* (Hensher, Rose, Green, 2005). Considerando quanto detto in precedenza riguardo il tipo di raccolta dati che abbiamo effettuato (questionario postale) abbiamo optato per sottoporre in modo casuale due esperimenti di scelta agli individui. La Tabelle 5, 6 e 7 evidenziano gli esperimenti "inviati", quelli "ricevuti" e le scelte che sono state effettuate. Questi dati ci permettono di stimare il nostro modello.

Tabella 7: Coppie di alternative ricevute e scelte

Alt.	Ricevute	%	Scelte	%
A = 10	136	13.0	78	57.4
B = 3	139	13.3	58	41.7
C = 7	115	11.0	70	60.9
D = 2	129	12.4	56	43.4
E = 11	81	7.8	40	49.4
F = 9	94	9.0	35	37.2
G = 5	121	11.6	62	51.2
H = 4	159	15.2	100	62.9
I = 12	70	6.7	23	32.9
Totale	1'040	100.0	520	

Tabella 8: Esempio di Choice Set

	Situazione A	Situazione B
Dist. parcheggio-domicilio	3 minuti a piedi	6 minuti a piedi
Costo parcheggio.	Fr. 30 mensili	Fr. 15 mensili
Difficoltà a trovare parcheggio.	Bassa	Bassa
LA SUA SCELTA →	Scelgo la A	Scelgo la B

3 Il modello teorico e il modello econometrico

3.1 Il modello teorico: il Random Utility Model (RUM)

L'utilizzo di questa tecnica (quella degli esperimenti di scelta) ci ha portati ad elaborare i dati mediante l'utilizzo di modelli a scelta discreta (*Discrete Choice Models*) che si basano sui principi del Random Utility Theory (RUM) vale a dire sull'idea di utilità come concetto latente (esistente nella mente dell'individuo ma che non può essere osservata direttamente dall'analista). A sostegno dei ricercatori è perciò stata introdotta una teoria della scelta probabilistica che

tiene conto dell'esistenza di svariate fonti di incertezza. L'analista non è in grado di stabilire con precisione l'utilità che una certa alternativa fornisce all'individuo intervistato ma può solo osservare un ordinamento delle alternative dal quale desumere l'influenza che i vari attributi hanno sull'utilità. Come conseguenza abbiamo che l'utilità si compone di due parti:

- l'utilità rappresentativa, definita attraverso una funzione solitamente di tipo lineare e additivo che lega i dati osservati dall'analista all'utilità dell'individuo;
- la componente stocastica vale a dire una variabile che cattura tutti quei fattori che influenzano l'utilità ma che non sono osservabili dal ricercatore.

Considerando quanto appena esposto è possibile definire la probabilità di scelta di un'alternativa come la probabilità che l'utilità di questa alternativa, per un determinato individuo, sia maggiore o uguale all'utilità delle altre alternative presenti nell'insieme delle scelte (nel Choice Set C_n).

Il RUM rappresenta perciò la base teorica sulla quale applicare i modelli a scelta discreta quali il Multinomial Logit e il Mixed Logit e con i quali cercheremo di analizzare i dati raccolti.

$$P(i|C_n) = P(U_{in} \geq U_{jn}, \forall j \in C_n) \quad (1)$$

Questa formula identifica la teoria dell'utilità casuale (Random Utility Theory) e il modello che si sviluppa è detto Random Utility Model.

In un modello a scelta binaria l'insieme di scelta (ovvero il Choice Set C_n) contiene solo due alternative (i e j) e di conseguenza un individuo (n) ha la seguente probabilità di scegliere l'alternativa i :

$$P_n(i) = P(U_{in} \geq U_{jn}) \quad (2)$$

dove U_{in} è l'utilità che percepisce l'individuo n scegliendo l'alternativa i , cioè quella che fra le due genera la maggiore utilità.

A questo punto dobbiamo evidenziare alcuni punti essenziali:

1. l'utilità è scomposta in due parti, una deterministica V e l'altra casuale ϵ ;

$$U_{in} = V_{in} + \epsilon_{in} \quad (3)$$

La parte deterministica è detta anche sistematica, rappresentativa o media dell'utilità.

2. L'analista cerca di mettere in relazione la parte deterministica dell'utilità che è formata dagli attributi che descrivono le alternative di scelta ($z_i n$) e

dalle caratteristiche individuali (S_n) che possono per esempio essere quelle socio-economiche dell'individuo:

$$V = V(z_{in}, s_n) \quad (4)$$

Mettendo assieme i due tipi di informazione otteniamo un vettore che abitualmente viene definito x_{in} . La equazione XXX diventa perciò:

$$U_{in} = V(Z_{in}, S_n) + \epsilon(Z, S) \quad (5)$$

che rappresenta la riscrittura nelle due componenti (deterministica e stocastica) della funzione di utilità.

Nel nostro caso (due alternative di scelta nel Choice Set) la probabilità di un individuo (n) di scegliere l'alternativa 1 può essere definita come:

$$P_{1n} = Prob(U_{1n} > U_{2n}) \quad (6)$$

$$P_{1n} = Prob(V_{1n} + \epsilon_{1n} > V_{2n} + \epsilon_{2n}) \quad (7)$$

$$P_{1n} = Prob(\epsilon_{2n} - \epsilon_{1n} < V_{1n} - V_{2n}) \quad (8)$$

La probabilità che la differenza dei due errori ($\epsilon_{2n} - \epsilon_{1n}$) prenda un valore inferiore alla differenza della parte osservata ($V_{1n} - V_{2n}$) può essere calcolata con una funzione di distribuzione cumulativa dopo avere specificato quali assunzioni si vogliono fare sugli errori (per esempio l'assunzione IID nel caso del Logit la quale porta ad una distribuzione di tipo Gumbel).

3. La forma funzionale più diffusa della parte deterministica dell'utilità è quella lineare nei parametri:

$$V_{in} = \sum_i \beta_i x_{jni} \quad (9)$$

4. L'assunzione sulla distribuzione della parte stocastica dell'utilità (ϵ_{in} , chiamata anche distribuzione degli errori) non è "definita automaticamente", ma dipende dal tipo di analisi che si vuole sviluppare e di conseguenza viene decisa dall'analista. Per questo motivo differenti assunzioni circa la distribuzione degli errori comportano l'utilizzo di modelli di scelta differenti. Va detto che la scelta della distribuzione del termine d'errore non è indipendente dalla specificazione di V ma essa è legata alla specificazione della parte sistematica dell'utilità⁶.

⁶Nella realtà spesso vengono omesse molte variabili nella parte sistematica in quanto l'utilizzo di un elevato numero di attributi potrebbe portare gli individui a scegliere fra alternative troppo complicate da valutare e di conseguenza la distribuzione di ϵ dipenderà dalla distribuzione congiunta delle variabili omesse

3.2 Il modello econometrico: il Logit Multinomial (MNL)

Sappiamo che il modello di scelta che può essere utilizzato dipende dalle assunzioni che si fanno sulla distribuzione della parte stocastica della funzione di utilità, vale a dire la distribuzione degli errori. Se assumiamo che ogni ϵ_{in} è indipendente e identicamente distribuito extreme value(IID) con la seguente funzione di densità:

$$f(\epsilon_{in}) = e^{\epsilon_{in}} e^{-e^{-\epsilon_{in}}} \quad (10)$$

otteniamo il modello logit. Questo passaggio è stato evidenziato da McFadden nel 1974 il quale ha messo in relazione il Random Utility Model con il modello logit. In particolare, riprendendo quanto evidenziato in precedenza, abbiamo che la probabilità che ogni individuo n scelga l'alternativa i è data da:

$$P_{in} = Prob(V_{in} + \epsilon_{in} > V_{jn} + \epsilon_{jn} \forall i \neq j) \quad (11)$$

$$P_{in} = Prob(\epsilon_{jn} - \epsilon_{in} < V_{in} - V_{jn} \forall i \neq j) \quad (12)$$

Se assumiamo che ϵ_{in} è dato, questa espressione è la distribuzione cumulativa per ogni ϵ_{in} valutato a $\epsilon_{jn} + V_{jn} - V_{in}$ vale a dire:

$$F(\epsilon_{in}) = \frac{1}{1 + e^{\epsilon_{in}}} \quad (13)$$

dove ϵ_{in} diventa $\exp(-\exp(-(\epsilon_{in} + V_{in} - V_{jn})))$. Dato che gli errori sono indipendenti, questa funzione di distribuzione cumulativa $\forall i \neq j$ è il prodotto delle distribuzioni cumulative individuali, vale a dire:

$$P_{in} | \epsilon_{in} = \prod_{i \neq j} e^{-e^{-(\epsilon_{in} + V_{in} - V_{jn})}} \quad (14)$$

Visto che ϵ_{in} , non è dato, la probabilità di scelta è perciò l'integrale dell'equazione precedente (14) per tutti i valori di ϵ_{in} ponderato per la sua densità (equazione 10):

$$P_{in} = \int \left(\prod_{i \neq j} e^{-e^{-(\epsilon_{in} + V_{in} - V_{jn})}} \right) e^{\epsilon_{in}} e^{\epsilon_{in}} d\epsilon_{in} \quad (15)$$

Risolvendo l'integrale otteniamo la seguente espressione in forma chiusa:

$$P_{in} = \frac{e^{V_{in}}}{\sum_j e^{V_{jn}}} \quad (16)$$

Nel modello logit binario, la probabilità che l'alternativa i venga scelta dall'individuo n è dunque pari a

$$P_n(i) = \frac{1}{1 + e^{(\beta' X_{in} - \beta' X_{jn})}} = \frac{e^{\beta' X_{in}}}{e^{\beta' X_{in}} + e^{\beta' X_{jn}}} \quad (17)$$

dove β' è il vettore trasposto dei parametri e X è il vettore degli attributi.

McFadden (1974) ha evidenziato che la formula del logit per le probabilità di scelta implica necessariamente che l'utilità non osservata (quella stocastica) sia distribuita extreme value.

Come evidenziato precedentemente i vari modelli di scelta (Choice Models) derivano dalle differenti assunzioni che vengono fatte sulla distribuzione degli errori, vale a dire la parte stocastica della funzione di utilità. Abbiamo evidenziato che il più semplice e nello stesso tempo utilizzato modello a scelta discreta è il logit: *ist popularity is due to the fact that the formula for teh choice probabilities takes a closed form and is readily interpretable (Train 2003)*.

Questo modello però ha dei limiti. Va precisato che non si tratta di limitazioni che escludono a priori l'utilizzo di questo modello nell'analisi dei dati raccolti. In effetti bisogna sempre considerare il contesto nel quale si vuole operare in quanto quelli che sembrano dei limiti potrebbero diventare dei "punti di forza" di questo modello. Train (2003) evidenzia che: *Three topics elucidate the power of logit models to represent choice behavior, as well as delineating the limits to that power. These topics are: taste variation, substitution patterns, and repeated choices over time.*

1. La variazione nei gusti (Taste Variation): il modello Logit assume omogeneità nei gusti come sottolineato anche da Ben-Akiva e Lerman (1985). Nella realtà sappiamo che i gusti possono variare da individuo a individuo in quanto ognuno può percepire da un determinato attributo o livello una differente "soddisfazione". Con il Logit si possono osservare soltanto i gusti che variano in modo deterministico, vale a dire quelli catturati tramite l'analisi delle variabili osservate (siano esse le variabili dell'esperimento o quelle che caratterizzano gli individui). Il Logit non permette di catturare quei gusti che non vengono espressi nelle variabili osservate o che sono semplicemente casuali. Per esempio due individui che hanno la stessa educazione e formazione e percepiscono lo stesso reddito potrebbero effettuare delle scelte differenti che rispecchiano il loro modo di pensare e di vivere. Questo limite del Logit diventa un vero problema se l'analista si aspetta che vi possono essere dei gusti che variano in funzione di variabili non osservate o semplicemente in maniera puramente casuale.
2. *Indipendenza dalle alternative irrilevanti - IIA (Substitution Patterns)*: La IIA è un'assunzione restrittiva che fa parte del modello Logit ed indica che il rapporto delle probabilità di scelta è indipendente dalla presenza o dall'assenza di altre alternative nel choice set. Non sempre questo limite può essere visto come qualche cosa di negativo (vedi Train, 2003, pagg. 52-53).
3. *Scelte ripetute (Panel Data)*: questo succede per esempio in un esperimento di scelte (stated preferences) dove agli individui viene chiesto di effettuare diversi esperimenti di scelta in modo da raccogliere per ognuno molti più dati (come nel nostro caso). Ogni situazione di scelta diventa perciò un'osservazione del dataset. Se l'utilità di ogni periodo viene fatta dipendere solo dalle variabili di quel periodo (per esempio $V_{int} = \beta' x_{int}$ dove x_{int} è un vettore di variabili che descrivono l'alternativa i la quale è stata scelta dall'individuo n nel periodo t , allora non vi è differenza fra il

modello logit con panel data e quello con i cross-sectional data. I dati che rappresentano scelte ripetute vengono chiamati anche Panel Data.

4. Considerando i limiti che abbiamo esposto precedentemente, abbiamo optato di stimare in ogni caso un modello Logit e in seguito di evidenziare empiricamente la differenza che potrebbe esserci con un'estensione di questo modello che nel nostro caso utilizza un modello Mixed Logit.

4 I dati

I dati utilizzati per questa analisi sono da definire "Stated Choice" in quanto ricavati da un esperimento di scelta (*Stated Choice Experiment*) che metteva a confronto due alternative (non etichettate - *unlabeled experiment* - semplicemente contrassegnate con A e B): il rispondente doveva dichiarare quale delle due alternative preferiva e questo per due esercizi di scelta. Le persone che hanno effettuato tale esperimento sono 270 su 387, vale a dire il 70%. Questo ci ha permesso di raccogliere un totale di 520 osservazioni.

4.1 La raccolta dei dati

I dati sono stati raccolti durante il mese di maggio 2007 tramite l'invio di un questionario postale contenente 25 domande. Gli invii sono stati effettuati a 710 cittadini scelti con il **metodo del campionamento casuale semplice stratificato** dove le variabili di stratificazione utilizzate erano l'età e gli anni di residenza nel Comune⁷.

La restituzione del questionario (*in forma anonima*) è avvenuta tramite busta allegata. Come detto, in totale sono rientrati 387 questionari validi (percentuale di risposta del 55% che per un questionario postale è un dato molto elevato). Questo sottolinea l'ottima collaborazione da parte dei cittadini nell'esporre le proprie opinioni in merito alla politica dell'amministrazione comunale.

Il questionario aveva lo scopo di raccogliere l'opinione dei cittadini su alcuni temi di politica comunale tra i quali il problema dei parcheggi. A questo tema abbiamo dedicato 5 domande specifiche (dalla domanda 17 fino alla domanda 21 compresa).

4.2 Il piano di campionamento e il campione

Per lo svolgimento del sondaggio abbiamo utilizzato la tecnica di campionamento casuale semplice stratificato la quale permette di introdurre delle informazioni ausiliarie conosciute a priori. Nel nostro caso le informazioni a priori che abbiamo utilizzato si riferiscono a due variabili: l'età e gli anni di residenza nel Comune. Teoricamente la stratificazione è uno dei migliori metodi per introdurre delle informazioni ausiliarie in un'inchiesta, al fine di aumentare la

⁷Per la scelta del campione sono stati presi in considerazione tutti i cittadini domiciliati nel Comune che avevano almeno 18 anni di età nel mese di aprile 2007

precisione degli stimatori. *Generalmente quando esistono delle informazioni a priori l'interesse dell'analista è quello di stratificare* (Tyllé, 2001).

4.3 Le risposte alle domande generali sui parcheggi

In questo paragrafo riportiamo alcune statistiche descrittive che riguardano le risposte ottenute alle domande 17 e 18 e le statistiche sulle variabili socio-economiche del campione intervistato.

La **domanda 17** chiedeva ai cittadini se utilizzano i parcheggi. Ecco le risposte:

Tabella 9: **Utilizzo dei parcheggi pubblici - domanda 17**

Si	No	Totale
258	129	387
66.6%	33.4%	100%

Con la **domanda 18** si voleva capire se chi utilizza i parcheggi (258 individui) ritiene che vi sia una sufficiente disponibilità. Ecco le risposte:

Tabella 10: **Disponibilità di parcheggi pubblici - domanda 18**

Disponibilità sufficiente	115	44.5%
Disponibilità insufficiente	138	53.4%
Non risposto	5	2.0%
Totale utilizzatori	258	100.0%

La Tabella 11 utilizza i dati della Tabella 10 suddividendo gli individui in base alla zona di residenza (abitano nel nucleo e non abitano nel nucleo).

Tabella 11: **Disponibilità di parcheggi pubblici**

Nucleo	Si	No	N. r.	Totale
Si	38.3%	60.0%	1.7%	100.0%
No	45.5%	52.0%	2.5%	100.0%
Totale	44.4%	53.2%	2.4%	100.0%

I dati che abbiamo utilizzato per questa analisi sono le risposte dei cittadini che hanno dichiarato di utilizzare i parcheggi pubblici (258 persone) più 12 cittadini che invece non hanno espresso un parere circa l'utilizzo dei parcheggi pubblici (270 individui).

4.4 Le statistiche descrittive

Il campione è composto da 141 donne (52.2%) e da 129 uomini (47.8%). Il 51.9% di dei rispondenti abita nel Comune da oltre 10 anni.

Tabella 12: Statistiche descrittive delle variabili socio-economiche

Sesso		%	Popol. %
Donne	141	52.2	51.6%
Uomini	129	47.8	48.4%

Etá		%	Popol. %
18-30	32	11.9	11.6
31-50	140	51.9	46.6
51-70	75	27.8	30.8
> 70	23	8.5	11.0

Anni residenza		%	Popol. %
0 - 5	74	27.4	26.6
6 - 10	53	19.6	15.9
>10	129	47.8	51.9
Nativi	14	5.2	5.6

Statuto		%	Popol. %
Dipendente	165	61.1	...
Indipendente	11	4.1	...
Studente	18	6.7	...
Casalinga	42	15.6	...
Pensionato	24	8.9	...
Altro	9	3.7	...

Residenza		%	Totale %
Nucleo	53	19.7	...
Non nucleo	217	80.3	...

4.5 Le variabili

4.5.1 La variabile spiegata

La variabile spiegata è di tipo binomiale vale a dire **Alternativa A** o **Alternativa B** e prende valori [1] se l'alternativa viene scelta e [0] altrimenti. La Tabella 13 evidenzia la percentuale di individui che sarebbe disposta ad accettare un parcheggio in base alle proposte effettuate. I risultati sono interessanti e coerenti nel senso che se la distanza aumenta, *a parità di condizioni*, vi è meno gente disposta ad accettare un parcheggio. Sempre a parità di condizioni, anche se il costo aumenta vi è meno gente disposta ad accettare.

Tabella 13: Percentuale di individui che accetterebbe un certo tipo di parcheggio

		Dispo. alta	Dispo. alta	Dispo. media	Dispo. media
		3 minuti	10 minuti	3 minuti	10 minuti
Costo	Gratis	...	57.4%	60.9%	59.8%
Costo	Fr. 15	49.4%	43.4%	51.2%	...
Costo	Fr. 30	41.7%	32.9%	37.2%	...

4.5.2 Le variabili esplicative

Le variabili utilizzate in questo nostro studio sono del seguente tipo:

1. **le variabili dell’esperimento:** sono state evidenziate nella tabella 2 e in principio ci siamo basati su alcuni lavori effettuati da Axhausen e Polak.
2. **le caratteristiche socio demografiche:** alcune le abbiamo raccolte direttamente presso l’Amministrazione Comunale, *l’età, gli anni di residenza nel Comune, il sesso e la zona di abitazione* (se nucleo o non nucleo). Altri dati invece li abbiamo raccolti con il questionario: *statuto professionale e luogo di lavoro*. Come sottolineano Hensher et al. (2005) *often socio-demographic characteristics are not thought to be important in utility/preference formation and the are therefore collected as a check of the sample comes from the correct population*.

5 I risultati

Qui di seguito vediamo dapprima di evidenziare i risultati che abbiamo ottenuto con la stima di un modello **Logit Multinomial con tutti gli attributi e tutti i livelli dell’esperimento**. In seguito metteremo a confronto due modelli, il **Logit Multinomial** e il **Mixed Logit considerando soltanto un coefficiente per ogni attributo**⁸.

5.1 La stima del modello Logit Multinomial con tutti gli attributi e tutti i livelli dell’esperimento

Come abbiamo visto, l’implementazione del Choice Model prevede dapprima l’osservazione delle scelte effettuate dagli individui e successivamente, in base ai dati raccolti, la stima dei parametri della funzione di utilità. Il metodo usato è quello della **massima verosimiglianza**: *Since the logit probabilities take a closed form, the traditional maximum-likelihood procedures can be applied* (Train 2003). I risultati sono riportati nella Tabella 14 (l’elaborazione è stata effettuata con il *free software packages BIOGEME*⁹).

⁸Questo ci ha costretti a modificare le colonne del dataset.

⁹BIOGEME: (Bierlaire Optimization Toolbox for GEV Model Estimation) is an object-oriented software package designed for the maximum likelihood estimation of Generalized Extreme Value (GEV) models. BIOGEME is distributed free of charge from biogeme.epfl.ch

Tabella 14: Risultati della stima 1 con il modello **Logit Multinomial**

Variabile		Coeff.	t-test	Robust Std. err.	Robust t- test
Distanza corta	Rif.
Distanza lunga	Mean	-1.228	-3.754	0.456	-2.693
Costo gratis	Rif.
Costo medio	Mean	-0.787	-2.791	0.381	-2.064
Costo alto	Mean	-0.824	-2.807	0.414	-1.988
Disponibilità sicura	Rif.
Disponibilità media	Mean	-0.681	-2.890	0.336	-2.026
Residenza ≤ 5 anni	Rif.
Residenza > 5 anni	Mean	-0.553	-2.319	0.331	-1.673
Nucleo	Mean	0.500	1.709	0.414	1.209
Costante	Mean	-1.250	-5.650	0.303	-4.130
Statistica riassuntiva					
No. individui = 267					
No. osservazioni = 518					
Parametri stimati = 7					
$L(0) = -359.05$					
$L(\hat{\beta}) = -296.01$					
$-2[L(0) - L(\hat{\beta})] = 126.07$					
$\rho^2 = 0.175$					
$\bar{\rho}^2 = 0.156$					

La forma dell'utilità che abbiamo scelto per la stima del modello Logit Multinomial è la seguente:

$$U_{A(MNL)} = Costante + \epsilon_0 \quad (18)$$

$$\begin{aligned} U_{B(MNL)} = & \beta_1 Distanza_{(lunga)} + \beta_2 Costo_{(Fr.15mese)} + \\ & + \beta_3 Costo_{(Fr.30mese)} + \beta_4 Disponib_{(media)} + \\ & + \beta_5 Residenza_{(>5anni)} + \beta_6 Nucleo_{(si)} + \epsilon_B \end{aligned} \quad (19)$$

Quale riferimento abbiamo: $Distanza_{(corta)} + Costo_{(gratis)} + Disponib_{(alta)} + Residenza_{(\leq 5anni)} + Nucleo_{(no)} + \epsilon_B$

Le stime dei coefficienti sono tutte significative ad un livello di significatività del 5% (variabile nucleo con un livello di significatività del 10%). In particolare i coefficienti stimati rappresentano il peso che *ogni livello degli attributi* assume nelle funzioni di utilità individuali; in altre parole, i coefficienti informano come le varie caratteristiche di un parcheggio pubblico influiscono sulla funzione di utilità del "consumatore" (in questo caso del cittadino che abita nel Comune).

Oltre alla significatività tutti i segni sono in accordo con l'intuizione economica di comportamento. In particolare il costo (calcolato in Fr. al mese) ha un coefficiente negativo in quanto un aumento della tassa di parcheggio genera una diminuzione di utilità agli individui. La distanza (calcolata in minuti a piedi dal luogo di domicilio al parcheggio) ha un coefficiente negativo evidenziando che più aumenta la "strada" da fare a piedi più diminuisce l'utilità degli individui. È inoltre interessante evidenziare che vi è un comportamento diverso fra chi abita il nucleo e chi invece non abita il nucleo. Il coefficiente è positivo (rispetto a chi non abita il nucleo) e questo significa che l'utilità di avere un parcheggio a disposizione è maggiore per chi abita il nucleo rispetto a chi non lo abita, *a parità di altre condizioni*. Questo è sicuramente un risultato molto interessante che spiega anche la petizione inoltrata in Municipio. Anche gli anni di residenza nel Comune sono significativi. Difficile interpretare questo coefficiente ma una possibilità d'interpretazione è che i "nuovi" cittadini del Comune hanno più bisogno di più parcheggi rispetto a chi abita da oltre 5 anni nel Comune in quanto hanno più auto rispetto al numero di parcheggi privati a disposizione.

5.2 Il modello delle previsioni dei comportamenti individuali con tutti gli attributi e i livelli dell'esperimento

Nella Tabella 15 vengono evidenziati le *previsioni individuali* della probabilità di accettare un parcheggio comunale date determinate caratteristiche dell'individuo e del parcheggio. Si tratta di probabilità medie individuali (e non aggregate) che utilizzano i dati stimati con il modello precedente.

Tabella 15: Stime delle disponibilità individuali ad accettare un certo tipo di parcheggio comunale

		Alta	Alta	Media	Media
		3 min.	10 min.	3 min.	10 min.
nucleo si	gratis	82.2%	62.8%	74.5%	46.0%
	Fr. 15	72.4%	43.4%	57.%	28.0%
	Fr. 30	71.6%	42.5%	56.1%	27.2%
nucleo no	gratis	77.7%	50.5%	63.9%	34.1%
	Fr. 15	61.4%	31.8%	44.6%	19.0%
	Fr. 30	60.5%	31.0%	43.7%	18.5%

Per esempio dalla Tabella 15) si può osservare che il 71.6% delle persone che abitano il nucleo sarebbero disposte ad accettare un parcheggio vicino a casa (massimo tre minuti a piedi), con un'alta probabilità di trovare un posto libero "sborsando" una tassa di Fr. 30 al mese. Se la probabilità di trovare un

parcheggio libero diminuisce, solo il 56.1% lo accetterebbe. Se si offrisse alle persone che non abitano il nucleo un parcheggio a 10 minuti a piedi dal domicilio con un costo di Fr. 30 al mese esso non sarebbe visto di buon occhio dalla popolazione (solo il 18.5% lo accetterebbe).

5.3 Multinomial Logit vs. Mixed Logit: un'evidenza empirica

A questo punto vogliamo approfondire da un punto di vista *empirico* una possibile estensione del modello Logit Multinomial per tentare di catturare l'eterogeneità che è implicita nei processi di scelta fra diverse alternative: *Originally, most applications were based on the use of the Multinomial Logit model (McFadden 1974), which, although it has important advantages in terms of ease of estimation, has certain disadvantages, notably in the form of flexible substitution patterns. Recently, the use of even more flexible model form, has increased dramatically, mainly thanks to improvements in the efficiency of simulation-based estimation processes, which are required when using this model form (Hess, Polak 2003).* In particolare il Mixed Logit assume che le preferenze dei soggetti varino in maniera continua nella popolazione. In altre parole questo modello permette ai coefficienti di variare in maniera puramente casuale fra gli individui vale a dire: *it allows for random taste variation across decision-makers (Hess, Polak 2003).*

Questo equivale a dire che: **in una popolazione, individui differenti hanno gusti differenti** il che è molto realistico. Questi gusti sono catturati in un processo di stima dei parametri (coefficienti) che seguono una determinata distribuzione e non sono perciò fissi (come nella stima che abbiamo fatto con il Logit Multinomial). Per stimare questi coefficienti bisogna perciò determinare quale tipo di distribuzione seguono i parametri. Questa distribuzione viene "arbitrariamente" scelta dall'analista *prima della stima del modello* e come tutte le distribuzioni avrà un valore medio (media) e una dispersione (varianza).

Le probabilità del Mixed Logit sono gli integrali del Logit Multinomial sulle densità dei parametri casuali. Formalmente la probabilità di un individuo n di scegliere l'alternativa i è data da:

$$P(n, i) = \int L_i(\beta, X_n) f(\beta|\theta) d\beta \quad (20)$$

dove X_n è una matrice degli attributi delle differenti alternative che dovrà scegliere l'individuo n . La funzione $L_i(\beta, X_n)$ rappresenta la probabilità di scelta del Logit Multinomial (questa volta condizionata su β che a sua volta è condizionato dalla distribuzione scelta a priori) che è data da:

$$L_i(\beta, X_n) = \frac{e^{\beta X_{ni}}}{\sum_{j=1}^I e^{\beta' X_{nj}}} \quad (21)$$

dove I è il numero totale delle alternative nel choice set dell'individuo. Il vettore β varia fra gli individui e riflette gli aspetti idiosincratichi delle preferenze dell'individuo e (i valori nel vettore β) sono distribuiti nella popolazione con

Tabella 16: Risultati della stima 2 con il modello **Logit Multinomial**

Variabile		Coeff.	t-test	Robust Std. err.	Robust t- test
Distanza	Mean	-1.020	-3.168	0.502	-2.034
Costo	Mean	-0.377	-2.048	0.433	-0.871
Dispon.	Mean	-0.678	-2.660	0.418	-1.621
Residenza.	Mean	-0.614	-1.883	0.791	-0.776
Nucleo	Mean	1.015	0.372	8.627	0.118
Costante	Mean	-1.308	-4.948	0.528	-2.479

Statistica riassuntiva
 No. individui = 267
 No. osservazioni = 518
 Parametri stimati = 6
 $L(0) = -359.05$
 $L(\hat{\beta}) = -297.63$
 $-2[L(0) - L(\hat{\beta})] = 122.83$
 $\rho^2 = 0.171$
 $\bar{\rho}^2 = 0.151$

densità $f(\beta|\theta)$ dove θ è pure un vettore di parametri che deve essere stimato. Esso comprende la media della popolazione e la deviazione standard del singolo coefficiente contenuto nel vettore β : *in general, two parameters are associated with each randomly distributed coefficient, representing the mean and spread in the coefficient's values across the population* (Hess, Polak 2003).

In questo nostro lavoro non intendiamo estendere la teoria di questo modello ma ci limitiamo semplicemente a fare un confronto empirico dei due modelli (**il Multinomial Logit e il Mixed Logit**) utilizzando gli stessi dati per la stima.

Per effettuare questo paragone abbiamo dapprima stimato il modello Logit Multinomial **utilizzando soltanto un coefficiente β per ogni attributo** e non sui livelli come fatto in precedenza.

Anche in questo caso i segni dei coefficienti sono coerenti con le nostre aspettative: costo, distanza e disponibilità hanno tutti segno negativo, residenza e nucleo hanno invece segno positivo. Con il modello Logit Multinomial il coef-

ficiente nucleo non è significativo mentre con il Mixed Logit tutti i coefficienti sono significativi al 5% e nucleo al 10%.

Per quanto riguarda i due coefficienti *Distanza* e *Costo*, con il modello Mixed Logit vengono stimati con l'assunzione che seguono una distribuzione di tipo normale (N). Considerando che questi parametri non risultano stimabili in forma chiusa occorre far riferimento a metodi di simulazione. La simulazione è stata effettuata con *1'000 random draws* utilizzando Biogeme.

Tabella 17: Risultati della stima 2 con il modello **Mixed Logit**

Variabile		Coeff.	t-test	Robust Std. err.	Robust t- test
Distanza (N)	Mean	-28.240	-2.134	2.419	-11.676
	Std dev.	52.831	2.273	2.645	19.974
Costo (N)	Mean	-9.941	-1.963	1.252	-7.581
	Std dev.	26.547	2.275	1.307	20.304
Dispon.	Mean.	-15.996	-2.109	1.353	-11.824
Residenza.	Mean.	-22.536	-2.129	2.541	-8.868
Nucleo	Mean	13.584	1.843	3.789	3.585
Costante	Mean	-45.828	-2.231	2.406	-19.046

Statistica riassuntiva

No. individui = 267
 No. osservazioni = 518
 Parametri stimati = 8
 $L(0) = -359.05$
 $L(\hat{\beta}) = -153.20$
 $-2[L(0) - L(\hat{\beta})] = 411.68$
 $\rho^2 = 0.573$
 $\bar{\rho}^2 = 0.548$

Il modello Mixed Logit riportato nella Tabella 17 è quello che sembra spiegare meglio le scelte effettuate in termini di verosimiglianza. Per l'attributo *costo* e per l'attributo *distanza* si è assunta una distribuzione di tipo normale (N). L'incremento nella log-verosimiglianza passando dal modello Logit Multinomial (con parametri fissi) a quello con parametri casuali è notevole (da - 297.63 a -

153.20) evidenziando così una forte eterogeneità nei processi di scelta. L'eterogeneità che abbiamo raccolto con il Mixed Logit si manifesta soprattutto in una sensibilità al costo (in Fr. al mese) e nella distanza (in minuti a piedi) che variano fra individuo e individuo.

5.4 La disponibilità a pagare: un confronto fra il modello Logit e il Mixed Logit

Il ratio fra il coefficiente della distanza e quello del costo è la misura dell'importo che il cittadino è disposto a pagare per risparmiare tempo (*a parità delle altre condizioni*). Numericamente abbiamo trovato che i cittadini del Comune sono disposti a pagare Fr. 2.85 in più al mese per avvicinare il parcheggio di un minuto verso il luogo di domicilio. Questo risultato ci sembra molto interessante in quanto l'autorità Comunale dovrebbe utilizzare anche questa informazione qualora dovesse rivedere la politica dei parcheggi comunali e in particolare l'eventuale applicazione di una tassa sui posteggi.

6 Conclusioni

L'obiettivo di questo nostro lavoro era quello di applicare un'analisi di tipo Stated Preference (chiamato anche Stated Choice) per cercare di valutare il comportamento dei cittadini di un Comune del Canton Ticino in merito alla politica dei parcheggi pubblici. Per ottenere queste informazioni abbiamo effettuato un sondaggio tra la popolazione scegliendo con il metodo del campionamento casuale semplice stratificato un campione di 710 cittadini maggiori di 18 anni. Agli individui è stato inviato in forma anonima un questionario postale con 25 domande di cui 5 riguardanti il tema dei posteggi pubblici. Due di queste domande riguardavano un esperimento di scelta. In particolare in ognuna delle due domande si sottoponevano i rispondenti ad un esercizio di scelta (choice) nei quali gli si proponevano due alternative non etichettate (*unlabeled alternatives*). L'individuo doveva scegliere l'alternativa che più preferiva (l'alternativa A o l'alternativa B). Questa scelta doveva essere fatta in base ai 3 attributi (*costo*, valutato in franchi al mese, *distanza*, valutata in minuti a piedi e *disponibilità*, valutata con probabilità elevata o media di trovare parcheggio), che componevano ogni alternativa. Gli attributi venivano fatti variare per ogni alternativa e per ogni esercizio di scelta. Da questo *Choice Experiment* abbiamo così potuto raccogliere i dati che abbiamo analizzato e stimato. La significatività statistica delle variabili dell'esperimento ci dimostra che, malgrado i nostri timori iniziali, tale esperimento ha portato buoni frutti e ci ha permesso di ottenere una risposta del comportamento economico al problema.

La prima stima che abbiamo effettuato è stata quella di un modello Logit Multinomial che prendeva in considerazione tutti gli attributi dell'esperimento (3) e tutti i livelli (7) con alcune caratteristiche socio-demografiche. Abbiamo così inserito 11 parametri di stima dove 4 li abbiamo mantenuti quale riferimento. I risultati ottenuti con questo modello sono soddisfacenti. In particolare guardando alle previsioni individuali (Tabella 15) notiamo che gran parte della popolazione accetterebbe (è disposta) a pagare per poter usufruire di un parcheggio pubblico. Questa disponibilità cresce se aumenta la sicurezza di

trovare il parcheggio e soprattutto se diminuisce la distanza da percorrere a piedi. La disponibilità a pagare è maggiore fra i cittadini del nucleo rispetto a quelli che non abitano nel nucleo e questo è un importante risultato che abbiamo ottenuto.

Dopo la stima di questo modello abbiamo voluto mettere a confronto il modello Logit Multinomial con il modello Mixed Logit. Quest'ultimo permette di catturare l'eterogeneità presente nei processi di scelta tra un numero discreto di alternative. Questo modello assume che le preferenze varino in maniera continua nella popolazione di riferimento e che siano quindi rappresentabili attraverso una funzione di densità continua.

La stima dei parametri del Mixed Logit richiede una complessità computazionale elevata che abbiamo risolto utilizzando il software Biogeme. La disponibilità a pagare (Willingness To Pay), calcolata in franchi al mese, per ogni minuto di strada a piedi risparmiato, è di Fr. 2.70 con il Logit Multinomial e di Fr. 2.85 con il Mixed Logit. L'utilizzo dei coefficienti distribuiti casualmente (randomly distributed) nel modello Mixed Logit permette di avere diverse indicazioni della variazione nei coefficienti (in questo caso distanza e costo) fra gli individui e di conseguenza del ratio.

Una ulteriore estensione dell'analisi sarebbe quella di incorporare l'intera distribuzione dei coefficienti nella calcolazione della distribuzione del valore del tempo. Un approccio potrebbe essere quello di: *producing a high number of draws (100'000) for the different coefficients, using the distributional assumptions resulting from the model fitting exercise (vedi Hensher, Greene (2001)).*

All the work was done using the Bierlaire's Optimization Toolbox for GEV Model Estimation¹⁰ and text was written with L^AT_EX.

¹⁰<http://www.biogeme.epfl.ch>

Riferimenti bibliografici

- [1] Axhausen, K. W., Polak, J. W. (1991), "Choice of parking: Stated preference approach", *Transportation* 18, pp. 59-81.
- [2] Bath, C. R. (2000), "A multi-level cross-classified model for discrete response variables", *Transportation Research*, 34B(7), pp. 567-582.
- [3] Bath, C. R., Castelar, S. (2002), "A Unified Mixed Logit Framework for Modelling Revealed and Stated Preferences: Formulation and Application to Congestion Pricing Analysis in the San Francisco Bay Area", *Transportation Research*, 36B(7), pp. 593-616.
- [4] Ben-Akiva, M., and Lerman, S. R. (1985), "Discrete Choice Analysis: Theory and Application to Travel Demand", Seventh Edition, *The MIT Press*, Cambridge MA, *Cambridge University Press*, Cambridge, MA.
- [5] Ben-Akiva, M., and Bierlaire, M. (1999), "Discrete choice methods and their applications to short-term travel decision", in R. Hall (ed.), *Handbook of Transportation Science*, Kluwer, pp. 5-34.
- [6] Bierlaire, M. (2003), "Biogeme: A free package for estimation of discrete choice models", *Proceeding of the 3th Swiss Transportation Research Conference*, Ascona, Switzerland.
- [7] Bierlaire, M. (2005), "An introduction to Biogeme Version 1.4", biogeme.epfl.ch.
- [8] Bradley, M., Kroes, E. and Hinloopen, E. (1993), "A joint model of mode/parking type choice with supply-constrained application", *Proceeding of the 21st Annual Summer PTRC Meeting on European Transport, Highways and Planning*, pp. 61-73.
- [9] Brownstone D., Train, K. (1999) "Forecasting new product penetration with flexible substitution patterns", *Journal of Econometrics*, vol. 89, pp. 109-129.
- [10] Erguen, G. (1971), "Development of a downtown parking model", *Highway Research Record* 369, pp. 118-134.
- [11] Fahrmeir, L., Tutz, G. (1994): "Multivariate statistical modelling based on generalized linear models". *Springer Series in Statistics*.
- [12] Gelhausen, M. C., (2006): "Airport and Access Mode Choice in Germany: A Generalized Neural Logit Model Approach". *Proceedings of the 2006 European Transport Conference*: pp. 1-32.
- [13] Haan, P., Uhlendorff, A., (2006), "Estimation of multinomial logit models with unobserved heterogeneity using maximum simulated likelihood", *DIW*, Berlin.
- [14] Hensher, D. A. and King, J. (2001), "Parking demand and responsiveness to supply, pricing and location in the Sydney central business district", *Transportation Research*, 35A(3), pp. 177-196.

- [15] Hensher, D., Greene, W. H., (2001), "The Mixed Logit Model: The State of Practice and Warning for the Unwary", *Institute of Transportation Studies*, The University of Sydney, Sydney, Australia.
- [16] Hensher, D. A., Rose, J. M. and Greene, W. H. (2005), "Applied Choice Analysis. A Primer", *Cambridge University Press*, Cambridge.
- [17] Hess, D. B. (2001), "The Effect of Free Parking on Commuter Mode Choice: Evidence from Travel Diary Data", *Transportation Research Record* 1753, pp. 35-42.
- [18] Hess, S., Polak, J. W, (2003), "An analysis of parking behaviour using discrete choice models calibrated on SP datasets", paper presented at the *3th European Transport Conference*, Strasbourg.
- [19] Kuhfeld, W. F. (2000), "Multinomial Model Discrete Choice Modeling", *SAS Institut*, Cary, NC
- [20] Lancaster, K. (1991), "Modern Consumer Theory", *Edward Elgar*.
- [21] Louviere, J. J., Hensher, D. A. and Swait, J. D. (2000), "Stated Choice Methods. Analysis and Application", *Cambridge University Press*, Cambridge.
- [22] McCluskey, J.J., and M.L. Loureiro. 2003. "Consumer Preferences and Willingness to Pay for Food Labeling: A Discussion of Empirical Studies." *Journal of Food Distribution Research* 34(1): 95-102.
- [23] McDonnell, S. T., Convery, F. J., Ferreira, S. (2007), "Impact of Modal Choice and Residential Location on Willingness to Pay for Bus Priority Provision: Evidence from Stated-Choice Survey of Catchment Area Residents in Ireland", *Transportation Research Board Annual Meeting 2007 Paper 07-2226*
- [24] McFadden, D. (1974), "Conditional logit analysis of qualitative choice behaviour", in P. Zarembka, ed., *Frontiers in Econometrics*, Academic Press, New York, pp. 105-142.
- [25] Montgomery, D. C. (2001), "Design and Analysis of Experiments", 5th Edition, *John Wiley and Sons*.
- [26] Polak, J. W., and Vythoulkas, P. (1993), "An Assesment of the state-of-the-art in the modelling of park behaviour", TSU Report 752, *Transport Studies Unit*, University of Oxford.
- [27] Spiess, H. (1996), "A Logit Parking Choice Model with Explicit Capacities", Working Paper, EMME/2 Support Center, Aegerten, Switzerland.
- [28] Teknomo, K. and Hokao, K. (1997), "Parking Behaviour in Central Business District: A. Case Study of Surabaya, Indonesia", *EASTS Journal*, Vol. 2.
- [29] Tillé, Y., (2001), "Théorie des sondages", *Dunod*, Paris.

- [30] Train, K. (1998), "Recreation demand models with taste differences over people", *Land Economics*, vol. 74, pp. 185-194.
- [31] Train, K. (2003), "Discrete Choice Methods with Simulation", *Cambridge University Press*, Cambridge, MA.
- [32] Van der Goot, D. (1982), "A model to describe the choice of parking places", *Transportation Research*, 16A, pp. 109-115.